University of Groningen

Neural mass model-based tracking of anesthetic brain states

Kuhlmann, Levin; Freestone, Dean R.; Manton, Jonathan H.; Heyse, Bjorn; Vereecke, Hugo Eric Marc; Lipping, Tarmo; Struys, Michel; Liley, David T. J.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Final author's version (accepted by publisher, after peer review)

Publication date:
2016

[Link to publication in University of Groningen/UMCG research database](#)

Neural mass model-based tracking of anesthetic brain states

Levin Kuhlmann[a,*], Dean R. Freestone[b], Jonathan H. Manton[a], Bjorn Heyse[c], Hugo E.M. Vereecke[d], Tarmo Lipping[e], Michel M.R.F. Struys[c,d], and David T.J. Liley[f]

[a] *Department of Electrical and Electronic Engineering, University of Melbourne, Parkville VIC 3010, Australia.*

[b] *Department of Medicine - St. Vincent's, University of Melbourne, Parkville VIC 3010, Australia.*

[c] *Department of Anesthesia, Ghent University Hospital, Gent, Belgium*

[d] *Department of Anesthesiology, University of Groningen, University Medical Center Groningen, The Netherlands*

[e] *Department of Information Technology, Pori Campus, Tampere University of Technology, Pori, Finland.*

[f] *Brain and Psychological Sciences Research Centre, Swinburne University of Technology, Hawthorn, VIC, Australia.*

Abstract

Neural mass model-based tracking of brain states from electroencephalographic signals holds the promise of simultaneously tracking brain states while inferring underlying physiological changes in various neuroscientific and clinical applications. Here, neural mass model-based tracking of brain states using the unscented Kalman filter applied to estimate parameters of the Jansen-Rit cortical population model is evaluated through the application of propofol-based anesthetic state monitoring. In particular, 15 subjects underwent propofol anesthesia induction from awake to anesthetised while behavioural responsiveness was monitored and frontal electroencephalographic signals were recorded. The unscented Kalman filter Jansen-Rit model approach applied to frontal electroencephalography achieved reasonable testing performance for classification of the anesthetic brain state (sensitivity: 0.51; chance sensitivity: 0.17; nearest neighbor sensitivity 0.75) when compared to approaches based on linear (autoregressive moving average) modelling (sensitivity 0.58; nearest neighbor sensitivity: 0.91) and a high performing standard depth of anesthesia monitoring measure, Higuchi Fractal Dimension (sensitivity: 0.50; nearest neighbor sensitivity: 0.88). Moreover, it was found that the unscented Kalman filter based parameter estimates of the inhibitory postsynaptic potential amplitude varied in the physiologically expected direction with increases in propofol concentration, while the estimates of the inhibitory postsynaptic potential rate constant did not. These results combined with analysis of monotonicity of parameter estimates, error analysis of parameter estimates, and observability analysis of the Jansen-Rit model, along with considerations of extensions of the Jansen-Rit model, suggests that the Jansen-Rit model combined with unscented Kalman filtering provides a valuable benckmark for future real-time brain state tracking studies. This is especially true for studies of more complex, but still computationally efficient, neural models of anesthesia that can more accurately track the anesthetic brain state, while simultaneously inferring underlying physiological changes that can potentially provide useful clinical information.

Keywords: neural mass models, unscented Kalman filtering, brain state tracking, parameter estimation, anesthesia.

*Corresponding author at: Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville VIC 3010, Australia. E-mail address: levink@unimelb.edu.au (L. Kuhlmann)

1. Introduction

Automated electroencephalography (EEG) based depth of anesthesia monitoring is a long-standing problem in the EEG literature (Bruhn et al., 2006; Voss & Sleigh, 2007; Palanca et al., 2009). Various commercial and non-commercial depth of anesthesia monitoring approaches have been developed (Kissin, 2000; Struys et al., 2002; Jordan et al., 2006; Ferenets et al., 2007; Liley et al., 2010; Shalbaf et al., 2013; Shoushtarian et al., 2015b,a) that primarily rely on extraction of features from the EEG to track anesthetic brain state. Despite significant history and recent work attempting to characterise the mulit-channel EEG and brain networks related to anesthesia in more detail (Cimenser et al., 2011; Lewis et al., 2012; Purdon et al., 2013; Kuhlmann et al., 2013; Lee et al., 2013), the international uptake of automated depth of anesthesia monitoring in the clinic is still lagging. Moreover, although commercial monitors can help to improve anesthetic delivery and postoperative recovery (Punjasawadwong et al., 2014; Kettner, 2014) their ability to reduce intraopertive awareness is still under question (Myles et al., 2004; Mashour et al., 2012), as is their ability to track effects of anesthetics with different molecular modes of action (Hirota, 2006). Therefore, improvements are needed.

Model-based approaches involving autoregressive moving average (ARMA) modelling have been shown to be able to disentangle the effects of simultaneously applied anesthetic and analgesic (Liley et al., 2010; Shoushtarian et al., 2015b,a). Therefore model-based approaches present an interesting alternative to the more standard extraction of EEG features employed in depth of anesthesia monitoring (Kissin, 2000; Jordan et al., 2006; Ferenets et al., 2007). In particular, neural mass or neural field model (Deco et al., 2008) approaches that attempt to take into account the meso-/macro-scale neurophysiology involved in the generation of EEG at rest and during anesthesia (Bojak & Liley, 2005; McCarthy et al., 2008; Ching et al., 2010; Steyn-Ross et al., 2012; Hindriks & van Putten, 2012; Hutt, 2013; Hutt & Buhry, 2014) present an interesting alternative to tracking the anesthetic brain state, while simultaneously allowing for estimation of underlying physiological variables that may have clinical significance.

The main idea behind these approaches is that different regions of parameter space of neural mass models describe different types of state dynamics of these models, such as limit cycles and fixed points, which in turn result in different types of modelled EEG amplitude spectra that can be related to real EEG data and spectra (Freestone et al., 2013; Moran et al., 2013). Methods that estimate the parameters of these neural mass models using real EEG data can then be used to infer and track key physiological variables, such as post-synaptic potential (PSP) amplitudes and rate constants, and these estimates in turn can be used to determine the current brain state (e.g. awake, anesthetised, asleep, or seizure) based on their relationship to the EEG signal dynamics and amplitude spectra (Freestone et al., 2013; Moran et al., 2013; Kuhlmann et al., 2015a).

Various approaches for parameter estimation of neural mass models have been developed (Wendling et al., 2002; Moran et al., 2013; Chong et al., 2011; Postoyan et al., 2012; Chong et al., 2012a,b, 2015; Freestone et al., 2011, 2013, 2014). Given that neural mass models are generally nonlinear and related noise sources can be considered to be Gaussian white noise (Liley et al., 2002; Nunez & Srinivasan, 2006), this paper presents the application of the unscented Kalman filter (UKF) to state and parameter estimation of a neural mass model of population activity in cerebral cortex. The goals are (1) to evaluate whether the parameter estimations can be used to efficiently track anes-

thetic brain states during anesthesia in humans using the gamma-amino-butyric-acid-A (GABA-A) receptor agonist propofol (Rudolph & Antkowiak, 2004), and (2) to evaluate if the approach can simultaneously provide information about the key changes in inhibitory post-synaptic potential (IPSP) amplitude and rate constant parameters known to vary with the level of propofol concentration (Collins, 1988; Bieda & MacIver, 2004; Wang et al., 2004; Ying & Goldstein, 2005; McCarthy et al., 2008; Wang, 2009; Hindriks & van Putten, 2012; Hutt, 2013). Such information may help anesthesiologists monitor the depth of anesthesia as well as understand the patient's underlying inhibitory systems during anesthesia and, once such methods have matured, provide improvements in general anesthesia and surgery outcomes.

2. Methods

Tracking of anesthetic brain state with neural mass model parameters was performed using a distribution-based approach (c.f. Zikov et al. (2006)) to classify the brain states of an individual by comparing the distribution of recently calculated neural mass model parameter estimates computed from single-channel frontal EEG data to distributions of the same variables for each brain state that have been computed from data from a set of 'training' subjects. The Jansen-Rit (JR) neural mass model with a 6 dimensional state space (Jansen & Rit, 1995) is the primary neural mass model considered here. The choice of this model lies in an Occam's razor approach where the model with the fewest assumptions that still describes the data and underlying physiology should be selected. This is consistent with the goal of using models that are simple enough to be part of computationally efficient brain state tracking algorithms. Thus one goal here is to begin to evaluate the ability of simple models of population activity in cerebral cortex, like the JR model, to track anesthetic brain states before considering more complex models. Although there are several other neural models that have been specifically developed to describe the EEG for different anesthetics (Bojak & Liley, 2005; McCarthy et al., 2008; Ching et al., 2010; Steyn-Ross et al., 2012; Hindriks & van Putten, 2012; Hutt, 2013; Hutt & Buhry, 2014) this paper focuses on the JR model (and extensions of this model that include inhibition of inhibition) to provide a benchmark for future studies with more complex models in order to find the simplest neural mass model that is accurate, informative and efficient enough for clinical application.

In order to benchmark the neural mass model approach it is compared to approaches based on (1) ARMA model parameter estimates obtained using the Broersen technique (Broersen, 2002, 2006) or (2) one of the best performing depth of anesthesia monitoring measures, Higuchi fractal dimension (HFD) (Higuchi, 1988; Ferenets et al., 2007). The evaluation of the approaches involves applying the distribution-based approach to classify the brain states, where the distributions are composed of (i) neural mass model parameter estimates, (ii) ARMA model parameter estimates, or (iii) HFD values. ARMA models provide a generic model-based benchmark and the Broersen technique for ARMA models is the primary estimation technique underlying the Brain anesthesia Response (BAR) monitor (Cortical Dynamics, Australia) (Liley et al., 2008, 2010; Shoushtarian et al., 2015b,a) developed in part by one of the authors. On the other hand, the HFD, when evaluated with the same dataset as used in this study, has been demonstrated to perform as well as or better than, three entropy-based features, two commercial entropy-based features, and three features used in the most common commercial depth of anesthesia monitoring device, the BIS

monitor (Covidian, Ireland) (Ferenets et al., 2007).

The methods underpinning our distribution-based classification approach proceed in two stages: first the methods underlying the JR model, neural mass model parameter estimation, ARMA modelling, ARMA model estimation and the HFD are described. Then the depth of anesthesia monitoring EEG data and the distribution-based tracking/classification approach are described. To complement the classification performance analysis of the different distribution-based approaches we further consider the following: (1) a monotonicity analysis of JR model parameter estimates to determine whether estimates of the parameters considered to be effected by propofol anesthesia vary in the same direction as known physiologically for increasing concentrations of propofol; (2) an observability and error analysis of the JR model states and parameters to further understand the influence of the parameter space on JR model state and parameter estimation; (3) an error analysis of extensions of the JR model that include inhibition of inhibition to assess if these extensions would be able to describe the power spectra for different anesthetic brain states.

2.1. Jansen-Rit (JR) neural mass model

The JR neural mass model (Jansen & Rit, 1995) of population activity in cerebral cortex, or modifications thereof, forms the basis of many current approaches to infer underlying physiological variables from sparsely sampled electrophysiological recordings (Wendling et al., 2002; Moran et al., 2013; Chong et al., 2011; Postoyan et al., 2012; Chong et al., 2012a,b, 2015; Freestone et al., 2011, 2013, 2014). This combined with the simplicity of the JR model makes it a suitable first choice in the search for the simplest neural mass model that is both accurate, informative and efficient enough for clinical application in anesthesia. The key features of the JR model considered here are illustrated in Figure 1(A). The model consists of three cortical populations: excitatory pyramidal cells, spiny stellate excitatory cells and inhibitory interneurons. The pyramidal population is driven by endogenous input activity and excites the spiny stellate excitatory cells and inhibitory interneurons, while the spiny stellate and inhibitory interneuron populations provide excitatory and inhibitory feedback, respectively, to the pyramidal population. The synaptic dynamics are governed by the population excitatory postsynaptic potential (EPSP) response kernel

$$h_e(t; A, a) = H(t) A a t \exp(-at), \tag{1}$$

and the population IPSP response kernel

$$h_i(t; B, b) = H(t) B b t \exp(-bt), \tag{2}$$

where t is time, A and B are the amplitude parameters and a and b are the rate constants for the EPSP and IPSP responses, respectively, and $H(t)$ is the Heaviside step function. The mean membrane potential, $v(t)$, of each population is converted into a population firing rate using a sigmoidal transfer function,

$$S(v(t)) = \frac{2e_0}{1 + \exp(r[v_0 - v(t)])} \tag{3}$$

4

where e_0 is a scaling parameter, r controls the slope of the sigmoid and v_0 is a threshold parameter. The configuration of EPSP and IPSP response kernels and firing rate transfer functions in the model is schematised in Figure 1(B). The endogenous noise driving the pyramidal population is captured by the input $\mu(t)$ and the EEG output signal of the model is considered to be mean membrane potential of the pyramidal population. The full set of equations defining the model are as follows

$$\dot{x}_1(t) = x_4(t) \tag{4}$$

$$\dot{x}_4(t) = AaS(x_2(t) - x_3(t)) - 2ax_4(t) - a^2 x_1(t) \tag{5}$$

$$\dot{x}_2(t) = x_5(t) \tag{6}$$

$$\dot{x}_5(t) = Aa\left[\mu(t) + c_2 S(c_1 x_1(t))\right] - 2ax_5(t) - a^2 x_2(t) \tag{7}$$

$$\dot{x}_3(t) = x_6(t) \tag{8}$$

$$\dot{x}_6(t) = Bb\left[c_4 S(c_3 x_1(t))\right] - 2bx_6(t) - b^2 x_3(t) \tag{9}$$

where x_1, x_2 and x_3 are the outputs of the three PSP kernels, c_1, c_2, c_3 and c_4, are connectivity constants and the single-channel EEG output signal is given by $y(t) = x_2(t) - x_3(t)$. Unless specified elsewhere the parameter values used are the same as those specified by Jansen & Rit (1995). For any forward simulations of the model Euler-Marayama numerical integration (Kloeden & Platen, 1992) was used with a time step of $\Delta t = 0.1$ ms and the input $\mu(t)$ was modelled as Gaussian white noise with a scaled standard deviation of $\sqrt{\Delta t} Aa\sigma$ with $\sigma = 5.74$. Moreover for simulations where the input mean was constant, this value was taken to be 220.

2.2. JR model state and parameter estimation

To track the states and selected parameters of the JR model estimated from frontal single-channel EEG measurements the time domain method of unscented Kalman filtering (UKF) (Simon, 2006; Freestone et al., 2013) is used. The UKF updates parameter estimates with each new EEG sample received and its computational efficiency as compared to more complicated filtering or estimation techniques (Simon, 2006) means that it holds potential for real-time tracking of brain states. Moreover, the UKF works well for nonlinear systems like the JR model and it assumes Gaussian data, however, this is a reasonable assumption for EEG data (Freestone et al., 2013). This approach has been successfully applied to the tracking of seizure dynamics through feeding single channel EEG into the UKF to invert the JR model and estimate its state and up to 5 parameters including the four PSP parameters A, a, B and b (Freestone et al., 2013). Here the same methods are employed, however, the focus is on estimating the key parameters that are most effected by increases in propofol concentration, the IPSP amplitude B and IPSP rate constant b (Collins, 1988; Krasowski et al., 1998; Bai et al., 1999, 2001; Siegwart et al., 2002; Dong & Xu, 2002; Kitamura et al., 2003; Wang et al., 2004; Feng & Macdonald, 2004; Bieda & MacIver, 2004; Ying & Goldstein, 2005; McCarthy et al., 2008; Drexler et al., 2009; Wang, 2009; Yue et al., 2011; Wakita et al., 2013; Hutt, 2013). These two parameters are considered to primarily capture the shape of the EEG amplitude spectrum, while a third parameter, the input to the model μ is also estimated to track scaling of the amplitude spectrum. All other model parameters

are treated as constants. Limiting the parameters to be estimated as $\boldsymbol{\theta} = \begin{bmatrix} B & b & \mu \end{bmatrix}^T$ somewhat constrains the ability of the JR model to fit the EEG data, however, it enables evaluation of whether the IPSP amplitude B and IPSP rate constant b increase and decrease, respectively, as is observed when the concentration of propofol anesthetic increases (Kitamura et al., 2003; Wang et al., 2004; Bieda & MacIver, 2004; Ying & Goldstein, 2005; McCarthy et al., 2008; Wang, 2009; Yue et al., 2011; Wakita et al., 2013; Hutt, 2013). This then can act as a benchmark for more complex and accurate models of the EEG linked to propofol anesthesia for the purposes of anesthesia monitoring and simultaneous estimation of known parameter changes.

The UKF approach relies on defining the JR model in the following discrete-time state space form

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, \boldsymbol{\theta}_t) + \boldsymbol{\varepsilon}_t \tag{10}$$

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t + \boldsymbol{\varphi}_t \tag{11}$$

$$y_t = \mathbf{C}\mathbf{x}_t + e_t, \tag{12}$$

where $\mathbf{x}_t \in \mathbb{R}^6$ is the vector of JR model states, and the disturbance terms $\boldsymbol{\varepsilon}_t$, $\boldsymbol{\varphi}_t$ and e_t capture model error, parameter drift and EEG measurement noise, respectively, and are all defined by independent and identically distributed Gaussian processes. Considering the parameters as constant plus a drift reflects the idea that the parameters are slowly varying compared to the states. The EEG output defined here is the same as described earlier for the JR model but now also includes the measurement noise. By augmenting the state vector with parameters that are to be estimated, the UKF algorithm can be applied and state and parameter estimation can be performed simultaneously. The augmented state vector is defined as

$$\mathbf{x}_t^{\theta} = \begin{bmatrix} \mathbf{x}_t^T & \boldsymbol{\theta}_t^T \end{bmatrix}^T \tag{13}$$

and the augmented state vector estimate is given by

$$\hat{\mathbf{x}}_t^{\theta} = \begin{bmatrix} \hat{\mathbf{x}}_t^T & \hat{\boldsymbol{\theta}}_t^T \end{bmatrix}^T \tag{14}$$

where $\hat{\boldsymbol{\theta}}_t = \begin{bmatrix} \hat{B}_t & \hat{b}_t & \hat{\mu}_t \end{bmatrix}^T$ and $\hat{\mathbf{x}}_t$ are the parameter and original state estimates, respectively. The time varying parameter estimates can then be found by determining the *a posteriori* augmented state estimate

$$\hat{\mathbf{x}}_t^{\theta} = \hat{\mathbf{x}}_{t+}^{\theta} = \mathbb{E}\left[\mathbf{x}_t^{\theta} | y_1, y_2, \ldots, y_t \right], \tag{15}$$

using the UKF. In this case the *a posteriori* augmented state estimate is the expected value of the current augmented state at time t given the history of all EEG measurements (or model output), $y_1, y_2, \ldots, y_t$, under Gaussian assumptions. The discrete-time UKF is a well known algorithm in the engineering literature and the reader is referred to Simon (2006) and Freestone et al. (2013) for further details.

In a spatial sense, this single-channel EEG analysis using the UKF with the JR model is a simplification of the more spatially complex forward and inverse modelling of electromagnetic source imaging (Nunez & Srinivasan, 2006).

Such an approach is warranted here as the focus is on clinically practical anesthetic state monitoring using neural mass models. In a temporal sense, this approach is more complex than more standard source imaging approaches (Nunez & Srinivasan, 2006).

2.3. ARMA models

ARMA time series models provide an accurate linear model description of single-channel EEG (Broersen, 2006). The basic form of an ARMA model is as follows:

$$y_t = -\sum_{j=1}^{p} a^{(j)} y_{t-j} + \sum_{k=1}^{q} b^{(k)} z_{t-k} + z_t, \tag{16}$$

where y_t is the observed (frontal EEG) signal, and $a^{(j)}$ and $b^{(k)}$ are the autoregressive (AR) and moving-average (MA) parameters, respectively, at time t. The constants p and q are the corresponding orders of the AR and MA parts, respectively, and z_t is the observation error or innovation process. The innovation process is assumed to be a Gaussian white noise process with zero mean and variance $\sigma_{z_t}^2$. Based on prior investigations a model order of $(p = 2, q = 1)$ was considered (Kuhlmann et al., 2015b). Although empirical determinations of optimal ARMA fits obtained with resting awake eyes-closed EEG suggest orders of $(p = 3 - 14, q = 2 - 5)$ (Schack & Krause, 1995; Tseng et al., 1995), such a low order model of $(p = 2, q = 1)$ reduces the degree of overfitting to the EEG and therefore provides good out-of-sample testing classification performance (Kuhlmann et al., 2015b).

2.4. ARMA estimation with the Broersen Technique

For a given epoch of single-channel data the Broersen technique was used to estimate an invertible and stationary ARMA model using a variant of Durbin methods with optimal intermediate autoregressive order on zero meaned data (Broersen, 2002, 2006). The Broersen method of ARMA model/parameter estimation is a well established method and is implemented directly in the ARMASA MATLAB Toolbox (Broersen, 2002, 2006). Subsequent estimates of the innovation variance, $\sigma_{z_t}^2$, are calculated as the standard deviation of the zero meaned signal epoch divided by the square root of the power gain of the derived filter/ARMA model. With regard to Equation 16 the AR and MA parameters are considered constant over a finite analysis epoch.

2.5. Higuchi fractal dimension - HFD

HFD is a nonlinear dynamical measure of the fractal dimension of a time-series that quantifies the scaling of the length of the time series, when viewed geometrically as a curve, as it is successively subsampled (Higuchi, 1988). HFD has very low computational complexity because it can be calculated in the time domain. Moreover as described above, HFD performs well as a feature for frontal-EEG-based depth of anesthesia monitoring. In particular, empirical investigations have established that HFD calculated on the 6-47 Hz frequency band gives the best performance (Ferenets et al., 2007). Here HFD was calculated on the 6-47 Hz frequency band for finite length epochs using the method outlined by Ferenets et al Ferenets et al. (2007).

7

Table 1: Responsiveness Scores of the Modified OAA/S Scale.

Score	Responsiveness
5	Responds readily to name spoken in normal tone
4	Lethargic response to name spoken in normal tone
3	Responds only after name is called loudly and/or repeatedly
2	Responds only after mild prodding or shaking
1	Responds only after painful trapezius squeeze
0	No response after painful trapezius squeeze

2.6. Depth of anesthesia monitoring data

The data analysed here has been used in previous studies on depth of anesthesia monitoring (Ferenets et al., 2007; Liley et al., 2010). Detailed information about the patient cohort and anesthetic protocols for this dataset can be found in Ferenets et al. (2007). The main details are summarized here. Frontal EEG data was recorded from 15 subjects undergoing propofol anesthesia from the awake to the anesthetised state using the M-Entropy module of the S/5 Anesthesia Monitor (GE Healthcare Finland Oy). The original EEG data were bandpass filtered at 0.5-118 Hz and sampled at 400 Hz. The standard entropy sensor of the S/5 Monitor was applied with a slightly modified positioning: the two recording electrodes of the sensor were positioned bilaterally on the forehead approximately 5 cm above the eyebrows and 4 cm from the midline in either direction. The ground electrode was positioned between the two recording electrodes. This montage was selected to minimize electromyographic activity that normally contributes to the calculation of the State Entropy and Response Entropy measures of the S/5 monitor. To start each subject study, a propofol effect site concentration of 0.75 μg/ml was targeted, then increased every 4 min by 0.25-0.30 μg/ml until loss of response to all relevant clinical measures of anesthetic depth was observed. Behavioural responsiveness of each subject was assessed using the modified Observers Assessment of Alertness/Sedation (OAA/S) score, a subjective clinical measure of arousal, alertness and sedation (Chernik et al., 1990). This behavioural scale ranges from 5 for the fully alert/awake state to 0 for complete unresponsiveness to a painful stimulus (see Table 1).

2.7. Resting state eyes-closed control data

To control against the possibility that UKF parameter estimates may by drifting due to unknown signal changes or saturation of the sigmoidal transfer function, resting state eyes-closed data from a single individual was used. This 5 minute segment of frontal EEG data was obtained from the FPz position referenced to linked mastoids in a previous study (Foster & Liley, 2011). The data was sampled at 500 Hz with a 0.1-70 Hz bandpass anti-aliasing filter and line noise and eye blink artifacts were subsequently removed (Foster & Liley, 2011).

2.8. Data pre-processing and artifact rejection

The data analysis for the JR and ARMA model estimation follows in part a previously described approach of the authors that involves the estimation of cortical state and cortical input using ARMA models in order to track

depth of anesthesia and analgesia, respectively (Liley et al., 2010). HFD was calculated according to the method detailed in Ferenets et al. (2007). The JR and ARMA model parameter estimates and the HFD were all calculated on contiguous 1 s non-overlapping data epochs. The 1 s epochs were assessed for artifact as outlined below and artifact containing epochs were ignored, otherwise epochs were accepted.

Because the UKF is an iterative algorithm it needs the data to be approximately continuous. The word 'approximate' is used here because if a 1-second window is removed or ignored due to artifact then for the next accepted 1-second window the UKF was initialised using values from the end of the last accepted window. Doing so avoids large artifact induced fluctuations of the parameter estimates that can affect the numerical stablity of the algorithm and classification sensitivity. The Broersen method used to estimate the ARMA ($p = 2, q = 1$) model parameters operates on a epoch-by-epoch basis and therefore estimation for the current window is not effected by estimates in the previous window. The same is true for the HFD.

For the UKF applied to the JR model approach, prior to windowing, the data was low pass filtered at 45 Hz with a 5th order Butterworth filter to remove 50 Hz line noise and resampled up to 1000 Hz to ensure numerical accuracy of the UKF. Application of the UKF with the JR model to real EEG data then required rescaling of the mean and standard deviation of the EEG signal to match that of the JR model for each considered region in parameter space. Moreover, when the UKF approach is applied to the OAA/S classification of real data the JR model and the initial parameter estimates, $\hat{\boldsymbol{\theta}}_0$, are set to the standard Jansen-Rit parameter values (Jansen & Rit, 1995). The effect of different regions of parameter space are considered in the additional analysis (see section 2.10).

For ARMA modelling, prior to windowing, the data was resampled to 80 Hz using a finite impulse response antialiasing filter with a sharp cutoff at 40 Hz, with the transition band made sufficiently sharp to minimize any aliasing. This was performed to avoid spurious fitting to 50-Hz line noise spectral peaks or any low-pass filter band edges.

For the HFD analysis, the raw data were filtered using a 6-47 Hz pass-band linear phase equiripple filter as previously described (Ferenets et al., 2007). The HFD was then calculated on 1-second epochs of this filtered data (Accardo et al., 1997).

Artifact rejection of epochs involved calculating the electromyogram (defined as the total power between 70-110 Hz excluding a notch at 98-102 Hz due to 50 Hz electric power harmonic at 100 Hz) using the raw EEG data. The root mean square (RMS) amplitude was also calculated from the EEG time series resampled at 80 Hz based on prior studies (Liley et al., 2010). Epochs were automatically excluded from further analysis if any of the following occurred: total electromyographic power greater than approximately 400 μV^2 or less than approximately 0.004 μV^2, RMS amplitude less than 5 μV or greater than 150 μV, or amplitude distributions were not normal (based on Lilliefors test at P<0.01). Given that the UKF was applied to the EEG time series resampled at 1000 Hz a second artifact rejection approach was also considered for the UKF analysis where the RMS amplitude and Lilliefors test was calculated using the EEG time series sampled at 1000 Hz instead of 80 Hz.

Table 2: Training and testing sets of cross-validation analysis.

Set	Subject indices	# OAA/S epochs						# patients per OAA/S					
		0	1	2	3	4	5	0	1	2	3	4	5
Train 1	1, 3, 4, 6, 7, 10, 11, 15	7	2	6	7	22	35	7	2	5	4	8	8
Test 1	2, 5, 8, 9, 12, 13, 14	6	2	6	9	24	36	6	2	5	3	7	7
Train 2	1, 5, 7, 8, 9, 10, 13, 15	7	2	6	7	23	38	7	2	6	4	8	8
Test 2	2, 3, 4, 6, 11, 12, 14	6	2	6	9	23	33	6	2	4	3	7	7
Train 3	2, 5, 6, 7, 11, 12, 13, 15	7	2	6	7	21	36	7	2	5	4	8	8
Test 3	1, 3, 4, 8, 9, 10, 14	6	2	6	9	25	35	6	2	5	3	7	7

OAA/S epochs: number of OAA/S epochs used for each OAA/S score for each set; # patients per OAA/S: number of patients contributing data for each OAA/S score for each set.

2.9. Tracking/classifying OAA/S state using distributions

The tracking/classification of the OAA/S score was performed using discrete distributions of either the estimated JR or ARMA model parameters (see sections 2.2 and 2.4) or Higuchi Fractal Dimension (see section 2.5). The primary goal is to track the OAA/S score for an arbitrary individual undergoing general anesthesia during surgery. Out-of-sample cross validation is applied here to evaluate the ability of the considered approaches to track/classify the OAA/S score of an arbitrary subject. This involves breaking up the data into training and testing sets that provide an approximately equal number of patients and number of OAA/S epochs for each OAA/S score. Moreover, the training sets were chosen such that no more than half of the patients are the same in each training set. This provides a balance between training and testing sets while preserving differences between the training sets. Three-fold cross validation was applied, meaning three train and test set pairs were created for validation. The data contained in each of these sets is detailed in Table 2. OAA/S epochs were defined to be 30-second segments prior to the OAA/S measurement times. This duration corresponds approximately to the amount of time it takes to obtain an OAA/S measurement. The OAA/S measurements were written down in the 10 seconds preceding each increase in propofol concentration. Only these OAA/S epochs were used in the cross-validation classification performance analysis, however, in practice the same training distributions can be used for tracking of brain states with continuous data. For an example case, tracking in this sense is considered here, where new parameter estimates are obtained every second and the current distribution of parameter estimates is computed over the last 30 seconds of data and is updated every second. Here 'current distribution' essentially refers to the distribution of the most recent parameter estimates obtained from scrolling EEG data.

For the case of the JR or ARMA models, training distributions of estimated parameters for each model for each

OAA/S score, $P_{ij..k}^{(o)}$, were obtained from the OAA/S epochs for all of the subjects in a training set pooled together for each OAA/S score, where o indexes the OAA/S score and $ij..k$ are the dimensions of the discrete probability distribution which correspond to the estimated parameters. For the case of HFD, univariate training distributions were constructed from the calculated HFD values. In this sense training does not involve any form of learning but rather the creation of distributions for each OAA/S score using the training set data. Then the current OAA/S score is estimated by comparing the distributions of the estimated parameters, or HFD values, of a current 30-second (OAA/S) epoch, $P_{ij..k}^{(t)}$, to the 6 training distributions of the estimated parameters, or HFD values, corresponding to the 6 OAA/S scores. This is achieved by taking the estimated OAA/S score for a current individual OAA/S epoch to be the minimum of the total variation (TV) between the current distribution and each of the 6 OAA/S score training distributions:

$$\text{TV}_t^{(o)} = \sum_{i,j,..,k} \frac{1}{2} \left| P_{ij..k}^{(o)} - P_{ij..k}^{(t)} \right|, \tag{17}$$

$$\hat{O}_t = \text{argmin}_o \text{TV}_t^{(o)}. \tag{18}$$

Total variation takes on values between 0 for exact matching distributions and 1 for non-overlapping distributions. The OAA/S index o with the smallest total variation (i.e. closest match to the different OAA/S level training distributions) is chosen to be the estimate, $\hat{O}_t$. This estimate is compared with the true OAA/S score for the 30-second epoch in order to assess classification sensitivity. This is done for each OAA/S epoch in the training and testing sets. Different measures, such as Jensen-Shannon Divergence (Lin, 1991) or earth mover's distance (Levina & Bickel, 2001), could be used to compare differences in probability distributions, however, the main focus here is to compare the JR and ARMA model and HFD approaches using the same distribution-based method.

2.9.1. Definition of discrete probability distributions

Discrete probability distributions were defined as multivariate functions of estimated parameters or univariate functions of HFD values. The discrete probability distributions were created by dividing each dimension into N bins where the bin centres were selected to uniformly span the range of values defined by the training set. In addition, all bin widths were equal except for bins on the edges of the distribution which were allowed to cover the remaining possible range of values that could be encountered in an arbitrary test set (i.e. $-\infty$ to ∞). In the analysis presented here $N = 5$ bins were considered for each JR or ARMA model parameter dimension. The number of bins was explored and $N = 5$ provided a balance between under and overfitting the data (results not shown). For the case of HFD, since there is only one feature dimension, this dimension was divided into 20 bins when computing the corresponding discrete probability distribution.

2.10. Additional Analysis

2.10.1. Monotonicity analysis of JR model parameter estimates

To assess whether the IPSP amplitude estimates, $\hat{B}_t$, and IPSP rate constant estimates, $\hat{b}_t$, respectively increase and decrease monotonically with increases in propofol concentration a monotonicity index was calculated on the

11

estimates as follows:

$$\Omega = \frac{0.5}{N_t} \sum_{k=1,...,N_t} \left[\frac{\sum_{i=1,..,k} \mathbb{1}_{\hat{B}_i < \hat{B}_k}(\hat{B}_i, \hat{B}_k)}{k} + \frac{\sum_{i=1,..,k} \mathbb{1}_{\hat{b}_i > \hat{b}_k}(\hat{b}_i, \hat{b}_k)}{k} \right] \tag{19}$$

where N_t is the number of parameter estimate samples for an individual, $k = 1, ..., N_t$ indexes this sequence of samples and

$$\mathbb{1}_{x<y}(x, y) = 1 \text{ if } x < y, \text{ or } 0 \text{ otherwise}, \tag{20}$$

and

$$\mathbb{1}_{x>y}(x, y) = 1 \text{ if } x > y, \text{ or } 0 \text{ otherwise}, \tag{21}$$

are indicator functions. The monotonicity index, Ω, takes on values close to 1 when $\hat{B}_t$ and $\hat{b}_t$ increase and decrease perfectly monotonically, respectively, as time progresses with increases in propofol concentration, while Ω shifts towards 0 when the parameter estimates no longer vary monotonically in the desired directions. The parameter space of the JR model was explored to find a region of the parameter space in which parameter estimates would vary monotonically in the desired directions when applied to subject 9, a subject who displayed all OAA/S behavioural levels. In particular, the EPSP amplitude A, EPSP rate constant a and the initial IPSP rate constant estimate $\hat{b}_0$ were varied for many different runs of the UKF. The rate constants were focused on as they primarily control the preferred frequencies of the model output (Jansen & Rit, 1995), while EPSP amplitude was varied to investigate what happens when the balance of excitation and inhibition is changed.

2.10.2. JR model observability and error analysis of UKF estimation

To further investigate how UKF parameter estimation depends on the region of the JR model parameter space, the mean square error (MSE) of the parameter estimates and the observability index was computed across multiple forward simulations of the JR model spanning the same region of parameter space considered in the monotonicity analysis: the EPSP amplitude A, EPSP rate constant a and the IPSP rate constant b. For the calculation of MSE, simulations were averaged over the cases when the initial parameter estimates $\hat{\theta}$ were set 20% above or below the true parameter values $\boldsymbol{\theta}$.

The observability index, δ, is a control theoretic concept that essentially tells us how well we can estimate the states of a system in a deterministic sense without consideration of the sources of noise (Hermann & Krener, 1977). Here the observability index is calculated for the augmented state space system defining the JR model in Equation (13). Given that the JR model is a nonlinear system, a method for nonlinear observability (Hermann & Krener, 1977) has been employed that follows the approach of Whalen et al. (2015). The reader is referred to this paper for mathematical details, however, it is briefly noted that the observability index is taken to be the average over many simulated samples and state trajectories of the absolute ratio of the minimum and maximum singular values of the inner product of the observability matrix evaluated at each sample using the Jacobian of the Lie derivative map of the JR model (Whalen et al., 2015). This definition produces an observability index with the range $0 \le \delta \le 1$ where values closer to 1 indicate full observability of the states.

2.10.3. Error analysis of extensions of the JR model

To consider whether extensions of the JR model may give more physiologically realistic estimation of the variation of IPSP amplitude and IPSP rate constant as propofol concentration increases, extensions of the JR model that include inhibition of inhibition were evaluated. Extending the model to include inhibition of inhibition is justified because previous anesthesia modelling studies have suggested that inhibition of the inhibitory population by itself at the cortical level is important for describing the EEG changes linked to anesthesia (Bojak & Liley, 2005; Hindriks & van Putten, 2012). Two types of extensions were considered. The first simply involved feeding the output of the IPSP kernel back to the inhibitory population input by way of an extra connectivty parameter c_5. The second involved a similar new connection, however, the connection was mediated by an additional IPSP kernel parametrised by the amplitude and rate constant parameters M and m, respectively. The equations defining the first extension are as follows

$$\dot{z}_1(t) = z_4(t) \tag{22}$$

$$\dot{z}_4(t) = AaS(z_2(t) - z_3(t)) - 2az_4(t) - a^2 z_1(t) \tag{23}$$

$$\dot{z}_2(t) = z_5(t) \tag{24}$$

$$\dot{z}_5(t) = Aa\left[\mu(t) + c_2 S(c_1 z_1(t))\right] - 2az_5(t) - a^2 z_2(t) \tag{25}$$

$$\dot{z}_3(t) = z_6(t) \tag{26}$$

$$\dot{z}_6(t) = Bb\left[c_4 S\left(c_3 z_1(t) - \frac{c_5}{c_4} z_3(t)\right)\right] - 2bz_6(t) - b^2 z_3(t) \tag{27}$$

where the $-\frac{c_5}{c_4} z_3(t)$ term in the last state equation is the principle difference with the original JR model.

The equations defining the second extension are as follows

$$\dot{v}_1(t) = v_5(t) \tag{28}$$

$$\dot{v}_5(t) = AaS(v_2(t) - v_3(t)) - 2av_5(t) - a^2 v_1(t) \tag{29}$$

$$\dot{v}_2(t) = v_6(t) \tag{30}$$

$$\dot{v}_6(t) = Aa\left[\mu(t) + c_2 S(c_1 v_1(t))\right] - 2av_6(t) - a^2 v_2(t) \tag{31}$$

$$\dot{v}_3(t) = v_7(t) \tag{32}$$

$$\dot{v}_7(t) = Bb\left[c_4 S\left(c_3 v_1(t) - v_4(t)\right)\right] - 2bv_7(t) - b^2 v_3(t) \tag{33}$$

$$\dot{v}_4(t) = v_8(t) \tag{34}$$

$$\dot{v}_8(t) = Mm\left[c_5 S\left(c_3 v_1(t) - v_4(t)\right)\right] - 2mv_8(t) - m^2 v_4(t) \tag{35}$$

where the addition of the extra IPSP block introduces two new dimensions to the state-space of the model. The block diagrams of the first and second JR model extensions are displayed in Figures 5(A) and (B), respectively.

To evaluate whether the extended models could give more realistic estimation of the IPSP amplitude and rate constant variation, for subject 9 who experienced all OAA/S levels, the average amplitude spectra of each OAA/S

13

level was compared using the MSE to the average amplitude spectra of forward simulations of the models as the level of inhibition of inhibition (i.e. c_5) was varied. For the first extended model, multiple forward simulations were performed across different values of the IPSP amplitude B and IPSP rate constant b to assess if there exists a trajectory through this parameter subspace where the parameters could move in the physiologically known directions as the OAA/S level decreases (i.e. propofol concentration increases). This was achieved by creating maps of the OAA/S score linked to the average data amplitude spectrum that best matches the average amplitude spectrum for multiple forward simulations for each point in the considered parameter subspace. For the second extended model the same approach was considered, however, the IPSP amplitude B was held fixed at the standard value of $B = 22$ (Jansen & Rit, 1995) while the IPSP rate constants b and m were varied given that other modelling studies suggest that the IPSP rate constants are critically important in modelling propofol EEG (Hindriks & van Putten, 2012; Hutt, 2013)

3. Results

It is demonstrated that the UKF reliably estimates parameters for forward simulations of the JR model. An example of tracking the OAA/S state using the UKF approach is then provided, along with a comparison of the classification performance of the JR model, ARMA model and HFD approaches. The monotonicity, observability and error analysis of the JR model/UKF approach is then considered, along with the error analysis of extensions of the JR model.

3.1. UKF estimation of time-varying JR model parameters

Figure 1(C) demonstrates that the UKF, when applied to an example forward simulation of the JR model, can produce parameter estimates (black) that reliably track the true parameters (red), IPSP amplitude B, IPSP rate constant b and input μ, as they vary over a 100 second simulation.

3.2. Example of UKF based OAA/S score tracking

An example of tracking the OAA/S score using the UKF approach with the JR model is given in Figure 2(A). The figure shows the raw EEG signal for subject 9 who experienced all OAA/S levels. The beginning of the recording displays significant artifact which disappears after the propofol induction is begun at the SI marker and the subject has settled. The estimates of the IPSP amplitude $\hat{B}$, IPSP rate constant $\hat{b}$ and input $\hat{\mu}$ are displayed below the raw data. It can be seen that the IPSP amplitude estimate increases with increasing propofol concentration until loss of responsiveness, as is consistent with observed physiology, however, the IPSP rate constant estimate does not decrease as is observed physiologically. Moreover, the input estimate increases with an overall increase in amplitude of the EEG signal as is expected. The fifth and sixth subplots of Figure 2(A) provide illustration of OAA/S score tracking using training set 2 to generate the training distribution. The fifth subplot displays the total variation time series for each OAA/S score, $\text{TV}_t^{(o)}$, illustrating how the lowest total variation values shift from corresponding to an OAA/S score of predominantly 5 (red) at the beginning through to predominantly 0 (pale pink) as the propofol
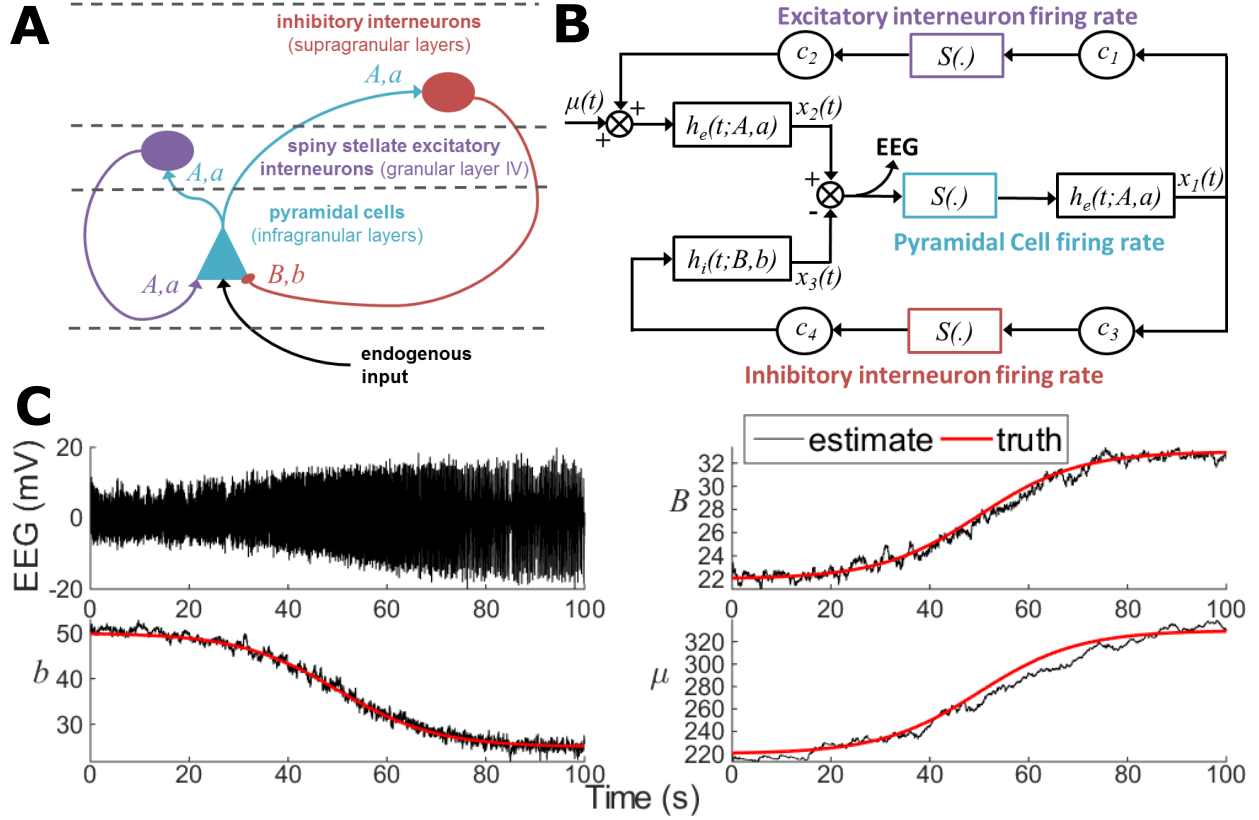
Figure 1: (A) Laminar cortical depiction of the JR model which consists of pyramidal, spiny stellate excitatory interneuron and inhibitory interneuron populations. The pyramidal population excites the stellate and inhibitory populations, is driven by endogenous input, is excited by the stellate population and inhibited by the inhibitory population. (B) The block diagram of the JR model. The model is driven by the input $\mu(t)$ and the mean membrane potential of the pyramidal population is considered to be the output and approximately equivalent to the EEG. Excitatory, $h_e(t; A, a)$, and inhibitory, $h_i(t; B, b)$, PSP response kernels provide synaptic contributions to mean population membrane potentials, while the sigmoidal firing rate function $S(.)$ produces the output firing rates of the stellate (purple), pyramidal (aqua) and inhibitory (red) populations. The connectivity strengths between populations are captured by c_1, c_2, c_3 and c_4. (C) UKF parameter estimation example for a forward simulation of the JR model with time-varying parameters. Top left: DC-filtered version of the JR model EEG output y_t; Top right: true IPSP amplitude time series B_t (red) and its estimates $\hat{B}_t$ (black); Bottom Left: true IPSP rate constant time series b_t (red) and its estimates $\hat{b}_t$ (black); Bottom right: true input time series μ_t (red) and its estimates $\hat{\mu}_t$ (black).

concentration increases. The sixth subplot shows how the estimate of the OAA/S score $\hat{O}_t$, and its moving average over 30 seconds $\bar{O}_t$, decreases as the actual OAA/S score decreases.

To verify the parameter estimates obtained in Figure 2(A), amplitude spectra of forward simulations of the model for the average parameter estimates corresponding to each OAA/S level were compared to the amplitude spectra of the data, y_t, and the data estimated by the UKF, $\hat{y}_t$, corresponding to each OAA/S level. Figure 2(B) plots this comparison where the forward simulations are performed by setting the average parameter estimates for each OAA/S level to be the actual parameter values used in the forward simulations. It can be seen that the average amplitude spectra of forward simulations (red) provides a reasonable match to the average amplitude spectra for the data (blue) for each OAA/S level by capturing the main shape and or alpha frequency peak. Note there is strong
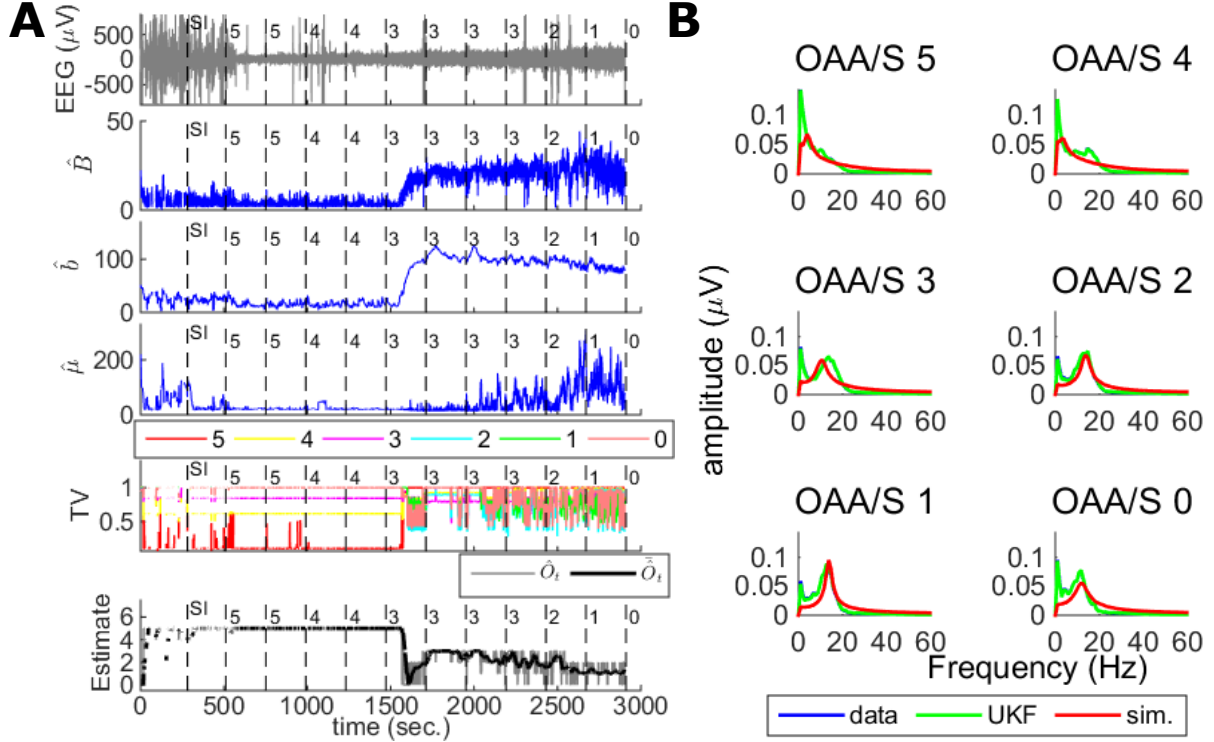
Figure 2: (A) Example of tracking depth of anesthesia using the UKF-based JR model parameter distribution approach. 1st row: Raw EEG signal for subject 9. 2nd row: IPSP amplitude estimate $\hat{B}$. 3rd row: IPSP rate constant estimate $\hat{b}$. 4th row: Input estimate $\hat{\mu}$. 5th row: Total variation time series for each OAA/S score. 6th row: OAA/S score estimate. Vertical dashed lines and numbers indicate time of OAA/S measurment and OAA/S score, respectively. SI: start of anesthetic induction. White gaps in all graphs indicate periods where artifact has been removed. TV: Total variation. (B) Amplitude spectra for the EEG data y_t (blue), the UKF data estimates $\hat{y}_t$ (green) and forward simulations using the average estimated parameters (red) for each OAA/S level for the same data analysis in (A).

overlap between the average amplitude spectra for the data (blue) and the UKF estimate of the data (green). This is why the green UKF trace makes it difficult to see the blue data trace. The reason for this strong overlap is because the UKF obtains the parameter estimates by trying to set its estimate of the data to be as close as possible to the actual EEG data signal. Therefore the best way to verify the parameter estimates is instead through the amplitude spectra of forward simulations.

3.3. Comparison of OAA/S classification approaches

The training and testing OAA/S classification performance for the comparison of the JR and ARMA model parameter estimate and HFD approaches within the distribution-based framework is shown in Tables 3 and 4, respectively. These results correspond to the three-fold cross-validation described in the methods. Here performance evaluation focuses on total sensitivity, as well as sensitivity for each OAA/S class. Sensitivities are given as the proportion of correctly classified OAA/S epochs and thus take values between 0 and 1. Given that there are six OAA/S classes, the chance level performance for each OAA/S score sensitivity is $1/6 = 0.17$. The 'nearest neighbour' sensitivity is also shown because there is strong overlap in the distributions of estimated parameters for adjacent

OAA/S scores and nearest neighbour sensitivity accounts for this. Nearest neighbour sensitivity is defined by checking if the current OAA/S score estimate is within range of the true OAA/S score ± 1. In Tables 3 and 4 the rows with bold sensitivities correspond to the best performing UKF method. In the following, training sensitivity and testing sensitivity refer to the sensitivities obtained for the training and testing sets, respectively.

Generally, the UKF method with the distribution dependent on the estimated parameters B and b and based on the second artifact rejection approach involving the RMS amplitude and Lilliefors test calculated on the 1000 Hz sampled signal gave the better testing sensistivity of the UKF based approaches with an average total testing sensitivity of 0.51 ± 0.03 (mean $\pm$ standard deviation; 'average' refers to average over cross-validation sets; 'total' refers to total over all OAA/S classes) and an average total nearest neighbour testing sensitivity of 0.75 ± 0.05. By comparison the same UKF method based on the first artifact rejection approach involving the RMS amplitude and Lilliefors test calculated on the 80 Hz sampled signal gave an average total testing sensitivity of 0.42 ± 0.02 and an average total nearest neighbour testing sensitivity of 0.69 ± 0.08.

Of all methods considered, the (2,1) order ARMA model gave the best overall performance with an average total testing sensitivity of 0.58 ± 0.05 and an average total nearest neighbour testing sensitivity of 0.91 ± 0.06. On the other hand, the HFD approach obtained an average total testing sensitivity of 0.50 ± 0.02, which is close to the best UKF case, and an average total nearest neighbour sensitivity of 0.88 ± 0.01, which is better than the best UKF case. Comparing Tables 3 and 4 it can be noted that the training sensitivtiy is generally better than the testing sensitivity as is expected.

3.4. Monotonicity, observability and error analysis of the JR model/UKF approach

The monotonicity analysis was performed to see if parameter estimates from real data vary in the physiologically known direction as propofol concentration increases when considering regions of parameter space other than the standard JR model parameters. This monotonicity analysis of the UKF parameter estimates across a large region of the JR model parameter space (spanning A, a and $\hat{b}_0$) is captured in Figure 3(A). In the figure the monotonicity index Ω is displayed for each point in the parameter space for which the UKF has been applied to estimate the JR model parameters using the EEG data from subject 9. It can be seen that the monotonicity index is closest to 1 in the yellow regions of the monotonicity maps. Therefore these regions are expected to give rise to parameter estimates for the IPSP amplitude B and the IPSP rate constant b that vary in the known physiological direction.

However, if one considers the MSE of the parameter estimates for JR model forward simulations shown in Figure 3(B) and effectively spanning the same parameter space as the monotonicity analysis, one can see that the MSE is lowest (most yellow) in the region close to the standard JR model parameters ($A = 3.25$, $a = 100$, $b = 50$). This suggests that the parameter regions shown to produce parameter estimates for the IPSP amplitude B and the IPSP rate constant b that vary in the known physiological direction may not be providing reliable parameter estimates.

The analysis of the observability index for the augmented-state JR model captured in Figure 3(C) further confirms this idea demonstrating that the nonlinear observability of the augmented-state is highest (most yellow) in a region close to the standard JR model parameters.

Table 3: Average training sensitivity of the distribution approach.

Method	Distrib. params.	**Sensitivity** OAA/S class						
		Tot.	0	1	2	3	4	5
UKF	B,b,μ	0.50	0.43	0.67	0.61	0.29	0.10	0.85
UKF	B,b	0.49	0.38	0.67	0.56	0.38	0.06	0.85
UKF	B	0.47	0.43	0.67	0.67	0.19	0.00	0.86
UKF	b	0.43	0.10	0.83	0.50	0.10	0.06	0.85
UKF2	B,b,μ	**0.57**	**0.71**	**1**	**0.66**	**0.29**	**0.23**	**0.81**
UKF2	B,b	0.55	0.81	1	0.67	0.10	0.20	0.81
ARMA	a^1,a^2,b^1	0.69	0.86	1	0.5	0.76	0.39	0.87
HFD	HFD	0.58	0.62	0.67	0.22	0.43	0.38	0.82

Method	Distrib. params.	**Nearest Neighbour Sensitivity** OAA/S class						
		Tot.	0	1	2	3	4	5
UKF	B,b,μ	0.71	0.67	0.83	0.89	0.43	0.58	0.85
UKF	B,b	0.69	0.57	0.83	0.89	0.52	0.52	0.85
UKF	B	0.70	0.52	0.83	0.83	0.52	0.47	0.92
UKF	b	0.67	0.67	1	0.89	0.48	0.40	0.85
UKF2	B,b,μ	**0.81**	**0.81**	**1**	**0.83**	**0.52**	**0.71**	**0.94**
UKF2	B,b	0.76	0.86	1	0.78	0.38	0.60	0.93
ARMA	a^1,a^2,b^1	0.92	0.95	1	0.72	0.90	0.90	0.96
HFD	HFD	0.88	0.95	1	0.61	0.71	0.85	0.98

Distrib. params.: parameters used to define the distributions; UKF: unscented Kalman filtering; UKF2: UKF using the second artifact rejection approach with RMS amplitude and Lilliefors test calculated on the 1000 Hz sampled signal; ARMA: autoregressive moving average; HFD: Higuchi fractal dimension.

To further assess if the parameter regions that give a high monotonicity index actually do coincide with unreliable parameter estimates, an example UKF estimation case that gave a high monotonicity index ($A = 1.25$, $a = 50$, $b = 300$) was considered and evaluated on control resting eyes-closed data. Figure 4(A) demonstrates the UKF parameter estimation result for this case for subject 9. It can be seen that the IPSP amplitude estimate $\hat{B}$ increases while the IPSP rate constant estimate $\hat{b}$ decreases as is desired. However, when the same UKF approach is applied to the resting control data the same variations in the parameter estimates occur. Thus demonstrating that these parameter estimates obtained in this region of parameter space are not useful for depth of anesthesia monitoring. This result can be attributed to the fact that the sigmoid firing rate function leads to a saturation of estimated states in this region of parameter space and this saturation reduces the ability of the UKF to infer the states (results not shown).

Table 4: Average testing sensitivity of the distribution approach.

| Method | Distrib. params. | **Sensitivity** | | | | | | |
		Tot.	OAA/S class					
			0	1	2	3	4	5
UKF	B, b, μ	0.41	0.22	0.17	0.50	0.04	0.03	0.90
UKF	B, b	0.42	0.28	0.17	0.50	0.04	0.04	0.90
UKF	B	0.43	0.39	0.17	0.44	0.00	0.06	0.91
UKF	b	0.41	0.22	0.17	0.56	0.04	0.03	0.90
UKF2	B, b, μ	0.49	0.50	0.17	0.50	0.07	0.22	0.87
UKF2	B, b	**0.51**	**0.67**	**0.17**	**0.61**	**0.11**	**0.21**	**0.87**
ARMA	a^1, a^2, b^1	0.58	0.72	0.17	0.28	0.33	0.49	0.82
HFD	HFD	0.50	0.56	0.00	0.00	0.26	0.36	0.83
Method	Distrib. params.	**Nearest Neighbour Sensitivity**						
		Tot.	OAA/S class					
			0	1	2	3	4	5
UKF	B, b, μ	0.71	0.61	1	0.89	0.56	0.48	0.91
UKF	B, b	0.69	0.56	1	0.78	0.59	0.47	0.91
UKF	B	0.69	0.67	1	0.72	0.41	0.48	0.93
UKF	b	0.68	0.61	1	0.94	0.44	0.43	0.91
UKF2	B, b, μ	0.**77**	**0.56**	**0.67**	**0.83**	**0.63**	**0.71**	**0.92**
UKF2	B, b	0.75	0.72	0.67	0.83	0.52	0.66	0.90
ARMA	a^1, a^2, b^1	0.91	1	1	0.78	0.96	0.85	0.95
HFD	HFD	0.88	0.94	0.83	0.67	0.78	0.86	0.96

Distrib. params.: parameters used to define the distributions; UKF: unscented Kalman filtering; UKF2: UKF using the second artifact rejection approach with RMS amplitude and Lilliefors test calculated on the 1000 Hz sampled signal; ARMA: autoregressive moving average; HFD: Higuchi fractal dimension.

3.5. Error analysis of extensions of the JR model

The error analysis of extensions of the JR model considered whether the inclusion of inhibition of inhibition could produce parameter trajectories in the extended model(s) that follow physiologically expected trajectories for propofol anesthesia (i.e. an increase in IPSP amplitude and a decrease in IPSP rate constant). Figure 5 captures this analysis for the first and second extensions of the JR model depicted in Figure 5(A) and (B), respectively.

For the first extension, Figure 5(C) displays the maps of the OAA/S score linked to the average data amplitude spectrum that best matches the average amplitude spectrum for multiple forward simulations for each point in the considered parameter subspace spanned by the IPSP amplitude B and IPSP rate constant b. It can be seen that for $c_5 = 0$ (corresponds to no inhibition of inhibition and is equivalent to the JR model) that there is a part of the parameter space (along $B = 22$) where the best matching OAA/S levels transition from 3 (aqua), to 2 (green), to 1 (yellow), to 0 (red) as the rate constant b decreases. This could be considered a potential match to physiology, however, as the rate constant continues to decrease the OAA/S level 5 begins to provide the best amplitude spectrum match to the model. Moreover, as the level of inhibition of inhibition increases the region over which interesting var-

A Monotonicity Index
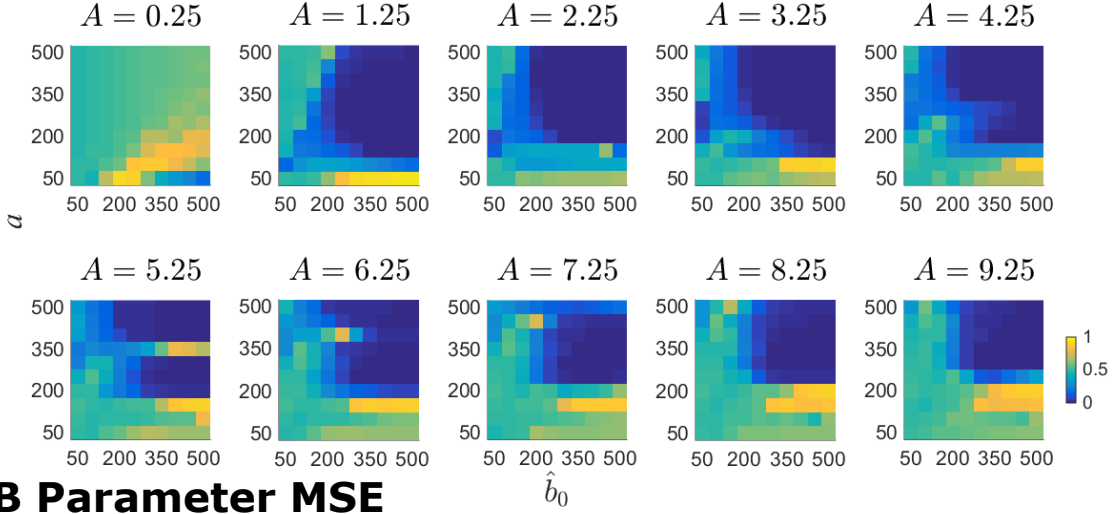
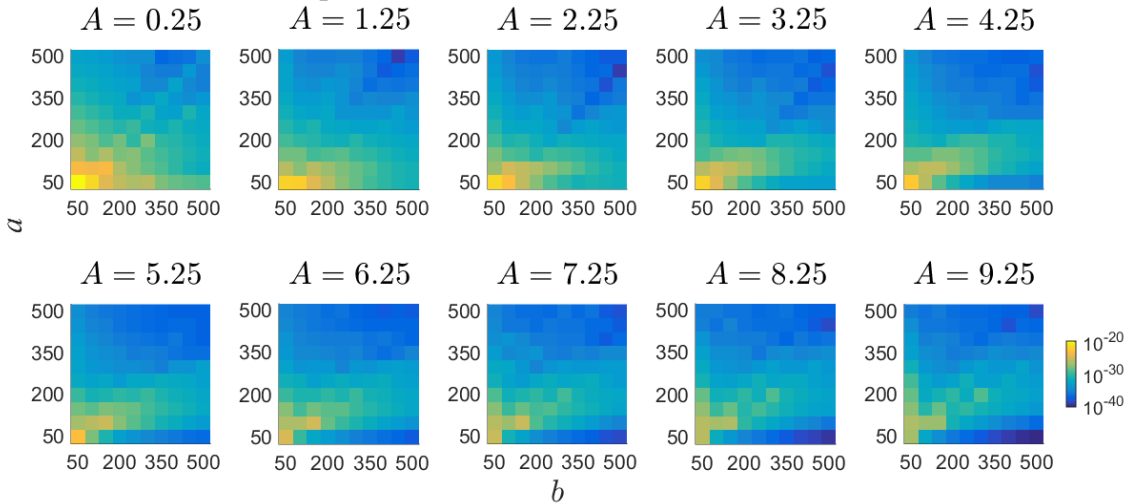# B Parameter MSE
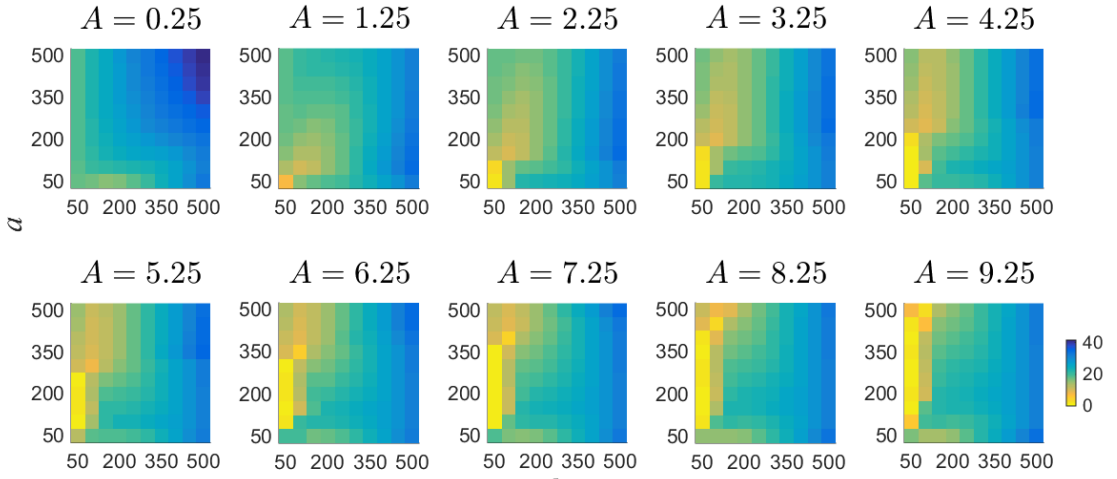
C Observability Index

Figure 3: (A) Maps of the monotonicity index, Ω, spanning a region of the JR model parameter space constrained by A, a and $\hat{b}_0$. The colorbar indicates the strength of the monotonicity index. (B) Maps of the MSE of the parameter estimates of forward simulations of the JR model spanning a region of the JR model parameter space constrained by A, a and b. The colorbar indicates the magnitude of the MSE. (C) Maps of the nonlinear observability index of the augmented-state JR model spanning a region of the JR model parameter space constrained by A, a and b. The colorbar indicates the magnitude of the observability index.
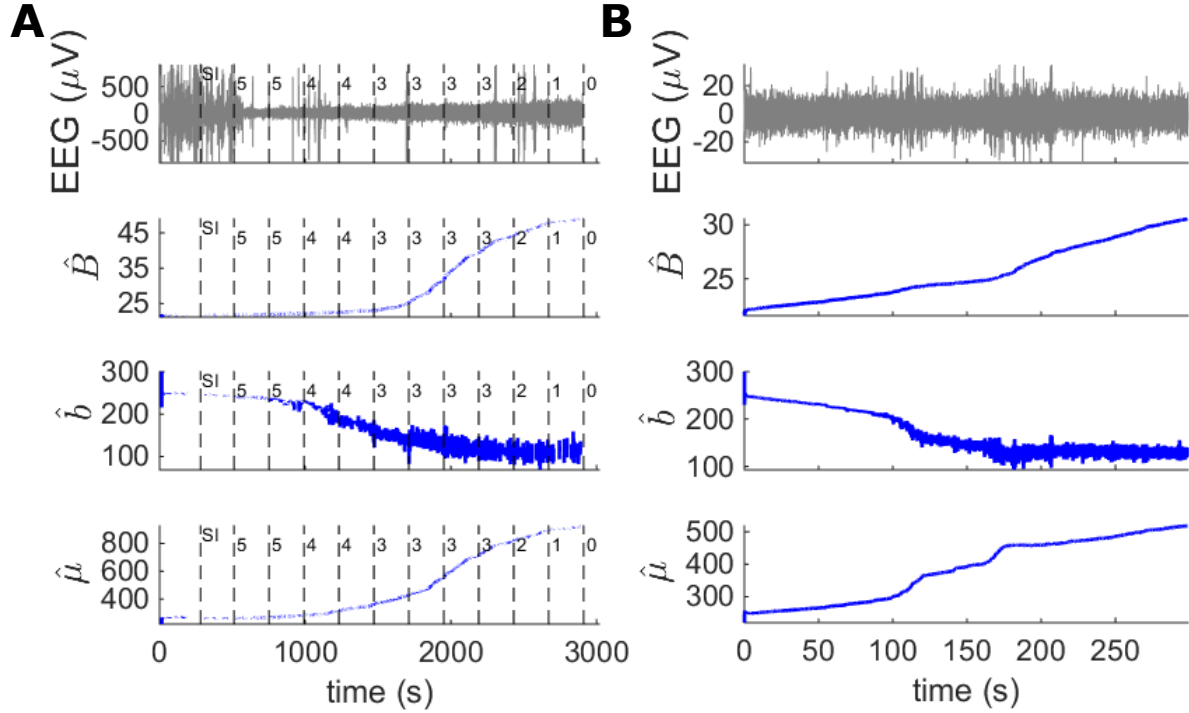
Figure 4: Example UKF estimation case that gave a high monotonicity index ($A = 1.25$, $a = 50$, $b = 300$). (A) UKF parameter estimation results obtained with the data from subject 9. (B) UKF parameter estimation results obtained with control resting-eyes closed data. In (A) and (B) - 1st row: EEG signal; 2nd row: IPSP amplitude estimate $\hat{B}$; 3rd row: IPSP rate constant estimate $\hat{b}$; 4th row: Input estimate $\hat{\mu}$. White gaps in all graphs indicate periods where artifact has been removed. In (A) vertical dashed lines and numbers indicate time of OAA/S measurment and OAA/S score, respectively. SI: start of anesthetic induction.

-iations in the OAA/S map occur become smaller and smaller.

For the second extension which includes the additional IPSP kernel, Figure 5(D) displays the maps of the OAA/S score linked to the average data amplitude spectrum that best matches the average amplitude spectrum for multiple forward simulations for each point in the considered parameter subspace spanned by the IPSP rate constant parameters m and b. Similar features are observed here as are seen for the first extension and inhibition of inhibition primarily contributes to reducing the region over which interesting variations in the OAA/S map occur.

4. Discussion

The key findings of this study are that the UKF applied to the JR model approach produces reasonable OAA/S score classification results when compared to a high performing standard depth of anesthesia monitoring feature HFD, however, its performance is inferior to the ARMA model approach. Therefore the UKF applied to the JR model provides a benchmark for future approaches that use neural mass or neural fields models to track anesthetic state while also providing physiologically meaningful information about underlying physiological parameter changes. Moreover, although the JR model can potentially be used to track anesthetic brain state, the parameter estimates derived from this model and its extensions that include inhibition of inhibition may not provide clear physiological
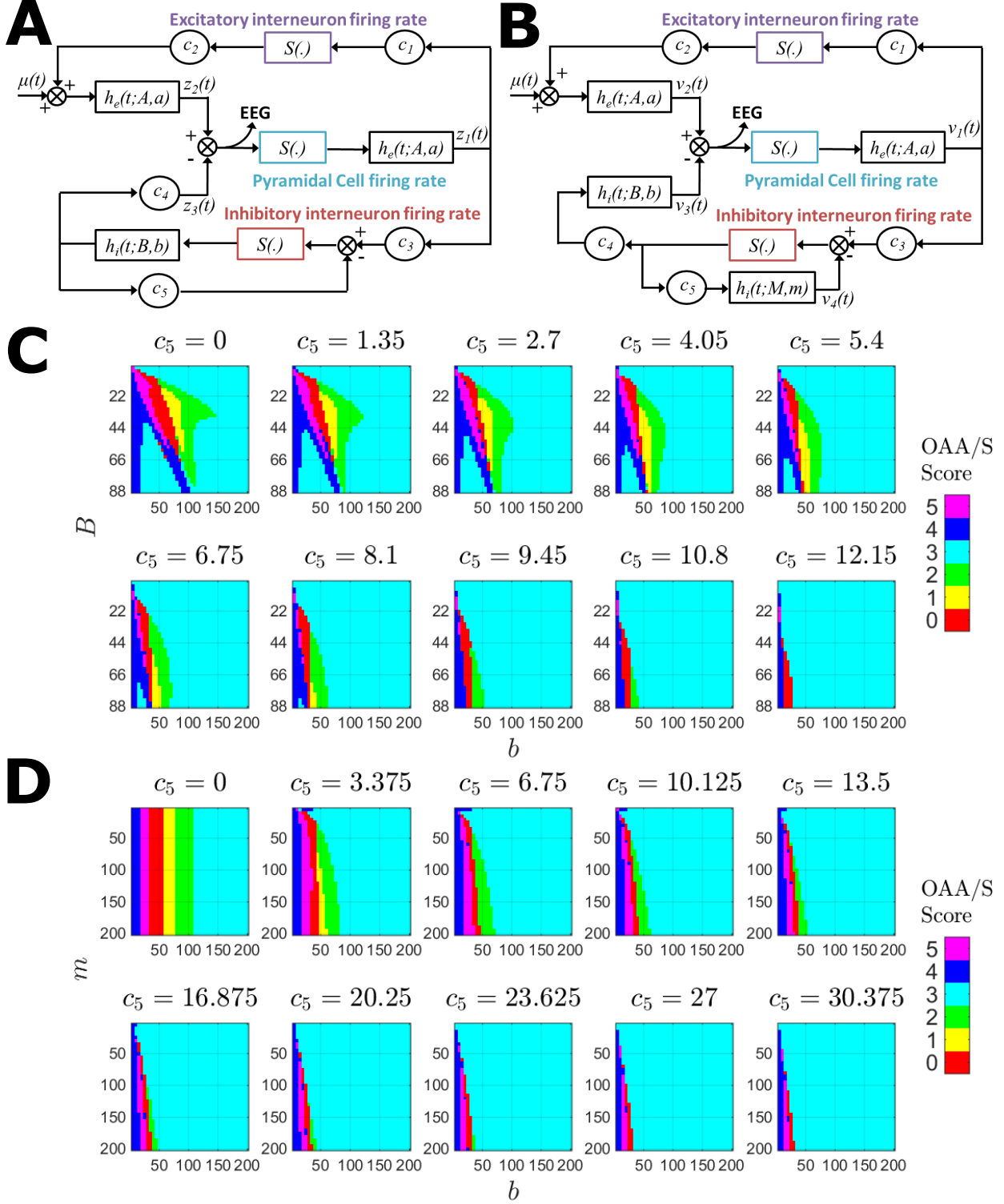
Figure 5: The block diagrams of the (A) first and (B) second extension of the JR model. In (A) the primary variation on the JR model is the additional feedback connection from the output of the IPSP kernel to the input of the inhibitory population via the connection weight c_5. In (B) the primary variation on the JR model is the additional feedback connection from the output of the inhibitory population to the input of the inhibitory population via the connection weight c_5 and an additional IPSP kernel, $h_i(t; M, m)$.(C) Best matching OAA/S score maps for the first extension of the JR model for the parameter subspace spanned by the IPSP amplitude B, IPSP rate constant b and the degree of inhibition of inhibition c_5. (D) Best matching OAA/S score maps for the second extension of the JR model for the parameter subspace spanned by the IPSP rate constants m and b ($B = 22$) and the degree of inhibition of inhibition c_5. In (C) and (D) the color code indicates the OAA/S score linked to the average data amplitude spectrum from subject 9 that best matches the average amplitude spectrum for multiple forward simulations for each given point in the considered parameter subspace.

22

interpretation, in particular in terms of the IPSP rate constant. Therefore future work with more complex, or appropriately designed, models (Bojak & Liley, 2005; Bojak et al., 2015) is warranted to devise a method that can both efficiently track anesthetic brain state and simultaneously infer realistic physiological changes linked to anesthesia.

4.1. OAA/S score classification

Although the UKF applied to the JR model approach produced reasonable OAA/S score classification testing sensitivty (0.51 ± 0.03) compared to the HFD (0.50 ± 0.02), the UKF approach was still weaker when considering nearest neighbour testing sensitivity (0.75 for UKF vs 0.88 for HFD). Moreover, HFD and the ARMA approach were evaluated on the first artifact rejection strategy where RMS amplitude and Lilliefors test were calculated on the 80 Hz sampled signal, while the best performing UKF case was evaluated on the second artifact rejection strategy where RMS amplitude and Lilliefors test were calculated on the 1000 Hz sampled signal. The second strategy led to greater artifact rejection of epochs and this may have contributed to improving the separability of the different OAA/S distributions. The key point, however, is that the UKF performs worse than the HFD and ARMA approach when the first artifact rejection strategy is applied. The weaker performance of the UKF, in particular with respect to the ARMA model (testing sensitivity of 0.51 for UKF vs 0.58 for ARMA), can be understood in part by the decision to only track the IPSP amplitude B, IPSP rate constant b and the input μ of the JR model-based on the physiological knowledge that the IPSP amplitude and rate constant are effected by propofol concentration (Kitamura et al., 2003; Wang et al., 2004; Bieda & MacIver, 2004; Ying & Goldstein, 2005; McCarthy et al., 2008; Wang, 2009; Yue et al., 2011; Wakita et al., 2013; Hutt, 2013). This decision puts more contraints on the UKF estimation process to fit the data. On the other hand, the ARMA model coefficients essentially capture the most significant aspects of the data and this leads to better classification accuracy.

It would be possible to use the UKF to track more parameters in order to produce a better performing UKF based classifier of the OAA/S score. We did explore the UKF applied to real data to track the EPSP and IPSP amplitudes and rate constants and sigmoid threshold of the JR model, but in this case the IPSP parameters did not both move in the desired directions and the case was not considered further (results not shown). Moreover, the goal here is not just to track anesthetic state, but rather to begin with a simple cortical model, the JR model, and see if it can simultaneously be used to track anesthetic brain state while also estimating realistic physiological changes in parameters that may provide additional information in the clinic. Here the JR model has been evaluated for this purpose and will provide a valuable benchmark for future studies with more complex and realistic models that are also simple enough to be employed efficiently in the clinical environment.

4.2. Monotonicity, observability and error analysis of the UKF/JR model

The monotonicity, observability and error analysis revealed that for the parameter subspace considered there are greater errors in UKF based estimation of JR model parameters when the parameters and their estimates don't operate close to the standard JR model parameters (Jansen & Rit, 1995). This can be understood in part by the fact that the standard parameters correspond to a region in parameter space involving an alpha rhythm limit cycle

(Jansen & Rit, 1995) and state observability is generally higher when limit cycle dynamics are present as compared say to a stable fixed point (Whalen et al., 2015). These results highlight the importance of considering monotonicity, observability and error analysis when investigating other neural mass model inversion procedures. Moreover, the demonstration that parameter estimates (IPSP amplitude and rate constant) that vary in the desired direction for anesthesia data can also move in the same direction with resting eyes-closed data also highlights the general importance for controlling for parameter estimate changes using control data.

4.3. JR model extensions, other models and realistic physiological inference

The exploration of the JR model extensions involving inhibition of inhibition revealed that the JR model without inhibition of inhibition already contains an interesting region of parameter subspace where if the IPSP amplitude is fixed then the IPSP rate constant would be expected to decrease as the OAA/S level shifts from 3 to 0 with an increase in anesthetic level as is observed physiologically. However, further descreases in the IPSP rate constant would then lead to a preferred match with an OAA/S score of 5. This suggests that the JR model could potentially be modified in some way to avoid this 'wrap around', however, it is not clear at present how to do this. The addition of inhibition of inhibition only led to a reduction of the region over which interesting variations of the best matching OAA/S score linked data spectra and model data spectra could be seen. Therefore this is not a likely full solution.

Local cortical inhibition of inhibition was demonstrated to be important in a thalamocortical model of propofol anesthesia (Hindriks & van Putten, 2012), in particular with respect to controlling the amplitude of the alpha rhythm. The results presented here are consistent with this, however, there are also critical differences between these models. In particular, the JR model is only a cortical model. Thalamocortical loops have been suggested to be important in creating the frontal alpha rhythm seen in propofol anesthesia (Ching et al., 2010), however, there are also cortical level models of anesthesia that can capture many of the important features of anesthesia (Bojak & Liley, 2005; Steyn-Ross et al., 2012), including propofol anesthesia (Hutt, 2013; Hutt & Buhry, 2014). Therefore the scope is currently open to evaluate the ability of new and existing models of anesthesia to find a model that is simple enough to be computationallly efficient, but complex enough for it to be able to characterise the key physiological changes linked to anesthesia within brain state tracking applications.

For the JR model, when the UKF was applied to real EEG data near the standard JR model parameter region, the UKF estimate of the IPSP amplitude increased with propofol concentration as is observed physiologically, while the IPSP rate constant increased also, which is inconsistent with known physiology. The actual degree of variation of IPSP parameters with propofol concentration observed physiologically is highly dependent on the brain region studied, the neurotransmittor receptors considered, and the experimental preparation, whether it be cell cultures or isolated cells (Krasowski et al., 1998; Bai et al., 1999, 2001; Siegwart et al., 2002; Dong & Xu, 2002; Feng & Macdonald, 2004; Drexler et al., 2009; Yue et al., 2011; Wakita et al., 2013), or tissue slices (Collins, 1988; Bieda & MacIver, 2004; Wang et al., 2004; Ying & Goldstein, 2005; Wang, 2009). In tissue slices, measurements focus more directly on PSP-level parameters and appear to indicate a greater sensitivity to propofol of the IPSP rate constant than for IPSP amplitude, however, there are still observed variations in both parameters depending on the preparation (Collins, 1988; Bieda & MacIver, 2004; Wang et al., 2004; Ying & Goldstein, 2005; Wang, 2009). The JR

24

model IPSP amplitude and rate constant parameters are lumped population level parameters that can be considered a combination of the various kinds of inhibitory receptors and subtypes influenced by propofol at the molecular scale (Rudolph & Antkowiak, 2004). Therefore it is expected that population models that take into account different inhibitory receptor types may be able to better characterise the effects of propofol on the EEG, however, such model extensions may compromise the goal of having the simplest possible model that is cliniclially useful while at the same time adequately characterises the data.

In a validation study of dynamical causal modelling applied to isoflurane anesthesia and involving a networked variation of the JR model (Moran et al., 2011) it was found that the parameter estimates of the EPSP and IPSP amplitude decreased and increased, respectively, as was observed in rats who underwent isoflurane anesthesia. The results presented here for the JR model demonstrating the IPSP amplitude increase for increases in propofol concentration are consistent with this finding, especially as isoflurane and propofol are believed to act through GABAergic receptors (Rudolph & Antkowiak, 2004; Garcia et al., 2010). It is also known, however, that like propofol, isoflurane also decreases the IPSP rate constant (Bojak & Liley, 2005). Therefore there is still scope to perform realistic physiological parameter inference using other models and considering different anesthetics with different molecular modes of action.

4.4. Other estimation approaches

Here the OAA/S classification/tracking performance of the UKF applied to the JR model was primarily derived using the standard JR model parameters and setting the initial parameter estimates to the standard values. Then the algorithm was let to run on the data and the parameter estimates were allowed to equilibrate to the data. This approach was applied as it would simplify the setup time in the clinic. An alternative approach could be to apply a dynamic causal modelling (Kiebel et al., 2009), particle swarm (Bojak & Liley, 2005) or other parameter estimation (Sorenson, 1980) approach to a short window of data in order to find the parameters (not necessarily just IPSP amplitude, IPSP rate constant and the input) that best fit the subject's resting data. Then allow the UKF to track variations of key parameters of interest starting from the region of parameter space found by the analysis of the resting segment.

An approach like this was recently applied using a new efficient window-based frequency-domain method of dynamical causal modelling involving Bayesian belief updating (Cooray et al., 2015). This algorithm was applied to continuous univariate EEG time series from epilepsy patients and provides an interesting alternative to compare with the time-domain UKF approach in future studies, in particular regarding computational efficiency. The run time of the UKF including artifact rejection for subject 9 in Figure 2 for 50.3 minutes of data is 13.1 minutes with an i7-3520M 2.9 GHz CPU and 8 GB of RAM, and the code has not been optimised. This indicates that the UKF framework presented here can be run in real-time applications. The Bayesian belief updating algortihm has been shown to run 10-20 times faster than standard dynamical causal modelling algortihms (Cooray et al., 2015), and therefore is a likely candidate for real-time applications also.

To apply the UKF to real data in this study the EEG signal mean and standard deviation was rescaled to match that of the JR model output. It may be possible to avoid this transformation by applying an appropriate high

pass filter to the JR model output. This was considered by augmenting the state of the JR model with a first order butterworth filter, and it was found that UKF parameter estimation worked for simulated data but was not reliable for the real EEG data (results not shown). Therefore this remains a topic of future research that could be solved through combined hardware/software solutions where the parameters of the EEG acquisition filters are known exactly. Here a commercial system was used for data collection and the exact filter specifications were not available.

Finally, it is worthwhile noting that the UKF approach presented here is similar in flavor to previous applications of Kalman filtering to multi-channel EEG source imaging (Galka et al., 2004; Barton et al., 2009), however, these approaches work with linear models while the UKF presented here deals with nonlinear models. It may be the case that multi-variate applications of the UKF to neural mass models (Freestone et al., 2014) or multi-model supervisory observers (Chong et al., 2015) can better characterise anesthetic brain state transitions, however, there are still difficulties with using multi-channel EEG in the clinical general anesthesia setting.

5. Conclusion

Tracking the state of the brain during anesthesia is a challenging problem and physiologically motivated approaches have the potential to provide improved solutions. Here it has been demonstrated using the UKF that the JR model can track anesthetic brain state with reasonably accuracy, however, more complex, or appropriately designed, models of anesthesia need to be investigated to determine if neural model-based approaches can outperform other model-based and standard depth of anesthesia monitoring methods, while simultaneously providing useful information about underlying physiological variable changes linked to anesthesia.

6. Acknowledgements

References

Accardo, A., Affinito, M., Carrozzi, M., & Bouquet, F. (1997). Use of the fractal dimension for the analysis of electroencephalographic time series. *Biological Cybernetics*, *77*, 339–350.

Bai, D., Pennefather, P. S., MacDonald, J. F., & Orser, B. A. (1999). The general anesthetic propofol slows deactivation and desensitization of gabaa receptors. *The Journal of Neuroscience*, *19*, 10635–10646.

Bai, D., Zhu, G., Pennefather, P., Jackson, M. F., MacDonald, J. F., & Orser, B. A. (2001). Distinct functional and pharmacological properties of tonic and quantal inhibitory postsynaptic currents mediated by γ-aminobutyric acida receptors in hippocampal neurons. *Molecular pharmacology*, *59*, 814–824.

Barton, M. J., Robinson, P., Kumar, S., Galka, A., Durrant-Whyte, H. F., Guivant, J., Ozaki, T. et al. (2009). Evaluating the performance of kalman-filter-based eeg source localization. *Biomedical Engineering, IEEE Transactions on*, *56*, 122–136.

Bieda, M. C., & MacIver, M. B. (2004). Major role for tonic gabaa conductances in anesthetic suppression of intrinsic neuronal excitability. *Journal of neurophysiology*, *92*, 1658–1667.

Bojak, I., & Liley, D. (2005). Modeling the effects of anesthesia on the electroencephalogram. *Physical Review E*, *71*, 041902.

Bojak, I., Stoyanov, Z. V., & Liley, D. T. (2015). Emergence of spatially heterogeneous burst suppression in a neural field model of electrocortical activity. *Frontiers in systems neuroscience*, *9*.

Broersen, P. (2006). *Automatic Autocorrelation and Spectral Analysis*. London: Springer-Verlag.

Broersen, P. M. (2002). Automatic spectral analysis with time series models. *Instrumentation and Measurement, IEEE Transactions on*, *51*, 211–216.

Bruhn, J., Myles, P., Sneyd, R., & Struys, M. (2006). Depth of anaesthesia monitoring: what's available, what's validated and what's next? *British journal of anaesthesia*, *97*, 85–94.

Chernik, D. A., Gillings, D., Laine, H., Hendler, J., Silver, J. M., Davidson, A. B., Schwam, E. M., & Siegel, J. L. (1990). Validity and reliability of the observer's: Assessment of alertness/sedation scale: Study with: Intravenous midazolam. *Journal of clinical psychopharmacology*, *10*, 244–251.

Ching, S., Cimenser, A., Purdon, P. L., Brown, E. N., & Kopell, N. J. (2010). Thalamocortical model for a propofol-induced α-rhythm associated with loss of consciousness. *Proceedings of the National Academy of Sciences*, *107*, 22665–22670.

Chong, M., Postoyan, R., Nesic, D., Kuhlmann, L., & Varsavsky, A. (2011). A nonlinear estimator for the activity of neuronal populations in the hippocampus. In *IFAC World Congress* (pp. 9899–9904). volume 18.

Chong, M., Postoyan, R., Nešić, D., Kuhlmann, L., & Varsavsky, A. (2012a). Estimating the unmeasured membrane potential of neuronal populations from the eeg using a class of deterministic nonlinear filters. *Journal of neural engineering*, *9*, 026001.

Chong, M., Postoyan, R., Nešić, D., Kuhlmann, L., & Varsavsky, A. (2012b). A robust circle criterion observer with application to neural mass models. *Automatica*, *48*, 2986–2989.

Chong, M. S., Nešić, D., Postoyan, R., & Kuhlmann, L. (2015). Parameter and state estimation of nonlinear systems using a multi-observer under the supervisory framework. *IEEE Transactions on Automatic Control*, *60*, 2336–2349.

Cimenser, A., Purdon, P. L., Pierce, E. T., Walsh, J. L., Salazar-Gomez, A. F., Harrell, P. G., Tavares-Stoeckel, C., Habeeb, K., & Brown, E. N. (2011). Tracking brain states under general anesthesia by using global coherence analysis. *Proceedings of the National Academy of Sciences*, *108*, 8832–8837.

27

Collins, G. (1988). Effects of the anaesthetic 2, 6-diisopropylphenol on synaptic transmission in the rat olfactory cortex slice. *British journal of pharmacology*, *95*, 939–949.

Cooray, G. K., Sengupta, B., Douglas, P. K., & Friston, K. (2015). Dynamic causal modelling of electrographic seizure activity using bayesian belief updating. *NeuroImage*, .

Deco, G., Jirsa, V. K., Robinson, P. A., Breakspear, M., & Friston, K. (2008). The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS computational biology*, *4*.

Dong, X.-P., & Xu, T.-L. (2002). The actions of propofol on γ-aminobutyric acid-a and glycine receptors in acutely dissociated spinal dorsal horn neurons of the rat. *Anesthesia & Analgesia*, *95*, 907–914.

Drexler, B., Jurd, R., Rudolph, U., & Antkowiak, B. (2009). Distinct actions of etomidate and propofol at $\beta3$-containing γ-aminobutyric acid type a receptors. *Neuropharmacology*, *57*, 446–455.

Feng, H.-J., & Macdonald, R. L. (2004). Multiple actions of propofol on $\alpha\beta\gamma$ and $\alpha\beta\delta$ gabaa receptors. *Molecular pharmacology*, *66*, 1517–1524.

Ferenets, R., Vanluchene, A., Lipping, T., Heyse, B., & Struys, M. M. (2007). Behavior of entropy/complexity measures of the electroencephalogram during propofol-induced sedation: dose-dependent effects of remifentanil. *Anesthesiology*, *106*, 696–706.

Foster, B. L., & Liley, D. T. (2011). Nitrous oxide paradoxically modulates slow electroencephalogram oscillations: implications for anesthesia monitoring. *Anesthesia & Analgesia*, *113*, 758–765.

Freestone, D., Aram, P., Dewar, M., Scerri, K., Grayden, D., & Kadirkamanathan, V. (2011). A data-driven framework for neural field modeling. *NeuroImage*, *56*, 1043–1058.

Freestone, D., Kuhlmann, L., Chong, M., Nesic, D., Grayden, D. B., Aram, P., Postoyan, R., Cook, M. J. et al. (2013). Patient-specific neural mass modelling: stochastic and deterministic methods. *Recent Advances in Predicting and Preventing Epileptic Seizures*, (pp. 63–82).

Freestone, D. R., Karoly, P. J., Nešić, D., Aram, P., Cook, M. J., & Grayden, D. B. (2014). Estimation of effective connectivity via data-driven neural modeling. *Frontiers in neuroscience*, *8*.

Galka, A., Yamashita, O., Ozaki, T., Biscay, R., & Valdés-Sosa, P. (2004). A solution to the dynamical inverse problem of eeg generation using spatiotemporal kalman filtering. *NeuroImage*, *23*, 435–453.

Garcia, P. S., Kolesky, S. E., & Jenkins, A. (2010). General anesthetic actions on gabaa receptors. *Current neuropharmacology*, *8*, 2.

Hermann, R., & Krener, A. J. (1977). Nonlinear controllability and observability. *IEEE Transactions on automatic control*, *22*, 728–740.

Higuchi, T. (1988). Approach to an irregular time series on the basis of the fractal theory. *Physica D: Nonlinear Phenomena*, *31*, 277–283.

Hindriks, R., & van Putten, M. J. (2012). Meanfield modeling of propofol-induced changes in spontaneous eeg rhythms. *Neuroimage*, *60*, 2323–2334.

Hirota, K. (2006). Special cases: ketamine, nitrous oxide and xenon. *Best Practice & Research Clinical Anaesthesiology*, *20*, 69–79.

Hutt, A. (2013). The anesthetic propofol shifts the frequency of maximum spectral power in eeg during general anesthesia: analytical insights from a linear model. *Frontiers in computational neuroscience*, *7*.

Hutt, A., & Buhry, L. (2014). Study of gabaergic extra-synaptic tonic inhibition in single neurons and neural populations by traversing neural scales: application to propofol-induced anaesthesia. *Journal of computational neuroscience*, *37*, 417–437.

Jansen, B., & Rit, V. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics*, *73*, 357–366.

Jordan, D., Schneider, G., Hock, A., Hensel, T., Stockmanns, G., & Kochs, E. F. (2006). Eeg parameters and their combination as indicators of depth of anaesthesia/eeg-parameter und deren kombination für das narkosemonitoring. *Biomedizinische Technik*, *51*, 89–94.

Kettner, S. (2014). Not too little, not too much: delivering the right amount of anaesthesia during surgery. *The Cochrane database of systematic reviews*, *6*, ED000084.

Kiebel, S. J., Garrido, M. I., Moran, R., Chen, C.-C., & Friston, K. J. (2009). Dynamic causal modeling for eeg and meg. *Human brain mapping*, *30*, 1866–1876.

Kissin, I. (2000). Depth of anesthesia and bispectral index monitoring. *Anesthesia & Analgesia*, *90*, 1114–1117.

Kitamura, A., Marszalec, W., Yeh, J. Z., & Narahashi, T. (2003). Effects of halothane and propofol on excitatory and inhibitory synaptic transmission in rat cortical neurons. *Journal of Pharmacology and Experimental Therapeutics*, *304*, 162–171.

Kloeden, P., & Platen, E. (1992). *Numerical solution of stochastic differential equations*. Springer–Verlag.

Krasowski, M. D., Koltchine, V. V., Rick, C. E., Ye, Q., Finn, S. E., & Harrison, N. L. (1998). Propofol and other intravenous anesthetics have sites of action on the γ-aminobutyric acid type a receptor distinct from that for isoflurane. *Molecular pharmacology*, *53*, 530–538.

Kuhlmann, L., Foster, B. L., & Liley, D. (2013). Modulation of functional eeg networks by the nmda antagonist nitrous oxide. *PloS one*, *8*, e56434.

Kuhlmann, L., Grayden, D. B., Wendling, F., & Schiff, S. J. (2015a). Role of multiple-scale modeling of epilepsy in seizure forecasting. *Journal of Clinical Neurophysiology*, *32*, 220–226.

Kuhlmann, L., Manton, J., Heyse, B., Vereecke, H. E., Lipping, T., Struys, M. M., & Liley, D. T. (2015b). Tracking electroencephalographic changes using distributions of linear models: application to depth of anesthesia monitoring. *Unpublished observations*, .

Lee, U., Ku, S., Noh, G., Baek, S., Choi, B., & Mashour, G. A. (2013). Disruption of frontal-parietal communication by ketamine, propofol, and sevoflurane. *Anesthesiology*, *118*, 1264.

Levina, E., & Bickel, P. (2001). The earth mover's distance is the mallows distance: Some insights from statistics. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on* (pp. 251–256). IEEE volume 2.

Lewis, L. D., Weiner, V. S., Mukamel, E. A., Donoghue, J. A., Eskandar, E. N., Madsen, J. R., Anderson, W. S., Hochberg, L. R., Cash, S. S., Brown, E. N. et al. (2012). Rapid fragmentation of neuronal networks at the onset of propofol-induced unconsciousness. *Proceedings of the National Academy of Sciences*, *109*, E3377–E3386.

Liley, D., Cadusch, P., & Dafilis, M. (2002). A spatially continuous mean field theory of electrocortical activity. *Network: Comput. Neural Syst.*, *13*, 67–113.

Liley, D., Sinclair, N., Lipping, T., Heyse, B., Vereecke, E., & Struys, M. (2010). Propofol and remifentanil differentially modulate frontal electroencephalographic activity. *Anesthesiology*, *113*, 1–13.

Liley, D. T., Leslie, K., Sinclair, N. C., & Feckie, M. (2008). Dissociating the effects of nitrous oxide on brain electrical activity using fixed order time series modeling. *Comput. Biol. Med.*, *38*, 1121–1130.

Lin, J. (1991). Divergence measures based on the shannon entropy. *Information Theory, IEEE Transactions on*, *37*, 145–151.

Mashour, G. A., Shanks, A., Tremper, K. K., Kheterpal, S., Turner, C. R., Ramachandran, S. K., Picton, P., Schueller, C., Morris, M., Vandervest, J. C. et al. (2012). Prevention of intraoperative awareness with explicit recall in an unselected surgical population: a randomized comparative effectiveness trial. *Anesthesiology*, *117*, 717.

McCarthy, M. M., Brown, E. N., & Kopell, N. (2008). Potential network mechanisms mediating electroencephalographic beta rhythm changes during propofol-induced paradoxical excitation. *The Journal of Neuroscience*, *28*, 13488–13504.

Moran, R., Pinotsis, D. A., & Friston, K. (2013). Neural masses and fields in dynamic causal modeling. *Frontiers in computational neuroscience*, *7*.

Moran, R. J., Jung, F., Kumagai, T., Endepols, H., Graf, R., Dolan, R. J., Friston, K. J., Stephan, K. E., & Tittgemeyer, M. (2011). Dynamic causal models and physiological inference: a validation study using isoflurane anaesthesia in rodents. *PloS one*, *6*, e22790–e22790.

Myles, P., Leslie, K., McNeil, J., Forbes, A., Chan, M., Group, B.-A. T. et al. (2004). Bispectral index monitoring to prevent awareness during anaesthesia: the b-aware randomised controlled trial. *The lancet*, *363*, 1757–1763.

Nunez, P. L., & Srinivasan, R. (2006). *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press.

Palanca, B. J. A., Mashour, G. A., & Avidan, M. S. (2009). Processed electroencephalogram in depth of anesthesia monitoring. *Current Opinion in Anesthesiology*, *22*, 553–559.

Postoyan, R., Chong, M., Nešić, D., & Kuhlmann, L. (2012). Parameter and state estimation for a class of neural mass models. In *IEEE Conference on Decision and Control* (p. Accepted). Maui, Hawaii.

Punjasawadwong, Y., Phongchiewboon, A., & Bunchungmongkol, N. (2014). Bispectral index for improving anaesthetic delivery and postoperative recovery. *The Cochrane Library*, .

Purdon, P. L., Pierce, E. T., Mukamel, E. A., Prerau, M. J., Walsh, J. L., Wong, K. F. K., Salazar-Gomez, A. F., Harrell, P. G., Sampson, A. L., Cimenser, A. et al. (2013). Electroencephalogram signatures of loss and recovery of consciousness from propofol. *Proceedings of the National Academy of Sciences*, *110*, E1142–E1151.

Rudolph, U., & Antkowiak, B. (2004). Molecular and neuronal substrates for general anaesthetics. *Nat. Rev. Neurosci.*, *5*, 709–720.

Schack, B., & Krause, W. (1995). Dynamic power and coherence analysis of ultra short-term cognitive processesa methodical study. *Brain Topogr*, *8*, 12736.

Shalbaf, R., Behnam, H., Sleigh, J. W., Steyn-Ross, A., & Voss, L. J. (2013). Monitoring the depth of anesthesia using entropy features and an artificial neural network. *Journal of neuroscience methods*, *218*, 17–24.

Shoushtarian, M., McGlade, D., Delacretaz, L., & Liley, D. (2015a). Evaluation of the brain anaesthesia response monitor during anaesthesia for cardiac surgery: a double-blind, randomised controlled trial using two doses of fentanyl. *Journal of Clinical Monitoring and Computing*, *In press.*.

Shoushtarian, M., Sahinovic, M., Absalom, A., Kalmar, A., Vereecke, H., Liley, D., & Struys, M. (2015b). Comparisons of electroencephalographically derived measures of hypnosis and antinociception in response to standardized stimuli during target-controlled propofol-remifentanil anesthesia. *Anesth Analg*, *In press.*.

Siegwart, R., Jurd, R., & Rudolph, U. (2002). Molecular determinants for the action of general anesthetics at recombinant $\alpha 2\beta 3\gamma 2\gamma$-aminobutyric acida receptors. *Journal of neurochemistry*, *80*, 140–148.

Simon, D. (2006). *Optimal State Estimation: Kalman, H-infinity and nonlinear approaches*. New Jersey, U.S.A.: Wiley.

Sorenson, H. W. (1980). *Parameter estimation: principles and problems* volume 9. M. Dekker.

31

Steyn-Ross, M. L., Steyn-Ross, D. A., & Sleigh, J. W. (2012). Gap junctions modulate seizures in a mean-field model of general anesthesia for the cortex. *Cognitive neurodynamics*, *6*, 215–225.

Struys, M., Jensen, E. W., Smith, W., Smith, N. T., Rampil, I., Dumortier, F., Mestach, C., & Mortier, E. P. (2002). Performance of the arx-derived auditory evoked potential index as an indicator of anesthetic depth: a comparison with bispectral index and hemodynamic measures during propofol administration. *Anesthesiology*, *96*, 803–816.

Tseng, S., Chen, R., Chong, F., & Kuo, T. (1995). Evaluation of parametric methods in eeg signal analysis. *Med Eng Phys*, *17*, 71–8.

Voss, L., & Sleigh, J. (2007). Monitoring consciousness: the current status of eeg-based depth of anaesthesia monitors. *Best practice & research Clinical anaesthesiology*, *21*, 313–325.

Wakita, M., Kotani, N., Nonaka, K., Shin, M.-C., & Akaike, N. (2013). Effects of propofol on gabaergic and glutamatergic transmission in isolated hippocampal single nerve-synapse preparations. *European journal of pharmacology*, *718*, 63–73.

Wang, X. (2009). Propofol and isoflurane enhancement of tonic gamma-aminobutyric acid type a current in cardiac vagal neurons in the nucleus ambiguus. *Anesthesia & Analgesia*, *108*, 142–148.

Wang, X., Huang, Z.-G., Gold, A., Bouairi, E., Evans, C., Andresen, M. C., & Mendelowitz, D. (2004). Propofol modulates gamma-aminobutyric acid-mediated inhibitory neurotransmission to cardiac vagal neurons in the nucleus ambiguus. *Anesthesiology*, *100*, 1198–1205.

Wendling, F., Bartolomei, F., Bellanger, J., & Chauvel, P. (2002). Epileptic fast activity can be explained by a model of impaired gabaergic dendritic inhibition. *European Journal of Neuroscience*, *15*, 1499–1508.

Whalen, A. J., Brennan, S. N., Sauer, T. D., & Schiff, S. J. (2015). Observability and controllability of nonlinear networks: The role of symmetry. *Physical Review X*, *5*, 011005.

Ying, S.-W., & Goldstein, P. A. (2005). Propofol suppresses synaptic responsiveness of somatosensory relay neurons to excitatory input by potentiating gaba. *Molecular pain*, *1*, 2.

Yue, L., Xie, A., Bruzik, K. S., Frølund, B., Qian, H., & Pepperberg, D. R. (2011). Potentiating action of propofol at gabaa receptors of retinal bipolar cells. *Investigative ophthalmology & visual science*, *52*, 2497.

Zikov, T., Bibian, S., Dumont, G., Huzmezan, M., & Ries, C. (2006). Quantifying cortical activity during general anesthesia using wavelet analysis. *IEEE Transactions on Biomedical Engineering*, *53*, 617–632.