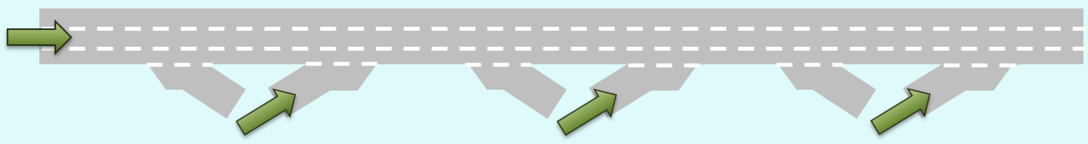
Fast Boundary Flow Prediction using Optimal Autoregressive Moving Average with Exogenous Inputs (ARMAX) Predictor

Paper-No.: TRB'14-4645
15 January 2014, 14:45 - 16:30, Marriott Salon 2, D03
TRB Session 841: "Advances in Traffic Flow Theory and Characteristics, Part 2" (AHB45)

Cheng-Ju Wu
Thomas Schreiter
Roberto Horowitz
Gabriel Gomes

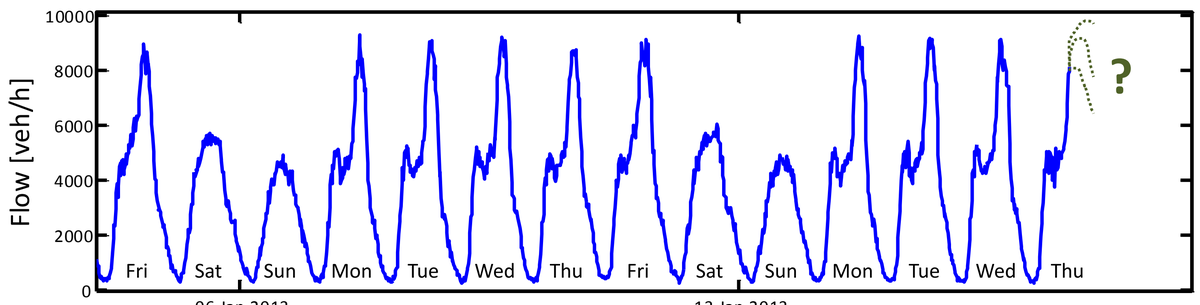
University of California, Berkeley
Partners for Advanced Transportation Technology
Berkeley, CA, USA
schreiter@berkeley.edu

Connected Corridors program by Berkeley PATH and Caltrans:
- Build decision support system for live TMC operations
- Short-term traffic prediction by cell-transmission model
- Requires short-term prediction of network boundary inflows



This paper: fast and reliable boundary flow predictor
- Based on months of historical data
- Combine historical average flow and recent flow
- Works in both regular and irregular conditions

Empirical Loop Traffic Data

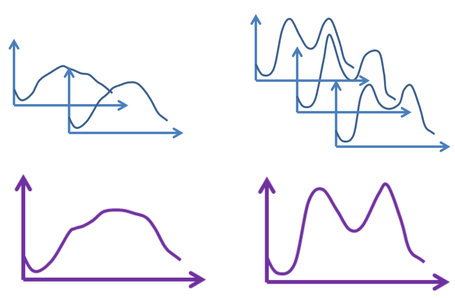


[California PeMS, Loop Sensor 1108595, I-15N]

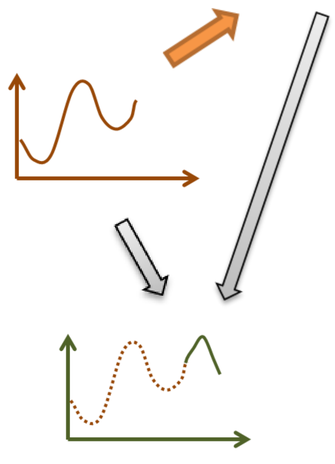
Phenomena:
- Traffic repeats daily and weekly
- Small deviations every day

Overview of Methodology

Calibration:
- Data: Months of *historical flow data* based on California PeMS sensors
- Algorithm: *Cluster historical data* by day of the week and calculate centroid by median



Live Cluster Matching:
- Data: *Live flow data* from midnight to now
- Algorithm: *Determine todays centroid* based on day of the week (trivial)



Forecast:
- Algorithm: *Combine live data with todays centroid by applying ARMAX model*

Details

Idea:
Traffic flow = historical average + structural deviation of today + error

Autoregressive Moving Average with Exogenous Input model (ARMAX):
Based on idea: $y^{pred}(k+1) = b_{-1}u(k+1) + b_0u(k) + b_1u(k-1) + \ldots$
$+ a_0y(k) + a_1y(k-1) + \ldots$
$+ c_0w(k) + c_1w(k-1) + \ldots$

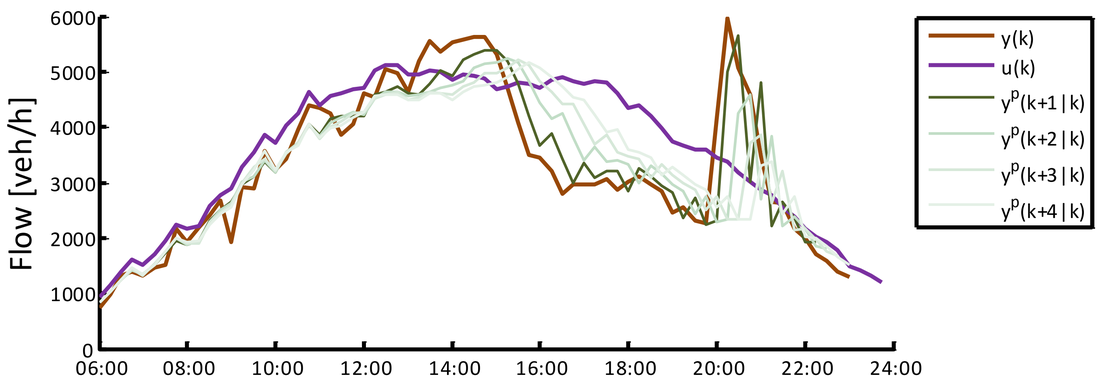
y^{pred}: prediction; y: todays data; u: centroid; w: estimation residual; k: time step; a_i, b_i, c_i: ARMAX parameters

Parameter estimation of a_i, b_i, c_i,: live in every time step by least squares

Features:
- ARMAX model reacts to structural deviations, e.g. caused by special events such as NFL Super Bowl game
- Prediction is fast, due to small amount of data
- Using median in calibration step is reliable w.r.t. outliers, e.g. caused by unusual traffic or sensor failures
- Calibration offline, which is slow due to large amount of data

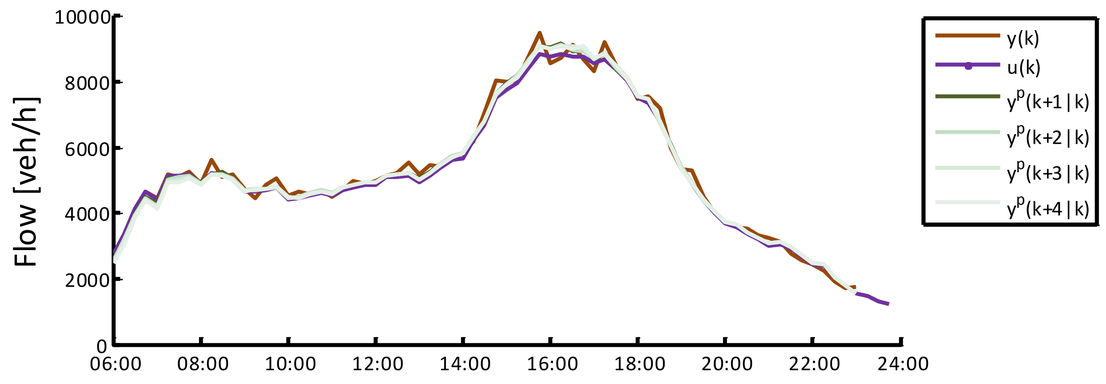
Results: One-hour Prediction of Traffic Flow at a Network Boundary

Day with irregular traffic pattern: Super Bowl Day



Mean Absolute Error of prediction by ARMAX: 501 veh/h
Mean Absolute Error of prediction by historical average only: 703 veh/h

Day with regular traffic pattern: a working day



Mean Absolute Error of prediction by ARMAX: 153 veh/h
Mean Absolute Error of prediction by historical average only: 160 veh/h