Adarules: Learning rules for real-time road-traffic prediction

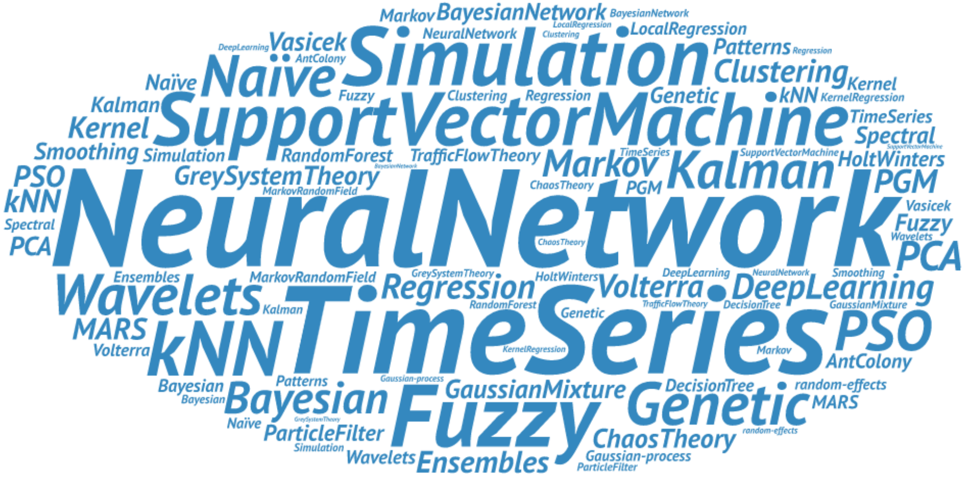
Rafael Mena-Yedra[1,2]
Ricard Gavaldà[2]
Jordi Casas[1]

[1] TSS-Transport Simulation Systems, Spain
[2] Universitat Politècnica de Catalunya, Spain

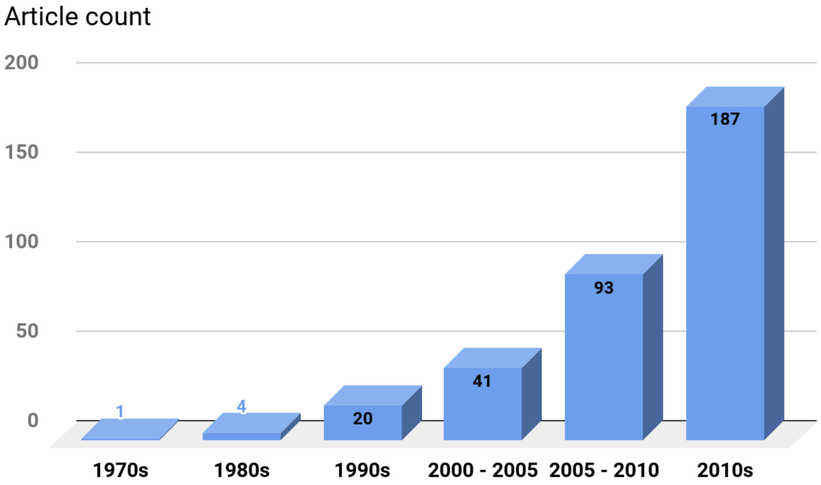
20th EURO Working Group on Transportation Meeting (EWGT 2017)
4-6 September 2017, Budapest, Hungary

Traffic (flow) prediction
How and what for?

Traffic prediction research



“Traffic flow prediction”



Why traffic prediction

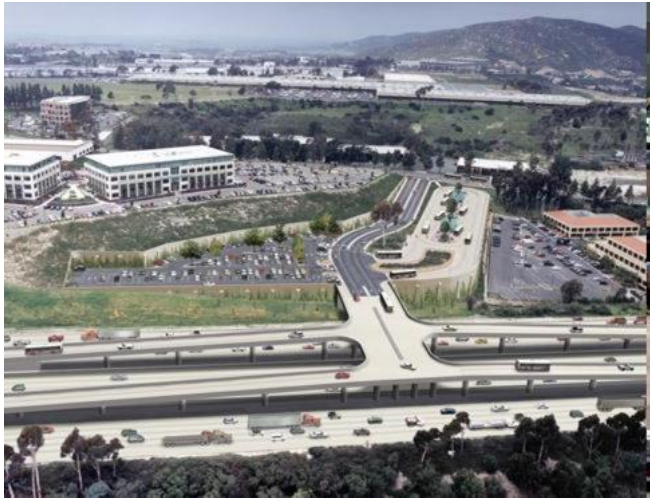
- Traveler Information Services
- Active Traffic Management

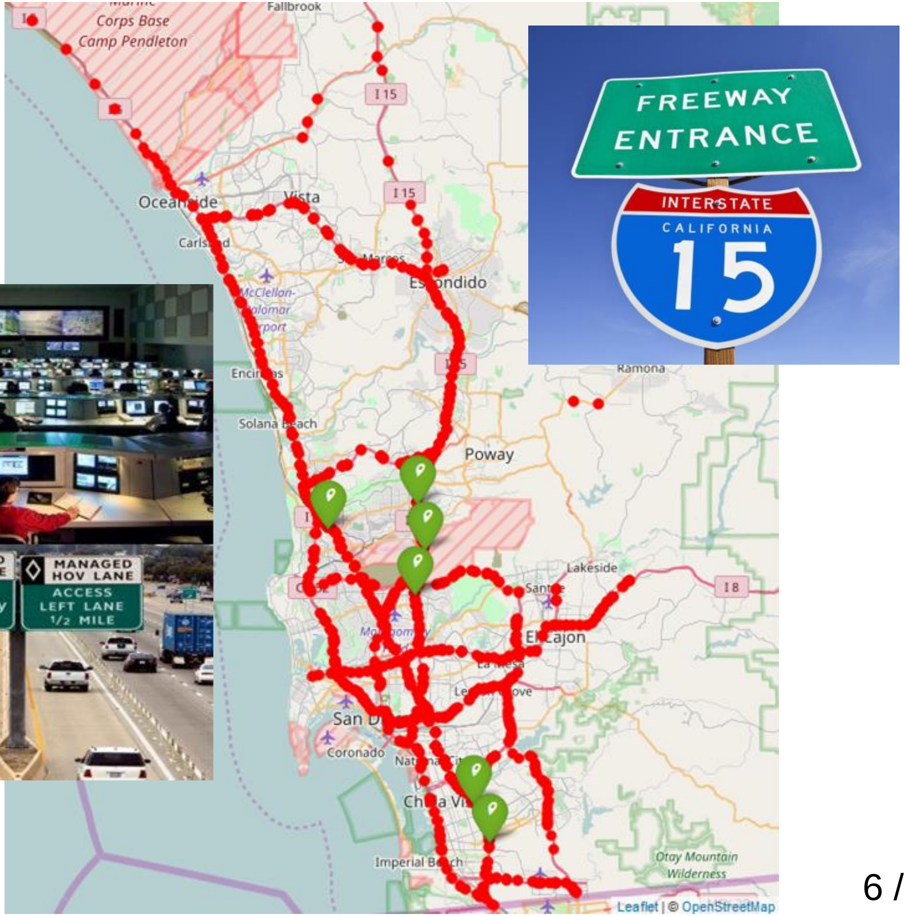
- Beneficial impact on the network performance in terms of throughput, congestion length and average network speeds.

- Decision support systems for real-time traffic management.
  - Example: Aimsun Online
- Valuable input for other processes: trend to merge both approaches, purely data-driven methods and simulation models.

Motivation

Case study: San Diego (I-15)

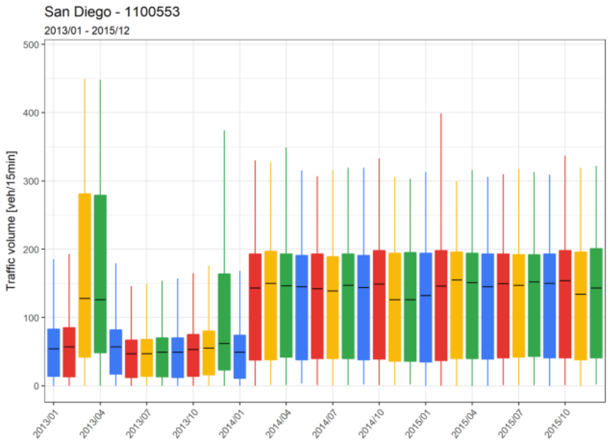


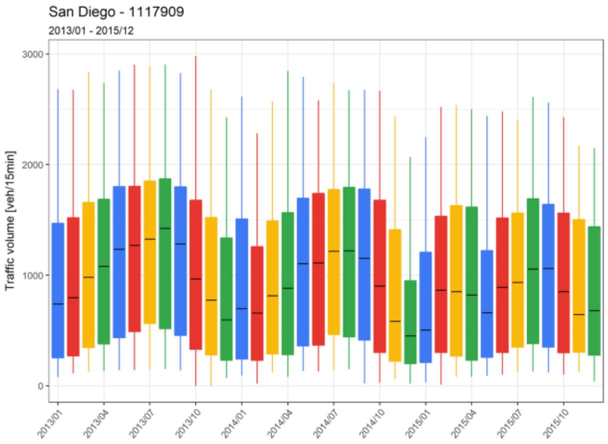


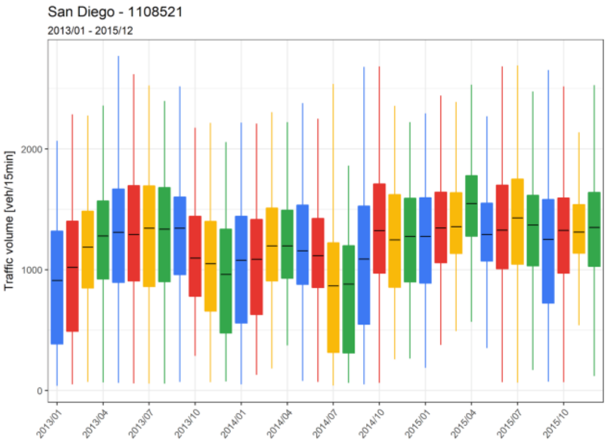
Data source: California Department of Transportation (Caltrans) Performance Measurement System (PeMS). State of California.

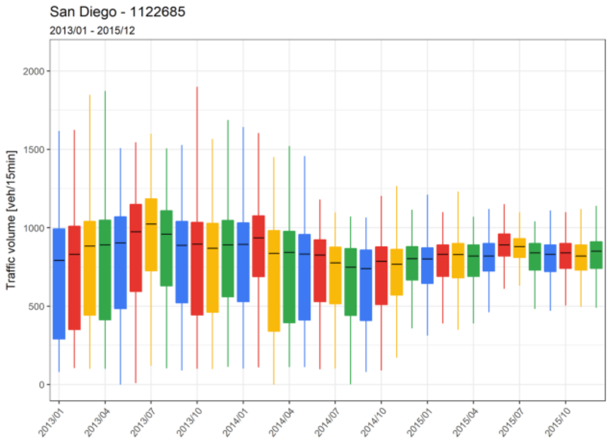
Case study: San Diego (I-15)

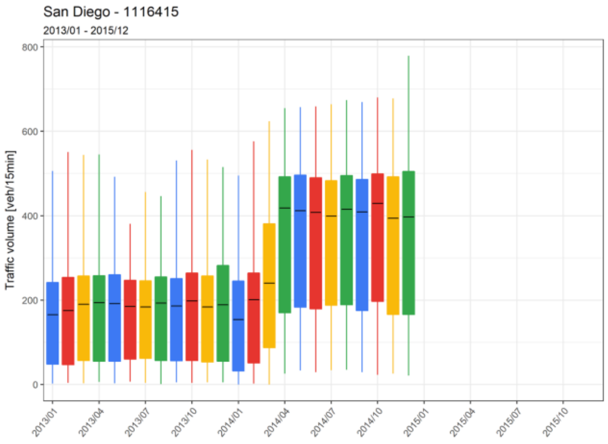
Data source: California Department of Transportation (Caltrans) Performance Measurement System (PeMS). State of California.

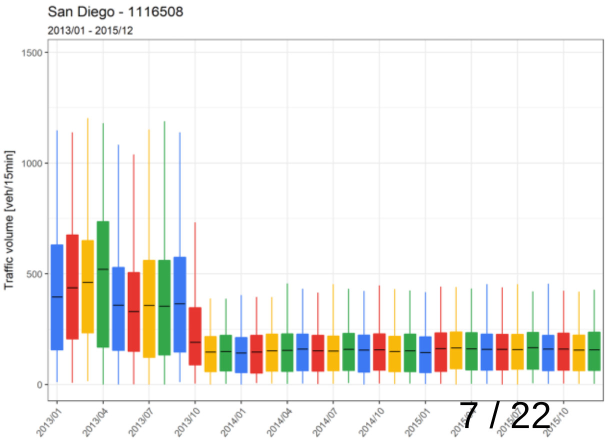












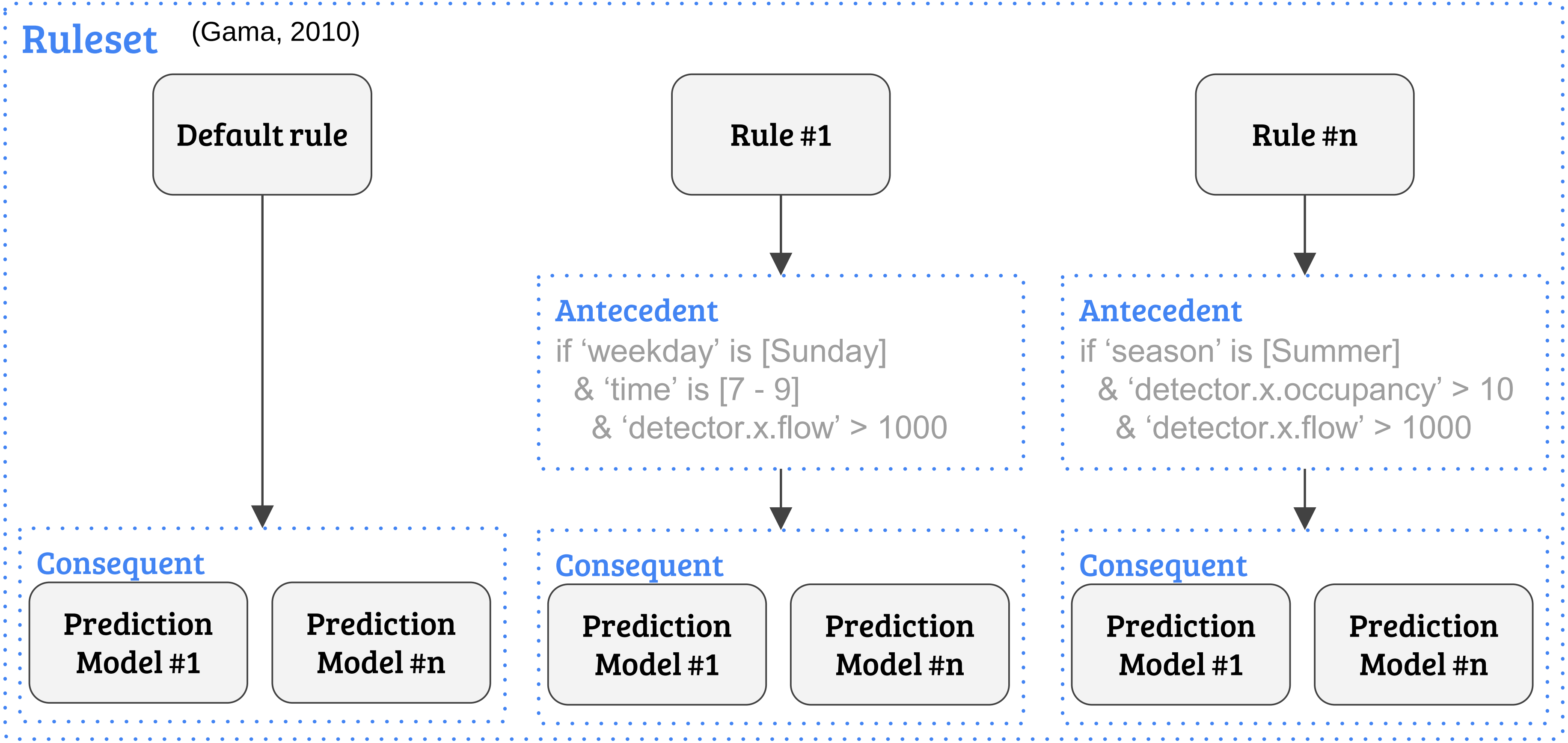
Identified issues

- Diversity (kind of network, or even within the same network)
- Sudden change
- Gradual change (drift)
- Missing data observations
- Dependence on the data scientist or traffic engineer criteria for each case

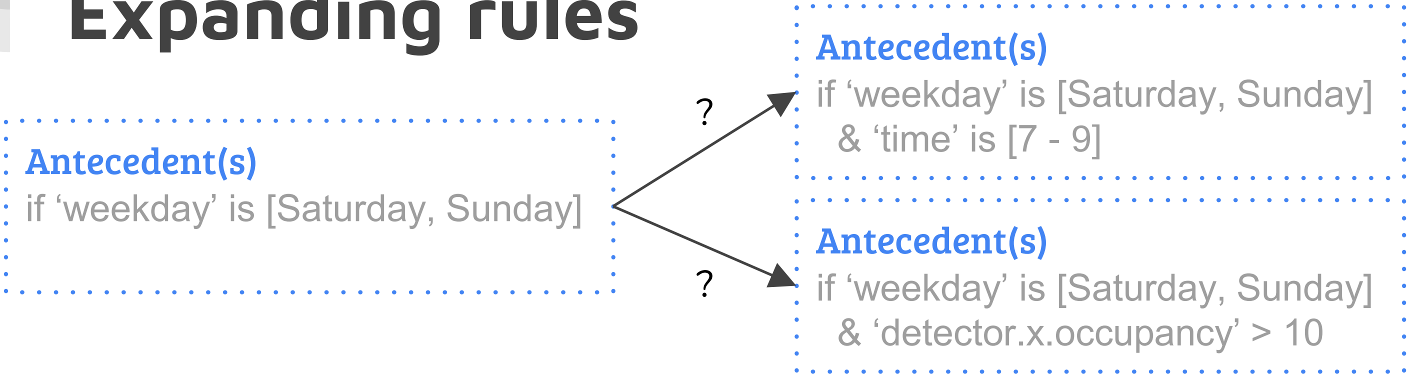
Our approach: learning adaptive rules

"Adarules"

Adarules


Ruleset
(Gama, 2010)
Default rule
Rule #1
Rule #n
Antecedent
if ‘weekday’ is [Sunday]
& ‘time’ is [7 - 9]
& ‘detector.x.flow’ > 1000
Antecedent
if ‘season’ is [Summer]
& ‘detector.x.occupancy’ > 10
& ‘detector.x.flow’ > 1000
Consequent
Prediction Model #1
Prediction Model #n
Consequent
Prediction Model #1
Prediction Model #n
Consequent
Prediction Model #1
Prediction Model #n

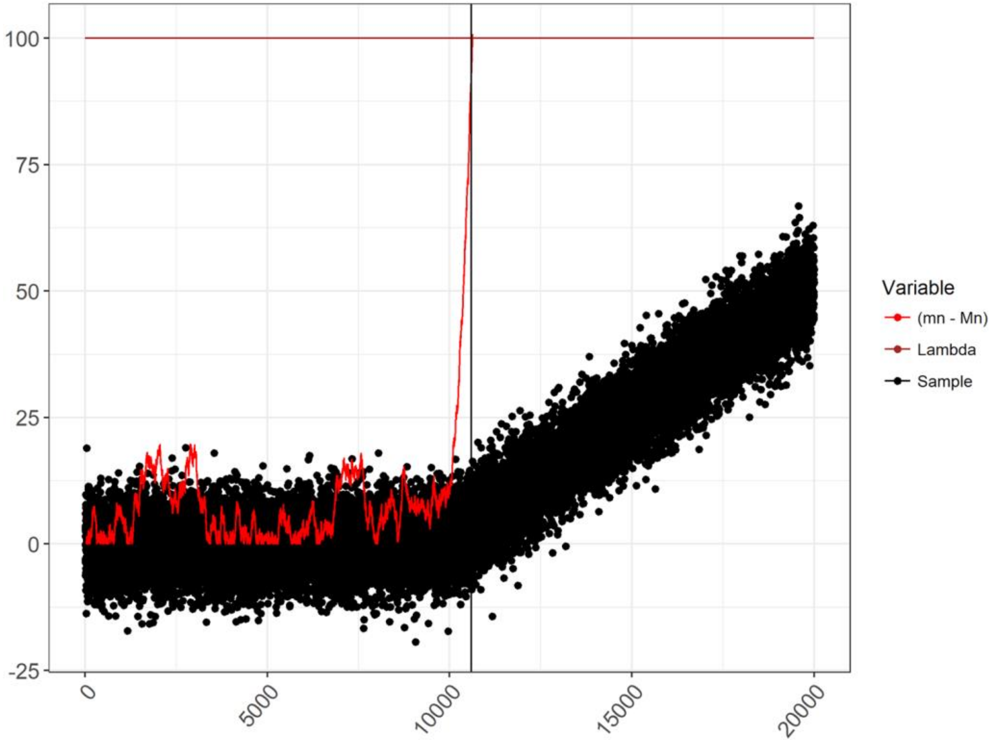
Expanding rules



- To further specialize a current rule after observing enough data
  - Select n combinations (random, smart guess...) of attributes/splitpoints
  - Calculate entropy (measuring the randomness of data) on the outcome distribution
  - Hoeffding bound (as in Gama, 2010); statistical test to decide if the best scored split significantly reduces the metric

★ Non-parametric approach (finding spatiotemporal patterns in the network)
★ Minimum number of assumptions (i.e. maximizing the outcome probability)
★ Better interpretability than black-box models

Online learning: Sudden change



- Concept drift detection. Algorithm used based on the Page-Hinkley test.
- It starts to monitor the rule's mean error when a new rule is built. Rule mean error should be located at 0.
- When a change is detected, the rule is removed from the ruleset.
  - Other approaches could be considered: changing the ruleset structure, merging rules...
- This kind of (sudden) change is handled at rule level

Rule prediction models

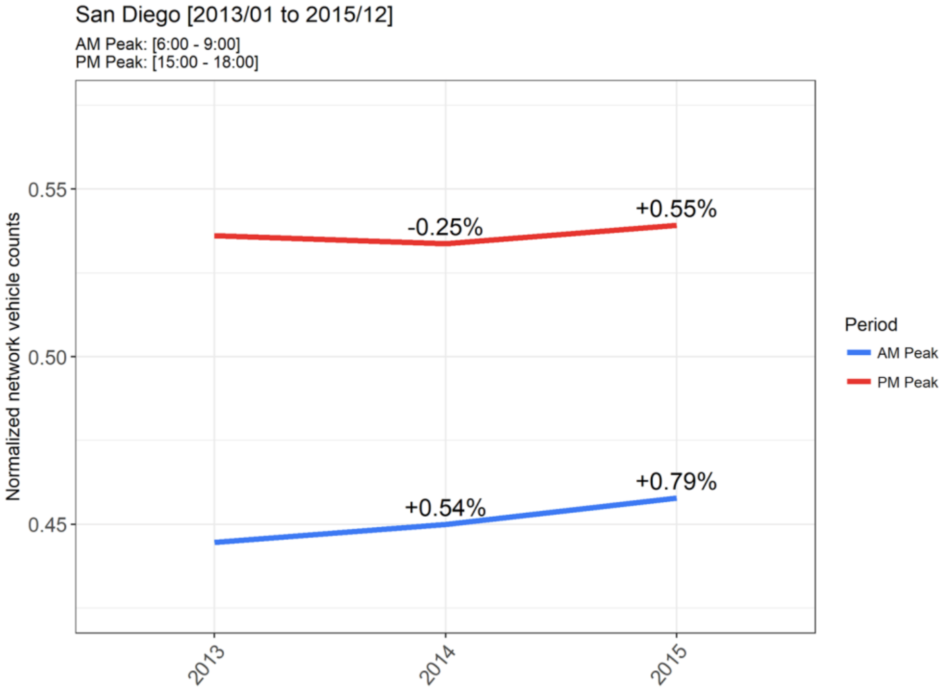
- **Weighted (historical) mean** (in the scope of the rule)
- **LASSO: Sparse linear regression** to capture the spatial dependencies in the network:

$$\min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{N} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1 \right\}$$

  - High-dimensional problem (San Diego district 11 has **+1500** detection stations)

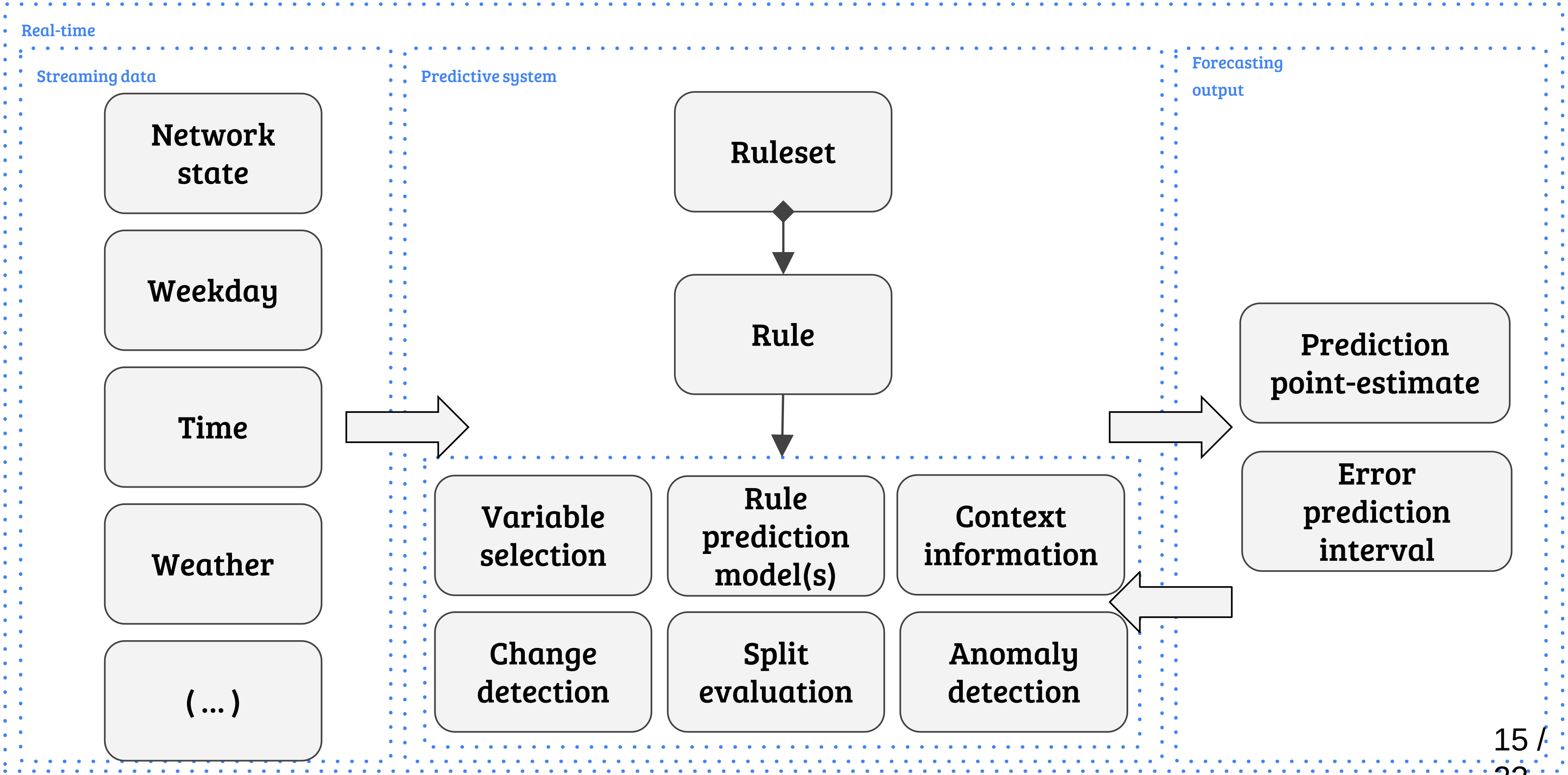
$$n \ll p$$

Online learning: Gradual change



- Seasonality, traffic demand growth...
- This kind of gradual change is handled at rule predictor level.
- Specific solution for each rule predictor
  - Weighted historical mean: age decaying factor
  - LASSO: coordinate-wise descent with soft-thresholding

Adarules

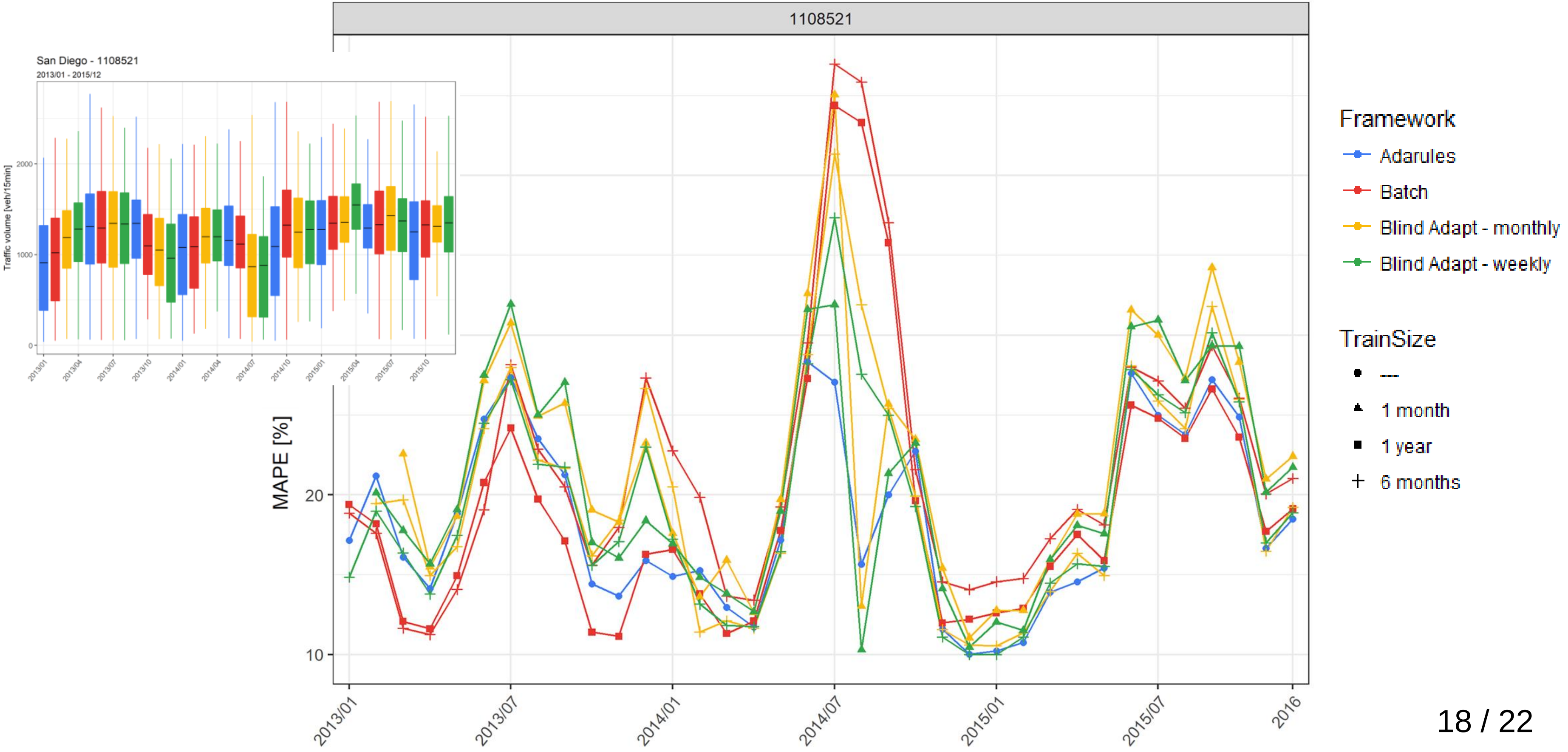

Real-time
Streaming data
Network state
Weekday
Time
Weather
(...)
Predictive system
Ruleset
Rule
Variable selection
Rule prediction model(s)
Context information
Change detection
Split evaluation
Anomaly detection
Forecasting output
Prediction point-estimate
Error prediction interval

Results

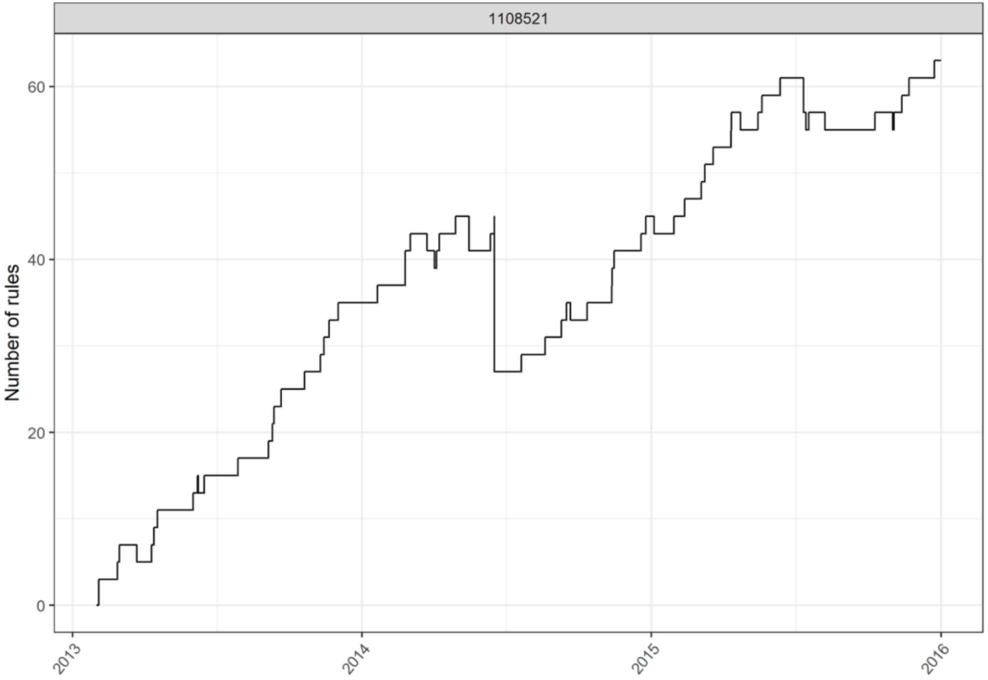
60-min traffic flow prediction

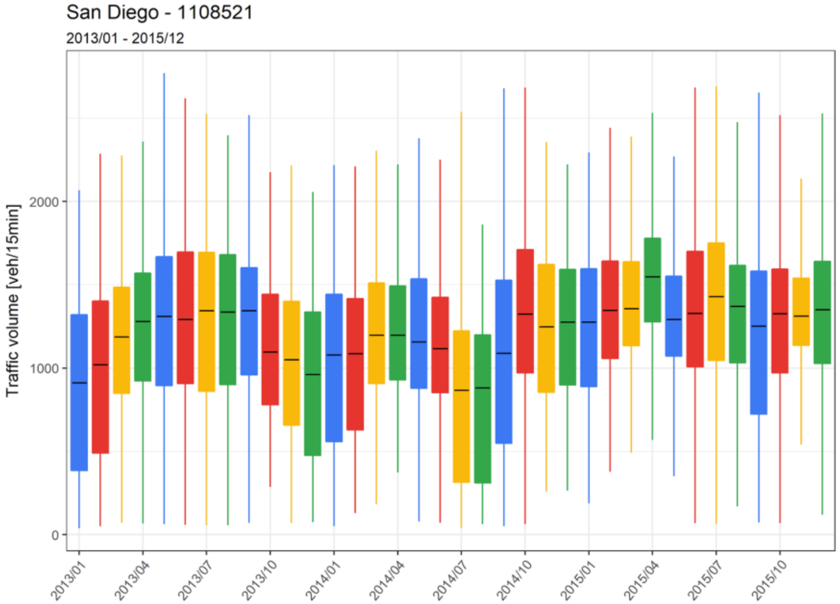
- Dataset: 2013/01 to 2015/12
- Tested approaches
  - **Adarules** (real-time)
  - Lassos for each 15-min interval trained in **batch** mode
    - **1 year** train data set (2013/01 to 2013/12)
    - **6 month** train data set (2013/01 to 2013/06)
  - Lassos for each 15-min interval **retrained** (blindly) every **month**
    - Using the last **6 month** as training data
    - Using the last **1 month** as training data
  - Lassos for each 15-min interval **retrained** (blindly) every **week**
    - Using the last **6 month** as training data
    - Using the last **1 month** as training data

60-min traffic flow prediction



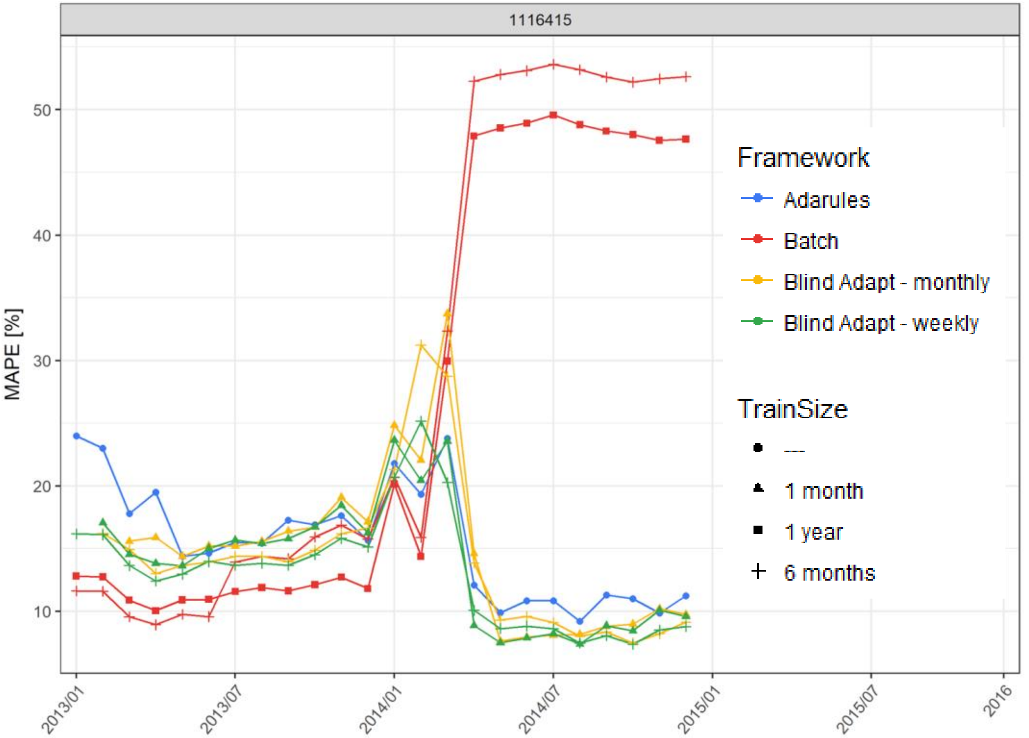
60-min traffic flow prediction

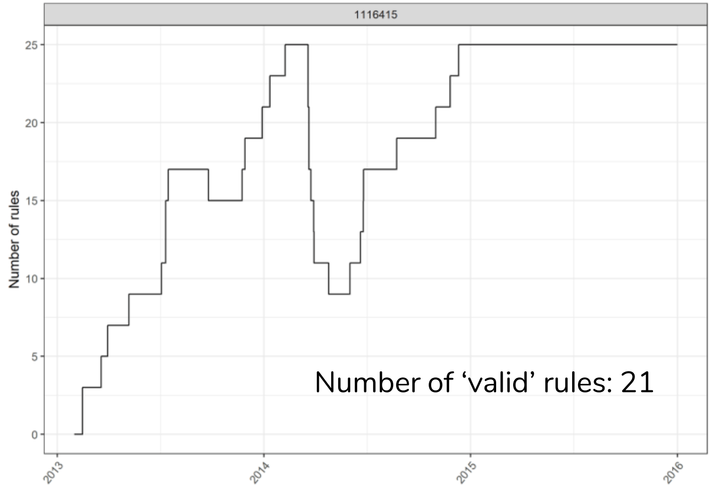


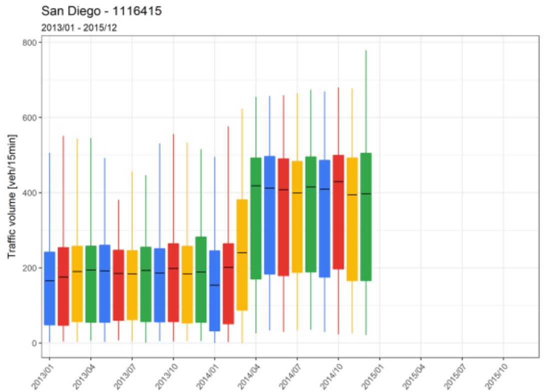


Number of 'valid' rules: 48

60-min traffic flow prediction







Conclusions & Future work

Conclusions

- ➔ Fast adaption to change
- ➔ Autonomy to decide the best decisions with more data
- ➔ Interpretable spatiotemporal patterns for traffic managers
- ➔ Prediction accuracy is important, but not the only criteria (Karlaftis and Vlahogianni, 2011; Kirby et al., 1997). Autonomy, maintenance and adaptation, interpretability

Future work

- ➔ Multi-task learning
- ➔ Incident management
- ➔ Improving real-time efficiency

Adarules:
Learning rules for real-time road-traffic prediction

Rafael Mena-Yedra[1,2]
Ricard Gavaldà[2]
Jordi Casas[1]

[1] TSS-Transport Simulation Systems, Spain
[2] Universitat Politècnica de Catalunya, Spain

20th EURO Working Group on Transportation Meeting (EWGT 2017)
4-6 September 2017, Budapest, Hungary

Aimsun

TSS TRANSPORT SIMULATION SYSTEMS

UPC UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

DOCTORATS INDUSTRIALS

References

- Gama, J., 2010. Knowledge discovery from data streams. CRC Press.
- Almeida, E., Ferreira, C., Gama, J., 2013. Adaptive Model Rules from Data Streams, in: Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases - Volume 8188, ECML PKDD 2013. Springer-Verlag New York, Inc., New York, NY, USA, pp. 480–492.
- Kirby, H.R., Watson, S.M., Dougherty, M.S., 1997. Should we use neural networks or statistical models for short-term motorway traffic forecasting? Int. J. Forecast. 13, 43–50.
- Vlahogianni, E.I., Karlaftis, M.G., Golias, J.C., 2014. Short-term traffic forecasting: Where we are and where we're going. Transp. Res. Part C Emerg. Technol., Special Issue on Short-term Traffic Flow Forecasting 43, Part 1, 3–19.
- Page, E., 1954. Continuous inspection schemes. Biometrika 41, 100–115.
- Hoeffding, W., 1963. Probability inequalities for sums of bounded random variables. J. Am. Stat. Assoc. 58, 13–30.
- Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 33, 1.
- Hastie, T., Tibshirani, R., Wainwright, M., 2015. Statistical Learning with Sparsity: The Lasso and Generalizations. Chapman and Hall/CRC.