Feature Selection for ECoG-based Motor Imagery Analysis

Analysis of cortical activity during motor imagery data is essential for designing Brain Computer Interfaces. This involves recognition of intended movements from the recorded brain activity. Here we use ElectroCorticoGraphic (ECoG) signals, measured during the food-tracking activity [1, 2], to predict the reconstruct movements of the subject.

Feature extraction for ECoG data [3]. The raw input data contains a number of the time series of voltage measurements (1000Hz) from N electrodes ($N = 64$ or $N = 32$). These measurements are highly correlated for neighbouring electrodes. To convert the time series into the data sample $D = \{\underline{\mathbf{X}}_m, \mathbf{y}_m\}$, select M time points t_i, $i = 1, \dots, M$ with step 100ms. Targets $\mathbf{y}_i$ are 3D spatial coordinates (measured at 120Hz) of a certain body part of the subject, correspondent to t_i. The feature tensors $\underline{\mathbf{X}}_m \in \mathbb{R}^{\times T \times F \times N}$ represent spatial and time-frequency information about the segment $[t_m - 1s, t_m]$ of the time series. The spatial component is represented by N electrodes. To obtain $T \times F$ features in time-frequency domain, use the following procedure.

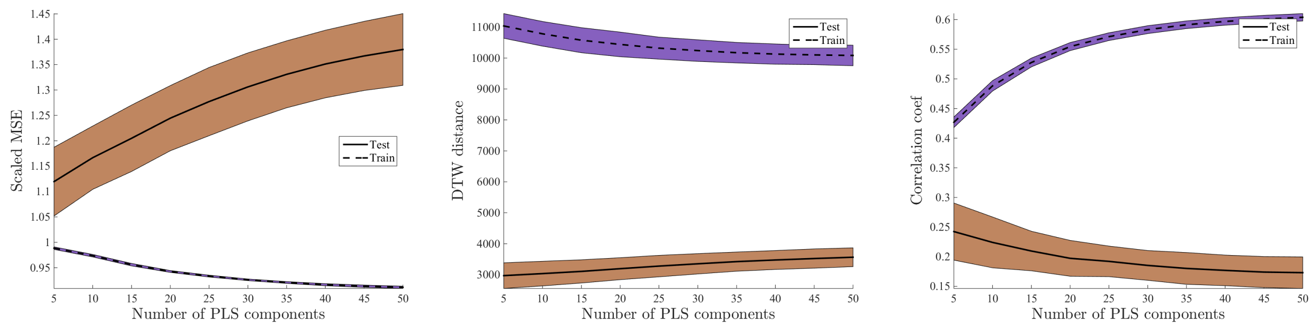
1. The time range $[t_m - 1s, t_m]$ is divided into consecutive intervals δt_i, $i = 1, \dots, 10 = T$, 100ms each (without overlap).
2. Ten basic logarithmically spaced frequencies (scales) s_j, $j = 1, \dots, 10 = F$ are chosen in the range 10 - 600 Hz.
3. For each time interval $\delta t_i = [t_{i1}, \dots, t_{i100}]$ and scale s_j the signal undergoes Morlet wavelet transform:
$$[W_\psi f](s_j, t_{ik}) = \frac{1}{\sqrt{|s_j|}} \int_{-\infty}^{\infty} \overline{\psi\left(\frac{x - t_{ik}}{s_j}\right)} f(x) dx.$$

For n-th electrode in $1, \dots, N$ the (i, j, n)-th element of tensor $\underline{\mathbf{X}}_m$ is given by $[W_\psi f]_{ij}$, where

$$[W_\psi f]_{ij} = \sum_{k=1}^{100} [W_\psi f](s_j, t_{ik}).$$

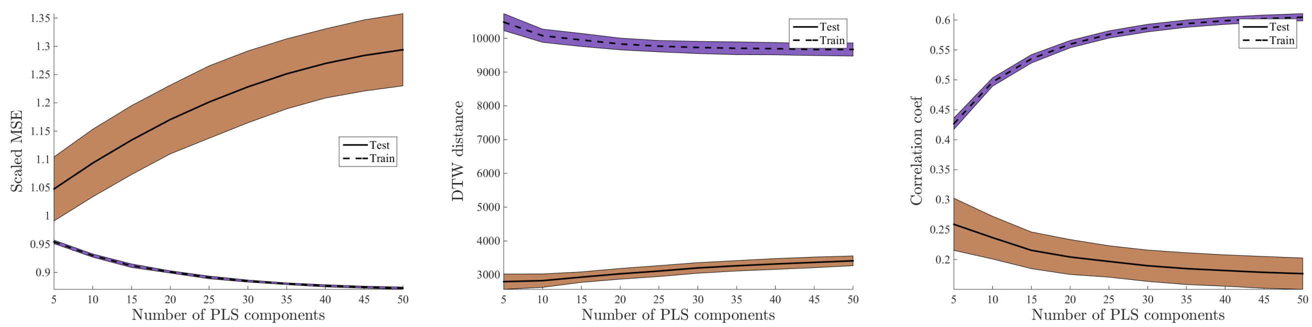
Quality measures include residue-based measures (RMSE, MAE, MAPE), correlation between predictions and the original data and DTW distance, since predictions may approximate the targets with some delay. Another criterion is smoothness of the resultant trajectory.

Since the data is high dimensional and highly correlated, various feature selection techniques are applied to improve the quality of solution. A widely used technique is PLS and its extensions for tensor data [4]. An alternative is recent approach to filtering feature selection by Katrutsa [5].



(a) (b) (c)

Рис. 1: Forecasting quality by the number of PLS component, motion marker "LSHO".



(a) (b) (c)

Рис. 2: Forecasting quality by the number of PLS component, motion marker "LELB".

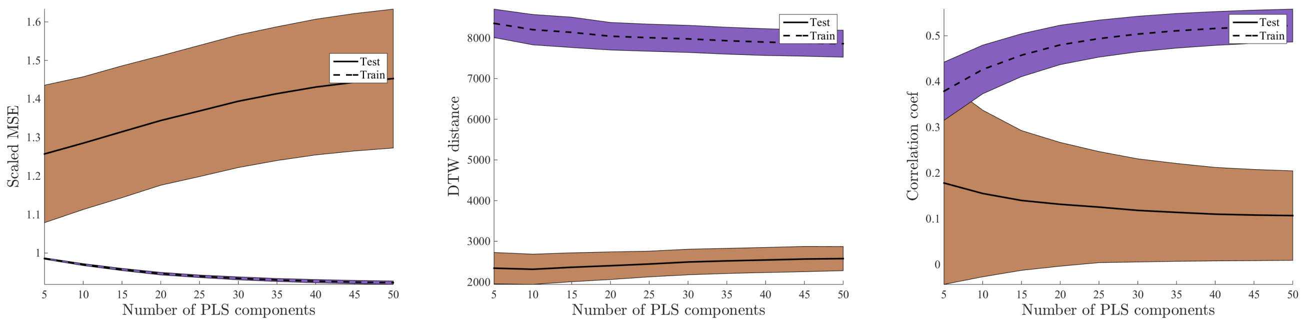
1 Experiments

Resampling with 0.05s Fig. 1–6 display dependance of forecasting quality on the number of PLS components for train and test data. The quality is measured as scaled MSE, DTW distance and correlation coefficient via 5-fold cross-validation. The original time series (0 to 950s) were split into segments with 0.05s step. Train data included segments from 5s to 650s.

Resampling with 10s. Fig. 8 display dependance of forecasting quality on the number of PLS components for train and test data. The quality is measured as scaled MSE, DTW distance and correlation coefficient. 'test' results, demonstrated at subfigures (a)–(c) are obtained via 10-fold cross-validation. Subfigure (d) displays correlation of predictions and the original time series for the hold-out sample. The original time series (0 to 950s) were split into segments with 10s step. Train data included segments from 5s to 650s. Hold-out data included segments from 655s to 950s.

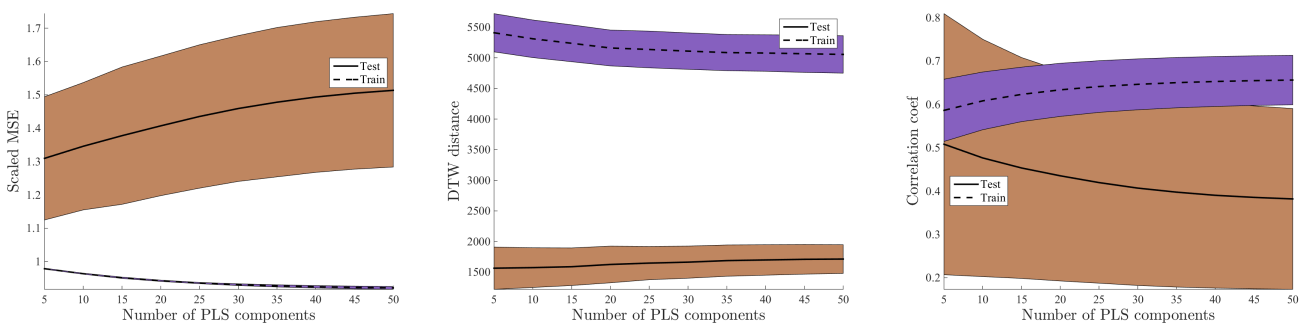
Список литературы

[1] Z.C. Chao, Y. Nagasaka, and N. Fujii. Long-term asynchronous decoding of arm motion using electrocorticographic signals in monkeys. *Frontiers in Neuroengineering*, 3:3, 2010.



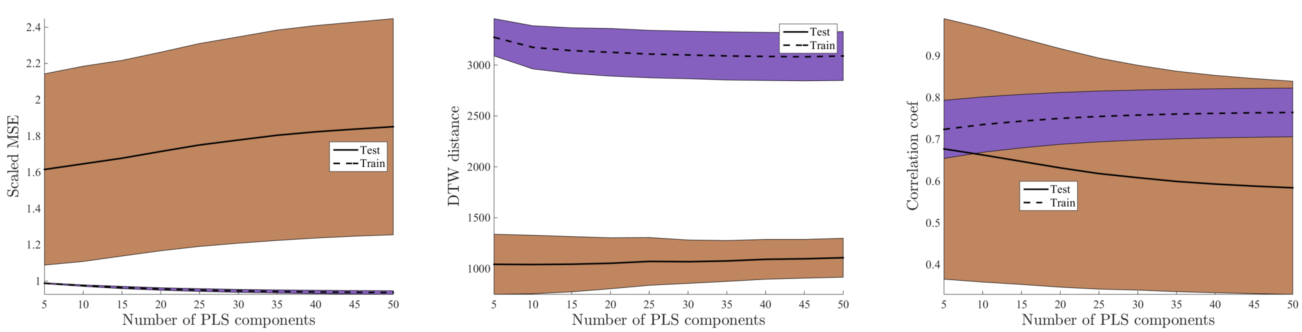
(a) (b) (c)

Рис. 3: Forecasting quality by the number of PLS component, motion marker “RSHO”.



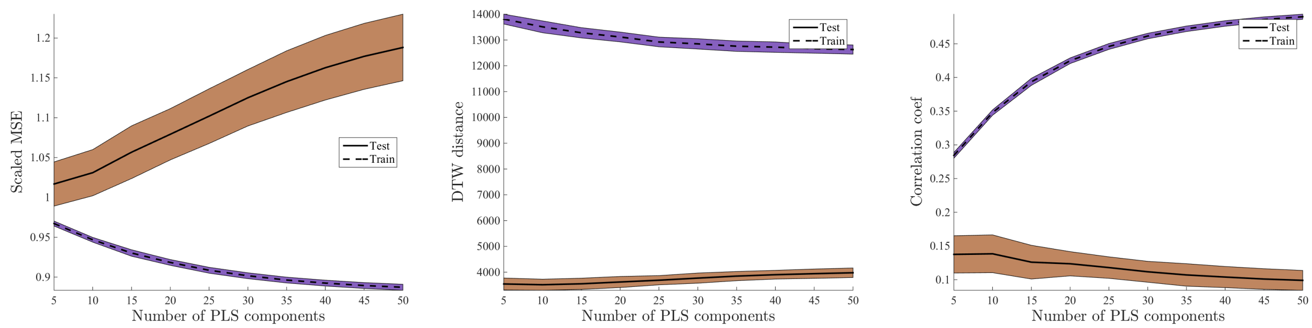
(a) (b) (c)

Рис. 4: Forecasting quality by the number of PLS components, motion marker “RELB”.



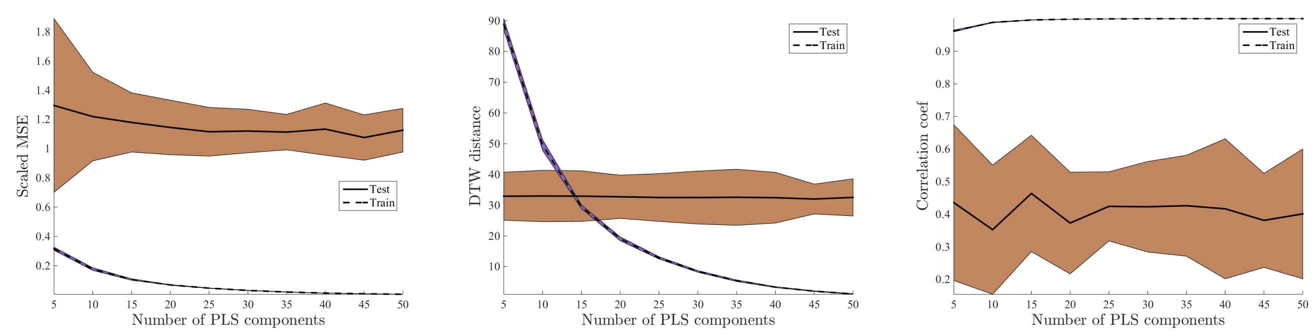
(a) (b) (c)

Рис. 5: Forecasting quality by the number of PLS components, motion marker “RWRI”.



(a) (b) (c)

Рис. 6: Forecasting quality by the number of PLS components, motion marker “RHND”.



(a) (b) (c)

Рис. 7: Forecasting quality by the number of PLS components, 11 motion markers.

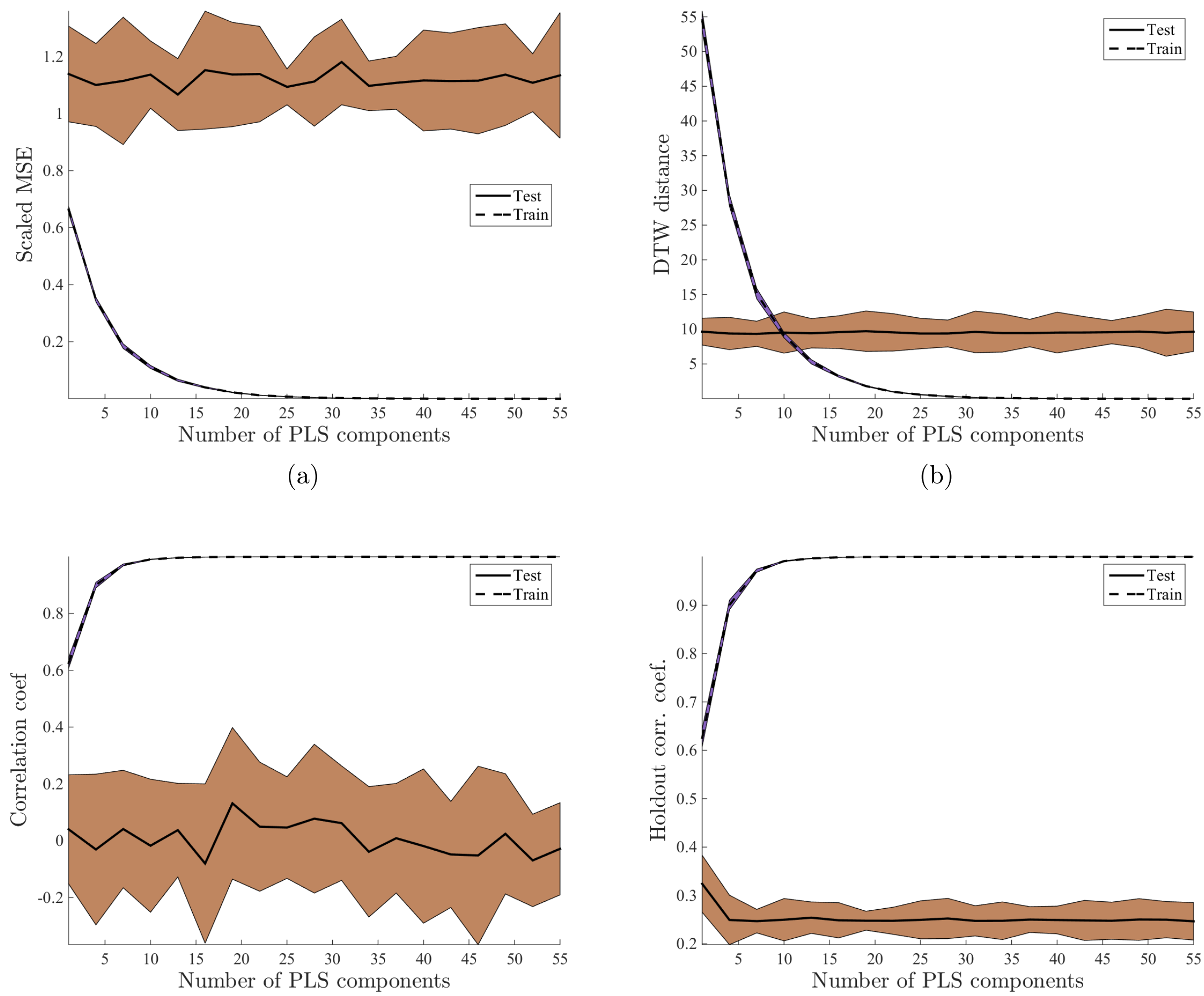


Рис. 8: Forecasting quality by the number of PLS component for right shoulder, right elbow and right wrist markers.

[2] http://neurotycho.org/food-tracking-task.

[3] Group374/Makarchuk2016ECoGSignals/doc/Makarchuk2016ECoGSignals.

[4] Andrey Eliseyev and Tetiana Aksenova. Penalized multi-way partial least squares for smooth trajectory decoding from lectrocorticographic (ecog). *PLoS ONE*, 11(5):e0154878, 2016.

[5] Aleksandr Katrutsa and Vadim Strijov. Comprehensive study of feature selection methods to solve multicollinearity problem according to evaluation criteria. *Expert Systems with Applications*, 2017.