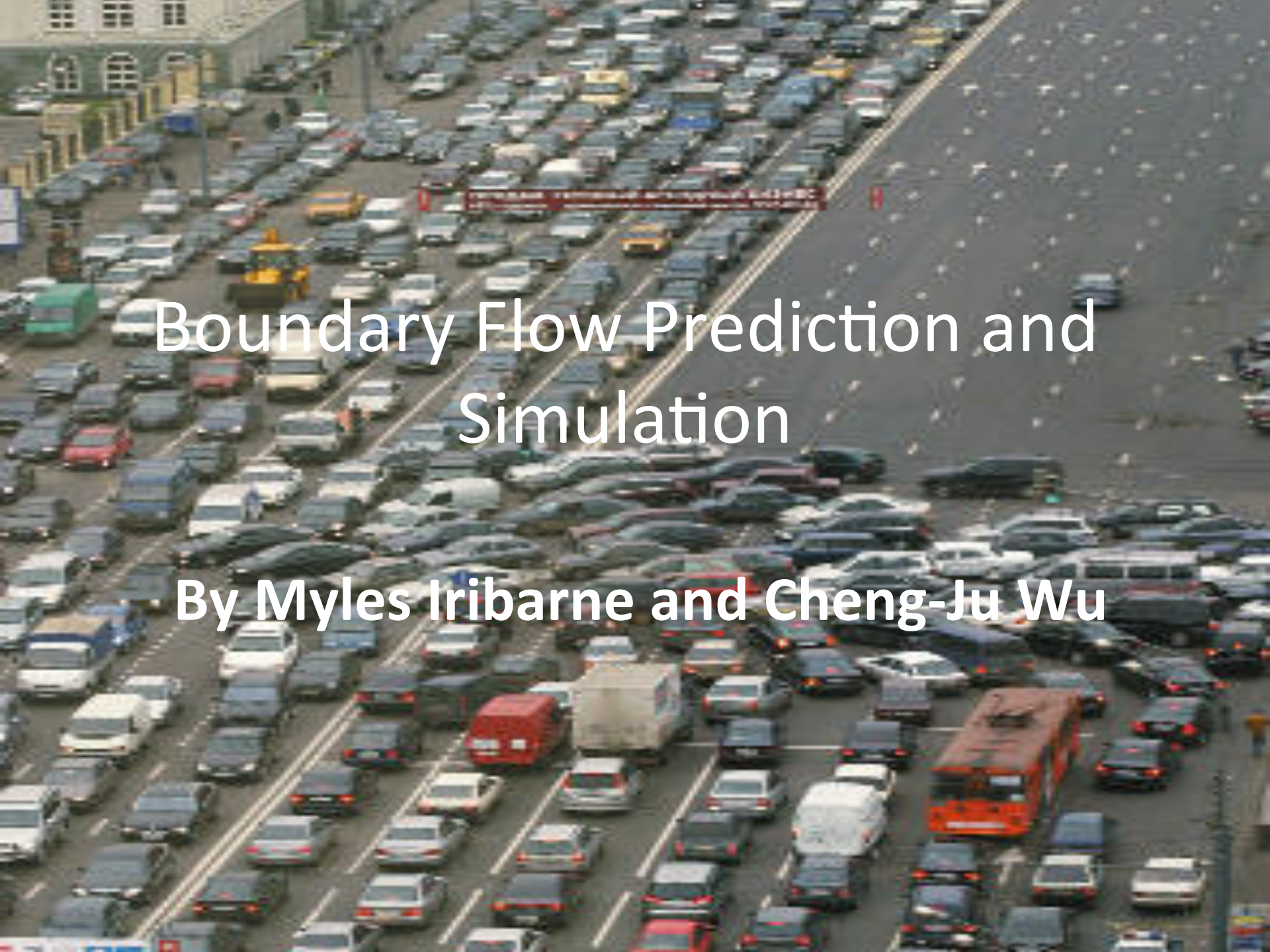


Boundary Flow Prediction and Simulation

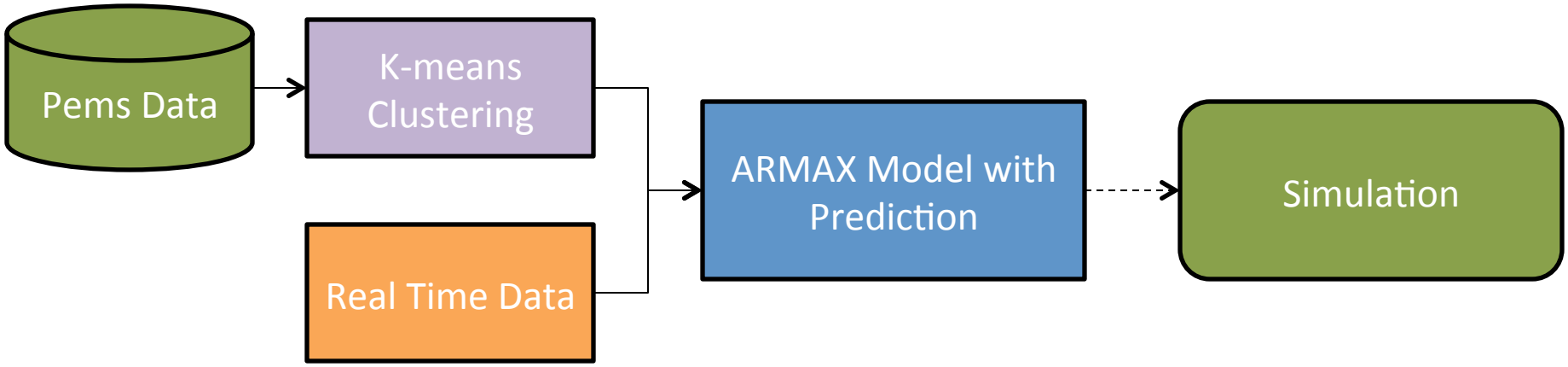
By Myles Iribarne and Cheng-Ju Wu

Why Traffic Engineering?

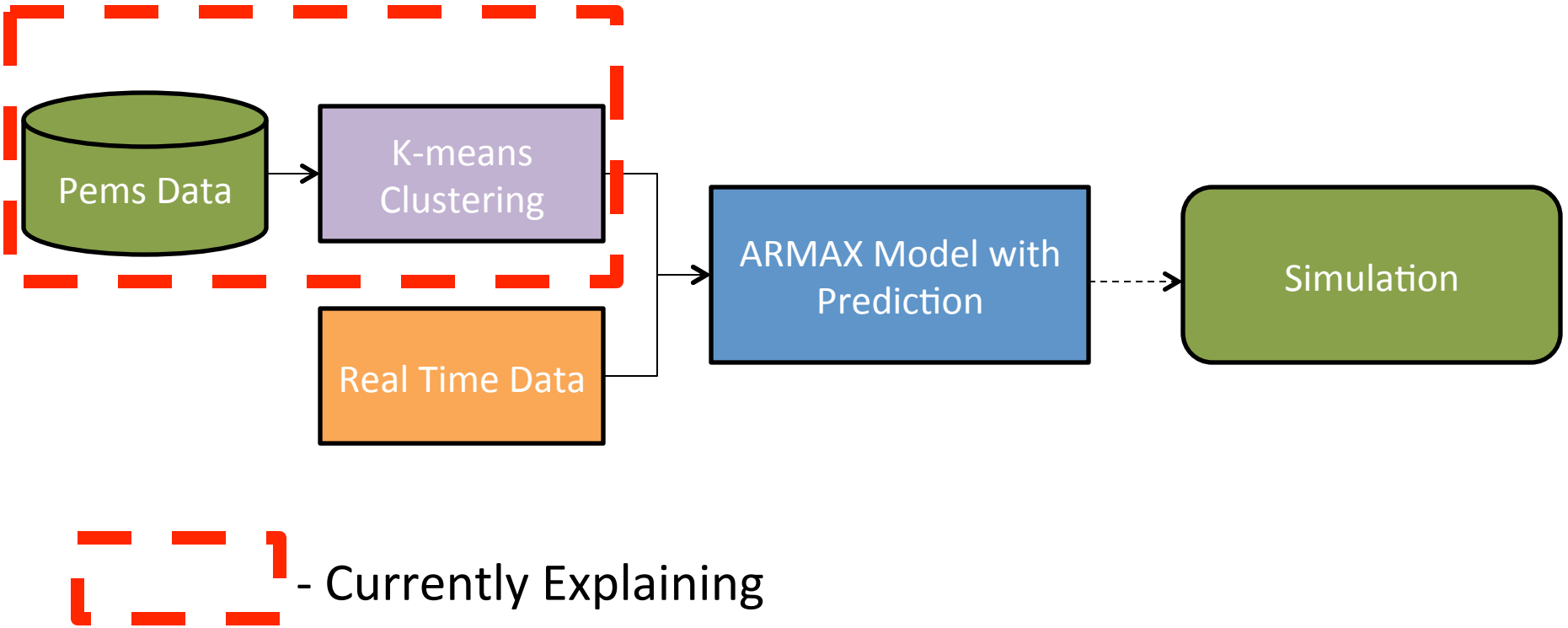
- Commuters spend 49 and 59 hours a year sitting in traffic in SF and LA, respectively [1]
- 1.9 billion gallons of gas wasted in each year in US sitting in traffic ~$7.6 billion [2]

Goal: to predict boundary traffic flow

- Motivation: boundary flow prediction needed for predictive control and simulations
- Strategy:



Strategy- K means Clustering



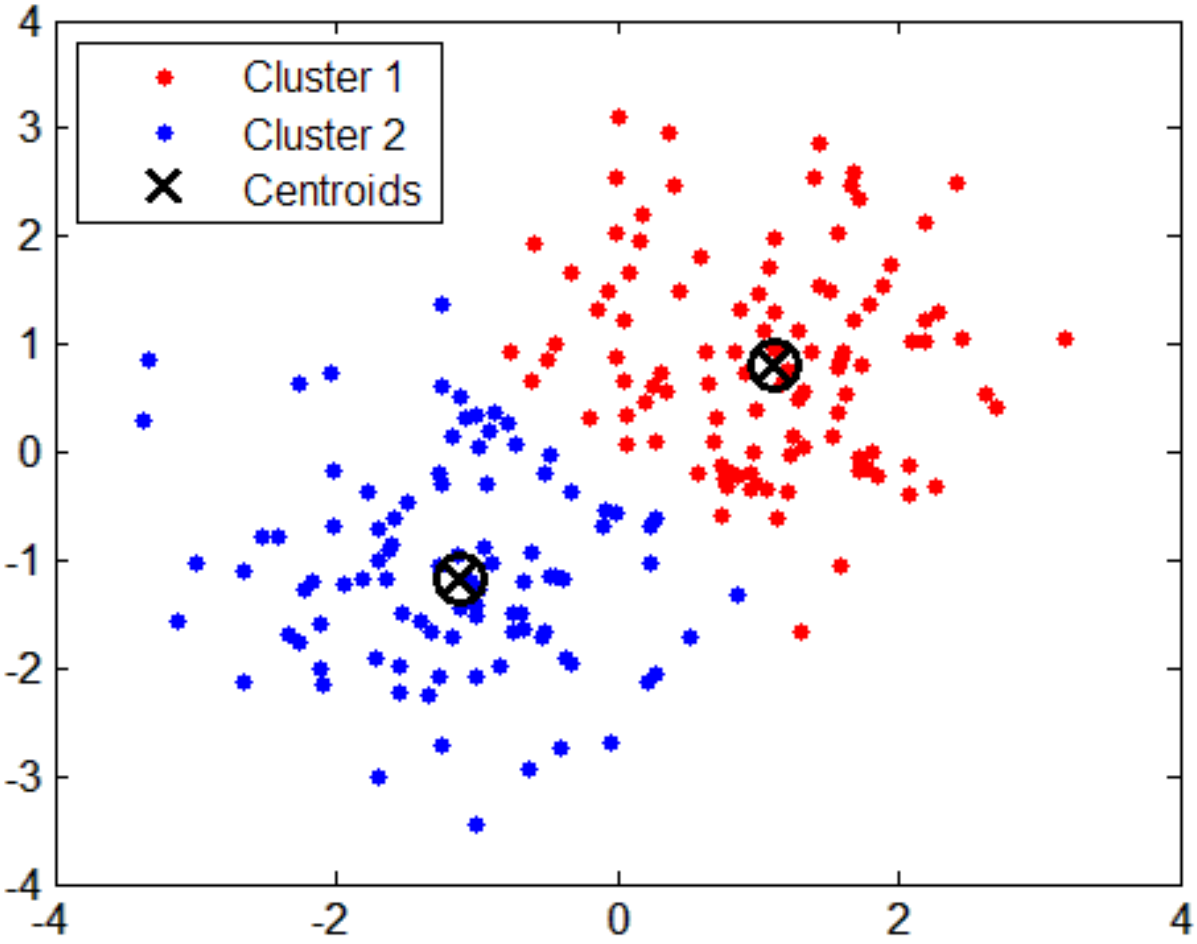
K means clustering

- Goal: partition data into K clusters
- In mathematics:

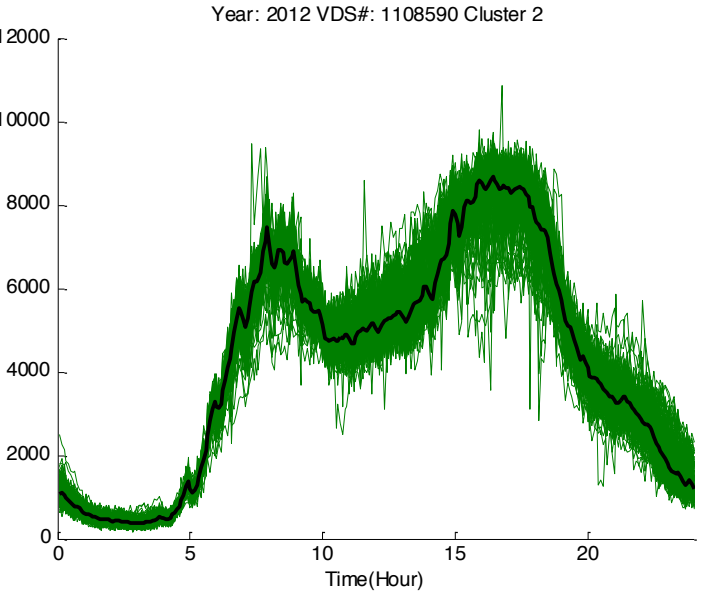
$$\min J = \sum_{i=1}^{N} \sum_{j=1}^{K} r_{ij} \left\| x_i - u_j \right\|^2$$

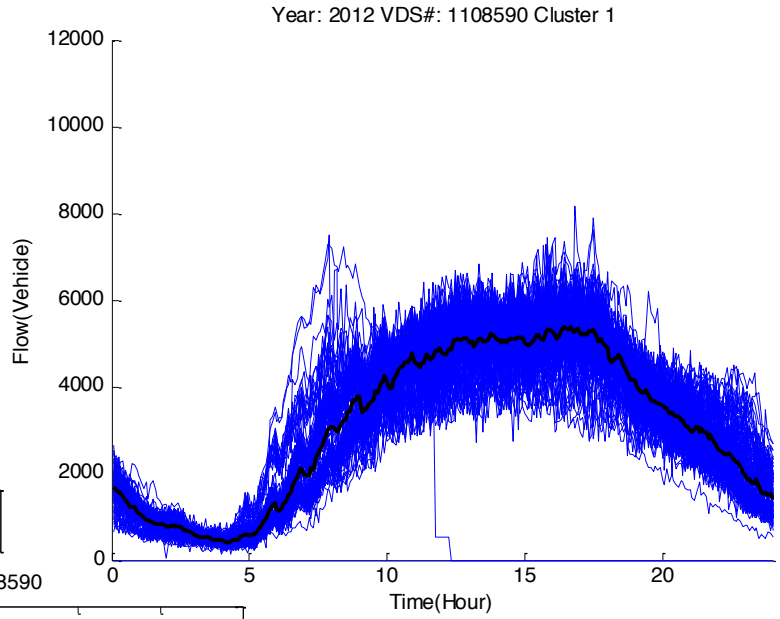
- K=# of clusters
- N=# of data points
- $$r_{ij} = \begin{cases} 1 & if\ data\ i\ belongs\ to\ cluster\ j \\ 0 & otherwise \end{cases}$$

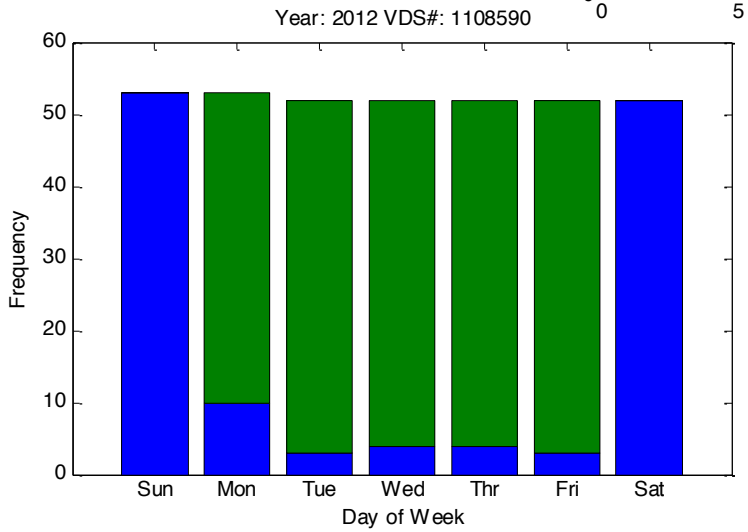
K means clustering- in a picture



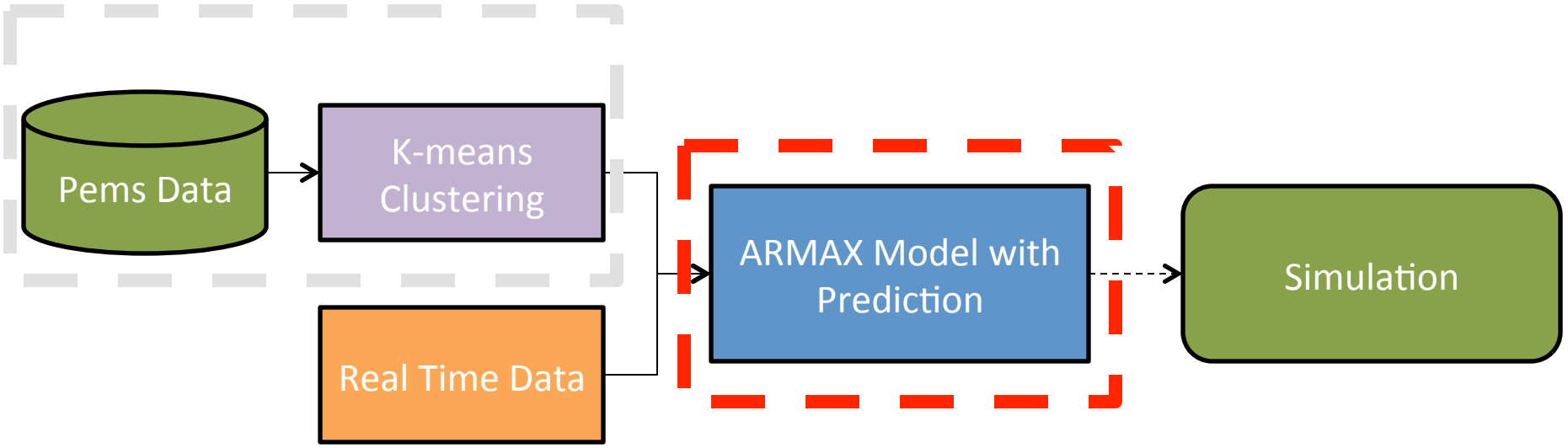
Clustering Traffic Flow Data







Strategy- ARMAX Model w/Prediction



- Previously covered

- Currently Explaining

ARMAX Model

- Auto-Regressive Moving-Average model with eXternal input
  - Is essentially the best discrete-time, linear model that fits given data:

$$y(k) = \frac{B(z^{-1})}{A(z^{-1})}u(k) + \frac{C(z^{-1})}{A(z^{-1})}e(k)$$

- A,B,C are polynomials w/input z^{-1}
  - i.e. $A(z^{-1})=1+a_1 z^{-1} +a_2 z^{-2} + a_3 z^{-3} \ldots\ldots a_n z^{-n}$
- z^{-1} represents a delay

ARMAX Model

$$y(k) = \frac{B(z^{-1})}{A(z^{-1})}u(k) + \frac{C(z^{-1})}{A(z^{-1})}n(k)$$

- A,B,C encode how previous inputs (u), states (y), and noise (n) affect the current state of the system (y(k))
- Example A=2nd order, B=0th order:
  - Looking only at u term and taking inverse z-transform, this equation tells us:

$$y(k) = -a_2 y(k-1) - a_1 y(k-2) + b_0 u(k)$$

ARMAX Model-Prediction proof

$$y(k) = \frac{B}{A}u(k) + \frac{C}{A}n(k) \rightarrow n(k) = \frac{A}{C}\left[y(k) - \frac{B}{A}u(k)\right] \quad [1]$$

Solving for n(k)

$$y(k+d) = \frac{B}{A}u(k+d) + \frac{C}{A}n(k+d) \quad [2]$$

Shifting [1] by d timesteps

$$\frac{C}{A}n(k+d) = Fn(k+d) + \frac{L}{A}n(k) \quad [3]$$

$$AF + z^{-d}L = C \quad [4]$$

Re-writing the second term in [2] where F and L are found by solving the Bezout equation seen in [4]

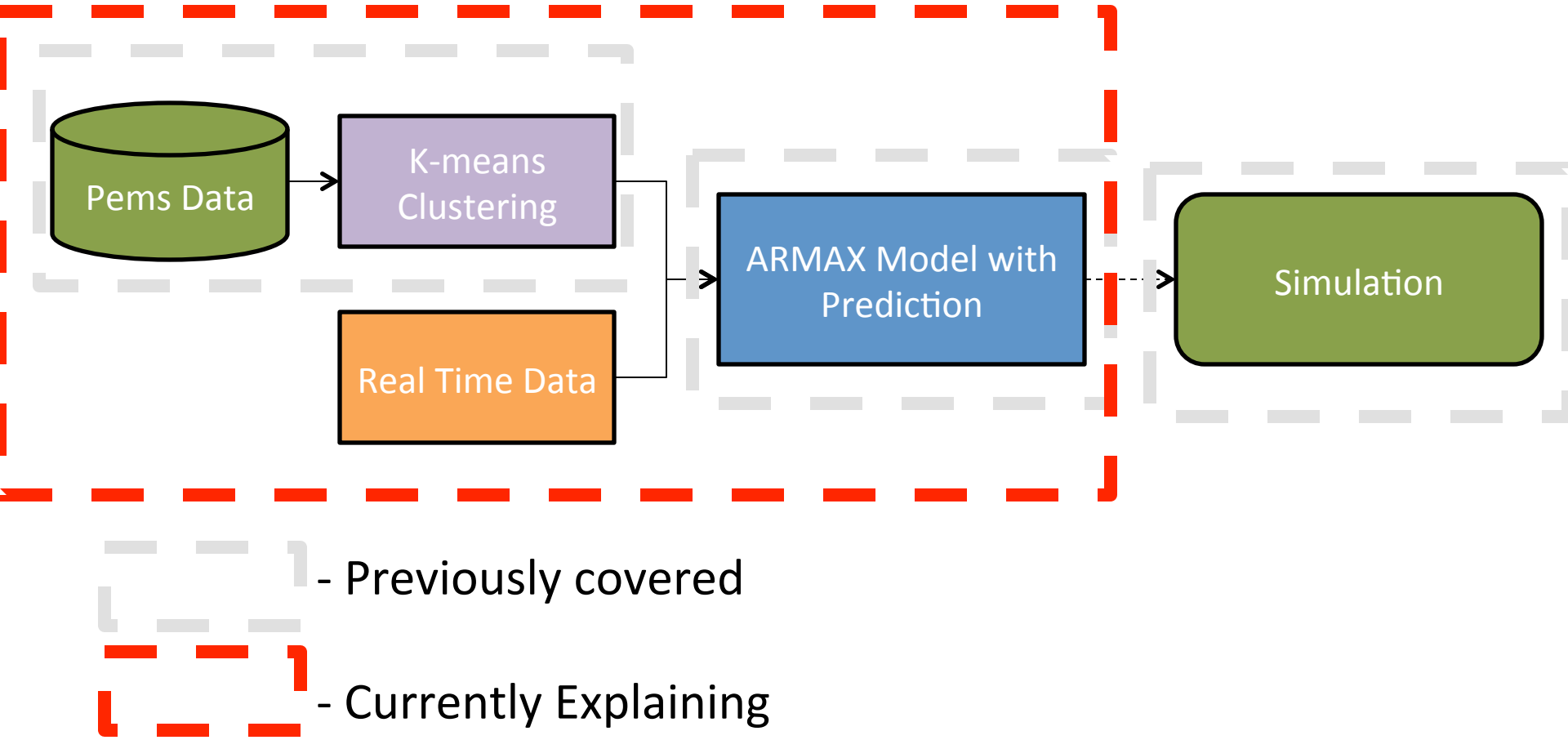
$$y(k+d) = \frac{L}{C}y(k) + \frac{BF}{C}u(k+d) + Fn(k+d) \quad [5]$$

Plugging [1] and [3] into [2] and rearranging carefully yields [5]

$$\hat{y}(k+d) = \frac{L}{C}y(k) + \frac{BF}{C}u(k+d) \quad [6]$$

➔ Excluding the noise term, eqn [6] Is the optimal d-step prediction

Strategy- Simulation



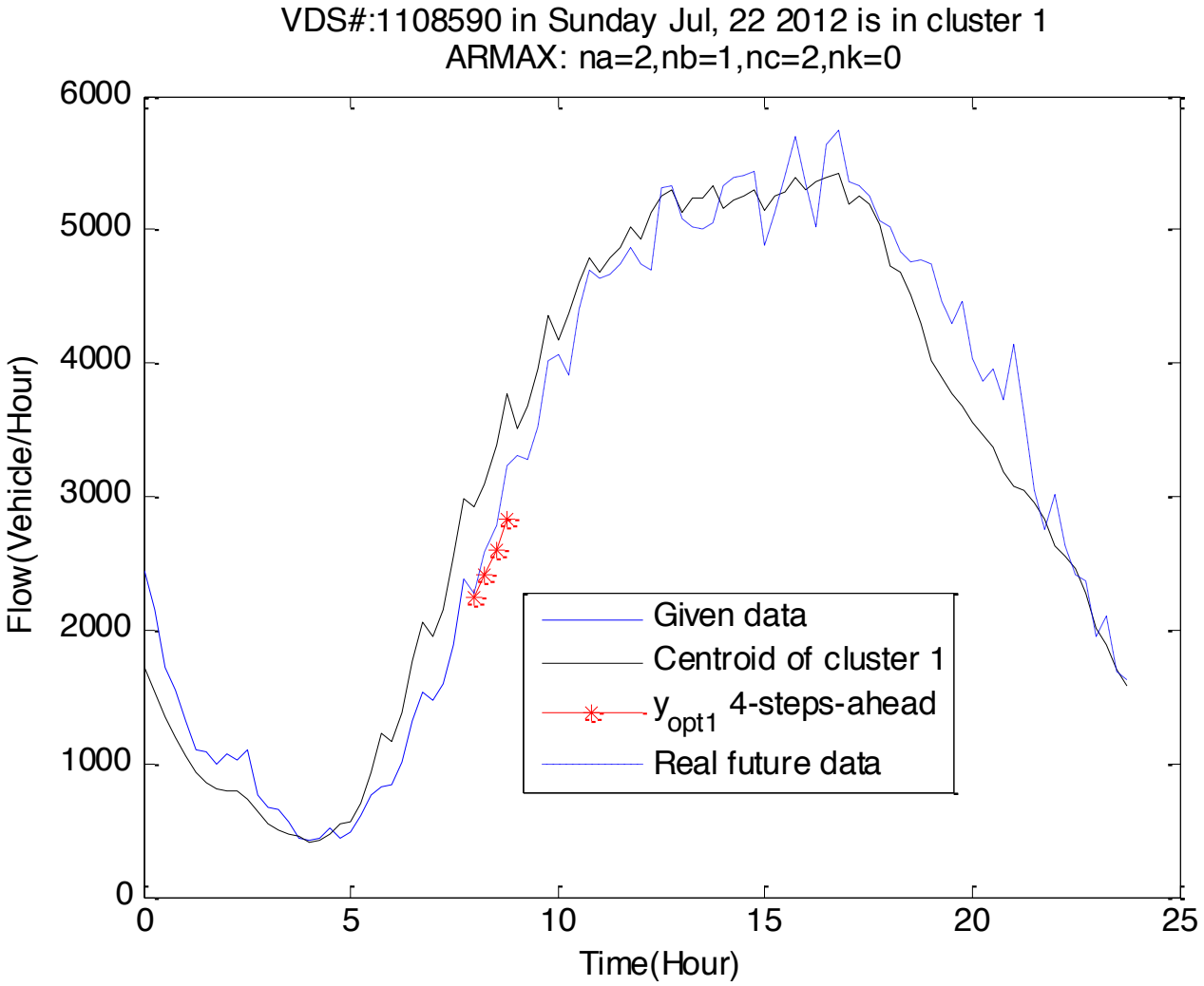
Putting it all together

1. Create clusters using a years worth of historical data
2. For each of the two clusters:
   - Create ARMAX model using cluster centroid and real-time data up until current time.
     - u(k)=centroid of cluster-
     - y(k)= flow at kth timestep
     - timestep=15 minutes
   - Predict future traffic flow
   - Update model w/new data as time passes for more accurate prediction

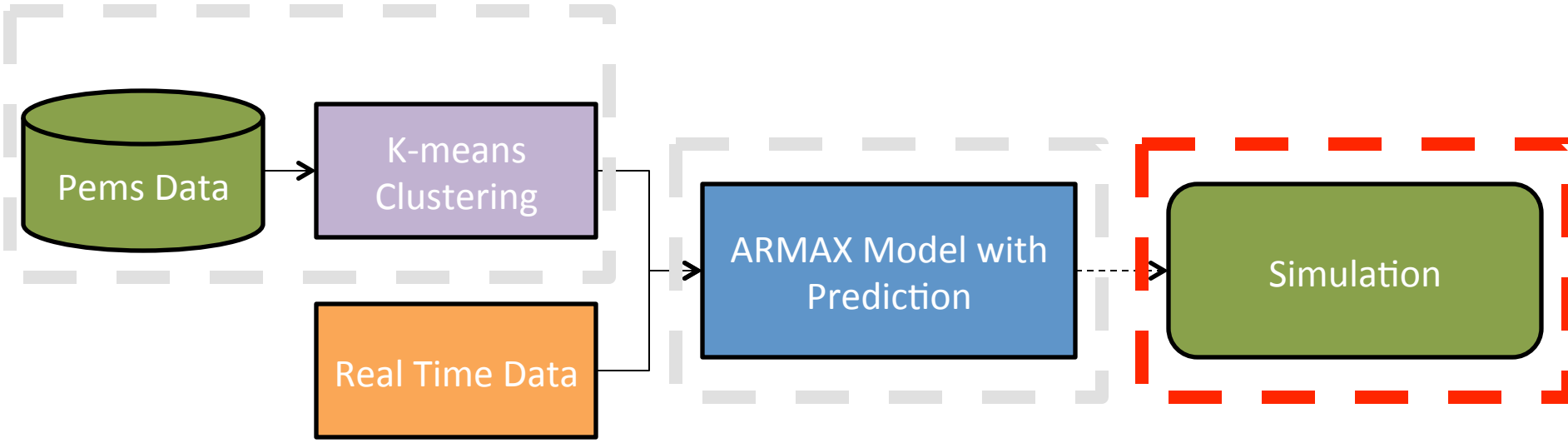
$$y(k) = \frac{B(z^{-1})}{A(z^{-1})}u(k) + \frac{C(z^{-1})}{A(z^{-1})}e(k)$$

** Example: use todays data from 12AM -8AM, along with cluster centroid over the same period, to create ARMAX model. This model could then be used for prediction from 8AM-12PM today.**

Prediction Results for Mainline #1108590



Strategy- Simulation

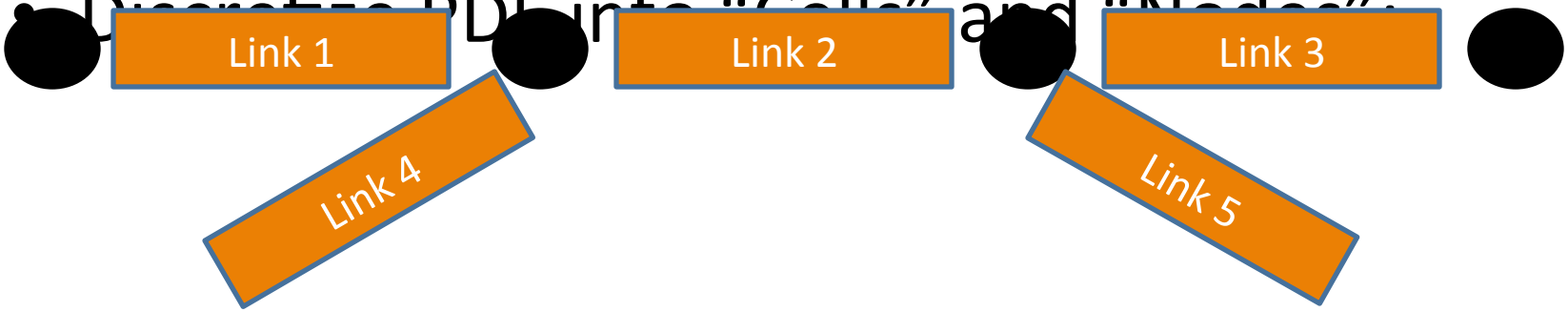


- Previously covered

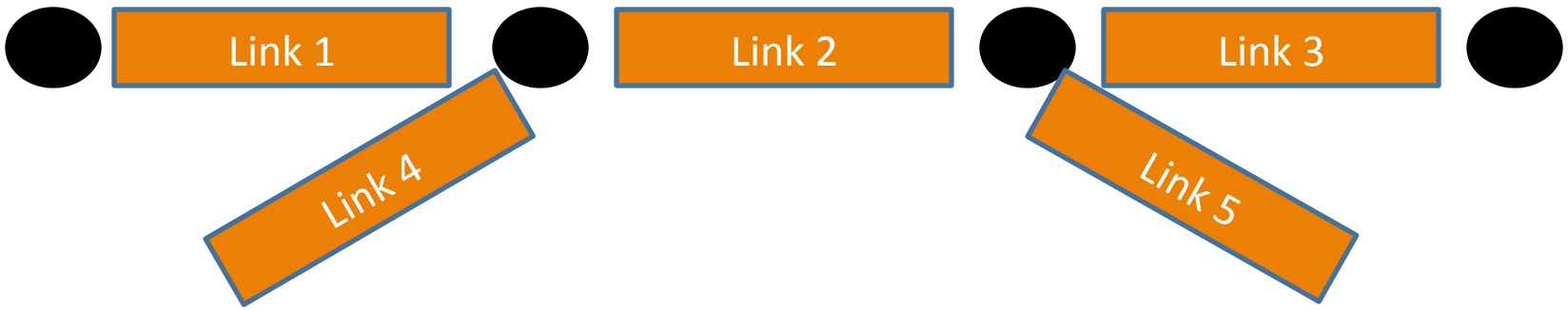
- Currently Explaining

Linked-Node CTM

- Stands for Linked-Node Cell Transmission Model
- Start with $\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial f(x,t)}{\partial t} = 0$
- Discretize PDE into "Cells" and "Nodes".

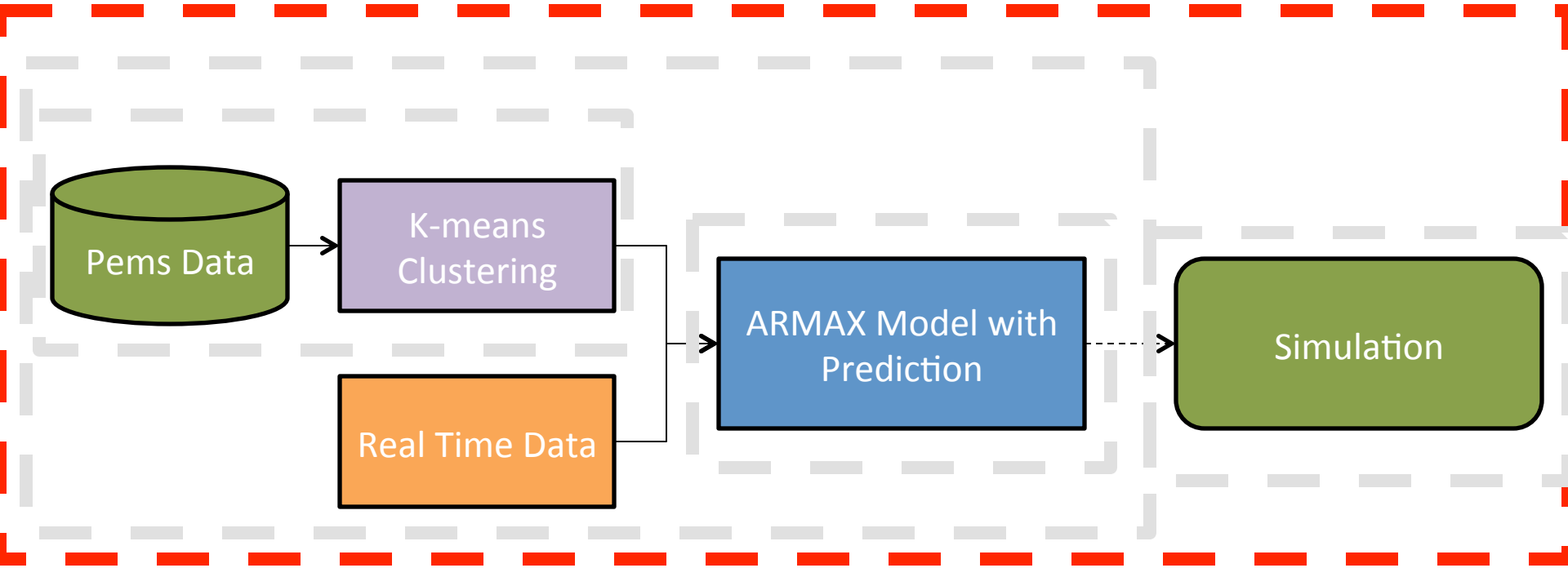


Linked-Node CTM



- At each time step, the density and flow in each link is updated
  - maximize flow between links while satisfying fundamental diagram constraints
- Nodes allow for arterials, on ramps, off ramps- don't store cars
- For off-ramps, split ratio (r) defines what portion of traffic flows to which link

Strategy- Testing it all together

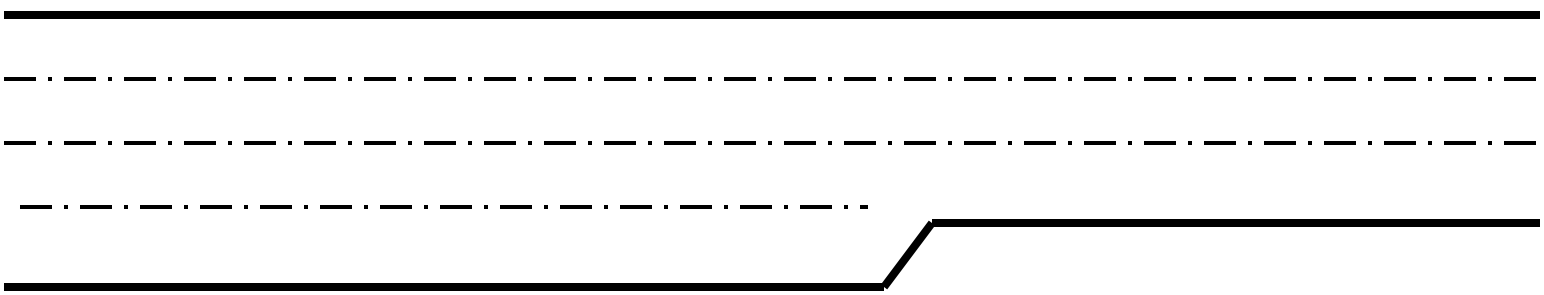


- Previously covered

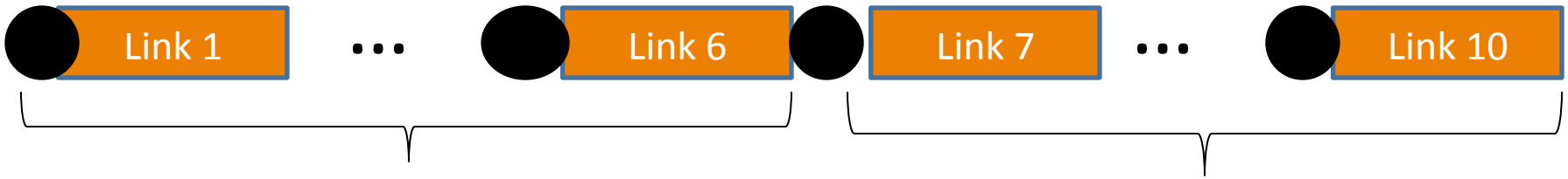
- Currently Explaining

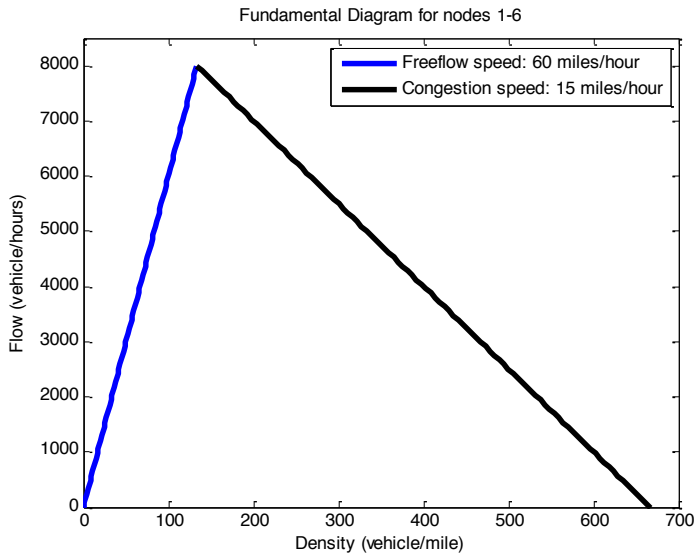
Using the Prediction in a Simulation

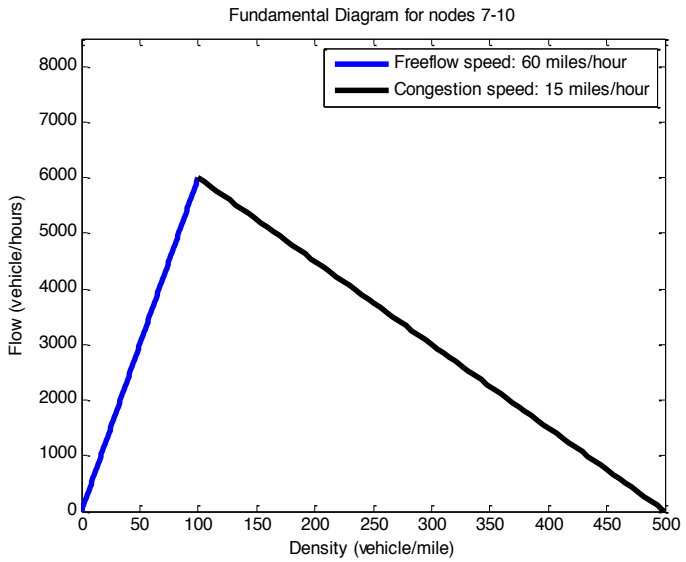
- In a linked-node CTM model, you can set the *demand* at the boundary
  - Boundary flow ≠ demand (if congested)
  - How to use predicted flow?
- Use simple "Merge" network with controlled characteristics to study this relationship:



Our network





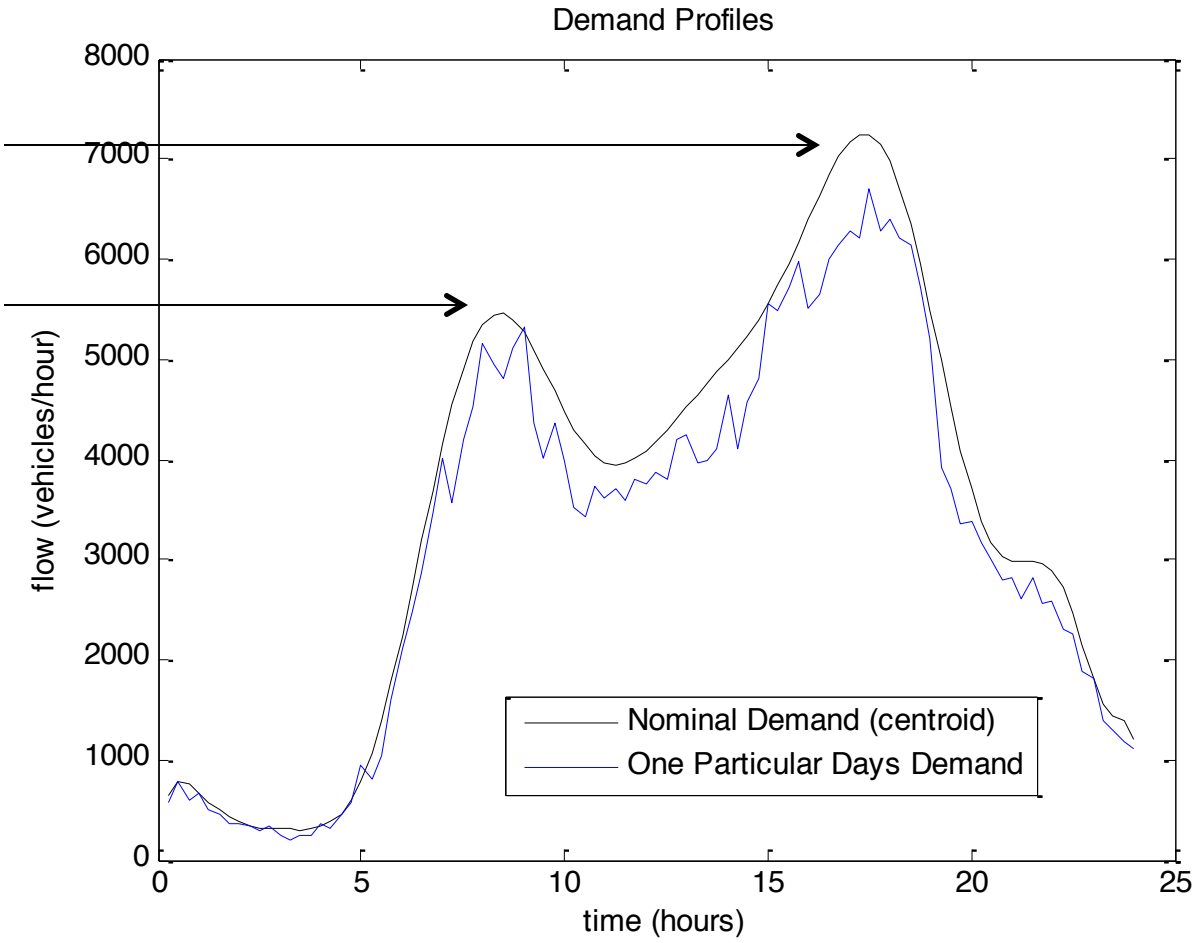


- Capacity in links 7-10 drops from 8000 to 6000
  - Reflects a “merge” from 4 lanes to three lanes

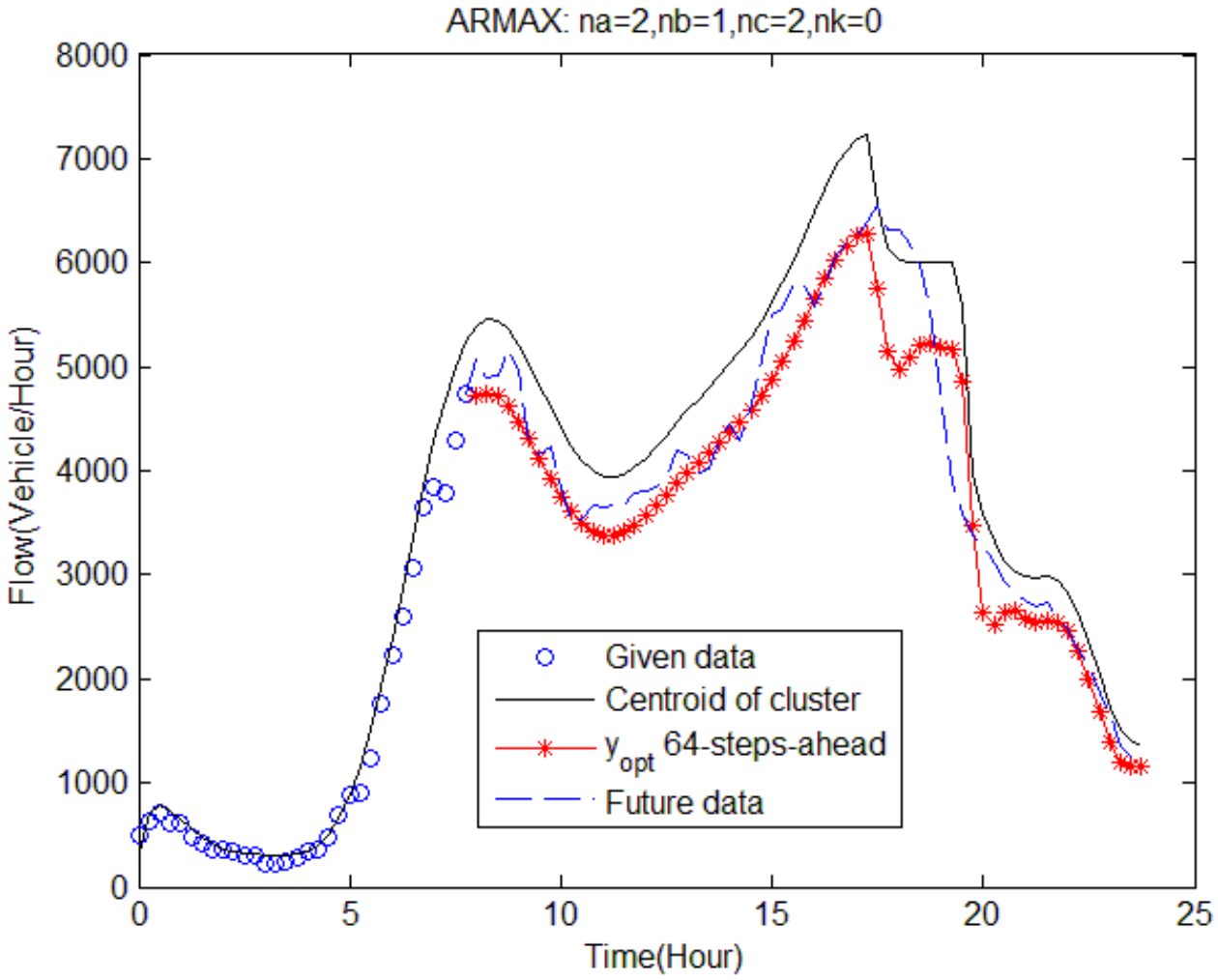
Creating Demand Profiles for use in Simulation

Evening Peak

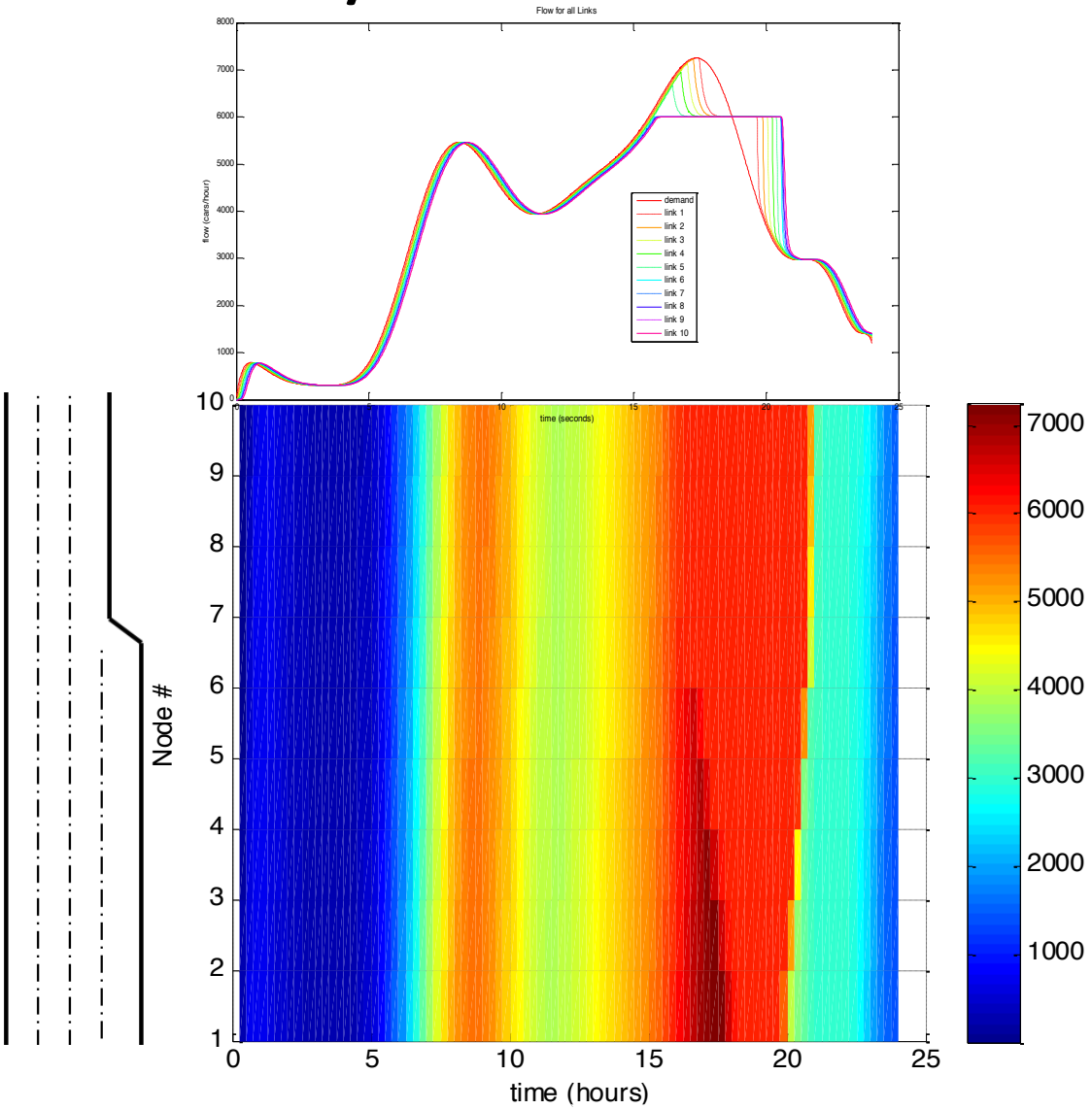
Morning Peak



Prediction Results

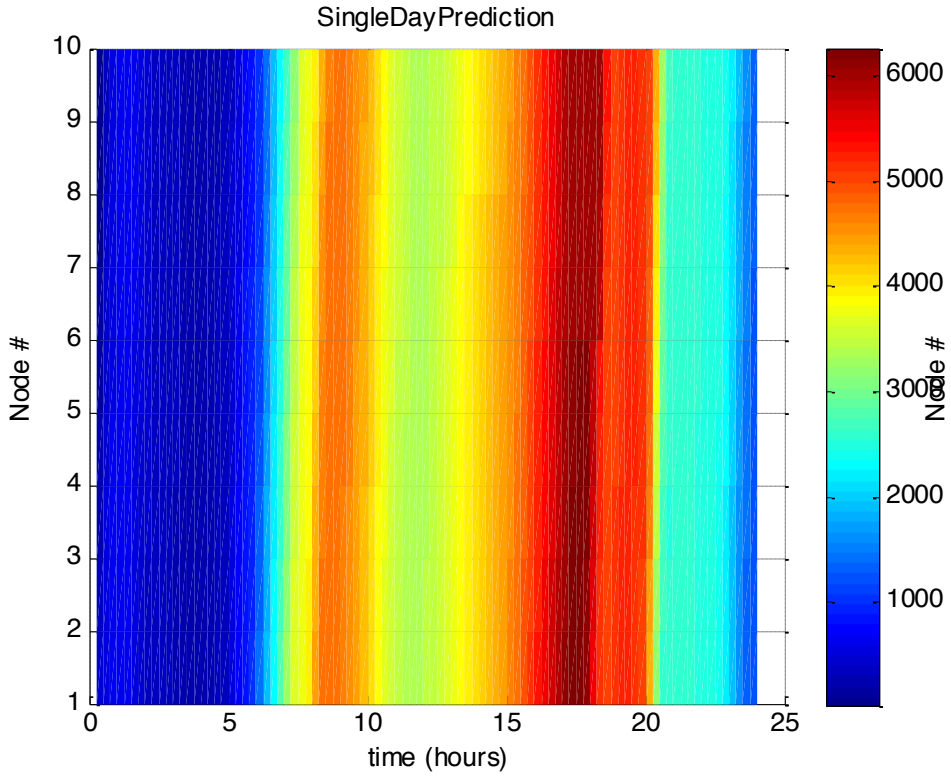


Flow contour w/Centroid Demand Profile

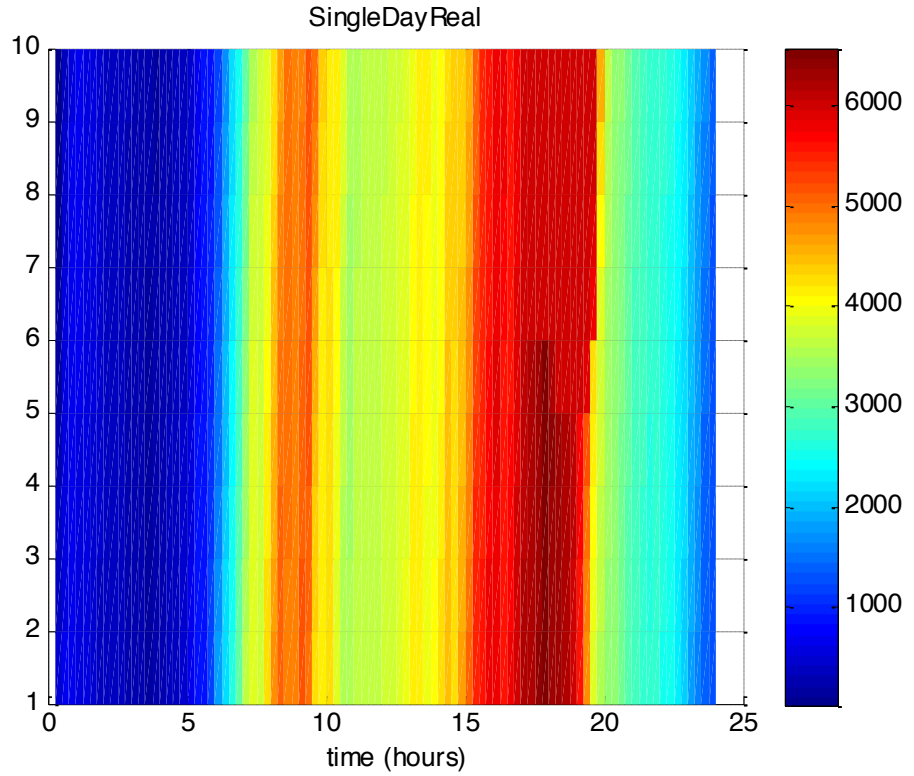


Flow Contour with Actual and Predicted boundary flow

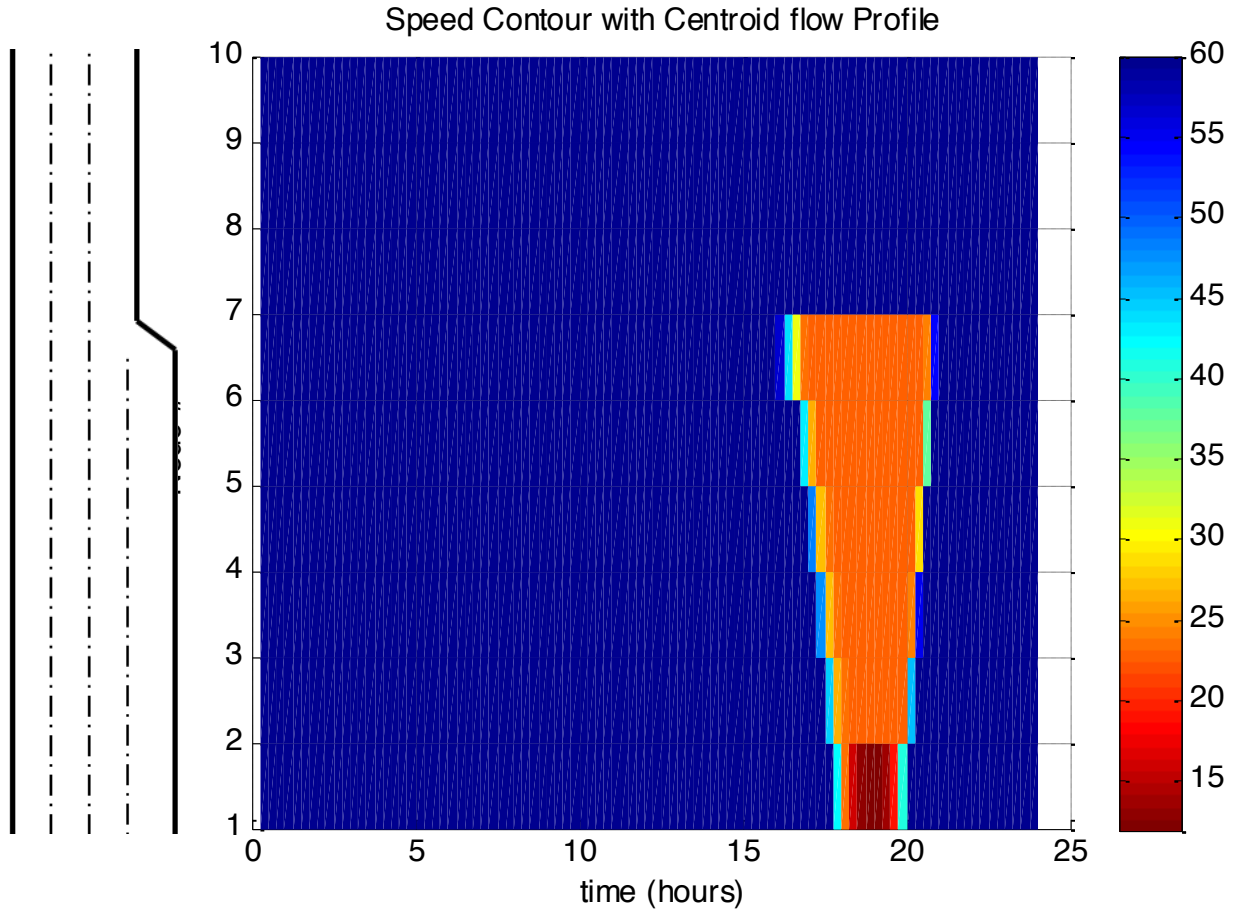
Predicted



Actual

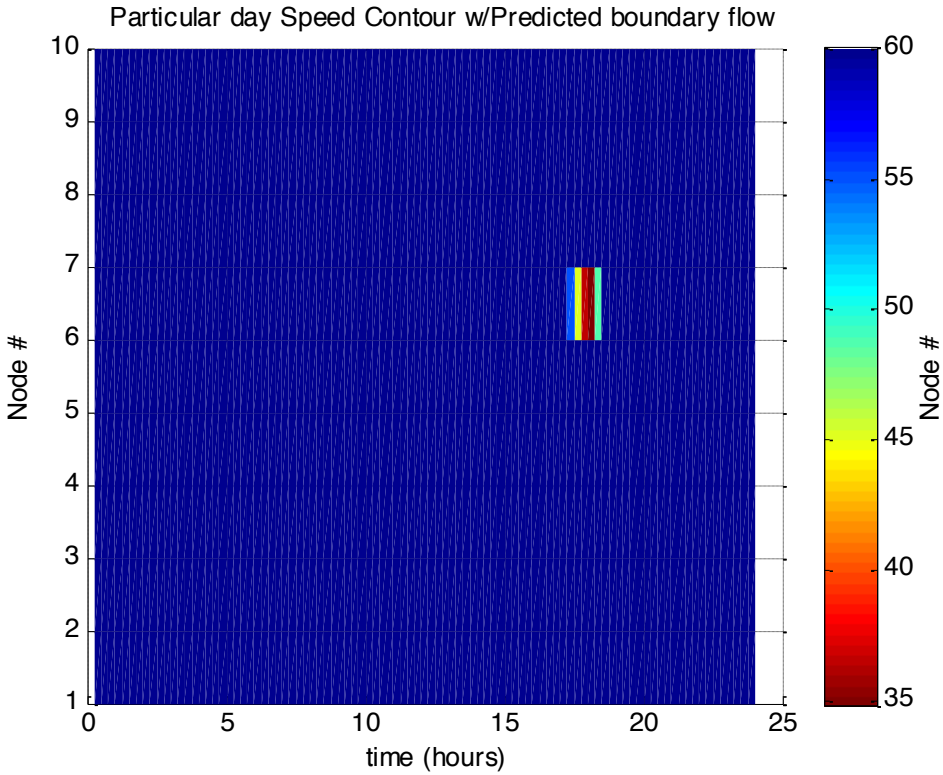


Speed contour using Centroid Demand Profile

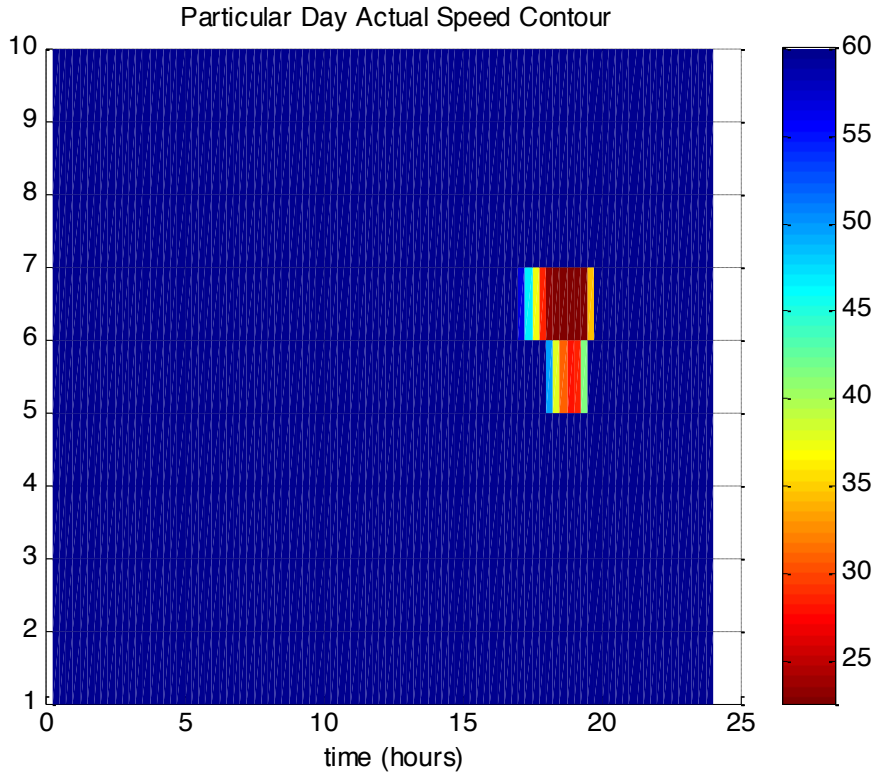


Speed Contour with Actual and Predicted boundary flow

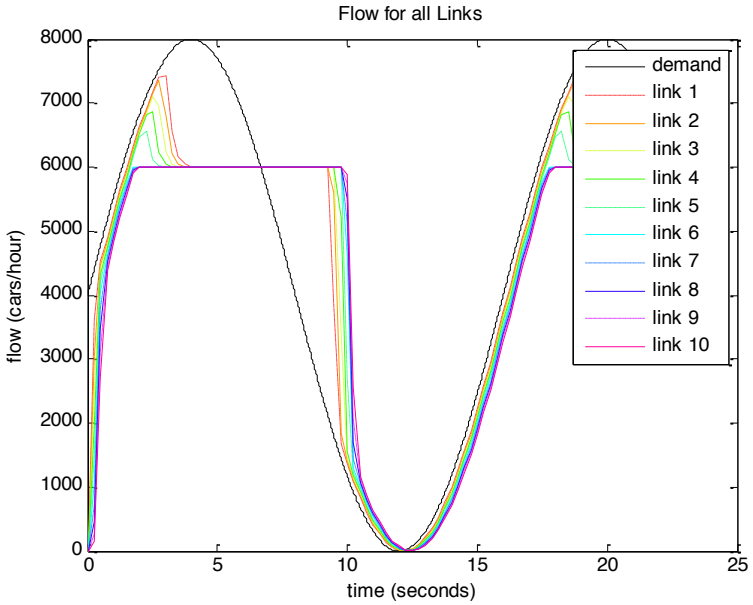
Predicted

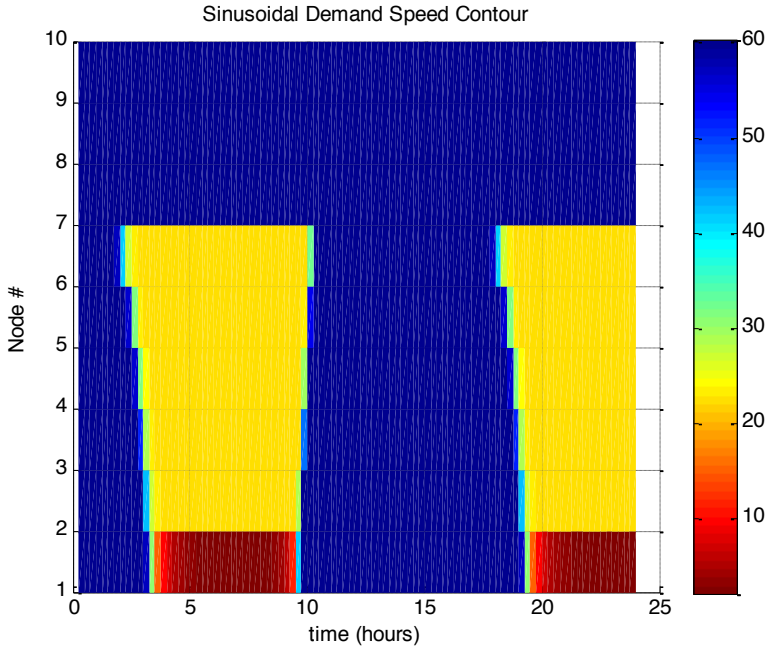


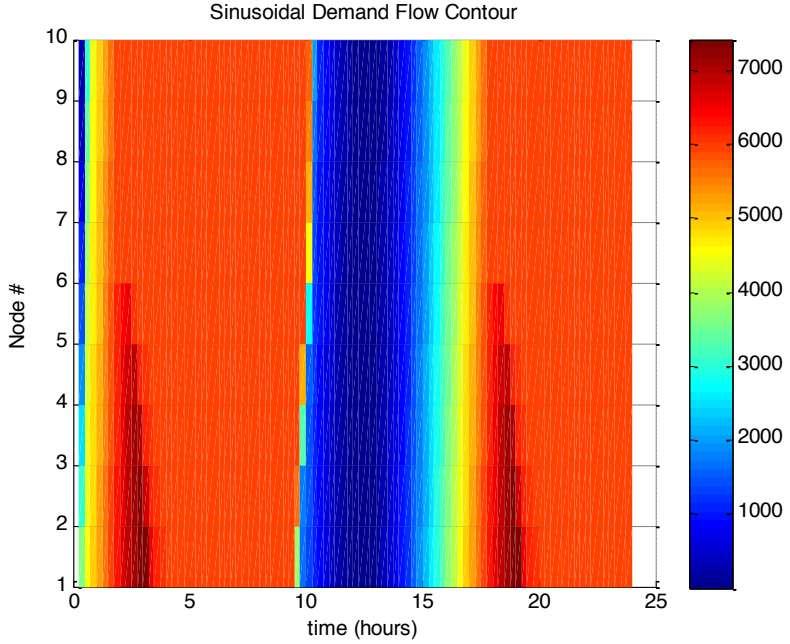
Actual



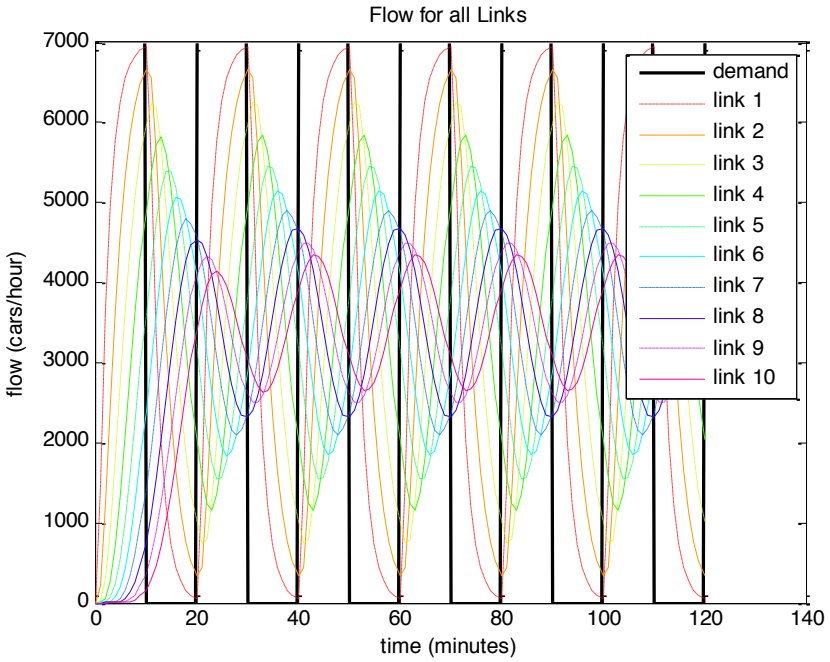
Sinusoidal Demand Profile (24 Hr period)

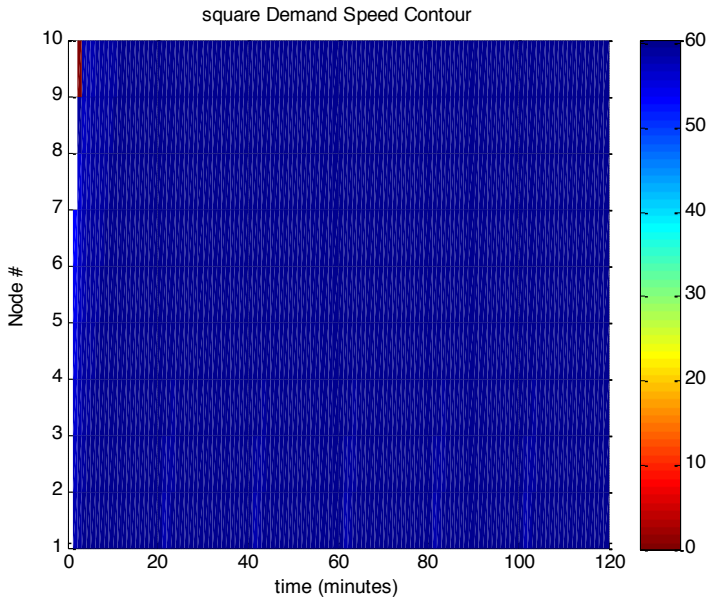


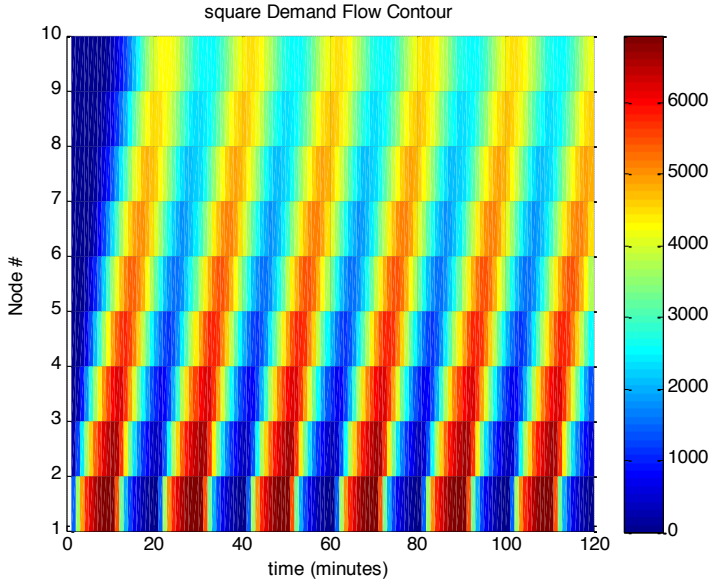




Square Wave Demand Profile (2 Hr period)







Conclusion and Next Steps

- LN-CTM demonstrates traffic dynamics well
- Cluster+ARMAX predictor useful - better than simply using historical data
- Next Steps
  - Try recursively defined ARMAX model
  - Validate with a real network and real data
  - Add on ramp and off ramps to model

References

[1] http://www.huffingtonpost.com/2013/04/25/worst-traffic-in-us-2013_n_3155823.html?ir=Los+Angeles#slide=2379965

[2] http://www.nbcnews.com/business/we-waste-1-9-billion-gallons-gas-sitting-traffic-558071