TU Delft

Delft University of Technology

Document Version
Final published version

Licence
CC BY

Citation (APA)
Tian, R., Yang, Y., van der Helm, F. C. T., & Dewald, J. P. A. (2018). A novel approach for modeling neural responses to joint perturbations using the NARMAX method and a hierarchical neural network. *Frontiers in Computational Neuroscience*, *12*, Article 96. https://doi.org/10.3389/fncom.2018.00096

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 06 December 2018
doi: 10.3389/fncom.2018.00096

A Novel Approach for Modeling Neural Responses to Joint Perturbations Using the NARMAX Method and a Hierarchical Neural Network

Runfeng Tian[1,2†], *Yuan Yang*[1,2*†], *Frans C. T. van der Helm*[1,2] *and Julius P. A. Dewald*[1,2,3]

[1] *Department of Physical Therapy and Human Movement Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States,* [2] *Department of Biomechanical Engineering, Northwestern University, Evanston, IL, United States,* [3] *Department of Biomechanical Engineering, Delft University of Technology, Delft, Netherlands*

The human nervous system is an ensemble of connected neuronal networks. Modeling and system identification of the human nervous system helps us understand how the brain processes sensory input and controls responses at the systems level. This study aims to propose an advanced approach based on a hierarchical neural network and non-linear system identification method to model neural activity in the nervous system in response to an external somatosensory input. The proposed approach incorporates basic concepts of Non-linear AutoRegressive Moving Average Model with eXogenous input (NARMAX) and neural network to acknowledge non-linear closed-loop neural interactions. Different from the commonly used polynomial NARMAX method, the proposed approach replaced the polynomial non-linear terms with a hierarchical neural network. The hierarchical neural network is built based on known neuroanatomical connections and corresponding transmission delays in neural pathways. The proposed method is applied to an experimental dataset, where cortical activities from ten young able-bodied individuals are extracted from electroencephalographic signals while applying mechanical perturbations to their wrist joint. The results yielded by the proposed method were compared with those obtained by the polynomial NARMAX and Volterra methods, evaluated by the variance accounted for (VAF). Both the proposed and polynomial NARMAX methods yielded much better modeling results than the Volterra model. Furthermore, the proposed method modeled cortical responded with a mean VAF of 69.35% for a three-step ahead prediction, which is significantly better than the VAF from a polynomial NARMAX model (mean VAF 47.09%). This study provides a novel approach for precise modeling of cortical responses to sensory input. The results indicate that the incorporation of knowledge of neuroanatomical connections in building a realistic model greatly improves the performance of system identification of the human nervous system.

Keywords: neural modeling, NARMAX, neural network, EEG, non-linear system identification

OPEN ACCESS

Edited by:
Yilei Zhang,
Nanyang Technological University, Singapore

Reviewed by:
Adam B. Barrett,
University of Sussex, United Kingdom
Ajitkumar Mulavara,
Universities Space Research Association (USRA), United States

***Correspondence:**
Yuan Yang
yuan.yang@northwestern.edu

†These authors have contributed equally to this work

Received: *08 September 2018*
Accepted: *21 November 2018*
Published: *06 December 2018*

Citation:
Tian R, Yang Y, van der Helm FCT and Dewald JPA (2018) A Novel Approach for Modeling Neural Responses to Joint Perturbations Using the NARMAX Method and a Hierarchical Neural Network. Front. Comput. Neurosci. 12:96. doi: 10.3389/fncom.2018.00096

INTRODUCTION

The human nervous system is an integrated, large-scale system consisting of connected neuronal networks. The coordination of neural activity across networks, between the periphery and the central nervous system, is essential for fulfilling our daily functions including movement control and sensory perception. Modeling and system identification of the human nervous system helps us understand how the brain processes sensory input and controls behavior. Efforts have been made through previous studies on building a mathematical model of the human nervous system from neurons and circuits to large-scale neural networks. As yet, challenges remain, because of substantial non-linearity and fast dynamics in the nervous system (Breakspear, 2017). The non-linearity of the nervous system enables the rich, task-relevant neural encoding, and communication, while the fast dynamics (including both linear and non-linear components) allow for efficient processing and transmission of neural information (Friston, 2000). However, these two basic properties of the human nervous system, as well as the poor signal-to-noise of measured neural signals, significantly increase the difficulty of building a precise model to describe the behavior of the human nervous system.

A typical way to investigate the input-output behavior of the human nervous system is to apply an external peripheral input and to measure the neural response from brain. Electroencephalography (EEG) is a non-invasive electro-neurophysiological technique widely applied to measure neural responses from the brain to external inputs. When providing periodic stimulation with specific frequencies, a linear system only generates phase-locked responses at the stimulated frequencies while a non-linear system can produce cross-frequency phase-coupled responses at non-stimulated frequencies (Langdon et al., 2011; Billings, 2013; Yang et al., 2016b). When applying angular position perturbations to the wrist joint, a recent study demonstrated that more than 80% signal power of brain activity resulting from the movement stimulation occurs in the non-stimulated frequencies and only 10% of brain activity could be explained using the best linear approximation model (Vlaar et al., 2017). As a result, a non-linear approach is necessary for modeling the brain response to wrist joint perturbations. Furthermore, previous EEG studies have reported sub-harmonic responses (i.e., the response frequency is a fraction of a stimulus frequency) in the brain activity following joint perturbations, indicating complicated non-linear dynamics of cortical oscillations (Yang et al., 2016a; Breakspear, 2017). Such an ill-posed non-linear phenomenon cannot be explained by classical non-linear models such as a Wiener and Hammerstein system configuration (i.e., a series connection of static non-linear blocks with a dynamic linear block; Crama and Schoukens, 2001; Paduart et al., 2012), though they have been previously used for system identification and modeling of the periphery system including human musculoskeletal systems (Westwick and Kearney, 2001; Dempsey and Westwick, 2004).

The human nervous system comprised of multiple neuronal circuits resulting in a complicated closed-loop non-linear system with fast dynamics (in the order of millisecond). Previous studies tried to use a simplified approach based on Volterra series to model the brain response to joint perturbations. This only explained around 40% of the measured EEG signal (Vlaar et al., 2018). Volterra series is a polynomial functional expansion similar to a Taylor series that provides an approximation of input-output relation in a non-linear system (Brockett, 1976). A serious limitation of Volterra series based approaches is that it completely ignores the closed-loop behaviors of the nervous system as they contain no autoregressive terms. To address this limitation, we proposed to use the non-linear autoregressive moving average with exogenous inputs (NARMAX) model (Chen and Billings, 1989; Chen et al., 1989). The NARMAX model is a powerful tool for black-box system identification problems, in particular when limited knowledge about the detailed model structure of the system is available. A wide range of non-linear systems can be well represented using the NARMAX method, including those with exotic non-linear behaviors such as subharmonics, bifurcations, and even chaos, as previously reported for the human nervous system (Breakspear, 2017).

The classical NARMAX model is based on polynomial expansions, which makes it difficult to precisely mimic properties of the nervous system such as the sigmoid function of a synapse. The model error may cumulatively increase when the neural pathway contains more than one synapse as in the somatosensory afferent pathways from muscles to the brain. To address this challenge, we replaced the polynomial non-linear terms with a hierarchical neural network. The hierarchical neural network is built based on known neuroanatomical connections and corresponding transmission delays in neural pathways. The proposed method is evaluated using an experimental dataset, where brain responses to joint perturbations at the wrist in able-bodied individuals were obtained by EEG signals.

DATA ACQUISITION AND PREPROCESSING

Data Acquisition

The experimental data were recorded from ten young able-bodied individuals (age range 22–25 years old; 6 men; all right-handedness). The experimental procedure was approved by the Human Research Ethics Committee of the Delft University of Technology, the Netherlands. All participants signed informed consent before the experiments.

During the experiment, subjects sat next to a wrist manipulator (WM), which is an actuated rotating device with a single degree of freedom to exert flexion and extension perturbations to wrist joint (Wristalyzer, Moog Inc., The Netherlands). The lower arm (in the dominant side) of subjects was strapped in the armrest while the hand closely touched the handle of the manipulator and fixed with velcro. The perturbations were applied as flexion and extension stretches to the participants as the external input to the nervous system via the handle of the manipulator. Participants were instructed to

relax their arm and fingers, and not react to the perturbations during the experiment.

The stimulation signals consisted of the sum of sinusoids with the frequencies of 1, 3, 5, 7, 9, 11, 13, 15, 19, and 23 Hz and with the period of 1 s. Seven different sum-of-sinusoidal signals (with the same frequencies) were generated based on different relative phases between sinusoids. All stimulation signals have similar statistical distributions with the same root-mean-square (i.e., the square root of the arithmetic mean of the squares of time series) that equals to 0.02 radians. The simulation signal was applied to wrist joint as angular position perturbations, therefore the unit of signal magnitude is in radians. The signals were designed to have the equal power on the first three frequency components (i.e., 1, 3, 5 Hz) and a decaying power spectrum (−20dB/decade slope) for rest frequency components. This design is a trade-off between reduced predictability of signal (to prevent the anticipation of participants during the experiment) and capabilities of the human wrist joint and the manipulator (Vlaar et al., 2017). Each signal was applied to stimulate the wrist joint for 7 trials of 36 s per trial.

EEG was recorded from the nervous system, using a 128-channel cap (5/10 systems, WaveGuard, ANT Neuro, Germany) with Ag/AgCl electrodes, using a common average reference. The EEG and movement stimulation were digitalized at 2,048 Hz using a Refa amplifier (Twente Medical Systems International B.V., the Netherlands) and stored for further analysis. Three seconds in the beginning and the end of each trial were removed from further analysis to reduce the effect of transient dynamics, resulting in $30 \times 7 = 210$ periods for each participant and each stimulation.

Data Preprocessing

We used Independent component analysis (ICA) (Makeig et al., 1996) to extract the cortical source activities as the output of the nervous system for modeling purposes. Before applying ICA, the continuous EEG signals were filtered by a 1–100 Hz zero-phase shift band-pass filter to remove possible high-frequency noise from neck muscles and slow trends in the data (e.g., blood pressure, heartbeat, breathing and sweat potentials). Notch filters implemented in Fieldtrip toolbox (Oostenveld et al., 2011) were applied to remove the 50 Hz line power noise and its harmonic distortions.

ICA was performed using the Infomax algorithm (Bell and Sejnowski, 1995) to decompose filtered EEG signals into independent source components by minimizing of mutual information among the data projections or maximizing joint entropy (Raimondo et al., 2012). Subsequently, all signals were resampled to 256 Hz ($N = 256$ samples) and then were segmented into 1-s periods according to the perturbation signal. The signal-to-noise ratio (SNR) of each ICA component was calculated using the algorithm developed by Vlaar et al. (2015). The ICA component with the highest SNR was selected as the system output for each dataset for the modeling. A dipole fitting algorithm implemented in the Fieldtrip toolbox (Oostenveld et al., 2011) was used to estimate the source locations of selected ICA components. The procedure is the same as the previous modeling study using Volterra model in the same datasets (Vlaar et al., 2018).

MODELING BASED ON NARMAX AND HIERARCHICAL NEURAL NETWORK

NARMAX Method

The non-linear autoregressive moving average with exogenous inputs (NARMAX) method provides a generalized framework for mathematical modeling of a non-linear closed-loop system (Chen and Billings, 1989; Chen et al., 1989). The input-output relationship of a non-linear dynamic system can be generally represented using the NARMAX method as follows:

$$\begin{aligned} y(k) = f(&y(k-1), \ldots, y(k-n_y), u(k-d), \\ &u(k-d-1), \ldots, u(k-d-n_u), \\ &e(k-1), \ldots, e(k-n_e)) \end{aligned} \quad (1)$$

where k is the time sample, $y(k)$, $u(k)$, and $e(k)$ are the output, input and prediction error, respectively; n_y, n_u, and n_e are the associated maximum memory lags, d is the shortest input-output time delay in the system and $f(\cdot)$ is a non-linear function. The NARMAX method has been previously applied to complex real-world problems in various research fields including medical (Billings et al., 2013) and neurophysiological sciences (Li et al., 2015). The classical NARMAX model uses a polynomial expansion (e.g., $u(k-d)\,u\,(k-d)+ u(k-d)\,u\,(k-d-1) + \ldots + u(k-d-n_u)\,u\,(k-d-n_u) + y(k-1)\,u\,(k-d) + \ldots + y\,(k-n_y)\,u(k-d-n_u) +$

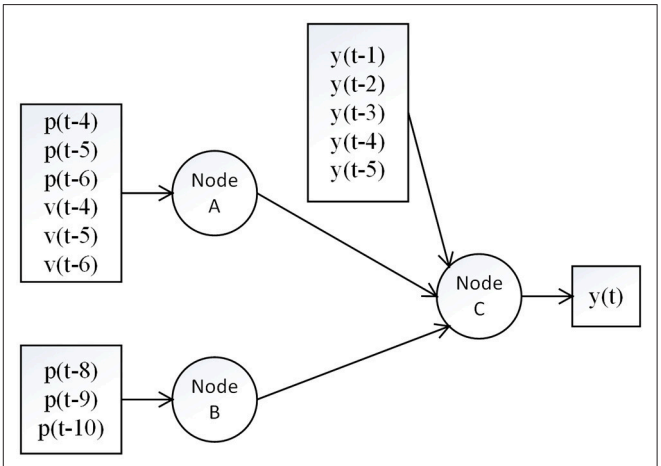

FIGURE 1 | Structure of the proposed model. The signals *p, v* and *y* represent position input, velocity input and feedback interaction, respectively. Node A and B are in the first layer at the medulla, and Node C is in the second layer at the thalamus. The pathway from A to C indicates the group Ia afferent pathway which transmits both velocity and position information with a short time delay, while the pathway from B to C indicates the group II afferent pathway which transmits only the position information with a long time delay.

$y(k-1)\,y(k-1) + y(k-1)\,y(k-2) + \ldots)$ as the non-linear function. Although most non-linear functions can be represented by this polynomial expansion, it is difficult to precisely mimic the neuronal behavior of the nervous system as the sigmoid function. Thus, we proposed to use a hierarchical neural network to replace the non-linear polynomial expansion in the NARMAX framework for modeling the neural response.

NARMAX Framework Based Hierarchical Neural Network (NARMAX-HNN) Model

The proposed model is built based on known neuroanatomical connections and corresponding transmission delays in neural pathways for processing the somatosensory information received from the proprioceptors in the periphery during the movement stimulation (Carpenter and Sutin, 1983; Standring, 2016). Both position and velocity of the applied joint perturbations are sensed by the muscle spindles and then transmitted by primary (group Ia) and secondary (group II) afferents to the central nervous system. Thus, we used both position (the perturbation signal) and velocity (its first-order derivative) signals as the inputs in the modeling. The Ia afferent pathway transmits both position (p) and velocity (v) information, while II afferent pathway transmits only position (v) information (Carpenter and Sutin, 1983; Standring, 2016). The NARMAX framework, therefore, is extended as:

$$\begin{aligned} y(k) = f(&y(k-1),\ldots,y(k-n_y),\, v(k-d_1),\\ &v(k-d_1-1),\ldots,v(k-d_1-n_1),\, p(k-d_1),\\ &p(k-d_1-1),\ldots,p(k-d_1-n_1),\, p(k-d_2),\\ &p(k-d_2-1),\ldots,p(k-d_2-n_2),\\ &e(k-1),\ldots,e(k-n_e)) \end{aligned} \quad (2)$$

for this dual-input-single-output model, where $v(k)$ represents the velocity signal and $p(k)$ the position signal, d_1 is the shortest delay in the group Ia afferent pathway for transmitting the velocity and position signals, and d_2 is the shortest delay in the group II afferent pathway for transmitting the position signal. The parameters of n_1 and n_2 indicate the system memory in group Ia and II afferent pathways, respectively. The Ia afferent pathway has the fastest nerve conduction velocity with the shortest delay from the periphery (wrist muscle receptors) to the brain. Previously electric-neurophysiological studies demonstrated that the shortest transmission delay of brain response to the somatosensory stimulation is around 16–20 ms for young able-bodied individuals (Desmedt and Cheron, 1980; Buchner et al., 1995; Campfens et al., 2015). The nerve conduction velocity for group Ia afferent pathway (80–120 m/s) is around twice of group II afferent pathway (33–75 m/s) (Buchthal and Rosenfalck, 1966; Siegel and Sapru, 2006). Thus, we set the $d_1 = 4$ (around 16 ms for sampling rate 256 Hz), $d_2 = 8$ (around 32 ms), $n_1 = n_2 = 2$ (around 8 ms).

There are two synaptic connections in the afferent pathways from muscle spindles to the brain. The first synaptic connection is in the medulla (nucleus cuneatus) and the second one is in the thalamus (ventral posterior nucleus; Carpenter and Sutin, 1983). Thus, we built a hierarchical neural network with two layers. The first layer represents the synaptic connection in the medulla and the second layer represents the synaptic connection in the thalamus. The position and velocity information is thought to be integrated in the thalamus (Kandel et al., 2000). The structure of the proposed model is provided in **Figure 1**.

The synaptic behavior was modeled with a sigmoid function S(u) = 1/(1+exp(-ρ(u-w))-1/(1+exp(ρw))) (Moran et al., 2007) with slope parameters ρ = 0.8 and firing threshold $w = 1.8$ for nodes A and B in the first layer, and ρ =1.6, $w = 1.8$ for the node C in the second layer (see **Figure 2**), which was in line with a previous neural modeling study in the human somatosensory

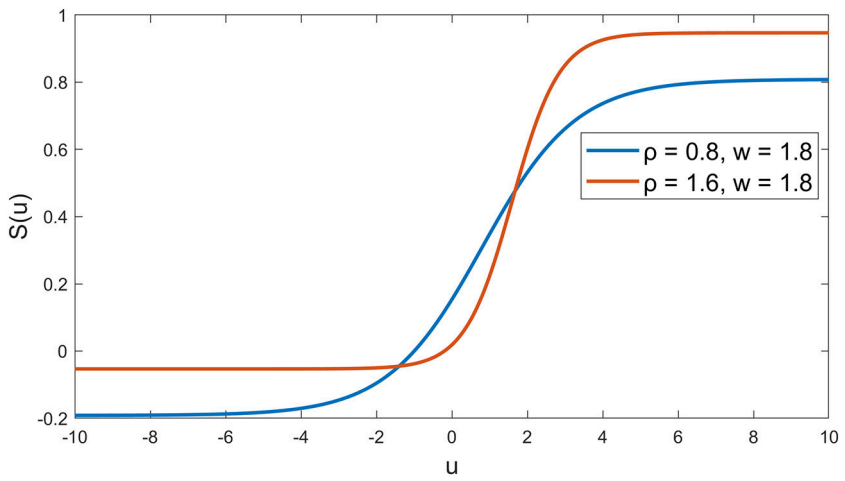

FIGURE 2 | Sigmoid functions for the nodes. Blue curve indicates the synaptic behavior in the first layer at the medulla and red curve indicates the synaptic behavior in the second layer at the thalamus (Marreiros et al., 2008).

pathway (Marreiros et al., 2008). The autoregressive terms y(k-1), …, y(k-n_y) was added to the second layer, which represents the feedback interaction in neural circuits of the thalamocortical radiation, where the n_y was set to be 5 (around 20 ms). The total number of parameters in our model is 19, including 6 weights for p and v through node A, 3 weights for p through node B, 5 weights for y and 2 weights for node A and B through node C, and another 3 bias parameters for nodes A, B and C.

The model was built and trained using Neural Network Toolbox in MATLAB R2018a. The training function was selected as the Scaled Conjugate Gradient (using MATLAB function *trainscg.m*), and the performance function was selected as the Mean Squared Error (using MATLAB function *mse.m*). The signals p, v, and y, as well as output of the nodes, were normalized to values between −1 and 1 to make them in the same scale.

Model Evaluation

Performances of the proposed model were evaluated by the Variance Accounted For (VAF) in a cross-validation test:

$$VAF = \left[1 - \frac{var\left(\hat{y} - y\right)}{var\left(y\right)}\right] \times 100\% \quad (3)$$

where $\hat{y}$ represents the estimated output by our model and y represents the measured output from the system. For each participant, six trials were used to train the model and the rest one trial was used to test the performance of the estimated model. For each trial, there is 30 periods of 1 s data with 256 points per second (sampling rate: 256 Hz). The training and testing procedures were repeated seven times, using each trial as the testing data in the cross-validation. The mean VAF across seven repetitions was reported for each participant as the performance of the proposed model.

The obtained results were compared with those from using the polynomial NARMAX (Chen and Billings, 1989) and Volterra (Vlaar et al., 2018) models. In the Volterra model, the output depends *only* on the history of the input to the non-linear system:

$$y\left(k\right) = f(u\left(k-d\right),\ u\left(k-d-1\right),\ldots,\ u\left(k-\ d-n_u\right),\\ e\left(k-1\right),\ldots,\ e\left(k-n_e\right)) \quad (4)$$

Thus, the Volterra model can be considered as a special case of the NARMAX model, which does not contain any autoregressive (e.g., y(k-1)y(k-1)) or interaction (e.g., y(k-1)u(k-d)) model terms.

For the NARMAX models (both the proposed model and the polynomial NARMAX model), we compared the estimated outputs from one-step-ahead (OSA) and multi-steps-ahead

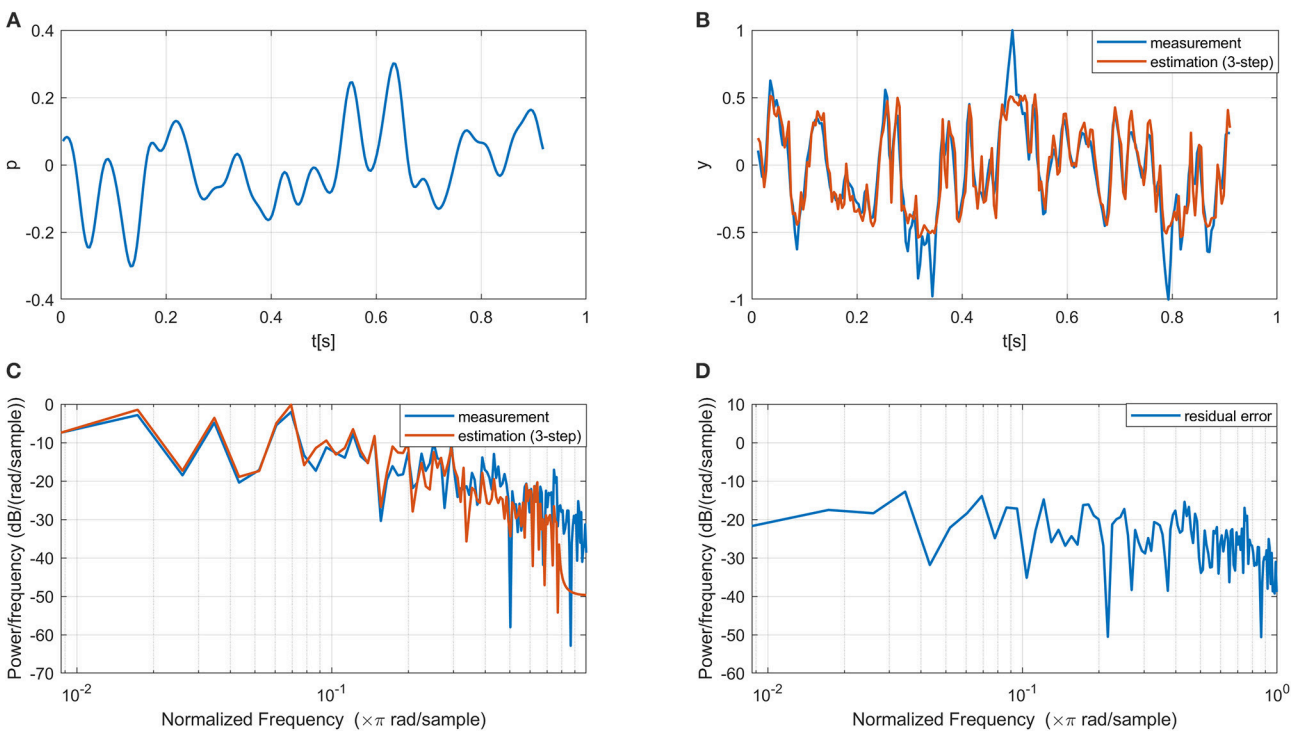

FIGURE 3 | Example of the input, measured output, estimated output, and residual error signals using our proposed NARMAX-HNN model. **(A)** A period of the input signal, **(B)** Comparison between estimated and measured output time series, **(C)** Comparison between estimated and measured output power spectrum density **(D)** Power spectrum density of the residual error.

TABLE 1 | Comparison of model performances, evaluated by variance accounted for in percentage (%), for the proposed NARMAX-HNN model, polynomial NARMAX (NARMAX-NP), and Volterra models.

Participants	NARMAX-HNN (OSA)	NARMAX-HNN (MSA, 3-step)	NARMAX-NP (OSA)	NARMAX-NP (MSA, 3-step)	Volterra
P1	94.37	63.44	95.52	57.08	38.37
P2	92.83	56.85	94.74	39.53	29.12
P3	90.95	67.16	92.95	31.17	32.18
P4	91.02	74.89	91.94	32.26	28.10
P5	92.58	82.31	94.04	61.57	53.74
P6	93.76	75.55	93.72	49.18	61.07
P7	93.08	74.32	95.73	65.35	54.30
P8	90.23	43.40	91.90	32.57	39.95
P9	90.36	77.16	92.24	37.98	26.35
P10	94.15	78.44	96.28	64.21	65.19
Mean	92.33	69.35	93.91	47.09	42.84
Std.	1.57	11.90	1.54	13.28	13.78

(MSA) predictions. The step-ahead prediction is defined as below:

(i) one-step-ahead model predicted output:

$$\begin{aligned}\hat{y}(k) = f(&y(k-1),\ldots,y(k-n_y),\\ &v(k-d_1),\ldots,v(k-d_1-n_1),p(k-d_1),\ldots,\\ &p(k-d_1-n_1),p(k-d_2),\ldots,p(k-d_2-n_2))\end{aligned} \tag{5}$$

(ii) two-step-ahead model predicted output:

$$\begin{aligned}\hat{y}(k+1) = f(&\hat{y}(k),y(k-1),\ldots,y(k-n_y),\\ &v(k+1-d_1),\ldots,v(k+1-d_1-n_1),\\ &p(k+1-d_1),\ldots,p(k+1-d_1-n_1),\\ &p(k+1-d_2),\ldots,p(k+1-d_2-n_2))\end{aligned} \tag{6}$$

(iii) three-step-ahead model predicted output:

$$\begin{aligned}\hat{y}(k+2) = f(&\hat{y}(k+1),\hat{y}(k),\\ &y(k-1),\ldots,y(k-n_y),v(k+2-d_1),\ldots,\\ &v(k+2-d_1-n_1),p(k+2-d_1),\ldots,\\ &p(k+2-d_1-n_1),p(k+2-d_2),\ldots,\\ &p(k+2-d_2-n_2))\end{aligned} \tag{7}$$

In this study, we evaluated the NARMAX model up to the three-step ahead prediction. Considering that dynamics of brain activity is typically in the order of few milliseconds, the three-step ahead prediction estimated 12 ms ahead EEG oscillations (based on the 256 Hz sampling rate), which is sufficient for testing the predictive performance of models. Noteworthy, for the Volterra model, there are no autoregressive variables. Consequently, it is not possible to generate the multi-step ahead prediction results for the Volterra model. The statistical significance of results is indicated by the paired t-test with $p < 0.05$.

RESULTS

In line with the previous modeling study on the same datasets (Vlaar et al., 2018), only one ICA component with the highest SNR was selected as the output of the nervous system for each dataset for modeling. All selected components have their sources located in the primary sensorimotor areas in the contralateral hemisphere (i.e., left hemisphere as all participants are right-hand dominant), indicating that the ICA results are neurophysiological plausible. **Figure 3** shows one period of the input and measured output signals from a typical subject. The estimated output signal from the multi-step (i.e., three-step) ahead prediction using the proposed NARMAX-HNN model is provided for comparison. The NARMAX-HNN captures the system's behavior well showing a similar waveform of the estimated output signal as the measured one. The residual error mainly appears in the high-frequency components, which is likely related to the fast dynamics of EEG oscillation as background noise in the system.

The model performances of NARMAX-HNN are summarized in the **Table 1** for all tested subjects. The results are compared with those obtained by using the polynomial NARMAX and Volterra methods. Both the proposed NARMAX-HNN method (VAF: 92.33% ± 1.57%) and polynomial NARMAX method (VAF: 93.91% ± 1.54%) generated significantly better results ($P < 0.0001$ for both NARMAX-HNN vs. Volterra and polynomial NARMAX vs. Volterra) than the Volterra method for the one-step ahead prediction (VAF: 42.84% ± 13.78%). Although the one-step prediction results are comparable for the NARMAX-HNN and the polynomial NARMAX, the proposed NARMAX-HNN (VAF: 69.35% ± 11.90%) method yielded a significant better ($P = 0.0015$) long-term prediction (three-step MSA) than the polynomial NARMAX method (VAF: 47.09% ± 13.28%). Furthermore, the NARMAX-HNN contains 19 parameters, which is less than the polynomial NARMAX (25 parameters) and Volterra model (46 parameters). Finally, the numbers of parameters, associated with the number of model terms in the polynomial NARMAX and Volterra model, are optimal numbers that were determined by using an

orthogonal forward regression (OFR) algorithm (Chen et al., 1989).

DISCUSSION AND CONCLUSION

This work proposes an advanced modeling approach based on the NARMAX framework and a hierarchical neural network to model cortical activity in response to flexion and extension stretch perturbations at the wrist. As the human nervous system is comprised of multiple neuronal circuits resulting in complicated closed-loop non-linear behaviors in response to an external input, a realistic model should contain autoregressive variables to consider this closed-loop behavior. The NARMAX framework contains autoregressive variables as well as the interaction between the feedback and the input, which better reflects the non-linear closed-loop behavior of the nervous system than the Volterra model. As we demonstrated, both the proposed method, i.e., NAMRAX-HNN, and polynomial NARMAX method generated significantly better results than the Volterra model.

As shown in **Table 1**, the long-memory effect of autoregressive model increases the accumulative error of the long-term prediction, which results in a drop in VAF for multi-step ahead prediction using NARMAX. It has to be pointed out that multi-step ahead prediction is still a recognized challenge in the field of time series forecasting (Hussein et al., 2016), especially for cortical activity due to its fast dynamics (as shown in **Figure 3**) and a poor signal-to-noise ratio (Breakspear, 2017). Different from the commonly used polynomial NARMAX method, the proposed approach replaced the polynomial non-linear terms with a hierarchical neural network. Beyond the classical non-linear system identification approaches based on artificial neural network (Nelles, 2013), the hierarchical neural network here is built based on known neuroanatomical connections and corresponding transmission delays in neural pathways [i.e., the dorsal columns, (Carpenter and Sutin, 1983)]. This biologically inspired innovation significantly improves the long-term prediction of NARMAX modeling, showing better performance than the polynomial NARMAX in the estimation of 12 ms ahead EEG oscillation. Furthermore, our results are also better than the previous modeling study on the same datasets using the Volterra model (Vlaar et al., 2018) as demonstrated in **Table 1**. Finally, our proposed method has lower model complexity with a smaller number of parameters than the polynomial NARMAX and Volterra models. Although one may improve the model prediction ability by using an advanced machine learning approaches, such as deep learning, the purely mathematical or data-driven approaches are not able to generate a biologically realistic model with a reduced number of parameters.

In conclusion, the proposed method provides a novel solution to modeling of neural responses in the human nervous system with greater precision than polynomial NARMAX and Volterra models, in particular for long-term predictions. The proposed method considers both neuroanatomical pathways and physiological properties of the human nervous system. Therefore, it allows for the generation of neuroanatomically realistic models, which goes beyond purely mathematical or data-driven approaches. This study therefore breaks new ground in neurobiological system identification and modeling. In the future, we will apply this method to investigate changes in sensory and motor pathways following a unilateral brain injury such as hemiparetic stroke or hemiparetic cerebral palsy.

AUTHOR CONTRIBUTIONS

RT performed the data analysis and modeling under the supervision of YY and JD. YY drafted the manuscript with the assistance of RT. YY, FvdH, and JD contributed in problem identification. FvdH provided the experimental data. YY and JD participated in editing the manuscript.

FUNDING

The data used in this study were collected in the Delft University of Technology, which was supported by the European Research Council under the ERC grant agreement n° 291339. The research leading to these results was supported by NIH R01HD039343, NIH UL1TR001422 Northwestern University (NU) Clinical and Translational Research Institute Voucher Program (Pilot study), and NU Department of Physical Therapy and Human Movement Sciences.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Martijn Vlaar for collecting and preprocessing the experimental data.

REFERENCES

Bell, A. J., and Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* 7, 1129–1159. doi: 10.1162/neco.1995.7.6.1129

Billings, C., Wei, H. L., Thomas, P., Linnane, S., and Hope-Gill, B. (2013). The prediction of in-flight hypoxaemia using non-linear equations. *Respir. Med.* 107, 841–847. doi: 10.1016/j.rmed.2013.02.016

Billings, S. A. (2013). *Nonlinear system identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*. Chichester: John Wiley & Sons, Ltd.

Breakspear, M. (2017). Dynamic models of large-scale brain activity. *Nat. Neurosci.* 20, 340–352. doi: 10.1038/nn.4497

Brockett, R. W. (1976). Volterra series and geometric control theory. *Automatica* 12, 167–176. doi: 10.1016/0005-1098(76)90080-7

Buchner, H., Adams, L., Müller, A., Ludwig, I., Knepper, A., Thron, A., et al. (1995). Somatotopy of human hand somatosensory cortex revealed by dipole source analysis of early somatosensory evoked potentials and 3D-NMR tomography. *Electroencephalogr. Clin. Neurophysiol.* 96, 121–134. doi: 10.1016/0168-5597(94)00228-7

Buchthal, F., and Rosenfalck, A. (1966). Evoked action potentials and conduction velocity in human sensory nerves. *Brain Res.* 3, v–vi,1–122. doi: 10.1016/0006-8993(66)90056-4

Campfens, S. F., Meskers, C. G., Schouten, A. C., Van Putten, M. J., and Van Der Kooij, H. (2015). Stretch evoked potentials in healthy subjects and after stroke: a potential measure for proprioceptive sensorimotor function. *IEEE Trans. Neural Syst. Rehabil. Eng.* 23, 643–654. doi: 10.1109/TNSRE.2015.2388692

Carpenter, M. B., and Sutin, J. (1983). *Human Neuroanatomy*. Baltimore, MD: Williams and Wilkins.

Chen, S., and Billings, S. A. (1989). Representations of non-linear systems: the NARMAX model. *Int. J. Control* 49, 1013–1032. doi: 10.1080/00207178908559683

Chen, S., Billings, S. A., and Luo, W. (1989). Orthogonal least squares methods and their application to non-linear system identification. *Int. J. Control* 50, 1873–1896. doi: 10.1080/00207178908953472

Crama, P., and Schoukens, J. (2001). "First estimates of Wiener and Hammerstein systems using multisine excitation," in *Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference*. (Budapest). 1365–1369.

Dempsey, E. J., and Westwick, D. T. (2004). Identification of Hammerstein models with cubic spline nonlinearities. *IEEE Trans. Biomed. Eng.* 51, 237–245. doi: 10.1109/TBME.2003.820384

Desmedt, J. E., and Cheron, G. (1980). Somatosensory evoked potentials to finger stimulation in healthy octogenarians and in young adults: wave forms, scalp topography and transit times of pariental and frontal components. *Electroencephalogr. Clin. Neurophysiol.* 50, 404–425. doi: 10.1016/0013-4694(80)90007-3

Friston, K. J. (2000). The labile brain. *I. Neuronal transients and nonlinear coupling. Philos. Trans. R Soc. Lond. B Biol. Sci.* 355, 215–236. doi: 10.1098/rstb.2000.0560

Hussein, S., Chandra, R., and Sharma, A. (2016). "Multi-step-ahead chaotic time series prediction using coevolutionary recurrent neural networks," in *IEEE Congress on Evolutionary Computation (CEC)* (Vancouver, BC), 3084–3091.

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Biochemistry, D. O., Jessell, M. B. T., Siegelbaum, S., et al. (2000). *Principles of Neural Science*. New York, NY: McGraw-hill.

Langdon, A. J., Boonstra, T. W., and Breakspear, M. (2011). Multi-frequency phase locking in human somatosensory cortex. *Prog. Biophys. Mol. Biol.* 105, 58–66. doi: 10.1016/j.pbiomolbio.2010.09.015

Li, Y., Wei, H. L., Billings, S. A., and Sarrigiannis, P. (2015). Identification of nonlinear time-varying systems using an online sliding-window and common model structure selection (CMSS) approach with applications to EEG. *Int. J. Syst. Sci.* 47, 2671–2681. doi: 10.1080/00207721.2015.1014448

Makeig, S., Bell, A.J., Jung, T.-P., and Sejnowski, T.J. (1996). "Independent component analysis of electroencephalographic data," in *Advances in Neural Information Processing Systems* (Princeton, NJ), 145–151.

Marreiros, A. C., Daunizeau, J., Kiebel, S. J., and Friston, K. J. (2008). Population dynamics: variance and the sigmoid activation function. *Neuroimage* 42, 147–157. doi: 10.1016/j.neuroimage.2008.04.239

Moran, R. J., Kiebel, S. J., Stephan, K., Reilly, R., Daunizeau, J., and Friston, K. J. (2007). A neural mass model of spectral responses in electrophysiology. *Neuroimage* 37, 706–720. doi: 10.1016/j.neuroimage.2007.05.032

Nelles, O. (2013). *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*. Berlin: Springer Science and Business Media.

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011:156869 . doi: 10.1155/2011/156869

Paduart, J., Lauwers, L., Pintelon, R., and Schoukens, J. (2012). Identification of a Wiener–Hammerstein system using the polynomial nonlinear state space approach. *Control Eng. Pract.* 20, 1133–1139. doi: 10.1016/j.conengprac.2012.06.006

Raimondo, F., Kamienkowski, J. E., Sigman, M., and Slezak, D. F. (2012). CUDAICA: GPU optimization of infomax-ICA EEG analysis. *Comput. Intell. Neurosci.* 2012:206972. doi: 10.1155/2012/206972

Siegel, A., and Sapru, H. N. (2006). *Essential Neuroscience*. Philadelphia, PA: Lippincott Williams and Wilkins.

Standring, S. (2016). *Gray's Anatomy: the Anatomical Basis of Clinical Practice, 41st Edn*. London: Elsevier Health Sciences

Vlaar, M. P., Birpoutsoukis, G., Lataire, J., Schoukens, M., Schouten, A. C., Schoukens, J., et al. (2018). Modeling the nonlinear cortical response in EEG evoked by wrist joint manipulation. *IEEE Trans. Neural Syst. Rehabil. Eng.* 26, 205–215. doi: 10.1109/TNSRE.2017.2751650

Vlaar, M. P., Solis-Escalante, T., Vardy, A. N., Van Der Helm, F. C., and Schouten, A. C. (2017). Quantifying nonlinear contributions to cortical responses evoked by continuous wrist manipulation. *IEEE Trans. Neural Syst. Rehabil. Eng.* 25, 481–491. doi: 10.1109/TNSRE.2016.2579118

Vlaar, M. P., Van Der Helm, F. C., and Schouten, A. C. (2015). Frequency domain characterization of the somatosensory steady state response in electroencephalography. *IFAC-PapersOnLine* 48, 1391–1396. doi: 10.1016/j.ifacol.2015.12.327

Westwick, D. T., and Kearney, R. E. (2001). Separable least squares identification of nonlinear Hammerstein models: application to stretch reflex dynamics. *Ann. Biomed. Eng.* 29, 707–718. doi: 10.1114/1.1385806

Yang, Y., Solis-Escalante, T., Van Der Helm, F. C., and Schouten, A. C. (2016a). A generalized coherence framework for detecting and characterizing nonlinear interactions in the nervous system. *IEEE Trans. Biomed. Eng.* 63, 2629–2637. doi: 10.1109/TBME.2016.2585097

Yang, Y., Solis-Escalante, T., Yao, J., Daffertshofer, A., Schouten, A. C., and Van Der Helm, F. C. (2016b). A general approach for quantifying nonlinear connectivity in the nervous system based on phase coupling. *Int. J. Neural Syst.* 26:1550031. doi: 10.1142/S0129065715500318

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Tian, Yang, van der Helm and Dewald. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.