Original Research Article:

Improving the accuracy of brain activation maps in the Group-Level analysis of fMRI data utilizing Spatiotemporal Gaussian Process model

Azam Saffar[1], Vahid Malekian[2], Majid Jafari Khaledi[3], Yadollah Mehrabi*[4]

1. ***Ms. Azam Saffar:*** Department of Biostatistics, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Email: azam.saffar66@gmail.com

2***. Dr.Vahid Malekian:*** Wellcome Centre for Human Neuroimaging, UCL Queen Square Institute of Neurology, University College London, London, United Kingdom: v.malekian@ucl.ac.uk

3. ***Dr.Majid Jafari Khaledi:*** Department of Statistics, Tarbiat Modares University, Tehran, Iran. Email: jafari-m@modares.ac.ir

4. Corresponding Author: Dr.Yadollah Mehrabi

Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran, ORCiD: https://orcid.org/0000-0001-9837-4956, Email: mehrabi@sbmu.ac.ir

CRediT author statement

Azam Saffar: Data curation, Writing- Original draft preparation, Software, Formal analysis. ***Vahid Malekian***: Validation, Writing review and editing. ***Majid Jafari Khaledi***: Formal Analysis, ***Yadollah Mehrabi:*** Supervision, Project administration, Validation, Writing review and editing.

High lights:

- Brain activation maps are critical tools for insight into the brain cognition system and the prevention, diagnosis, and treatment of many brain disorders.
- Incorporating spatiotemporal dependencies result in a more accurate brain activation map.
- The STGP model considers the non-separable unstructured spatiotemporal correlation in the calculation.
- The STGP model resulted in more accurate and precise maps in exprimental and simulated data compared to two common scenarios of fMRI data analysis, i.e., the GLM and Bayesian approaches.

Abstract:

Objective: Accuracy and precision of the statistical analysis methods used for brain activation maps are essential. Adjusting models to consider spatiotemporal correlation embedded in fMRI data may increase their accuracy, but it also introduces a high computational cost. The present study aimed to apply and assess the spatiotemporal Gaussian process (STGP) model to improve accuracy and reduce cost.

Methods: We applied the spatiotemporal Gaussian process (STGP) model for both simulated and experimental memory tfMRI data and compared the findings with fast, fully Bayesian, and General Linear Models (GLM). To assess their accuracy and precision, the models were fitted to the simulated data (1000 voxels,100 times point for 50 people), and an average of accuracy indexes of 100 repetitions was computed. Functional and activation maps for all models were calculated in experimental data analysis.

Results: STGP model resulted in a higher Z-score in the whole brain, in the 1000 most activated voxels, and in the frontal lobe as the approved memory area. Based on the simulated data, the STGP model showed more accuracy and precision than the other two models. However, its computational time was more than the GLM, as the price of model correction, but much less than that of the fast, fully Bayesian model.

Conclusion: Spatiotemporal correlation further improved the accuracy of the STGP compared to the GLM and fast, fully Bayesian model. This can result in more accurate activation maps. Moreover, the STGP model's computational speed appears to be reasonable for model application.

Keywords:

fMRI data analysis, Brain mapping, Spatiotemporal Gaussian Process Model, GLM, Accuracy Assessment

1. Introduction:

Functional magnetic resonance imaging (fMRI) is a neuroimaging technique for brain mapping that measures changes in oxygenated and deoxygenated blood magnetization recorded in Blood Oxygenated Level Dependent (BOLD) signals [1,2]. Brain mapping is used to associate mental function and its spatial location in the brain's anatomical structure. BOLD and the identification of the neural activation relationship make fMRI a tool for mapping brain activity [3]. Contrary to most previous brain imaging methods, fMRI is a non-invasive way to map brain function. Such maps are critical tools for insight into the brain cognition system and the prevention, diagnosis, and treatment of many brain disorders [1,4].

There are two main approaches for fMRI experiments: task-based (tfMRI), in which participants either perform a specific task or receive stimuli in an MRI scanner while BOLD signals are recorded, and resting-state fMRI (rsfMRI), which includes no stimulus events and the participants' BOLD signals are recorded as they rest quietly in the scanner [2,3,5]. As tfMRI enables researchers to link intended behavioral or cognitive tasks and brain activations, it has become a crucial tool for brain mapping [2].

Statistical analysis applied to make inferences about the significance of the effect of manipulated experiments on brain activation is divided into two categories: data-driven and model-based methods [1]. Data-driven approaches need no task information. However, model-based methods are extensively used when there is information about the task and stimulation. Model-based techniques are generally easier to implement and interpret[6].

The most common models for analyzing tfMRI data are based on General Linear Models (GLM) family that was first applied in this field by Friston [7]. As fMRI experiments make a sequence of three-dimensional images at each time point during scanning, spatiotemporal correlations are unquestionably embedded in acquired 4D data [1,3,8]. Models that are used in common analysis pipelines, like those obtained using the FSL software package (FMRIB, University of Oxford), AFNI (Analysis of Functional Neuro-Images software), and SPM (Statistical Parametric Mapping software) do not consider this complex spatiotemporal correlation structure. Instead, they try to account for such structure in the preprocessing stage or post-process corrections [9,10]. Ignoring

this residual correlation structure may lead to an underestimation of parameters' variability and consequently increase the false-positive rate (FPR) up to 70%, especially in the group-level analysis [1,10–13]. Moreover, it may restrict the reliability and accuracy of brain activation findings [1]. Z-score is defined as an estimated parameter divided by its standard error, used to determine brain activation. Therefore, activation maps known as Z-score (or statistical parametric) maps are based on the significance level of the estimated regression coefficients. Their accuracy and precision depend on the properties of the estimated model coefficients and their variances.

One of the most common methods in the context of model-based fMRI data analyses is the Bayesian method [14–16]. Bayesian approaches are popular for their estimation ability, and the preparation of appropriate activation maps without pre-smoothing [17]. In the spatiotemporal Bayesian framework, various prior distributions for spatial and temporal correlation (almost separated) are considered to model the data [18]. Incorporating spatiotemporal correlations results in a computational burden of high-dimensional parameters in the posterior distributions [19]. To simplify these calculations, certain prior structure correlations such as Markov Random Field (MRF) and Conditional autoregressive (CAR) are considered. These restricted correlation structures may not accurately approximate the actual structure, which is used to estimate and produce activation maps [18].

So far, several studies have been conducted on modeling fMRI data with spatiotemporal correlations. Different ideas have been proposed to define the brain voxel spatial neighborhood, such as six physically adjacent voxels in three orthogonal directions[20], contiguous first-order neighboring voxels[21], all pairs of exchangeable voxels within functionally grouped voxels [22], and finally, using neuroanatomic parcellation to define groups of voxels [23]. Since previous researches have shown that functional brain activation can occur in separate brain areas, the last two neighborhood definitions for brain activation findings seem to be more suitable than the others [15].

Spatiotemporal correlations can be considered as separable or inseparable structures, depending on time and space. Separability means that the total covariance matrix can be factorized as a product of the time covariance with the spatial covariance matrixes. Separability is an underlying assumption of spatiotemporal processes that must be assessed. Failure to establish this hypothesis leads to misspecification of the mean or covariance structure. [24]. Aston *et al.* [25] suggested a

method for separability hypothesis testing in high-dimensional and hyper-surfaces. However, when separability is considered, it does not allow any interaction between time and space and leads to discontinuity in correlations [24]. As an example of using separable structure modeling, Lorenzi *et al.* [26] considered a spatiotemporal structure using the kernel convolution of a white noise Gaussian process. Furthermore, Lorenzi *et al.* [27] applied a Kronecker product of spatial and temporal covariance. As an example of non-separable structure modeling, Long et al. [28] used a non-parametric Functional Data Analysis (FDA) method to consider non-stationarity with time-varying property in an fMRI data non-separable covariance structure. Non-separability is often more realistic, especially when space and time covariance come from Fourier transformation like fMRI data [24].

For ultra-high dimension 4D data, models usually introduce numerous estimated parameters especially, when incorporating non-separable spatiotemporal correlations. So, their application is limited by computational complexity and cost. Most statistical models for fMRI data have been designed for single subject or region of interest (ROI) based data analysis to prevent this kind of complexity [29]. Statistical power can be improved by multi-subject studies that include possible heterogeneity among subjects [15,30]; however, their high computational cost and complexity have been a prominent challenge for group-level data analysis in Cognitive Neuroscience and the medical image analysis field [31].

The GLM does not incorporate spatial or temporal correlations, which results in imprecise estimations. Moreover, although the Bayesian approach can take the spatiotemporal correlation into account, these methods are faced with the problem of covariance structure and also time-consuming computations. In this work, we applied the spatiotemporal Gaussian process (STGP) model introduced by Hyung *et al.* [4] to analyze positron emission tomography (PET) and MRI longitudinal data. Due to its ability to handle massive data incorporating spatiotemporal correlation for group-level analysis, we customized the model to analyze fMRI data. The STGP can be a computationally efficient framework for approximating unstructured non-separable huge spatiotemporal covariance in group fMRI data analysis [4,18]. Using this model, we expect to produce a more accurate activation map due to its residual structure adjustment for the non-separable spatiotemporal correlation in group-level data. This decreases FPR and leads to achieving more accuracy in functional maps. In addition, it is computationally efficient when using

FPCA and parcellation. A simulation study was conducted to evaluate the model's accuracy, precision, and computational speed compared with an ordinary GLM and a fast, fully Bayesian spatiotemporal model for fMRI introduced by Musgrove [29].

The rest of the article is organized as follows: In Section 2, we introduce the experimental tfMRI data used in this study and the basics of the STGP model. In Section 3, we explain the estimation process, the details of the STGP model application, and the simulation scenarios and its results. In Section 4, we showed the results of fitting the STGP model, spatiotemporal Bayesian model, and standard Group-Level-GLM to real data. All results are discussed in Section 5, and finally, the last section is the conclusion.

2. Materials and Methods:

We considered memory task-based data by Lewis-Peapock [32] that are published online at https://openneuro.org/datasets/ds001497/versions/1.0.2. They acquired whole-brain fMRI images with a 3.0 Tesla scanner for ten volunteers, including seven men and three women aged 19-32 years old, all right-handed with no history of medical, neurological, or psychiatric illness. Each participant had a high-resolution structural image (30 axial slices, 0.937×0.937 ×4mm) and a functional image (30 axial slices, 3.75×3.75 ×4mm) of BOLD signals acquired using a gradient-echo echo-planner sequence (TR=2000ms, TE=50ms) that lasted for 6 minutes and 50 seconds [32]. The stimulation task was done via a block design: 30 seconds rest and 30 seconds memory task. During the task time, the participants viewed 15 non-repeated random pictures of famous locations, people, or objects. Then they indicated their answers on a 4-point Likert scale using a stimulus-response box to show the degree of attraction to the celebrity, visiting the location, or encountering the object every day [18,32]. The prefrontal cortex in the frontal lobe was expected to be the main activated area in the brain, as the previous study showed [32].

We preprocessed the data using the FSL software package [33] because of its more accurate estimation in single-level computations compared to other analysis tools such as SPM, AFNI, etc. [12]. The following steps were taken for preprocessing: the Brain Extraction Tool (BET) was used to remove non-brain tissues, motion correction, slice timing correction, spatial smoothing, and registration of the subject's structural and Montreal Neurological Institute templates (MNI

atlas) images. Then, we analyzed the data using the spatiotemporal Gaussian process (STGP) model. Figure 1 shows the steps in STGP modeling.

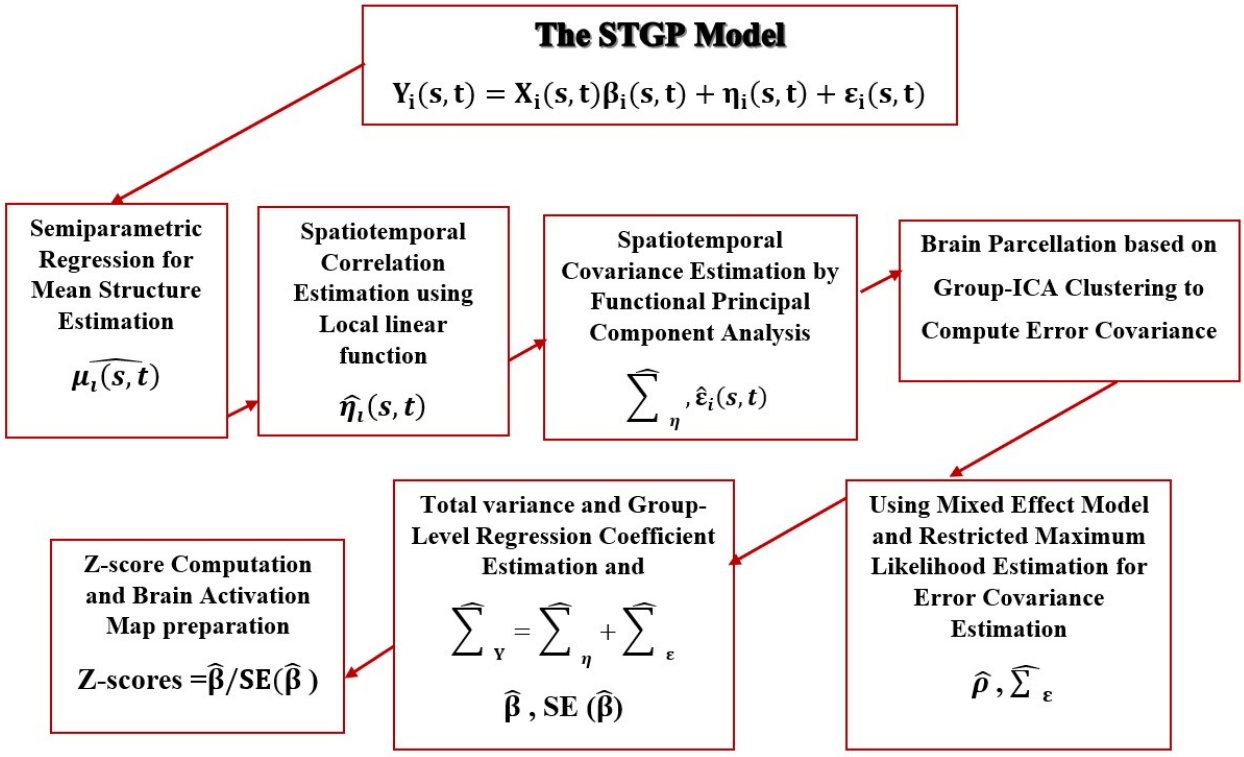

Fig.1. Diagram of the STGP model used to analyze the tfMRI data.

More details about calculation and model assessment by simulation are explained in the next section.

We used Matlab 2019b and R version 3.6.1 and FSL software to conduct the analyses. The computation system RAM was 8 Gigabytes with a Core-i7 CPU

3. Theory and calculation:

3.1. Spatiotemporal Gaussian process (STGP) model calculation:

Consider a set of fMRI data for a group of n subjects. For each subject, we observed the intensity of the BOLD signal denoted by $Y_i(s,t)$ (i=1,2,…,n) at each three-dimensional (3D) voxels and each time point (4D data); t shows the measuring time-point (t∈Ω={1,2,…,T}) and s is the spatial

location of the voxel ($s \in \mathcal{S}=\{1,2,\ldots,S\}$). At each time point, we have the whole-brain (all voxels) intensity and a vector of p covariates. The STGP model for fMRI data is given by:

$$Y_i(s,t) = \mu_i(s,t) + \eta_i(s,t) + \varepsilon_i(s,t) \quad (1)$$

where

$$\mu_i(s,t) = X_i(s,t)\beta_i(s,t) \quad (2)$$

and $\beta_i(s,t)$ is the effect of the covariate $X_i(s,t)$ and $\mu_i(s,t)$ is the mean structure for characterizing this effect for the i^{th} subject across (s, t). For the i^{th} subject $\eta_i(s,t)$ are random functions characterizing the spatiotemporal correlation structure across (s, t), and $\varepsilon_i(s,t)$ are the measurement errors. It is assumed that $\eta_i(s,t)$ and $\varepsilon_i(s,t)$ are mutually independent and identically distributed as $GP(0, \sum_\eta)$ and $GP(0, \sum_\varepsilon)$, respectively; where $GP(\mu, \sum)$ denotes a Gaussian Process with the mean function $\mu(s,t)$ and covariance function $\sum((s,t),(s',t'))$.

Incorporating spatiotemporal correlation, we want to adjust Y_i covariance estimation for a heterogeneous structure that is the sum of the spatiotemporal random function covariance and the measurement error covariance. Ordinary estimation of β_i and its covariance must be corrected with respect to spatiotemporal correlation, considering inverse covariance of Y_i as the estimation weight [34]. Then, a new Z-score will be computed based on these adjusted estimations to reach a more accurate activation map. The STGP model uses the following steps to estimate the true covariance of Y_i based on spatiotemporal observations and measurement errors:

1. We estimated the mean structure μ_i using a semiparametric regression model and obtained the residuals: $r_i(s,t) = Y_i(s,t) - \widehat{\mu_i(s,t)}$;
2. These residuals were used to estimate spatiotemporal random functions $\widehat{\eta_i}(s,t)$ using a local linear function by smoothing r_i. Having more than one hundred time points, we faced dense sampling data, so an optimum boundary of a univariate kernel function k_h:

$$k_h(s_m - s) = h^{-3} \prod_{j=1}^{S} k((s_{m,j} - s_j)/h) \quad (3)$$

was used to estimate η_i with respect to β_0 and β_1:

$$\hat{\eta}_i(s,t) = \sum_{m=1}^{S}\sum_{j=1}^{T}\left\{r_i(s_{m,j},t_{ij}) - \beta_0 - \beta_1^T\left((s_{m,j}-s_j)^T, t_{ij}-t\right)^T\right\}^2 k_h(s_{m,j}-s_j)k_{h_1}(t_{ij}-t) \quad (4)$$

where:

$$k_{h_1}(t_{ij}-t) = k(\frac{(t_{ij}-t)}{h_1})/h_1 \quad (5)$$

The optimal boundaries (h, h_1) were selected by minimizing the generalized cross-validation (GCV) score.

3. To estimate the covariance of spatiotemporal Gaussian random functions, $\sum_{\eta}$, we first checked its separability hypotheses using the Aston *et al*. test [25], which was rejected (P-value=<0.001, 0.022, and 0.038). In this non-separable unstructured spatiotemporal covariance matrix, we faced a daunting 23,961,600 parameters (122880 voxels × 195-time points) for our data. By taking it into standard brain space, they became 176,012,655 parameters (902629 voxels ×195 time points) in an MNI standard format (average152T1). The STGP model uses the Functional Principal Component (FPC) model to consider global spatiotemporal dependencies. Using FPC allows regression coefficients to vary across the brain because it estimates spatiotemporal basis functions directly. In addition, it enables STGP to quickly and easily estimate the huge spatiotemporal dependencies matrix. We used FPC's of an average of the product of the $\hat{\eta}_i$ of all subjects single spectral decomposition [35]:

$$\widehat{\sum}_{\boldsymbol{\eta}}\left((s,t),(s',t')\right) = n^{-1}\sum_{i=1}^{n}\hat{\eta}_i(s,t)\hat{\eta}_i(s',t') \quad (6)$$

$$\widehat{\sum}_{\boldsymbol{\eta}} = \sum_{k=0}^{\infty}\lambda_k\psi_k(s_i,t_i)\psi_k(s_j,t_j) \quad (7)$$

Where $\lambda_1 > \lambda_2 > \lambda_3 > \cdots 0$ are the ordered eigenvalues of linear operation determined by $\sum_{\eta}$ and $\psi_k(s,t)$ are the corresponding orthonormal eigenfunctions:

$$\hat{\eta}_i(s,t) = \sum_{k=0}^{k_0} \xi_{i_k} \psi_k(s,t) \quad (8)$$

Where:

$$\xi_{i_k} = \int_0^T \int_0^S \hat{\eta}_i(s,t)\ \psi_k(s,t) dt\, dv(s) \quad (9)$$

ξ_{i_k} is referred to k^{th} FPC score for the i^{th} subject, and $dv(s)$ is the Lebesgue measure [36]. ξ_{i_k} is a random variable with $E(\xi_{i_k}) = 0$ and $E({\xi_{i_k}}^2) = \lambda_k$. Some first ordered eigenvalues were used to approximate the spatiotemporal random function covariance structure, $\widehat{\sum}_{\boldsymbol{\eta}}$, because they explained at least 70% of the total relative variation [37]. Here, $\widehat{\sum}_{\boldsymbol{\eta}}$ is the first element of the total variance of Y_i estimation, $\widehat{\sum}_{\boldsymbol{Y}}$. Afterward, measurement errors were estimated by:

$$\hat{\varepsilon}_i(s,t) = Y_i(s,t) - \hat{\mu}_i(s,t) - \sum_{k=0}^{m} \xi_{i_k} \psi_k(s,t) \quad (10)$$

4. This model used a partitioned covariance estimation to capture error dependencies. The same parcellation, anatomical or functional methods are usually used to partition all subjects' brains. Since distant voxels can have a similar activity pattern in studies such as the present one, activity-based (functional) parcellation is more appropriate[15]. Therefore, Group Independent Component Analysis (GICA) was conducted using the FSL software package (melodic) for parcellation of all subjects' brains together [33]. Using GICA, we identified some brain clusters that have the same activity pattern. These clusters were used to define brain parcellations or regions of interest (ROI) for error covariance estimation.
5. To approximate the error covariance structure, a mixed-effects model was implemented on each subject's brain ROI,(developed in the previous step. This model was used to estimate the measurement error distribution covariance parameters by restricted maximum likelihood estimation (RMLE) for the first-order autoregressive model in each area. Errors $\varepsilon_i(s,t)$ were assumed to be independent across the partitioned brain regions, while the spatiotemporal correlation was preserved within each subdivision. Then the estimation of $\sum_\varepsilon$ became a block-

diagonal matrix, with each block belonging to one of the ROIs in an AR(1) format, and the ROIs are considered independent.

6. The total variance of Y_i or the STGP model, $\widehat{\sum}_{\mathbf{Y}}$, was computed by adding up $\widehat{\sum}_{\eta}$ to $\widehat{\sum}_{\varepsilon}$;

$$\widehat{\sum}_{\mathbf{Y}} = \widehat{\sum}_{\boldsymbol{\eta}} + \widehat{\sum}_{\varepsilon} \tag{11}$$

After total variance computation, the adjusted estimation of β_i was computed by the Weighted Least Square (WLS) method with taking the total variance as inverse weight; then, the model estimation of observation is calculated as follows:

$$\hat{\beta} = (\mathrm{X}' \widehat{\sum}_{\mathrm{Y}}^{-1} \mathrm{X})^{-1} \mathrm{X}' \widehat{\sum}_{\mathrm{Y}}^{-1} \mathrm{Y} \tag{12}$$

$$\widehat{Y_i}(s,t) = \mathrm{X_i}(\mathrm{s,t})\widehat{\beta_i}(\mathrm{s,t}) + \widehat{\eta_i}(\mathrm{s,t}) + \widehat{\varepsilon_i}(\mathrm{s,t}) \tag{13}$$

Calculating the $\hat{\beta}$ variance adjusted for spatiotemporal correlation, the new corrected Z-score was calculated.

$$cov(\widehat{\beta}) = (\mathrm{I} - \mathrm{X}\left(\mathrm{X}' \widehat{\sum}_{\mathrm{Y}}^{-1} \mathrm{X}\right)^{-1} \mathrm{X}')\mathrm{Y}\left(\mathrm{X}' \widehat{\sum}_{\mathrm{Y}}^{-1} \mathrm{X}\right)^{-1} \tag{14}$$

$$Z - score = \hat{\beta}\,(cov(\widehat{\beta}))^{-1/2} \tag{15}$$

The corrected brain activity pattern map was computed using the STGP model and then compared with Group-Level-GLM output from the FSL software package (Feat-second level analysis) and the fast, fully Bayesian spatiotemporal model. They were visualized using xjView (https://www.alivelearn.net/xjview) and FSL eye toolboxes.

3.2. Simulation study

A simulation study was conducted to evaluate the STGP model and to compare its properties with the Group-Level GLM and fast, fully Bayesian spatiotemporal models. Imaging data with an

S=10×10×10 set of voxels were generated for n=50 subjects at T=100 time points. We considered a fixed design matrix of block task of 30 seconds rest and 30 seconds task that was convolved by a double-gamma canonical hemodynamic response function (HRF). The random function of the spatiotemporal correlation structure

$$\eta_i(s,t) = \sum_{k=0}^{k_0} \xi_{i_k} \psi_k(s,t)$$

were set for k_0=3 and

$$\xi_{1_k} \sim N(0, 7^2), \qquad \xi_{2_k} \sim N(0, 3^2), \qquad \xi_{3_k} \sim N(0, 1.5^2).$$

Based on three eigenfunctions of applied real data [4,18,35], the eigenfunctions were considered as:

$$\psi_{1i}(s,t) = \frac{1}{5\sqrt{30}} \sin\left(\frac{\pi s}{5}\right) \cos\left(\frac{\pi t}{18}\right) \tag{16}$$

$$\psi_{2i}(s,t) = \frac{1}{5\sqrt{30}} \cos\left(\frac{\pi s}{5}\right) \cos\left(\frac{\pi t}{18}\right) \tag{17}$$

$$\psi_{3i}(s,t) = \frac{1}{2\sqrt{255}} \left(0.9 - \frac{s}{5}\right) \cos(\pi t) \tag{18}$$

Ten parcels were defined for error simulation, and at each parcel, $\varepsilon_i(s,t)$ was considered as a zero-mean Gaussian Process. The non-separable space-time covariance function proposed by Bourotte [38] was defined as the covariance structure.

The GLM, STGP and Spatiotemporal Bayesian models were fitted to each simulated dataset. Then, we assessed the models using measures of accuracy (the percentage of voxels whose status in terms of activation or inactivity is correctly estimated by the model), precision (the percentage of active voxels that the model has correctly estimated to be active), false positive rate, FPR (the ratio of active voxels that are incorrectly detected by the model as active voxels to the total number of active voxels), false negative rate, FNR (the percentage of inactive voxels that are wrongly seen by the model as inactive voxels to the total number of inactive voxels). Also, Mean of Absolute

Error (MAE), Root Mean Square of Error (RMSE) and Normalized Mean Square Error (NMSE) were computed and reported. It is clear that models with higher accuracy and precision and lower FPR, FNR, MAE, RMSE, and NMSE≅ 1 are preferred. By concatenating Y_is as Y=($Y_1, Y_2, \dots, Y_n$) and setting N= $S \times T \times n$, we calculated MAE, RMSE, and NMSE using the following formulas:

$$MAE = (Absolute(Y - \hat{Y}))(1'_N 1_N)^{-1} \quad (19)$$

$$RMSE = \left((Y - \hat{Y})'(Y - \hat{Y})(1'_N 1_N)^{-1}\right)^{1/2} \quad (20)$$

$$\text{NMSE} = (Y - \hat{Y})'(Y'Y)^{-1}(Y - \hat{Y}) \quad (21)$$

This process was repeated 100 times and the mean of the calculated measures are reported in Table 1.

Table 1. The model comparison indexes for GLM, Bayesian Spatiotemporal Model and STGP Model based on the simulation study.

Models	Accuracy (%)	Precision (%)	FPR (%)	FNR (%)	MAE	RMSE	NMSE
Group-Level-GLM	86.7	78.32	4.35	3.27	3.576	3.85	2.32
Bayesian ST model	93.8	95.65	1.08	0.9	2.952	2.509	0.432
STGP Model	99.5	99.95	0.41	0.1	1.352	1.745	0.934

4. Results:

The STGP model, the ordinary Group-GLM, and a fast, fully Bayesian spatiotemporal model were fitted to tfMRI (memory task) data, and their performances were compared. Preprocessing steps were similar for all models. First level and Group-Level-GLM analyses were conducted using FSL software.

Using the Spatiotemporal Gaussian Process (STGP) model, $\hat{\mu}_i$ and r_i were computed for each subject. According to the FPCA results, the first three eigenvalues could explain about 90% of the total relative variation (Figure 2).

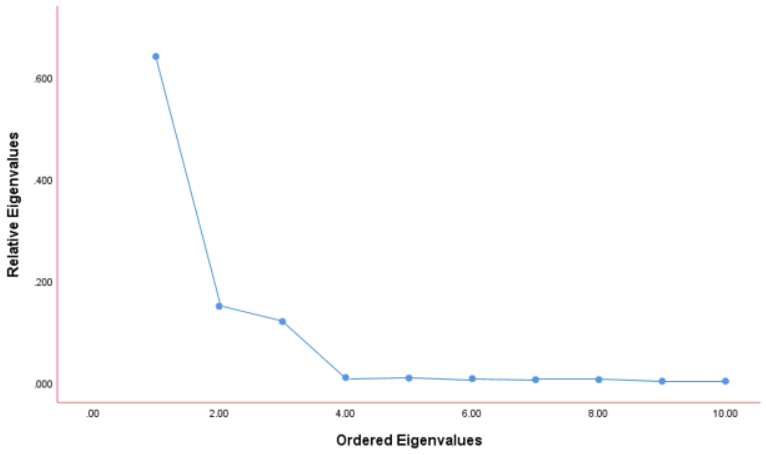

Fig. 2. The first ten eigenvalues of $\widehat{\sum}_{\eta}$ for the experimental tfMRI memory data.

All-subject brain parcellation was conducted using GICA in the FSL(Melodic toolbox) software package. Based on GICA results, 11 activity-based regions of interest (ROI) were selected, one of which consisted of all inactive voxels. Therefore, the error correlation was estimated using the spatiotemporal autoregressive model with the REML method[39]. In the next step, total variance estimation, adjusted $\hat{\beta}$, and its variance were computed. Figure 3 represents the voxel-wise Z-scores ($\hat{\beta}/\mathrm{SE}(\hat{\beta})$) for the three models.

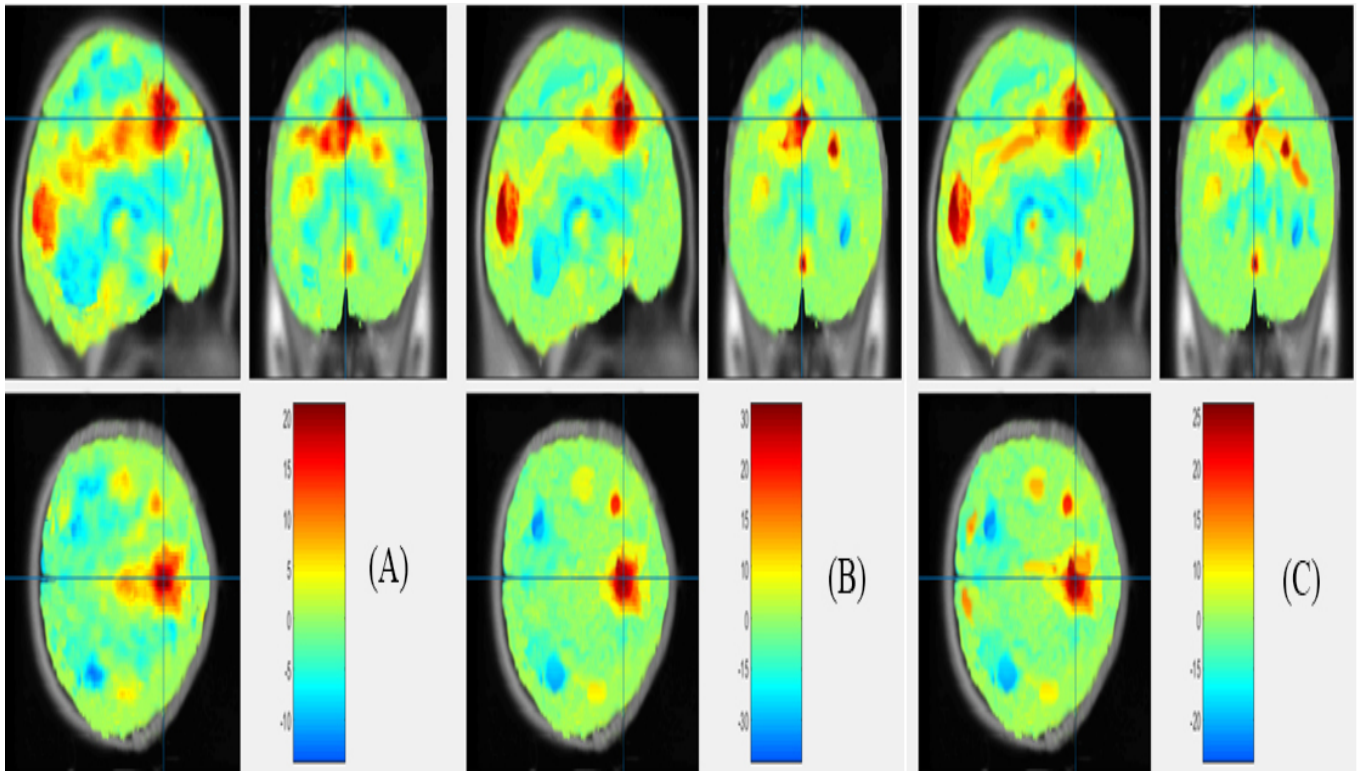

Fig. 3. Comparison between activation (significant voxel) maps of selected slides of the functional image: (A) Group-Level-GLM, (B) STGP Model, and (C) Fast, fully Bayesian model. The main activated area was the same for all models, but adjustment for spatiotemporal correlation in the STGP and Fast, fully Bayesian model lead to surpassed areas with higher Z-score.

The absolute minimum and maximum Z-scores in the STGP were considerably higher than those of the Group-Level GLM and somewhat higher than the fast, fully Bayesian spatiotemporal model (Table 2).

Table2. Z-scores in the Group-Level-GLM, Bayesian Spatiotemporal Model, and STGP Model.

Models	Minimum	Maximum	Mean Absolute in 1000 higher activated voxels	Mean in the main activated area in Prefrontal cortex
Group-Level-GLM	-14.8	20.96	10.178	12.1
Bayesian Spatiotemporal Model	-20	22.32	18.2	22.23
STGP Model	-33.1	32.2	31.91	30.3

The computed family-wise error rate (FWER) threshold was 5.99. Based on this threshold, the brain activation maps are shown in Figure 4.

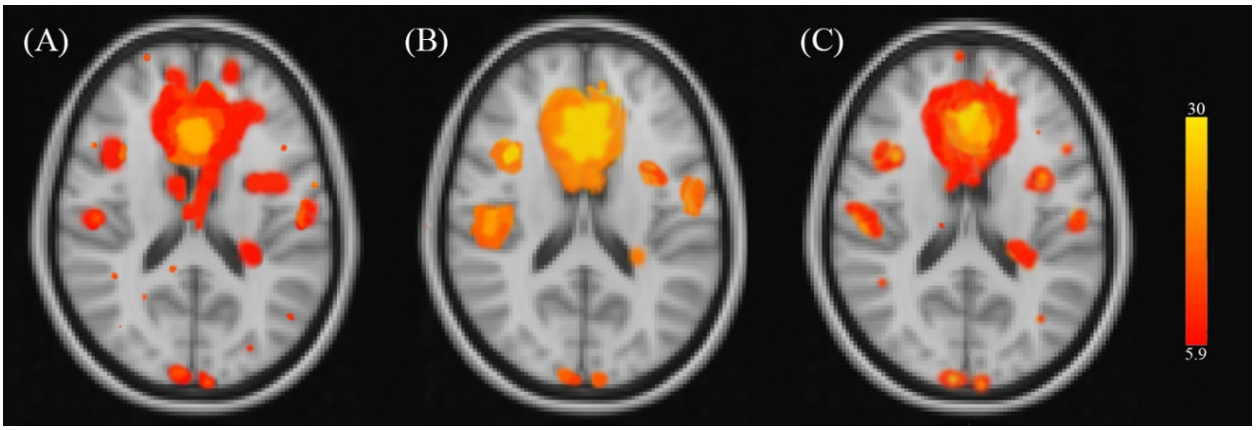

Fig. 4. The axial view of activation (Z-score) maps for : (A) Group-Level-GLM, (B) STGP Model, and (C) Fast, fully Bayesian model. The main activated area in all models was the same in the super motor area. The STGP model had a higher Z-score but less activated voxels compared to the group-level GLM and Fast, fully Bayesian model. The STGP model had a higher Z-score but less activated voxel, especially in diffused voxels.

The STGP model, Group-Level GLM, and fast, fully Bayesian spatiotemporal model found 14.5%, 25%, and 19% activated voxels in the whole brain, respectively. Mean computational time consumed for single, and group-level (for real and simulated data) calculations for the models are reported in Table 3.

Table 3. Time (in minutes) consumed for data analysis by the GLM, Bayesian Spatiotemporal Model, and STGP Model.

Models	Single-level analysis	Group-level analysis (10 subjects)	Simulated data Group-level analysis (50 subjects,1000 voxels)
Group-Level-GLM	17	192	116
Bayesian ST model	780	7998	4793
STGP Model	112	1312	702

5. Discussion

This research aimed to assess the efficiency and accuracy of the STGP model compared to two common scenarios of fMRI data analysis, i.e., the GLM and Bayesian approaches. As statistical

analysis is the basis of reported brain activation (Z-score) maps, it is obvious that inaccurate analysis can lead to imprecise brain maps [10].

fMRI data are acquired in three-dimensional space during the scanning time, so when the spatiotemporal correlation is embedded, it makes it four-dimensional. Ignoring this correlation may lead to underestimating the variability and an increased false-positive rate [1]. Additionally, when fMRI data moves from a k-space inverse Fourier transformation to the time-space, a more realistic spatiotemporal covariance structure would be non-separable [24]. However, incorporating an ultra-high-dimension non-separable unstructured spatiotemporal covariance matrix causes difficulties in computations. Hence, a spatiotemporal model with the ability to consider huge non-separable unstructured spatiotemporal covariance for fMRI data with a lower computational burden and more accuracy seemed necessary. We applied the proposed STGP model, which considers spatiotemporal correlation implanted in the fMRI data. The STGP model used FPCA for spatiotemporal covariance and parcellation for error covariance estimation. To evaluate this model in comparison to the GLM and fast, fully Bayesian models, we fitted them to experimental tfMRI memory and also simulated datasets.

In the real data, the main activated areas were the same for all three models, the occipital and frontal lobes, which is in agreement with results in previous work on this data [32]. Moreover, these results were also confirmed by previous works on memory task-based stimuli [32,40–42]. Recent studies have shown that the frontal lobe and the posterior sensory-motor area, where the actual memory is demonstrated, were retrained [40].

Z-score is a measure that shows how changes in the BOLD signal of a voxel follow a normal pattern. A higher absolute Z-score shows more likely activation. After adjusting for spatiotemporal correlation, it is expected that the variance estimation of regression coefficients, $\mathrm{cov}(\widehat{\beta})$, will increase, and the Z-score will decrease. Then, the false activation areas must also decrease. On the other hand, spatiotemporal adjustment corrects the regression coefficient estimation, $\widehat{\beta}$. Increasing the absolute Z-scores, considering these two corrections together, shows stronger activation findings.

In this study, the STGP model increased the absolute maximum Z-score by about 18 and 13 units compared to the group-level GLM and fast, fully Bayesian model. The adjusted Z-score for

spatiotemporal correlation in 1000 highest activated voxels in the STGP model increased up to 21 and 13.7 units compared to the group-level GLM and fast, fully Bayesian model, respectively. The average Z-score in the frontal lobe, a specific confirmed area, was about 18 and 8 units more than the group-level GLM and fast, fully Bayesian model, respectively (Table.2.). An increasing Z-score showed stronger activation and suppressed activated areas in comparison with the two other models; this may be due to false activation removal. Activated voxels in the whole brain were observed to be ten percent less in the STGP model, especially the sporadic voxels in the GLM. This may be because the spatiotemporal correlation increases precision and accuracy while reduced diffused unrelated activated voxels, which were practically considered as noise in the computations.

In the simulation study, the STGP model resulted in more accurate results (higher accuracy and precision; NMSE$\cong$ 1; lower: FPR, FNR, MAE, and RMSE) compared to the other two models, presumably due to the spatiotemporal correlation adjustment (Table.1.).

On the aspect of computation time, the GLM for single and group-level analysis was the fastest model both on real and simulated data. Moreover, the fast, fully Bayesian model was more time-consuming than the two other models in both situations (Table.3.). This computational burden makes it very difficult when analyzing big group data. The lower fitting speed of the STGP and fast, fully Bayesian models is because of the spatiotemporal justification, as mentioned in previous studies [3,8,9,17,19]. Considering the non-separable unstructured spatiotemporal correlation, the calculation time consumption of the activation map findings for the STGP is lower than even a fast version of the Bayesian models. The STGP model showed better results for real and simulated data (more accurate with less FPR, higher absolute Z-score maps). However, its computational time is more than a simple model like the GLM.

There are several limitations to the STGP model application. All steps in the STGP model must be supervised, and the setting at each step depends on the result of the previous step. All following processes must also be supervised, which is not as easy as implementing a group-level GLM analysis. In addition, the STGP model is based on the Gaussian assumption in fMRI data. It would be better to evaluate Gaussian Process assumptions and adjust the model for cases where it would not be acceptable.

6. Conclusion

To conclude, the STGP model, which considers spatiotemporal dependencies in a single subject and group level fMRI data, resulted in more accurate and precise maps in real and simulated data. There is a tradeoff between accuracy, precision, and computational cost. The STGP is not as fast as GLM but provides less FPR and more accurate group-level voxel-wise task-based fMRI maps.

Declaration:

Competing Interest:

The authors declare that they have no conflict of interest.

Ethics approval:

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of Shahid Beheshti University of Medical Sciences (No. IR.SBMU.PHNS.REC.1397.033).

Acknowledgment:

The authors would like to thank Prof. Hongtu Zhu and Prof. Yimei Li for their guidance and for introducing us to the STGP model. Data collection for this project was supported by the NIH grant MH064498 (B.R.P) and was shared on www.Openneuro.org (https://openneuro.org/datasets/ds001497/versions/1.0.2). The authors would also like to thank the Shahid Beheshti University of Medical Sciences for grant No 15481.

References:

[1] S. Bollmann, A.M. Puckett, R. Cunnington, M. Barth, Serial correlations in single-subject fMRI with sub-second TR, Neuroimage. 166 (2018) 152–166. https://doi.org/10.1016/j.neuroimage.2017.10.043.

[2] L.K. Hansen, Statistical models for comprehensive meta-analyses of neuroimaging studies To cite this version : ` de doctorat Th ese, (2020).

[3] L. Zhang, M. Guindani, F. Versace, J.M. Engelmann, M. Vannucci, A spatiotemporal nonparametric Bayesian model of multi-subject fMRI data, Ann. Appl. Stat. 10 (2016) 638–666. https://doi.org/10.1214/16-AOAS926.

[4] J.W. Hyun, Y. Li, C. Huang, M. Styner, W. Lin, H. Zhu, STGP: Spatio-temporal Gaussian process models for longitudinal neuroimaging data, Neuroimage. 134 (2016) 550–562. https://doi.org/10.1016/j.neuroimage.2016.04.023.

[5] J. Ye, N.A. Lazar, Y. Li, Geostatistical analysis in clustering fMRI time series, Stat. Med. 28 (2009) 2490–2508. https://doi.org/10.1002/sim.3626.

[6] M. Ernst, S. Torrisi, N. Balderston, C. Grillon, E.A. Hale, FMRI functional connectivity applied to adolescent neurodevelopment, Annu. Rev. Clin. Psychol. 11 (2015) 361–377. https://doi.org/10.1146/annurev-clinpsy-032814-112753.

[7] K.J. Friston, A.P. Holmes, J.B. Poline, P.J. Grasby, S.C.R. Williams, R.S.J. Frackowiak, R. Turner, Analysis of fMRI time-series revisited, Neuroimage. 2 (1995) 45–53. https://doi.org/10.1006/nimg.1995.1007.

[8] C.A. Nader, N. Ayache, P. Robert, M. Lorenzi, Monotonic Gaussian Process for Spatio-Temporal Trajectory Separation in Brain Imaging Data, (2019). http://arxiv.org/abs/1902.10952.

[9] P. Sidén, F. Lindgren, D. Bolin, A. Eklund, M. Villani, Spatial 3D Mat\'ern priors for fast whole-brain fMRI analysis, (2019). http://arxiv.org/abs/1906.10591.

[10] L.L. Wald, J.R. Polimeni, Impacting the effect of fMRI noise through hardware and acquisition choices – Implications for controlling false positive rates, Neuroimage. 154 (2017) 15–22. https://doi.org/10.1016/j.neuroimage.2016.12.057.

[11] R.W. Cox, G. Chen, D.R. Glen, R.C. Reynolds, P.A. Taylor, FMRI Clustering in AFNI: False-Positive Rates Redux, Brain Connect. 7 (2017) 152–171. https://doi.org/10.1089/brain.2016.0475.

[12] A. Eklund, T. Nichols, H. Knutsson, Can parametric statistical methods be trusted for fMRI based group studies?, (2015). https://doi.org/10.1073/pnas.1602413113.

[13] K. Li, L. Guo, J. Nie, G. Li, T. Liu, Review of methods for functional brain connectivity detection using fMRI, Comput. Med. Imaging Graph. 33 (2009) 131–139. https://doi.org/10.1016/j.compmedimag.2008.10.011.

[14] L. Zhang, M. Guindani, M. Vannucci, Bayesian models for functional magnetic resonance imaging data analysis, Wiley Interdiscip. Rev. Comput. Stat. 7 (2015) 21–41. https://doi.org/10.1002/wics.1339.

[15] F.D. Bowman, Brain Imaging Analysis, Annu. Rev. Stat. Its Appl. 1 (2014) 61–85. https://doi.org/10.1146/annurev-statistics-022513-115611.

[16] G. Derado, F.D. Bowman, L. Zhang, Predicting brain activity using a Bayesian spatial model, Stat. Methods Med. Res. 22 (2013) 382–397. https://doi.org/10.1177/0962280212448972.

[17] P. Sidén, A. Eklund, D. Bolin, M. Villani, Fast Bayesian whole-brain fMRI analysis with spatial 3D priors, Neuroimage. 146 (2017) 211–225. https://doi.org/10.1016/j.neuroimage.2016.11.040.

[18] J.W. Hyun, Y. Li, J.H. Gilmore, Z. Lu, M. Styner, H. Zhu, SGPP: Spatial Gaussian predictive process models for neuroimaging data, Neuroimage. 89 (2014) 70–80. https://doi.org/10.1016/j.neuroimage.2013.11.018.

[19] M. Bezener, J. Hughes, G. Jones, Bayesian spatiotemporal modeling using hierarchical spatial priors, with applications to functional magnetic resonance imaging (with discussion), Bayesian Anal. 13 (2018) 1261–1313. https://doi.org/10.1214/18-BA1108.

[20] K. Katanoda, Y. Matsuda, M. Sugishita, A spatio-temporal regression model for the analysis of functional MRI data, Neuroimage. 17 (2002) 1415–1428. https://doi.org/10.1006/nimg.2002.1209.

[21] W.D. Penny, N.J. Trujillo-Barreto, K.J. Friston, Bayesian fMRI time series analysis with spatial priors, Neuroimage. 24 (2005) 350–362. https://doi.org/10.1016/j.neuroimage.2004.08.034.

[22] F.D. Bowman, Spatio-temporal modeling of localized brain activity, Biostatistics. 6 (2005) 558–575. https://doi.org/10.1093/biostatistics/kxi027.

[23] F.D.B. Bowman, B. Caffo, S.S. Bassett, C. Kilts, A Bayesian hierarchical framework for spatial modeling of fMRI data, Neuroimage. 39 (2008) 146–156. https://doi.org/10.1016/j.neuroimage.2007.08.012.

[24] W. Chen, M.G. Genton, Y. Sun, Space-time covariance structures and models, Annu. Rev. Stat. Its Appl. 8 (2021) 191–215. https://doi.org/10.1146/annurev-statistics-042720-115603.

[25] J.A.D. Aston, D. Pigoli, S. Tavakoli, Tests for separability in nonparametric covariance operators of random surfaces, Ann. Stat. 45 (2017) 1431–1461. https://doi.org/10.1214/16-AOS1495.

[26] M. Lorenzi, G. Ziegler, D.C. Alexandr, O. Sebastien, Efficient Gaussian Process-Based Modelling and Prediction of Image Time Series, in: Inf. Process. Med. Imaging, 2015: pp. 626–638. https://doi.org/10.1007/978-3-319-19992-4 Library.

[27] M. Lorenzi, G. Ziegler, D.. Alexander, S. Ourselin, Modelling Non-stationary and Non-separable Spatio-Temporal Changes in Neurodegeneration via Gaussian Process Convolution ., in: Mach. Learn. Meets Med. Imaging, 2015: pp. 35–45. https://doi.org/10.1007/978-3-319-27929-9.

[28] C.J. Long, E.N. Brown, C. Triantafyllou, I. Aharon, L.L. Wald, V. Solo, Nonstationary noise estimation in functional MRI, Neuroimage. 28 (2005) 890–903. https://doi.org/10.1016/j.neuroimage.2005.06.043.

[29] D.R. Musgrove, J. Hughes, L.E. Eberly, Fast, fully Bayesian spatiotemporal inference for fMRI data, Biostatistics. 17 (2016) 291–303. https://doi.org/10.1093/biostatistics/kxv044.

[30] B.B. Risk, D.S. Matteson, R.N. Spreng, D. Ruppert, Spatiotemporal mixed modeling of multi-subject task fMRI via method of moments, Neuroimage. 142 (2016) 280–292. https://doi.org/10.1016/j.neuroimage.2016.05.038.

[31] Y. Zhao, X. Li, H. Huang, W. Zhang, S. Zhao, M. Makkie, M. Zhang, Q. Li, T. Liu, 4D Modeling of fMRI Data via Spatio-Temporal Convolutional Neural Networks (ST-CNN), IEEE Trans. Cogn. Dev. Syst. (2019). https://doi.org/10.1109/TCDS.2019.2916916.

[32] J.A. Lewis-Peacock, B.R. Postle, Temporary activation of long-term memory supports working memory, J. Neurosci. 28 (2008) 8765–8771. https://doi.org/10.1523/JNEUROSCI.1953-08.2008.

[33] M. Jenkinson, C.F. Beckmann, T.E.J. Behrens, M.W. Woolrich, S.M. Smith, Review FSL, Neuroimage. 62 (2012) 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015.

[34] J. Fox, Applied Regression Analysis and Generalized Linear Models (2015, SAGE Publ.).pdf, 2015.

[35] J.L. Wang, J.M. Chiou, H.G. Müller, Functional Data Analysis, Annu. Rev. Stat. Its Appl. 3 (2016) 257–295. https://doi.org/10.1146/annurev-statistics-041715-033624.

[36] C. Heil, Lebesgue Measure BT - Introduction to Real Analysis, in: C. Heil (Ed.), Springer International Publishing, Cham, 2019: pp. 33–86. https://doi.org/10.1007/978-3-030-26903-6_2.

[37] R. Cangelosi, A. Goriely, Component retention in principal component analysis with application to cDNA microarray data, Biol. Direct. 2 (2007) 1–21. https://doi.org/10.1186/1745-6150-2-2.

[38] M. Bourotte, D. Allard, E. Porcu, A flexible class of non-separable cross-covariance functions for

multivariate space–time data, Spat. Stat. 18 (2016) 125–146. https://doi.org/10.1016/j.spasta.2016.02.004.

[39] C. Partlett, R.D. Riley, Random effects meta-analysis: Coverage performance of 95% confidence and prediction intervals following REML estimation, Stat. Med. 36 (2017) 301–317. https://doi.org/10.1002/sim.7140.

[40] A.H. Lara, J.D. Wallis, The role of prefrontal cortex in working memory: A mini review, Front. Syst. Neurosci. 9 (2015) 1–7. https://doi.org/10.3389/fnsys.2015.00173.

[41] J. Eriksson, E.K. Vogel, A. Lansner, F. Bergström, L. Nyberg, Neurocognitive Architecture of Working Memory, Neuron. 88 (2015) 33–46. https://doi.org/10.1016/j.neuron.2015.09.020.

[42] S. Funahashi, Working memory in the prefrontal cortex, Brain Sci. 7 (2017). https://doi.org/10.3390/brainsci7050049.