Real-time prediction of bus arrival using joint models for vehicle and road state

Tom Elliott

Supervisor: Professor Thomas Lumley



Overview

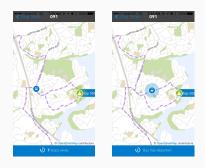
- 1. A quick motivation
- 2. Two real-time models: vehicle (particle filter) & road (Kalman filter)
- 3. Predicting arrival times

Prediction inaccuracy

- Prediction inaccuracy
- Prone to errors



- Prediction inaccuracy
- Prone to errors



- Prediction inaccuracy
- Prone to errors



- Prediction inaccuracy
- Prone to errors



- Prediction inaccuracy
- Prone to errors
- Recent modelling frameworks <u>don't</u> make use of all real-time vehicle data



 $\textbf{Goal:} \ \, \text{use observations of bus location (GPS)} \, \dots$

Goal: use observations of bus location (GPS) . . .

$$\mathbf{Y}_k = egin{bmatrix} \phi_k \\ \lambda_k \\ t_k \end{bmatrix} = egin{bmatrix} ext{latitude (degrees)} \\ ext{longitude (degrees)} \\ ext{timestamp} \end{bmatrix}$$

Goal: use observations of bus location (GPS) . . .

$$\mathbf{Y}_k = egin{bmatrix} \phi_k \\ \lambda_k \\ t_k \end{bmatrix} = egin{bmatrix} ext{latitude (degrees)} \\ ext{longitude (degrees)} \\ ext{timestamp} \end{bmatrix}$$

... to infer unobservable vehicle state ...

Goal: use observations of bus location (GPS) ...

$$\mathbf{Y}_k = egin{bmatrix} \phi_k \\ \lambda_k \\ t_k \end{bmatrix} = egin{bmatrix} ext{latitude (degrees)} \\ ext{longitude (degrees)} \\ ext{timestamp} \end{bmatrix}$$

... to infer unobservable vehicle state ...

$$\mathbf{X}_k = egin{bmatrix} d_k \\ v_k \\ s_k \\ \vdots \end{bmatrix} = egin{bmatrix} \operatorname{distance into trip (meters)} \\ \operatorname{velocity/speed } (ms^{-1}) \\ \operatorname{last visited stop} \\ \vdots \end{bmatrix}$$

Goal: use observations of bus location (GPS) ...

$$\mathbf{Y}_k = egin{bmatrix} \phi_k \\ \lambda_k \\ t_k \end{bmatrix} = egin{bmatrix} ext{latitude (degrees)} \\ ext{longitude (degrees)} \\ ext{timestamp} \end{bmatrix}$$

... to infer unobservable vehicle state ...

$$\mathbf{X}_k = egin{bmatrix} d_k \\ v_k \\ s_k \\ \vdots \end{bmatrix} = egin{bmatrix} \operatorname{distance into trip (meters)} \\ \operatorname{velocity/speed } (ms^{-1}) \\ \operatorname{last visited stop} \\ \vdots \end{bmatrix}$$

...in real time.



Example: Route 274, Britomart to Three Kings

- Represent \mathbf{X}_k by a <u>sample</u> of point-estimates (particles) $\mathbf{x}_k^{(i)}$
- Flexible modeling framework, fewer assumptions
- Better coverage of possible states (multimodality)
- Intuitive likelihood function

- Represent \mathbf{X}_k by a <u>sample</u> of point-estimates (particles) $\mathbf{x}_k^{(i)}$
- Flexible modeling framework, fewer assumptions
- Better coverage of possible states (multimodality)
- Intuitive likelihood function

Step 1: predict

 \Rightarrow generate sample of possible vehicle states

- Represent \mathbf{X}_k by a <u>sample</u> of point-estimates (particles) $\mathbf{x}_k^{(i)}$
- Flexible modeling framework, fewer assumptions
- Better coverage of possible states (multimodality)
- Intuitive likelihood function

Step 1: predict

 \Rightarrow generate sample of possible vehicle states

Step 2: update

 \Rightarrow compare predictions to observation, remove those no longer plausible

Step 1: predict

- Start with "known" vehicle state, $\mathbf{X}_{k-1} = \{\mathbf{x}_{k-1}^{(i)}: i=1,\dots,N\}$

- Start with "known" vehicle state, $\mathbf{X}_{k-1} = \{\mathbf{x}_{k-1}^{(i)}: i=1,\ldots,N\}$
- Transition each particle independently

$$\mathbf{x}_k^{(i)} = f(\mathbf{x}_{k-1}^{(i)}, \sigma_v^2, \ldots)$$

Step 1: predict

- Start with "known" vehicle state, $\mathbf{X}_{k-1} = \{\mathbf{x}_{k-1}^{(i)}: i=1,\ldots,N\}$
- Transition each particle independently

$$\mathbf{x}_k^{(i)} = f(\mathbf{x}_{k-1}^{(i)}, \frac{\sigma_v^2}{\sigma_v}, \ldots)$$

1. Add system noise

$$v_k^{(i)} \sim \mathcal{N}_T(v_{k-1}^{(i)}, \frac{\sigma_v^2}{\sigma_v}), \qquad 0 \leq v_k^{(i)} \leq \mathbf{v}_{\text{max}}$$

 $\textbf{v}_{\text{max}} \approx \text{road speed limit}$

- Start with "known" vehicle state, $\mathbf{X}_{k-1} = \{\mathbf{x}_{k-1}^{(i)}: i=1,\ldots,N\}$
- Transition each particle independently

$$\mathbf{x}_k^{(i)} = f(\mathbf{x}_{k-1}^{(i)}, \sigma_v^2, \ldots)$$

- 1. Add system noise
- 2. Move particles along route (Law of Motion)

$$d_k^{(i)} = d_{k-1}^{(i)} + (t_k - t_{k-1})v_k^{(i)}$$

Step 1: predict

- Start with "known" vehicle state, $\mathbf{X}_{k-1} = \{\mathbf{x}_{k-1}^{(i)}: i=1,\ldots,N\}$
- Transition each particle independently

$$\mathbf{x}_k^{(i)} = f(\mathbf{x}_{k-1}^{(i)}, \sigma_v^2, \ldots)$$

- 1. Add system noise
- 2. Move particles along route

$$d_k^{(i)} \le d_{k-1}^{(i)} + (t_k - t_{k-1})v_k^{(i)}$$

3. What about intermediate stop(s)?

- Start with "known" vehicle state, $\mathbf{X}_{k-1} = \{\mathbf{x}_{k-1}^{(i)}: i=1,\ldots,N\}$
- Transition each particle independently

$$\mathbf{x}_k^{(i)} = f(\mathbf{x}_{k-1}^{(i)}, \sigma_v^2, \ldots)$$

- 1. Add system noise
- 2. Move particles along route
- 3. What about intermediate stop(s)?
 - Does the particle stop? $p_{s_k}^{(i)} \sim \operatorname{Bernoulli}(\pi_{s_k})$

- Start with "known" vehicle state, $\mathbf{X}_{k-1} = \{\mathbf{x}_{k-1}^{(i)} : i = 1, \dots, N\}$
- Transition each particle independently

$$\mathbf{x}_k^{(i)} = f(\mathbf{x}_{k-1}^{(i)}, \sigma_v^2, \ldots)$$

- 1. Add system noise
- 2. Move particles along route
- 3. What about intermediate stop(s)?
 - Does the particle stop? $p_{s_k}^{(i)} \sim \text{Bernoulli}(\pi_{s_k})$
 - If so, for how long? $\overline{t}_{s_k} \sim \mathcal{E}(au_{s_k})$

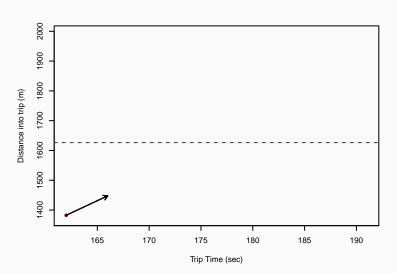
- Start with "known" vehicle state, $\mathbf{X}_{k-1} = \{\mathbf{x}_{k-1}^{(i)}: i=1,\ldots,N\}$
- Transition each particle independently

$$\mathbf{x}_k^{(i)} = f(\mathbf{x}_{k-1}^{(i)}, \sigma_v^2, \ldots)$$

- 1. Add system noise
- 2. Move particles along route
- 3. What about intermediate stop(s)?
 - Does the particle stop? $p_{s_k}^{(i)} \sim \mathrm{Bernoulli}(\pi_{s_k})$
 - If so, for how long? $\overline{t}_{s_k} \sim \mathcal{E}(au_{s_k})$
 - **Dwell time** = $p_{s_k}^{(i)}(\gamma + \overline{t}_{s_k})$ $\gamma =$ minimum dwell time (deccelerate/accelerate, open/close doors) $\overline{t}_{s_k} =$ passengers on/off

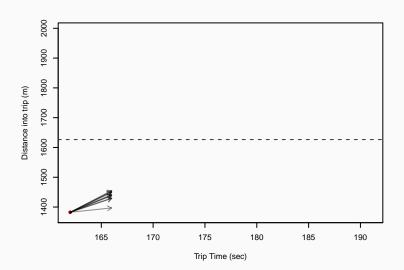
Step 1: predict

Example (N = 10 particles)



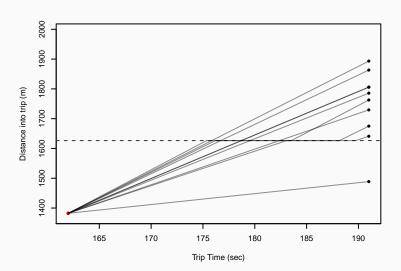
Step 1: predict

Example (N = 10 particles)



Step 1: predict

Example (N = 10 particles)



Step 2: update

Step 2: update

• Likelihood function $\ell(\mathbf{Y}_k|\mathbf{x}_k^{(i)})$

Step 2: update

- Likelihood function $\ell(\mathbf{Y}_k|\mathbf{x}_k^{(i)})$
 - Transform particles onto flat plane

$$\mathbf{z}_k^{(i)} = g(h(\mathbf{x}_k^{(i)})|\mathbf{Y}_k)$$

Step 2: update

- Likelihood function $\ell(\mathbf{Y}_k|\mathbf{x}_k^{(i)}, \mathbf{h})$
 - Transform particles onto flat plane

$$\mathbf{z}_k^{(i)} = g(\mathbf{h}(\mathbf{x}_k^{(i)})|\mathbf{Y}_k)$$

h: measurement function (distance into trip \rightarrow lat/lon)

Step 2: update

- Likelihood function $\ell(\mathbf{Y}_k|\mathbf{x}_k^{(i)},h,\mathbf{g})$
 - Transform particles onto flat plane

$$\mathbf{z}_k^{(i)} = \mathbf{g}(h(\mathbf{x}_k^{(i)})|\mathbf{Y}_k)$$

 $g(\cdot|\mathbf{Y}_k)$: projection

centered on \mathbf{Y}_k , 1 unit = 1 meter in all directions

Step 2: update

- Likelihood function $\ell(\mathbf{Y}_k|\mathbf{x}_k^{(i)},h,g)$
 - Transform particles onto flat plane

$$\mathbf{z}_k^{(i)} = g(h(\mathbf{x}_k^{(i)})|\mathbf{Y}_k)$$

• Bivariate normal likelihood, $g(Y_k|Y_k) = 0$

$$\mathbf{Y}_k | \mathbf{z}_k^{(i)} \sim \mathbf{z}_k^{(i)} | \mathbf{Y}_k \sim \mathcal{N}_2(\mathbf{0}, \sigma_y^2 I_2)$$
 $(\sigma_y^2 = \mathsf{GPS} \; \mathsf{error})$

Step 2: update

- Likelihood function $\ell(\mathbf{Y}_k|\mathbf{x}_k^{(i)},h,g)$
 - Transform particles onto flat plane

$$\mathbf{z}_k^{(i)} = g(h(\mathbf{x}_k^{(i)})|\mathbf{Y}_k)$$

• Bivariate normal likelihood, $g(Y_k|Y_k) = 0$

$$\mathbf{Y}_k | \mathbf{z}_k^{(i)} \sim \mathbf{z}_k^{(i)} | \mathbf{Y}_k \sim \mathcal{N}_2(\mathbf{0}, \sigma_y^2 l_2)$$
 $(\sigma_y^2 = \mathsf{GPS} \; \mathsf{error})$

• For each particle

$$\ell(\mathbf{Y}_k|\mathbf{x}_k^{(i)},h,g)\propto e^{-\frac{1}{2\sigma^2}\left((\mathbf{z}_k^{(i)})^T\mathbf{z}_k^{(i)}\right)}$$

Step 2: update

Likelihood function

$$\ell(\mathbf{Y}_k|\mathbf{x}_k^{(i)},h,g) \propto e^{-\frac{1}{2\sigma^2}\left((\mathbf{z}_k^{(i)})^T\mathbf{z}_k^{(i)}\right)}$$

Compute weights

$$w_i = \frac{\ell(\mathbf{Y}_k|\mathbf{x}_k^{(i)})}{\sum_{j=1}^N \ell(\mathbf{Y}_k|\mathbf{x}_k^{(j)})}$$

Step 2: update

Likelihood function

$$\ell(\mathbf{Y}_k|\mathbf{x}_k^{(i)},h,g) \propto e^{-\frac{1}{2\sigma^2}\left((\mathbf{z}_k^{(i)})^T\mathbf{z}_k^{(i)}\right)}$$

Compute weights

$$w_i = \frac{\ell(\mathbf{Y}_k | \mathbf{x}_k^{(i)})}{\sum_{j=1}^N \ell(\mathbf{Y}_k | \mathbf{x}_k^{(j)})}$$

Weighted resample with replacement

Step 2: update

Likelihood function

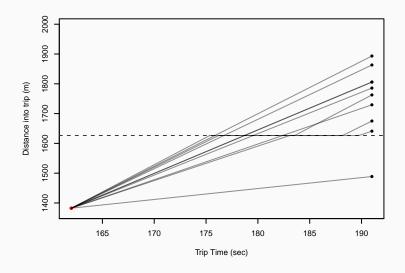
$$\ell(\mathbf{Y}_k|\mathbf{x}_k^{(i)},h,g)\propto e^{-\frac{1}{2\sigma^2}\left((\mathbf{z}_k^{(i)})^T\mathbf{z}_k^{(i)}\right)}$$

Compute weights

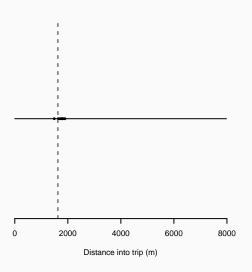
$$w_i = \frac{\ell(\mathbf{Y}_k | \mathbf{x}_k^{(i)})}{\sum_{j=1}^{N} \ell(\mathbf{Y}_k | \mathbf{x}_k^{(j)})}$$

- Weighted resample with replacement
 - \Rightarrow keep particles plausible given data

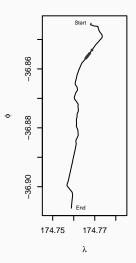
Step 2: update



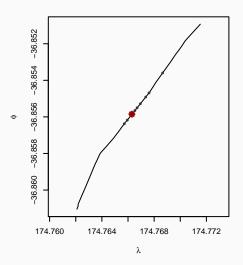
Step 2: update



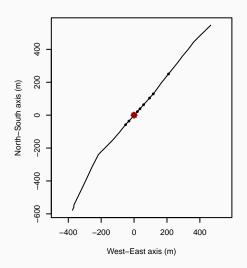
Step 2: update



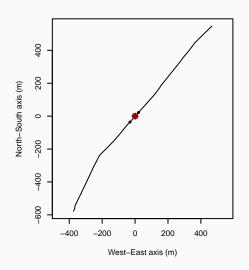
Step 2: update



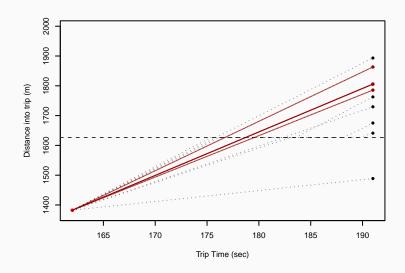
Step 2: update



Step 2: update



Step 2: update



1. Particle filter \Rightarrow speed estimates for a given bus

- 1. Particle filter \Rightarrow speed estimates for a given bus
- 2. Identify segments of road common to multiple routes

- 1. Particle filter \Rightarrow speed estimates for a given bus
- 2. Identify segments of road common to multiple routes
- 3. Use speed information from <u>all busses</u> to estimate speed along a given road segment

- 1. Particle filter ⇒ speed estimates for a given bus
- 2. Identify segments of road common to multiple routes
- 3. Use speed information from <u>all busses</u> to estimate speed along a given road segment

2. Identify segments of road common to multiple routes

2. Identify segments of road common to multiple routes ⇒ between intersections



2. Identify segments of road common to multiple routes ⇒ between intersections



3. Use speed information from <u>all busses</u> to estimate speed along a given road segment

3. Use speed information from <u>all busses</u> to estimate speed along a given road segment

⇒ Kalman filter

3. Use speed information from <u>all busses</u> to estimate speed along a given road segment

⇒ Kalman filter

Road state: mean speed for all M road segments at time t_ℓ

$$\boldsymbol{\nu}_{\ell} = \begin{bmatrix} \nu_{1\ell} \ \nu_{2\ell} \ \cdots \ \nu_{M\ell} \end{bmatrix}^{\mathsf{T}}$$

with associated covariance matrix

$$\Xi_{\ell} = \begin{bmatrix} \xi_{1\ell} & 0 & \cdots & 0 \\ 0 & \xi_{2\ell} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \xi_{M\ell} \end{bmatrix}$$

3. Use speed information from <u>all busses</u> to estimate speed along a given road segment

⇒ Kalman filter

no complex model necessary (Normal distribution adequate)

3. Use speed information from <u>all busses</u> to estimate speed along a given road segment

⇒ Kalman filter

- no complex model necessary (Normal distribution adequate)
- updated using particle filter estimates

$$v_\ell = \nu_\ell + r_\ell$$

- v_ℓ: mean speed of particles
- $r_\ell \sim \mathcal{N}(0,R_\ell)$, R_ℓ : variance of particle speeds

- 1. Schedule
- 2. Schedule deviation (AT?)
- 3. Vehicle state
- 4. Road state

- 1. Schedule
- 2. Schedule deviation (AT?)
- 3. Vehicle state
- 4. Road state

Some notation:

- $S_j^t =$ scheduled arrival time at stop j
- $\hat{A}_j = (\text{predicted})$ arrival time at stop j
- $\tilde{T}^a_{s_k}, \tilde{T}^d_{s_k} =$ observed arrival/departure delay at last stop (from Auckland Transport's API)

1. Schedule

- $\hat{A}_j = S_j^t$
- Baseline for other predictors

2. Schedule deviation

$$\hat{A}_j = \begin{cases} S_j^t + \tilde{T}_{s_k}^d & \text{if departed stop } s_k \\ S_j^t + \tilde{T}_{s_k}^a & \text{if not departed stop } s_k \end{cases}$$

• **OR** use particle estimates of arrival/departure delay, $\tilde{A}_{s_k}^{(i)}$ and $\tilde{D}_{s_k}^{(i)}$

3. Vehicle state

• $S_j^d = \text{distance along route of stop } j$

- S_j^d = distance along route of stop j
- $\hat{A}_j = t_k + \frac{S_j^d d_k}{v_k}$

- S_j^d = distance along route of stop j
- $\hat{A}_{j}^{(i)} = t_{k} + \frac{S_{j}^{d} d_{k}^{(i)}}{v_{k}^{(i)}}$
- Prediction for each particle

- S_j^d = distance along route of stop j
- $\hat{A}_{j}^{(i)} = t_{k} + \frac{S_{j}^{d} d_{k}^{(i)}}{v_{k}^{(i)}} + \sum_{\ell=s_{k}+1}^{j-1} p_{\ell}^{(i)} (\gamma + \overline{t}_{\ell}^{(i)})$
- Prediction for each particle
- Allow for dwell time uncertainty

- S_j^d = distance along route of stop j
- $\hat{A}_j^{(i)} = t_k + \frac{S_j^d d_k^{(i)}}{v_k^{(i)}} + \sum_{\ell=s_k+1}^{j-1} p_\ell^{(i)} (\gamma + \overline{t}_\ell^{(i)})$
- Prediction for each particle
- Allow for dwell time uncertainty
- Potentially multi-modal

4. Road state

r_k = route segment index
 R_b = distance along route of start of segment b

- r_k = route segment index
 R_b = distance along route of start of segment b
- $\bullet \hat{A}_j = t_k +$

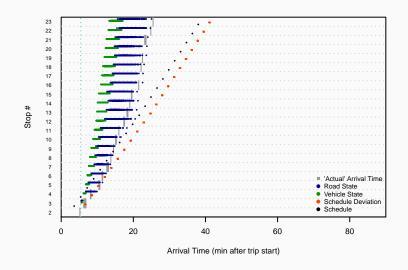
- r_k = route segment index
 R_b = distance along route of start of segment b
- $\bullet \hat{A}_j = t_k + \frac{R_{s_k+1} d_k}{\nu_{s_k}} +$
- travel time until end of current segment

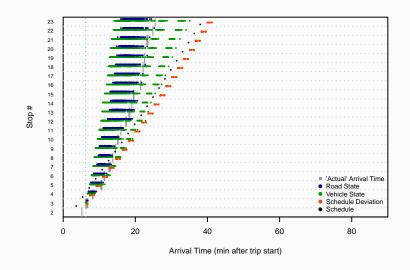
- r_k = route segment index
 R_b = distance along route of start of segment b
- $\hat{A}_j = t_k + \frac{R_{s_k+1} d_k}{\nu_{s_k}} + \sum_{b \in B^*} \frac{R_{b+1} R_b}{\nu_b} + \frac{R_{b+1} R_b}{\nu_b}$
- travel time through intermediate segments (B^*)

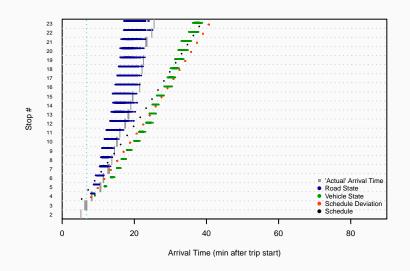
- r_k = route segment index
 R_b = distance along route of start of segment b
- $\hat{A}_j = t_k + \frac{R_{s_k+1} d_k}{\nu_{s_k}} + \sum_{b \in B^*} \frac{R_{b+1} R_b}{\nu_b} + \frac{S_j^d R_{b'}}{\nu_{b'}}$
- lacktriangleright travel time along segment b' to stop j

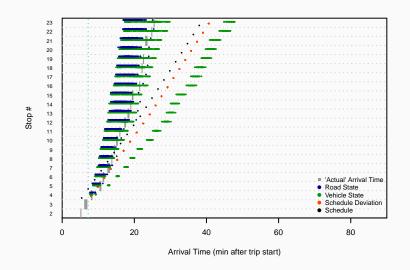
- r_k = route segment index
 R_b = distance along route of start of segment b
- $\hat{A}_{j}^{(i)} = t_k + \frac{R_{s_k(i)+1} d_k(i)}{\nu_{s_k(i)}} + \sum_{b \in B^*} \frac{R_{b+1} R_b}{\nu_b^{(i)}} + \frac{S_j^d R_{b'}}{\nu_{b'}^{(i)}}$
- For each particle, sample $v_b^{(i)} \sim \mathcal{N}(\nu_b, \xi_b)$

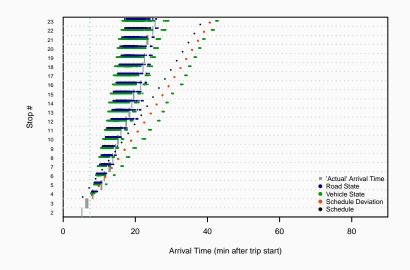
- r_k = route segment index
 R_b = distance along route of start of segment b
- $\hat{A}_{j} = t_{k} + \frac{R_{s_{k}(i)+1} d_{k}^{(i)}}{v_{s_{k}(i)}} + \sum_{b \in B^{*}} \frac{R_{b+1} R_{b}}{v_{b}^{(i)}} + \frac{S_{j}^{d} R_{b'}}{v_{b'}^{(i)}} + \frac{\sum_{\ell=s_{k}+1}^{j-1} p_{\ell}^{(i)} (\gamma + \overline{t}_{\ell}^{(i)})}{t_{\ell}^{(i)}}$
- For each particle, sample $v_b^{(i)} \sim \mathcal{N}(\nu_b, \xi_b)$
- Allow for dwell time uncertainty

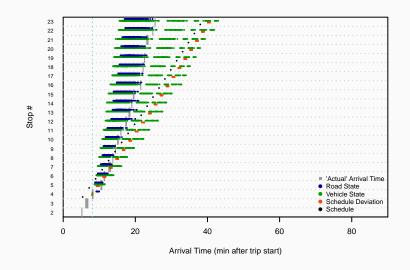


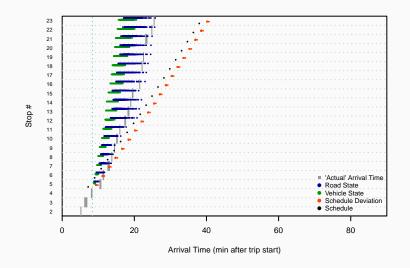


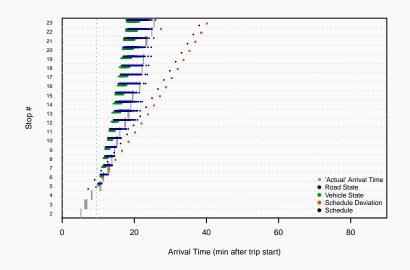


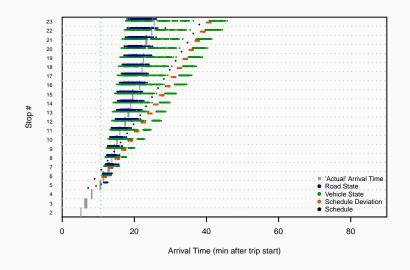


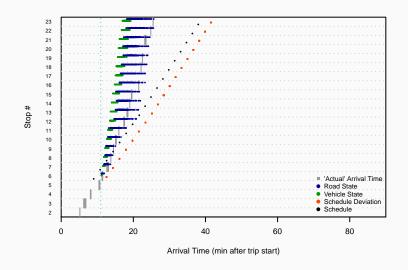


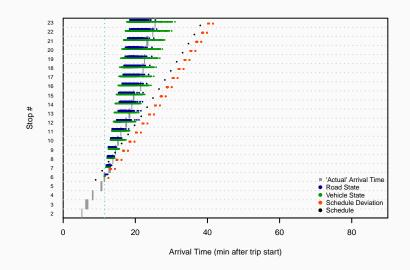


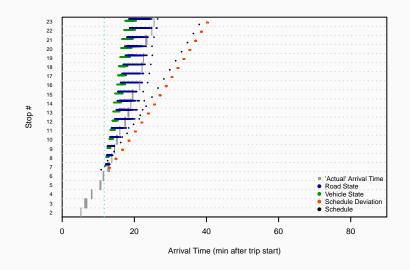


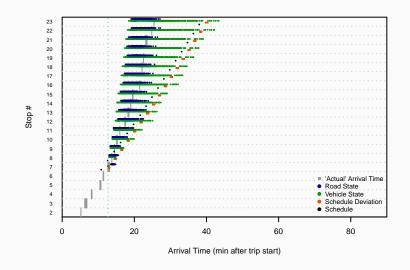


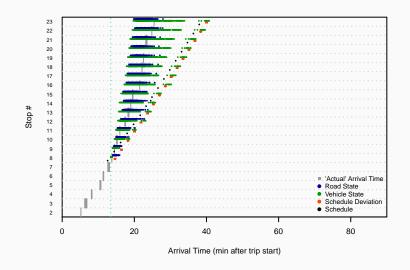


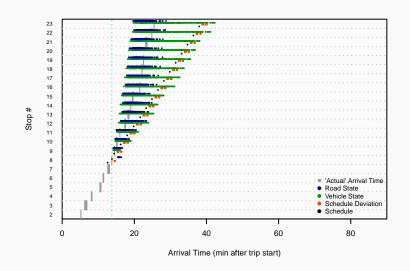


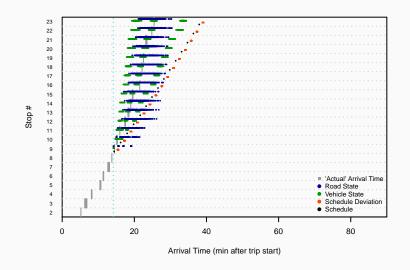


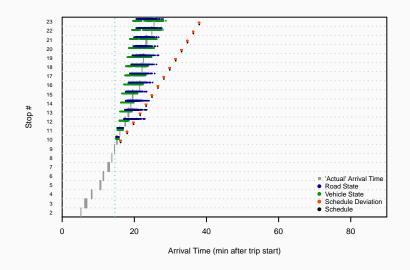


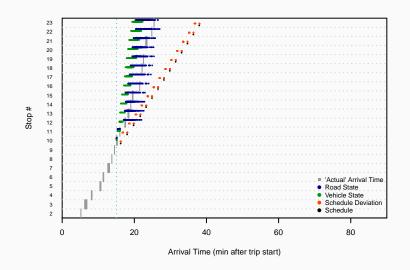


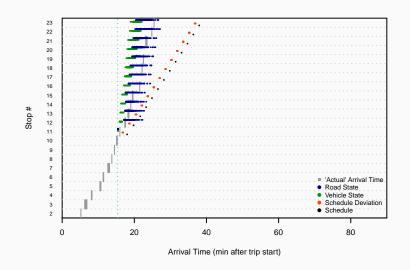


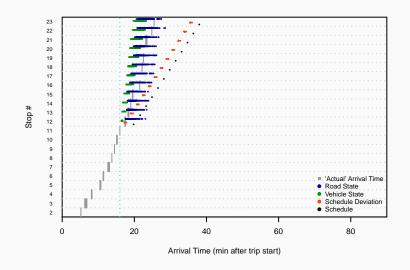


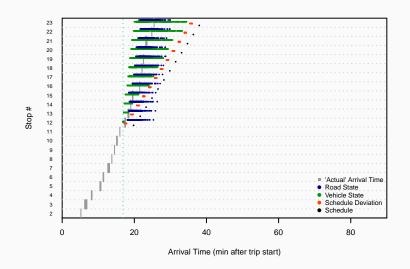


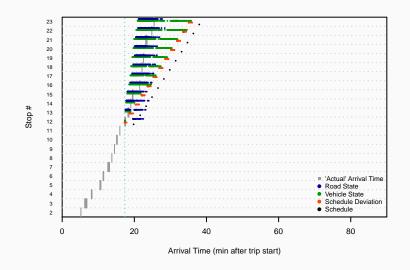


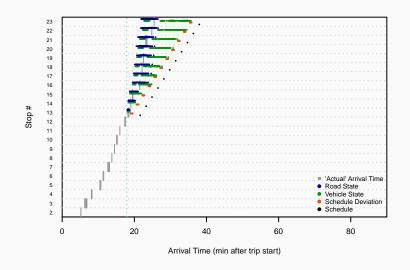


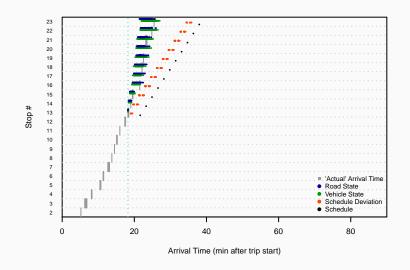


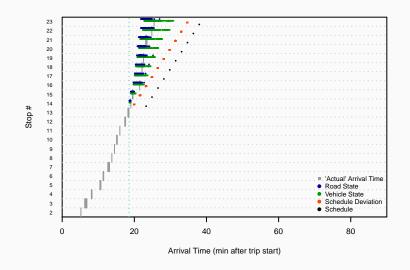


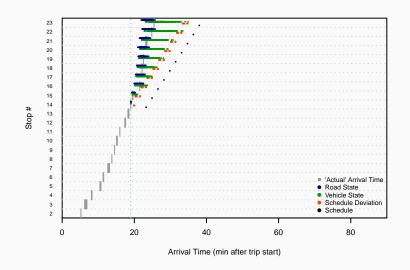


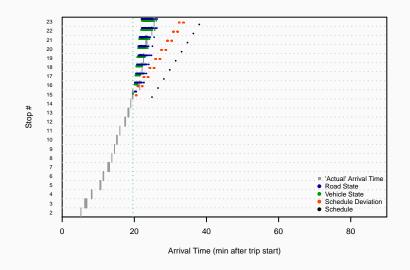


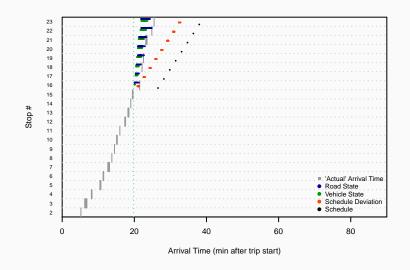


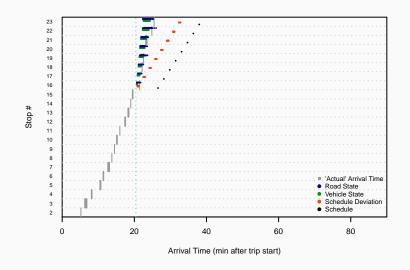


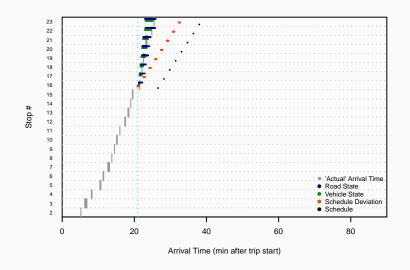


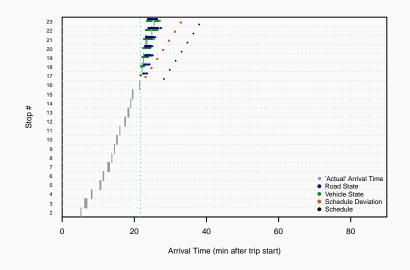


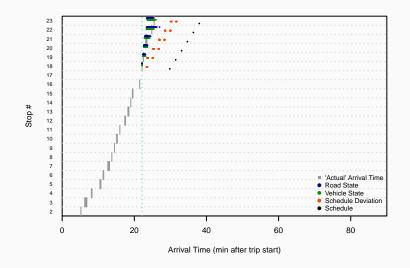


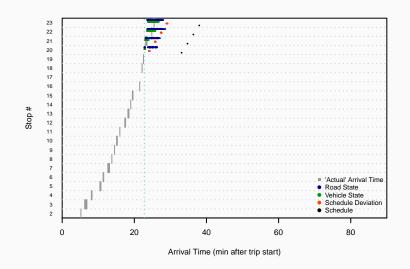


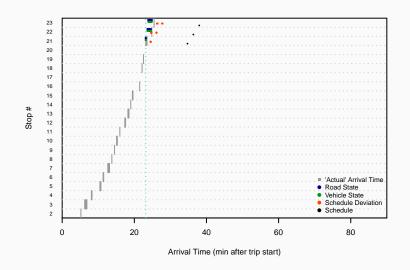


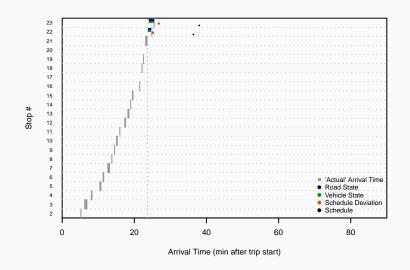


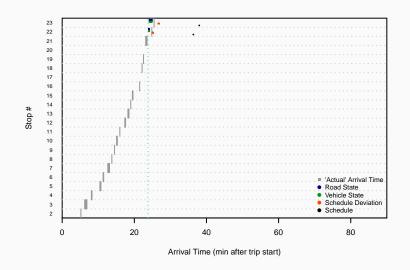


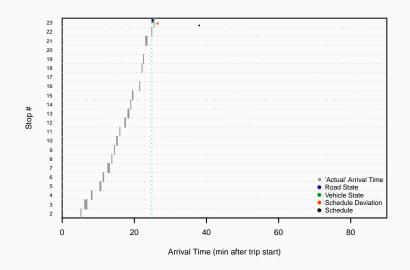


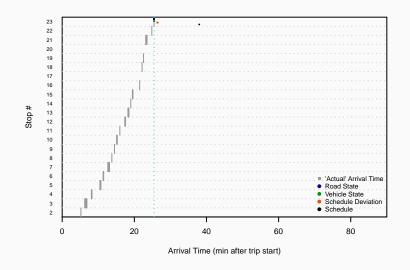












Conclusions:

• Schedule: ...

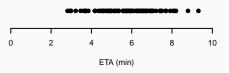
- Schedule: ...
- Schedule deviation: OK for very short-range prediction; relies on time table accuracy

- Schedule: ...
- Schedule deviation: OK for very short-range prediction; relies on time table accuracy
- Vehicle state: high variability; OK for short-range prediction

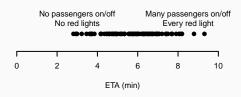
- Schedule: ...
- Schedule deviation: OK for very short-range prediction; relies on time table accuracy
- Vehicle state: high variability; OK for short-range prediction
- Road state: little variability; performs well at long-range prediction

- Schedule: ...
- Schedule deviation: OK for very short-range prediction; relies on time table accuracy
- Vehicle state: high variability; OK for short-range prediction
- Road state: little variability; performs well at long-range prediction

How do we communicate estimate + uncertainty to commuters?

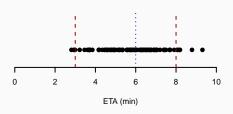


How do we communicate estimate + uncertainty to commuters?



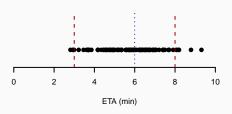
How do we communicate estimate + uncertainty to commuters?

⇒ Prediction intervals



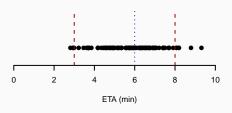
How do we communicate estimate + uncertainty to commuters?

- **⇒ Prediction intervals**
 - easy to compute from particle sample



How do we communicate estimate + uncertainty to commuters?

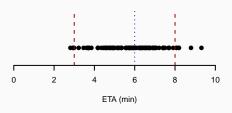
- **⇒ Prediction intervals**
 - easy to compute from particle sample
 - intuitive: ETA 6 min (mean) versus ETA 3-8 min



How do we communicate estimate + uncertainty to commuters?

⇒ Prediction intervals

- easy to compute from particle sample
- intuitive: ETA 6 min (mean) versus ETA 3–8 min
- Biased to reduce chance of missing bus



Add more routes

- Add more routes
- Collect historical data to estimate parameters

- Add more routes
- Collect historical data to estimate parameters
 - Dwell times and stopping probabilties

- Add more routes
- Collect historical data to estimate parameters
 - Dwell times and stopping probabilties
 - Segment speed covariance matrix (including off-diagonals)

- Add more routes
- Collect historical data to estimate parameters
 - Dwell times and stopping probabilties
 - Segment speed covariance matrix (including off-diagonals)
- Scale up: $\underline{\mathsf{ALL}}$ routes/busses \Rightarrow computational speed

Thank you!

