



# Predicting Bus Dwell Times in Dense Urban Transit Networks Using Bayesian Hierarchical Modeling (RISS 2017)

Conor Igoe, Eli Bronstein,  
Dr. Isaac Isukapati, Prof. Stephen Smith



## Motivation

- Public transit service quality must improve as urban societies seek to reduce their carbon footprint and rely on less energy-intensive technologies.
- Current efforts involve Transit Signal Priority (TSP) systems, but these fail to make strategic traffic signal control decisions that accommodate both transit schedules and non-transit vehicle flow.
- A holistic TSP approach is complicated because transit vehicles have frequent stops with uncertain dwell times.
- We aim to predict bus dwell times in real-time to form a more accurate characterization of bus behavior and empower the development of TSP systems that maximize transit quality for all passengers.

## Approach

- We analyze a dataset of over 100 million timestamped records of dwell times and number of onboarding and alighting passengers for 2012-14 in the Pittsburgh area.
- We develop a real-time, lightweight, robust Bayesian hierarchical model consisting of an Onboarding Model and Dwell Time Model.

## Results

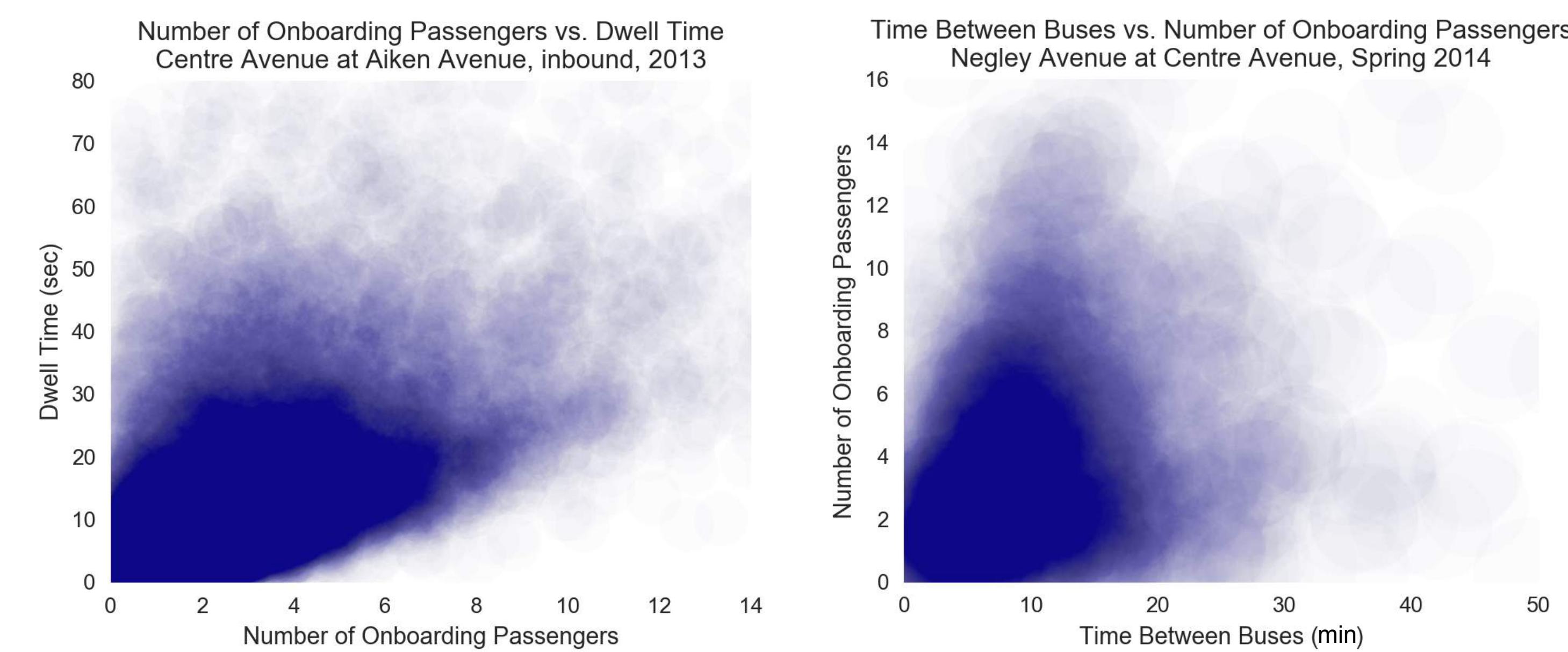
- Time between bus arrivals is positively correlated with onboardings, as are onboardings and alightings with dwell time. This correlation is weaker for larger values.
- When predicting onboardings for an approaching bus, we achieve a median absolute error of 1.95 passengers and a median error of -0.84 passengers.
- When predicting dwell times with known onboardings and alightings, we achieve a median absolute error of 4.08 seconds and a median error of 1.0 seconds.

## Future Work

- Improve Onboarding Model; develop an Alighting Model; combine the Onboarding, Alighting, and Dwell Time Models into a complete hierarchical model; make a classification model to predict whether the bus will stop; and conduct a field test of the hierarchical model to assess generalizability.

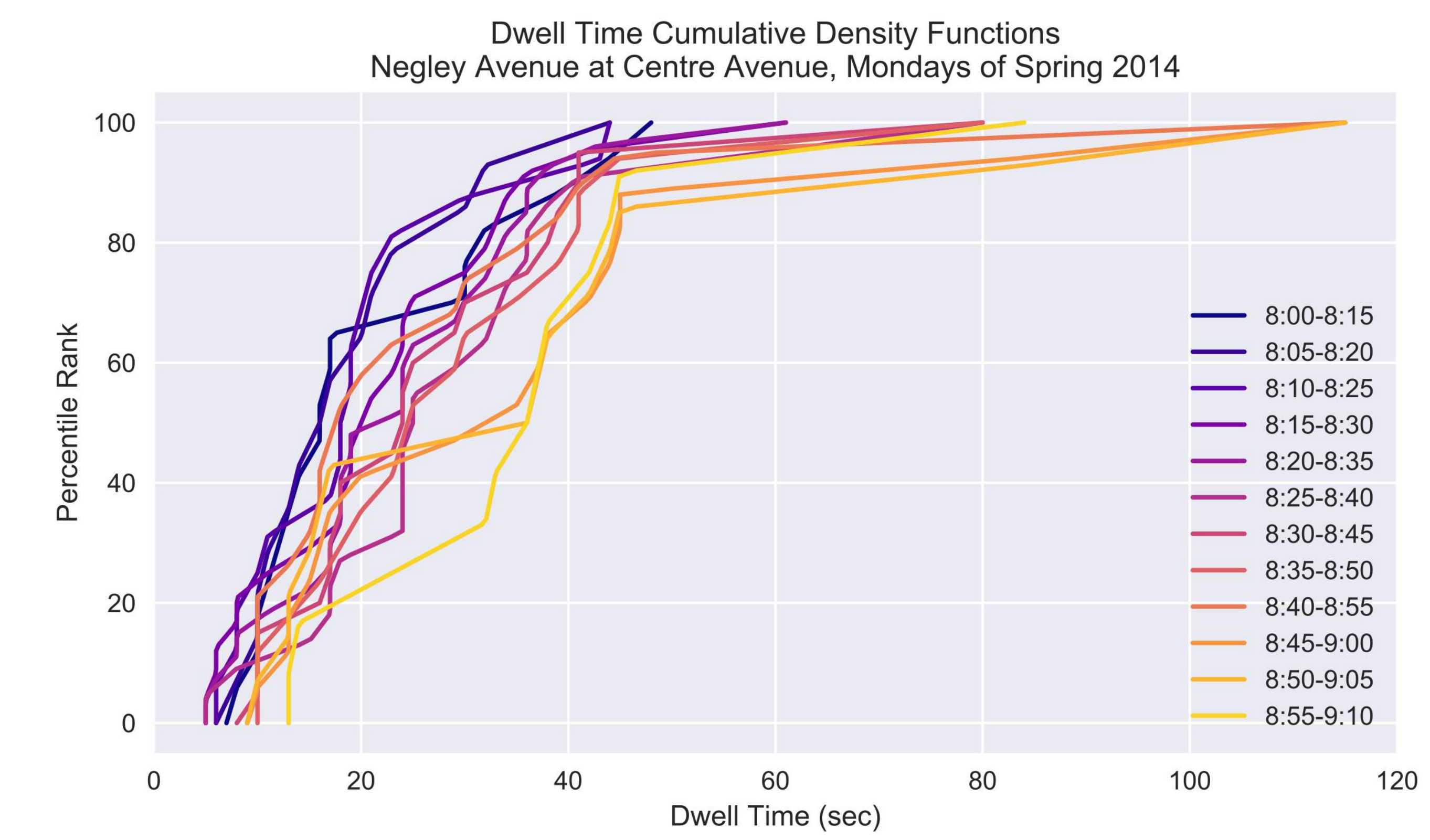
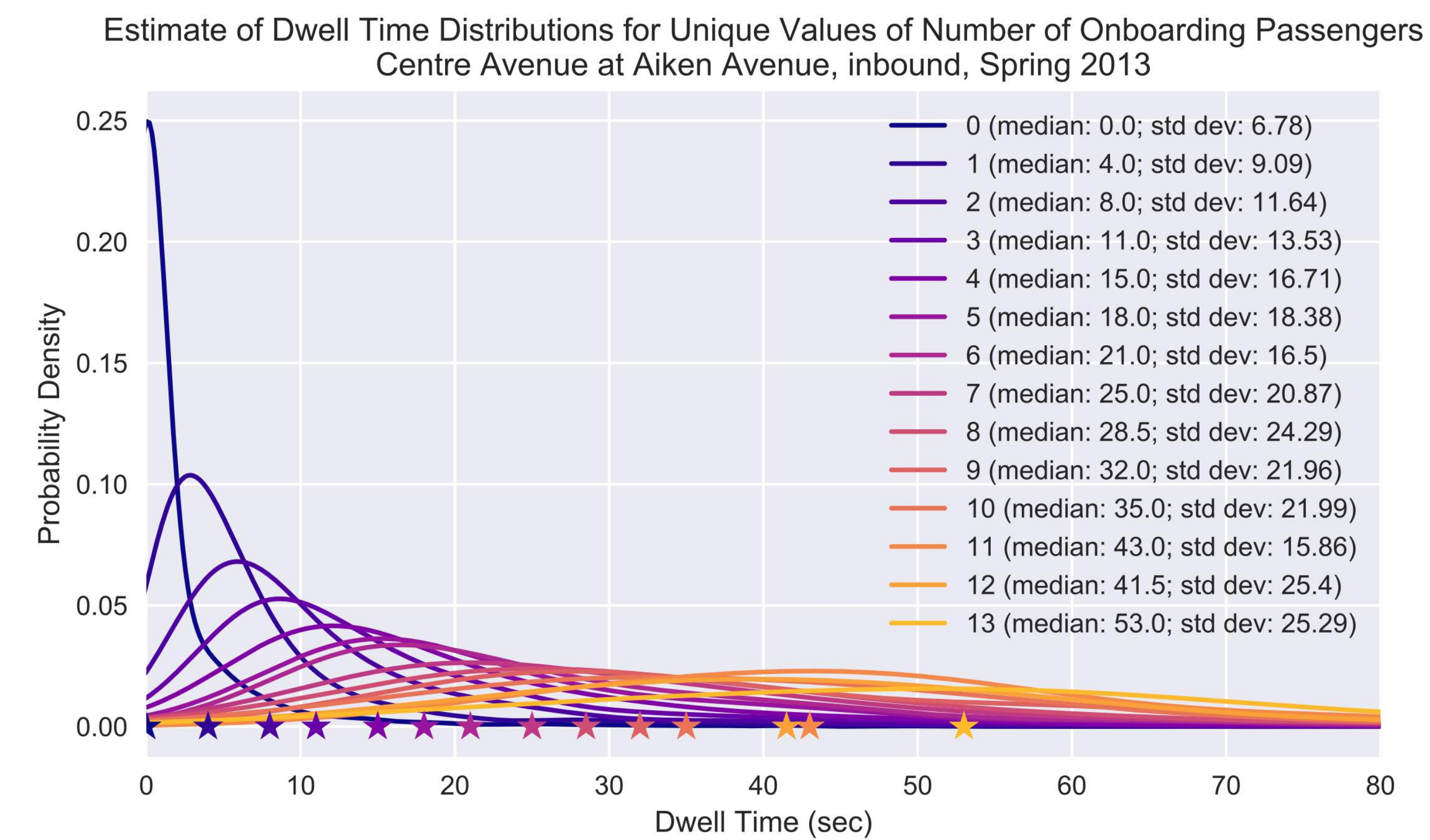
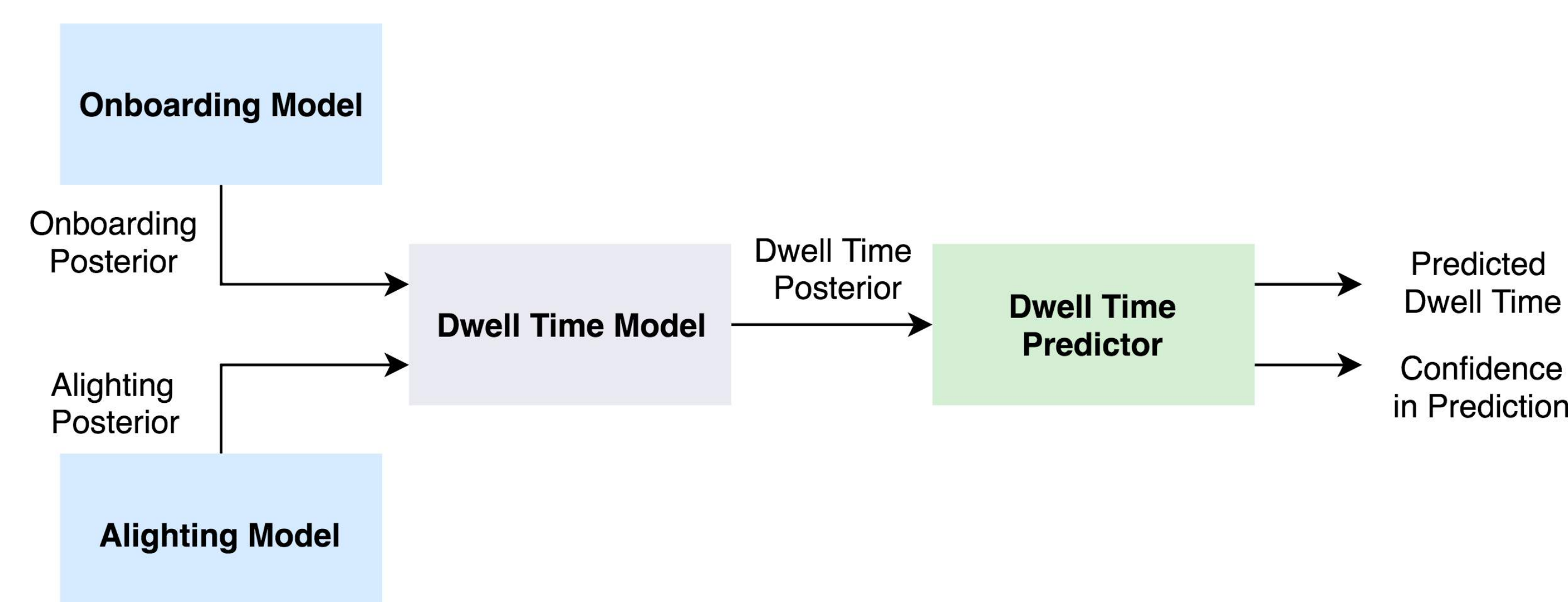
## Data Analysis

- The dwell time distribution can change significantly in a span of 5 minutes. As time goes on, this distribution remains similar for lower percentiles (e.g. < 70), but is unstable for higher percentiles.
- We smooth the discrete dwell time distribution using Kernel Density Estimation (KDE) to obtain a continuous distribution.
- We fit various analytical distributions to the dwell time KDE, and assess goodness of fit using max-deviation tests.
- Log-normal, Non-central F, Log-logistic, and Burr distributions are the best fits.



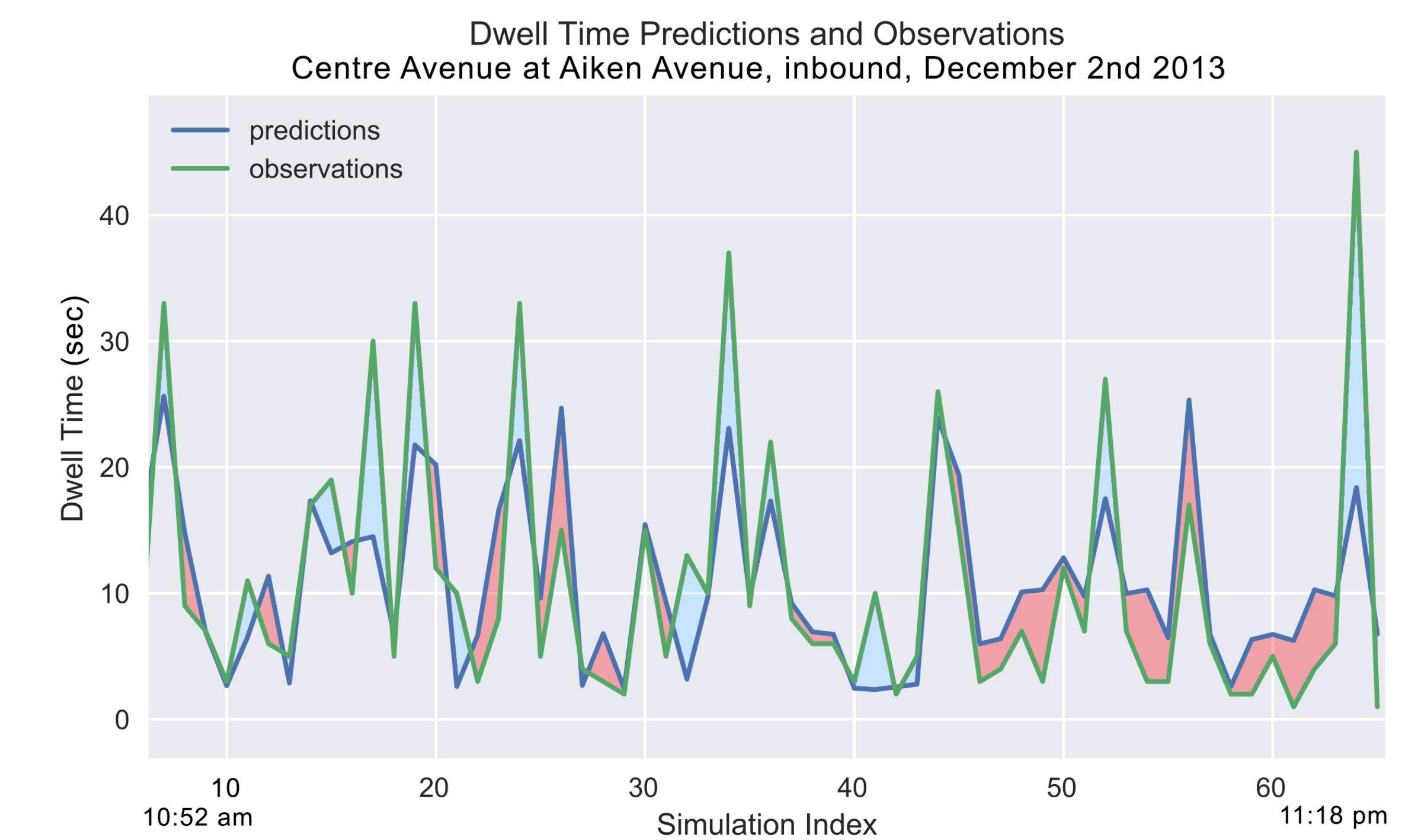
## Onboarding Model

- We model the arrival rate of passengers at a bus stop using a Poisson distribution and derive the analytical posterior Gamma distribution of the Poisson parameter.
- To predict onboardings for an approaching bus, we multiply the predicted arrival rate by the difference between the previous bus departure time and approaching bus arrival time.
- With each new observation, we analytically update the arrival rate distribution using the observed arrival rate.



## Dwell Time Model

- We model dwell times using a Log-normal distribution, where the parameters are defined by the covariates (onboardings and alightings) and their regression coefficients. We use a power kernelization to account for the non-linear relationship between dwell time and covariates.
- For an approaching bus, we use posterior predictive distributions from the Onboarding and Alighting Models as covariates. Then we sample the posterior dwell time distribution using Markov chain Monte Carlo (MCMC) methods and predict a point estimate.
- With each new dwell time observation, we update the posterior distribution of regression coefficients using MCMC sampling.



## Implementation

- We simulate a real-time system to predict dwell times for individual bus stops by stepping through the data chronologically.
- The individual sub-models of the overall hierarchical model are implemented and assessed separately.
- We test both models assuming error-free covariate prediction (i.e. Onboarding Model uses correct time between buses and Dwell Time Model uses correct number of onboardings and alightings).

## Acknowledgements

Thank you to Dr. Isukapati, Prof. Smith, Rachel Burcin, Prof. Dolan, Cormac O'Meadhra, and the entirety of the RISS team.

