Delay Differential Analysis of Human Schizophrenia MMN Data

Erin Brown & Debha Amatya

14 August 2014

Delay Differential Analysis of MMN Data

- Mismatch Negativity
- The Data Set
- Review of Delay Differential Analysis
- Methods
- Preliminary Results
- Future Work

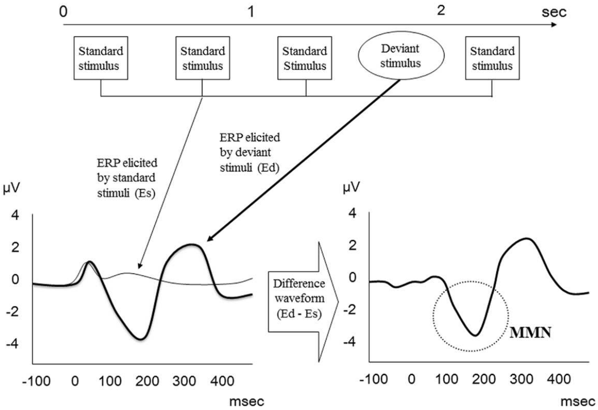
Schizophrenia: the need for biomarkers

- Schizophrenia is one of the most disabling and costly disease worldwide (top 10)
- The current standard of care suffers from a lack of objective and quantitative biomarkers
- Diagnoses are obtained through patient questionnaires and behavioral observation
- Mismatch negativity (MMN) is an EEG biomarker that shows promise in classifying schizophrenia

Mismatch Negativity

- The ERP response to deviant stimuli in a series of standard stimuli (ssdssdssssdsss)
- Auditory MMN
  - Use stimuli that differ in duration or pitch
  - Observe subject's change detection response
    - Subject does not have to be attentive
  - Two decades of research shows that schizophrenia patients consistently have reduced amplitude ERP to deviant tones
    - Correlations between amplitude reduction and symptom severity
- Other benefits of MMN
  - Well-tolerated
  - Non-invasive

Mismatch Negativity



Nagai, Tatsuya, Mariko Tada, Kenji Kirihara, Tsuyoshi Araki, Seiichiro Jinde, and Kiyoto Kasai. "Mismatch Negativity as a "Translatable" Brain Marker Toward Early Intervention for Psychosis: A Review." Frontiers in Psychiatry 4 (2013).

The Data Set

- 1630 human subjects from 5 sites
  - 753 controls
  - 877 schizophrenia patients (no other Axis 1 psych diagnoses)
- EEG MMN data
  - Auditory "oddball duration" paradigm
    - 50 ms standard tone and 100 ms deviant tones
    - 450 ms between tones
  - 2 channel recordings

Delay Differential Analysis–Review

What is delay differential analysis (DDA)?

Tool for non-linear dynamical classification of time series

What is the basic process?

1. Roughly fit a delay differential equation (DDE) to a time series
2. Use coefficients and root-mean-square error as distinguishing features for classification
3. Try different DDEs to come up with best classification

Why DDA?

- DDA is reliable even for short time series
- computations are fast
- DDA is noise insensitive
- no pre-processing of data
- small feature space compared to traditional techniques
- 3-4 features: no overfitting

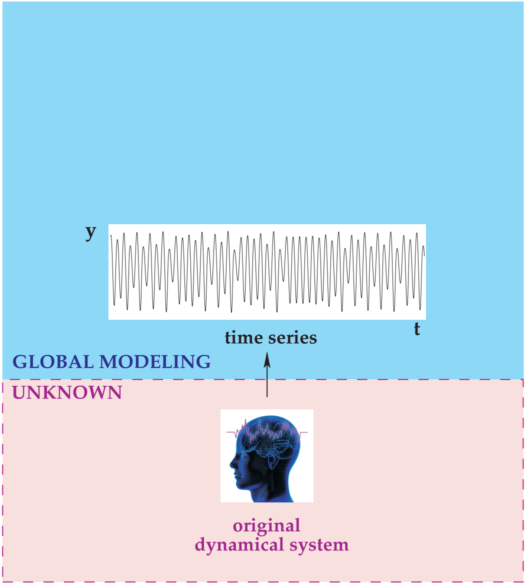
Dynamical System Analysis

Dynamical System Analysis

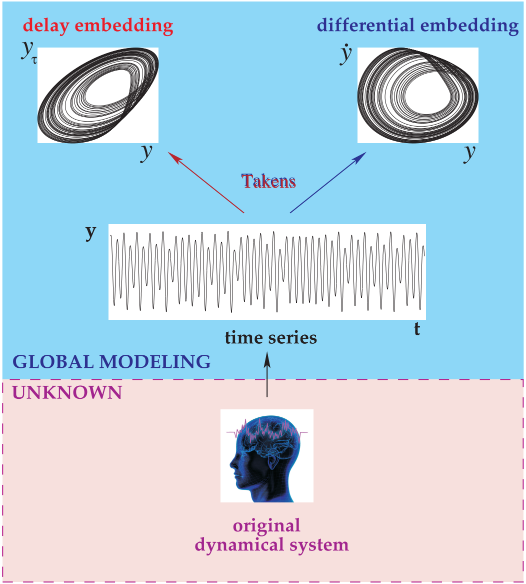
UNKNOWN

original
dynamical system

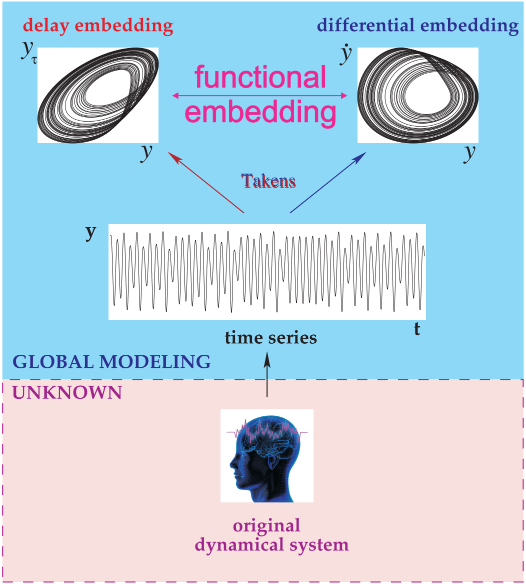
Dynamical System Analysis



Dynamical System Analysis

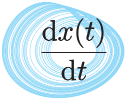


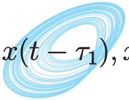
Dynamical System Analysis



Delay differential analysis is done in the time domain,

not in the spectral domain!



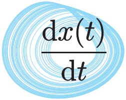


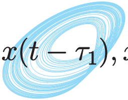
$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f(a, x(t-\tau_1), x(t-\tau_2), \dots)$$

Delay Differential Analysis (DDA)

Delay differential analysis is done in the time domain,

not in the spectral domain!





$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f(a, x(t-\tau_1), x(t-\tau_2), \ldots)$$

linear terms: dominant time scales (frequencies)

non-linear terms: frequency/phase couplings, feedback

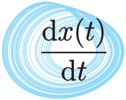
DDE has

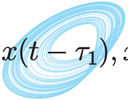
- n delays: $\tau_1, \tau_2, \ldots \tau_n$
- l terms with coefficients $a_1, a_2, \ldots, a_l$
- degree m of nonlinearity

$\Rightarrow$ coefficients a_k and model error ρ as features to identify

dynamical differences in data

General Model





$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f(a, x(t-\tau_1), x(t-\tau_2), \ldots)$$

$$\begin{aligned}
\dot{x} \quad = \quad & a_1 x_{\tau_1} + a_2 x_{\tau_2} + a_3 x_{\tau_1}^2 + a_4 x_{\tau_1} x_{\tau_2} \\
& + a_5 x_{\tau_2}^2 + a_6 x_{\tau_1}^3 + a_7 x_{\tau_1}^2 x_{\tau_2} + a_8 x_{\tau_1} x_{\tau_2}^2 \\
& + a_9 x_{\tau_2}^3 + a_{10} x_{\tau_1}^4 + a_{11} x_{\tau_1}^3 x_{\tau_2} \\
& + a_{12} x_{\tau_1}^2 x_{\tau_2}^2 + a_{13} x_{\tau_1} x_{\tau_2}^3 + a_{14} x_{\tau_2}^4
\end{aligned}$$

$$x_{\tau_j} = x(t - \tau_j)$$

Structure (Model) Selection

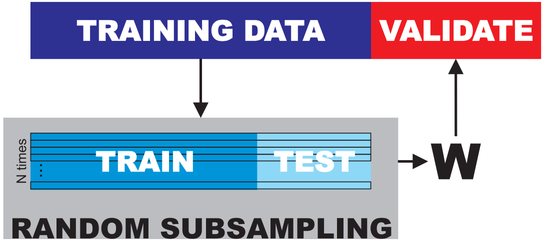
Select **one** fixed model with a fixed set of delays
(e.g. $\dot{x} = a_1 x_{\tau_1} + a_2 x_{\tau_2}^2$)

supervised or unsupervised

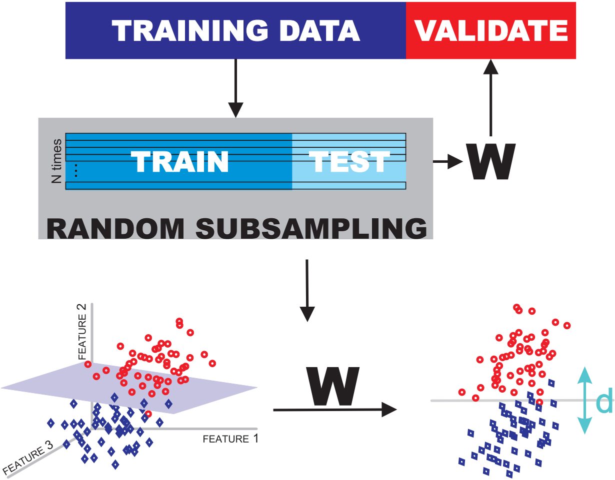
classification: best separation between classes of data

lowest error DDE model from data — dynamical deviation from that stage

Repeated Random Subsampling

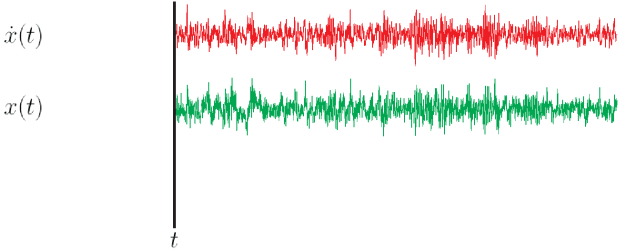


Repeated Random Subsampling



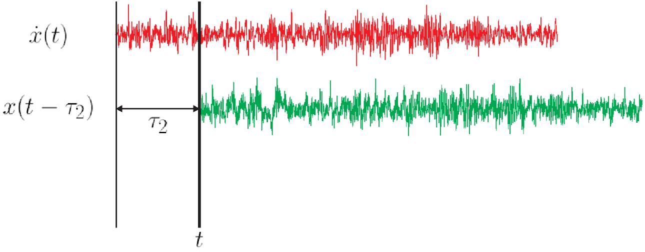
Delay Differential Equation

$$\dot{x}(t) = a_1\, x(t-\tau_2) + a_2\, x(t-\tau_1)^2$$



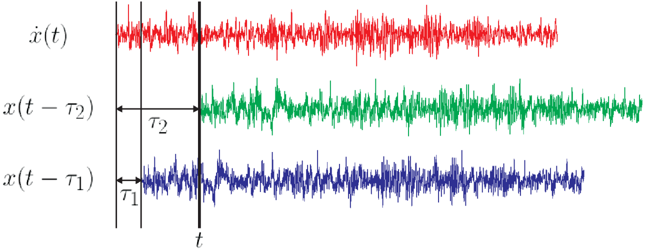
Delay Differential Equation

$$\dot{x}(t) = a_1\, x(t-\tau_2) + a_2\, x(t-\tau_1)^2$$



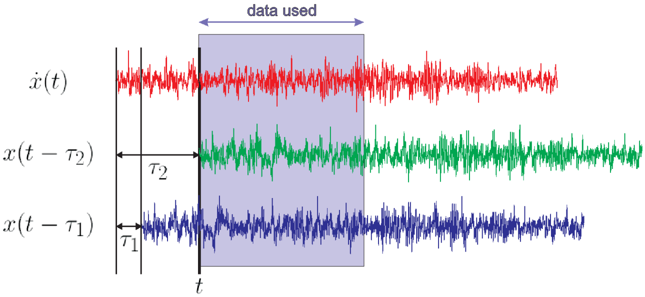
Delay Differential Equation

$$\dot{x}(t) = a_1\, x(t-\tau_2) + a_2\, x(t-\tau_1)^2$$



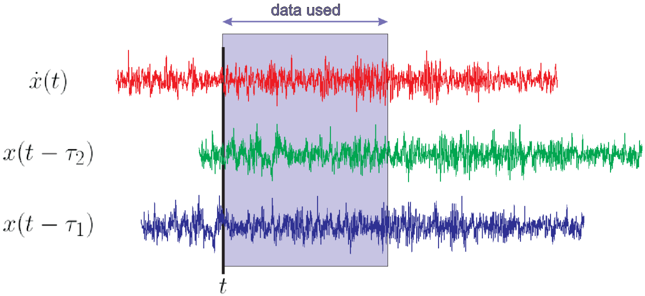
Delay Differential Equation

$$\dot{x}(t) = a_1\, x(t-\tau_2) + a_2\, x(t-\tau_1)^2$$



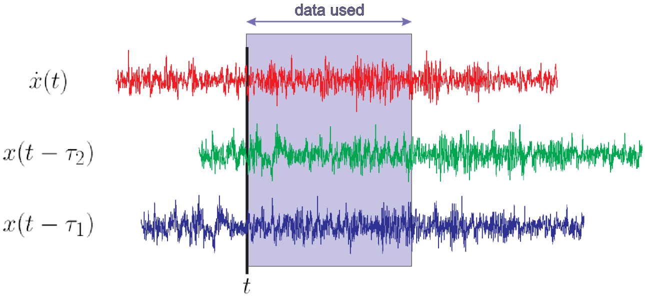
Delay Differential Equation

$$\dot{x}(t) = a_1\, x(t - \tau_2) + a_2\, x(t - \tau_1)^2$$



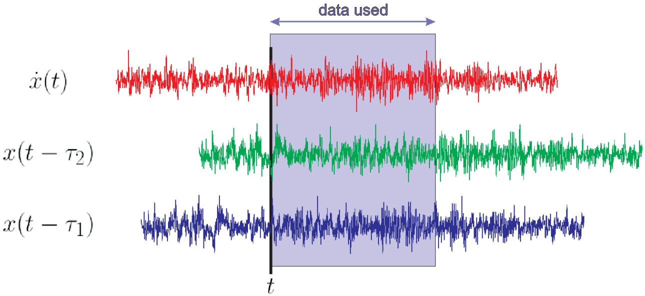
Delay Differential Equation

$$\dot{x}(t) = a_1\, x(t - \tau_2) + a_2\, x(t - \tau_1)^2$$



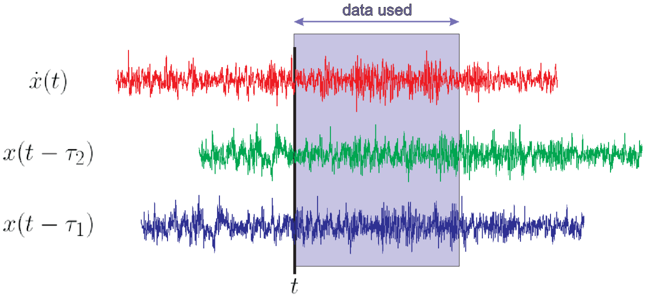
Delay Differential Equation

$$\dot{x}(t) = a_1\, x(t-\tau_2) + a_2\, x(t-\tau_1)^2$$



Delay Differential Equation

$$\dot{x}(t) = a_1\, x(t - \tau_2) + a_2\, x(t - \tau_1)^2$$



Cross-Validation Processes for Windowing across Trials

Three Kinds:

1. Single-Matrix Time Independent (SMTI) cross-validation
2. Multi-Matrix Time Dependent (MMTD) cross-validation
3. Single-Matrix Time Dependent (SMTD) cross-validation

Each kind of cross-validation is characterized by the **length of the CV feature space** and the **effective number of trials**.

Single-Matrix Time Independent CV

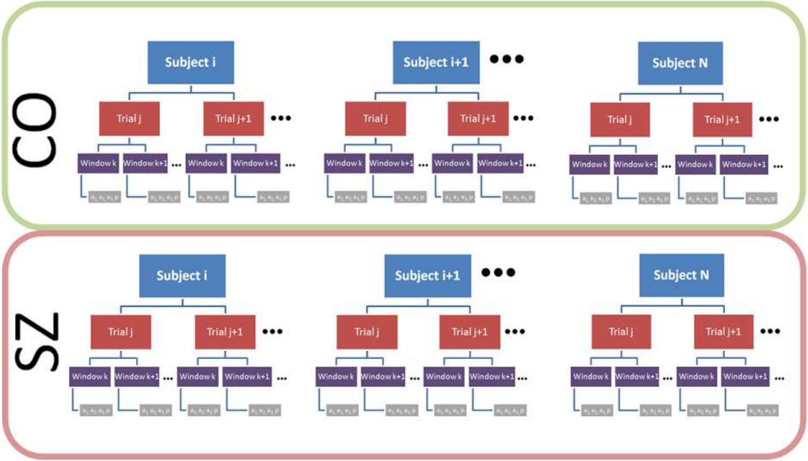
Idea: Treat each window as a separate trial.

$$\text{effective number of trials} = \text{NTrials} \times \text{WN}$$
$$\text{length of CV feature space} = \text{NTerms} + 1$$

Best suited for:

picking out time-invariant differences in dynamics between classes.

Cross-Validation Feature Space



Multi-Matrix Time Dependent CV

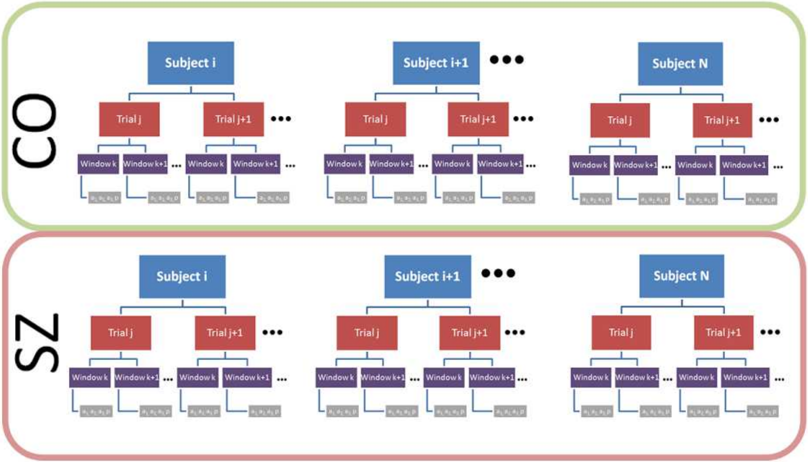
Idea: A separate feature matrix is formed for each window number. Cross-validation is performed on each feature matrix individually.

$$\text{effective number of trials} = \text{NTrials}$$
$$\text{length of CV feature space} = \text{NTerms} + 1$$

Best suited for:

identifying regions of particularly pronounced dynamical difference between conditions.

Cross-Validation Feature Space



Single-Matrix Time Dependent CV

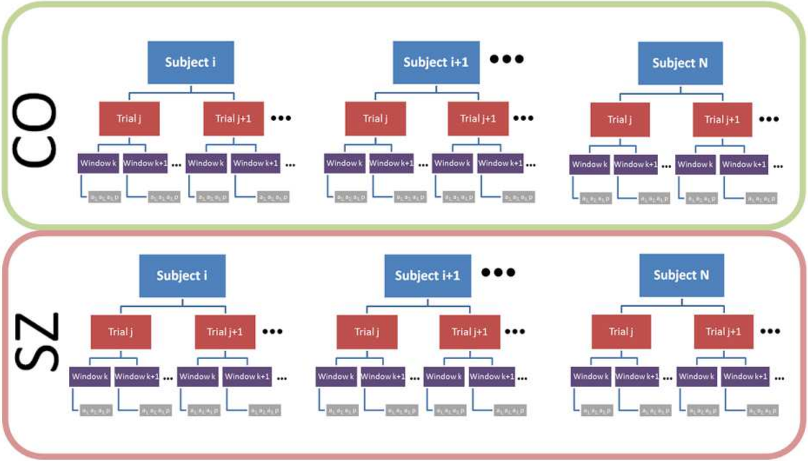
Idea: Concatenate features from each window to create extended cross-validation feature space

$$\text{effective number of trials} = \text{NTrials}$$
$$\text{length of CV feature space} = (\text{NTerms} + 1) \times \text{WN}$$

Best suited for:

classifying between conditions with time-dependent but differing dynamics

Cross-Validation Feature Space



Results–PRELIMINARY

CV Type	Model	Tau 1	Tau 2	A' Ind. Trials
SMTD	4	26	19	0.6890
SMTI	4	26	19	0.6745
MMTD 1	2	19	26	0.6770
MMTD 2	4	26	19	0.6824
MMTD 3	4	25	18	0.6693

- Only **10** randomly chosen subjects for each class
- Segments following deviant tones

Continuing Work

- Classification between conditions of time segments following standard tones, deviant tones
- Classification between segments following standard tones and segments following deviant tones within subjects
- Explore windowing options
- Use longer data segments containing multiple tones
- How best can DDA be implemented specifically for MMN?
- Expand analysis to whole data set

Acknowledgements

Claudia Lainscsek

Greg Light

Terry Sejnowski

CNL

HHMI Capstone Program