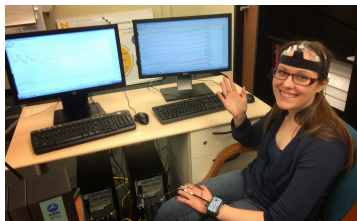# Regression Spline Mixed Models for Analyzing EEG Data and Event-Related Potentials

Karen E. Nielsen[1]   Rich Gonzalez[1]

[1]University of Michigan
Department of Statistics & BioSocial Methods Collaborative
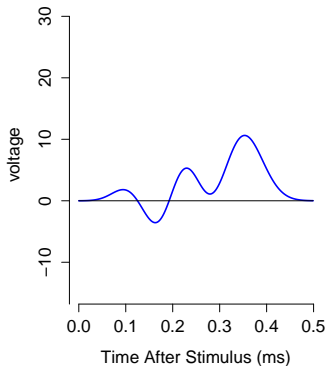
UseR! 2015

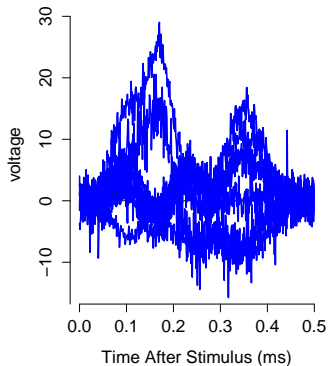# Event-Related Potential (ERP) Data

Figure: Simulated ERP Data

Figure: Actual ERP Data - Messy!

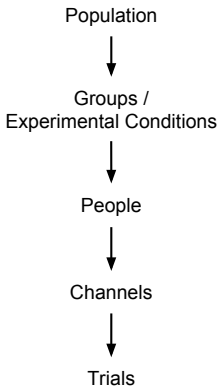Simulation using *eegkit* package in R.

Theory

Population

↓

Groups /
Experimental Conditions

↓

People

↓

Channels

↓

Trials

Figure: Data is organized in a hierarchy

Practice

Population

↓

Groups /
Experimental Conditions

↓

People

↓

Channels

↓

Trials
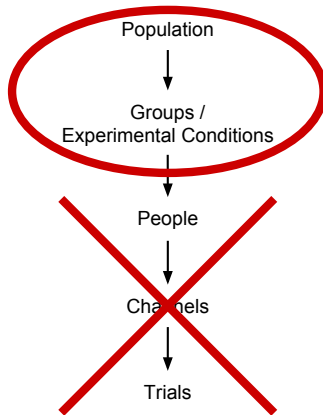
Figure: Current methods ignore lower levels

Theory
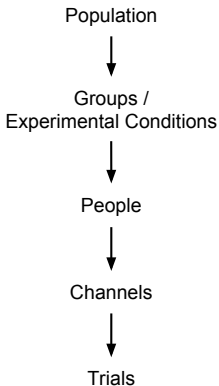
Population

↓

Groups /
Experimental Conditions

↓

People

↓

Channels

↓

Trials

Figure: Data is organized in a hierarchy

Practice

Population

↓

Groups /
Experimental Conditions

↓

People

↓

Channels

↓

Trials
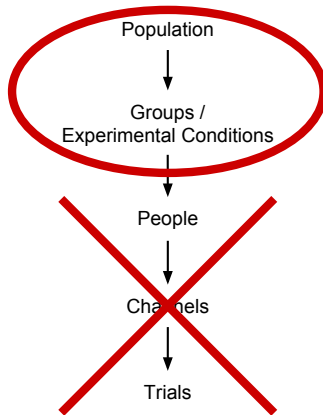
Figure: Current methods ignore lower levels
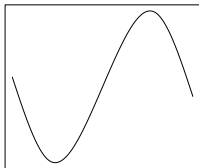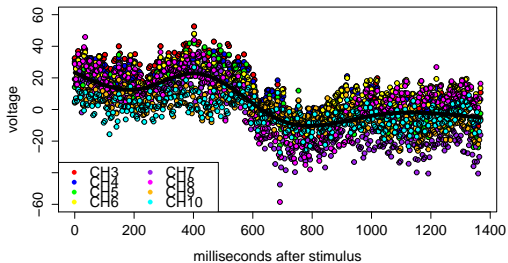
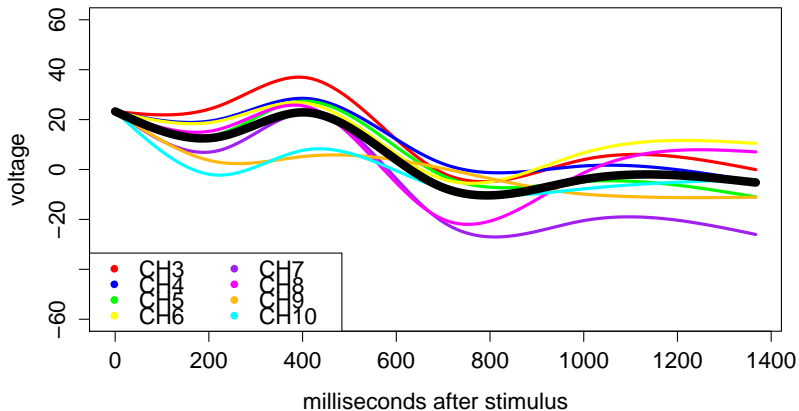Figure: A Cubic Spline

Figure: Actual ERP Data

Figure: Regression Splines for each channel (colors) and overall fit (solid)
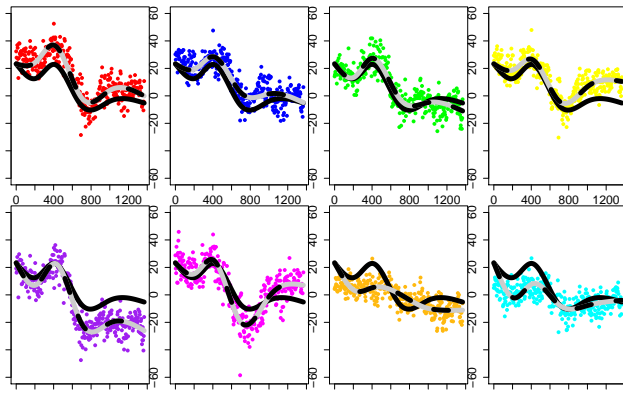
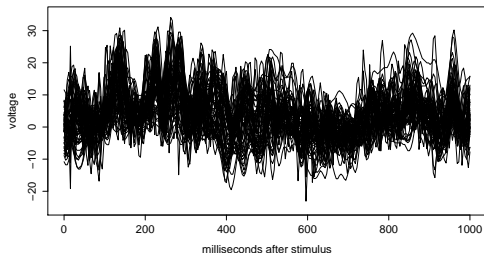Figure: Regression Splines for each channel (dashed) and overall fit (solid)

Figure: ERP data with time- and amplitude-varying clusters
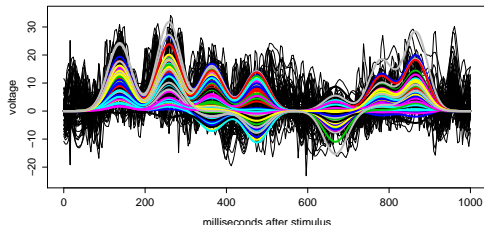

Figure: 7 normal kernels are fit to the channels, with fixed time locations.

```
reshape(data,

    varying=c("CH2", ..., "CH64"),

    v.names="voltage",

    timevar="channel",

    times=c("CH2", ..., "CH64"),

    direction="long"))
```
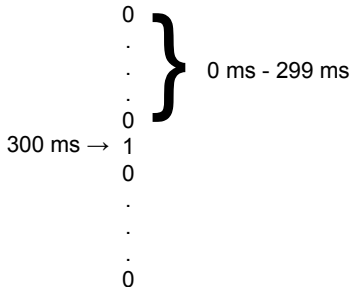
subj	trial	channel	time	voltage
1	1	1	1	7.749
1	1	1	2	5.231
1	1	1	3	-12.071
1	1	1	4	2.810
1	1	1	5	14.702
1	1	2	1	-1.472
1	1	2	2	-1.961
1	1	2	3	4.795
1	1	2	4	0.493
1	1	2	5	-4.327
1	2	1	1	-9.247
1	2	1	2	8.576
1	2	1	3	-6.839
1	2	1	4	12.489
1	2	1	5	11.944
1	2	2	1	-10.081
1	2	2	2	1.587
1	2	2	3	-8.330
1	2	2	4	-2.444
1	2	2	5	7.906
2	1	1	1	-1.764
2	1	1	2	6.341
2	1	1	3	20.570

Figure: Data in long format

Basis Set:

R Users have a greater P300 peak amplitude.

$\Longrightarrow$

0
.
.
.
0 } 0 ms - 299 ms
300 ms → 1
0
.
.
.
0
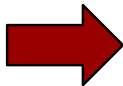
R stuff

lmer(voltage

    basis1 + ... + basisN +

    (basis1+...+ basisN | group),

    data)

```
> model1
Nonlinear mixed-effects model fit by maximum likelihood
  Model: (voltage) ~ mixn1(times, b1, m1)
  Data: simdata
  Log-likelihood: -12237.35
  Fixed: b1 + m1 ~ 1
         b1        m1
  1.6751424 0.3506333

Random effects:
 Formula: b1 ~ 1 | group
            b1  Residual
StdDev: 9.213658 5.148868

Number of Observations: 4000
Number of Groups: 2
```
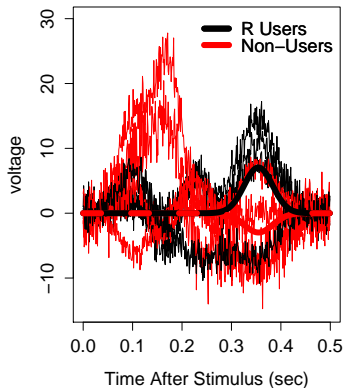
Figure: Example output

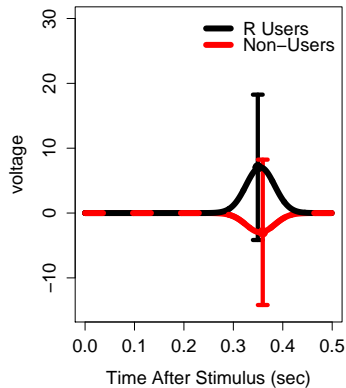Figure: R users have a greater P300 peak amplitude...

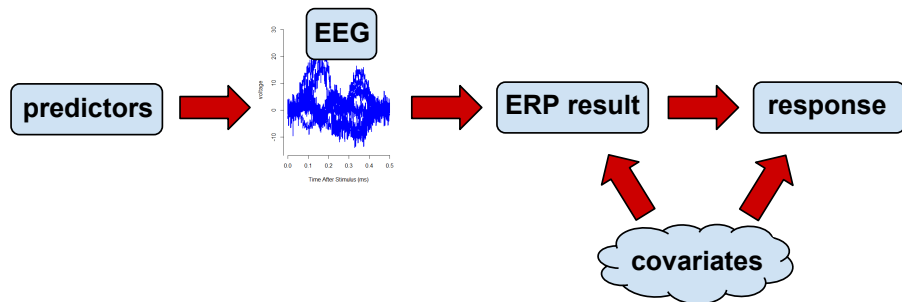Figure: But the difference is not significant.

Benefits of R Workflow

- Open source (not pre-packaged, black-box)
- Stats-oriented
- Area of opportunity

Future Work

- Go beyond model creation
- Include other covariates
- Spatial correlations for channels
- Standard processing stream for ERPs

Thanks!

Contact: karenen@umich.edu

Web: http://sites.lsa.umich.edu/karenen/useR-talk/