Data assimilation for parameter estimation of a single-compartment Type 1 diabetes ODE model

Christina Catlett[1], Daniel Shenker[2], Rachel Wander[3], Maya Watanabe[4]

[1]Scripps College, [2]Johns Hopkins University, [3]Harvey Mudd College, [4]University of Massachusetts Amherst

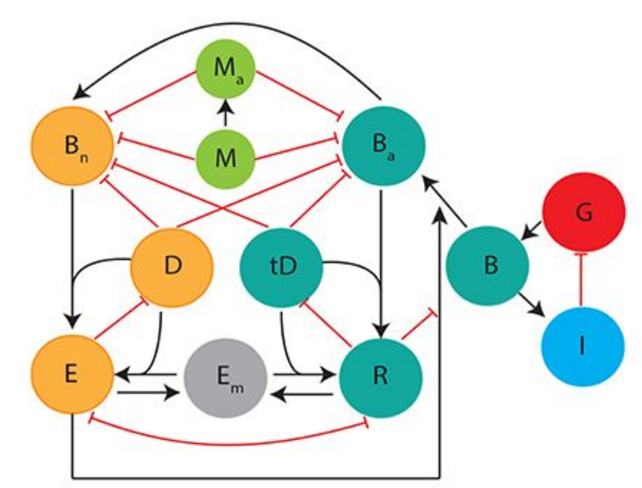
Motivation

Mathematical modeling allows us to simulate cell interactions in Type 1 diabetes onset. By connecting existing mouse glucose data to our dynamical systems model, we estimate realistic parameters for the system, allowing it to act as an *in silico* testbed for emerging treatments. We explore three data assimilation techniques for parameter estimation.

Model Description

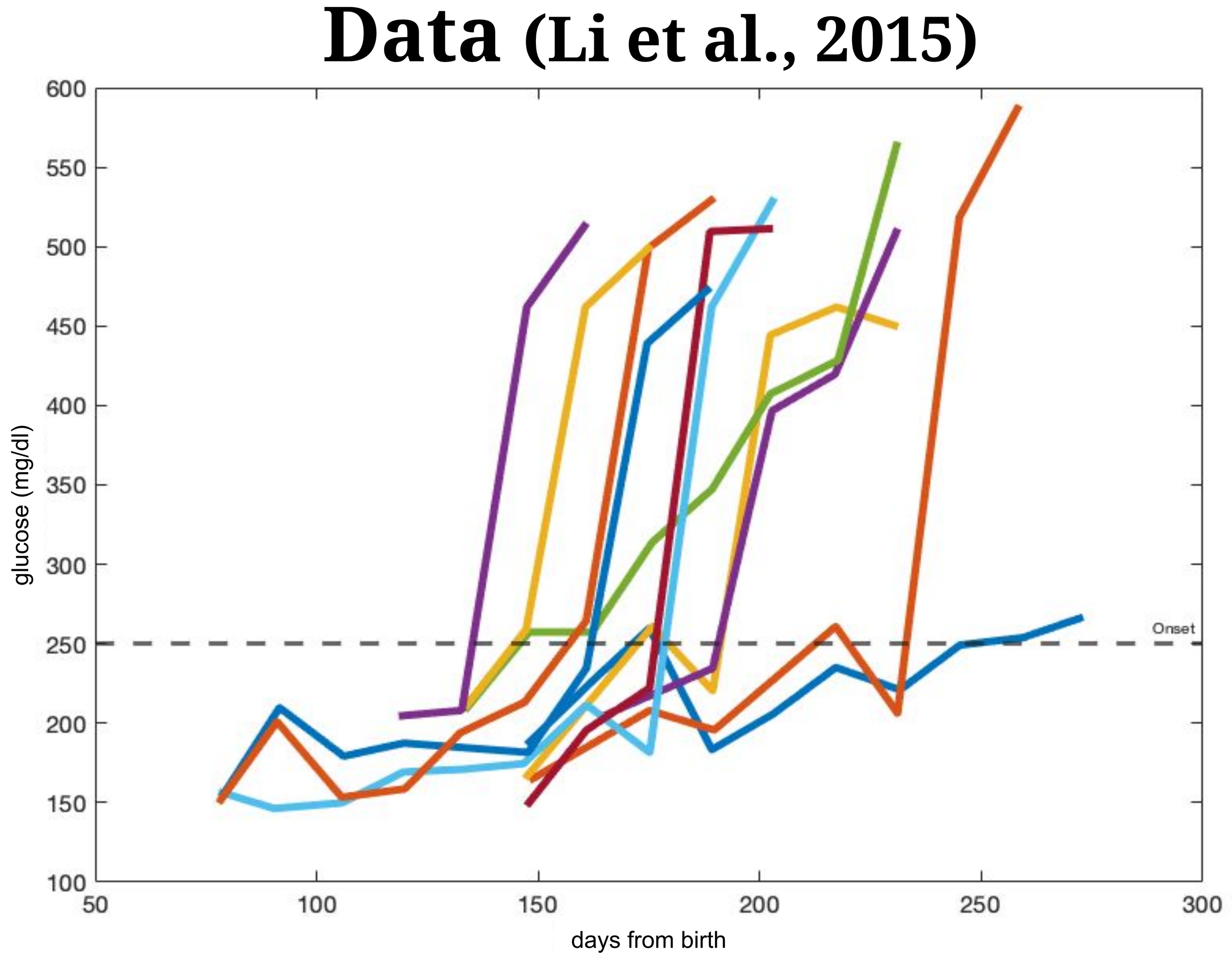
Single compartment non-linear ODE model from Shtylla et al., 2019:

- 12 states describe immune cell populations, glucose, insulin of pancreas over time
- Simulates *apoptotic wave* as hypothesized catalyst for diabetes onset
- Parameters sit 'on the edge' of diabetes, where small variation sends a mouse from non-diabetic to diabetic state



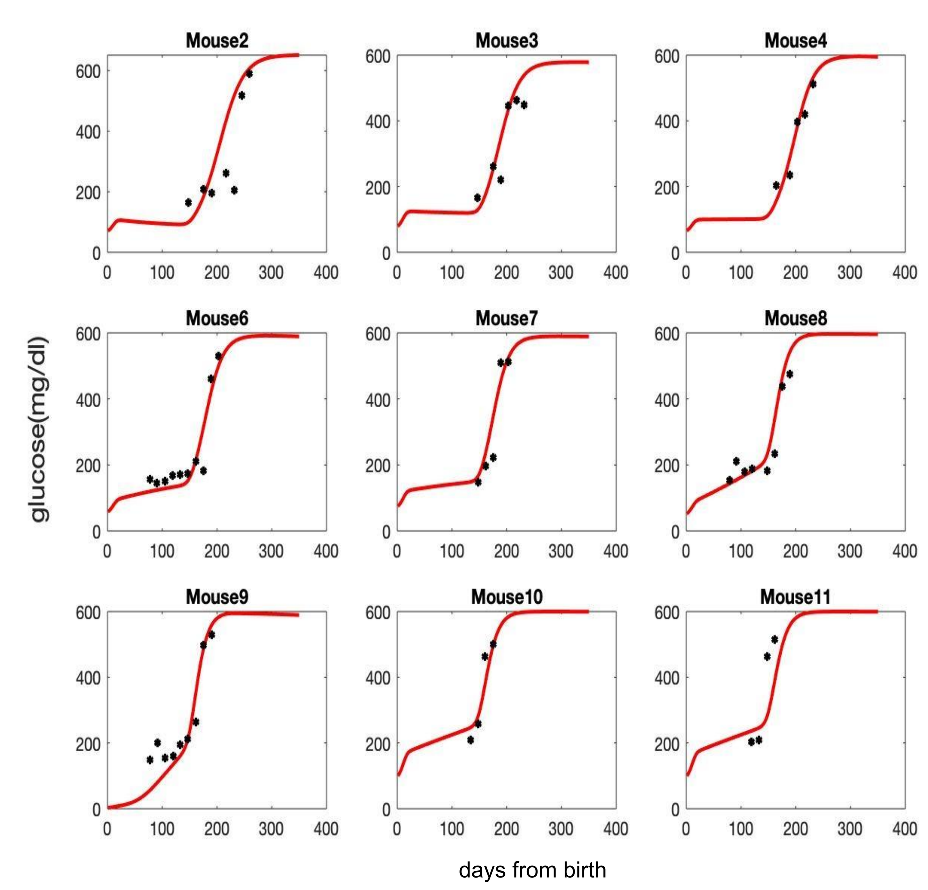
Variable	Description
B	Healthy β-cell population
Ba	Apoptotic β-cell population
Bn	Necrotic β-cell population
D	Immunogenic DC population
tD	Tolerogenic DC population
E	Effector T cell population
Em	Memory T cell population
R	Regulatory T cell population
Ma	Active macrophage population
M	Macrophage populations
G	Glucose levels
I	Insulin levels

Data (Li et al., 2015)

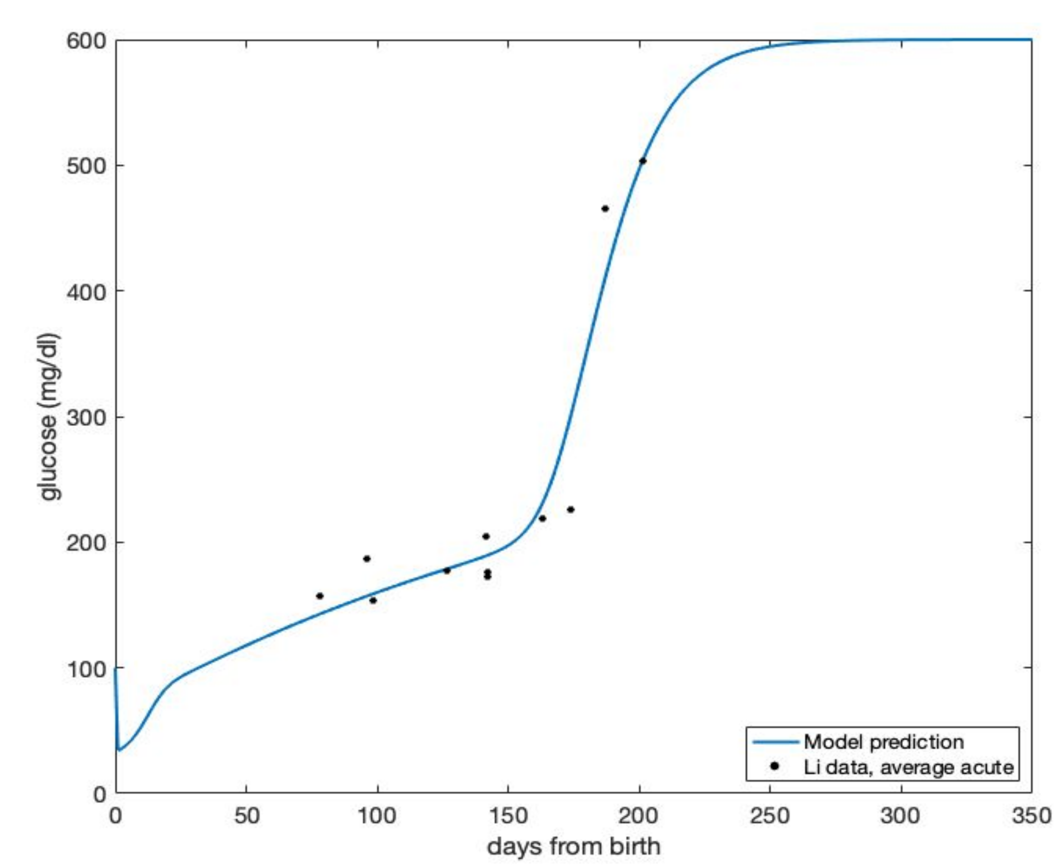


Results

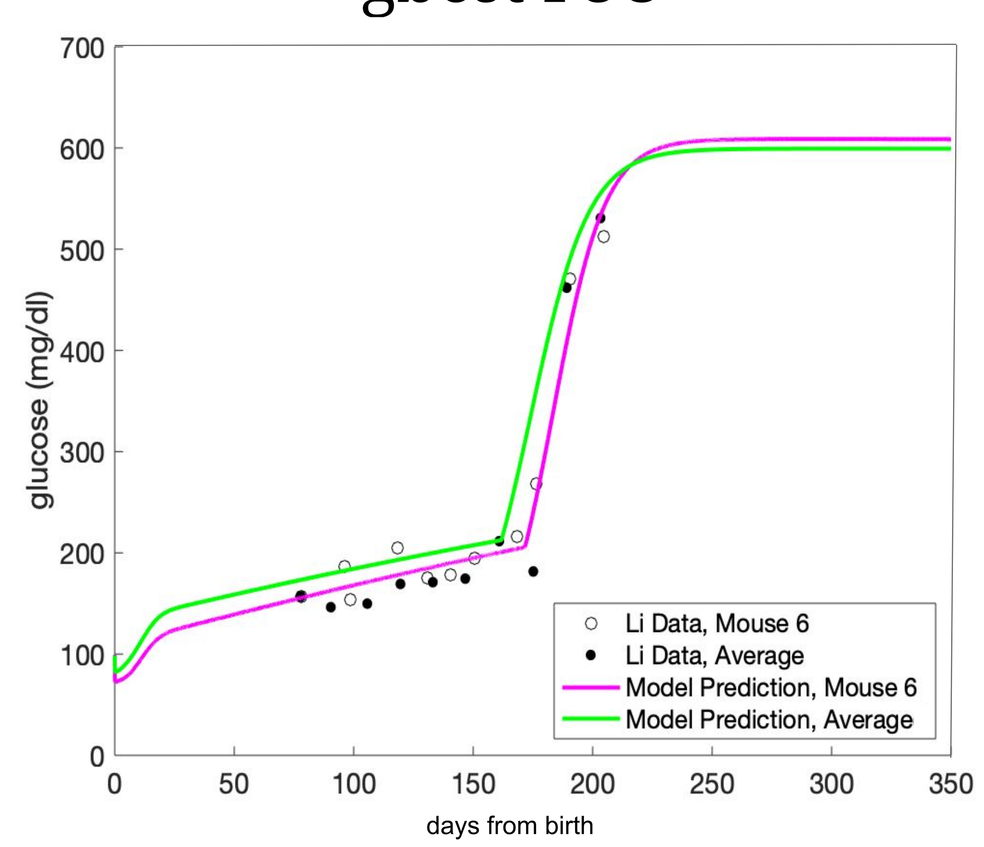
Dual UKF (Individual)



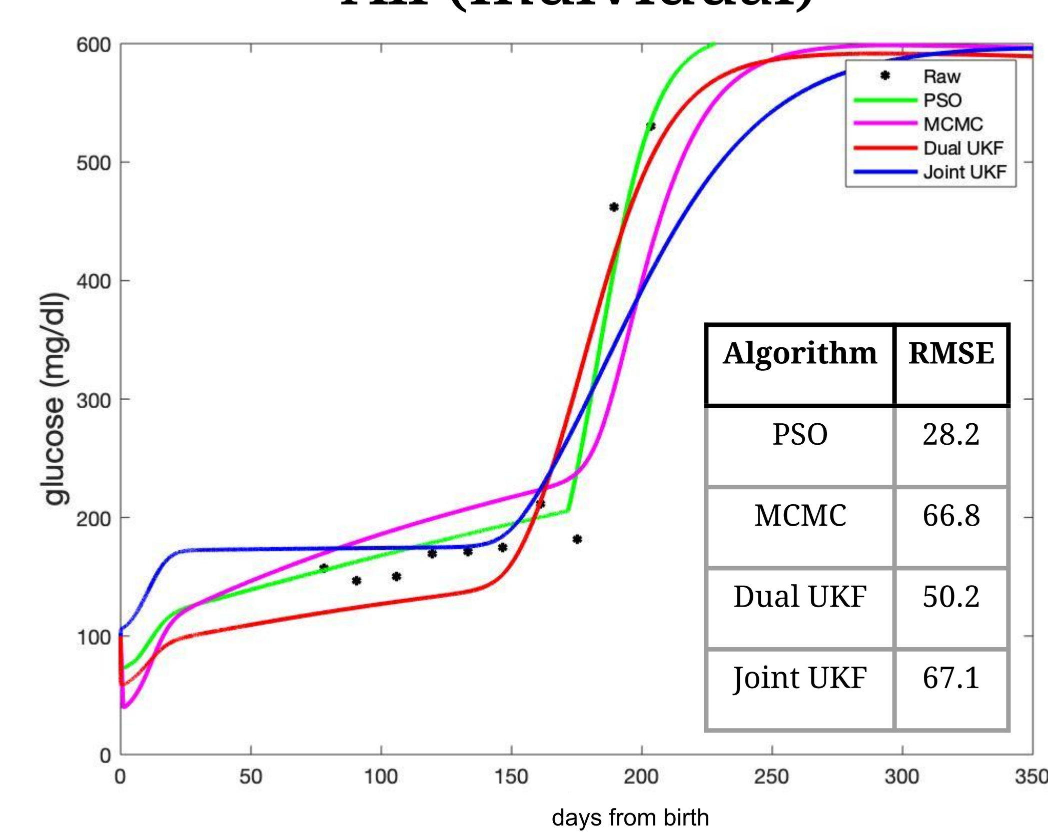
DRAM MCMC (Averaged)



gbest PSO



All (Individual)



Algorithm	RMSE
PSO	28.2
MCMC	66.8
Dual UKF	50.2
Joint UKF	67.1

Conclusions

We are interested in understanding which parameter estimation technique performs best under certain data constraints, in particular whether the dataset is population or individual level. On the individual level, the *UKF* performs best. This is due to its superior ability to handle the noise associated with a single individual.

When data is population-level (e.g., averaged across individuals), the *MCMC* technique performs best. Because MCMC considers all of the available data at once, we hypothesize that this makes it superior for deriving holistic interpretations of a dataset, assuming viable prior and likelihood distributions are selected.

Finally, *PSO* shows no preference towards individual or population-level data. We hypothesize that this is due to its agent-based approach and low number of design parameters in comparison to other techniques tested.

Methods

Unscented Kalman Filters (UKF)

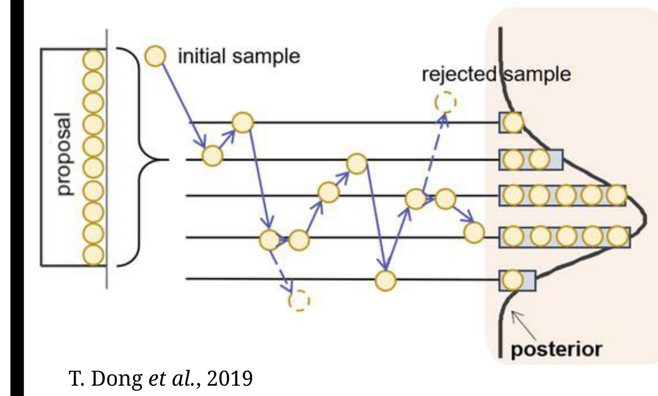
The UKF is a recursive filtering method, rooted in state space formulation of dynamical systems. The filter brings in observed data points one by one, predicting the model's unobserved states at each step. This process is tuned as more data are considered. The algorithm has two main steps:

1. Prediction: Evaluate model using previous estimate
2. Update: Use error between prediction and data to update estimate.

We ran our data on both Dual and Joint UKFS. We find the Dual UKF most accurate for our system.

Markov Chain Monte Carlo Methods (MCMC)

MCMC methods propose and evaluate a large series of candidate parameter sets. Candidates are accepted or rejected according to acceptance criterion based on their Bayesian probability (prior, likelihood). Accepted candidates create a posterior distribution from which optimal parameters can be selected.



T. Dong *et al.*, 2019

Delayed rejection adaptive Metropolis (DRAM) proposes multiple candidates before rejection and adjusts acceptance criterion during the run.

Particle Swarm Optimization (PSO)

Biologically inspired, PSO is an agent-based algorithm for optimization. Agents stochastically update velocity vectors to traverse parameter space according to knowledge gained from the evaluation of an objective function by both the individual and the swarm. The most classical implementation, we used global best (gbest) PSO, where the network of agents is fully connected.

Acknowledgements

We thank Professors Blerta Shtylla, Lisette de Pillis, and Christina Edholm for their endless guidance, and An Do for her previous work on the model. Additionally, the NSF (#1757952) and the HMC for their support.