Parameter Estimation Using Unscented Kalman Filter on the Gray-Box Model for Dynamic EEG System Modeling

Sun-Hee Kim, Hyung-Jeong Yang, Ngoc Anh Thi Nguyen, Raja Majid Mehmood, and Seong-Whan Lee, *Fellow, IEEE*

***Abstract*— Model parameters' estimation is one of the most important tasks in the analysis and design process of a nonlinear dynamic system in real time, especially in the presence of noise. This article presents a novel approach in estimating important parameters of gray-box model for such a system on real nonlinear EEG to simulate efficiently the dynamic characteristics of neurons. Specifically, the proposed methodology exploits unscented Kalman filter (UKF) that is combined with chaos neural population model to formulate the interaction between the cortical areas. The proposed methodology is compared with the state-of-the-art parameters' estimation techniques to verify the efficiency of the UKF on the gray-box model. Experimental results show that the proposed method demonstrates the lowest error value of root mean square error (RMSE) among existing parameter estimation methods. The robustness of the proposed approach is further validated in its convergence and automation, with minimum error relatively than others and without any user-specified input, respectively.**

***Index Terms*— Electroencephalogram, gray-box model, nonlinear dynamic system, parameter estimation, unscented Kalman filter (UKF).**

Manuscript received August 27, 2019; revised December 15, 2019; accepted December 29, 2019. Date of publication January 17, 2020; date of current version August 11, 2020. This work was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIT) under Grant NRF-2018R1A2B6006046, in part by a Korea University Grant, in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant NRF-2017R1A4A1015559, and in part by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea Government under Grant 2017-0-00451 (Development of BCI Based Brain and Cognitive Computing Technology for Recognizing User's Intentions Using Deep Learning). The Associate Editor coordinating the review process was Anirban Mukherjee. *(Corresponding author: Seong-Whan Lee.)*

Sun-Hee Kim and Seong-Whan Lee are with the Department of Brain and Cognitive Engineering, Korea University, Seoul 02841, South Korea (e-mail: sunheekim@korea.ac.kr; sw.lee@korea.ac.kr).

Hyung-Jeong Yang is with the Department of Computer Science, Chonnam National University, Gwangju 61186, South Korea (e-mail: hjyang@jnu.ac.kr).

Ngoc Anh Thi Nguyen is with the Faculty of Information Technology, The University of Da Nang—University of Education, Da Nang 550000, Vietnam (e-mail: ngocanhnt@ued.udn.vn).

Raja Majid Mehmood is with the Department of Information and Communication, School of Electrical and Computer Engineering, Xiamen University Malaysia, Sepang 43900, Malaysia (e-mail: rmeex07@gmail.com).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIM.2020.2967138

I. Introduction

In a dynamic system, system states and parameters are important for modeling, performance analysis, and prediction due to their decisive effect on the precision, stability, and control of the system. Most dynamic systems involve nonlinearity that does not satisfy the properties of superposition and homogeneity [1]. With those challenges, the state estimation of such system that exhibits nonlinearity is more complicated and difficult compared to a linear system [2]. Moreover, parameter values cannot be obtained, while the system is deteriorating or changing over time, or cannot be measured directly. In this case, parameters affect the control and monitoring of a system. Therefore, parameter estimation is an ongoing important research area on such a nonlinear dynamic system recently. Many various methods have been investigated to estimate parameters [3] [4]. The well-known parameter estimation methods are used, including the maximum-likelihood estimators (MLEs), the Bayes estimators, method of moments' estimators (MMEs) [5], least squares estimator (LSE) [6], nonlinear least squares (NLS) [7], and minimum variance unbiased estimator (MVUE) [8].

In more detail, MLE is the estimation method that maximizes the likelihood function for parameters [9]. When applied appropriately, it demonstrates excellent qualities for characteristics of consistency, normality, efficiency, sufficiency, invariance, and so on [10]. However, the MLE method might fail if the likelihood function is not well-defined or in the case of data loss [11]. MLE is often nontrivial mathematically, particularly if confidence intervals are desired for the parameters. In addition, the numerical estimation is usually nontrivial [12]. In contrast, NLS minimizes the sum of squares of differences between the empirical cumulative distribution function (CDF) and the theoretical CDF [7]. The advantage of NLS is that it can be applied to any distribution. However, its major drawbacks lie in optimization and high dependence on the initial values [13].

Another alternative approach for parameter values optimization in nonlinear dynamic systems, which is an extension of the Kalman filter, also known as the extended Kalman filter (EKF), is employed for parameter estimation [14]. Sun *et al.* [15],

Bavdekara *et al.* [16], and Ebrahimian *et al.* [17] estimated a state value and parameter using the EKF on their nonlinear system for a biochemical networks model, a structural finite element model, and a benchmark continuous fermenter model, respectively. Moradkhani and Hsu [18] estimated the parameter and state value of a drainage facility model with the particle filter (PF). Wang *et al.* [19] applied the Monte Carlo method in order to use PF in a dynamic system to estimate states. Kantas *et al.* [20] employed the PF to perform static parameter estimation in state-space models. Armando *et al.* [21] proposed the cubature Kalman filter (CKF) to determine a change in the state and parameters of the model during status epilepticus. Liu and Gao [22] suggested using UKF in order to observe states and estimate parameters in a neural mass model (NMM). In summary, among nonlinear methods, UFK is widely used as the best method for successfully estimating the state and a specific parameter of the model [23].

The human brain, composed of numerous neurons, is one of the most complex physiological systems. EEG measures the electrical activity occurring due to biochemical interactions between brain cells. Hence, the EEG has been used as the most significant means for capturing changes in brain activity and forecasting seizures in epilepsy studies [24]. However, the EEG is easily affected by noise and artifacts, such as heartbeat, eye blinking, eye movements, and muscular movement during measurements. Moreover, it has highly complex nonlinear characteristics that can cause statistical characteristics to change over time, exhibiting temporal dependence [25]. Therefore, to analyze the EEG, nonlinear dynamic systems have been employed [26]. In this article, we propose the adoption of a UKF as the parameter estimation method of a gray-box model, with the aim of modeling nonlinear EEG dynamic systems. The proposed method employs a chaos neural population model to structure the model of the nonlinear EEG system. Consequently, the UKF is exploited to estimate parameters of the model based on observed data. UKF does not require linearization as nonlinear transformation using the unscented transformation (UT). It estimates the optimal value by repeatedly performing the updation process of an error covariance, which removes the measurement noise of the system [27] [28]. Our contributions are as follows.

1) *Good Model Design:* We introduce a gray-box model, a dynamic system modeling method comprises a mean-field model modeling the dynamics of the brain activity and an unscented Kalman filter estimating parameters.
2) *Parameter Free:* Our gray-box model can be executed automatically without user intervention. It provides a model's parameter values minimizing the modeling error rate.
3) *Accuracy:* Our model offers good accuracy with respect to the root mean square error (RMSE) between the observed and predicted data.
4) *Scalability:* The run time of the proposed method grows linearly with the total input data. Our model can generate a new signal that can describe the trends of the input signal.

The rest of this article is organized as follows. Section II presents the proposed parameter estimation method and the chaos neural population model for the gray-box model. Section III provides the experimental results obtained with the proposed method. Section IV discusses existing methods and compares them with the proposed method. Finally, the conclusions are drawn in Section V.

II. Materials and Methods

Parameter estimation algorithms are used to determine a state or parameter value in a dynamic system. In this section, a chaos neural population that describes a nonlinear EEG dynamic system in a gray-box model is introduced, followed by the UKF, which is applied for the parameter estimation of the model.

A. Gray-Box Model

Dynamical modeling can be separated into three broad categories: white-box modeling, black-box modeling, and gray-box modeling. White-box modeling is conducted by 1-D principles of physical or chemical laws. Black-box modeling determines the structure of a model based on the observed data. Gray-box modeling is a mixture of white-box modeling and black-box modeling. It works by determining the structure of a system and estimates the parameters of the system using observed data. The gray-box model has the advantage of providing both physical knowledge to model the system structure and data knowledge to identify the parameters in the model. Thus, the gray box itself preserves the benefits of both the white-box and black-box models [29].

1) Chaos Neural Population Model: Neurons make up the structural and functional units of the nervous system. They transfer and process information by exchanging neurotransmitters with other neurons. A neuron is composed of the neuron body and the synapse. The neuron body is composed of a pair comprising an excitatory cortical cell and inhibitory cortical cell. In the case of excitatory neurons, neurotransmitters are gathered and sent to connected neurons. However, inhibitory neurons do not send the gathered currents to connected neurons. A neural network is a mathematical model that simplifies biological neural information transmission processes [24], [30]. This is applied to solve problems, such as image processing, voice recognition, and optimization. However, neural networks meet the limitation of modeling by using only particular neurons (i.e., only the excitatory neuron) with regard to biological functions. To overcome this problem, the chaos neural population model, which can describe complicated and diverse dynamic characteristics of neurons, is proposed [31].

A brain is regarded as a chaotic dynamical system. The generated EEG signals from this brain are also chaotic due to the amplitude changes randomly with time. We adopt the chaos neural population model based on the mean-field model (MEM) for dynamic EEG system modeling. It is a well-known nonlinear chaotic system that can exhibit the chaotic behavior of EEG [32]. The chaos neural population model comprises a group of excitatory and inhibitory neurons that are connected through the synapse. Each group is composed of two chaotic neurons and four synapses [31], [33]. The states of two chaotic neurons in the chaos neural population model are presented as $e_k(t), k = 1, \ldots, N_e$, for excitatory neurons

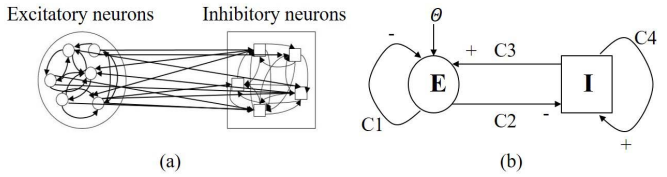


Fig. 1. Chaos neural population model. (a) Two populations of neurons of two types. (b) Simplified MEM.

[see the small circles in Fig. 1(a)] and $i_l(t), l = 1, \ldots, N_i$ for inhibitory neurons [see the small squares in Fig. 1(a)]. The activities of $e_k(t)$ and $i_l(t)$ in the neuronal population can described, as shown in the following equations:

$$\frac{de_k}{dt} = -e_k + S\left(\frac{1}{N_e}\sum_{k=1}^{N_e} u_{kl}e_l - \frac{1}{N_i}\sum_{l=1}^{N_i} v_{kl}i_l - \theta_k^e + p_k\right) \tag{1}$$

$$\frac{di_k}{dt} = -i_k + S\left(\frac{1}{N_e}\sum_{k=1}^{N_e} w_{kl}e_l - \frac{1}{N_i}\sum_{l=1}^{N_i} q_{kl}i_l - \theta_k^i\right) \tag{2}$$

where parameters u, v, w, and q are the strengths of the connections between the populations. The neuronal population models exhibit several types of interactions that involve self- and cross-interactions expressed by u_{kl}, v_{kl}, w_{kl}, q_{kl}. θ_k^e, and θ_k^i are the firing thresholds and should be sufficiently large. $S(x) = (1 + \exp(-x))^{-1}$ represents the sigmoid function of nonlinear data [31].

In this article, we exploited the chaos neural population model to simplify dynamic neuron populations as Fig. 1(b). This model consists of a network of coupled neuronal populations that use one or two steady-state variables to indicate the mean activity of the population. This so-called mean-field mode to indicate the mean activity of each group is given by the following equations:

$$E(t) = \left(\frac{1}{N_e}\right)\sum_{k=1}^{N_e} e_k(t) \tag{3}$$

$$I(t) = \left(\frac{1}{N_i}\right)\sum_{l=1}^{N_i} i_l(t). \tag{4}$$

Given a mean activities $E_r(t)$ and $I_r(t)$ of each group at location r and time t, this can be represented with follows:

$$\hat{E}_r(t) = -E_r(t) + S(c_1E_r(t) - c_2I_r(t) - \varepsilon^e + \Theta) \tag{5}$$

$$\hat{I}_r(t) = -I_r(t) + S(c_3E_r(t) - c_4I_r(t) - \varepsilon^i) \tag{6}$$

where $\varepsilon^e = \Sigma\theta_k^e/N_e$, $\varepsilon^i = \Sigma\theta_k^i/N_i$, and $\Theta = p_k/N_e$. The chaos neural population model is connected by four interactions c_1, c_2, c_3, and c_4. These parameters express the strength of connection with the interactions (all pairs do not have a value of 0). The chaos neural population model includes an external input value Θ of an excitatory neuron, and thresholds ε^e and ε^i have large values. $\hat{E}_r(t)$ and $\hat{I}_r(t)$ obtained by (5) and (6) are considered as the new activities for the excitatory and inhibitory neurons at time t. The final output of the model is given by the following equation [34]:

$$\hat{z}_r(t) = \hat{E}_r(t) - \hat{I}_r(t). \tag{7}$$

In this article, we employ the chaos neural population model, which most similarly reproduces the electric potential activity of a neuron on the gray-box model for nonlinear EEG system modeling. To estimate the parameters of an EEG model, such as the strength of self- and cross-interactions, c_1, c_2, c_3, and c_4, threshold, ε^e and ε^i, and external input value, Θ, UKF is employed. In the gray-box model, the initial parameter values are used in (5) and (6), $\hat{z}_r(t)$ is calculated by (7), and $\hat{z}_r(t)$ is again used in (8) to compute the error value between the observed data $z_r(t)$ and the predicted data $\hat{z}_r(t)$ by the model

$$e_r(t) = z_r(t) - \hat{z}_r(t). \tag{8}$$

Our proposed method automatically updates the parameter set to using UKF to minimize the values of (8), $e_r(t)$. Therefore, the gray-box model does not require any user-intervention for the initial parameter setting.

2) Unscented Kalman Filter: The UKF using a UT is suggested to solve the linearization problem of the EKF, which applies the Jacobian transformation [15], [27]. UKF is a technique that recursively estimates the states of a nonlinear dynamic system. It uses UT that calculates the state variable and covariance using generated sigma points and weights [35]. If the random variable x is n-dimensional, the sigma point $\chi_i (i = 1, \ldots, 2n + 1)$ with the number of $2n + 1$ on time t is calculated as shown in (9). The calculated sigma point χ_t is applied on nonlinear model function $f(\chi_t)$, and a random variable x propagated through nonlinear transformation

$$\begin{aligned}
\chi_t^0 &= \widehat{x}_t \\
\chi_t^i &= \widehat{x}_t + (\sqrt{(n+\lambda)P_t})_i, i = 1, \ldots, n \\
\chi_t^{i+n} &= \widehat{x}_t - (\sqrt{(n+\lambda)P_t})_i, i = 1, \ldots, n \\
\widehat{x}_t &= \boldsymbol{E}[x] \\
P_t &= \boldsymbol{E}[(x - \widehat{x})(x - \widehat{x})^T]
\end{aligned} \tag{9}$$

where n presents the number of dimensions of a state variable x, $(\sqrt{(n+\lambda)P_t})_i$ is the ith column of square root matrix (lower triangular Cholesky factorization), $\boldsymbol{E}[\cdot]$ is the expectation operator, and it generally requires taking expectations of a nonlinear function of the prior state variable. In case of the UKF, it assumes that x has initial mean and covariance. λ is given by the following equation:

$$\lambda = \alpha^2(n + \kappa) - n \tag{10}$$

where α is a scalar parameter that calculates spreading of the sigma point and always has a small value ($10^{-3} \leq \alpha \leq 1$). κ is usually set to 0 as it is the secondary scale parameter. Sigma point $2n + 1$ calculated in (9) is applied in nonlinear function f as (11), which results in prediction of the mean and covariance of x by (12) and (13)

$$Z_i = f(\chi_i), \quad i = 1, \ldots, 2n \tag{11}$$

$$\widehat{z} = \sum_{i=0}^{2n} W_i^{(m)} Z_i \tag{12}$$

$$P_i = \sum_{i=0}^{2n} W_i^{(c)} (Z_i - \widehat{z})(Z_i - \widehat{z})^T \tag{13}$$

where weights $W_i^{(m)}$ and $W_i^{(c)}$ are constants that determine the specific gravity of each sigma point when calculating the mean and covariance, given in the following equation:

$$\begin{aligned} W_0^{(m)} &= \lambda/(n+\lambda) \\ W_0^{(c)} &= \lambda/(n+\lambda) + (1-\alpha^2+\beta) \\ W_i^{(m)} &= W_i^{(c)} = 1/2(n+\lambda) \end{aligned} \tag{14}$$

where superscripts of m and c represent the mean and covariance, and β is a preliminary knowledge of the random variable distribution. If a random variable has a Gaussian distribution, the optimal value would be $\beta = 2$.

B. Parameter Estimation in a Gray-Box Model

Generally, dynamic systems are nonlinear, and in the case of experimental data, measuring errors can arise. Therefore, it is very important to know whether predicted data from the actual system can be employed. In this article, we employ actual EEG data in the gray-box model to implement parameter estimation using UKF. The UKF is based on UT to recursively estimate the system state. The following is the UKF procedure for parameter estimation in the EEG dynamic system.

First, the state variable $\widehat{x}_0$ and the error covariance matrix P_0 are initialized as (15) and (16) for EEG dynamic system

$$\widehat{x}_0 = \boldsymbol{E}[x_0] \tag{15}$$

$$P_0 = \boldsymbol{E}[(x_0 - \widehat{x}_0)(x_0 - \widehat{x}_0)]. \tag{16}$$

Second, the weight is calculated as (14), and the sigma point is calculated using state variable and covariance

$$\chi_{t-1}^i = [\widehat{x}_{t-1}, \widehat{x}_{t-1} \pm \sqrt{(n+\lambda)P_{t-1}}]. \tag{17}$$

Third, sigma points of $t-1$ calculated by (17) are propagated through the function F of (18). F is the nonlinear system function that is already given by (5) and (6). The transformed points are used to compute the mean and covariance of the prediction value of x_t as (19) and (20)

$$\chi_{t|t-1}^i = F(\chi_{t-1}^i) \tag{18}$$

$$\hat{x}_{t|t-1} = \sum_{i=0}^{2n} W_i^{(m)} \chi_{t|t-1}^i \tag{19}$$

$$P_{t|t-1} = \sum_{i=0}^{2n} W_i^{(c)} (\chi_{t|t-1}^i - \hat{x}_{t|t-1})(\chi_{t|t-1}^i - \hat{x}_{t|t-1})^T. \tag{20}$$

Fourth, we propagate then the sigma points through the measurement model function H. H is given by (7). With the transformed observations by (21), their mean and covariance (innovation covariance) are computed with (22) and (23)

$$Z_{t|t-1}^i = H(\chi_{t|t-1}^i) \tag{21}$$

$$\hat{z}_{t|t-1} = \sum_{i=0}^{2n} W_i^{(m)} Z_{t|t-1}^i \tag{22}$$

$$P_t^{zz} = \sum_{i=0}^{2n} W_i^{(c)} (Z_{t|t-1}^i - \hat{z}_{t|t-1})(Z_{t|t-1}^i - \hat{z}_{t|t-1})^T. \tag{23}$$

Fifth, the cross covariance between $\hat{x}_{t|t-1}$ and $\hat{z}_{t|t-1}$ can be obtained by (24), and it is used to calculate the Kalman Gain from (25). Subsequently, the Kalman Gain K, actual observed value z_t, and predicted measuring variable $\hat{z}_t$ are used to update the state estimation variable and covariance matrix using (26) and (27)

$$P_t^{xz} = \sum_{i=0}^{2n} W_i^{(c)} (\chi_{t|t-1}^i - \hat{x}_{t|t-1})(Z_{t|t-1}^i - \hat{z}_{t|t-1})^T \tag{24}$$

$$K = P_t^{xz} (P_t^{zz})^{-1} \tag{25}$$

$$\hat{x}_t = \hat{x}_{t|t-1} + K_t(z_t - \hat{z}_{t|t-1}) \tag{26}$$

$$P_t = P_{t|t-1} - K_t (P_t^{zz}) K_t^T \tag{27}$$

where P_t is the state error covariance. $W_i^{(m)}$ and $W_i^{(c)}$ are weights calculated in (14). In the gray-box model for EEG dynamic modeling, the optimum state is estimated by repeatedly updating the mean and error covariance. In this article, we use the least squares method (LSM) to generate the data that are most similar to observed data and estimate unknown parameters of nonlinear model. LSM obtains minimized errors between the observed and predicted data from a model [36]. For estimating the optimal parameter set, we measure the error value between the predicted data by the model $\hat{z}_R$ and observed data z_R by (28). R means the iteration number for the parameter estimation

$$e_R = z_R - \hat{z}_R = \triangle z. \tag{28}$$

Here, we used the LSM to find the optimal parameter set to minimize e_R. Our method automatically finds the seven parameters: $c_1, c_2, c_3, c_4, \varepsilon^e, \varepsilon^i$, and Θ, which needs the EEG nonlinear dynamic model by performing the LSM. In this article, we set lower and upper bounds of model as $-10 \leq Q \geq 10$. The gray-box model using UKF in order to the parameter estimation that is proposed in this article is expressed in Table I.

TABLE I
GRAY-BOX MODEL USING UKF

Input: EEG signal, $z_i, i = 1, ..., t$, and an initial parameter set, Q
Output: a complete set of parameters:
$Q = (C_1, C_2, C_3, C_4, \varepsilon^e, \varepsilon^i, \Theta)$

for $q = 1 : t$ **do** // t is length of a signal
  while minimizing the error cost **do**
    // maximum repetition number: r
    /* Update parameter set */
    for $j = 1 : n$ **do** // n is dimension of state variables
      1. Initialization state with noise and covariance by Eqs. (15)-(16)
      2. Calculate the sigma point by Eq. (17)
      3. Perform prediction transformation by Eqs. (18)-(20)
      4. Perform measurement transformation by Eqs. (21)-(23)
      5. Update the estimated state variable and covariance by Eqs. (24)-(27)
    end for
    Calculate the error value e^r by Eq. (28)
    $e_r = argmin(\hat{z}_i | Q = (C_1, C_2, C_3, C_4, \varepsilon^e, \varepsilon^i, \Theta))$
    Update parameter set $Q \leftarrow Q_{opt}$
  end while
end for
return $Q = (C_1, C_2, C_3, C_4, \varepsilon^e, \varepsilon^i, \Theta)$

III. EXPERIMENTAL RESULTS

A. Epilepsy Real Data Set

We used a data set of intracranial EEG (iEEG) taken from sources of "https://www.ieeg.org/" and Kaggle platform [37].

TABLE II
DATA DESCRIPTION

State	Dog1	Dog2	Dog3
interictal	4 hours	7 hours	12 hours
preictal	4 hours	7 hours	12 hours
60 minute * 60 s * 400Hz=1,440,000 time points			

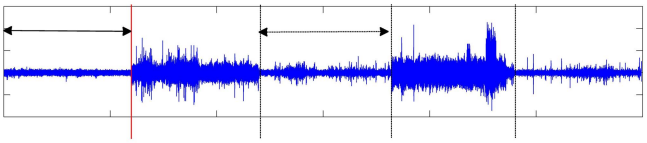


Fig. 2. Description of preictal and interictal.

This data set was collected from dogs with naturally occurring epilepsy, recorded over a long duration, ranging for multiple months to a year, with 16 electrodes employing a sampling rate of 400 Hz. The electrode array consisting of 16 electrodes included the top $1\sim8$ electrodes located on the left hemisphere, whereas the remaining $9\sim16$ electrodes were placed on the right hemisphere. Each data set contains sequential interictal and preictal data segments. The interictal data segments depict nonseizure activity recorded before or after seizures, while preictal data were the data before the seizure onset (see Fig. 2). In this article, preictal data segments were organized into 10-min EEG clips, recorded from 1 h prior to seizure to 5 min before seizure onset (i.e., from 1:05 to 0:05 before seizure onset). Similarly, interictal data segments were chosen randomly as 10-min EEG clips from the full data, with the restriction that they occurred at least one week before or after any seizure. This length could be predicted with sufficient warning to allow the administration of fast-acting medications before the seizure onset [38]. This iEEG data set was annotated and verified by experts in all subclinical and clinical seizures, and it has been validated in several studies [39], [40]. We used three canine iEEG data sets (see Table II), which contained a sufficient number of seizures and long-lasting interictal signals suitable for analysis [39].

B. Parameter and State Estimation

UKF adopted UT to improve the errors due to linearization on the EKF. The UKF can estimate states without the linearization process using the UT. We proposed the adoption of the UKF for parameter estimation on a gray-box model. Finding optimal parameters is advantageous, as it creates data closest to the observations. Fig. 3 shows three types of synthetic and predicted data through the proposed method. Solid lines in Fig. 3(a), (c), and (e) represent actual observed data, while dotted lines indicate the data predicted by the estimated parameters using the UKF. Consequently, the predicted data from the proposed method is highly similar to the synthetic data (see Fig. 3). Fig. 3(b), (d), and (f) shows state values of each synthetic data, which are estimated in the process performing UKF over time points. State values 1 and 2 reflect as values of the excitatory neuron and the inhibitory neuron in our gray-box model, respectively, and the initial state value is set with noise.

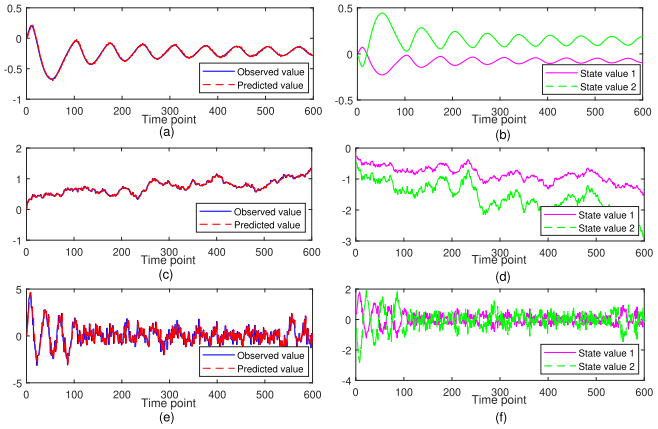


Fig. 3. Comparison of synthetic and predicted data from a gray-box model based on (a), (c), and (e) UKF and (b), (d), and (f) state values.

State variables contain information of the past or present, which describes the characteristics and dynamic behavior of a system. A dynamic system can express a model that includes state variables. However, in the case of nonlinear systems, state variables cannot be expressed using linear models because of the complicated dynamic behavior of the system. State variables and parameters that comprise a model should be estimated simultaneously. Most nonlinear systems are not directly measurable; thus, the probability distribution of initial state variable x_0 assumes a normal distribution. However, the probability distribution of a propagated state variable by nonlinear models does not follow the normal distribution [36]. Therefore, the probability distribution of state variables cannot be briefly described by its mean and covariance. For this reason, the nonlinear state estimation technique uses an approximation method of the probability density function for state variables. For example, a probability distribution is not a normal distribution; however, it is converted into a normal distribution using the approximation method, where the mean and the covariance can be expressed.

In the case of the UKF, if the probability distribution follows a normal distribution, the UKF obtains optimal results when $\lambda = 3 - L$ (L is state dimension). If λ is a negative number, one or more elements can be negative in the covariance calculation, and the probability distribution loses its property of a normal distribution. Therefore, the UKF uses a square root filter to solve the negative number problem of covariance calculation [41]. Here, unknown state variables are estimated during the process of parameter estimation by the UKF. Fig. 4 shows the probability distribution of estimated state variables by the nonlinear model. The blue line indicates the estimated state variable $\widehat{x}_1$, and the red line indicates the estimated state variable $\widehat{x}_2$. Fig. 4(a) shows the probability distribution of the estimated state variables, and $\widehat{x}_1$ and $\widehat{x}_2$ in Fig. 3(a) are synthetic data. $\widehat{x}_1$ has a mean value of -0.0811 and a covariance value of 0.0023. Moreover, $\widehat{x}_2$ has a mean value of 0.1581 and a covariance value of 0.0087. Fig. 4(b) and (c) shows a probability distribution of Fig. 3(c) and (e). Resultantly, the probability distribution of the estimated state variable in each data set indicates a normal distribution.

TABLE III
ESTIMATED PARAMETERS AND RMSE

Parameters	c_1	c_2	c_3	c_4	ε^e	ε^i	Θ	RMSE
Data1(Fig. 3 (a))	-0.0426	0.0051	-0.0536	0.0143	0.0364	-0.0147	-0.1445	0.0096
Data2(Fig. 3 (c))	0.0070	-0.0075	0.0042	0.0066	0.0062	0.0017	0.0018	0.0101
Data3(Fig. 3 (e))	0.0089	-0.0127	0.1101	0.0804	0.0533	0.0963	0.0352	0.0104

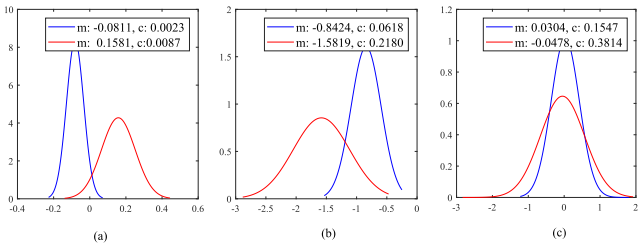


Fig. 4. Probability (m: mean and c: covariance). Synthetic (a) data1, (b) data2, and (c) data3.

In particular, we confirmed that the UKF expresses the estimated state variable with a mean and a covariance through approximation of the conditional probability density function using the square root filter.

Table III lists optimal parameters acquired from synthetic data [see Fig. 3(a), (c), and (e)]. These parameters are used to minimize the error between observed and predicted data using UKF in the gray-box model. Since UKF is very strong in a nonlinear system, linearization is not needed and errors do not occur. The UKF repeats updating the error covariance to estimate optimum parameter values [27]. In Table III, Data 1 [see Fig. 3(a)] shows RMSE of 0.0096, while Data 2 [see Fig. 3(c)] and Data 3 [see Fig. 3(e)] show the RMSE of 0.0101 and 0.0104, respectively. Therefore, our method employing the UKF to estimate the optimal parameter produces data with the least error compared with the actual observed data.

C. Robust Parameter Estimation Method on Gray-Box Model

In this article, to prove the efficiency of the gray-box model, which is a combination of the chaos neural population model and the UKF, we compared the parameter estimation methods generally used in a dynamic system. For the performance evaluation, we used the EKF and PF for parameter estimation in the gray-box model. The EKF is a nonlinear extension of the Kalman filter, and PF is a simulation prediction technique, which is a continuing Monte Carlo method. The latter method can be used on the nonlinear system, similar to the UKF [42], [43]. The particle number in PF for this experiment was set equal to 500. In the filter design, computational efficiency and estimation accuracy need to be considered. Thus, a large number of particles are computationally ineffective and request high memory [44]. Therefore, we selected the number of particles through the experimental results that provided the best performance among particle numbers, including 100, 300, 500, 700, 900, and 1000.

For the comparison of the similarity between the observed and predicted data, we measured the RMSE. Fig. 5 shows the randomly selected data from the actual EEG data obtained from experiments conducted on dogs, predicted data obtained by the popular parameter estimation methods, and the error values between the real EEG and predicted data. In Fig. 5, the blue line depicts the real EEG data, the red line denotes the UKF, the green line is the EKF, and finally, the yellow line represents the PF result. We confirm that the predicted data using UKF is similar to the observed data to the naked eye through observing Fig. 5(a) and (c). Moreover, they obtained the lowest RMSE in comparison to the other parameter estimation methods. More detailed results are shown in Fig. 5(b) and (d), which shows the error values between the observed and predicted data obtained by each method. The UKF in Fig. 5(a) and (d) has an error value close to 0. In the case of the EKF, Fig. 5(a) and (c) seems very similar to the actual observed data; however, through Fig. 5(b) and (d), it can be observed that they have a larger error compared with the UKF. In addition, PF has a high error rate in several areas. Thus, when the UKF is used on the gray-box model to estimate parameters, it predicts data that are most similar to the observations.

We observed that PF could not produce accurate data in comparison to the EKF and UKF, as shown in Fig. 5. PF can be used on a system with nonlinearity, and it is resistant to noise. Nevertheless, it is highly dependent on the number of particles and requires a long computation time [45]. PF has emerged as the most successful algorithm for nonlinear state-space models [20], but the performance in the gray-box model using the chaos neural population model to structure a nonlinear EEG system dramatically decreased. Therefore, we used the EKF, which exhibits better efficiency than PF in our experiments. Namely, we compared the proposed method and EKF with significantly more data to verify that our method is more adequate for optimal parameter estimation in the gray-box model. Fig. 6 shows the predicted data obtained using the proposed method and the EKF from EEG data using dogs. Fig. 6(a), (c), and (e) is a part of Dog 1, Dog 2, and Dog 3 signals measured in preictal areas, while Fig. 6(b), (d), and (f) is the dog signals measured from interictal areas. We used the EKF and UKF on these data to estimate the parameters. In Fig. 6, the actual observed data are the blue line, the UKF is the red line, and the EKF is the green line.

In the case of data generated using the UKF to estimate parameters, results were very similar to the signal in reality. In contrast, in the case of the EKF, there was also similarity to the observed signals. However, the range of error was wider than that of the UKF. Fig. 6(e) shows that both the EKF and UKF methods are very similar to real observed data. The EKF estimates states and parameters by employing primary linearization on a nonlinear system. This is the reason

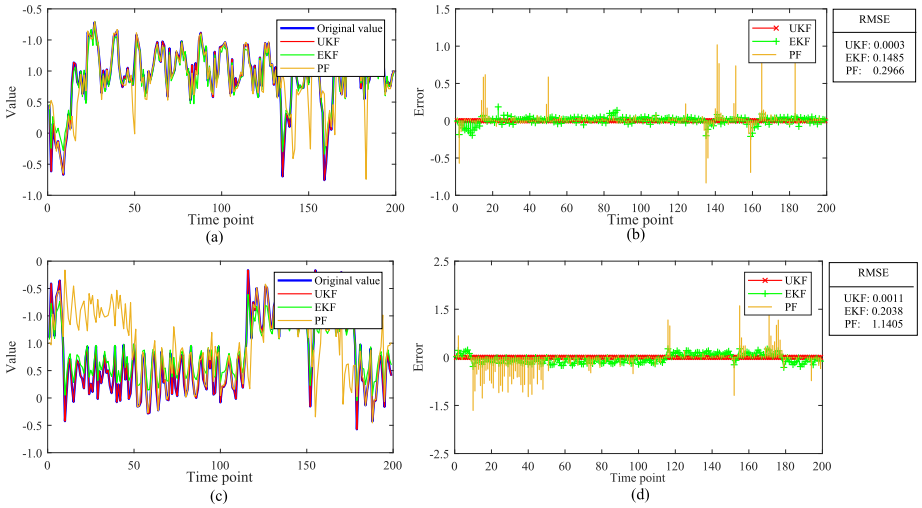


Fig. 5. Comparison of proposed method and popular parameter estimation method. (a) and (c) Results of predicted data by proposed method, EKF, and PK parameter estimation method. (b) and (d) Error values (y-axis) at each time point (x-axis).

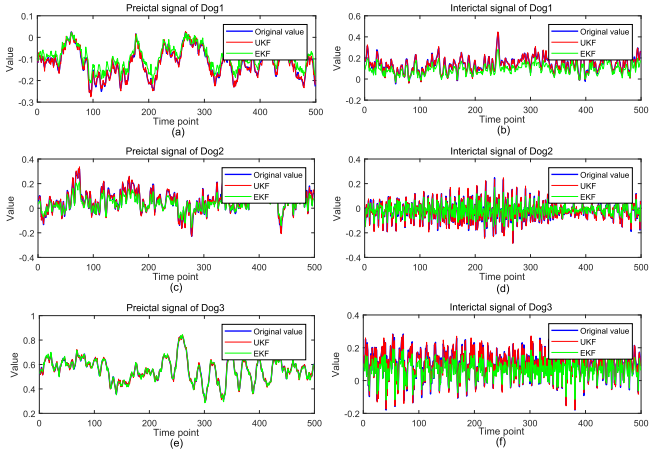


Fig. 6. Comparison UKF. Result of (a) and (b) dog 1, (c) and (d) dog 2, and (e) and (f) dog 3.

the EKF is efficiently applied in several fields. However, Fig. 6(a)–(d) and (f) shows, if the data are nonlinear and contain significant noise, the EKF has a high tendency to generate greater error percentages between the observed data and the data from the model obtained by linearization, compared with the UKF.

To verify the rate of convergence, we measured and tabulated the convergence time and iteration number of convergence of EKF and UKF, as shown in Table IV. As a result, our method using UKF is converging at faster speed than EKF in all data sets. Moreover, the iteration number for convergence appears lower in our method than in EKF. The EKF uses the Jacobian to linearize the nonlinear dynamic system and the nonlinear measurement model. Therefore, EKF has some problems that damage the stability of the filter because the incurred by linearization is large and provides a highly inexact prediction about the state of the system. Furthermore, the linearization using a Jacobian requires considerable time and computational ability [15]. In contrast, the UKF is a nonlinear filter using UT. Therefore, it reduces the error incurred by linearization, as the Jacobian calculation is unnecessary, and it can converge quickly at the optimized parameters [27], [28]. Consequentially, UKF shows the ability to rapidly estimate optimal parameters in the nonlinear EEG data set, as shown the Table IV, when it is applied on the nonlinear dynamic model.

D. Scalability

To verify scalability of the proposed method, we measured the CPU run-time required to estimate optimum parameters that minimize the error rate between the observed and predicted data. We measured the CPU run time according to data length by increasing the length of input data by 10 000 (25 s) and set the maximum repeating number for optimizing parameter to 500. To measure the CPU run time according to the length of data, we used MATLAB (version R2014a) in a Windows Server 2008-64-bit environment. Moreover, the detailed experiment environment is as follows: CPU is Intel Xeon E7450 @ 2.40 GHz, and RAM is 128 GB. Table V shows the CPU run-time required to estimate optimum parameters using the EKF and UKF in the data each for interictal and preictal data from three dogs. Consequently, depending on the increasing size of the observed data, the run time for estimating optimal parameters gradually increased for both the EKF and UKF. These results equally appeared in all dogs data sets. In a more detailed comparison of the UKF and EKF with both interictal and preictal data, the UKF yielded the lowest run time despite its similarities with EKF complexity $O(n^3)$ [46]. That is, our method using UKF quickly converges on the parameter estimation than EKF.

The method employed to estimate optimal parameters using the UKF on a gray-box model is faster than the EKF. In addition, in interictal data, the execution time for optimizing parameters was on average slower than the time required for preictal data. Hence, the execution time of preictal data for optimizing parameters was faster than time required for interictal data ranging from 0.3 to 2 times faster. This is

TABLE IV

ITERATION NUMBER OF CONVERGENCE AND CONVERGENCE TIME OBTAINED BY PARAMETER ESTIMATION IN TEN RUNS

Data	States	EKF		UKF	
		Iteration Number of Convergence	Convergence Time(s)	Iteration Number of Convergence	Convergence Time(s)
Dog1	Interictal	103.0	4.22	83.2	3.76
	Preictal	121.2	5.37	91.0	4.00
Dog2	Interictal	87.4	3.89	87.0	3.66
	Preictal	114.6	4.62	85.8	3.77
Dog3	Interictal	87.8	3.75	84.4	3.73
	Preictal	127.8	8.36	99.4	4.31

TABLE V

COMPARISON OF EXECUTION TIME FOR OPTIMIZING PARAMETERS ON REAL EEG DATA OF DOGS (UNIT: S)

Length of Data	Dog1				Dog2				Dog3			
	Interictal		Preictal		Interictal		Preictal		Interictal		Preictal	
	EKF	UKF	EKF	UKF	EKF	UKF	EKF	UKF	EKF	UKF	EKF	UKF
20,000	40.1	24.1	17.3	29.4	28.2	11.6	78.0	36.3	42.5	44.5	107.7	70.9
30,000	71.1	61.0	54.3	61.6	47.1	28.6	67.1	42.6	38.6	55.4	109.4	113.1
40,000	91.2	79.3	60.2	90.3	59.1	101.0	336.3	153.9	80.1	42.7	80.9	108.6
50,000	112.3	63.5	106.3	105.7	105.2	30.7	222.5	205.0	48.6	91.6	138.1	146.4
60,000	135.7	114.5	235.2	135.4	134.6	115.0	258.3	198.8	78.7	207.7	2710.4	183.6
70,000	146.8	73.0	126.2	139.6	95.9	189.4	161.6	238.7	115.6	152.7	406.4	306.1
80,000	265.6	213.9	85.5	186.6	422.8	57.9	502.4	357.1	314.1	250.3	936.0	313.8
90,000	128.0	63.2	541.6	111.8	149.4	123.3	316.3	312.9	289.3	148.7	1586.4	324.1
100,000	175.1	201.2	304.5	148.1	144.6	191.2	498.3	429.7	278.3	284.0	420.4	308.6
Average	**129.5**	**99.3**	**170.1**	**112.1**	**131.9**	**94.3**	**271.2**	**219.5**	**142.9**	**142.0**	**721.7**	**208.4**

because using the UKF to estimate parameters is suitable for complicated properties or nonlinear data with significant noise [28], [35]. Based on these results, we can identify that our proposed method, employing the UKF, is scalable in terms of execution time. Accordingly, our method can be applied in various domains, such as modeling, data analysis, and prediction of the real nonlinear dynamic system.

Fig. 7 shows the RMSE between the observed and predicted data by the proposed model on the data from three dogs, based on data size (as Table V). Fig. 7(a) illustrates a measuring result of data set Dog 1, and when the EKF was applied, it showed a higher error rate compared with the UKF. Hence, when the UKF was applied, it showed an average error rate of 0.01, while the average EKF error rate is 0.03. The Dog 2 data set is shown in Fig. 7(b). When the UKF was applied, a result similar to that of Dog 1 was obtained. With the EKF, interictal data produced less error the than preictal data. Fig. 7(c) depicts data set Dog 3, and when UKF was applied, an average error rate of 0.01 was achieved in both the interictal and preictal data. However, when the EKF was applied on interictal and preictal data, average error rates of 0.02 and 0.09 are obtained, respectively. Thus, instead of employing the EKF method to determine the optimized parameters, employing the UKF, which optimizes parameters by UT, is more suitable.

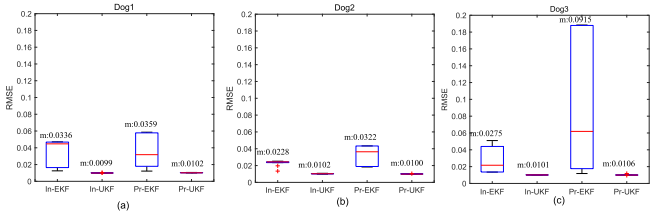


Fig. 7. RMSE size. (a) Data set Dog 1. (b) Data set Dog 2. (c) Data set Dog 3.

IV. DISCUSSION

In general, the models of neural activity can be divided into NMMs and neural-field models (NFMs) [47]. These models have been used as nonlinear models that express synaptic and neural populations that constitute brain active networks [48]. NMMs are a simple approach for modeling the activity of large populations of neurons and synapses since the 1970s. It formulates a set of ordinary differential equations describing the states of neural populations and simulates the state dynamics of the system describing the overall connectivity between populations of a given cortical area. NMMs are especially useful in understanding the brain signal produced by electrical activity within the brain, such as the EEG and magnetoencephalogram (MEG). However, they can be only applicable for data that reflect the average behavior of neuronal populations. In addition, the dynamics of NMMs in response to a rapid change in the input does not correctly reproduce a microscopic simulation of the population of neurons [49].

NFMs can describe the spatiotemporal evolution and represent a MEMs [50]. In this article, we used the MEM to explain the model. The MFMs not only are widely used in statistical physics but also are carried out in computationally or analytically unmangeable problems because of its effectiveness. MFMs are based on stochastic differential equations that include random fluctuations and consider the Gaussian approximation to the population. These models describe the dynamical behavior of the network and allow the explicit simulation of neuronal and cortical activity and behavior (e.g., performance and reaction time) [48]. MFMs trace the activity of an excitatory population and an inhibitory population

and estimate the mean and variance of population activity, in contrast to NMMs that consider only the mean activity of neuronal populations [51]. These MFMs has been applied in various areas, including statistical physics [52], queueing theory [53], and game theory [54]. Therefore, MFMs for nonlinear dynamic models are appropriate to the data that reflect the behavior of a population of neurons, such as EEG. Consequently, we adopted MEMs for modeling the nonlinear EEG system.

For modeling the nonlinear EEG dynamic system, we need a nonlinear model, which reflects brain activity, and a method, which estimates the unknown parameters of a model. In this article, we used flexible MFMs than NMMs and UKF that were widely used for parameter estimation on the gray-box modeling for dynamic EEG system. Nonlinear models usually contain many unknown parameters, and parameter estimation problems on these nonlinear systems can often be solved by choosing a model and then iteratively improving it until good estimated values are obtained. Parameter estimation is an important step and acts as a descriptive measure of an entire population of MFMs.

EKF, PF, and UKF were used as the typical method for parameter estimation on nonlinear models. However, EKF is very complicated to implement, and it is limited to linear systems on time scaling during update period [35]. In addition, the EKF for parameter estimation provides primary approximation on optimum nonlinear estimation through linearization of nonlinear systems. This approximation can contain large errors when obtaining posterior mean and covariance of the changed Gaussian random variable [55]. Since PF uses probability to estimate the next state and performs the calculation in proportion to the number of particles, the particle number affects the performance and simulation time. PF can be applied to nonlinear or non-Gaussian systems without restrictive conditions. However, this method has the drawback that it augments computational complexity in proportion to the number of particles.

In this article, we compared the result with existing methods, EKF and PF, and the proposed method used UKF on gray-box model. As a conclusion, with respect to the comparison with UKF, EKF required the high run-time because of linearization by the Jacobian Transformation. In the case of the PF, it needs a number of particles for good performance, thus leading to long execution time. In this article, we used 500 as the particle number according to experimental results. Nevertheless, PF on gray-box model shows the high rms error rates than using EKF and UKF. UKF estimates parameter values using the observed data from nonlinear dynamic systems, then the estimated parameter values are then applied to the mathematical equation, and finally, new data were generated. We calculated the RMSE between the observed and predicted data that are generated by the gray-box model. As a result, the proposed method using the UKF showed faster convergence speed and excellent estimation performance.

The gray-box model has been applied in various domains, such as indoor temperature prediction [56], transformer winding phenomena analysis [57], computer systems performance prediction [58], electrical power consumption-to-temperature identification [59], cardiovascular disease prediction [60], and predictive analytics of the smart manufacturing field [61]. Setting initial values for parameters is not only a challenging problem but also very important step of parameter estimation of gray-box model. Since the final parameter estimation can change depending on the initial value setting, this matter should be treated carefully. The improper initial value of parameters may cause convergence to a local optimum, or it may delay convergence. Therefore, significant research has been performed to solve the initial problem arising from parameter estimation processes. However, these problems require user intervention, as well as complex formula calculations, and numerous throughputs. The UKF is widely known as a method to assure a global optimum in the parameter estimation. The method proposed in this article does not require any user intervention for the initial parameter value setting, as we estimate all parameter values on a gray-box model using the UKF. This method guaranteed rapid convergence, and it provided optimal results.

V. Conclusion

Most real dynamic systems are nonlinear. For the monitoring of nonlinear systems, such as the EEG, precise mathematical equations are based on real system requirements. In order to improve the performance of a system, it is crucial to determine the exact parameter values in equations system. In this article, the chaos neural population model was used in a gray-box model to model a nonlinear EEG dynamic system. This simplified neural information processing and the UKF were proposed to estimate the model parameters. The proposed method estimated optimal unknown parameters of EEG dynamic model and demonstrated outstanding performance compared to other methods. Furthermore, it was possible to estimate the parameter value fitting in the EEG dynamic system despite the initial parameter values being unknown. Therefore, our method is not affected by the initial value problem of parameters that occur frequently in dynamic systems. The benefit of the proposed method can estimate unknown parameters in the model with only observed data. This leads to not only eliminate the difficulties of decomposing the system to obtain the parameter values but also gain a more accurate dynamic system. In future work, we forward aim to extend our method for future value prediction in the nonlinear EEG dynamic system.

References

[1] T. E. Kam, H. I. Suk, and S. W. Lee, "Non-homogeneous spatial filter optimization for ElectroEncephaloGram (EEG)-based motor imagery classification," *Neurocomputing*, vol. 108, no. 2, pp. 58–68, 2013.

[2] B. J. Sohns, J. Allison, H. K. Fathy, and J. L. Stein, "Efficient parameterization of large scale dynamic models through the use of activity analysis," in *Proc. Int. Conf. Mech. Eng. Congr. Expo.*, Chicago, IL, USA, Nov. 2006, pp. 1147–1154.

[3] F. Ding, "Combined state and least squares parameter estimation algorithms for dynamic systems," *Appl. Math. Model.*, vol. 38, no. 1, pp. 403–412, Jan. 2014.

[4] F. Ding, L. Xu, F. E. Alsaadi, and T. Hayat, "Iterative parameter identification for pseudo-linear systems with ARMA noise using the filtering technique," *IET Control Theory Appl.*, vol. 12, no. 7, pp. 892–899, May 2018.

[5] F. Bouchereau and D. Brady, "Method-of-moments parameter estimation for compound fading processes," *IEEE Trans. Commun.*, vol. 56, no. 2, pp. 166–172, Feb. 2008.

[6] V. Strejc, "Least squares parameter estimation," *Automatica*, vol. 16, no. 5, pp. 535–550, Sep. 1980.

[7] K. F. Chen, "Estimating parameters of a sine wave by separable nonlinear least squares fitting," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 12, pp. 3214–3217, Dec. 2010.

[8] S. H. Chan. *Lecture 6: Minimum Variance Unbiased Estimators*. Accessed: Jan. 25, 2020. [Online]. Available: https://engineering.purdue.edu/ChanGroup/ECE645Notes/StudentLecture06.pdf

[9] S. Zheng, *Statistical Theory II Maximum Likelihood Estimation*. Accessed: Jan. 25, 2020. [Online]. Available: http://people.missouristate.edu/songfengzheng/Teaching/MTH541/Lecture%20notes/MLE.pdf

[10] Y. Xia, S. Kanna, and D. P. Mandic, "Maximum likelihood parameter estimation of unbalanced three-phase power signals," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 3, pp. 569–581, Mar. 2018.

[11] T. S. Ferguson, "An inconsistent maximum likelihood estimate," *J. Amer. Stat. Assoc.*, vol. 77, no. 380, pp. 831–834, Dec. 1982.

[12] S. Chandrasekaran, *Offshore Structural Engineering: Reliability and Risk Assessment*, Boca Raton, FL, USA: CRC Press, 2017.

[13] D. Gorinevsky, "An approach to parametric nonlinear least square optimization and application to task-level learning control," *IEEE Trans. Autom. Control*, vol. 42, no. 7, pp. 912–927, Jul. 1997.

[14] X. Wang *et al.*, "Performance enhancement of pharmacokinetic diffuse fluorescence tomography by use of adaptive extended Kalman filtering," *Comput. Math. Methods Med.*, vol. 2015, no. 1, pp. 1–15, 2015.

[15] X. Sun, L. Jin, and M. Xiong, "Extended Kalman filter for estimation of parameters in nonlinear state-space models of biochemical networks," *PLoS ONE*, vol. 3, no. 11, p. e3758, Nov. 2008.

[16] V. A. Bavdekar, A. P. Deshpande, and S. C. Patwardhan, "Identification of process and measurement noise covariance for state and parameter estimation using extended Kalman filter," *J. Process Control*, vol. 21, no. 4, pp. 585–601, Apr. 2011.

[17] H. Ebrahimian, R. Astroza, and J. P. Conte, "Extended Kalman filter for material parameter estimation in nonlinear structural finite element models using direct differentiation method," *Earthquake Eng. Struct. Dyn.*, vol. 44, no. 10, pp. 1495–1522, Aug. 2015.

[18] H. M. Moradkhani and K. L. Hsu, "Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter," *Water Resour. Res.*, vol. 41, no. 5, pp. 1–17, 2005.

[19] W. P. Wang, S. Liao, and T. W. Xing, "Particle filter for state and parameter estimation in passive ranging," in *Proc. Int. Conf. Comput. Inf. Sci.*, Shanghai, China, Nov. 2009, pp. 257–261.

[20] N. Kantas, A. Doucet, S. S. Singh, J. Maciejowski, and N. Chopin, "On particle methods for parameter estimation in state-space models," *Statist. Sci.*, vol. 30, no. 3, pp. 328–351, Aug. 2015.

[21] A. López-Cuevas, B. Castillo-Toledo, L. Medina-Ceja, and C. Ventura-Mejía, "State and parameter estimation of a neural mass model from electrophysiological signals during the status epilepticus," *NeuroImage*, vol. 113, pp. 374–386, Jun. 2015.

[22] X. Liu and Q. Gao, "Parameter estimation and control for a neural mass model based on the unscented Kalman filter," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 88, no. 4, 2013, Art. no. 042905.

[23] E. P. Lara, G. O. Jordi, and J. P. Anto, "Extracranial estimation of neural mass model parameters using the unscented Kalman filter," *Frontiers Appl. Math. Stat.*, vol. 4, no. 46, pp. 1–15, 2018.

[24] C. Sahin, S. N. Ogulata, K. Aslan, H. Bozdemir, and R. Erol, "A neural network-based classification model for partial epilepsy by EEG signals," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 22, no. 5, pp. 973–985, Aug. 2008.

[25] J.-H. Kim, F. Biessmann, and S.-W. Lee, "Decoding three-dimensional trajectory of executed and imagined arm movements from electroencephalogram signals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 5, pp. 867–876, Sep. 2015.

[26] T. Mckenna, T. Mcmullen, and M. Shlesinger, "The brain as a dynamic physical system," *Neuroscience*, vol. 60, no. 3, pp. 587–605, Jun. 1994.

[27] S. Julier and J. K. Uhlmann, "A new extension of the Kalman filter to nonlinear system," in *Proc. SPIE*, vol. 3068, pp. 182–193, 1997.

[28] S. Julier and J. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.

[29] J. Hauth, "Grey-box modelling for nonlinear systems," Ph.D. dissertation, Univ. Kaiserslautern, Kaiserslautern, Germany, 2008. [Online]. Available: https://kluedo.ub.uni-kl.de/files/2045/diss.pdf

[30] N.-S. Kwak, K.-R. Müller, and S.-W. Lee, "A convolutional neural network for steady state visual evoked potential classification under ambulatory environment," *PLoS ONE*, vol. 12, no. 2, Feb. 2017, Art. no. e0172578.

[31] R. Sole, B. Goodwin, and R. Sol, *Signs of Life: How Complexity Pervades Biology*. New York, NY, USA: Basic Books, 2000.

[32] P. J. Garc and C. C. L. Pe, "Analysis of EEG signals using nonlinear dynamics and chaos: A review," *Appl. Math. Inf. Sci.*, vol. 9, no. 5, pp. 2309–2321, 2015.

[33] H. G. Schuster and P. Wagner, "A model for neuronal oscillations in the visual cortex: 1. Mean-field theory and derivation of the phase equations," *Biol. Cybern.*, vol. 64, no. 1, pp. 77–82, Nov. 1990.

[34] H. R. Wilson and J. D. Cowan, "Excitatory and inhibitory interactions in localized populations of model neurons," *Biophysical J.*, vol. 12, no. 1, pp. 1–24, Jan. 1972.

[35] S. Julier, J. Uhlmann, and H. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. Autom. Control*, vol. 45, no. 3, pp. 477–482, Mar. 2000.

[36] D. Robertson and J. Lee, "A least squares formulation for state estimation," *J. Process Control*, vol. 5, no. 4, pp. 291–299, Aug. 1995.

[37] Kaggle Inc. *American Epilepsy Society Seizure Prediction Challenge*. Accessed: Jan. 25, 2020. [Online]. Available: https://www.kaggle.com/c/seizure-prediction

[38] F. J. Muñoz-Almaraz, F. Zamora-Martínez, P. Botella-Rocamora, and J. Pardo, "Supervised filters for EEG signal in naturally occurring epilepsy forecasting," *PLoS ONE*, vol. 12, no. 6, Jun. 2017, Art. no. e0178808.

[39] J. J. Howbert *et al.*, "Forecasting seizures in dogs with naturally occurring epilepsy," *PLoS ONE*, vol. 9, no. 1, Jan. 2014, Art. no. e81920.

[40] B. H. Brinkmann *et al.*, "Forecasting seizures using intracranial EEG measures and SVM in naturally occurring canine epilepsy," *PLoS ONE*, vol. 10, no. 8, Aug. 2015, Art. no. e0133900.

[41] K. Dróżdż and K. Szabat, "Application of unscented Kalman filter in adaptive control structure of two-mass system," in *Proc. IEEE Int. Power Electron. Motion Control Conf.*, Varna, Bulgaria, Sep. 2016, pp. 1150–1154.

[42] A. Doucet, J. F. G. Freitas, and N. J. Gordon, "An introduction to sequential Monte Carlo methods," in *Sequential Monte Carlo Methods in Practice*. New York, NY, USA: Springer-Verlag, 2001.

[43] C. Nemeth, P. Fearnhead, and L. Mihaylova, "Sequential Monte Carlo methods for state and parameter estimation in abruptly changing environments," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1245–1255, Mar. 2014.

[44] X. Wang, T. Li, S. Sun, and J. Corchado, "A survey of recent advances in particle filters and remaining challenges for multitarget tracking," *Sensors*, vol. 17, no. 12, p. 2707, Nov. 2017.

[45] J. W. Park, C. K. Yang, and D. S. Shim, "Particle filter performance for ultra-tightly GPS/INS integration," *J. Inst. Control, Robot. Syst.*, vol. 14, no. 8, pp. 785–791, 2008.

[46] M. Raitoharju and R. Piche, "On computational complexity reduction methods for Kalman filter extensions," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 34, no. 10, pp. 2–19, Oct. 2019.

[47] D. Pinotsis, P. Robinson, P. B. Graben, and K. Friston, "Neural masses and fields: Modeling the dynamics of brain activity," *Frontiers Comput. Neurosci.*, vol. 8, no. 149, pp. 1–3, 2014.

[48] G. Deco, V. K. Jirsa, P. A. Robinson, M. Breakspear, and K. Friston, "The dynamic brain: From spiking neurons to neural masses and cortical fields," *PLoS Comput. Biol.*, vol. 4, no. 8, Aug. 2008, Art. no. e1000092.

[49] T. Schwalger, M. Deger, and W. Gerstner, "Towards a theory of cortical columns: From spiking neurons to interacting neural populations of finite size," *PLoS Comput. Biol.*, vol. 13, no. 4, Apr. 2017, Art. no. e1005507.

[50] S. Coombes. *Neural Fields*. Accessed: Jan. 25, 2020. [Online]. Available: http://www.scholarpedia.org/article/Neural_fields

[51] A. S. C. C. Marreiros, "Dynamic models of brain imaging data and their Bayesian inversion," Ph.D. dissertation, Univ. College London, London, U.K., 2010.

[52] L. P. Kadanoff, "More is the same; phase transitions and mean field theories," *J. Stat. Phys.*, vol. 137, nos. 5–6, pp. 777–797, Dec. 2009.

[53] N. D. Vvedenskaya, R. L. Dobrushin, and F. I. Karpelevich, "Queueing system with selection of the shortest of two queues: An asymptotic approach," *Problemy Peredachi Informatsii*, vol. 32, no. 1, pp. 20–34, 1996.

[54] J.-M. Lasry and P.-L. Lions, "Mean field games," *Jpn. J. Math.*, vol. 2, no. 1, pp. 229–260, 2007.

[55] E. A. Wan and R. van der Merwe, *The Unscented Kalman Filter, Kalman Filtering and Neural Networks*. Hoboken, NJ, USA: Wiley, 2001.
[56] P. Hietaharju, M. Ruusunen, and K. Leiviskä, "A dynamic model for indoor temperature prediction in buildings," *Energies*, vol. 11, no. 6, p. 1477, Jun. 2018.
[57] E. Rahimpour, V. Rashtchi, and R. Aghmasheh, "Parameters estimation of transformers gray box model," in *Proc. Int. Conf. Modern Electr. Energy Syst.*, Kremenchuk, Ukraine, 2017, pp. 372–375.
[58] D. Didona, "Gray box performance modeling of in-memory distributed transactional platforms," Ph.D. dissertation, Dept. Inf. Syst. Comput. Eng., IST Univ. Lisbon, Lisbon, Portugal, 2015. [Online]. Available: https://www.gsd.inesc-id.pt/didona/website/DiegoDidona-PhDThesis.pdf
[59] F. Sossan, V. Lakshmanan, G. T. Costanzo, M. Marinelli, P. J. Douglass, and H. Bindner, "Grey-box modelling of a household refrigeration unit using time series data in application to demand side management," *Sustain. Energy, Grids Netw.*, vol. 5, pp. 1–12, Mar. 2016.
[60] M. Hwang, C. H. Leem, and E. B. Shim, "Toward a grey box approach for cardiovascular physiome," *Korean J. Physiol. Pharmacol.*, vol. 23, no. 5, pp. 305–310, 2019.
[61] Z. Yang *et al.*, "Investigating grey-box modeling for predictive analytics in smart manufacturing," in *Proc. Int. Conf. Design Eng. Tech. Comput. Inf. Eng.*, Cleveland, OH, USA, 2017, Art. no. V02BT03A024.

Sun-Hee Kim received the B.S. degree in multimedia from the Korean Educational Development Institute, Seoul, South Korea, in 2004, the M.S. degree in computer science from Dongguk University, Seoul, in 2006, and the Ph.D. degree in computer science from Chonnam National University, Gwangju, South Korea, in 2011.

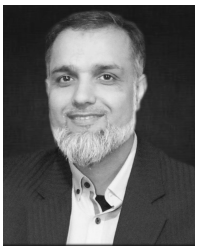
From 2011 to 2013, she held a post-doctoral position at Carnegie Mellon University, Pittsburgh, PA, USA, where she worked on data mining problems in bio-signal with the Database Group. She is currently a Research Professor with the Department of Brain and Cognitive Engineering, Korea University, Seoul. Her current research interests are data mining, machine learning, bioinformatics, and its applications.

Hyung-Jeong Yang received the B.S., M.S., and Ph.D. degrees from Chonbuk National University, Jeonju, South Korea.

She is currently an Associate Professor with Chonnam National University, Gwangju, South Korea. Her main research interests include multimedia data mining, pattern recognition, artificial intelligence, e-learning, and e-design.

Ngoc Anh Thi Nguyen received the B.S. degree from the Faculty of Mathematics and Informatics, The University of Da Nang, Da Nang, Vietnam, the M.S. and the Ph.D. degrees in electronics and computer engineering from the Department of Electronics and Computer Engineering, Chonnam National University, Gwangju, South Korea.

She is currently a Lecturer with the Faculty of Information Technology, The University of Da Nang—University of Education. Her research interests include data mining, machine learning, data analysis, signal processing, pattern recognition, and mathematical modeling, mainly working on time series mining, applying on bioinformatics and neuroscience.



Raja Majid Mehmood received the B.S. degree in computer science from Gomal University, Dera Ismail Khan, Pakistan, in 2004, the M.S. degree in software technology from Linnaeus University, Vaxjo, Sweden, in 2010, and the Ph.D. degree in computer engineering from the Division of Computer Science and Engineering, Chonbuk National University, Jeonju, South Korea.

From April 2011 to February 2014, he was a Lecturer with the Software Engineering Department, King Saud University, Riyadh, Saudi Arabia. He has served as a Research Professor with the Department of Brain and Cognitive Engineering, Korea University, Seoul, South Korea. He is currently an Assistant Professor with the School of Electrical and Computer Engineering, Information and Communication Department, Xiamen University Malaysia, Sepang, Malaysia. His main research interests include affective computing, brain–computer interfaces, information visualization, image processing, and pattern recognition.

Seong-Whan Lee (Fellow, IEEE) received the B.S. degree in computer science and statistics from Seoul National University, Seoul, South Korea, in 1984 and the M.S. and Ph.D. degrees in computer science from the Korea Advanced Institute of Science and Technology, Seoul, in 1986 and 1989, respectively.

He is currently the Hyundai-Kia Motor Chair Professor and the Head of the Department of Brain and Cognitive Engineering, Korea University, Seoul. His research interests include pattern recognition, artificial intelligence, and brain engineering.

Dr. Lee is also a fellow of the International Association for Pattern Recognition (IAPR) and the Korea Academy of Science and Technology.