Deconvolutional Time Series Regression: A Technique for Modeling Temporally Diffuse Effects

Cory Shain and William Schuler

shain.3@osu.edu | Ohio State University | schuler.77@osu.edu

Problem

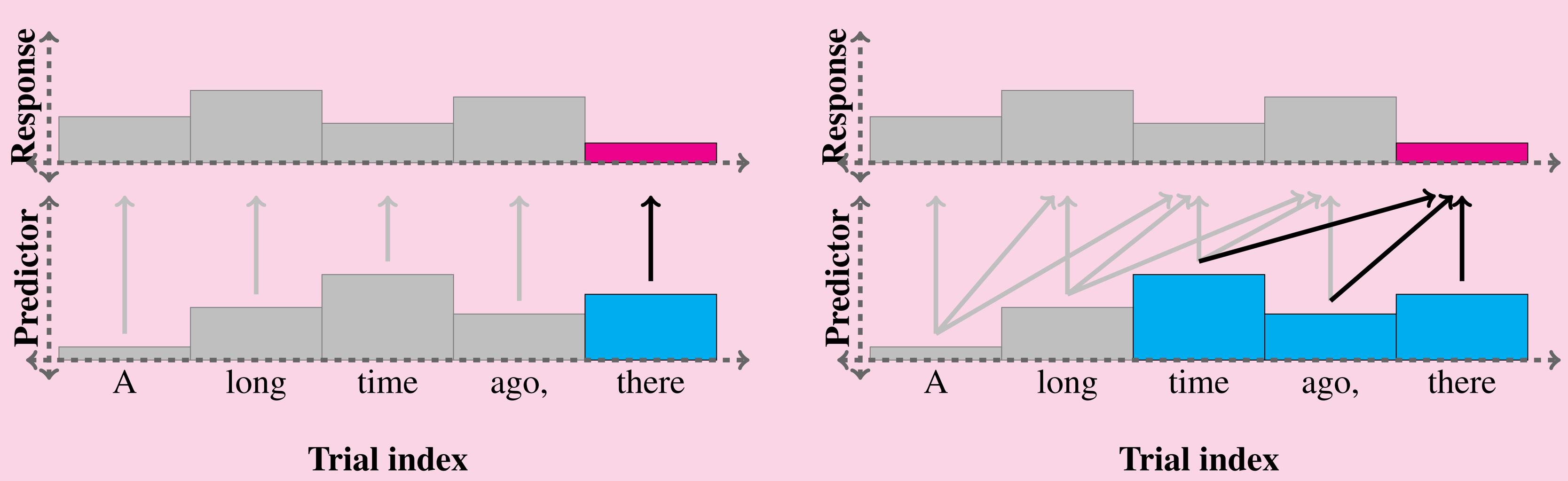
Temporal diffusion of effects may be present in data generated by human subjects. While deconvolutional models directly estimate temporal diffusion, major deconvolutional frameworks are discrete-time and difficult to apply to variably-long naturalistic stimuli without distortion and/or sparsity.

- Linear models with “spillover” (Erlich & Rayner 1983)
- Finite impulse response models (Dayal et al. 1996)
- Vector autoregression (Sims 1980)

Failure to control temporal diffusion presents a serious risk of obtaining a misleading model.

Case in point: Effects of dependency locality and constituent wrap-up reported in (Shain et al. 2016) vanish when one baseline predictor (PCFG surprisal) is spilled over one position.

Linear time series models

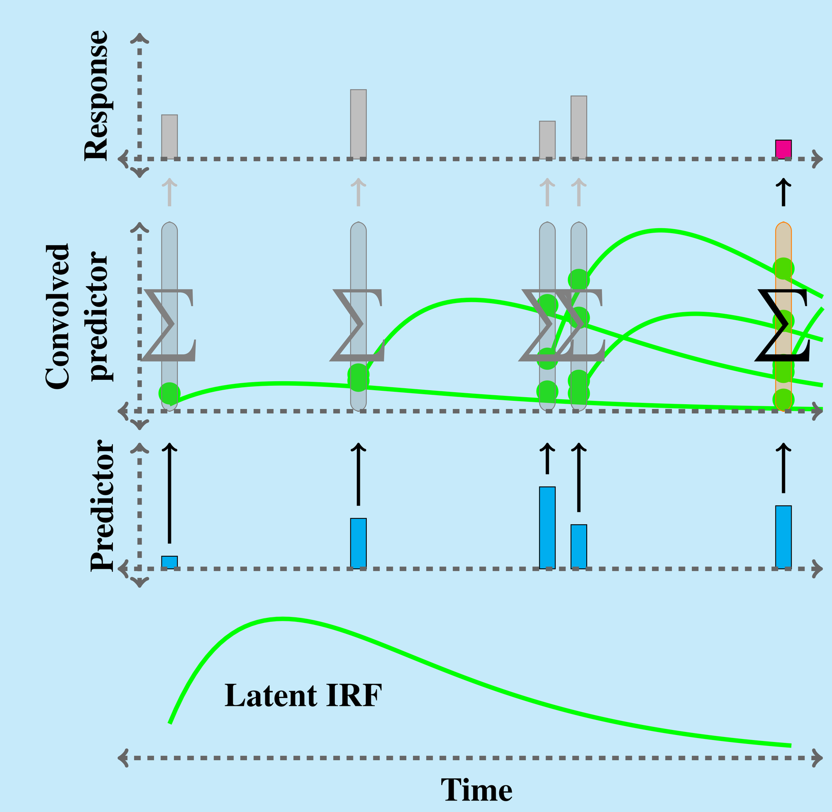


(a) **No spillover.** Responses are independent of preceding stimuli.

(b) **Two spillover positions.** Distance between stimuli is ignored.

Proposal

Deconvolutional time series regression (DTSR), a continuous-time mixed-effects regression technique for discovering parametric impulse response functions (IRFs) in arbitrary time series. DTSR jointly fits (1) IRFs with which to convolve predictors and (2) a linear model of the response on the convolved predictors.

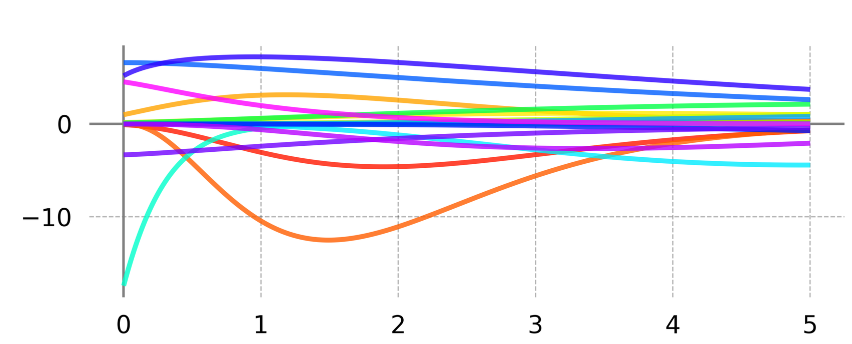


DTSR model

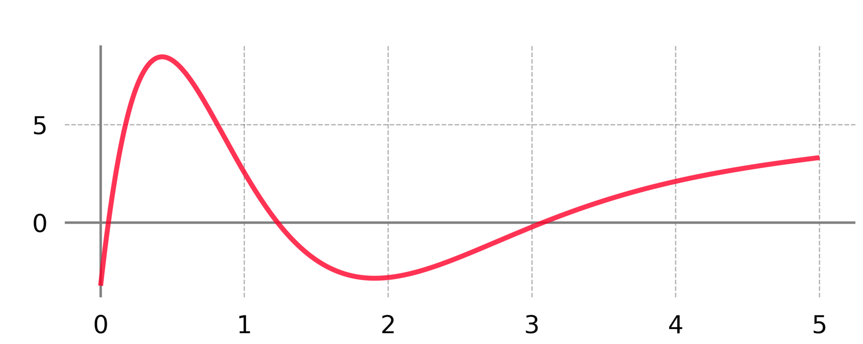
Benefits:

- Produces high-resolution continuous estimates of IRFs
- Autodiff, no need to derive estimators
- $O(1)$ model complexity on num. timesteps
- Supports:
  - Variably-spaced data
  - Unsynchronized data
  - Mixed-effects models
  - Arbitrary parametric IRF kernels
  - Non-parametric IRFs through spline kernels
  - Composition of IRF kernels
  - Variational Bayesian inference
- Documented open-source Python package written in Tensorflow (Abadi et al. 2015) and Edward (Tran et al. 2016): `https://github.com/coryshain/dtsr`

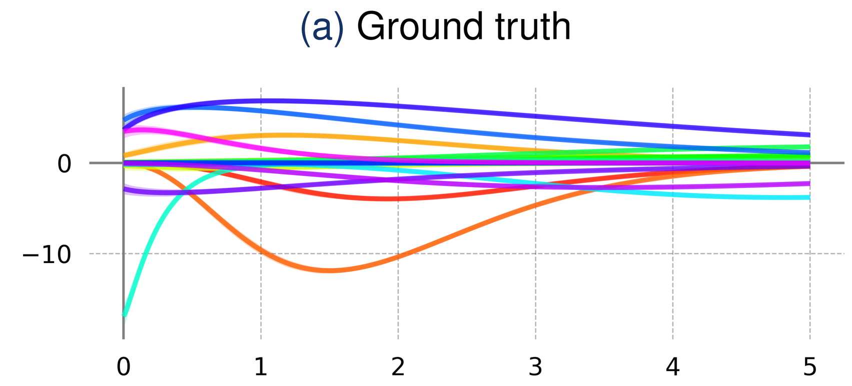
Synthetic Evaluation



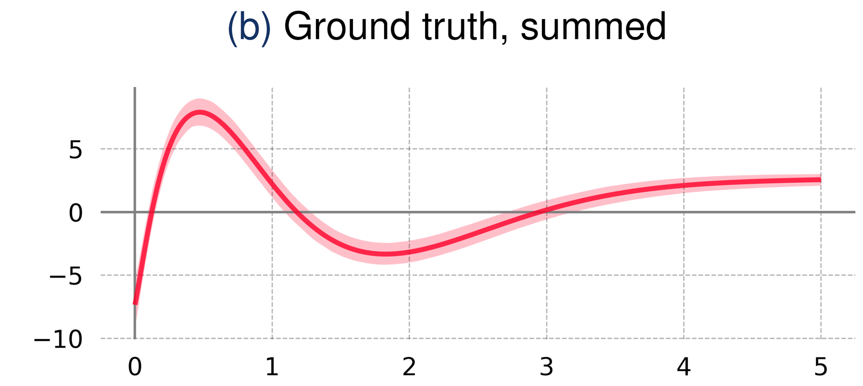
(a) Ground truth



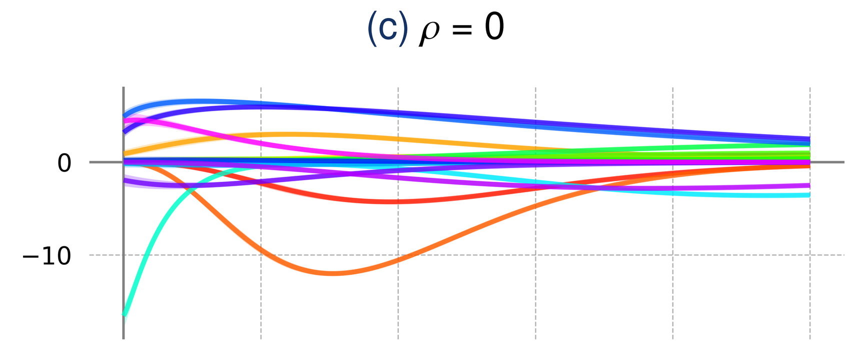
(b) Ground truth, summed



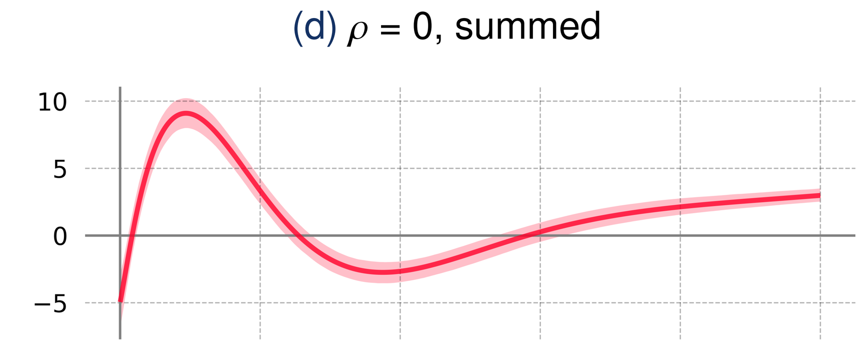
(c) $\rho = 0$



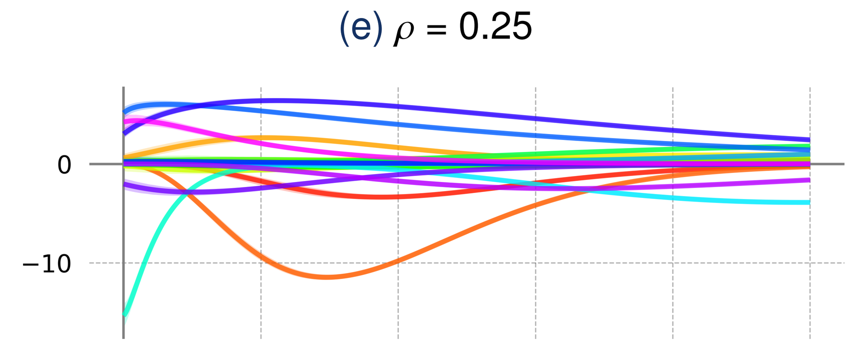
(d) $\rho = 0$, summed



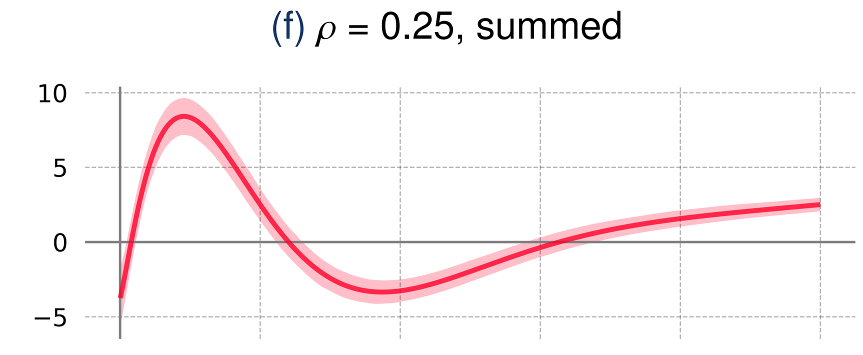
(e) $\rho = 0.25$



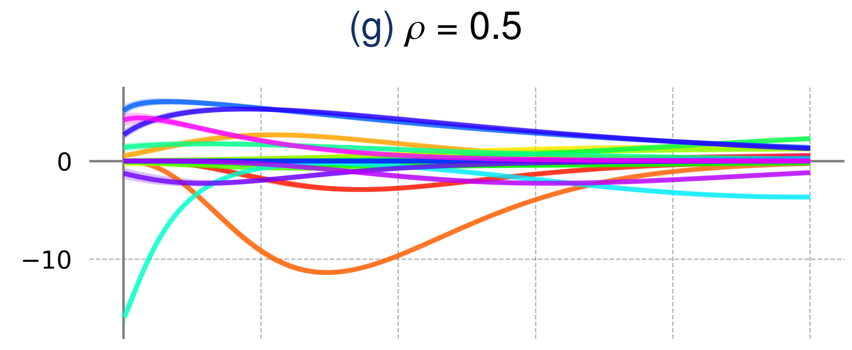
(f) $\rho = 0.25$, summed



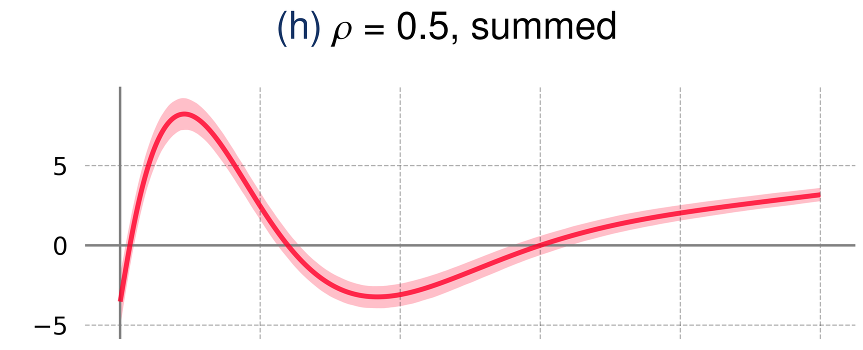
(g) $\rho = 0.5$



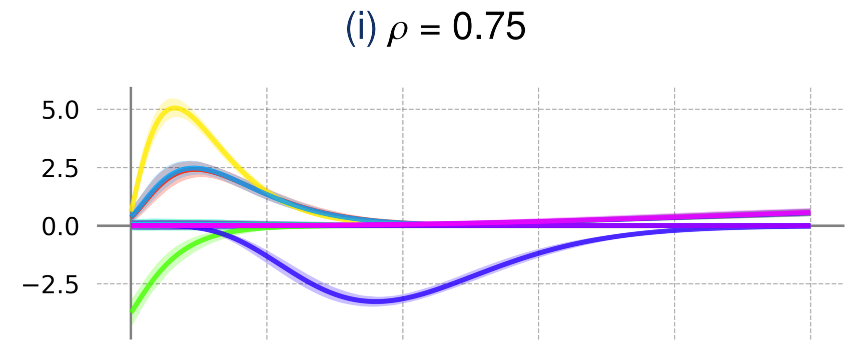
(h) $\rho = 0.5$, summed



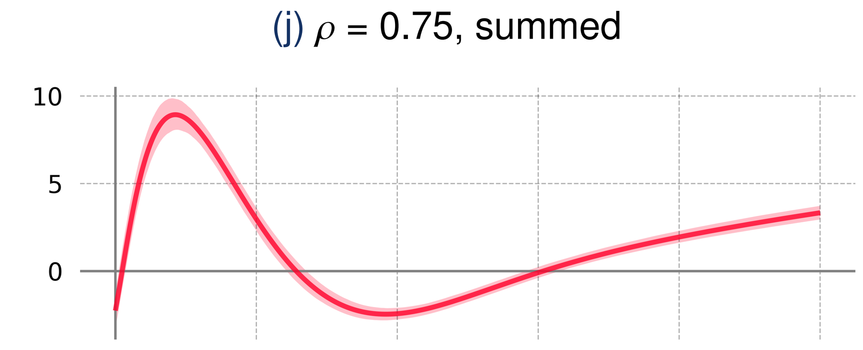
(i) $\rho = 0.75$



(j) $\rho = 0.75$, summed



(k) $\rho = 1$



(l) $\rho = 1$, summed

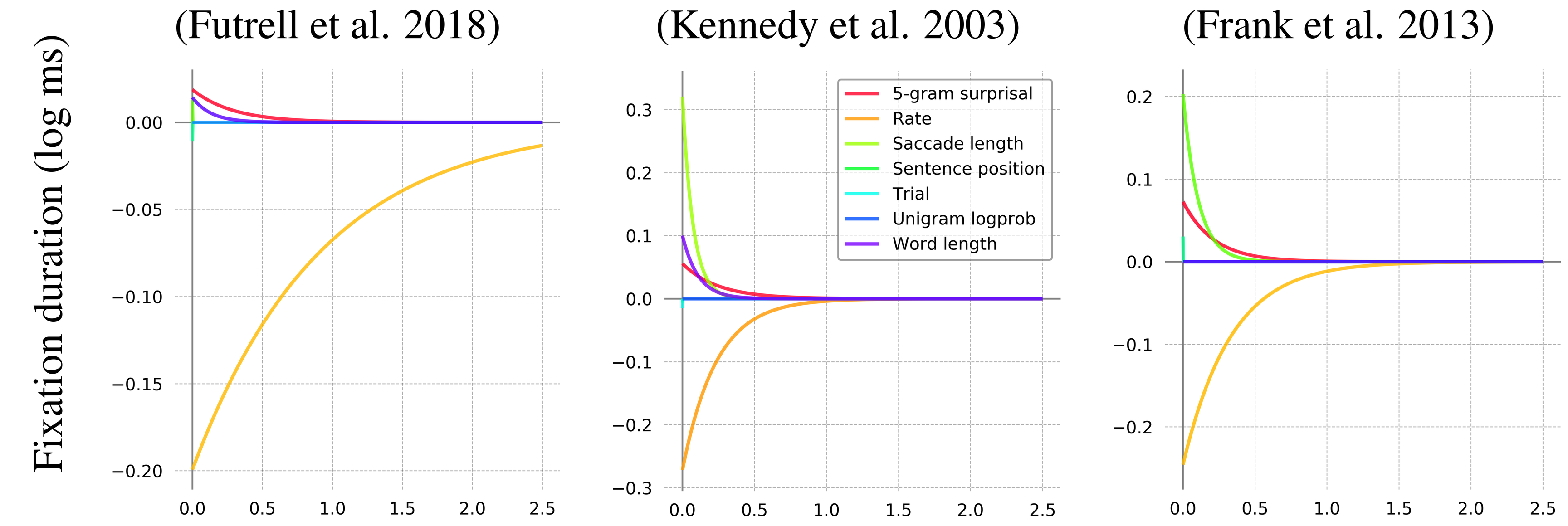
Time (s)

DTSR faithfully recovers ground truth impulse response structure, even with high multicolinearity ($\rho \leq 0.75$).

Naturalistic Evaluation: Reading Time Modeling

Estimated IRFs

Self-Paced Reading Natural Stories (Futrell et al. 2018) | **Eye-Tracking** Dundee (Kennedy et al. 2003) | **Eye-Tracking** UCL (Frank et al. 2013)

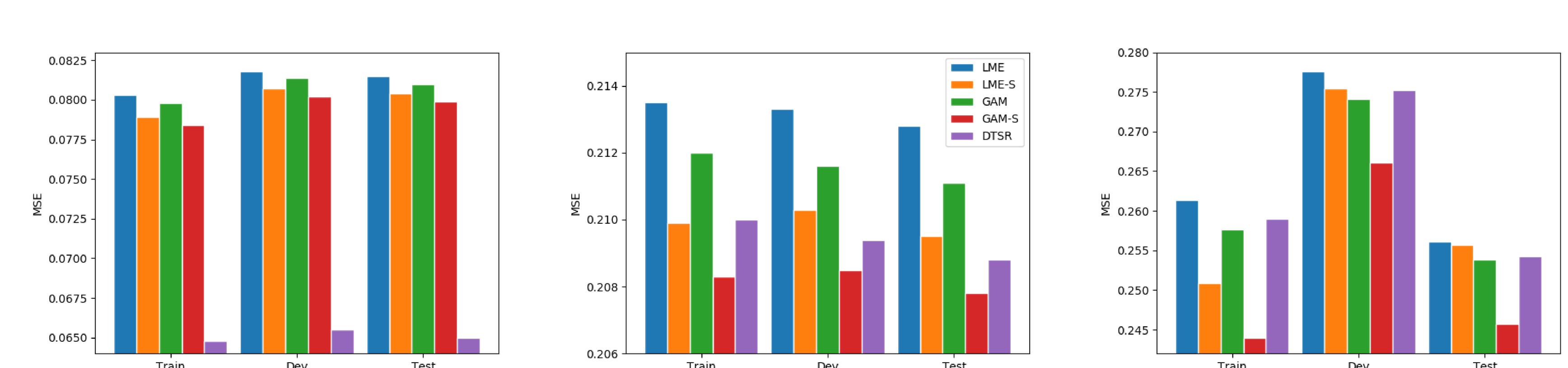


Time (s)

- Diffusion mostly restricted to first second after stimulus presentation
- Large negative influence of *Rate* (convolved intercept) suggests inertia
- Top-down response slower than bottom-up (surp vs. word/sac. len)
- Similar temporal profile across eye-tracking corpora

Prediction error

Natural Stories | Dundee | UCL



DTSR provides comparable or improved prediction quality to widely-used baselines. Baselines with "-S" had 3 spillover positions per predictor. Significant ($p = 0.0001^{***}$) overall improvement against all baselines.

Conclusion

- Results validate fitted IRFs:
  - Closely recovers synthetic model, even with high multicolinearity
  - Comparable or improved prediction performance on human data to widely-used baselines
- IRF estimates reveal important patterns that are not easily detectible without DTSR