STUDY clustering overview

STEP 1
Build a STUDY

STEP 2
Precompute the data

STEP 3
Precluster the data

STEP 4
Cluster the data

STEP 5
Edit/view the clusters
STUDY clustering overview

STEP 1
Build a STUDY

STEP 2
Precompute the data

STEP 3
Precluster the data

STEP 4
Cluster the data

STEP 5
Edit/view the clusters
Memory options

STUDY options (set these checkboxes if you intend to work with studies):
- If set, keep at most one dataset in memory. This allows processing hundreds of datasets within studies.
- If set, save not one but two files for each dataset (header and data). This allows faster data loading in studies.
- If set, write ICA activations to disk. This speeds up loading ICA components when dealing with studies.

Memory options
- If set, use single precision under Matlab 7.x. This saves RAM but can lead to rare numerical imprecisions.
- If set, use memory mapped array under Matlab 7.x. This may slow down some computation.

ICA options
- If set, precompute ICA activations. This requires more RAM but allows faster plotting of component activations.
- If set, scale ICA component activities to RMS (Root Mean Square) in microvolt (recommended).

Folder options
- If set, when browsing to open a new dataset assume the folder/directory of previous dataset.

Memory options should change when using STUDY vs single dataset.
Build a STUDY

![EEGLAB GUI](image)

- **Import data**
  - Import epoch info
  - Import event info
  - Export

- **Load existing dataset**
  - Save current dataset(s)
  - Save current dataset as
  - Clear dataset(s)

- **Create study**
  - **Using all loaded datasets**
  - **Browse for datasets**

- **Memory and other options**
  - Save history

- Quit
Build a STUDY, cont'd

Create a new STUDY set
STUDY set name:
STUDY set task name:
STUDY set notes:

dataset filename  | browse | subject | session | condition | group | Select by r.v. |
1  | ... | ... | ... | ... | ... | Clear |
2  | ... | ... | ... | ... | ... | Clear |
3  | ... | ... | ... | ... | ... | Clear |
4  | ... | ... | ... | ... | ... | Clear |
5  | ... | ... | ... | ... | ... | Clear |
6  | ... | ... | ... | ... | ... | Clear |
7  | ... | ... | ... | ... | ... | Clear |
8  | ... | ... | ... | ... | ... | Clear |
9  | ... | ... | ... | ... | ... | Clear |
10 | ... | ... | ... | ... | ... | Clear |

Important note: Removed datasets will not be saved before being deleted from EEGLAB memory.

Update dataset info - datasets stored on disk will be overwritten (unset = do not overwrite).
Delete cluster information to allow loading new datasets, set new component.
Save this study to a disk file named: ...

Choose dataset to add to STUDY – pop_study()
## Edit dataset info

- **STUDY set name:**
- **STUDY set task name:**
- **STUDY set notes:**

### Dataset Information Table

<table>
<thead>
<tr>
<th>Dataset Filename</th>
<th>Browse</th>
<th>Subject</th>
<th>Session</th>
<th>Condition</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDY/S01/S01_attend1.pos1.set</td>
<td>...</td>
<td>S01</td>
<td>1</td>
<td>TargetAttL</td>
<td>normals</td>
</tr>
<tr>
<td>IDY/S01/S01_attend1.pos5.set</td>
<td>...</td>
<td>S01</td>
<td>1</td>
<td>NonTargetAttL</td>
<td>normals</td>
</tr>
<tr>
<td>IDY/S01/S01_attend5.pos1.set</td>
<td>...</td>
<td>S01</td>
<td>1</td>
<td>TargetAttR</td>
<td>normals</td>
</tr>
<tr>
<td>IDY/S01/S01_attend5.pos5.set</td>
<td>...</td>
<td>S01</td>
<td>1</td>
<td>NonTargetAttR</td>
<td>normals</td>
</tr>
<tr>
<td>IDY/S02/S02_attend1.pos1.set</td>
<td>...</td>
<td>S02</td>
<td>1</td>
<td>TargetAttL</td>
<td>normals</td>
</tr>
<tr>
<td>IDY/S02/S02_attend1.pos5.set</td>
<td>...</td>
<td>S02</td>
<td>1</td>
<td>NonTargetAttL</td>
<td>normals</td>
</tr>
<tr>
<td>IDY/S02/S02_attend5.pos1.set</td>
<td>...</td>
<td>S02</td>
<td>1</td>
<td>TargetAttR</td>
<td>normals</td>
</tr>
<tr>
<td>IDY/S02/S02_attend5.pos5.set</td>
<td>...</td>
<td>S02</td>
<td>1</td>
<td>NonTargetAttR</td>
<td>normals</td>
</tr>
<tr>
<td>IDY/S03/S03_attend1.pos1.set</td>
<td>...</td>
<td>S03</td>
<td>1</td>
<td>TargetAttL</td>
<td>normals</td>
</tr>
<tr>
<td>IDY/S03/S03_attend1.pos5.set</td>
<td>...</td>
<td>S03</td>
<td>1</td>
<td>NonTargetAttL</td>
<td>normals</td>
</tr>
</tbody>
</table>

**Important note:** Removed datasets will not be saved before being deleted from EEGLAB memory.

- **Update dataset info:** datasets stored on disk will be overwritten (unset = keep study info separate).
- **Delete cluster information:** (to allow loading new datasets, set new components for clustering, etc.)
- **Re-save STUDY:** Uncheck and use menu File > Save study as to save under a new filename.

---

![Image of EDIT DATASET window](image.png)
ICs to cluster
STUDY structure

STUDY =

    name: 'Synonyms'
    task: 'Word Recognition'
    notes: ''
    filename: 'workshop.study'
    cluster: [1x1 struct]
    history: [1x6654 char]
    datasetinfo: [1x10 struct]
    filepath: '/data/STUDY'
    subject: {'S02' 'S05' 'S07' 'S08' 'S10'}
    group: {}
    session: []
    condition: {'non-synonyms' 'synonyms'}
    setind: [2x5 double]
    etc: [1x1 struct]
    preclust: [1x1 struct]
    saved: 'no'
    changrp: []
Subject info in STUDY structure

```matlab
>> STUDY.datasetinfo

ans =

1x10 struct array with fields:
    filepath
    filename
    subject
    session
    condition
    group
    comps
    index
```

Gives information for each subject
STUDY clustering overview

STEP 1
Build a STUDY

STEP 2
Precompute the data

STEP 3
Precluster the data

STEP 4
Cluster the data

STEP 5
Edit/view the clusters
Precompute data measures
Precompute data measures

TIP: Compute all measures so you can test different combinations for clustering

Recommend: 'alpha', .01 (time-consuming)
STUDY clustering overview

STEP 1
Build a STUDY

STEP 2
Precompute the data

STEP 3
Precluster the data

STEP 4
Cluster the data

STEP 5
Edit/view the clusters
Precluster the data
Precluster the data

![Image of EEG data processing interface](image)

**Build pre-clustering matrix for STUDY 'Attention'**

Select the cluster to refine during sub-clustering (any existing sub-hierarchy will be overwritten).

- **ParentCluster 1 (181 ICs)**

(not only measures that have been precomputed may be used)

<table>
<thead>
<tr>
<th>Load</th>
<th>Dims</th>
<th>Norm</th>
<th>Rel Wt</th>
<th>Freq range [Hz]</th>
<th>Time range [ms]</th>
<th>Use channel values</th>
<th>Freq range [Hz]</th>
<th>Time range [ms]</th>
</tr>
</thead>
<tbody>
<tr>
<td>spectra</td>
<td>10</td>
<td>✓</td>
<td>✓</td>
<td>3.25</td>
<td>0.600</td>
<td>✓</td>
<td>3.45</td>
<td>2.30</td>
</tr>
<tr>
<td>ERPs</td>
<td>10</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dipoles</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>scalp maps</td>
<td>10</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERSPs</td>
<td>20</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITCs</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final dimensions</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Save STUDY to file**: `/home/julie/WorkshopSD2007/STUDY/attention.study`
STUDY clustering overview

STEP 1
Build a STUDY

STEP 2
Precompute the data

STEP 3
Precluster the data

STEP 4
Cluster the data

STEP 5
Edit/view the clusters
Cluster components

Performing clustering on cluster 'ParentCluster 1'

Clustering algorithm:
Number of clusters to compute:
- Separate outliers: 3 std. dev. from any cluster center
- Save STUDY set to disk: /home/julie/WorkshopSD2007/STUDY/attempt...
Cluster info in 'STUDY'

```matlab
>> STUDY.cluster
ans =
    1x26 struct array with fields:
        name
        parent
        child
        comps
        sets
        algorithm
        centroid
        preclust
        dipole
        topo
        topox
        topoy
        topoall
        topopol

>> STUDY.cluster(2)
ans =
    name: 'Cls 2'
    parent: {'ParentCluster 1'}
    child: []
    comps: [9 10 21 18 26 20 27]
    sets: [2x7 double]
    algorithm: {'Kmeans' [12]}
    centroid: []
    preclust: [1x1 struct]
```

Gives information for each cluster (1st cluster is always the parent cluster)

More on cluster info in later tutorials
STUDY clustering overview

STEP 1
Build a STUDY

STEP 2
Precompute the data

STEP 3
Precluster the data

STEP 4
Cluster the data

STEP 5
Edit/view the clusters
View and edit clusters

[Image of EEGLAB v6.0b interface with options: Edit study info, Precompute channel measures, Plot channel measures, Precompute component measures, Build preclustering array, Cluster components, Edit/plot clusters]
Plot cluster data

Study ‘Attention’: 181 of 181 components clustered

Select cluster to plot
- All cluster centroids
- ParentCluster 1 (181 ICs)
- outlier 2 (1 ICs)
- Cls 3 (5 ICs)

Select component(s) to plot
- outlier 2′ comp. 1 (S12 IC12)
- Cls 3 comp. 1 (S01 IC11)
- Cls 3 comp. 2 (S05 IC11)
- Cls 3 comp. 3 (S06 IC15)

- Plot scalp maps
- Plot dipoles
- Plot ERPs
- Plot spectra
- Plot ERSPs
- Plot IT-Cs

Create new cluster
Rename selected cluster
Merge clusters

Save STUDY set to disk
/home/julie

Plot mean scalp maps for easy reference

EEGLAB Workshop VII, Apr. 20-22, 2009, Bloomington, IN: Julie Onton – STUDY GUI 21
Plot cluster data

Choose which cluster

Choose which components

Study ‘Attention’: 181 of 181 components clustered

Select cluster to plot

Cls 6 (3 ICs)
Cls 7 (10 ICs)
Cls 8 (5 ICs)
Cls 9 (12 ICs)

Select component(s) to plot

All components
S01 TC6
S05 IC9
S06 IC12

Plot scalp maps
Plot dipoles
Plot ERPs
Plot spectra
Plot ERSPs
Plot ITCs
Plot cluster properties

Create new cluster
Rename selected cluster
Merge clusters

Cancel
Help

EEGLAB Workshop VII, Apr. 20-22, 2009, Bloomington, IN: Julie Onton – STUDY GUI 22