A comparative analysis of implicit models for real-time short-term traffic predictions

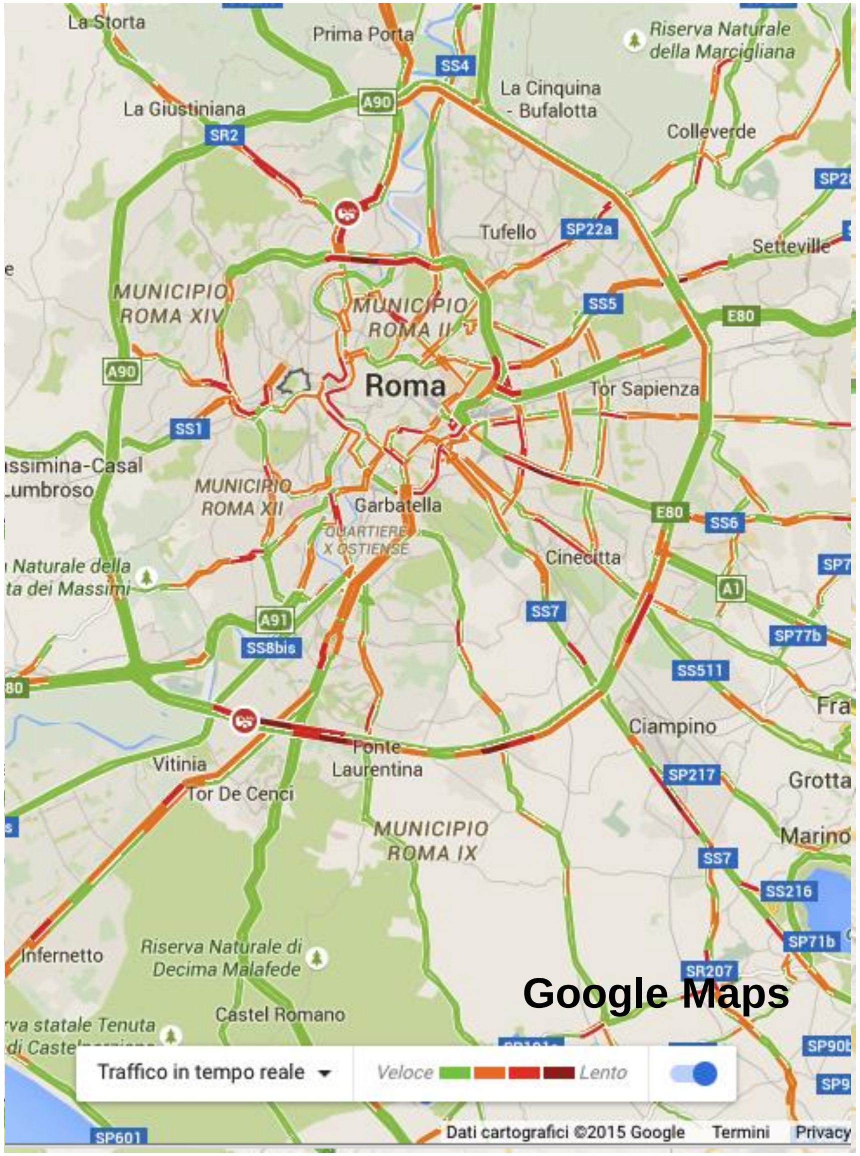
SAPIENZA
UNIVERSITÀ DI ROMA

Authors:

Gaetano Fusco
Chiara Colombaroni
Natalia Isaenko

Introduction

- Traffic maps are becoming familiar to most travelers.
- Tracking connected personal devices allows major web providers (Google, Apple) and specialized companies (TomTom, Inrix) supplying real-time traffic information.

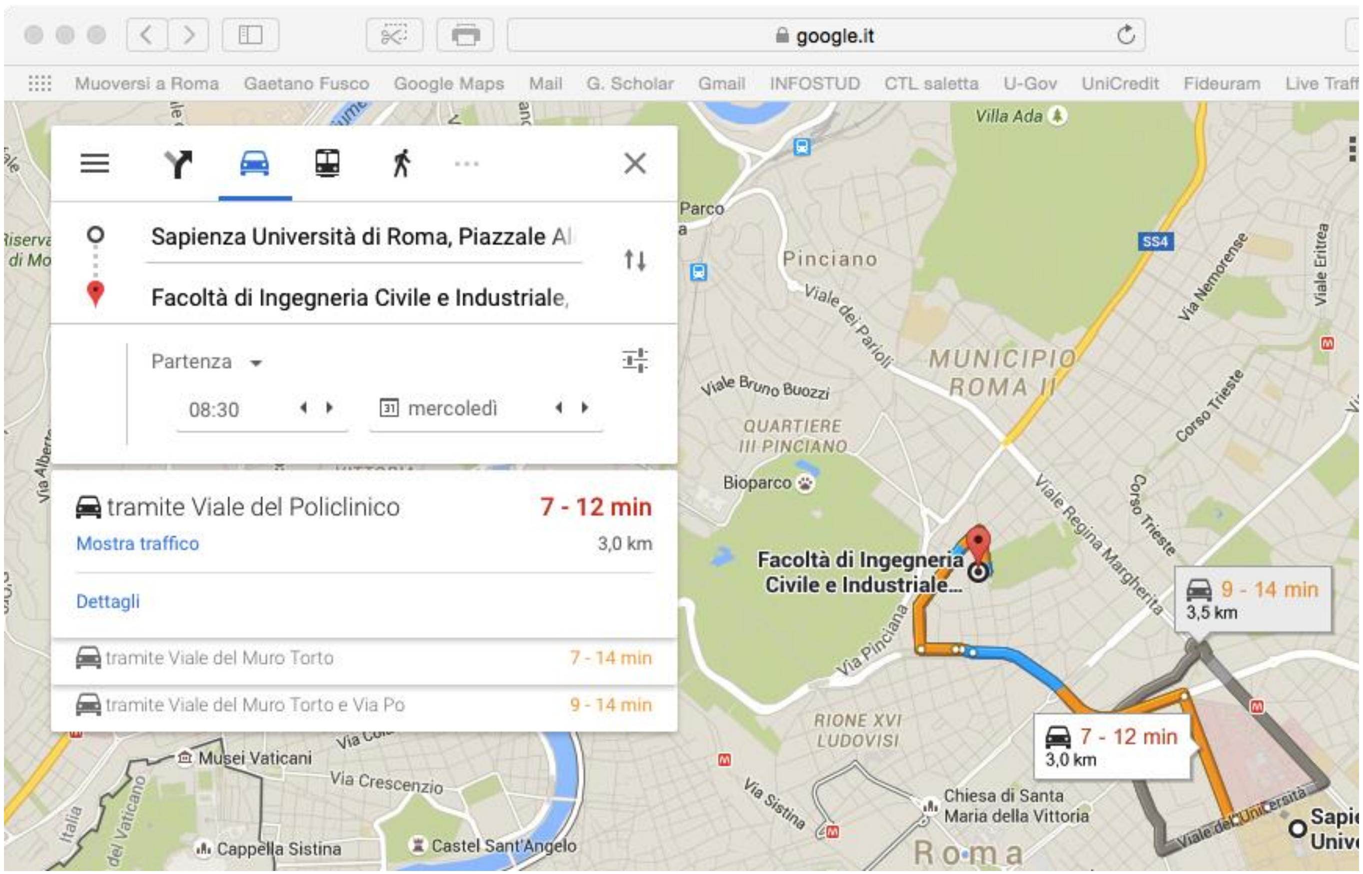


Introduction

- On-board and web traffic information services display both real-time and usual traffic conditions.
- They also supply users with route suggestions based on the desired departure time.
- However, methods for predicting future travel times are not known.

Introduction

- Precision of prediction methods is not yet fully satisfactory.

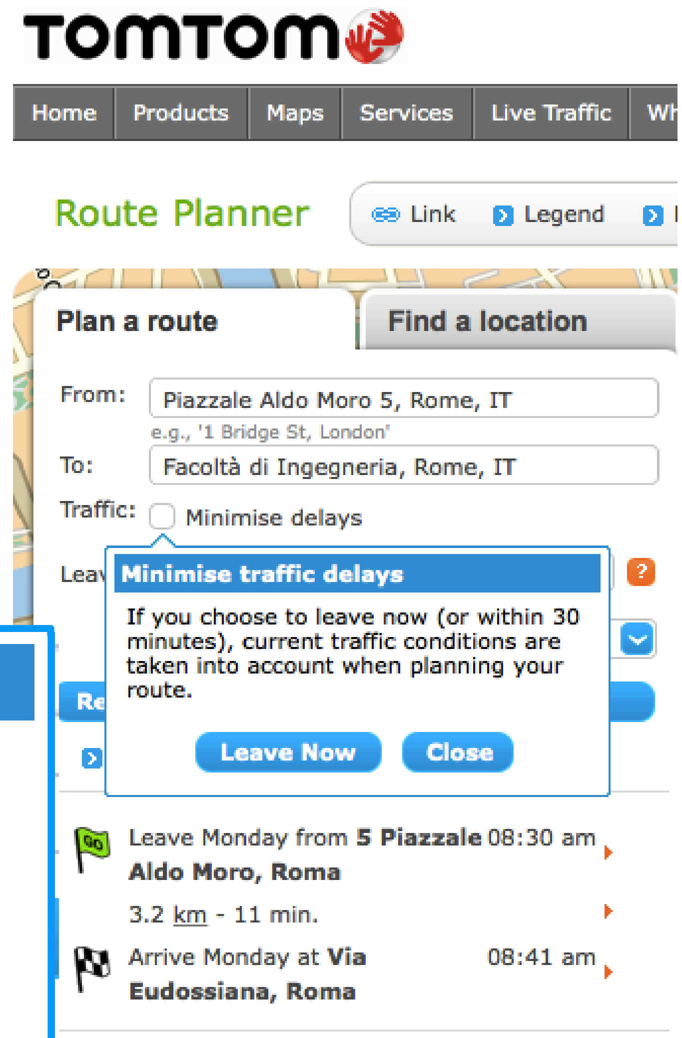


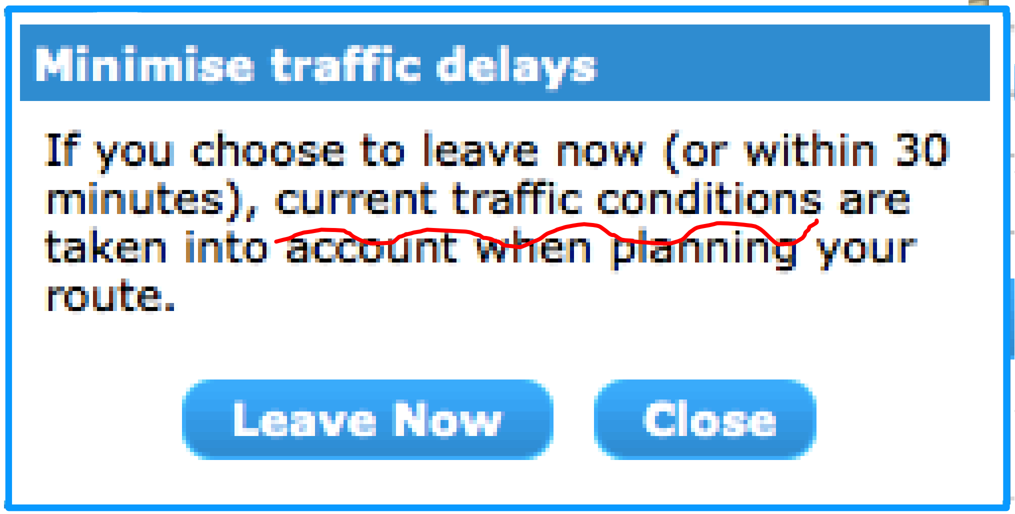
Introduction

- Since ‘70s, academics developed many prediction methods to allow Intelligent Transportation Systems proactive features.
- Explicit models simulate demand-supply interactions: they can estimate unobserved conditions and prevent overreaction.
- Implicit models extrapolate short-term predictions from observed trend (data-driven): they are often applied to short-term predictions on freeways, because of their simple physical structure.

Introduction

- However, these models are not used in commercial applications.
- TomTom doesn’t supply dynamic traffic predictions.





Object of the study

- We want to assess implicit models in short-term traffic predictions on urban networks.
- Advantages of implicit models:
  - Standard automatable calibration methods
  - Fast computation for real-time applications
  - Easy parallelization for dealing with large graphs
  - Independence from specific conditions for **easy generalization**
- Drawbacks of implicit models:
  - Lack of correspondence with the physics of vehicle traffic
- Specific goal:
  - **Relate model structure to network topology**

State of the art of implicit models

- Time-series models (AR, ARMA, ARIMA, SARIMA)
- Artificial Neural Networks (ANN): Feed-Forward, Kohonen Maps, Fuzzy NN, NAR, NARX)
- Bayesian networks
- Non-parametric estimates (clustering, nearest neighbor, fuzzy classifiers)
- Combination of different methods (Bayesian committee, Bayesian combination, ARIMA-NN)

Definitions

- Short-term implicit prediction model

$$\hat{v}^a_{t+1} = f\left(v^a_t, v^a_{t-s}, v^a_{t-2s}, \ldots, v^a_{t-hs}\right)$$

– a: prediction link; t: current time
– h: rolling horizon; s: rolling step

- Multi-input model (data from different links)

$$\hat{v}^a_{t+1} = f\left(\mathbf{V}_t, \mathbf{V}_{t-s}, \mathbf{V}_{t-2s}, \ldots, \mathbf{V}_{t-hs}\right);\ \mathbf{V}_t = \left\{v^a_t, v^b_t, \ldots, v^m_t\right\};\ a, b, \ldots, m \in L$$

- Multi-input with exogenous variables

$$\hat{v}^a_{t+1} = f\left(\mathbf{v}_t, \mathbf{v}_{t-s}, \mathbf{v}_{t-2s}, \ldots, \mathbf{v}_{t-hs}; \mathbf{X}_t, \mathbf{X}_{t-s}, \mathbf{X}_{t-2s}, \ldots, \mathbf{X}_{t-h's}\right)$$

$$\mathbf{X}_t = \left\{x^a_t, x^b_t, \ldots, x^m_t; y^a_t, y^b_t, \ldots, y^m_t; \ldots; z^a_t, z^b_t, \ldots, z^m_t\right\}$$

- Multi-step prediction

$$\left\{\hat{v}^a_{t+1}, \hat{v}^a_{t+2}, \ldots, \hat{v}^a_{t+p}\right\} = f\left(\mathbf{v}_t, \mathbf{v}_{t-s}, \mathbf{v}_{t-2s}, \ldots, \mathbf{v}_{t-hs}; \mathbf{X}_t, \mathbf{X}_{t-s}, \mathbf{X}_{t-2s}, \ldots, \mathbf{X}_{t-h's}\right)$$

Feed-Forward Neural Network

- Well known A.I. paradigm inspired by natural brain biology

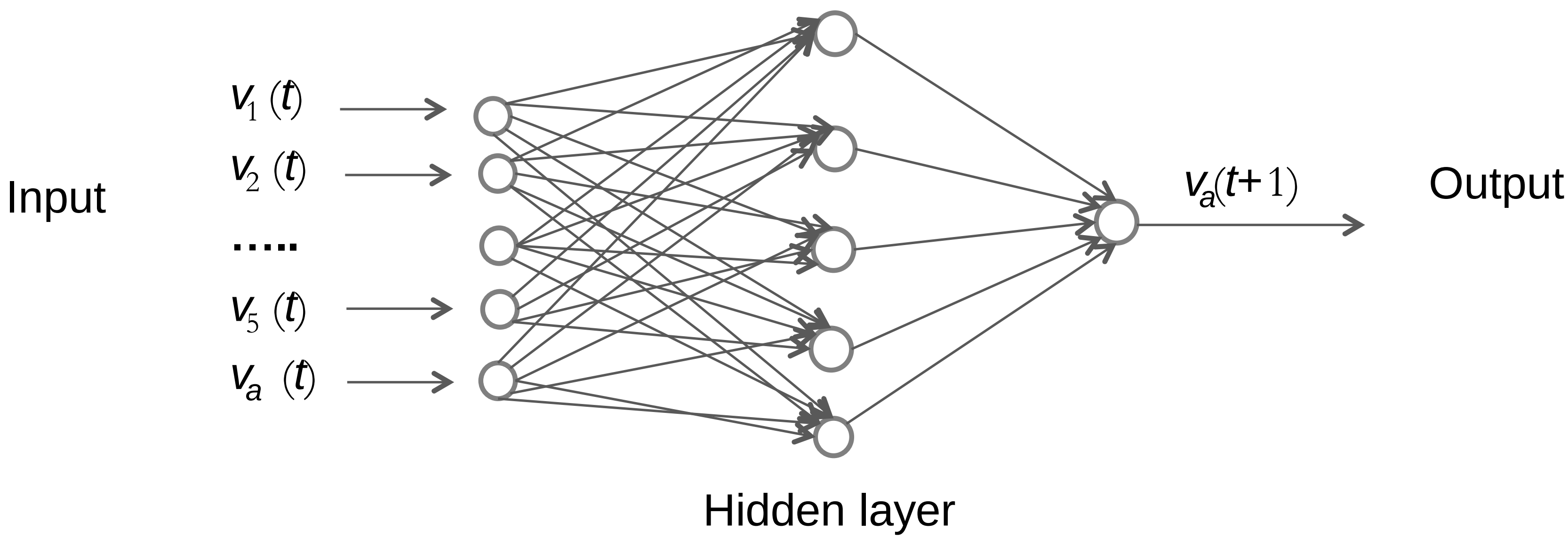
$$\hat{v}_a(t+1) = f\left(\mathbf{Cz} + J_C\right)$$

$$\mathbf{z} = f\left(\mathbf{Bv}\left(t, t-1, \ldots, t-h\right) + J_D\right)$$

– f: nonlinear function

□ ϑ_C, ϑ_D: thresholds

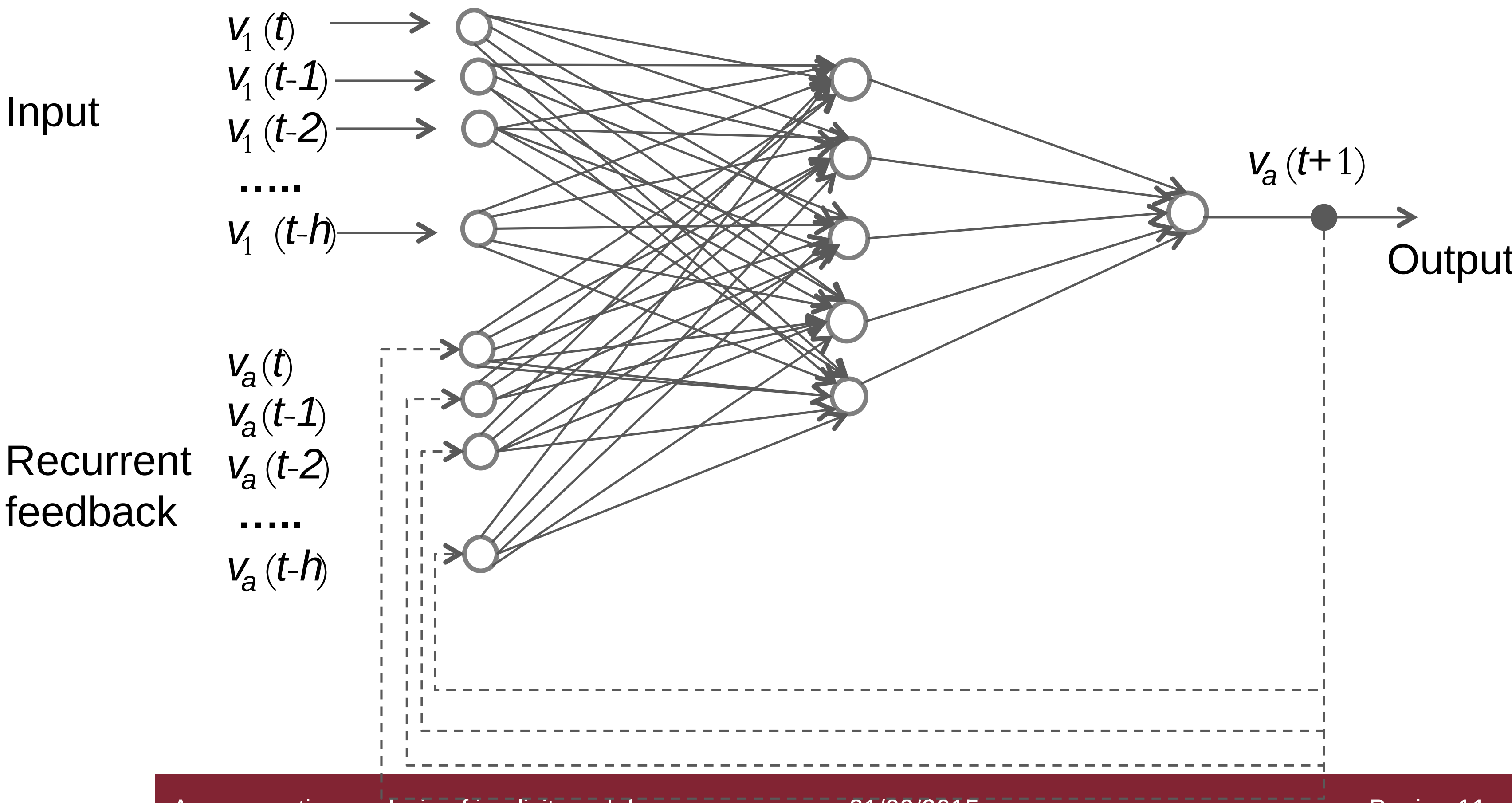
– $\mathbf{B}$, $\mathbf{C}$: weight matrices



Time dependent Neural Network

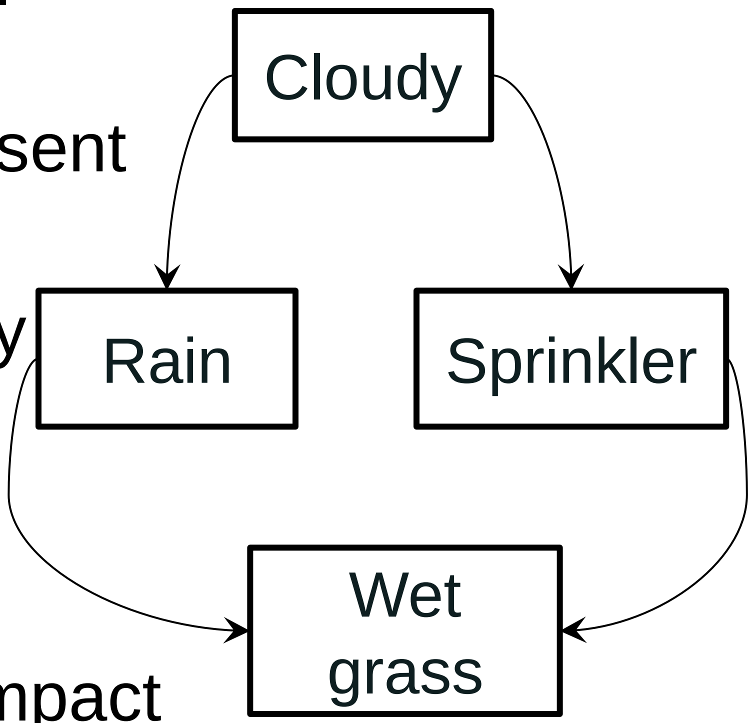
$$\hat{v}_a(t+1) = f\left(\hat{v}_a(t), \hat{v}_a(t-1), \ldots, \hat{v}_a(t-h); \mathbf{v}(t), \mathbf{v}(t-1), \ldots, \mathbf{v}(t-h)\right)$$

Non AutoRegressive eXogenous model (NARX)



Bayesian Networks (BN)

- Bayesian Networks are graphical models whose nodes represent random variables and arcs represent conditional assumptions.
- Unlinked nodes are stochastically independent variables.
- Link direction indicate cause-effect relationship.
- Network structure provides a compact representation of joint probability:

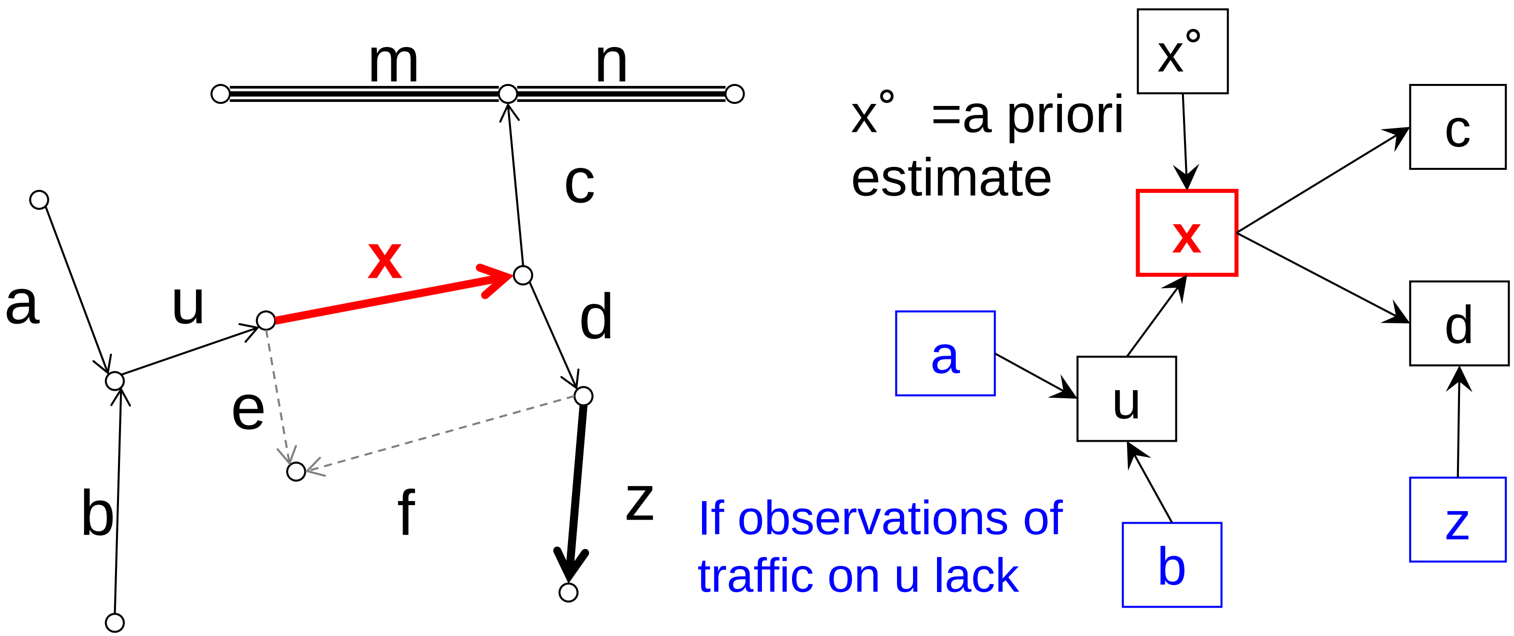


$$P(C, S, R, W) = P(C) \cdot P(S|C) \cdot P(R|C) \cdot P(W|S,R)$$

Application of BN to traffic prediction

- Nodes of the Bayesian Network are traffic variables
- Bayesian Network architecture should reproduce traffic interaction on the road network
- Cause-effect mechanisms are hidden and stochastic independence prevails

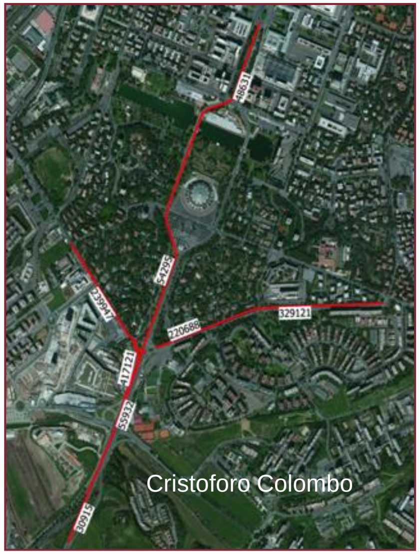
“z” affects “x” because it is parent of its children



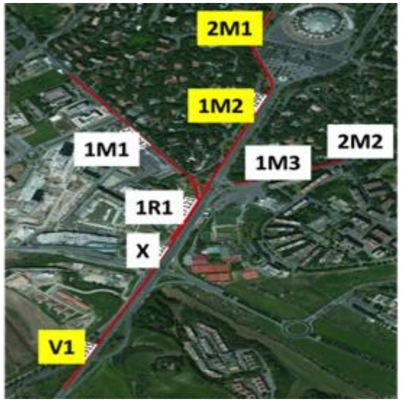
Experimental application

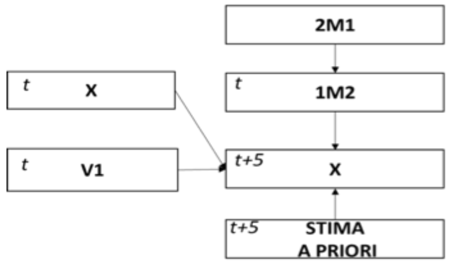
- TomTom HD data (speed):
  - 6 sub-areas in the town of Rome
  - 3 months of observations (March 2014 – May 2014)
  - 5-min. intervals speed aggregation
  - qData reliability expressed by “confidence factor”
- Forecasts for 5-15-30-45-60 minutes:
  - ANN NARX
  - ANN Feed Forward
  - Bayesian Network
  - Naïve method
  - Historical Average
- Research project with DUEL SpA (funded by FILAS)

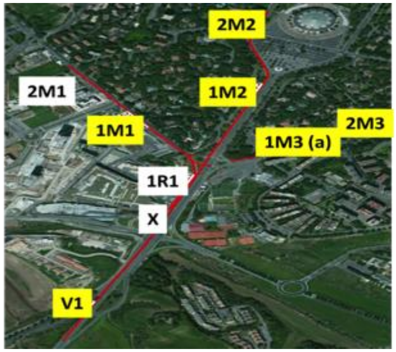
Real case studies: TomTom Dataset

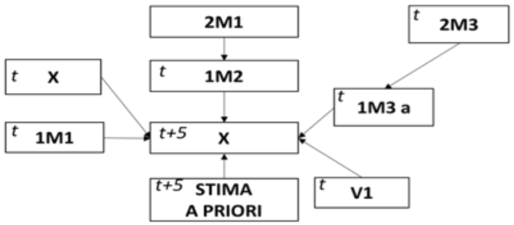


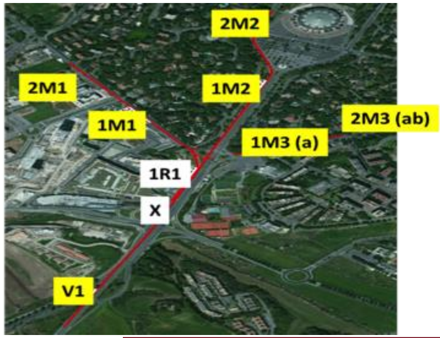
Bayesian Network Architecture

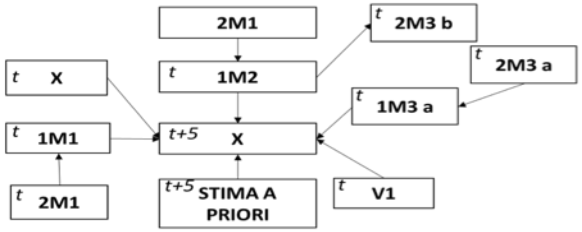








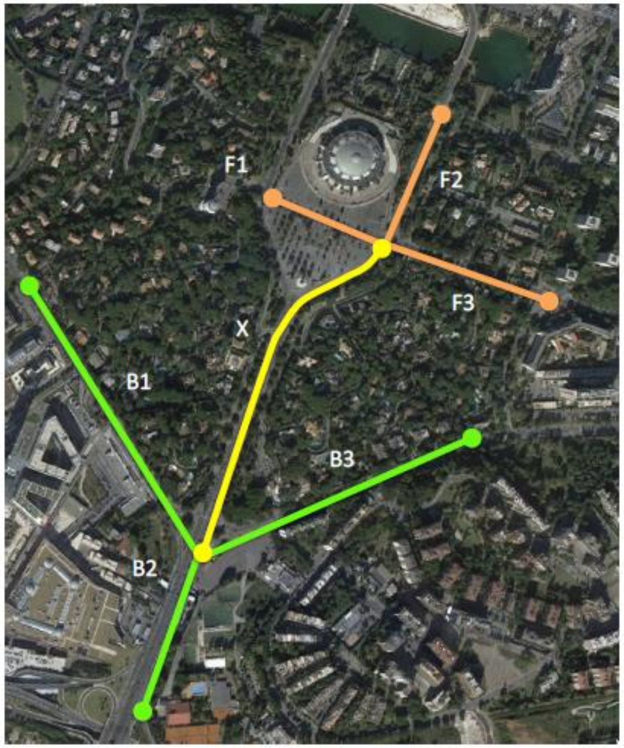


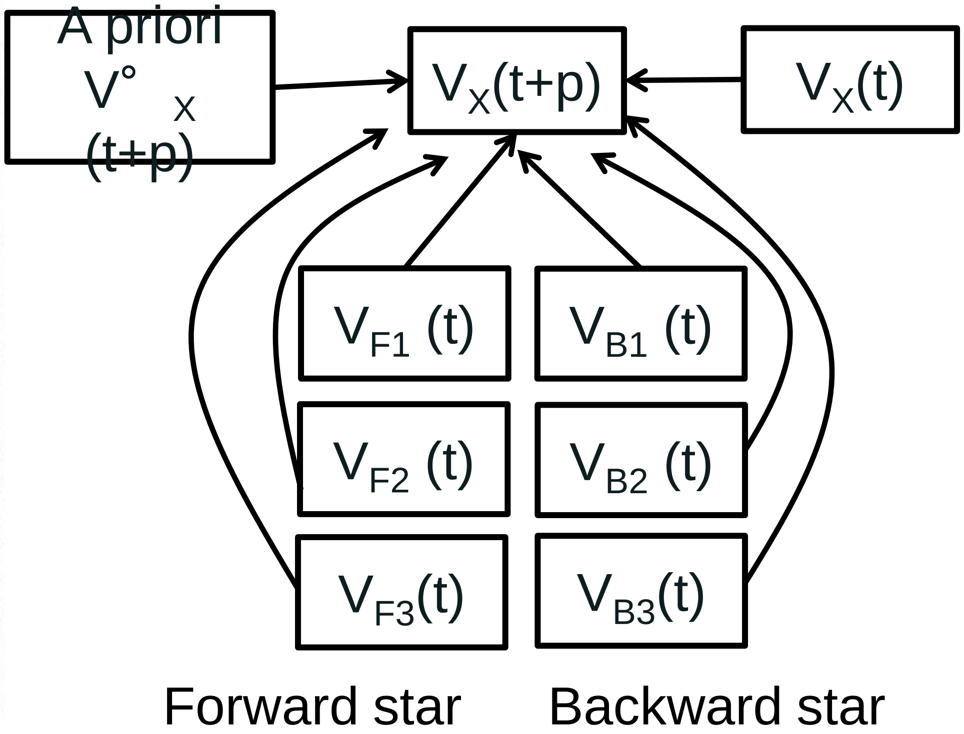


- Different architectures were tested to reflect physical relationships between close links.
- Serial links are useful when some data of a parent node are missing

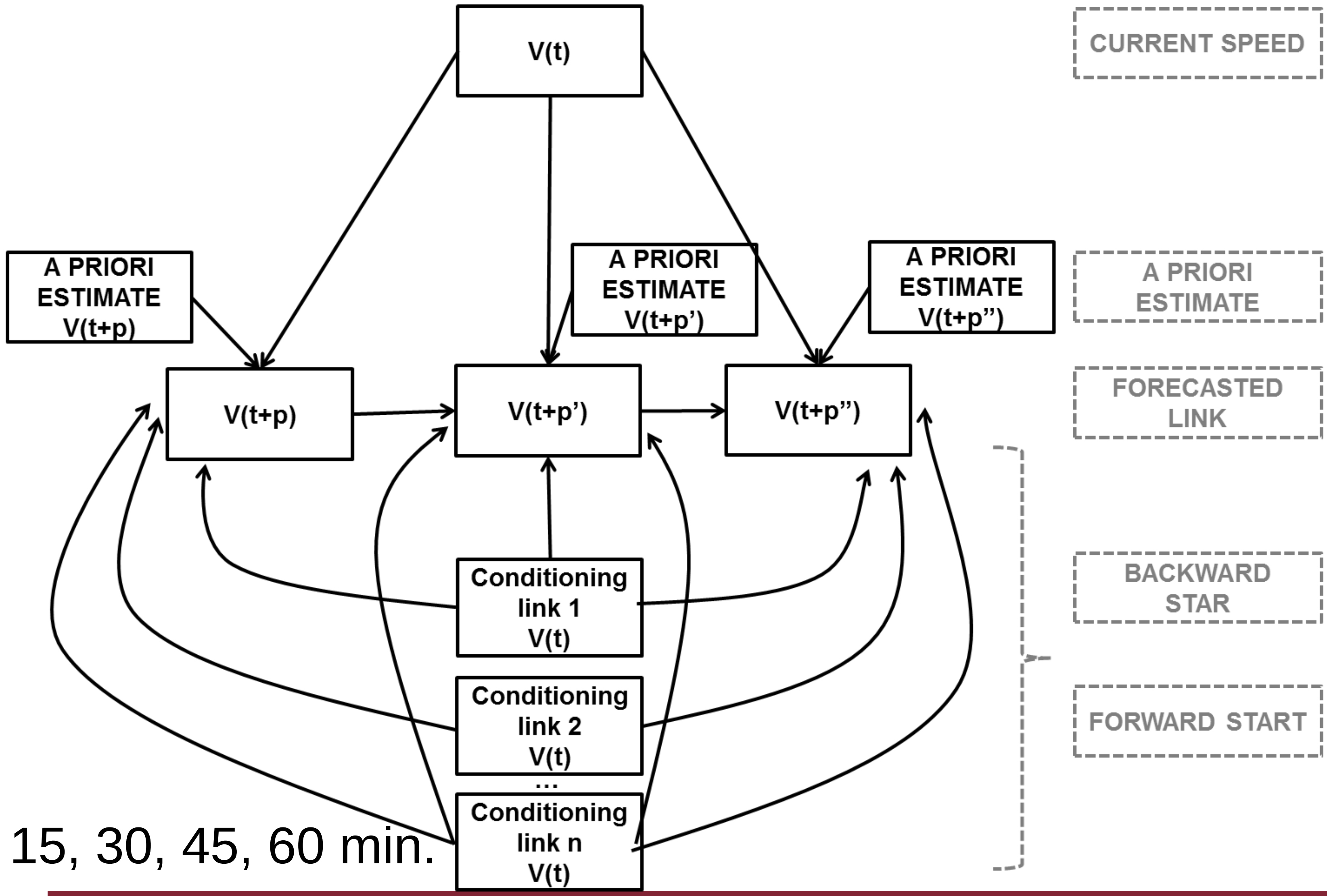
Model test

- Different model network architecture related to road network topology
- Different number of parameters and training algorithms





Multistep BN prediction horizon



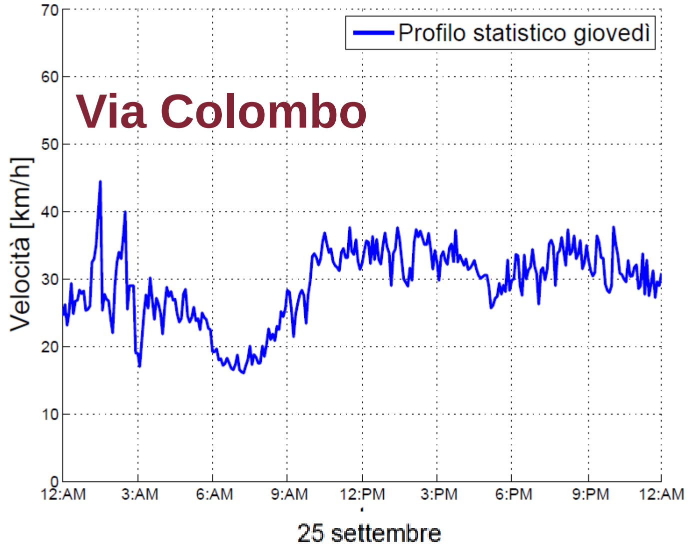
p=5, 15, 30, 45, 60 min.

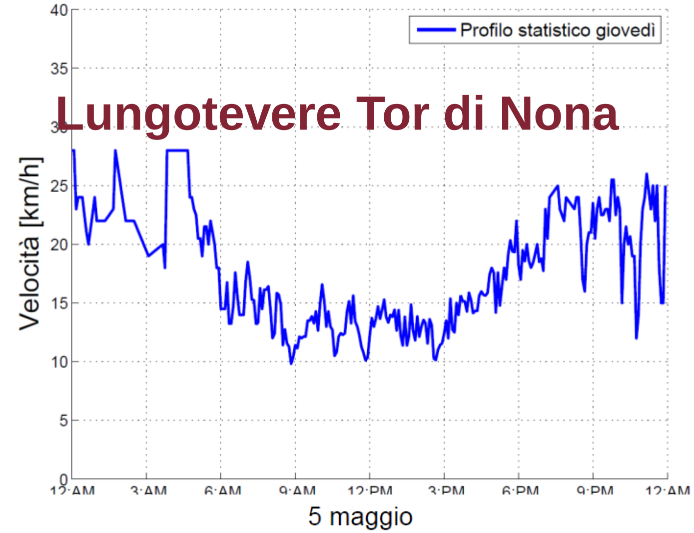
Application of Neural Networks

- Neural Network architecture:
  - One hidden layer;
  - 2-40 hidden neurons;
  - 4 different architectures considered;
  - Training algorithms: classical Levenberg-Marquardt and Levenberg-Marquardt with Bayesian regularization;
  - Input-output correlation, convergence and autocorrelation examined.
- The final structure: 12 hidden neurons; input data collected in 6 previous time intervals on forward and backward star links

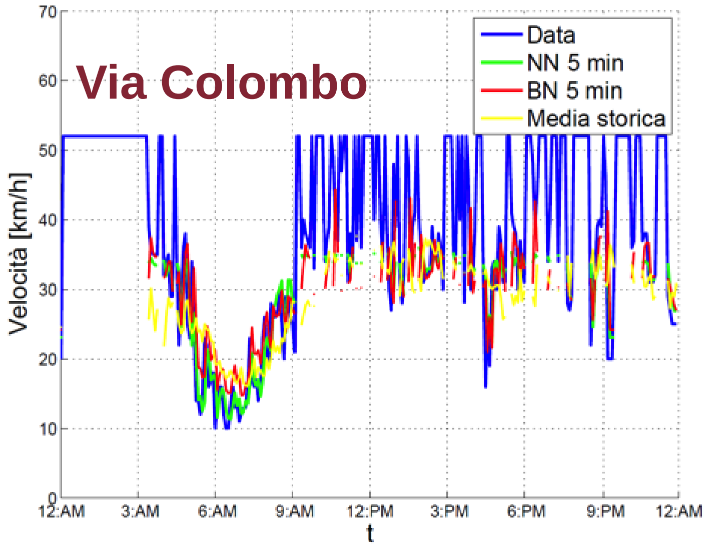
From a priori to a posteriori

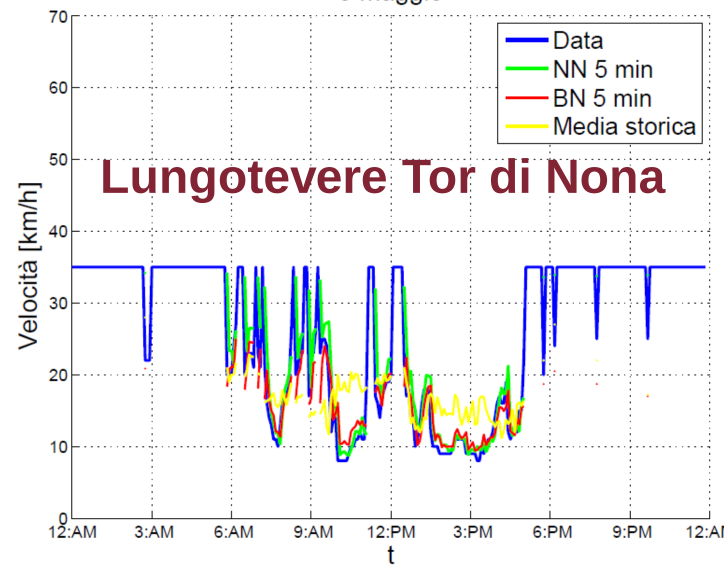
A priori estimate



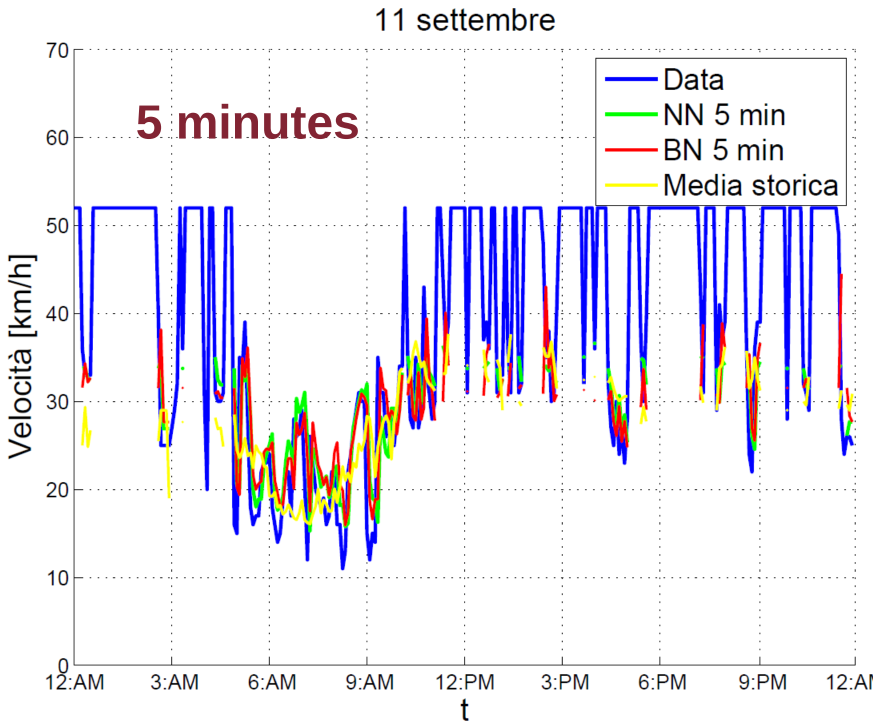


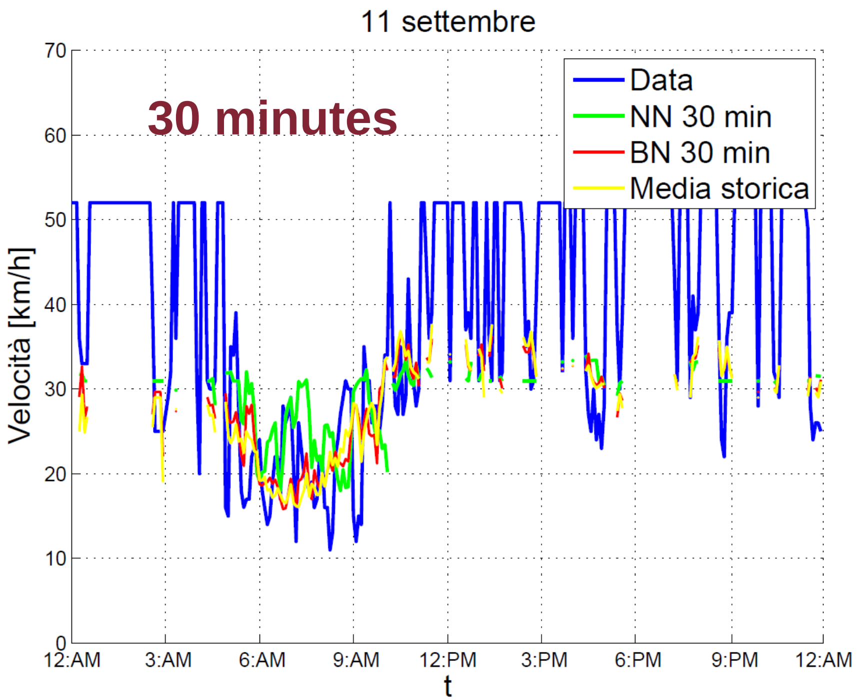
A posteriori estimate





Results: recurrent noisy congestion

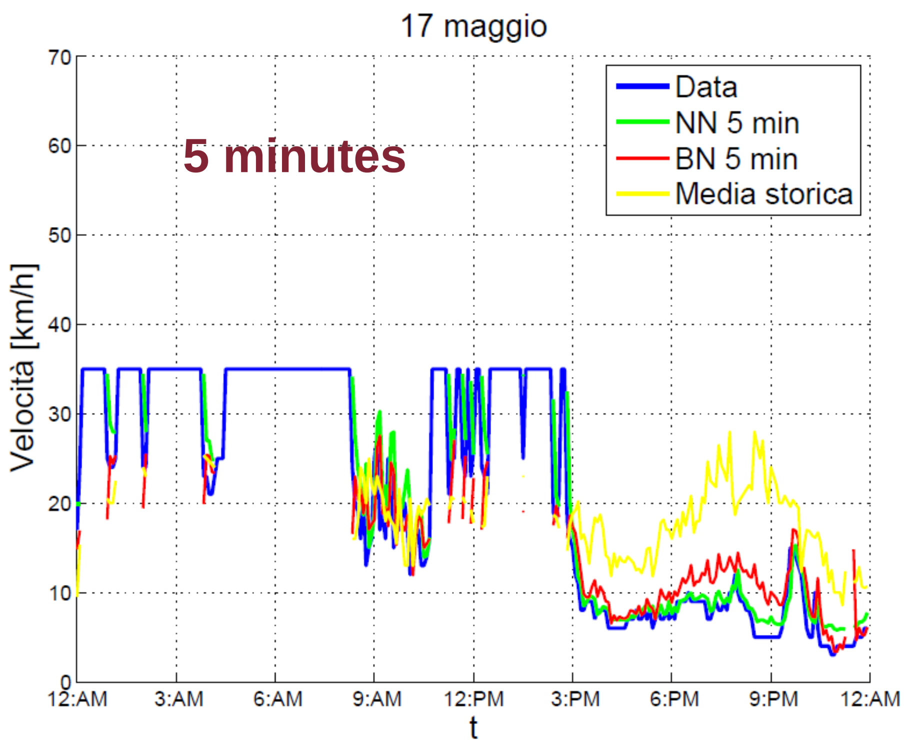


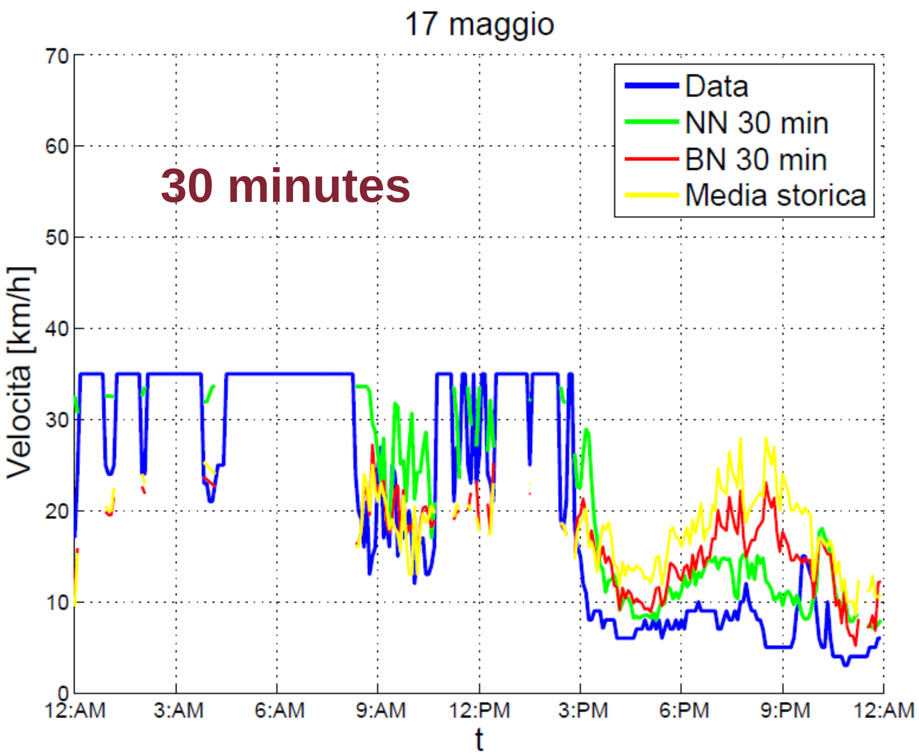


- for 5 min. interval both models provide good estimates
- for 30 min. noisy data lead to overestimate by Neural Network.

Results: non-recurrent congestion

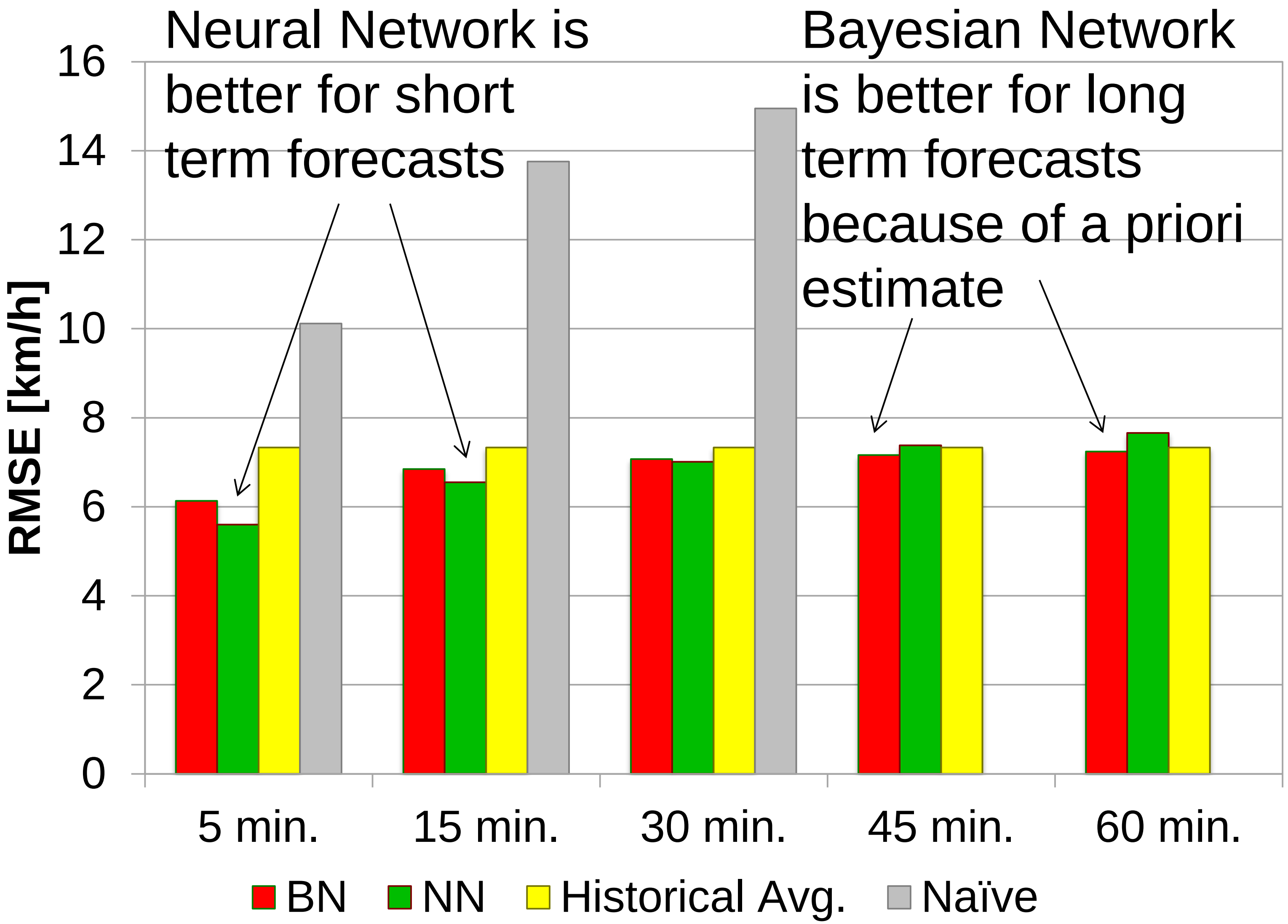
Lungotevere Tor di Nona





- for 5 minutes interval both models provide good estimates;
- for 30 minutes only Neural Network follows the actual speed trend.

Error evaluation (via C.Colombo)



Conclusions

- Application of Neural and Bayesian Networks shows that an appropriate graph structure of models can improve forecast performances with respect to historical average or naïve predictions
- Accuracy worsens as prediction horizon increases:
- FFW Neural networks can be applied up to 15-minute predictions
- Bayesian Network with statistical a priori forecast is better in standard conditions for longer predictions
- Time-dependent Neural Network outperforms BN in anomalous conditions
- A supervisor can be introduced to choose the best model depending on observed data pattern