



Reducing Brain Signal Noise in the Prediction of Economic Choices: A Case Study in Neuroeconomics

Raanju R. Sundararajan¹, Marco A. Palma^{2*} and Mohsen Pourahmadi¹

¹ Department of Statistics, Texas A&M University, College Station, TX, United States, ² Department of Agricultural Economics, Texas A&M University, College Station, TX, United States

OPEN ACCESS

Edited by:

Paul E. M. Phillips,
University of Washington,
United States

Reviewed by:

Dino J. Levy,
Tel Aviv University, Israel
Vijay Mohan K. Nambodiri,
University of North Carolina at Chapel
Hill, United States

*Correspondence:

Marco A. Palma
mapalma@tamu.edu

Specialty section:

This article was submitted to
Decision Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 19 July 2017

Accepted: 30 November 2017

Published: 14 December 2017

Citation:

Sundararajan RR, Palma MA and
Pourahmadi M (2017) Reducing Brain
Signal Noise in the Prediction of
Economic Choices: A Case Study in
Neuroeconomics.
Front. Neurosci. 11:704.
doi: 10.3389/fnins.2017.00704

In order to reduce the noise of brain signals, neuroeconomic experiments typically aggregate data from hundreds of trials collected from a few individuals. This contrasts with the principle of simple and controlled designs in experimental and behavioral economics. We use a frequency domain variant of the stationary subspace analysis (SSA) technique, denoted as DSSA, to filter out the noise (nonstationary sources) in EEG brain signals. The nonstationary sources in the brain signal are associated with variations in the mental state that are unrelated to the experimental task. DSSA is a powerful tool for reducing the number of trials needed from each participant in neuroeconomic experiments and also for improving the prediction performance of an economic choice task. For a single trial, when DSSA is used as a noise reduction technique, the prediction model in a food snack choice experiment has an increase in overall accuracy by around 10% and in sensitivity and specificity by around 20% and in AUC by around 30%, respectively.

Keywords: choice behavior, neuroeconomics, EEG data, multi-dimensional time series, stationary subspace analysis

JEL Codes: C32, D87

1. INTRODUCTION

The interest of economists and other social scientists to integrate neurophysiological data to study human behavior has dramatically increased. Neuroeconomics has opened the door to research aiming to explain behavioral models of decision making (Camerer et al., 2004). However, neuroeconomics has only had limited reception into mainstream economics, perhaps due to the limitation of brain processes for improving the prediction of economic behavior (see Harrison, 2008; Bernheim, 2009; Konovalov and Krajbich, 2016). After all, economists are ultimately interested in predicting behavior (Gul and Pesendorfer, 2008; Fehr and Rangel, 2011).

Behavioral and experimental economics seem to be very much interrelated with neuroeconomics. However, behavioral and experimental economics rely on simple and controlled experiments to infer causality. Brain data, by nature, are very noisy. This is due to the fact that subjects react to the presented stimuli and process it in their brain visually (i.e., colors, shapes, etc.), physically (i.e., moving their eyes and muscles), emotionally, engaging in memory and other processes that simultaneously activate different regions of the brain. The data collected from an individual on a single trial measures the activity of the brain for the stimuli, but it also captures the noise from all other activity unrelated to the task of the experiment. Brain experiments typically

implement hundreds of trials that when aggregated filter out noisy signals (Plassmann et al., 2007; Hare et al., 2009; Milosavljevic et al., 2010). There is a trade-off between the experimental economics principle of designing simple experiments to assess causality and the neuroeconomics need for a large number of trials to reduce the noise in brain signals.

An emerging literature in neuroeconomics uses brain signals to directly explain choice behavior. One of the models used to explain decision making is the Neural Random Utility Model (NRUM; Webb et al., 2013). EEG data have been used to predict purchase decisions (Ravaja et al., 2013), consumer's future choices (Telpaz et al., 2015), predict preferences (Khushaba et al., 2012, 2013) and response to advertisements (Boksem and Smidts, 2015; Venkatraman et al., 2015).

EEG signals from different electrodes measuring brain activity have, in the past, been regarded as a multi-dimensional nonstationary time series; see Ombao et al. (2005) and von Bünau et al. (2010) for examples. Kaplan et al. (2005) regard the nonstationarity as the “unavoidable noise” in the brain signal. Here the nonstationary sources in the brain signal contributes to the noise in the EEG data and removing this nonstationarity is extremely useful for prediction purposes in brain related experiments. We use the words noise and nonstationarity interchangeably because in our setup the nonstationary sources contribute to parts of the signal that are unrelated to the task related activity in the experiment. Hence eliminating nonstationarity reduces noise in the brain signal. See section 3.2 and **Figures 1, 2** for illustrations of the signal before and after noise reduction. von Bünau et al. (2009) and von Bünau et al. (2010) associate *alpha* oscillations in the data as a nonstationary source. These oscillations appear usually in the range of 8–12 Hz and are associated with blinking, fatigue or tiredness. Over the course of the experiment such changes in the EEG time series are unrelated to the experimental task and corrupt the signal.

Transforming a multi-dimensional nonstationary time series into stationarity through linear transformations is a nontrivial problem of fundamental importance in many application areas. Stationary subspace analysis (SSA) (von Bünau et al., 2009) is a recent technique that attempts to find stationary linear transformations, in lower dimensions, of multi-dimensional nonstationary time series, where nonstationarity means independent heterogenous data with mean and variance smoothly varying across time.

In this paper, we apply dependent SSA (DSSA, for short) method proposed in Sundararajan and Pourahmadi (2017) for the general class of multi-dimensional nonstationary time series to data from a neuroeconomics case study. DSSA relies on the asymptotic uncorrelatedness of the discrete Fourier transform (DFT) of a second-order stationary time series at different Fourier frequencies and unlike the well-known *cointegration* theory that is restricted to parametric models such as VAR, DSSA avoids parametric model assumptions.

We employ the DSSA technique, as a noise reduction step to separate stationary (useful signal) and nonstationary sources to reduce noise in the EEG brain signal. This is important because using this process may move neurophysiological responses to become more aligned with the design of traditional economics

experiments. In other words the nonstationary sources in the brain signal are associated with variations in the mental state that are unrelated to the experimental task at hand. Hence the DSSA technique can be useful in reducing the number of trials needed from each participant in neuroeconomic experiments. More importantly, the technique greatly improves the prediction performance of an incentivized economic food choice task. In addition, the DSSA technique performs a formal test of stationarity that ensures there is a statistically significant reduction in nonstationarity (noise) in the observed signal. The ability of DSSA and ISSA in detecting the true dimension of the stationary subspace process is illustrated for different sample sizes and dimensions and it was observed that in most cases DSSA performs better than ISSA. See section 3 of Sundararajan and Pourahmadi (2017) for more details.

The rest of the paper has the following content: Section 2 introduces the SSA model setup, describes the existing SSA technique and then the DSSA approach. Section 3 discusses an empirical application of decisions for purchasing food snacks. Subjects were presented with 10 food snack choice questions and the observed EEG signal from 9 electrodes is treated as a 9-dimensional time series. The various steps taken for noise reduction are described in section 3.2 and the “noise reduced” series is then analyzed. Finally, prediction models (section 3.2.2) of the decision made by the subjects regarding various food choices are constructed and their performance is illustrated.

2. STATIONARY SUBSPACE ANALYSIS (SSA)

We begin with the SSA model setup and the method in von Bünau et al. (2009) which deals with independent and heteroscedastic data and is denoted by ISSA. Then a SSA method for dependent data, denoted by DSSA, for finding the stationary subspace process (noise reduced signal) of the observed second-order nonstationary process (observed EEG signal) is described.

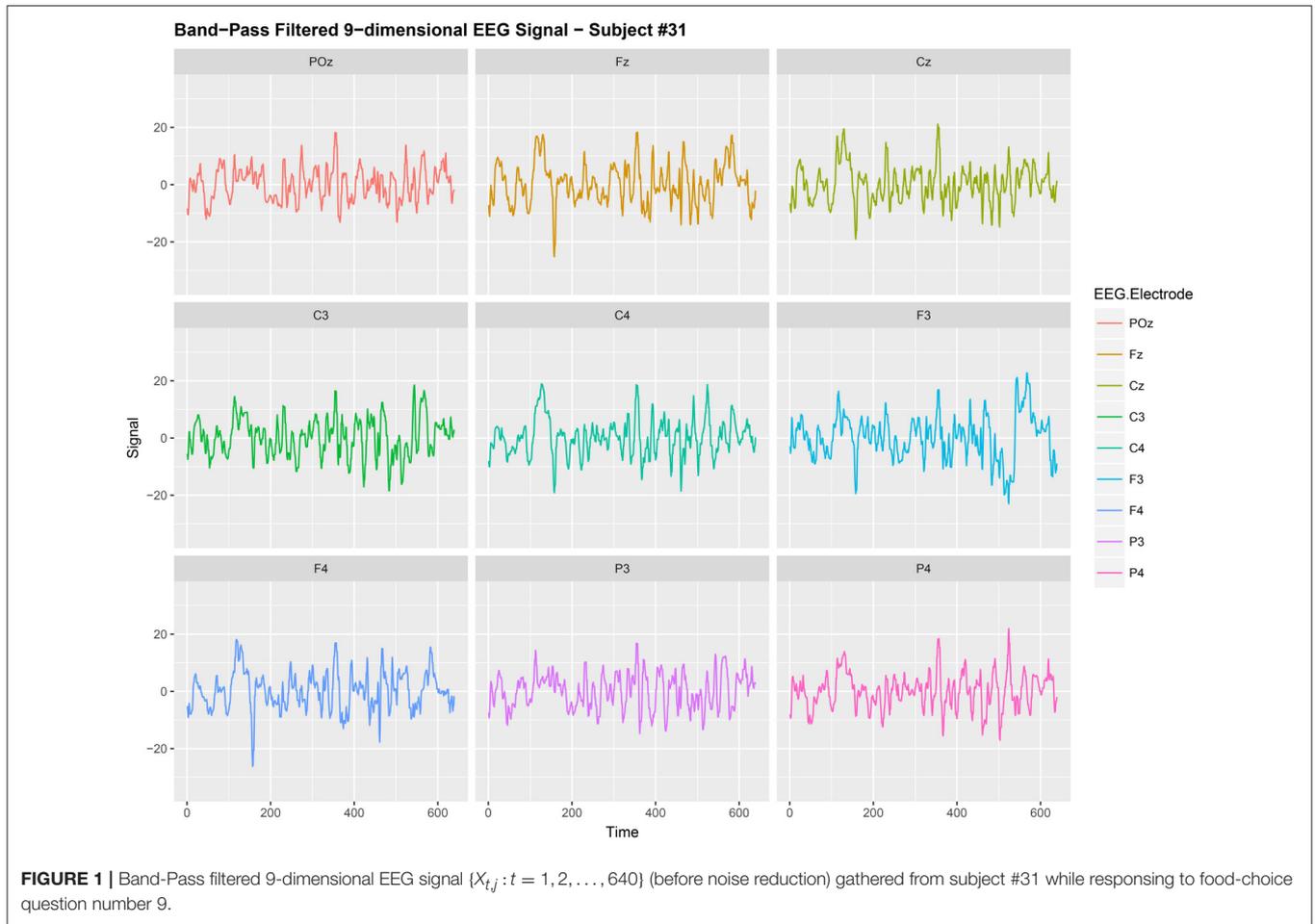
2.1. ISSA

Let $\{X_t\}$ be the observed p -dimensional nonstationary time series that is linearly generated by d stationary sources Y_t^s and $p - d$ nonstationary sources Y_t^n . More precisely,

$$X_t = AY_t = [A_s | A_n] \begin{bmatrix} Y_t^s \\ Y_t^n \end{bmatrix}, \quad (1)$$

where A is the unknown $p \times p$ (invertible) mixing matrix, A_s and A_n are $p \times d$ and $p \times (p - d)$ matrices, respectively. The dimension d is unknown and needs to be estimated. In ISSA, the notion of stationarity is with respect to the first two moments (i.e., mean and lag-0 covariance).

The goal is to estimate the demixing matrix $B = A^{-1}$ so that $Y_t = BX_t$ is separated into stationary and nonstationary sources. The nonstationary source Y_t^n contributes to the noise (for example, fatigue and tiredness of participants during the experiment) in the observed EEG signal and separating it from the stationary source Y_t^s is useful in improving performance in prediction models.



The matrix B is assumed to be an orthogonal matrix. The matrix B is estimated first by dividing the time series data into N epochs and then minimizing, as a function of B , the Kullback-Leibler (KL) divergence between Gaussian distributions across these segments. Let $\hat{\mu}_i, \hat{\Sigma}_i, i = 1, 2, \dots, N$, be the estimated mean and covariance of the data for the i^{th} segment, respectively. Let the $d \times p$ matrix B_1 be the first d rows of B . It follows that the mean and variance of $Y_i^s = B_1 X_t$ on the i^{th} segment are

$$\hat{\mu}_i^s = B_1 \hat{\mu}_i \text{ and } \hat{\Sigma}_i^s = B_1 \hat{\Sigma}_i B_1^T, \quad i = 1, 2, \dots, N. \quad (2)$$

The matrix B is then chosen so that the means and covariances vary the least across all epochs. This leads to an estimate B_1 such that $B_1 X_t$ is the target stationary source. The objective function is the sum of the Kullback-Leibler (KL) divergences between the $N(\hat{\mu}_i^s, \hat{\Sigma}_i^s), i = 1, \dots, N$, on each segment and a normal distribution with their grand averages as its parameters, namely $N(\bar{\mu}^s, \bar{\Sigma}^s)$ where $\bar{\mu}^s = 1/N \sum_{i=1}^N \hat{\mu}_i^s$ and $\bar{\Sigma}^s = 1/N \sum_{i=1}^N \hat{\Sigma}_i^s$:

$$\begin{aligned} L(B) &= \sum_{i=1}^N D_{KL} [N(\hat{\mu}_i^s, \hat{\Sigma}_i^s) || N(\bar{\mu}^s, \bar{\Sigma}^s)] \\ &= \sum_{i=1}^N \left(-\log \det \hat{\Sigma}_i^s + (\hat{\mu}_i^s - \bar{\mu}^s)^T (\hat{\mu}_i^s - \bar{\mu}^s) \right). \quad (3) \end{aligned}$$

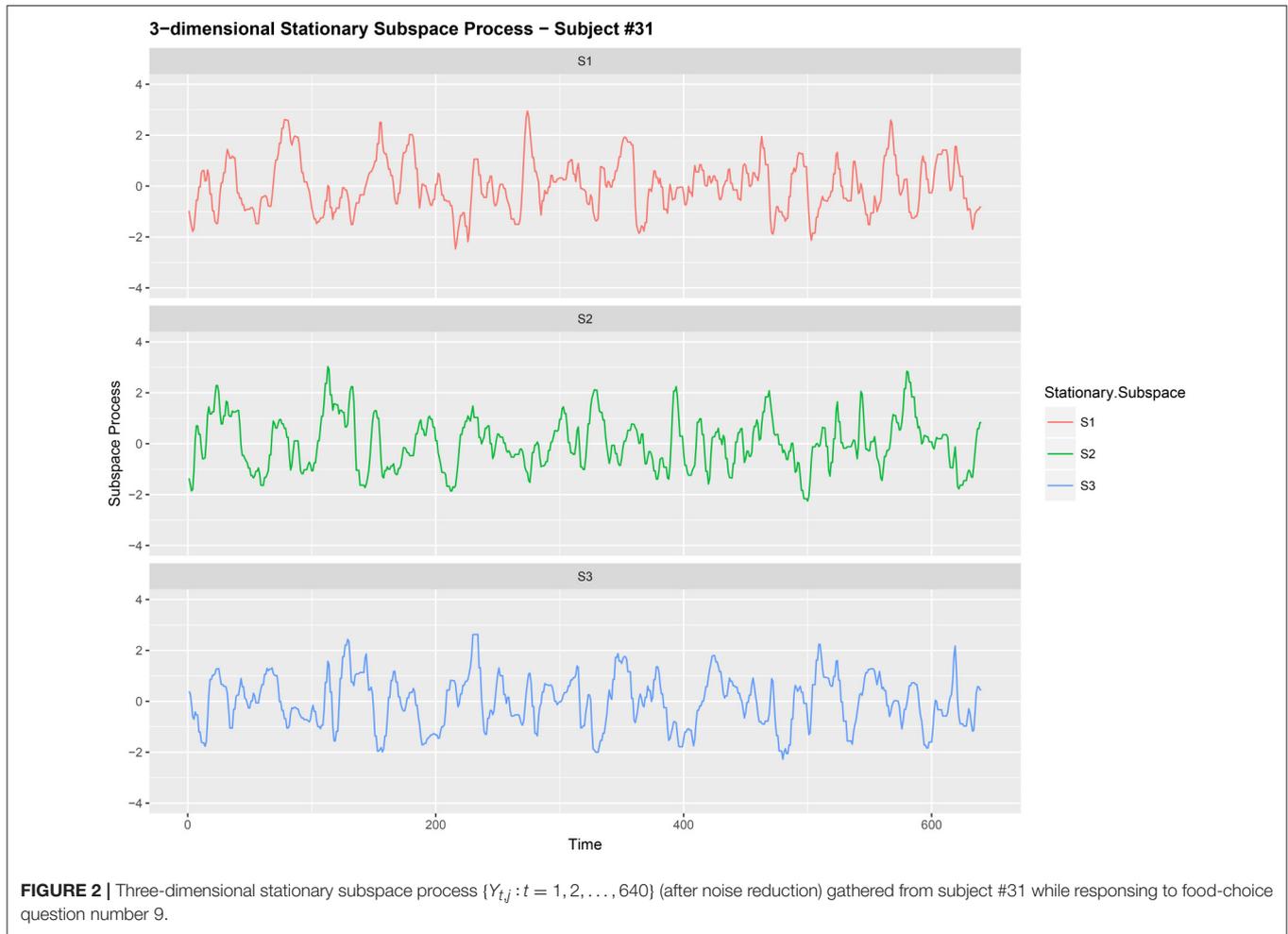
The matrix B is estimated by minimizing $L(B)$; see von Bünau et al. (2009) for more details. In the above methodology, the partitioning of the time series into N segments comes with some disadvantages. In addition to the independence assumption, ISSA works under the assumption that nonstationarity in the data is only with respect to the first two moments (mean and variance) which can be restrictive.

2.2. Dependent SSA (DSSA)

Here we describe the DSSA approach to finding a stationary subspace process of given multi-dimensional second-order nonstationary processes satisfying Equation (1) using properties from the frequency domain. This method does not require dividing the time series data into several segments and utilizes a test of stationarity for determining whether the estimated source is stationary. This would be useful in not only finding the stationary subspace process (noise reduced signal) but to also ensure there is a statistically significant reduction in the nonstationarity (noise) in the observed EEG signal.

Recall that the discrete Fourier transform (DFT) of any d -variate series $\{Z_t\}, 1 \leq t \leq T$, is given by

$$J_Z(\omega_k) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^T Z_t \exp(-it\omega_k), \quad (4)$$



where $\omega_k = \frac{2\pi}{T}k$, $k = 1, 2, \dots, T$. Treating the DFT as a time series indexed by k , its lag- r autocovariance is the $d \times d$ complex valued matrix given by

$$\hat{\Gamma}_r^Z = \frac{1}{T} \sum_{k=1}^T J_Z(\omega_k) J_Z(\omega_{k+r})^*, \quad (5)$$

where $J_Z(\cdot)^*$ denotes the complex conjugate transpose and $r = 0, 1, \dots$. It is well known (Theorem 4.3.1 of Brillinger, 2001) that for a second-order stationary time series $\{Y_t^s\}$, its DFTs are asymptotically uncorrelated when $\omega_i \neq \omega_j$, i.e.,

$$\text{cov}(J_{Y^s}(\omega_i), J_{Y^s}(\omega_j)) = O\left(\frac{1}{T}\right), \quad (6)$$

where $J_{Y^s}(\cdot)$ denotes the DFT of $\{Y_t^s\}$. Thus, for a given positive integer m , based on the magnitudes of the first few autocovariances of the DFTs of $Y_t^s = B_1 X_t$, we construct the following objective function as a measure of departure from stationarity:

$$D_Y(B) = \sum_{r=1}^m \|\hat{\Gamma}_r^{Y^s}(B)\|_F^2, \quad (7)$$

where for a matrix $A \in \mathbb{R}^{d \times d}$, $\|A\|_F = \sqrt{\sum_{i,j=1}^d |a_{ij}^2|}$ denotes its Frobenius norm. A solution \hat{B} is obtained by minimizing $D_Y(B)$ subject to the orthonormality assumption $BB^T = I_p$, see Sundararajan and Pourahmadi (2017) for more details. In section 2.2.5 of the previous work a sequential technique for estimating the unknown dimension d of the stationary subspace is described.

The previous work includes theoretical justifications for DSSA to correctly identify the dimension d of the stationary source. Also, numerous simulation examples with different types of stationary and nonstationary sources are simulated and the ability of DSSA and ISSA to identify a stationary subspace process $\{Y_t^s\}$ and its dimension d has been discussed. In these examples, for the nonstationary sources $\{Y_t^n\}$ in Equation (1), independent Gaussian components with changing variances and dependent Gaussian components with changing variances are simulated. The dimensions were allowed to vary from 1 to 5. Other simulation examples include stationary vector autoregressive processes (VAR) with no nonstationary sources ($d = p$), time-varying vector moving average processes (VMA) with no stationary sources ($d = 0$) and nonstationary unit-root VAR processes with the number of nonstationary sources varying from 1 to $p - 1$.

3. EXPERIMENT OF ECONOMIC CHOICES: A CASE STUDY

A total of 181 right-handed students participated in a food snack choice decision experiment conducted in the Texas A&M Human Behavior Laboratory. The sample consisted of about 50% females and 50% males. The subjects were presented with 10 food choice task questions (10 trials). Each choice consisted of two food products, product A and product B. The two products within each alternative had the same features relative to brand, price, packaging and flavor. The only difference between each pair of products was that one of them had fewer calories, making it a healthier choice (original strawberry Jello –70 calories- vs. sugar-free strawberry Jello –10 calories)¹. Subjects were asked to fast for 3 h prior to the experiment, and received a compensation fee of \$20 in exchange for their participation. In order to incentivize and make the food choice task real, one of the 10 tasks was randomly selected to be binding and participants had to eat the food snack before being paid and leaving the laboratory. The displayed picture of each item was the actual photo available for purchase in Walmart's website; however, the participants were not aware that the products were purchased in Walmart.

The experimental design proceeds as follow. At the beginning of the experiment, a blank slide with a fixation point in the middle of the computer screen was presented for 2 s. Then, for each food choice task, the actual product images were presented in the following screen for 8 s. A separate decision slide asked participants which of the two food snacks they prefer to eat. After each decision, an inter-stimulus slide was presented for 0.75 s. The order of the products was randomized across trials in the experimental design; however, all subjects completed the task in the same order.

3.1. Data Acquisition

The participant was fitted with a proper size EEG headset (B-Alert X10, Advanced Brain Monitoring, Inc.) with 9 electrodes to record brain activity from the pre-frontal (F3, F4, FZ), central (C3, C4, CZ), and parietal (P3, P4, POZ) cortices and a linked mastoid reference. An electrode impedance test was performed to ensure proper conductivity of the electrodes. The impedance level threshold was 20 k Ω . An EEG calibration procedure was implemented before the data collection. The EEG calibration incorporated choice tasks (unrelated to the study), psychomotor, and auditory psychomotor vigilance tasks. The EEG data was collected at a sampling rate of 256 Hz. The experiment was presented using the iMotions software platform.

3.2. Data Analysis

For any given food choice task, say product A vs. product B, we gathered the 9-dimensional EEG signal from the 9 electrodes from the start of the stimuli when the product images are shown to 2.5 seconds after the start. On the digital signal scale, this constitutes 640 observations (2.5×256). More precisely, for each

subject $j = 1, 2, \dots, 181$, the data comprises of 640 observations across time.

Given the raw 9-dimensional EEG time series obtained in this case study, we proceed according to the following algorithm to obtain the prediction results:

The Prediction Algorithm.

-
- Step 1: Filter the raw 9-dimensional signal using a 0.5 Hz high-pass and 45 Hz low-pass filter. Denote the filtered series as $\{X_{t,j}\}$ where $j = 1, 2, \dots, 181$ and $t = 1, 2, \dots, 640$.
 - Step 2: Pre-whiten $\{X_{t,j}\}$. For convenience in notation, we denote $X_{t,j}$ as the band-pass filtered signal that has been pre-whitened.
 - Step 3: Noise reduction: Apply SSA to $\{X_{t,j}\}$ to obtain $\{Y_{t,j}\}$ (Section 3.2.1).
 - Step 4: Feature Selection and Prediction Models (Section 3.2.2).
 - Step 5: Assessing Prediction Performance (Section 3.2.3).
-

In Step 2 we pre-whiten $\{X_{t,j}\}$ before further analysis by computing the 9×9 sample covariance matrix S_j and then transform the data to $(S_j)^{-0.5}X_{t,j}$. This standardization reduces the cross-sectional correlation in $\{X_{t,j}\}$.

3.2.1. Noise Reduction via SSA

It is common to treat data like $\{X_{t,j}\}$ as a nonstationary time series (Ombao et al., 2005; Park et al., 2014). The words noise and nonstationarity are used interchangeably because in our setup the nonstationary sources contribute to parts of the signal that are unrelated to the food choice task. Hence eliminating nonstationarity reduces noise in the brain signal. As an illustration, we make a plot of the 9-dimensional EEG signal $X_{t,j}$ (before noise reduction) in **Figure 1**. In **Figure 2**, we then plot a 3-dimensional stationary subspace process obtained after application of DSSA. The presence of nonstationarity (noise) in $X_{t,j}$ was confirmed by carrying out formal tests of stationarity (Jentsch and Subba Rao, 2015). Hence we resort to the SSA technique for removing this nonstationarity from the signal and this is described in this section.

As a pre-processing technique to reduce noise, we apply DSSA and ISSA described in section 2 to obtain a d dimensional stationary subspace process where $d < 9$, denoted by $\{Y_{t,j}\}$. Since the actual dimension d is unknown, we present the results for $d = 4, 5, 6, 7, 8$. We also applied the sequential technique in Sundararajan and Pourahmadi (2017) to detect d for each subject and each food choice task. Here we obtained a mode of $d = 8$ as an estimate of the dimension of the stationary subspace.

3.2.2. The Prediction Models

We discuss three prediction models based on logistic regression with different derived features. The aim of the prediction models discussed below is to fit a model to predict product choice (A or B) based on the input signal. While building prediction models M_1 and M_2 , only Step 2 of the algorithm is used, for prediction model M_3 both Steps 2 and 3 are needed. Note that model M_2 assumes that $\{X_{t,j}\}$ is stationary whereas model

¹The product list and amount of calories is listed in section A.1 in Appendix. The focus of this paper is reducing EEG noise to improve the prediction of which of the two food snacks participants would choose, irrespective of the product's identity.

M_3 assumes that $\{X_{t,j}\}$ is nonstationary and applies SSA before extracting features and estimating the prediction model.

3.2.2.1. Model M_1

A standard model similar to Telpaz et al. (2015) is based on the importance of the pre-frontal EEG channels in explaining choice behavior in individuals. Following their aggregation technique to reduce the noise when computing preference scores for products, we take average of the signals from the 3 pre-frontal channels (F3, F4, FZ) over the 2.5 s. The signal here is a 3-dimensional band-pass filtered signal that was pre-whitened. The average is taken per subject per food choice question (say product A vs. product B). For subject $j, j = 1, 2, \dots, 181$, this average denoted by the scalar \bar{X}_j is used as a feature in the following logistic regression model:

$$P(c_{j,AB} = 1 \mid \bar{X}_j) = \frac{\exp(a_0 \bar{X}_j)}{1 + \exp(a_0 \bar{X}_j)}, \tag{8}$$

for $j = 1, 2, \dots, 181$. In the model above we have denoted 1 for product A and 0 for product B and the model predicts the class (0 or 1) based on the derived feature \bar{X}_j .

3.2.2.2. Model M_2

In this approach, to distinguish between the two classes denoted as 1 for product A and 0 for product B, we take $\{X_{t,j}\}$ the pre-whitened 9-dimensional band-pass filtered signal. We then focus on the covariance structure of $X_{t,j}$ for each of the two classes (0 and 1). The aim is to derive features that bring out the differences between the two classes based on the covariance structure of the signal. This is achieved by computing the average spectral density matrices for the two classes over the Fourier frequencies:

$$\bar{g}^i(\omega_k) = \frac{1}{n_i} \sum_{j \in \text{Class } i} g_j(\omega_k), \quad i = 0, 1, \tag{9}$$

where $g_j(\omega_k)$ is the estimated 9×9 spectral matrix for subject j using observations $\{X_{t,j}\}$, n_i for $i = 0, 1$ is the number of subjects in the two classes and $\omega_k = \frac{2\pi k}{640}$, $k = 1, 2, \dots, 640$, are the fundamental Fourier frequencies. The spectral matrix was estimated using a Daniell kernel with smoothing window length 25 (approximately $\sqrt{640}$); see Example 10.4.1 in Brockwell and Davis (1991).

In order to train the classifier, for every subject $j \in \{1, 2, \dots, 181\}$, a distance vector $p_{j,AB} = (p_{0,j,AB}, p_{1,j,AB})$ is computed where

$$p_{i,j,AB} = \frac{1}{640} \sum_{k=1}^{640} \|g_j(\omega_k) - \bar{g}^i(\omega_k)\|_F^2 \quad i = 0, 1,$$

and $\|\cdot\|_F$ is the Frobenius norm of a matrix. It measures the distance to the center of each of the two classes and serves as our two-dimensional feature vector used in constructing the following logistic regression model (prediction model):

$$P(c_{j,AB} = 1 \mid p_{j,AB}) = \frac{\exp(\alpha_0 p_{0,j,AB} + \alpha_1 p_{1,j,AB})}{1 + \exp(\alpha_0 p_{0,j,AB} + \alpha_1 p_{1,j,AB})}, \tag{10}$$

for $j = 1, 2, \dots, 181$ and $c_{j,AB}$ is the class indicator (1 for product A or 0 for product B) for subject j .

3.2.2.3. Model M_3

Here we apply Step 2 on the raw 9-dimensional EEG signal to obtain $\{X_{t,j}\}$. Then we obtain on the d -variate stationary subspace processes, $\{Y_{t,j}\}$, using DSSA/ISSA (Step 3). Similar to the approach in model M_2 , we aim to capture the differences between the two classes based on the covariance structure of the signal. Unlike model M_2 , we apply DSSA and ISSA described in section 2 to obtain a d -dimensional stationary subspace process where $d < 9$, denoted by $\{Y_{t,j}\}$. Features to be fed into the prediction model will be based on $\{Y_{t,j}\}$ as opposed to model M_2 wherein $\{X_{t,j}\}$ was used. Then, proceeding as in model M_2 , we compute the average spectral density matrices for the two classes over the Fourier frequencies:

$$\bar{f}^i(\omega_k) = \frac{1}{n_i} \sum_{j \in \text{Class } i} f_j(\omega_k), \quad i = 0, 1, \tag{11}$$

where $f_j(\omega_k)$ is the estimated $d \times d$ spectral matrix for subject j using observations $\{Y_{t,j}\}$, n_i for $i = 0, 1$ is the number of subjects in the two classes and $\omega_k = \frac{2\pi k}{640}$, $k = 1, 2, \dots, 640$ are the fundamental Fourier frequencies. The spectral matrix was estimated using a Daniell kernel with smoothing window length 25 (approximately $\sqrt{640}$).

In order to train the classifier, for every subject $j \in \{1, 2, \dots, 181\}$, a distance vector $d_{j,AB} = (d_{0,j,AB}, d_{1,j,AB})$ is computed where

$$d_{i,j,AB} = \frac{1}{640} \sum_{k=1}^{640} \|f_j(\omega_k) - \bar{f}^i(\omega_k)\|_F^2 \quad i = 0, 1,$$

and $\|\cdot\|_F$ is the Frobenius norm of a matrix. It measures the distance to the center of each of the two classes and serves as our two-dimensional feature vector used in constructing the following logistic regression model (prediction model):

$$P(c_{j,AB} = 1 \mid d_{j,AB}) = \frac{\exp(\beta_0 d_{0,j,AB} + \beta_1 d_{1,j,AB})}{1 + \exp(\beta_0 d_{0,j,AB} + \beta_1 d_{1,j,AB})}, \tag{12}$$

for $j = 1, 2, \dots, 181$ and $c_{j,AB}$ is the class indicator (1 for product A or 0 for product B) for subject j .

3.2.3. Prediction Performance

We assess the performance by computing the overall prediction accuracy and the average sensitivity and specificity. Using the

TABLE 1 | Confusion matrix.

		Prediction		
		Product A	Product B	Total
Actual	Product A	C_A	I_B	T_A
	Product B	I_A	C_B	T_B

confusion matrix given in **Table 1**, we compute two prediction accuracy measures given by

$$A_1 = \frac{C_A + C_B}{T_A + T_B}, \quad A_2 = \frac{\frac{C_A}{T_A} + \frac{C_B}{T_B}}{2}, \quad (13)$$

where A_1 is the overall prediction accuracy of the model and A_2 , in a binary classification context, is the average of sensitivity (true positive rate) and specificity (true negative rate) of the prediction models. Finally, we present an estimate of the AUC: area under the ROC curve (LeDell et al., 2015) for the 10 food choice questions for each of the 3 models and this measure is denoted as A_3 . The ROC curve plots the true positive rate against the false positive rate and is a useful measure of model performance. The area under the ROC curve (known as AUC) varies between 0 and 100% with a value of 50% as baseline (uninformative classifier).

In **Table 2**, we shuffle the class labels randomly and fit the prediction models and assess the performance measures. The shuffling of labels is done 500 times, each time fitting the prediction models, and the average performance measure over the 500 runs across the 10 food choice questions is presented. This enables us to identify a baseline for the 3 performance

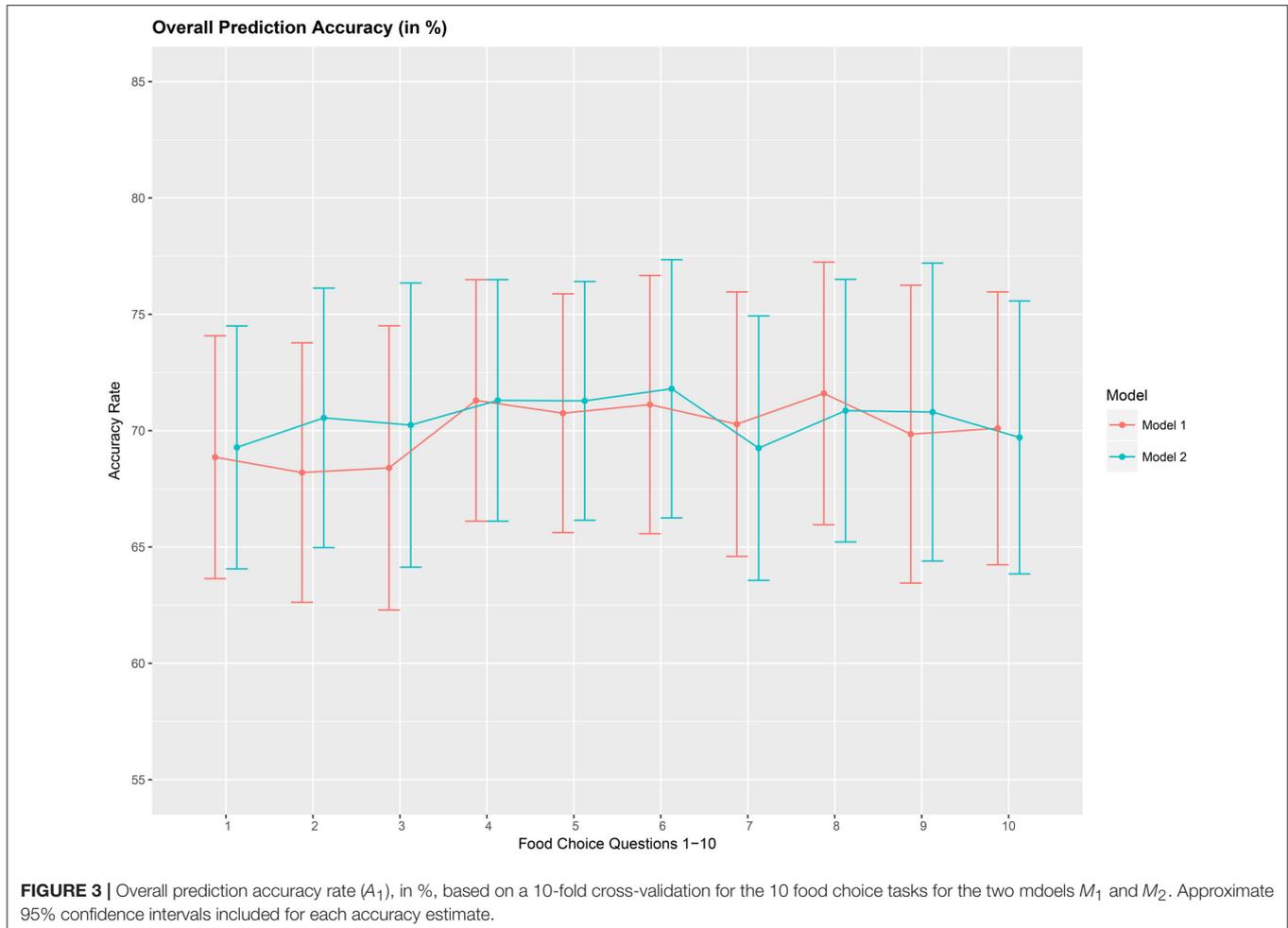
measures (70% for performance measure A_1 and 50% for performance measures A_2 and A_3).

These accuracy rates are computed using a 10-fold cross-validation technique where the data is randomly divided into 10 nearly equal parts. Each part is removed, in turn, while the remaining data are used to fit the prediction models M_1, M_2, M_3 and the predictions are carried out for the left out part. More precisely, the computed accuracy rates are the out-of sample

TABLE 2 | Prediction performance of the 3 models with shuffled labels: the average of the 3 performance measures A_1 , A_2 , and A_3 (AUC) taken across the 10 food choice questions for the three competing models M_1 , M_2 , and M_3 .

Model	Overall accuracy - A_1	Avg. of sensitivity and specificity - A_2	AUC - A_3
M_1	69.54	48.97	53.26
M_2	69.05	51.81	53.25
M_3 - DSSA	69.12	51.34	54.20
M_3 - ISSA	69.51	51	52.08

For model M_3 the choice of d is taken as 8.



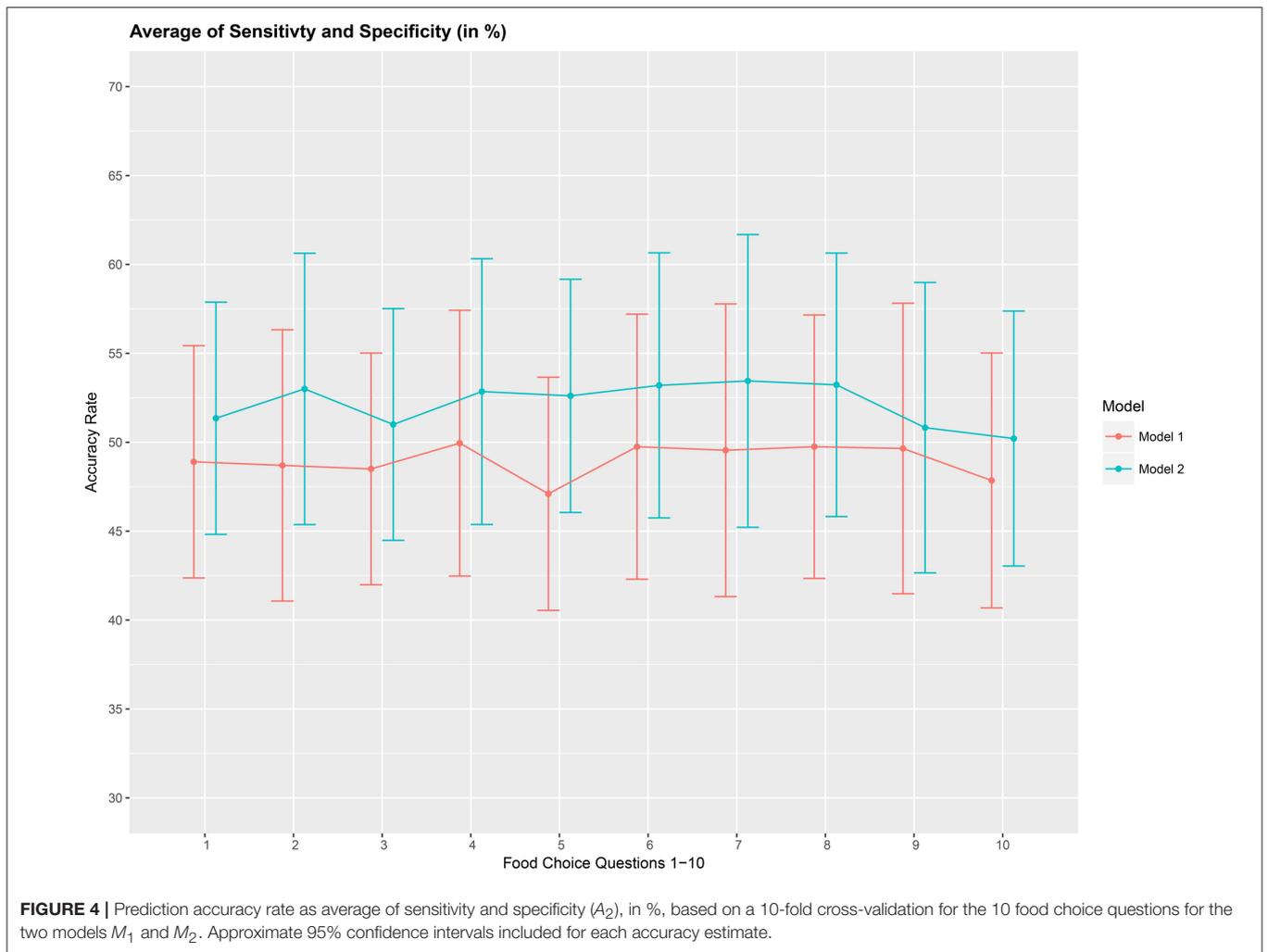


TABLE 3 | Ten-fold cross-validation overall prediction accuracy (in %) for the 10 questions Q1–Q10 corresponding to $d = 4, 5, 6, 7, 8$ for DSSA and ISSA (model M_3).

d		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
4	DSSA	72.80	73.65	74.22	71.53	74.00	73.83	73.48	71.38	74.50	72.97
	ISSA	69.60	71.20	75.10	71.40	66.50	76.20	70.18	68.45	72.37	73.13
5	DSSA	73.50	73.21	75.76	73.00	68.30	77.85	73.55	72.40	74.60	78.75
	ISSA	74.00	75.48	71.80	71.80	75.10	75.10	75.20	74.62	72.40	70.70
6	DSSA	74.65	75.00	78.46	72.85	77.35	75.10	75.70	74.52	77.45	78.61
	ISSA	71.94	77.30	74.00	79.62	78.50	73.50	75.70	76.80	74.00	74.25
7	DSSA	76.12	80.27	78.56	79.15	79.20	76.15	79.65	80.11	79.12	80.50
	ISSA	75.70	74.00	76.20	81.80	80.12	79.80	78.50	80.80	77.90	76.80
8	DSSA	78.98	81.35	81.52	84.12	80.24	79.11	82.30	80.58	80.13	81.94
	ISSA	76.90	75.10	80.16	79.60	75.10	79.00	79.00	84.75	82.00	75.68

Significant results (instances of at least a 1% improvement in DSSA) are highlighted in bold.

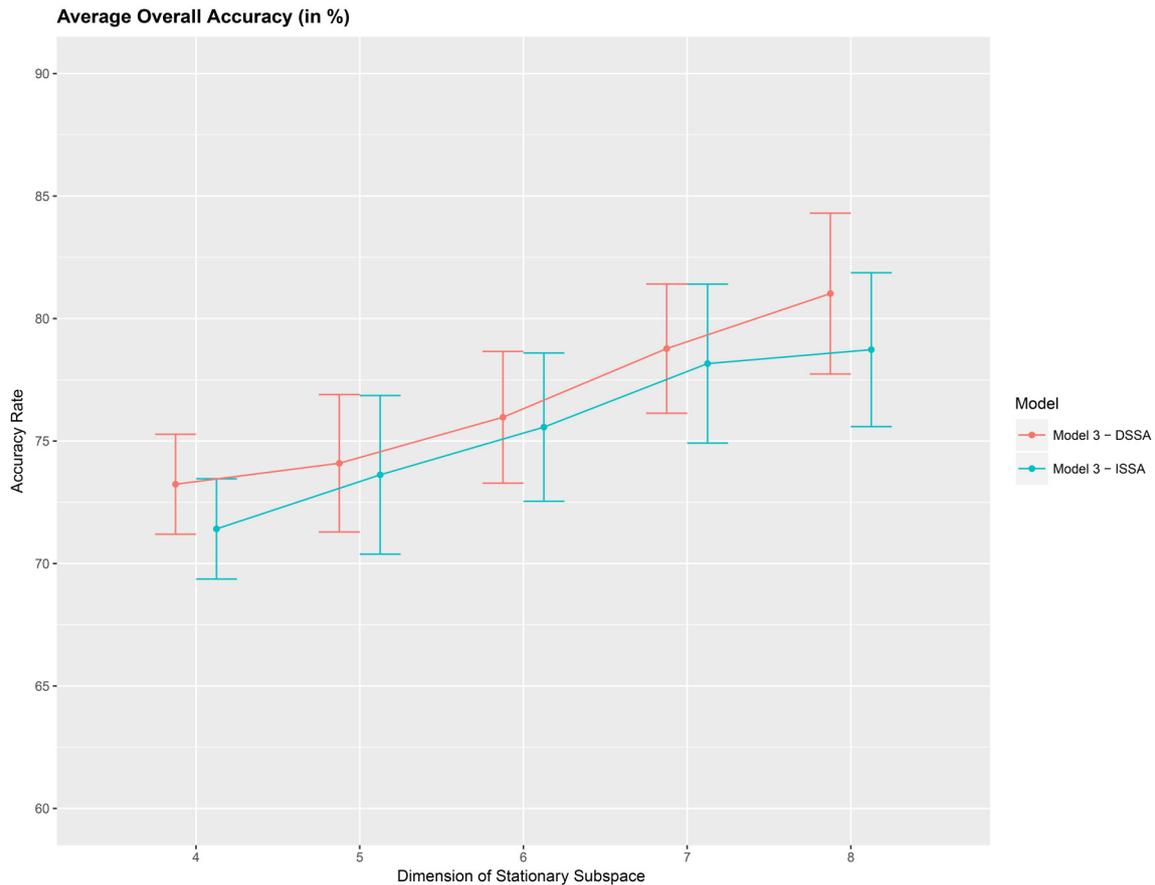


FIGURE 5 | Average 10-fold cross-validation overall accuracy rate (in %) for the 10 food choice questions (y-axis) vs. dimension of the stationary subspace (x-axis). Approximate 95% confidence intervals included for each accuracy estimate.

estimates wherein for any given pair of products A and B, the prediction model is fit based on roughly 90% of the subjects and the predictions are carried out for the remaining subjects.

The overall accuracy rate (A_1) for models M_1 and M_2 computed and plotted in **Figure 3** shows that it varies between 69 and 72% for both models. Next, we look at the performance measure A_2 as an average of the sensitivity and specificity of models M_1 and M_2 . We notice from **Figure 4** that both methods perform poorly with accuracy rates around 50%. Note that as opposed to averaging over the signal across the 3 channels in model (Equation 8), we also assessed the performance of the logistic regression models fitted individually with each of the pre-frontal channels. We obtained rates (not presented here) similar to that seen in **Figures 3, 4** in terms of overall prediction accuracy and average of sensitivity and specificity.

Next, we study the overall prediction accuracy (A_1) after applying the pre-processing techniques DSSA and ISSA and removing the nonstationarity (noise) in the EEG signal, and fitting the prediction model (Equation 12 of model M_3). Since the actual dimension of the stationary subspace is unknown, in **Table 3** we present the results for dimensions $d = 4, 5, 6, 7, 8$,

which show that DSSA performs better than ISSA in most cases. The average overall accuracy rate based on the 10 food choice tasks for each value of d is given in **Figure 5**. It is seen that the 10-fold cross-validation accuracy rate is around 80% for each of the 10 tasks when the dimension $d = 8$. This rate is roughly 10% more than the accuracy rate from **Figure 3** wherein no SSA-type pre-processing technique is applied. We also notice that as the dimension of the stationary subspace d increases, the accuracy rate also increases. This phenomenon was also observed in von Büнау et al. (2010) and confirms the improvements in prediction accuracy when there are fewer nonstationary sources (noise) in model (Equation 1). The DSSA/ISSA turns out to be a very useful tool for reducing the noise (nonstationarity) in the EEG signal.

We then assess the performance measure A_2 which is an average of the sensitivity and specificity for model M_3 . We set the dimension of the stationary subspace at $d = 8$. **Figure 6** shows that DSSA performs slightly better than ISSA in most cases. More importantly, we note that in comparison to **Figure 4**, DSSA has roughly a 20% increase in the performance measure A_2 .

Finally, we present a cross-validation estimate of the AUC for the 3 competing models in **Figure 7**. We again notice

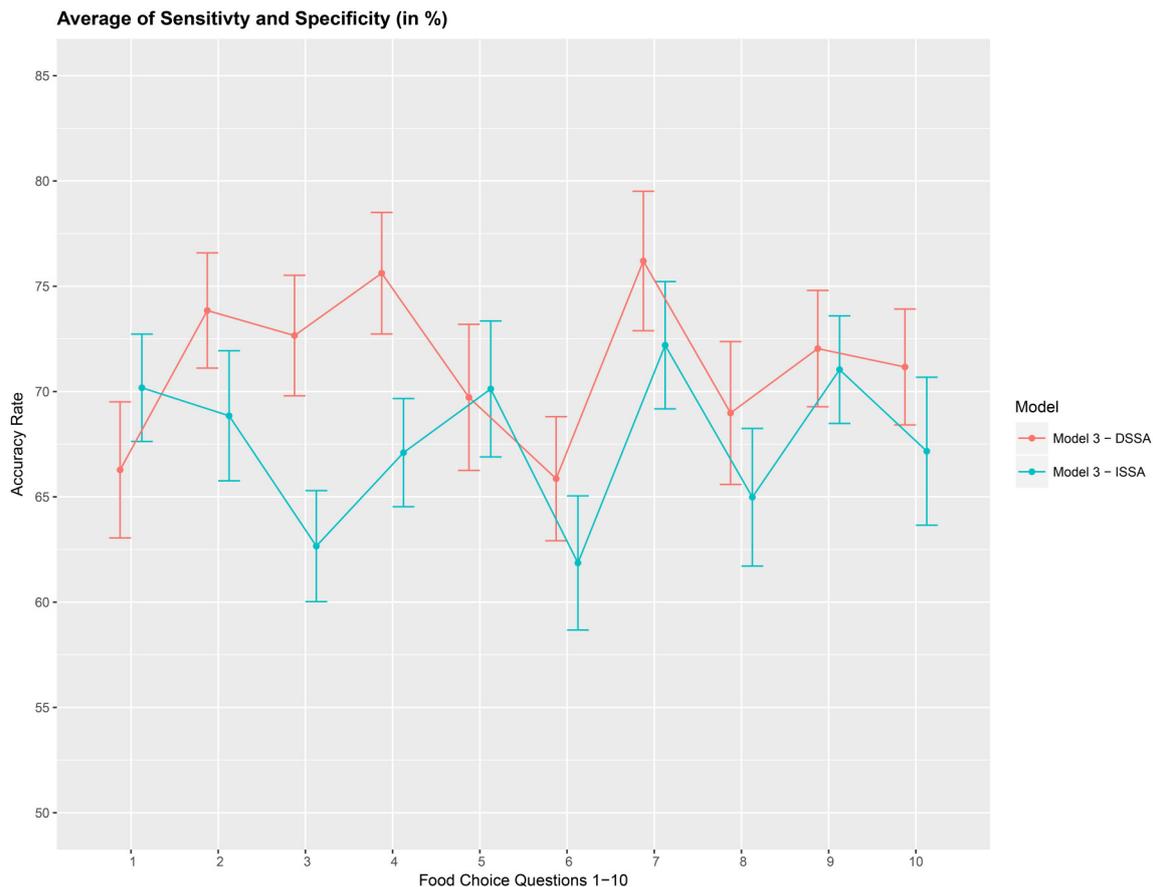


FIGURE 6 | Prediction accuracy rate: average of sensitivity and specificity (A_2), in %, based on a 10-fold cross-validation for the 10 food choice questions. Model M_3 was used with $d = 8$. Approximate 95% confidence intervals included for each accuracy estimate.

roughly a 20% increase when using DSSA/ISSA (Model M_3) as a noise reduction technique before constructing the prediction model.

The average of the 3 performance measures A_1 , A_2 , and A_3 (AUC) taken across the 10 food choice questions for the three competing models M_1 , M_2 , and M_3 is reported in **Table 4**. We note that for models M_1 and M_2 the overall prediction accuracy (A_1) is roughly 70% which is treated as a baseline for this measure. However, the performance measures A_2 and A_3 (AUC) are only around 50% which suggests a poor performance. In contrast for model M_3 , overall accuracy rate increased by roughly 10%, the measure A_2 is higher by around 20% and measure A_3 (AUC) is significantly higher (increase of roughly 30%) than models M_1 and M_2 .

4. CONCLUDING REMARKS

EEG records the electrical activity of the brain directly in the scalp. EEG signals have high temporal resolution, thus providing rich time series data of brain activity. We concentrate on EEG because it is a less expensive method to obtain brain data, making it more accessible. Brain data, however, is inherently

noisy, because it captures the brain activity for the stimuli, along with other activity unrelated to the task of the experiment. In order to filter out the noise, neuroeconomic experiments typically aggregate data from hundreds of trials from each participant. In addition to potential fatigue effects, we point out a tradeoff between the basic experimental economics principle of simplicity with the neuroeconomic need for a large number of trials to reduce brain signal noise. We apply a new statistical technique to a food choice task and show its potential for reducing the noise and hence the number of trials needed for EEG experiments. Based on the results presented in section 3.2.3, we notice that the overall accuracy rate is around 80% for each of the 10 trials when noise reduction is carried out through SSA (model M_3). More importantly, the overall prediction accuracy from a single trial increased by around 10%, the average of sensitivity and specificity increased by around 20% and the AUC increased by roughly 30% when the DSSA/ISSA technique was used to reduce signal noise.

The improvement in the prediction results in this case study by implementing noise reduction via DSSA/ISSA indicates the dynamic behavior of the brain processes. This time-varying behavior leads to nonstationarity in the observed multi-dimensional EEG signal. This phenomenon has been observed

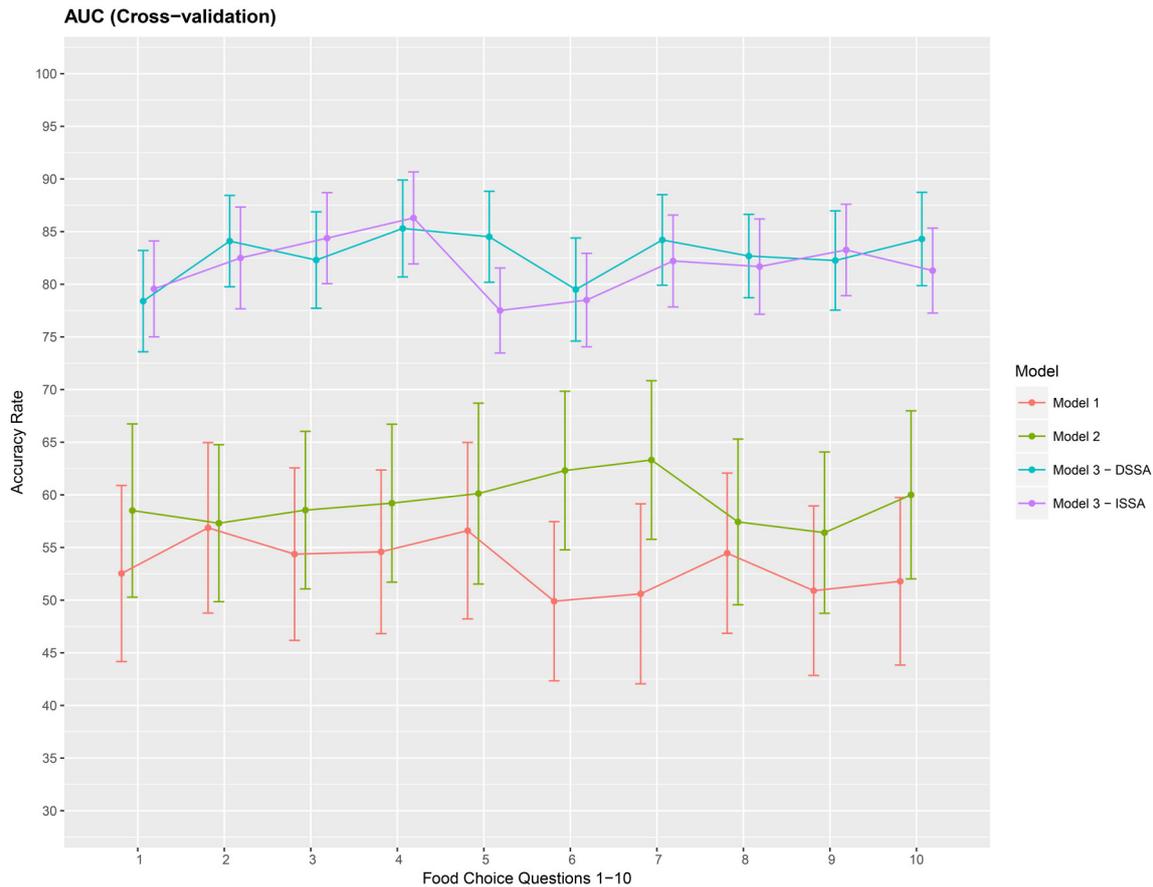


FIGURE 7 | Cross-validation estimate of the AUC in % (Area under the ROC curve) for the 3 models M_1 , M_2 , and M_3 . Approximate 95% confidence intervals included for each accuracy estimate. For model M_3 we take $d = 8$.

TABLE 4 | The average of the 3 performance measures A_1 , A_2 , and A_3 (AUC) taken across the 10 food choice questions for the three competing models M_1 , M_2 , and M_3 .

Model	Overall accuracy - A_1	Avg. of sensitivity and specificity - A_2	AUC - A_3
M_1	70.04	48.97	53.26
M_2	70.52	52.17	59.31
M_3 - DSSA	81.02	71.23	82.75
M_3 - ISSA	78.73	67.61	81.72

For model M_3 the choice of d is taken as 8.

and studied in other works (Ombao et al., 2005; Demiralp et al., 2007; von Bünau et al., 2009; Sundararajan and Pourahmadi, 2017, to name a few) wherein the observed EEG signal is treated as a multi-dimensional nonstationary time series. Hence removing nonstationarity from the EEG signal is seen to improve prediction performance in the above cited works and also in the current case study. The improvement in the prediction performance of EEG brain data shown in this case study is encouraging. Future work can build upon our procedure

to develop other post-processing techniques to further refine neurophysiological predictors of choice behavior.

ETHICS STATEMENT

The Study was approved by the IRB protocol number IRB2016-0059D.

AUTHOR CONTRIBUTIONS

The research problem and data came from the second author (Human Behavior Laboratory at Texas A&M). The methodology and computations were performed by the first and third authors. Contributions toward presentation of the results and connecting the research problem to the methodology came from all three authors. All authors have made direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

This study was carried out in accordance with the recommendations of the Texas A&M Institutional Review

Board with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Texas A&M IRB board protocol IRB2016-0059D. The first author was

supported by the National Science Foundation DMS-1513647. The work of the third author was supported by the National Science Foundation DMS-1612984. We thank Michelle Segovia for assistance with data collection.

REFERENCES

- Bernheim, B. D. (2009). On the potential of neuroeconomics: a critical (but hopeful) appraisal. *Am. Econ. J. Microecon.* 1, 1–41. doi: 10.1257/mic.1.2.1
- Boksem, M. A., and Smidts, A. (2015). Brain responses to movie trailers predict individual preferences for movies and their population-wide commercial success. *J. Market. Res.* 52, 482–492. doi: 10.1509/jmr.13.0572
- Brillinger, D. (2001). *Time Series: Data Analysis and Theory*. Society for Industrial and Applied Mathematics. doi: 10.1137/1.9780898719246
- Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*. Berlin: Springer-Verlag.
- Camerer, C. F., Loewenstein, G., and Prelec, D. (2004). Neuroeconomics: why economics needs brains. *Scand. J. Econ.* 106, 555–579. doi: 10.1111/j.0347-0520.2004.00377.x
- Demiralp, T., Bayraktaroglu, Z., Lenz, D., Junge, S., Busch, N. A., Maess, B., et al. (2007). Gamma amplitudes are coupled to theta phase in human EEG during visual perception. *Int. J. Psychophysiol.* 64, 24–30. doi: 10.1016/j.ijpsycho.2006.07.005
- Fehr, E., and Rangel, A. (2011). Neuroeconomic foundations of economic choice—recent advances. *J. Econ. Perspect.* 25, 3–30. doi: 10.1257/jep.25.4.3
- Gul, F., and Pesendorfer, W. (2008). “The case for mindless economics,” in *The Foundations of Positive and Normative Economics: A Handbook*, Vol. 1, eds A. Caplin and A. Schotter (Oxford University Press), 3–42.
- Hare, T. A., Camerer, C. F., and Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* 324, 646–648. doi: 10.1126/science.1168450
- Harrison, G. W. (2008). Neuroeconomics: a critical reconsideration. *Econ. Philos.* 24, 303–344. doi: 10.1017/S0266267108002009
- Jentsch, C., and Subba Rao, S. (2015). A test for second order stationarity of a multivariate time series. *J. Econ.* 185, 124–161. doi: 10.1016/j.jeconom.2014.09.010
- Kaplan, A. Y., Fingelkurts, A. A., Fingelkurts, A. A., Borisov, S. V., and Darkhovskiy, B. S. (2005). Nonstationary nature of the brain activity as revealed by EEG/MEG: methodological, practical and conceptual challenges. *Signal Process.* 85, 2190–2212. doi: 10.1016/j.sigpro.2005.07.010
- Khushaba, R. N., Greenacre, L., Kodagoda, S., Louviere, J., Burke, S., and Dissanayake, G. (2012). Choice modeling and the brain: a study on the electroencephalogram (eeg) of preferences. *Exp. Syst. Appl.* 39, 12378–12388. doi: 10.1016/j.eswa.2012.04.084
- Khushaba, R. N., Wise, C., Kodagoda, S., Louviere, J., Kahn, B. E., and Townsend, C. (2013). Consumer neuroscience: assessing the brain response to marketing stimuli using electroencephalogram (EEG) and eye tracking. *Exp. Syst. Appl.* 40, 3803–3812. doi: 10.1016/j.eswa.2012.12.095
- Kononov, A., and Krajbich, I. (2016). Over a decade of neuroeconomics: what have we learned? *Organ. Res. Methods*. doi: 10.1177/1094428116644502
- LeDell, E., Petersen, M., and van der Laan, M. (2015). Computationally efficient confidence intervals for cross-validated area under the ROC curve estimates. *Electron. J. Statist.* 9, 1583–1607. doi: 10.1214/15-EJS1035
- Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., and Rangel, A. (2010). The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgm. Decis. Making* 5, 437–449. doi: 10.2139/ssrn.1901533
- Ombao, H., von Sachs, R., and Guo, W. (2005). Slex analysis of multivariate nonstationary time series. *J. Am. Stat. Assoc.* 100, 519–531. doi: 10.1198/01621450400001448
- Park, T., Eckley, I. A., and Ombao, H. C. (2014). Estimating time-evolving partial coherence between signals via multivariate locally stationary wavelet processes. *IEEE Trans. Signal Process.* 62, 5240–5250. doi: 10.1109/TSP.2014.2343937
- Plassmann, H., O’Doherty, J., and Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27, 9984–9988. doi: 10.1523/JNEUROSCI.2131-07.2007
- Ravaja, N., Somervuori, O., and Salminen, M. (2013). Predicting purchase decision: the role of hemispheric asymmetry over the frontal cortex. *J. Neurosci. Psychol. Econ.* 6:1. doi: 10.1037/a0029949
- Sundararajan, R. R., and Pourahmadi, M. (2017). Stationary subspace analysis of nonstationary processes. *J. Time Series Anal.* doi: 10.1111/jtsa.12274. [Epub ahead of print].
- Telpaz, A., Webb, R., and Levy, D. J. (2015). Using EEG to predict consumers’ future choices. *J. Market. Res.* 52, 511–529. doi: 10.1509/jmr.13.0564
- Venkatraman, V., Dimoka, A., Pavlou, P. A., Vo, K., Hampton, W., Bollinger, B., et al. (2015). Predicting advertising success beyond traditional measures: new insights from neurophysiological methods and market response modeling. *J. Market. Res.* 52, 436–452. doi: 10.1509/jmr.13.0593
- von Bünau, P., Meinecke, F. C., Scholler, S., and Müller, K. R. (2010). “Finding stationary brain sources in EEG data,” in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (Buenos Aires), 2810–2813. doi: 10.1109/IEMBS.2010.5626537
- von Bünau, P., Meinecke, F. C., Király, F. C., and Müller, K.-R. (2009). Finding stationary subspaces in multivariate time series. *Phys. Rev. Lett.* 103:214101. doi: 10.1103/PhysRevLett.103.214101
- Webb, R., Glimcher, P. W., Levy, I., Lazzaro, S. C., and Rutledge, R. B. (2013). Neural random utility: relating cardinal neural observables to stochastic choice Behaviour’. *SSRN*. doi: 10.2139/ssrn.2143215

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 © 2017 Sundararajan, Palma and Pourahmadi. 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) or licensor 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.

A. APPENDIX

A.1. List of Food Snack Products Used in the Experiment

A.1.1. Experimental Instructions

The Food Choice Task will proceed as follows:

1. This stage consists of 10 choice situations.
2. In each trial, you will be presented with two food products.
3. You need to choose which of the two products you would prefer to eat.
4. Your decisions are real. At the conclusion of the experiment, one decision will be randomly selected to be binding.
5. You will receive one single unit of the food product you chose and will have to eat it at the end of today's session.

TABLE A1 | Food snack choices.

Choice	Product A	Product B
1	Low calorie Jell-O gelatin (10 cal.)	Original Jell-O gelatin (70 cal.)
2	Oven baked Lays potato chips (120 cal.)	Classic Lays potato chips (160 cal.)
3	No sugar added Dole peaches (25 cal.)	Original Dole peaches (70 cal.)
4	Light Yoplait yogurt (90 cal.)	Original Yoplait yogurt (150 cal.)
5	Fat free Pringles potato chips (70 cal.)	Original Pringles potato chips (150 cal.)
6	Sugar free Snack Pack pudding (70 cal.)	Original Snack Pack pudding (110 cal.)
7	Reduced fat Sargento string cheese (50 cal.)	Original Sargento string cheese (80 cal.)
8	Non-fat Oikos Greek yogurt (120 cal.)	Traditional Oikos Greek yogurt (150 cal.)
9	Reduced fat Cheez-It baked crackers (130 cal.)	Original Cheez-It baked crackers (150 cal.)
10	Diet Lipton green tea (0 cal.)	Traditional Lipton green tea (100 cal.)

The pictures of all products were taken from the WalMart's website where they were purchased.