Australasian Transport Research Forum 2024 Proceedings
27-29 November, Melbourne, Australia
Publication website: https://australasiantransportresearchforum.org.au/

Cordon-based pricing for multi-modal networks: Introducing a data-driven method

Nasser Parishad[1], Mehmet Yildirimoglu[1]

[1] The University of Queensland, School of Civil Engineering, Brisbane QLD 4072, Australia

N.Parishad@uq.edu.au

Abstract

Congestion pricing is a promising approach to manage traffic flow. However, developing a real-time dynamic pricing method that considers demand elasticity and travellers heterogeneity, is a challenging task. While current methods often depend on accurate estimation of network parameters, this study aims to develop a data-driven, cordon-based pricing methodology using reinforcement learning to optimize traffic flow. A multi-modal, trip-based Macroscopic Fundamental Diagram (MFD) simulation, which accounts for travellers' heterogeneity and mode choice decisions, along with a Deep Q-Network (DQN) agent, are employed to maximizes network outflow. Without any prior knowledge of network parameters, the results indicate that the agent maintains car user accumulation at critical levels. Sensitivity analyses further demonstrate that even with a 20% margin of error in input values, the agent remains focused on alleviating traffic congestion.

1 Introduction

Traffic congestion is a significant global issue, imposing substantial economic and environmental burdens and resulting in billions of dollars lost annually. The extensive body of research in this field underscores the importance of addressing this issue. To mitigate traffic congestion, various control strategies have been proposed, including congestion pricing (Zheng et al. 2012; Li and Ramezani 2022), perimeter flow control (Geroliminis et al. 2012; Zhou & Gayah 2021), and route guidance (Yildirimoglu et al. 2015; 2018; Koh et al. 2020; Leclercq et al. 2021).

Congestion pricing, among these strategies, stands out as a promising approach (Small 1997). Previous studies have shown that pricing can influence travellers' behaviour. By imposing higher costs on car users, pricing methods can encourage the adoption of alternative departure times (Amirgholy and Gao 2017; Miralinaghi et al. 2019; Liu et al. 2023), routes (Li and Ramezani, 2022; Genser and kouvelas, 2022), or modes of transportation (Basso and Jara-Díaz 2012; Zheng et al. 2016; Li et al. 2019). Motivating people to choose more sustainable modes of transportation through pricing strategies not only alleviates congestion but also indirectly contributes to the reduction of carbon emissions (Balzer and Leclercq 2022).

Investigating control methods in large-scale networks requires a network-level representation of the city. Macroscopic Fundamental Diagrams (MFDs) provide a parsimonious regional scale representation of urban traffic dynamics (Geroliminis and Daganzo 2008). The concept of representing traffic dynamics using MFDs was first introduced by Godfrey (1969) and then further developed analytically and measured experimentally by Daganzo and Geroliminis (2008) and Geroliminis and Daganzo (2008) respectively.

While traditional accumulation-based MFDs have been effective, they lack granularity. Considering the average trip length for all travellers, accumulation-based MFDs cannot represent individual attributes. To address this limitation, trip-based simulation was proposed initially by Arnott (2013) and further investigated by Fosgerau (2015), Daganzo and Lehe (2015), Lamotte and Geroliminis 2016. Trip-based MFDs offer improved modelling capabilities by providing different trip lengths for every traveller. Nevertheless, the absence of a closed-

form analytical expression in trip-based MFD remains the challenge of finding optimal solutions.

To overcome this challenge, data-driven approaches have been introduced (Kouvelas et al. 2017; Ren et al. 2020; Wang and Paccagnan 2022). Data-driven methods offer promise in alleviating some of the limitations of traditional model-based approaches. Reinforcement Learning (RL) has emerged as a promising solution for addressing complex challenges, thanks to its model-free nature, which requires minimal prior knowledge of environment dynamics (Silver et al. 2014; Mnih et al. 2015; Sutton and Barto 2018; Hou and Xiong 2019; Szepesvári 2022). The limited effectiveness of early reinforcement learning algorithms such as TD learning, SARSA, and Q-learning in large-scale and complex environments due to their dependence on tabular representations of value functions has prompted the integration of reinforcement learning with deep learning techniques (Schmidhuber 2015). Despite these advancements, the combination of nonlinear function approximators, off-policy learning, and bootstrapping introduced challenges such as instability and divergence. The introduction of the Deep Q-Networks (DQN) algorithm (Mnih et al. 2015) marked a significant milestone in addressing these challenges by employing techniques like experience replay buffer and target network.

This study utilizes RL and trip-based MFD to develop a novel cordon-based pricing scheme able to consider travellers heterogeneity. The main objective is to maximize network outflow (minimize total travel time) by introducing a dynamic, proactive tolling strategy that adapts in real-time without any prior network parameters knowledge, accounting for demand elasticity. Using a trip-based MFD as the RL environment, the study employs a DQN agent to find the optimal cordon-based tolling strategy.

2 Traffic dynamics and modelling

2.1 Traffic dynamics

A trip based MFD framework has been developed for a single region, representing a congested city centre. All trips originate and end within the network. The simulation considers travellers' mode choice and offers three main modes: car, public transport (bus), and walking. Figure 1 illustrates the link-level representation of a segment of the network for each mode considered in this study. While car and walking modes can effectively use the entire network, bus stops are located every three nodes both in the horizontal and vertical directions. This enables a realistic representation of a multi-modal decision-making process.

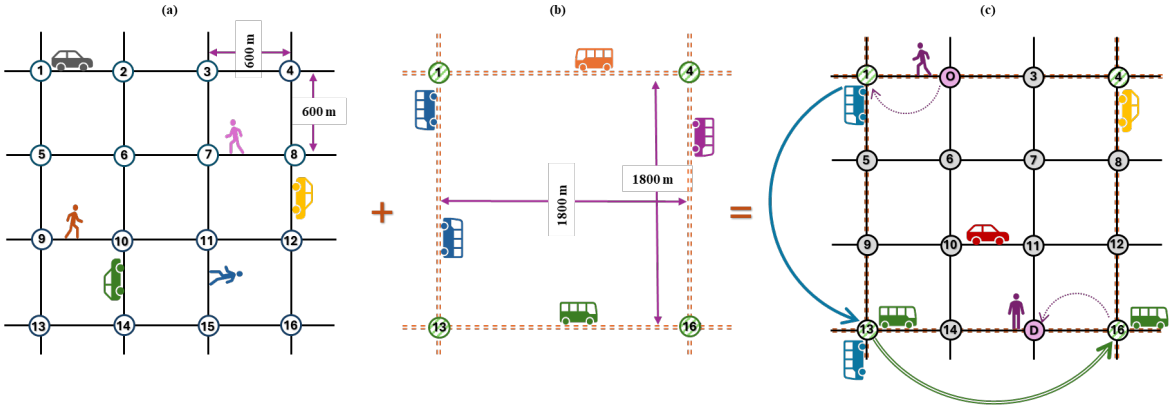

Figure 1. A portion of network layout, a) car/pedestrian network, b) public transport network, c) integrated network

A single-peaked demand profile is developed to simulate traffic congestion. Traveller heterogeneity is considered by randomly generating origin, destination, and departure time for each traveller. By assigning unique origins and destinations to each traveller, their varying trip lengths can be calculated using the Dijkstra shortest path algorithm (Dijkstra, 1959). The travel time (T_i^{car}) of the car user i, with trip length l_i^{car}, departure time t_i, and the speed V_i^{car} calculated by:

$$l_i^{car} = \int_{t_i}^{t_i+T_i^{car}} V_i^{car} dt \quad (1)$$

The buses are assumed to operate on a designated line with a fixed headway. A reasonable constant speed is considered for both the bus and walking modes based on literature (Balzer and Leclercq 2022; Aghabayk et al. 2021).

2.2 Mode choice

Below are the monetary costs associated with using different modes. The decision-making process follows a multinomial logit model, which assumes that users perceive costs independently, with an added error term following a Gumbel distribution.

$$\begin{cases} C_i^{car} &= -\alpha T_i^{car} - \tau \\ C_i^{walk} &= -\alpha T_i^{walk} \\ C_i^{bus} &= -\alpha (T_i^{bus} + T_i^{wait} + T_i^{Transit} + T_i^{access}) \end{cases} \quad (2)$$

Where α is the Value of Time (VoT) and τ represents the toll price. T_i^{wait}, T_i^{Transit} represent the required time that traveller i must wait for the bus and the required time for traveler i to change direction and board another bus, respectively. Both are uniformly distributed between 0 and the bus headway (Daganzo, 2010). Additionally, T_i^{access} equals sum of entrance and egress time.
Assuming that mode alternatives are independent, and the unobserved cost term follows a Gumbel distribution, this study considers the decision-making process as a multinomial logit model. The probability of traveller i choosing the car mode $Pr_i^{car}(t)$ is represented by the following equation:

$$Pr_i^{car}(t) = \frac{e^{\mu C_i^{car}(t)}}{e^{\mu C_i^{car}(t)} + e^{\mu C_i^{bus}} + e^{\mu C_i^{walk}}}$$

In the logit model, the parameter $\mu > 0$ scales the variance of the unobserved portion of utility. As μ approaches zero, choices become more random, while they become more deterministic as μ approaches infinity (Ben-Akiva and Lerman 1985). After calculating probability of choosing each mode, travellers will be assigned to their mode using the Monte-Carlo simulation algorithm.

3 RL Formulation

A deep Q-network (DQN) agent is used to learn a policy that maximizes the expected return from the start time. Figure 2 provides a schematic diagram, illustrating the operation of DQN. The structure of the reinforcement learning (RL) problem is defined as follows:

- **State space:** at each time step the agent considers the demand value of the next time step, car, bus, and pedestrian accumulation, and the toll amount from previous time step as input data.
- **Action space:** the agent is allowed to take three different actions: increasing toll value by $\Delta\tau$, decreasing toll value by $\Delta\tau$, and keep the toll value constant.
- **Transition dynamics:** these are expressed by traffic dynamics described in section 2.1.

- **Reward function:** the scalar reward is defined as the normalized value of network outflow at each step.

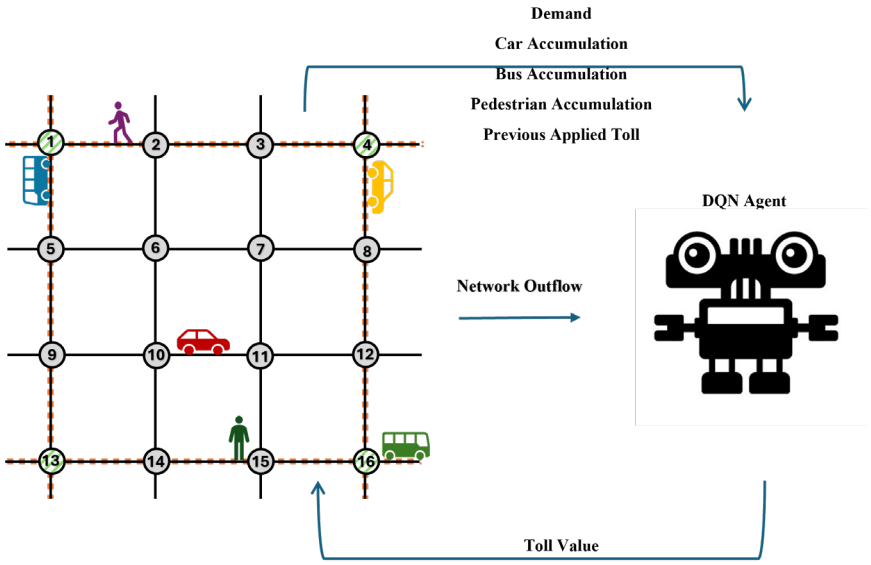

Figure 2. RL structure

4 Experiment

4.1 Experiment setup

A trip-based MFD grid network, with a size of 10 by 10, has been developed to represent a congested city centre at a real scale. A link length of 600 meters is considered for the car/walk network, while 1800 meters is considered for the bus network, reflecting the finer granularity of the car/walk mode compared to the bus mode (Figure 1). Figure 3.a shows the speed-MFD figure of the region.

A 90-minute demand profile with a single peak has been developed to project traffic congestion in city centre (Figure 3.b). Random origin and destination nodes alongside random departure time are assigned to each traveller, representing heterogeneity of the network. An average speed of $3\,^{m}/_{s}$ ($10.8\,^{km}/_{h}$) and $1.18\,^{m}/_{s}$ ($4.3\,^{km}/_{h}$) considered for bus and walk network respectively (Balzer and Leclercq 2022; Aghabayk et al. 2021).

With these settings, the evolution of accumulations under No Control (NC) strategy is presented in Figure 3.c, showing severe congestion in the region. However, this congestion does not escalate into a traffic jam. The VoT is chosen based on the work of Tirachini et al. (2014) and Chen et al. (2021).

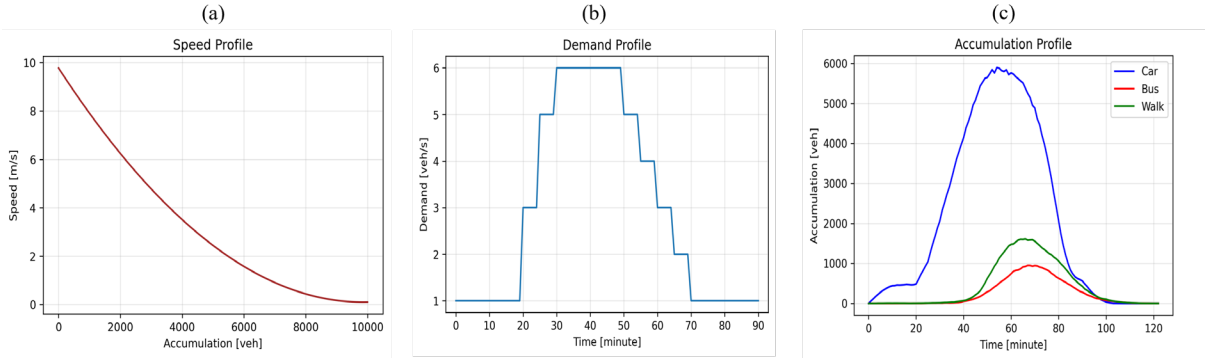

Figure 3: (a) Speed-MFD of the region, (b) 90-minutes demand profile and (c) no-control (NC) scenario accumulation profile.

4.2 Experiment results

The results of training the DQN agent and its application on the developed network are presented in Figure 4. The agent's evolution, based on the total time spent by all travellers, is illustrated in Figure 4.a. It demonstrates a reduction in the total travel time spent by all users, from approximately 2.1×10^7 seconds in the NC scenario to approximately 1.5×10^7 seconds after 250 episodes.

Additionally, the agent identifies the optimal toll profile within the first 150 episodes, and by episode 250, it has converged to the optimal solution. It is worth noting that the slight noise in the total travel time is due to the random selection of origin and destination nodes for each traveller.

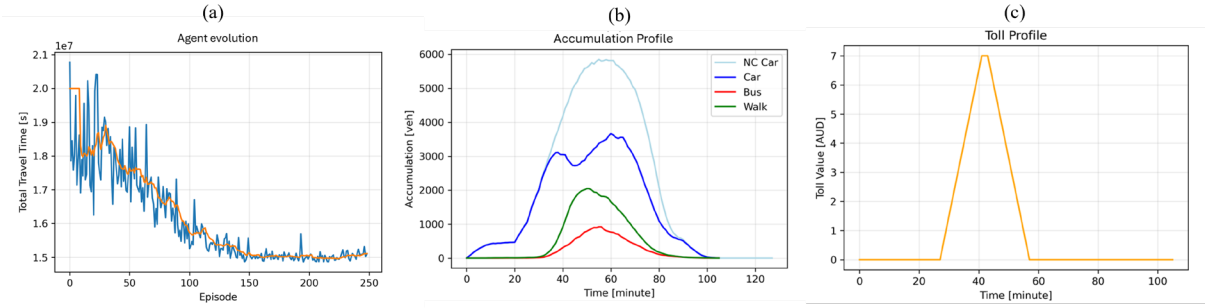

Figure 4: (a) agent evolution, (b) optimized accumulation profile, (c) toll profile.

Figure 4.b illustrates the accumulation of all modes under DQN strategy, while Figure 4.c displays the DQN toll profile. As depicted in Figure 4.b, the agent aims to maintain the car accumulation around the critical level ($n_{cr} \cong 3333$) and maximize network throughput by applying the toll profile on the network. The light blue line shows the car accumulation under NC strategy. It is noteworthy that comparing all previously developed methods such as feedback controller (Zheng et al. 2012) and simulation-based optimisation (Gu and Saberi 2019), the agent identifies the toll profile without any prior knowledge of critical accumulation.

Given the complexity of accurately predicting demand, it is essential to understand how much the proposed DQN agent depends on accurate demand predictions when determining the toll profile for the region. Moreover, although car accumulation can be calculated in real-time using loop detectors or other data sources, this calculation is prone to errors, leading to either over-prediction or under-prediction of car accumulation. To evaluate sensitivity of the trained agent to its input variable, the trained agent is tested under various test scenarios by introducing a 20% margin of error in demand and car accumulation input data.

Figure 5 shows the sensitivity of the agent to the demand profile. In Figure 5.a, the dark blue line represents the proposed demand values provided to the model, while the actual demand values is the light blue noisy line. Under the DQN toll profile (Figure 5.c), the accumulation of different modes is depicted in Figure 5.b.

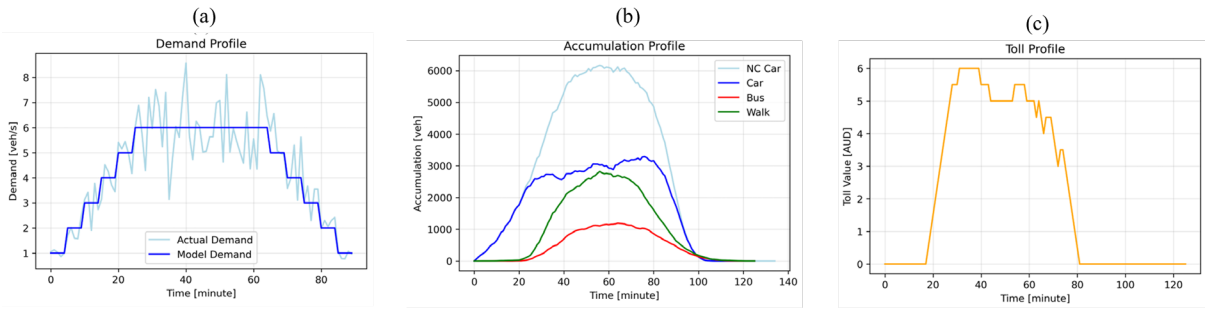

Figure 5: Sensitivity to demand: (a) noisy demand profile, (b) accumulation profile, (c) toll profile.

While the agent is capable of handling entirely unseen demand profiles compared to the training demand profile, introducing 20% error in the agent's demand input does not prevent it from maintaining car accumulation below the critical level.

Moreover, Figure 6 shows the sensitivity of the agent to the car accumulation. For this analysis a different demand profile is also provided to make a more severe situation for the agent. The provided scenario considers that the agent is under-predicting the car accumulation by 20% throughout of the simulation. Considering this scenario the agent still able to reduce traffic congestion significantly. Figure 6.a shows the different demand profile, and Figure 6.b shows the users accumulation. While dark blue line shows the actual car accumulation, the light blue line shows the car accumulation that the agent sees. Figure 6.c shows the proposed toll profile by DQN agent. Even under this severe scenario, the agent maintains the car accumulation around the critical accumulation and prevent traffic congestion.

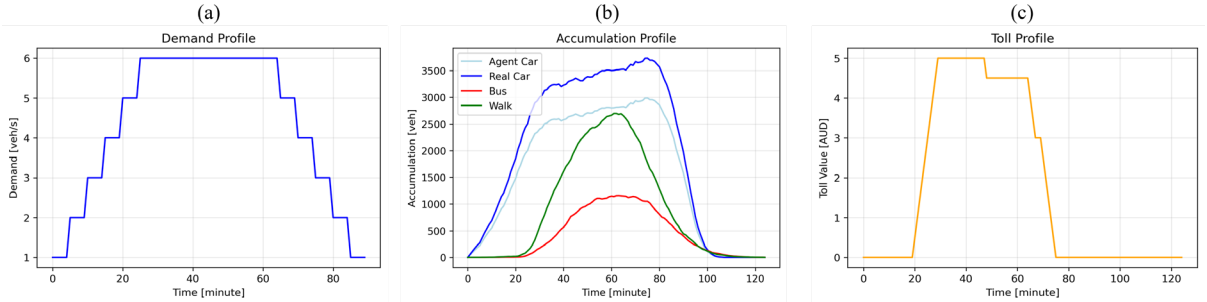

Figure 6: Sensitivity to accumulation and demand (a) different demand profile, (b) accumulation profile, (c) toll profile.

5 Conclusion and Discussion

This study proposes a cordon pricing strategy based on reinforcement learning (RL), utilizing a trip-based Macroscopic Fundamental Diagram (MFD) model to simulate individual mode choice decisions. The objective is to improve network efficiency by maximizing throughput (minimizing total travel time) and considering travellers' heterogeneity. The multi-modal trip-based simulation incorporates three main modes: car, public transport (bus), and walking, with each mode having a distinct network layout to enhance the realism of traffic dynamics. By employing a Deep Q-Network (DQN) agent, a data-driven optimization method has been developed, addressing travellers' heterogeneity. The proposed toll profile not only significantly alleviates traffic congestion but also aims to maintain car accumulation around the critical level and perform proactively.

To perform sensitivity on demand value, 20% noise is introduced to actual demand values while the agent unable to see the noise. Moreover, for car accumulation sensitivity analysis, it is assumed the agent is under-predicting car accumulation throughout of the simulation. The results show that the agent is still able to effectively alleviate congestion and maintain critical accumulation levels.

To further this research, considering other travellers' decision options, such as departure time choice and route choice, can be a valuable direction for future study. Additionally, previous studies indicate that distance-based and time-based pricing methods could be more effective than cordon-based pricing. Moreover, to reduce computation costs, this study employed trip-based simulation as the RL environment. However, applying the proposed method to more realistic simulations, such as agent-based or microscopic simulation, can be a promising avenue for future research.

6 References

Aghabayk, K., Parishad, N., & Shiwakoti, N. (2021). Investigation on the impact of walkways slope and pedestrians physical characteristics on pedestrians normal walking and jogging speeds. Safety Science, 133, 105012.

Amirgholy, M., & Gao, H. O. (2017). Modeling the dynamics of congestion in large urban networks using the macroscopic fundamental diagram: User equilibrium, system optimum, and pricing strategies. Transportation Research Part B: Methodological, 104, 215-237.

Arnott, R. (2013). A bathtub model of downtown traffic congestion. Journal of Urban Economics, 76, 110-121.

Balzer, L., & Leclercq, L. (2022). Modal equilibrium of a tradable credit scheme with a trip-based MFD and logit-based decision-making. Transportation Research Part C: Emerging Technologies, 139, 103642.

Basso, L. J., & Jara-Díaz, S. R. (2012). Integrating congestion pricing, transit subsidies and mode choice. Transportation Research Part A: Policy and Practice, 46(6), 890-900.

Ben-Akiva, Moshe E., and Steven R. Lerman. 1985. Discrete Choice Analysis: Theory and Application to Travel Demand. Vol. 9. Cambridge, MA: MIT Press.

Chen, Y., Zheng, N., & Vu, H. L. (2021). A novel urban congestion pricing scheme considering travel cost perception and level of service. Transportation Research Part C: Emerging Technologies, 125, 103042.

Daganzo, C. F., & Geroliminis, N. (2008). An analytical approximation for the macroscopic fundamental diagram of urban traffic. Transportation Research Part B: Methodological, 42(9), 771-781.

Daganzo, C. F. (2010). Structure of competitive transit networks. Transportation Research Part B: Methodological, 44(4), 434-446.

Daganzo, C. F., & Lehe, L. J. (2015). Distance-dependent congestion pricing for downtown zones. Transportation Research Part B: Methodological, 75, 89-99.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. Numerische Mathematik, 1(1), 269-271.

Fosgerau, Mogens. 2015. “Congestion in the Bathtub.” Economics of Transportation 4 (4): 241–255.

Genser, A., Kouvelas, A., 2022. Dynamic optimal congestion pricing in multi-region urban networks by application of a multi-layer-neural network. Transp. Res. C 134, 103485.

Geroliminis, N., & Daganzo, C. F. (2008). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. Transportation Research Part B: Methodological, 42(9), 759-770.

Geroliminis, N., Haddad, J., & Ramezani, M. (2012). Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: A model predictive approach. IEEE Transactions on Intelligent Transportation Systems, 14(1), 348-359.

Godfrey, J., 1969. The mechanism of a road network. Traffic Eng Control 8 (8).

Gu, Z., & Saberi, M. (2019). A simulation-based optimization framework for urban congestion pricing considering travelers' departure time rescheduling. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC) (pp. 2557-2562). IEEE.

Hou, Z., Xiong, S., 2019. On model-free adaptive control and its stability analysis. IEEE Trans. Automat. Contr. 64, 4555–4569. https://doi.org/10.1109/ TAC.2019.2894586.

Koh, S., Zhou, B., Fang, H., Yang, P., Yang, Z., Yang, Q., ... & Ji, Z. (2020). Real-time deep reinforcement learning based vehicle navigation. Applied Soft Computing, 96, 106694.

Kouvelas, A., Saeedmanesh, M., & Geroliminis, N. (2017). Enhancing model-based feedback perimeter control with data-driven online adaptive optimization. Transportation Research Part B: Methodological, 96, 26-45.

Lamotte, R., & Geroliminis, N. (2016). The morning commute in urban areas: Insights from theory and simulation (No. 16-2003).

Leclercq, L., Ladino, A., & Becarie, C. (2021). Enforcing optimal routing through dynamic avoidance maps. Transportation Research Part B: Methodological, 149, 118-137.

Li, Y., Guo, Y., Lu, J., & Peeta, S. (2019). Impacts of congestion pricing and reward strategies on automobile travelers' morning commute mode shift decisions. Transportation research part A: policy and practice, 125, 72-88.

Li, Y., & Ramezani, M. (2022). Quasi revenue-neutral congestion pricing in cities: Crediting drivers to avoid city centers. Transportation Research Part C: Emerging Technologies, 145, 103932.

Liu, R., Chen, S., Jiang, Y., Seshadri, R., Ben-Akiva, M., & Lima Azevedo, C. (2023). Managing network congestion with a trip-and area-based tradable credit scheme. Transportmetrica B: Transport Dynamics, 11(1), 434-462.

Miralinaghi, M., Peeta, S., He, X., & Ukkusuri, S. V. (2019). Managing morning commute congestion with a tradable credit scheme under commuter heterogeneity and market loss aversion behavior. Transportmetrica B: Transport Dynamics, 7(1), 1780-1808.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D., 2015. Human-level control through deep reinforcement learning. Nature 518, 529–533.

Ren, Y., Hou, Z., Sirmatel, I. I., & Geroliminis, N. (2020). Data driven model free adaptive iterative learning perimeter control for large-scale urban road networks. Transportation Research Part C: Emerging Technologies, 115, 102618.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural networks, 61, 85-117.

Small, K.A., 1997. Economics and urban transportation policy in the United States. Regional Science and Urban Economics 27, 671–691.

Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., Riedmiller, M., 2014. Deterministic Policy Gradient Algorithms. In: 31st International Conference on Machine Learning. Beijing, China.

Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning: An Introduction. MIT Press.

Szepesvári, C. (2022). Algorithms for reinforcement learning. Springer nature.

Tirachini, A., Hensher, D. A., & Rose, J. M. (2014). Multimodal pricing and optimal design of urban public transport: The interplay between traffic congestion and bus crowding. Transportation research part b: methodological, 61, 33-54.

Wang, Y., & Paccagnan, D. (2022, December). Data-driven robust congestion pricing. In 2022 IEEE 61st Conference on Decision and Control (CDC) (pp. 4437-4443). IEEE.

Yildirimoglu, M., Ramezani, M., & Geroliminis, N. (2015). Equilibrium analysis and route guidance in large-scale networks with MFD dynamics. Transportation Research Procedia, 9, 185-204.

Yildirimoglu, M., Sirmatel, I. I., & Geroliminis, N. (2018). Hierarchical control of heterogeneous large-scale urban road networks via path assignment and regional route guidance. Transportation Research Part B: Methodological, 118, 106-123.

Zheng, N., Rérat, G., & Geroliminis, N. (2016). Time-dependent area-based pricing for multimodal systems with heterogeneous users in an agent-based environment. Transportation Research Part C: Emerging Technologies, 62, 133-148.

Zheng, N., Waraich, R. A., Axhausen, K. W., & Geroliminis, N. (2012). A dynamic cordon pricing scheme combining the macroscopic fundamental diagram and an agent-based traffic model. Transportation Research Part A: Policy and Practice, 46(8), 1291-1303.

Zhou, D., & Gayah, V. V. (2021). Model-free perimeter metering control for two-region urban networks using deep reinforcement learning. Transportation Research Part C: Emerging Technologies, 124, 102949.