# For your eyes only: Computational modeling of gaze behavior

# Bayesian parameter estimation for dynamical models of eye-movement control using adaptive Markov Chain Monte Carlo simulations

*R.* Engbert<sup>1,3,4</sup>, *M. M.* Rabe<sup>1,4</sup>, *S. A.* Seelig<sup>1,3</sup>, *S.* Reich<sup>2,3</sup>, <sup>1</sup>Department Psychologie, <sup>2</sup>Institut für Mathematik, <sup>3</sup>SFB 1294, <sup>4</sup>SFB 1287, Universität Potsdam

# In Short

- We study parameter inference for cognitive models of eye-movement control during reading.
- For the class of dynamical cognitive models being investigated a likelihood function can be computed numerically for experimental data sets.
- Using adaptive Markov Chain Monte Carlo (MCMC) techniques, we implemented a Bayesian approach for model parameter inference.
- We aim at increasing the applicability of this approach by improving means of constructing the likelihood function as well as estimation algorithms.
- In perspective, this approach might be exploited for gaze prediction in everyday tasks (e.g. humanmachine interaction, driving).

In cognitive modeling, we investigate mathematical and computational approaches to explain human behavior based on assumptions on information processing of mind and brain. Model fitting and comparison are increasingly important due to growing model complexity and competing modeling approaches.

Experimental data are available as fixation sequences from eye-tracking experiments (time series), i.e., sequences of spatial fixation locations and fixation durations of the eyes within the stimulus display. We developed a procedure for computing the likelihood of fixation sequences for our models given a set of model parameters. This likelihood function is fundamental for a Bayesian approach to parameter inference. Since we work with time-ordered data, this framework is called data assimilation.

In a Bayesian approach, advanced Markov Chain Monte Carlo (MCMC) methods exploit the likelihood function to generate a posterior sampling distribution of model parameters for a given data set based on priors on model parameters. Feeding the obtained parameters back into the generative models allows us to compare empirical with simulated data. This provides extremely sensitive information for model evaluation and improvement. We successfully demonstrated the viability of our approach [1,2] and seek to advance its applicability by improving on the means of constructing the likelihood function, as well as researching more effective estimation algorithms.

Since the eyes are both sensory and motor systems, they are among the best experimental measures of ongoing cognition. We investigate several dynamical models which generate sequences of eye movements during reading [1], scene perception [2] and maintained fixation [3]. The project currently focuses on dynamical models of eye-movement control during reading, which is an important area of cognitive modeling.

As high-acuity visual perception is limited to the very center of the visual field (the fovea) and visual information processing is only efficient if the eye is stationary, our visual system has to apply a discrete sampling strategy. The eye performs a rapid sequence of fast eye movements (saccades, about 3-4 times per second) to minimize flight time, while information processing is done during fixations (average duration 200 ms). In reading, approximately one fixation is generated per word, however, there is statistical variability depending on word length and word difficulty. Even during reading of simple texts, only 50% of all saccades move the gaze from one word to the next (see Fig. 1). Frequently, refixations of the same word, skippings, and regression (saccades that go against reading direction) are observed.



**Figure 1:** Saccades during reading. Only about 50% of the saccades during reading move the eye from word n to n+1. Other types of saccades are skippings, refixations, and regressions.

The SWIFT model of eye-movements during reading [1] is a dynamical model that generates all types of saccades from a theoretically coherent framework using a temporally evolving neural activation field (Fig. 2). The model combines well-established cognitive assumptions on sensorimotor processing, attention, and word recognition with neurophysiological plausibility (e.g. partial independence of spatial and temporal control of saccades).

The model uses stochastic internal (latent) states that need to be approximated in the numerical computation of the likelihood function. Since the log-



**Figure 2:** SWIFT model. Visualization of a simulated gaze trajectory from the SWIFT model of eye-movement control during reading. The three curves on the left part (grey) represent internal processes of saccade programming, while the right curves (colors) indicate word recognition processes during the reading process.

likelihood function is a sum of statistically independent likelihoods for single experimental trials (100 to 1,000 for a typical experiment), the numerical calculations are ideally implemented via parallel computations.

An important challenge for parameter inference of cognitive models is the fact that statistical random effects are observed due to inter-individual differences (e.g., reading span of experimental participants) and text material (e.g., text genre). The Bayesian approach can account for both of these random effects simultaneously by using hierarchical priors on the model parameters.

The complexity of the posterior distributions and high dimensionality of the parameter space requires advanced sampling strategies within the MCMC framework. We successfully employed approximate Bayesian computation, pseudo-marginal likelihood methods and differential evolution algorithms to increase computational performance.

Using the HLRN-IV computing facilities, we implemented a fully Bayesian framework for parameter inference [2] and used an adaptive MCMC procedure, the DREAM framework with improvements [4]. For parameter estimation, we used eye tracking data of 36 participants who read 150 single sentences each. For every participant only 70% of the data were needed during the estimation (training data), so that the remaining 30% of the data (test data) were available for simulated data sets which were based on point estimates of the obtained posterior parameter distributions. We compared typical measures of fixation durations (contingent on saccade programming) and fixation probabilities (relating to oculomotor behavior and target selection). The comparisons indicate a remarkable agreement of simulated and experimental data (Fig. 3).



**Figure 3:** Relationship between fixation probabilities (left) and mean fixation durations (right) of simulated and experimental data. Each datapoint represents one participant.

We will explore additional techniques for synthetic and approximate likelihood computations to improve the efficiency of our simulations. In perspective, our approach to model identification for individual observers could be used to implement gaze prediction across tasks on the basis of mobile eye-tracking devices. A potential application is the surveillance of human-machine interaction.

#### www

https://www.sfb1294.de/research/research-areab/b03/

#### **More Information**

- S. A. Seelig, M. M. Rabe, N. Malem-Shinitski, S. Reich, & R. Engbert, *Manuscript under review* https://arxiv.org/abs/1901.11110
- [2] H. H. Schütt, L. O. M. Rothkegel, H. A. Trukenbrod, S. Reich, F. A. Wichmann, & R. Engbert, *Psychol. Rev.* **124**(4), 505–524 (2017). doi: 10.1037/rev0000068
- [3] R. Engbert, K. Mergenthaler, P. Sinn, & A.
  Pikovsky, *PNAS* 108(39), 765–770 (2011). doi: 10.1073/pnas.1102730108
- [4] J. A. Vrugt, C. ter Braak, C. Diks, B. A. Robinson, J. M. Hyman, & D. Higdon *IJNSNS* 10(3), 271–288 (2009). doi:10.1515/IJN-SNS.2009.10.3.273

## **Project Partners**

DFG Sonderforschungsbereich 1287 "Limits of Variability in Language", Project B03

### Funding

DFG Sonderforschungsbereich 1294 "Data Assimilation", Project B03