Reinforcement Learning and its Application in Variable Speed Limit Control

Krešimir Kušić, Martin Gregurić, Edouard Ivanjko

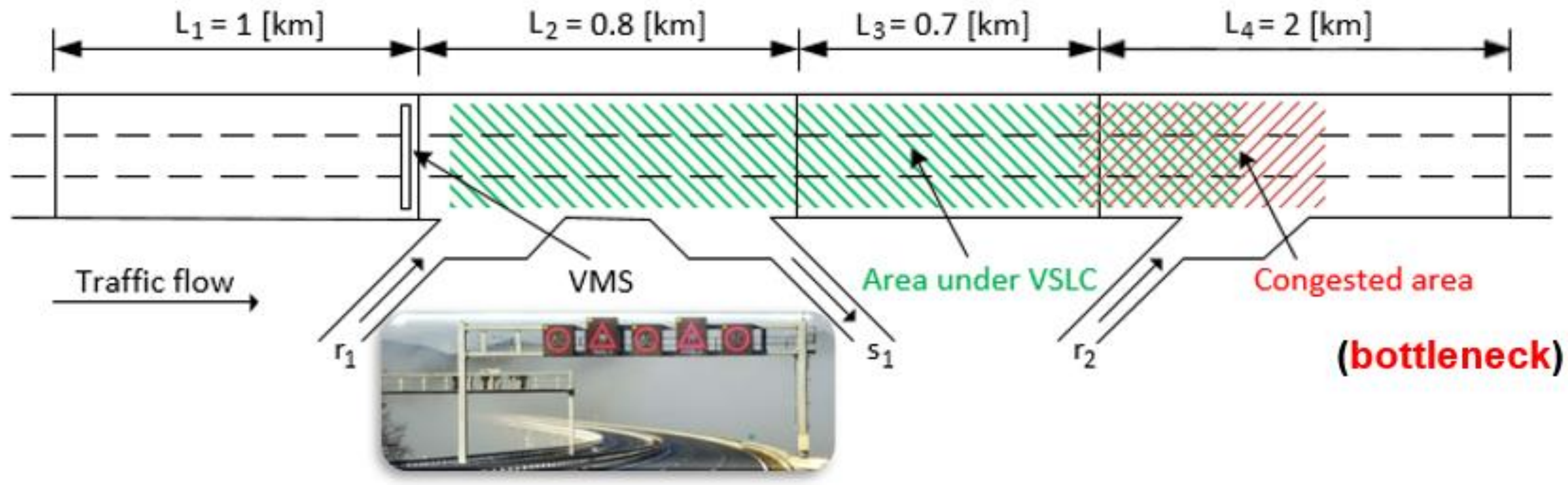
University of Zagreb, Faculty of Transport and Traffic Sciences

www.unizg.hr, http://www.fpz.unizg.hr

kresimir.kusic@fpz.hr; martin.greguric@fpz.hr; edouard.ivanjko@fpz.hr

MOTIVATION FOR VARIABLE SPEED LIMIT CONTROL

- ***Variable Speed Limit Control (VSLC) directly affects mainline traffic flow entering the bottleneck by adjusting the speed limits posted upstream of the controlled section on variable message sign (VMS)***
- ***VSLC stabilizes and homogenizes traffic flow (reducing):***
  - Number of accidents
  - Air pollution
  - Traffic jams



VSLC BASED ON REINFORCEMENT LEARNING

- **Q-learning with function approximation method for continuous state variables:**
- The normalized state vector is defined as:

$$\vec{S}_t = \left(\frac{a(t-1)}{v_f}, \frac{a(t-2)}{v_f}, \frac{\rho_2(t)}{\rho_j}, \frac{\rho_3(t)}{\rho_j}, \frac{\rho_4(t)}{\rho_j}, \frac{v_2(t)}{v_f}, \frac{v_3(t)}{v_f}, \frac{v_4(t)}{v_f}\right)$$

  - Components are previously executed speed limit value (a), current density (ρ) in section *i*, and current speed (v) in section *i* respectively
- Important properties from the state vector $\vec{s}_t$ are captured into the feature vector:

$$\vec{\phi}_{s,a} = (1 + \phi_1(s,a_1), \dots, \phi_l(s,a_1), \dots, \phi_1(s,a_7), \dots, \phi_l(s,a_7))$$

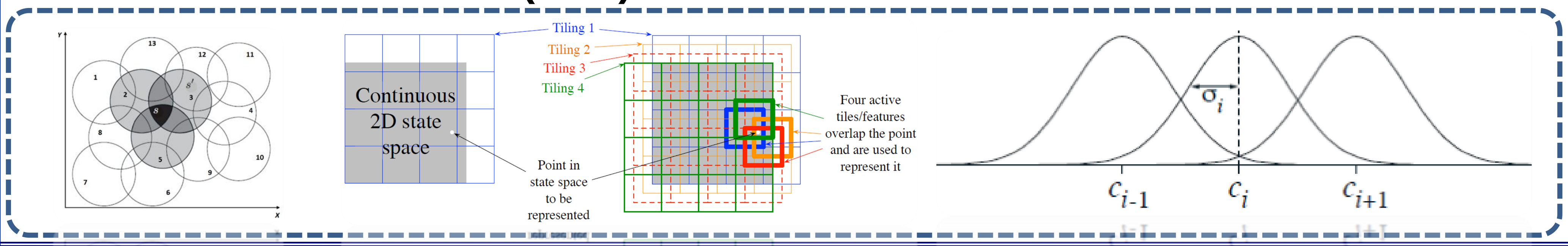
- Q-value approximated by linear combination of features and parameters:

$$Q_\theta(s,a) = \vec{\theta}^T \vec{\phi}_{s,a} \approx Q(s,a)$$

- **Goal** is to find parameter vector $\vec{\theta}$ by applying stochastic gradient descent:

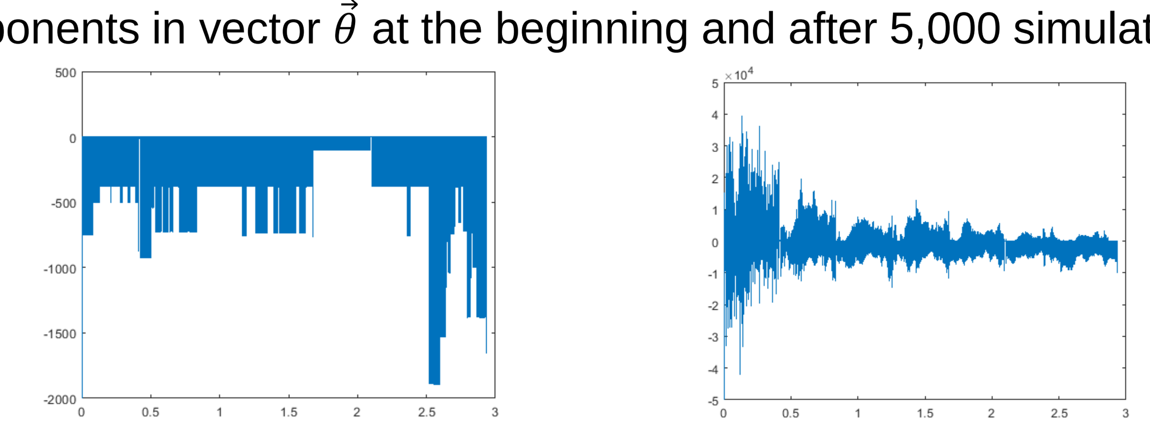
$$\vec{\theta}_{t+1} = \vec{\theta}_t + \alpha_n\left(r_{t+1} + \gamma max_{a' \in A}(\vec{\theta}_t^T \vec{\phi}_{s_{t+1},a'}) - \vec{\theta}_t^T \vec{\phi}_{s_t,a_t}\right)\vec{\phi}_{s_t,a_t}$$

- Three feature construction methods are used: **Coarse** and **Tile coding,** and **Radial Basis Function (RBF)**



- **IA technology used:** Reinforcement learning (RL) based on ***Markov Decision Process (MDP). MDP in VSLC is described by a 5-tuple*** $\langle S, A, \pi, R, \gamma\rangle$
- ***Q-learning with lookup table:***
- ***States*** *(traffic density, average speed, speed limit value from previous control time interval)*
- ***Actions*** *(changing the speed limit values in the current control time step)*
- **Policy function:** Learned during operation $\pi: S \rightarrow A$
- ***Rewards*** (Proportional to TTS) depend on: Chosen action or **Policy function**
- **γ** - influence of future rewards over the agent's behavior
- **Q-learning** with **function approximation method** is used as the RL method

Components in vector $\vec{\theta}$ at the beginning and after 5,000 simulations

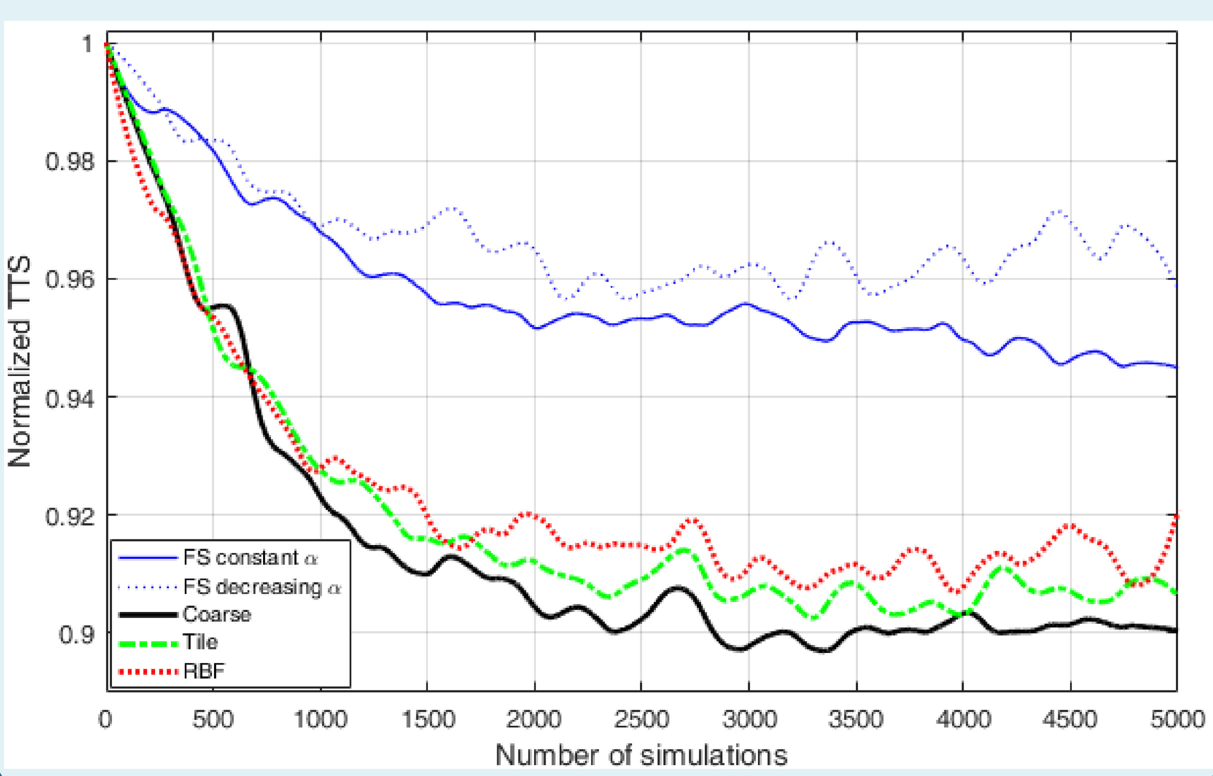


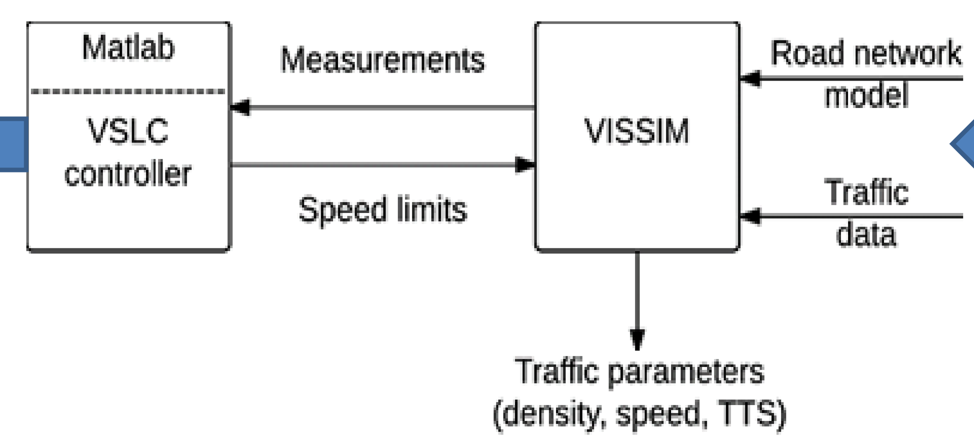
SIMULATION FRAMEWORK AND RESULTS

Evaluation results

(5,000 simulations, individual duration 2.5 $[h]$)

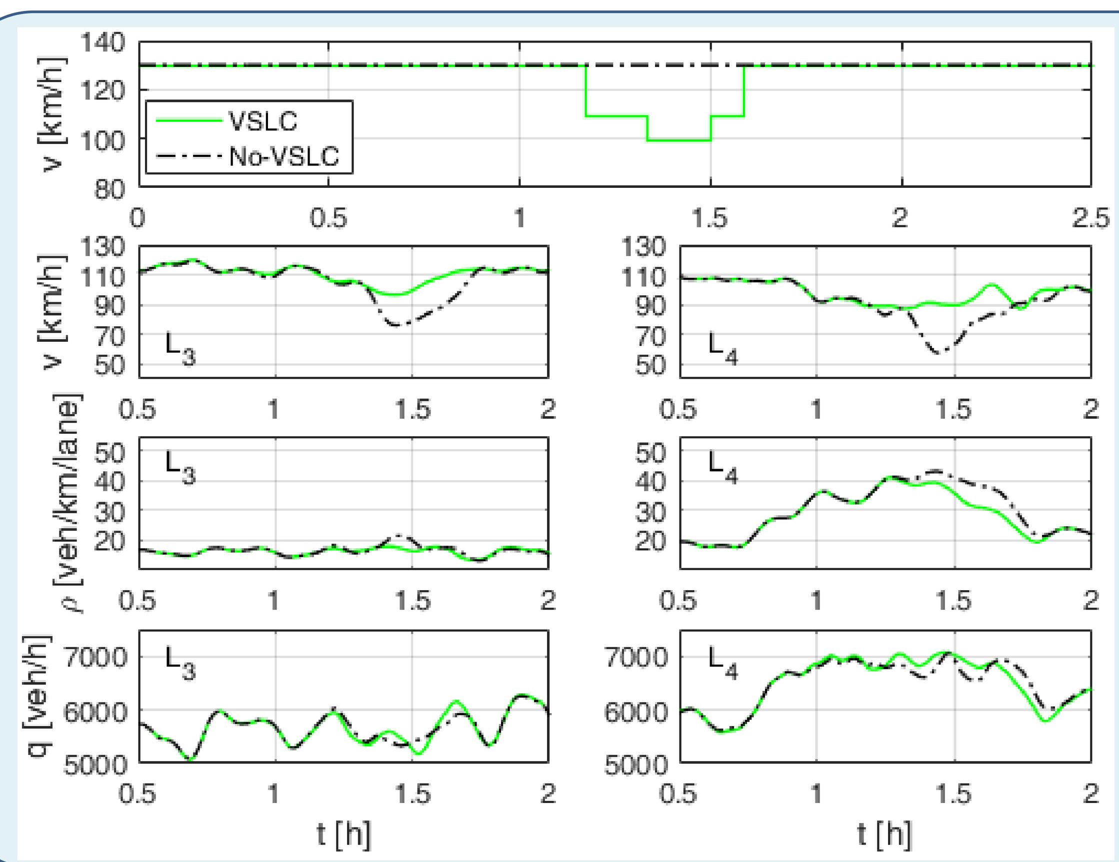
- **Q-Learning applied technique**
  - Convergence of the normalized TTS during the learning process
  - Linear function approximation approaches have a steeper decrease rate (coarse and tile coding, and RBF)





Relevant parameters for VSLC control

- **Underlying traffic process:**
  - Urban motorway traffic (input traffic data)
  - Initialization of parameters of the learning algorithm



- **Impact of VSLC on the mainstream traffic parameters is most evident in sections L_3 and L_4**
  - Typical gradual decrease and increase of the speed limit without large unallowed changes has been learned
  - Positive effect of VSLC is evident with actions that actively reduce the density compared with the case of no-control
  - Timely applied sequence of speed limits keeps the traffic flow speed at a higher value compared to the no-control case
  - Gradual reduction of the speed of vehicles coming into section L_4 allows the congestion in it to dissolve more quickly than in the case of no-control

University of Zagreb, Faculty of Transport and Traffic Sciences

Future work:

- Implemented Q-learning VSLC will be augmented to a multiagent approach to assign speed limits to several consecutive sections on the controlled urban motorway
- Deep neural networks will be used to cover the hidden feature patterns from traffic data to enable more efficient state generalization and faster learning

Publications related with VSLC and Reinforcement Learning

- Kušić, Krešimir; Ivanjko, Edouard; Gregurić, Martin. A Comparison of Different State Representations for Reinforcement Learning Based Variable Speed Limit Control // Proceedings of MED-2018. Zadar, Croatia, 19-22 June 2018, pp. 266-271
- Z. Li, P. Liu, C. Xu, H. Duan, and W. Wang, "Reinforcement learning based variable speed limit control strategy to reduce traffic congestion at freeway recurrent bottlenecks," IEEE Transactions on Intelligent Transportation Systems, vol. 18, no. 11, pp. 3204–3217, Nov. 2017
- T. Schmidt-Dumont and J. V. Vuuren, "Decentralised reinforcement learning for ramp metering and variable speed limits on highways," IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 8, pp. 1–10, 2015.