Monitoring building occupancy to build data-driven models: how long is enough?

Liam O'Brien, Ph.D., P.Eng.

Associate Professor, Civil & Environmental Engineering

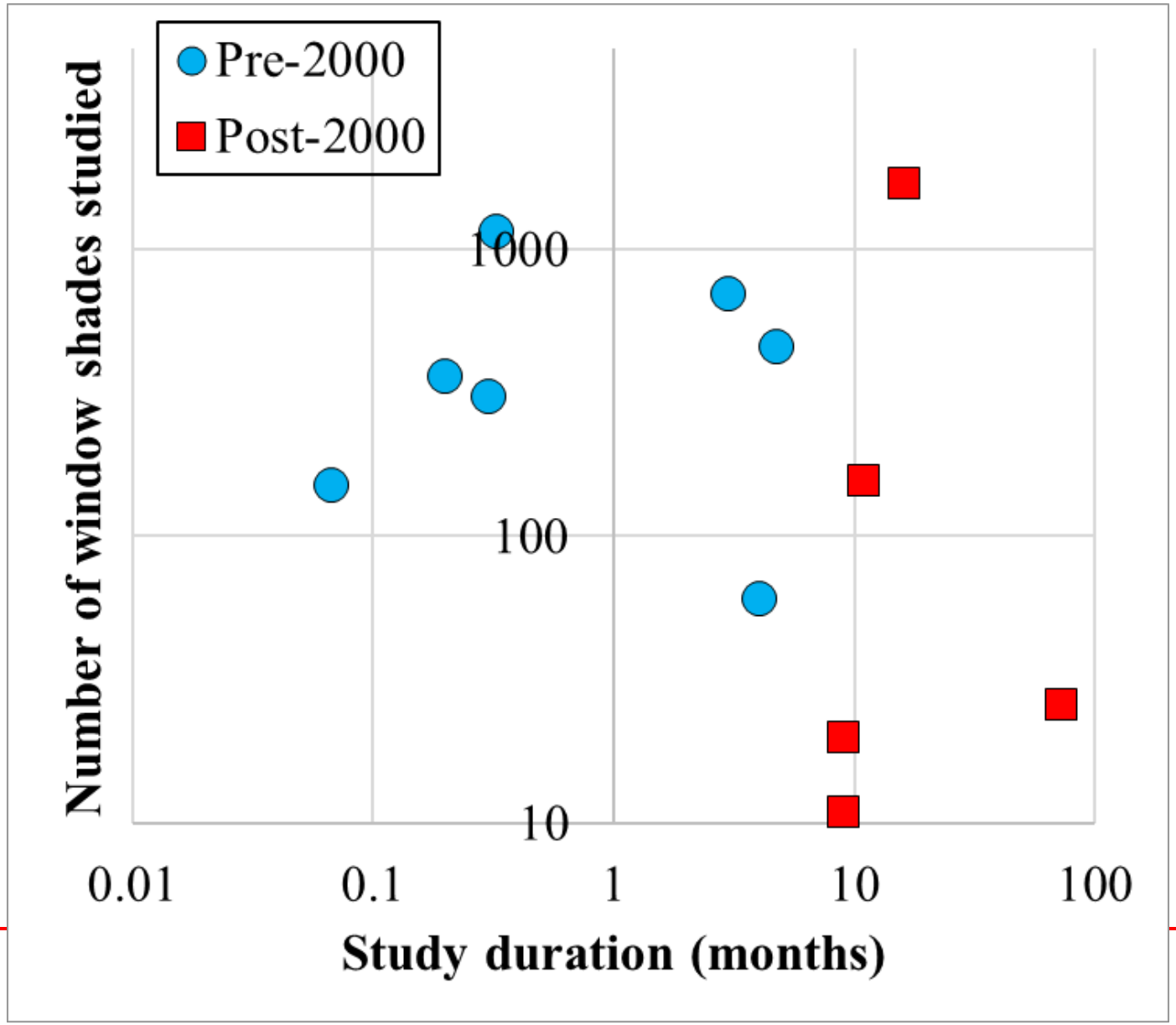
PI, Human Building Interaction Laboratory

Carleton University

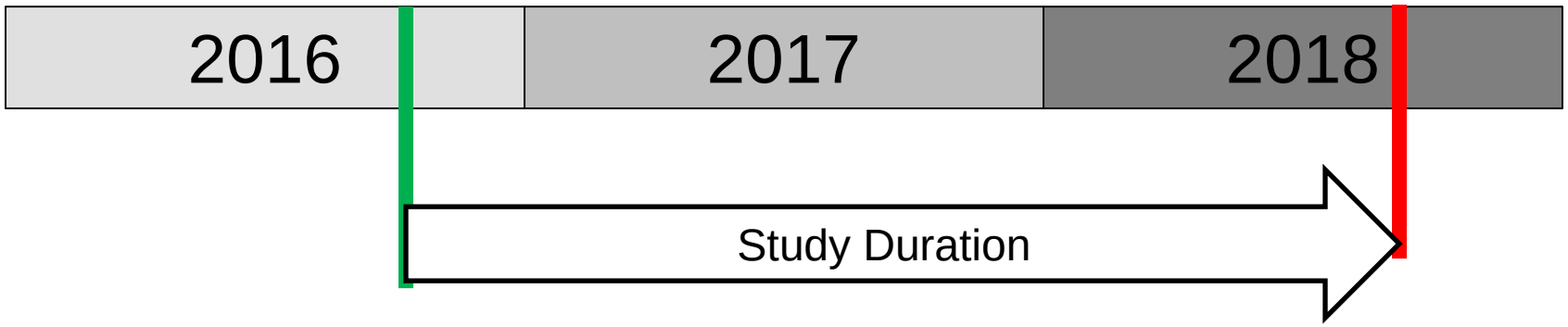
Background and motivation

- Occupant monitoring is time-consuming and often expensive
- The existing literature has studies lasting from days to decades
  - Long studies tend to have smaller sample size
- Previous research indicated ~10^2 occupants should be monitored

Number of occupants vs. study duration



Study start day and duration



- How long is enough to build reliable occupant behaviour/occupancy models? And does the starting time matter?
  - What statistical tools can be used to answer this question?

Methodology: the building and data

- 46 private offices' occupancy recorded for 1 to 5 years
  - Professors and admin
- PIR sensors used, so event-based data converted to timestep data

Methodology: three approaches to answer the question

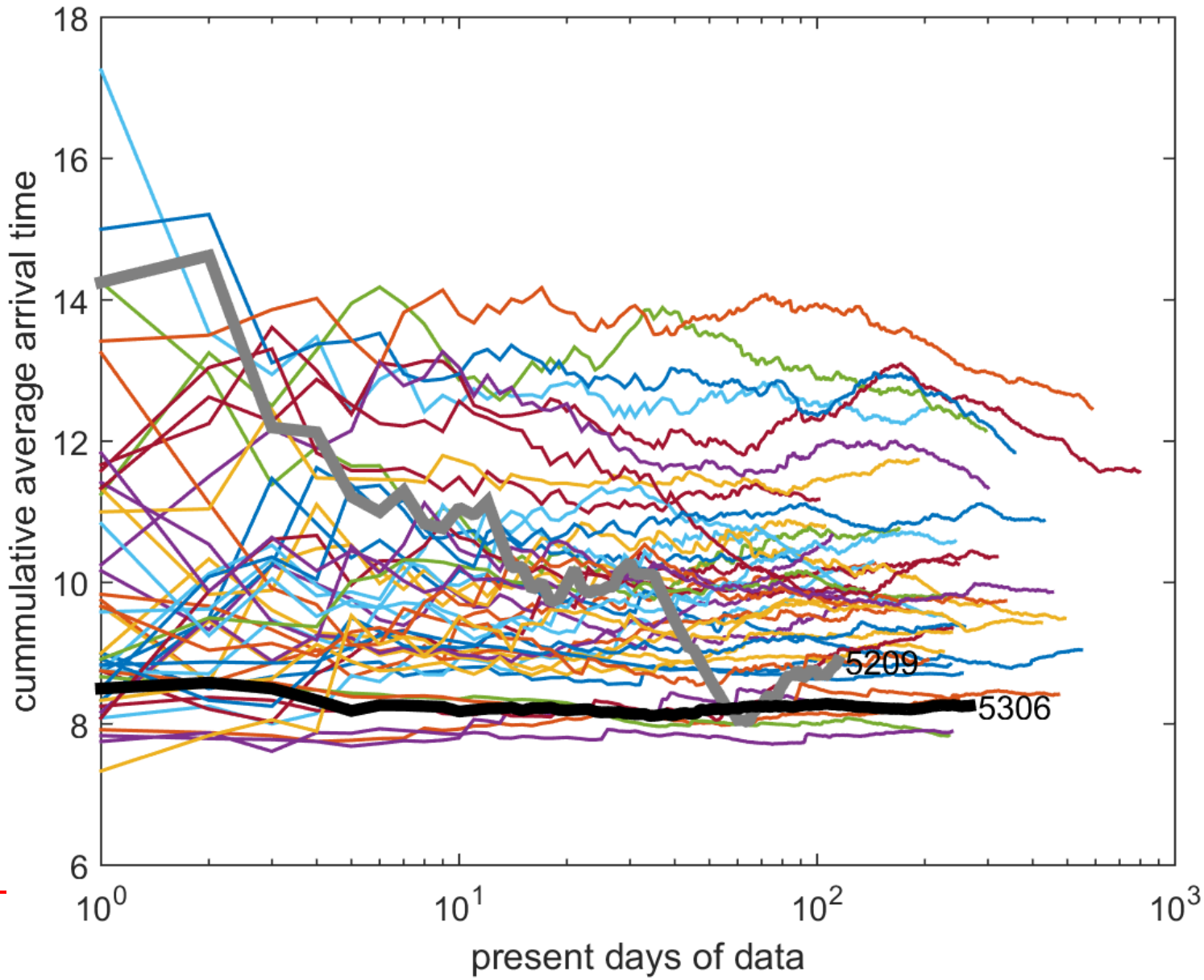
1. Descriptive statistics (trends, autocorrelation, daily, weekly, monthly, annual patterns)
2. Parameter reliability
3. Modelling study to confirm results of reliability

Key occupant traits of interest

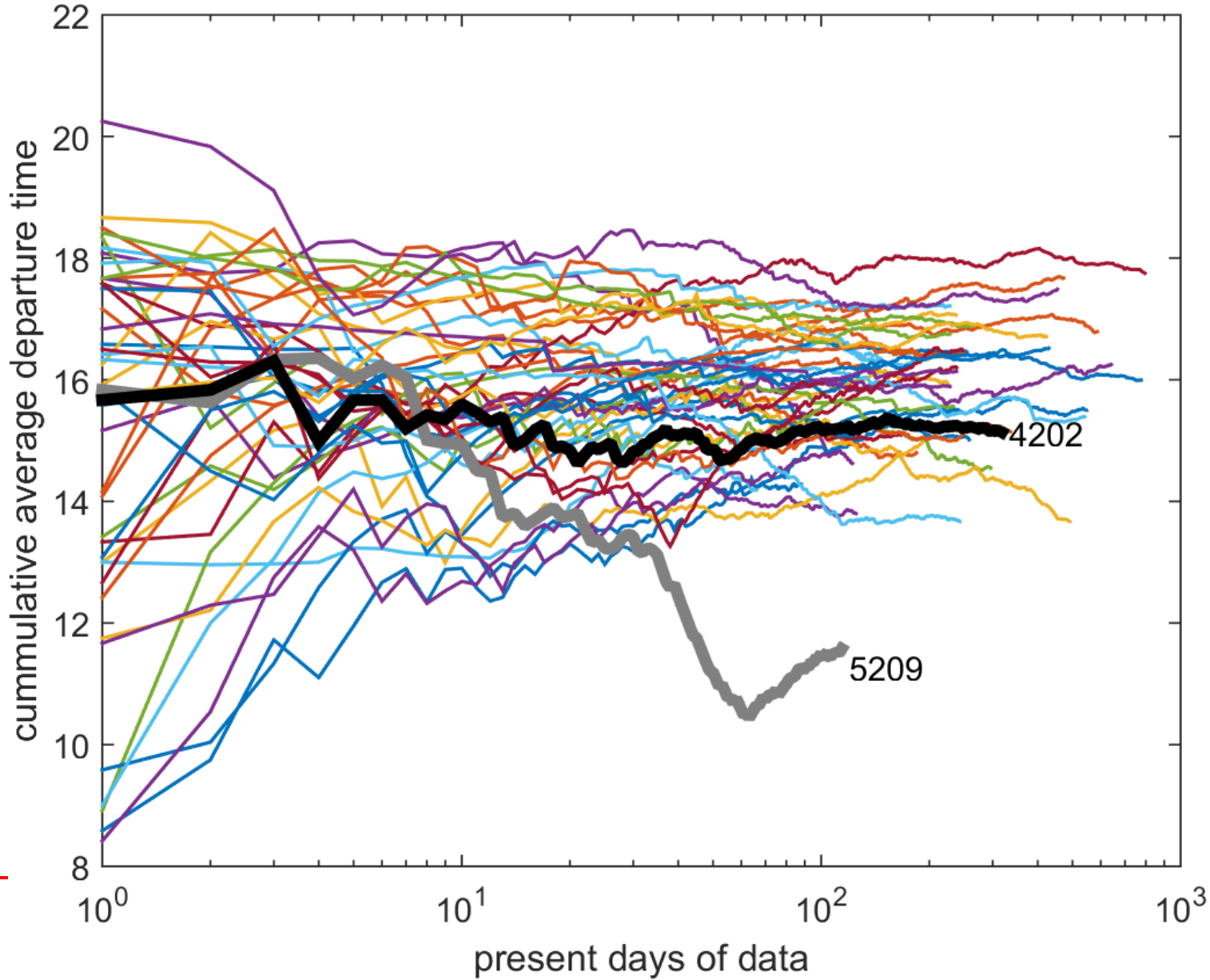
- mean first daily arrival time
- mean last daily departure time
- mean day length (first arrival to last departure)
- mean daytime mobility parameter (ratio of change to no change in presence state)
- long absence probability and duration
- fraction of days absent

1. Descriptive Statistics

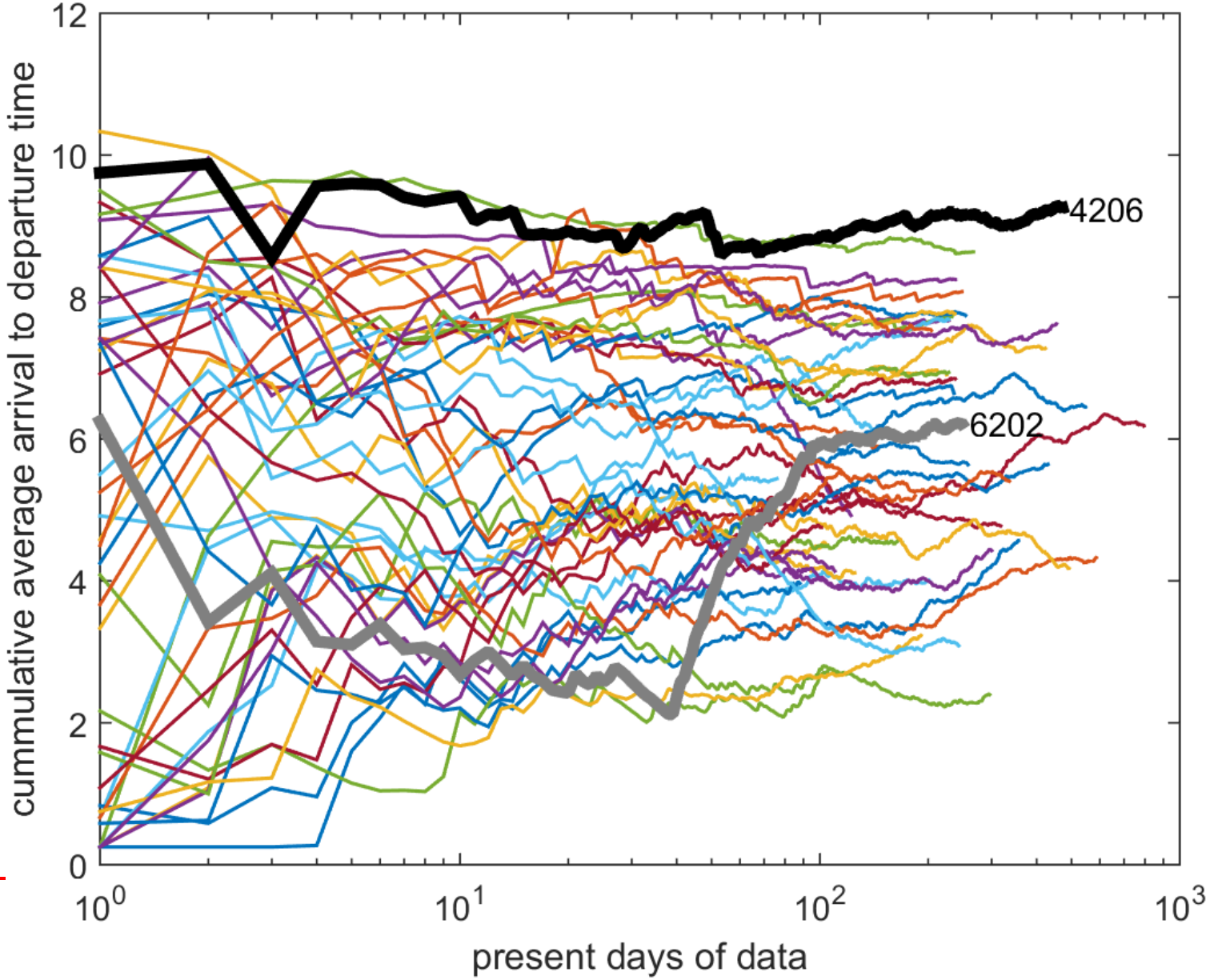
Mean Arrival Time Trend



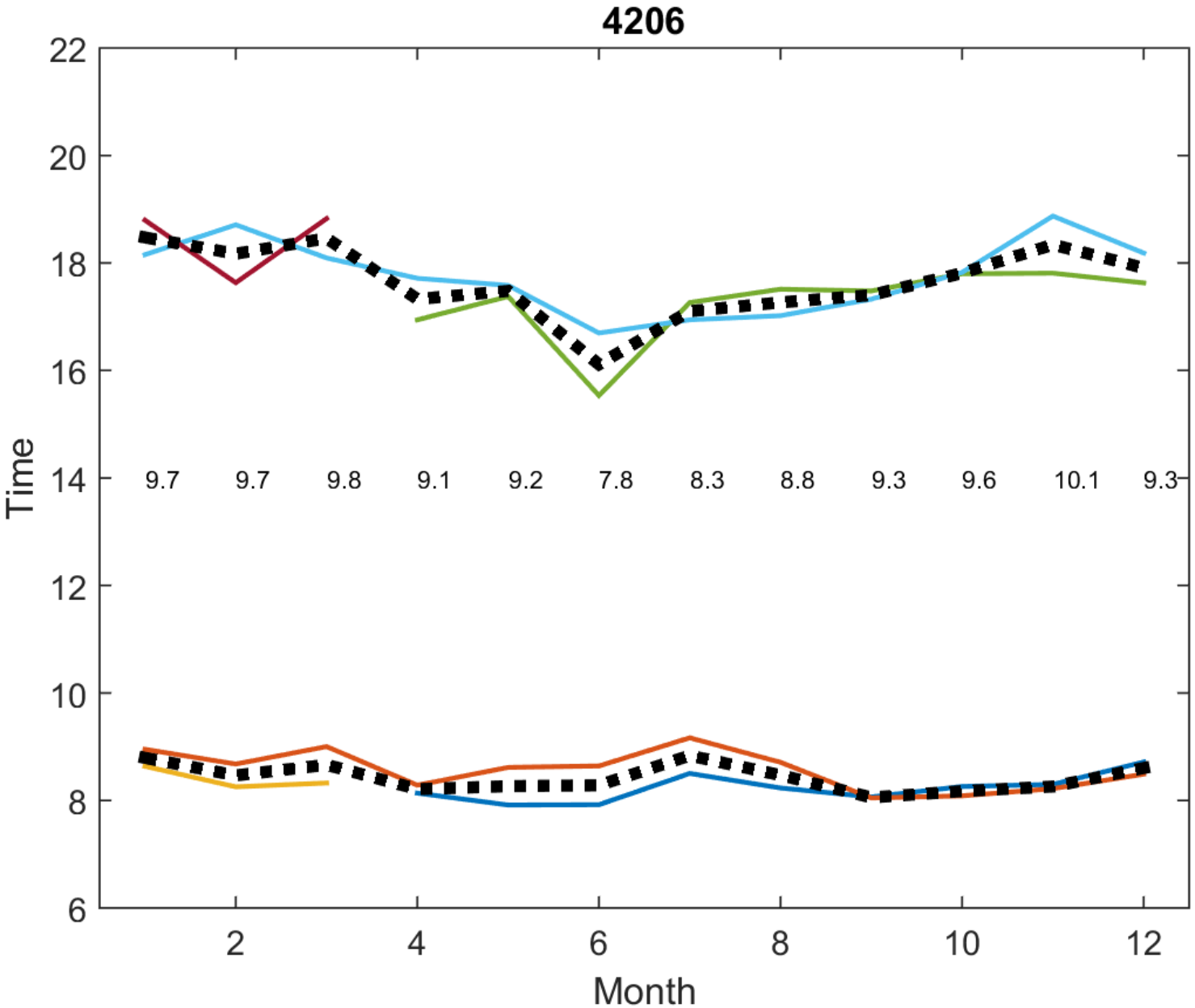
Mean Departure Time Trend



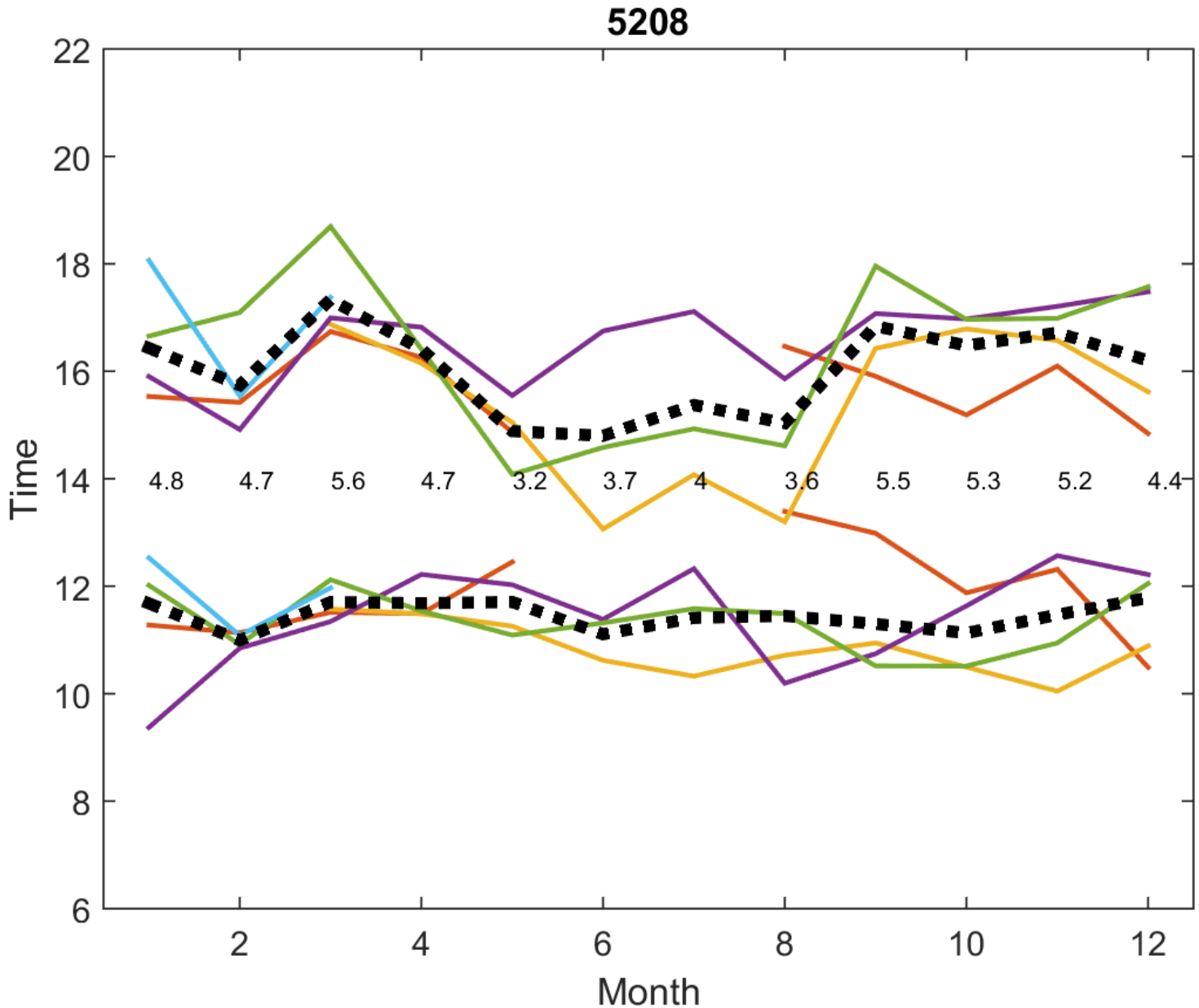
Arrival to Departure Time



Monthly averages: arrival and departure time

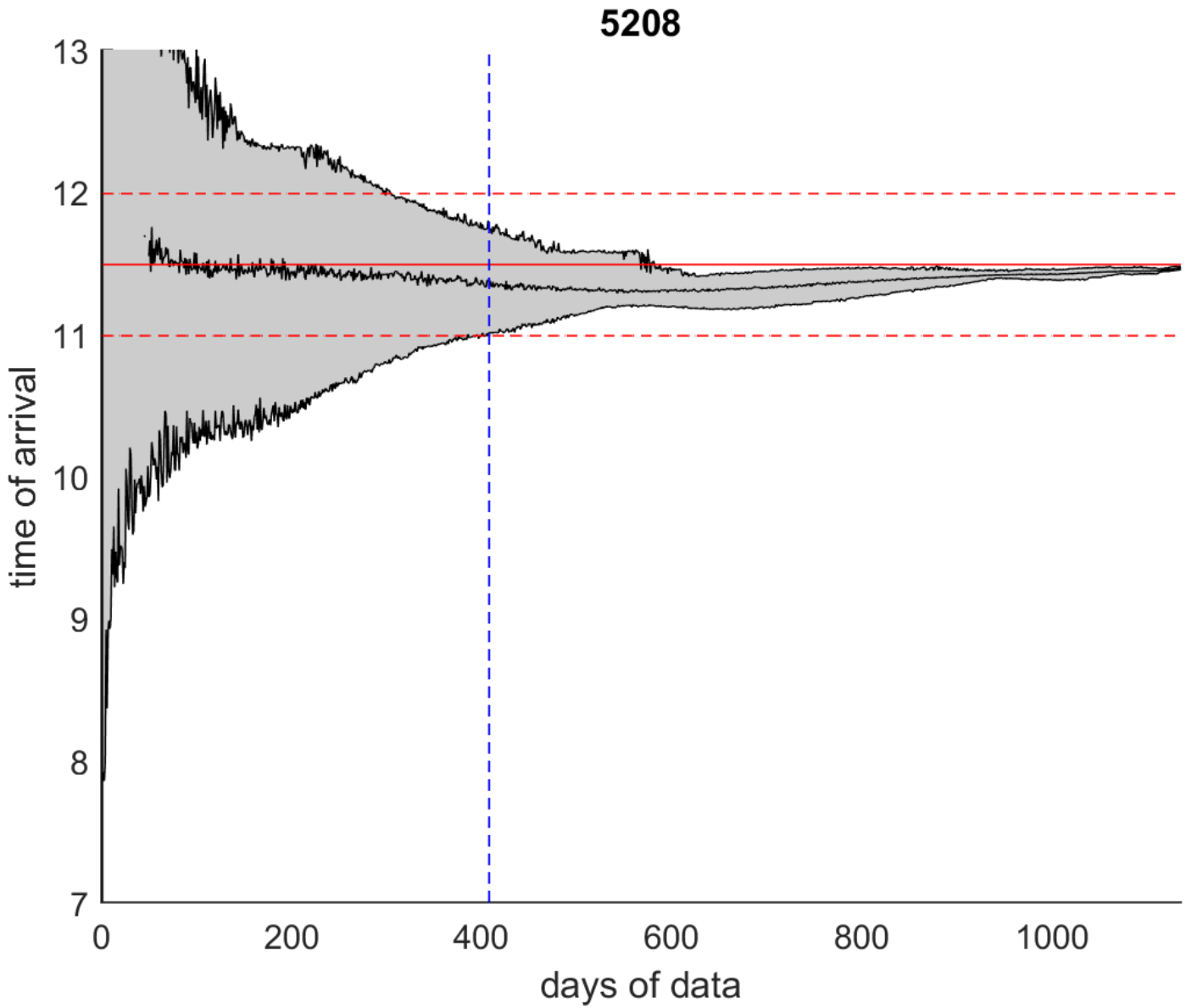


Monthly averages: arrival and departure time

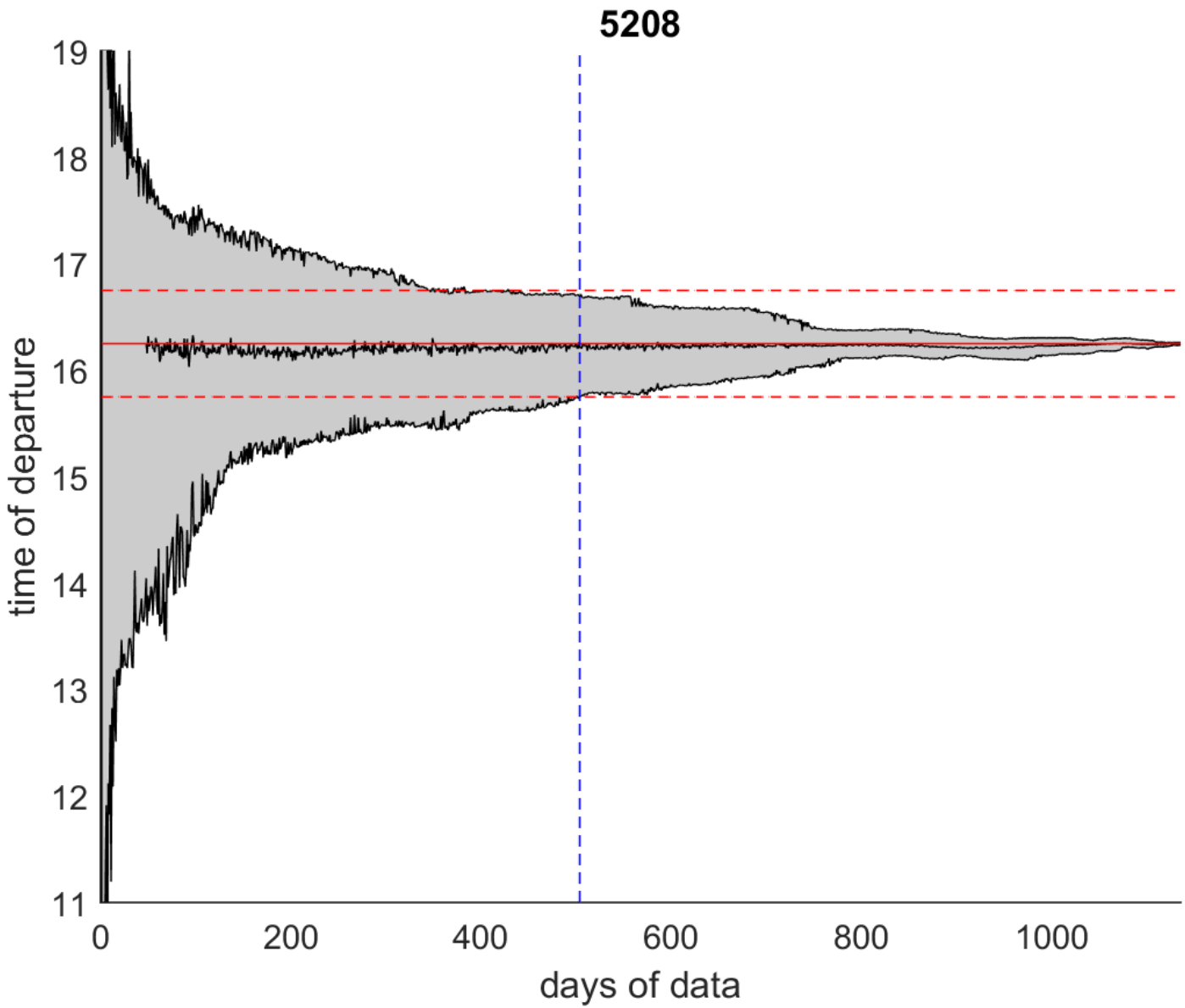


2. Subsampling & predicting key traits

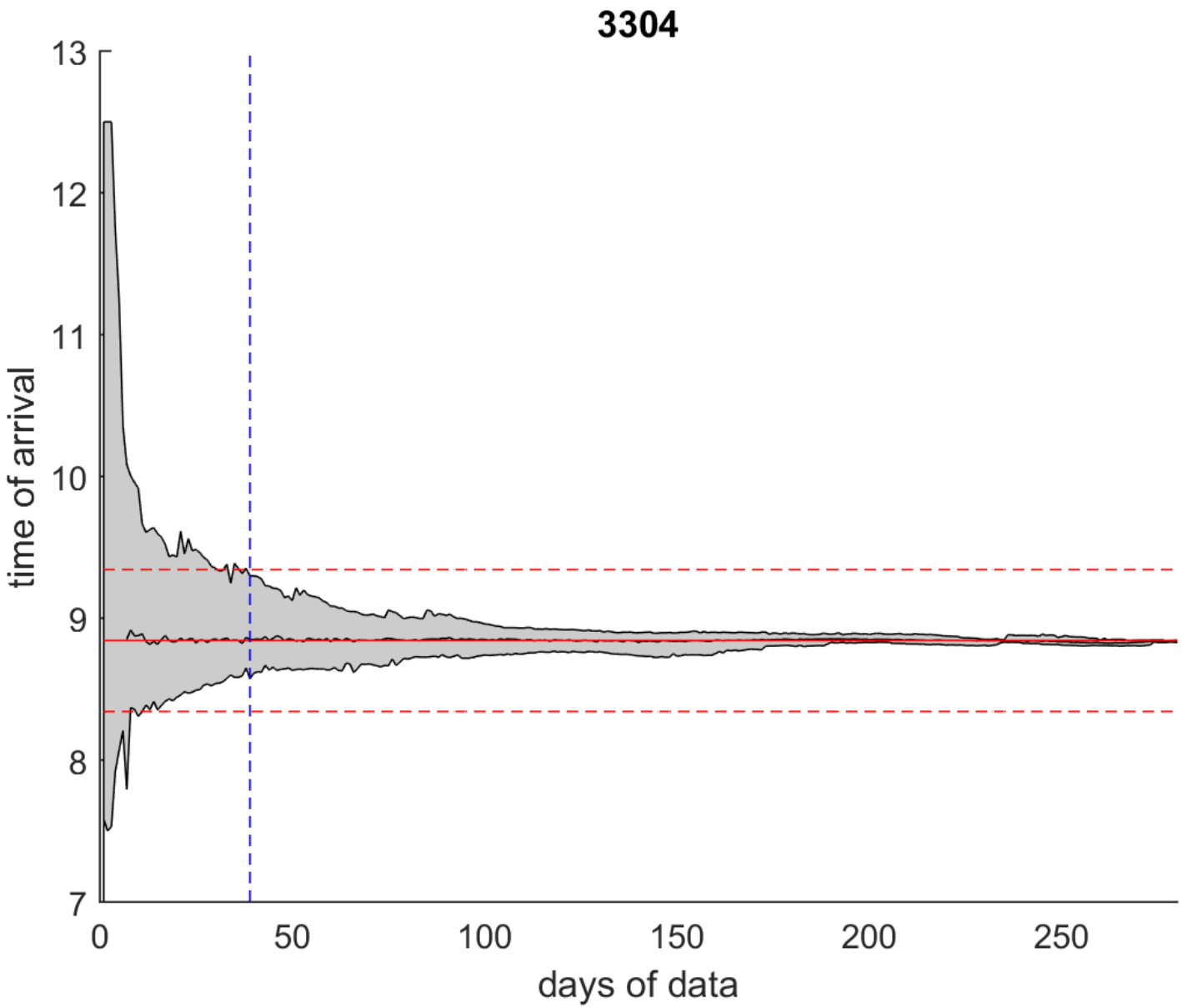
Sampling to predict arrival time: me



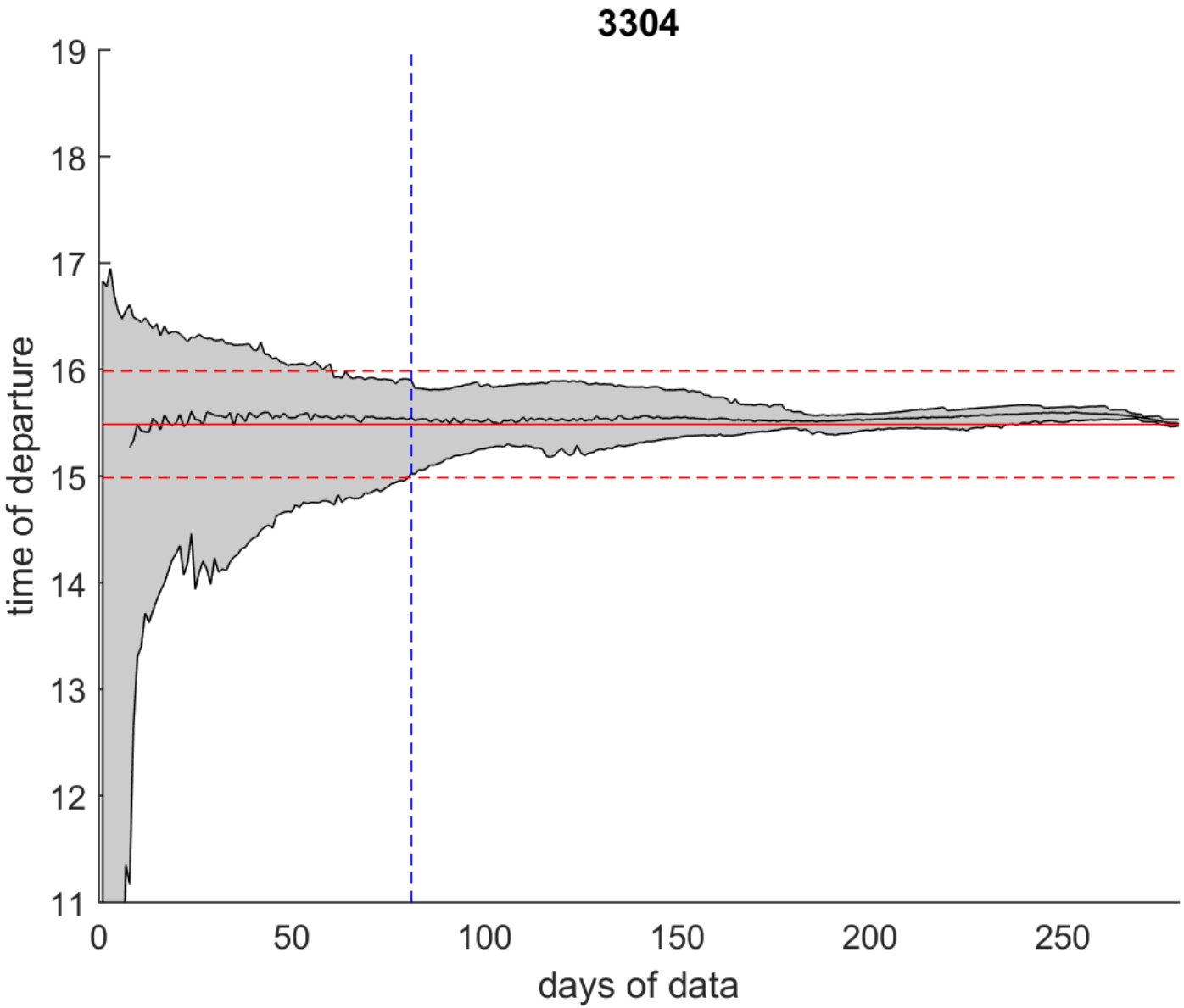
Sampling to predict departure time: me



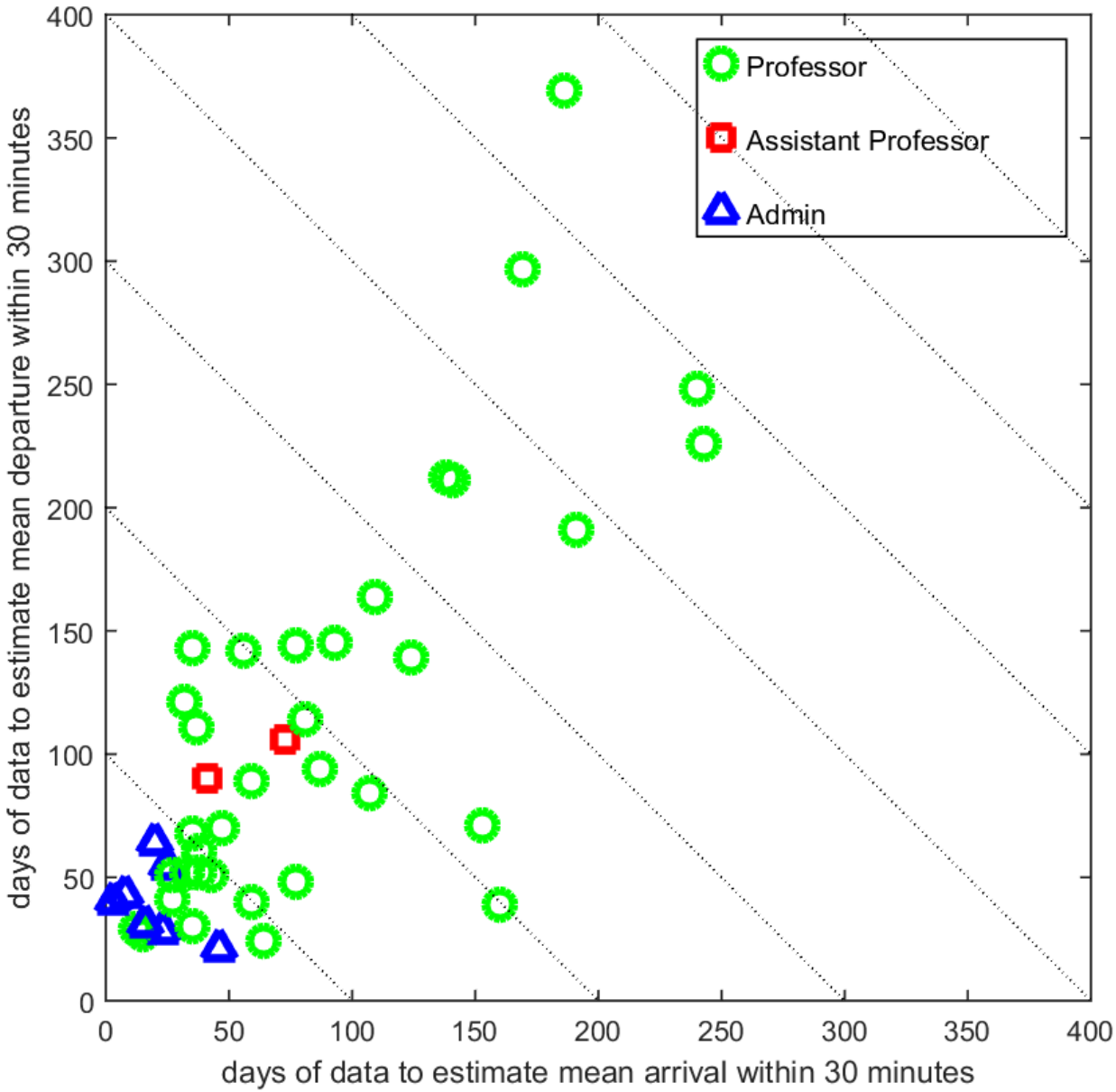
Sampling to predict arrival time: admin



Sampling to predict departure time: admin

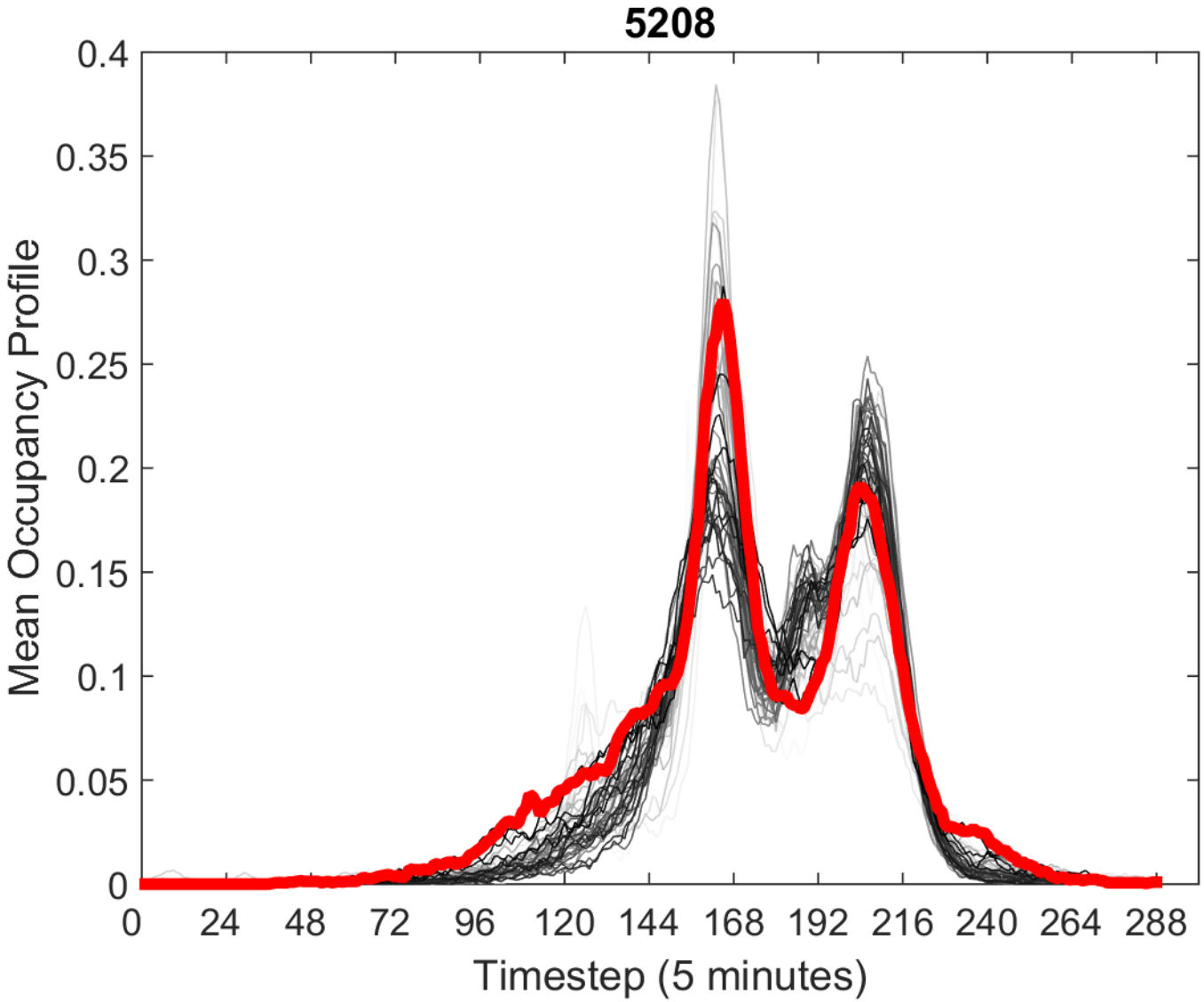


Summary of subsampling results

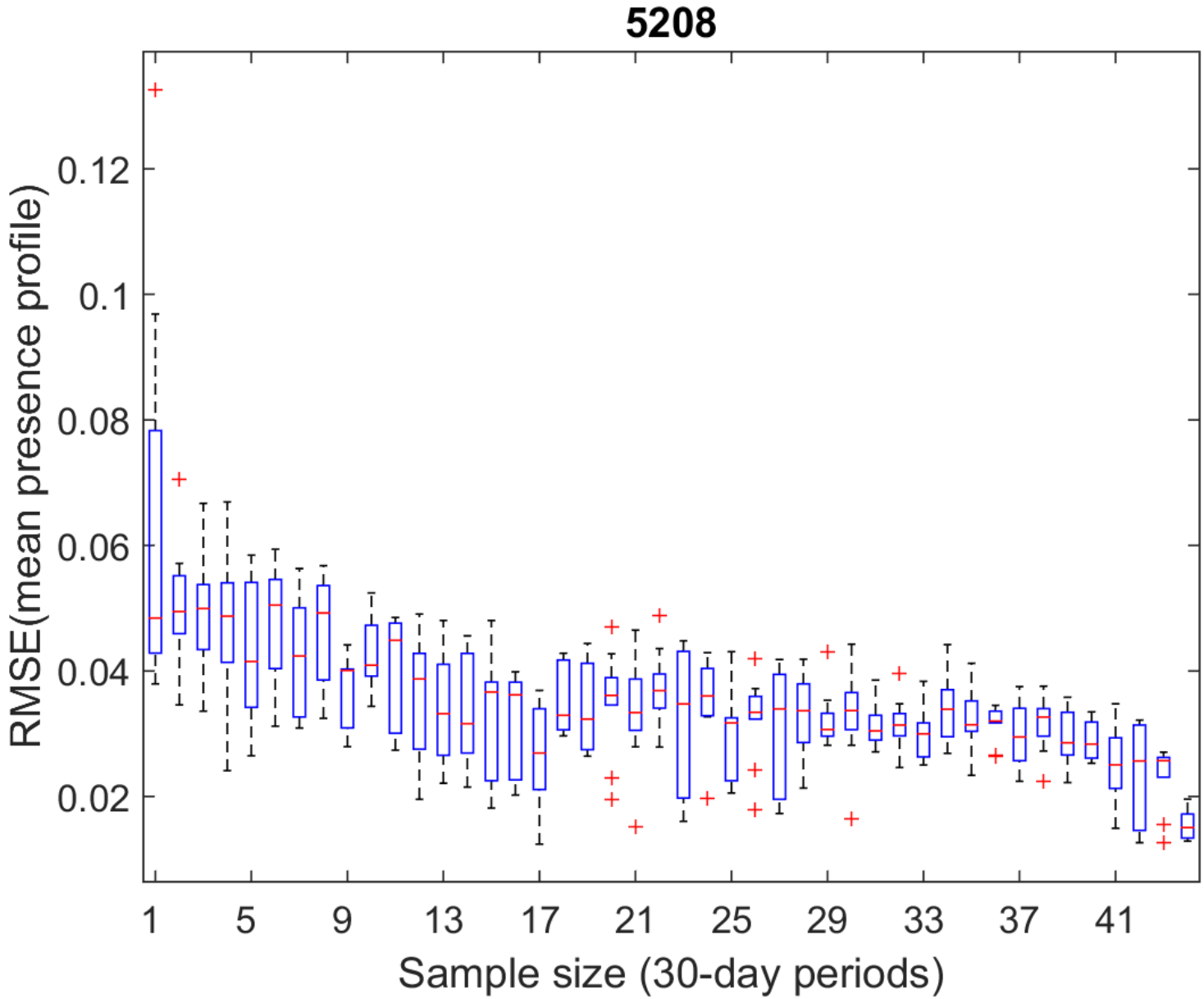


3. Subsampling & modelling

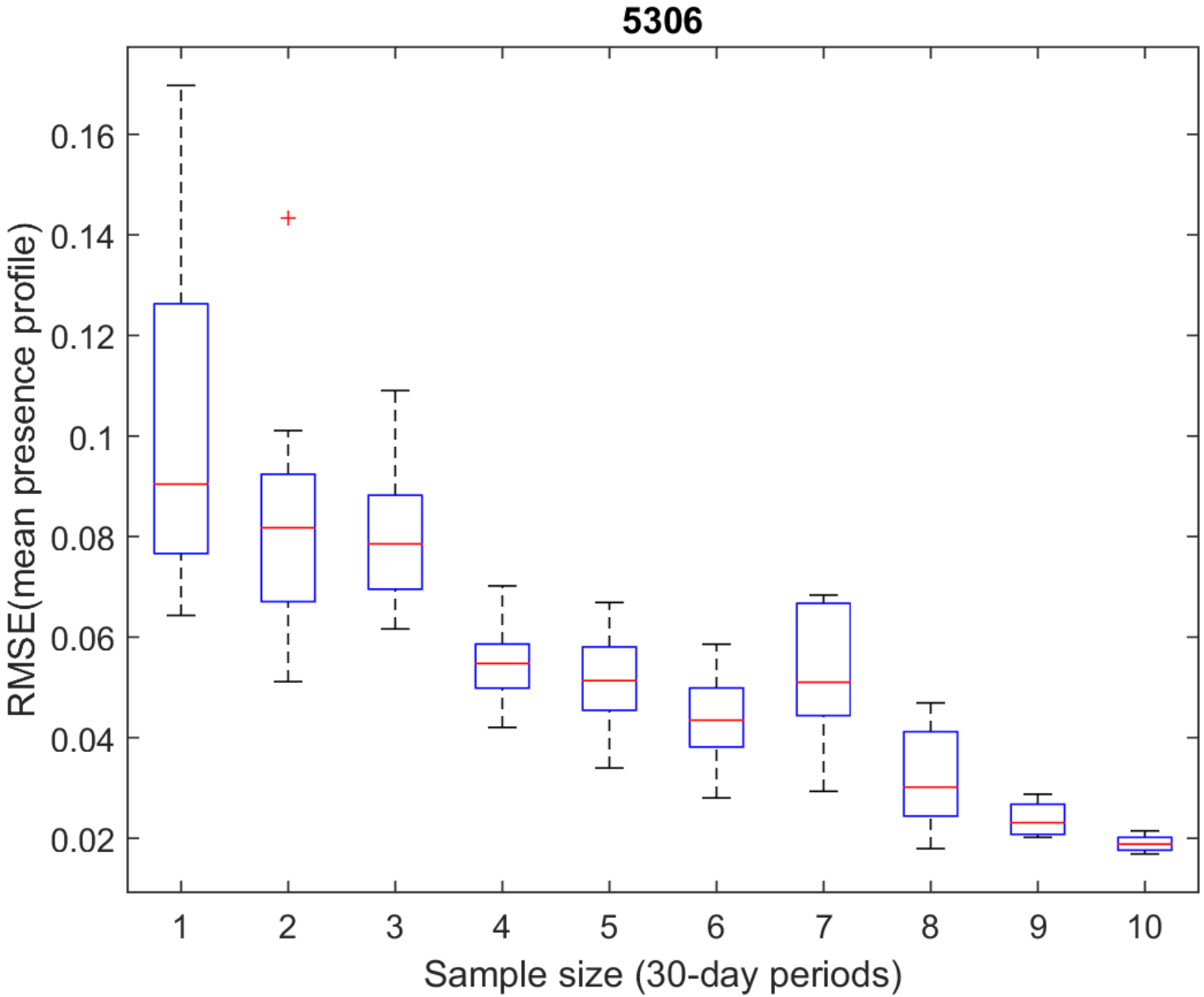
Mean occupancy profile: me



Mean occupancy profile: me



Mean occupancy profile: admin



Conclusions…so far

- Necessary study duration is highly dependent on the consistency of occupant and trait of interest
  - Mean presence profile can be learned within weeks for administrators
  - Long absences for professors may require years of data to characterize
- Clearly we do not often have the privilege of years of data
  - So we need some guidelines on approximate adequate study duration
  - Reliability metrics could be quantified during study to indicate when it can stop
- Ultimately, model accuracy requirements depends on modelling objective