Impacts of Urban Consolidation Centres for Sustainable City Logistics Using Adaptive Dynamic Programming Based Multi-Agent Simulation

N Firdausiyah1*, E Taniguchi2, A G Qureshi3

1Urban and Regional Planning Department, Brawijaya University, Malang
2Resilience Research Unit, Kyoto University, Japan
3Department of Urban Management, Kyoto University, Japan

*email: firdausiyah@ub.ac.id

Abstract. This paper aims at evaluating the impacts of Urban Consolidation Centers (UCC) for sustainable city logistics using Adaptive Dynamic Programming (ADP) based multi-agent simulation (MAS-ADP). Economic efficiency and environment friendliness criteria were used to evaluate the sustainability of UCC. The results proved that the implementation of UCC as a sustainable city logistics scheme is efficient in reducing 8% of the total delivery cost for freight carrier and reducing 36% of the total emissions released to the environment. It is also showed that the use of learning agents is essential to demonstrate the successful implementation of the UCC, as it is only in the learning-based simulation, UCC operator could get a profit. Our simulation analysis also confirmed that compared to widely used reinforced learning algorithms (Q-learning), MAS-ADP brought increased accuracy to the evaluations’ outcomes of UCC.

Keywords: sustainability, city logistics, urban consolidation centre, adaptive dynamic programming, multi-agent system

1. Introduction
Sustainable city logistics has become an important issue in urban and transportation planning due to high population density in urban areas as well as due to the social, economic, and environmental problems associated with it. City logistics is defined as the process of fully optimizing the logistics and transport activities with the support of advanced information systems in urban areas considering the traffic environment, the traffic congestion, the traffic safety, and the energy savings within the framework of a market economy [1]. The harmonization of economic efficiency and environmental friendliness in city logistics is essential for ensuring sustainable development in urban areas [2], which faces two difficult problems. First is the efficiency of goods delivery within the uncertain environment (due to the parking issues, traffic congestion, and other restrictions in the urban area) that directly effects the operational cost as well as the action selection in presence of optional solutions or policies. The second issue is the involvement of multiple agents in city logistics system, such as freight carriers, shippers, customers, and administrator. All these key stakeholders in urban freight transport have their own specific objectives and tend to behave in a different manner to any urban freight policy [3]. These stakeholders also interact and influence each other in the city logistics environment, which makes the
environment unpredictable. Therefore, the main challenge for the city logistics is to provide a sustainable urban freight transportation while considering multi-agent problems within the uncertain environment.

In order to achieve these aims, numerous city logistics initiatives have been proposed and implemented in several cities, including the Urban Consolidation Centers (UCC) [4]. It is important to evaluate the city logistics policies before they can be effectively deployed due to their manifold implications on different city logistics stakeholders [5]. For that purpose, decision support tools (DST) are needed to help public decision makers and practitioners make decisions acceptable to all parties. These DSTs are mainly based on modeling, optimization, simulation, and evaluation procedures.

There have been many attempts to develop multi-agent simulations to analyze decision-making process of various stakeholders in city logistics, but almost all of them rely on Q-learning [6], [7], [8]. However, based on previous research experiences, which will be described in more detail in the next section, it has been found that ADP-based learning performed better in the accuracy of the outcomes when agents need to interact in uncertain environment, such as city logistics. Therefore, in order to have an accurate evaluation of the UCC, an ADP-based multi-agent simulation has also been developed, which can be used as a DST to achieve better outcomes in the decision process of designing and implementing sustainable city logistics policies.

1.1. Evaluation models for evaluating city logistics measures

Multi-agent systems (MAS) based on the reinforced learning (RL) algorithms have been used for evaluating the behavior of stakeholders, who are affected by the implementation of a city logistics policy. In MAS environment, multiple agents come together and interact, cooperate, coordinate, and negotiate with each other to reach their intended objectives. Various other city logistics policies have been evaluated using MAS with Q-learning such as load factor control and road pricing [6], e-Commerce [7], truck ban and motorway toll discounts [8]. These researches used Q-learning to model evolving behavior of the key stakeholders, namely the carriers, shippers, administrator, and residents relating to urban freight transport. The MAS with Q-learning algorithms have also been used to evaluate the dynamic usage of UCC [9]. The results indicated that the main policy measures that contributed to the successful functioning of the UCC are road pricing, operational subsidies, and the application of time windows. Another study evaluated the Joint Delivery System (JDS) with parking restrictions using MAS with Q-learning [10] and the results showed that JDS with UCC and car parking management have the potential for improving environmental issues related with the urban freight. The results indicated that implementing a truck ban in environmentally damaged areas and discounting motorway tolls in the urban motorway network will have a large environmental impact, resulting in an acceptable environment for all stakeholders.

It can be observed that most of the MAS research in city logistics use Q-learning to represent the decision making of the agents. Some comparative studies modeled the freight carriers and UCC operator behavior in JDS with UCC using on policy algorithm named MAS-ADP based RL [11], [12]. Their result showed that MAS-ADP based RL have better capability than MAS-Q-learning to adjust and adapt the uncertain environment, thus improving the accuracy of decision-making. This study, therefore, would like to apply the MAS-ADP based RL for evaluating the impacts of JDS with UCC for sustainable city logistics.

1.2. MAS-ADP for evaluating city logistics measures

ADP is a learning model in RL part that can be used in the simulation field and optimal control field. As an optimal control tool, [13] has described that ADP scheme is suitable for applications to the systems with strong coupling, strong nonlinearity, and high complexity. It has also been concluded that the ADP is capable to deal with uncertainty [14]. ADP has been widely implemented at the confluence of control problem (15), intelligence traffic systems [16], and robotics [17]. However, none of these previous researches has used the ADP in the multi-agent simulation field, particularly in the
area of city logistics, which represents a highly uncertain environment. Therefore, this study will use the ADP based multi-agent simulation (MAS-ADP) to evaluate the UCC for sustainable city logistics.

2. Methods

2.1. Objective of the learning agents
Recognizing agent’s objective is important in modeling its learning behavior. The freight carrier’s objective is minimizing the total cost of delivering goods to customers. These costs include fixed vehicle utilization cost, the travel cost, and the penalty cost, which are calculated by the Vehicle Routing Problem with Soft Time Windows (VRPSTW) [18].

The UCC operator has been considered as a private or public company that consolidates and delivers the goods from the UCC to customers. Therefore, the UCC operator’s objective is to maximize the profit, which is the difference between revenue (obtained by multiplying the UCC fee with the total demand (parcels) that the freight carrier gives to the UCC) and delivery cost (calculated by the VRPSTW).

2.2. Actions and rewards of the learning agents
The freight carrier can choose two possible actions, viz., direct delivery (DD) or Joint Delivery System (JDS). The operational cost and the total additional parking cost are the immediate rewards for a freight carrier associated with DD action. While UCC fee is the consequence of choosing the JDS. The second learning agent, the UCC operator has three options of managing the UCC fee, which are price going up, flat price, and price going down. The UCC operator will receive the immediate profits associated with the selected action.

2.3. Model for the learning agents
This research applied the Multi-agent-, Simulation-Adaptive-, Dynamic-Programming-based Reinforcement Learning (MAS-ADP based RL) to evaluate the impacts of Joint Delivery Systems (JDS) measured with Urban Consolidation Center. For more details on the MAS-ADP based RL formulation, readers are referred to Firdausiyah et al., [11], [12]. As mentioned earlier, Q-learning-based MAS has been extensively used in the evaluation of the city logistics policies (including the UCC) and because one of the objectives of this research is to come up with a more accurate evaluation tool (i.e. the ADP-based MAS), this research compared the MAS-ADP with the Q-learning [19].

In addition, this research used the VRPTW (Vehicle routing problem with soft-time window) model to calculate the transportation cost either for freight carriers or for UCC operator. For more details on the VRPSTW formulation and solution algorithms, readers are referred to Qureshi et al., [18].

2.4. Environmental emissions model
To evaluate the benefit of using UCC for the environment, this research calculated the emissions using qualitative appraisal index for the basic unit of carbon dioxide (CO\(_2\)), oxides of nitrogen (NO\(_x\)) and suspended particulate matter (SPM), produced by the freight trucks. For more details on the formulation, readers are referred to NILIM [20].

3. Results and Discussions
A square topology-based, hypothetical network (Figure 1) was used for the evaluation of the UCC based on the simulations using ADP and Q-learning models within MAS. Four carriers (A, B, C, D), one UCC, and 20 customers were involved in this network. The agents will make decisions as actions by considering the fluctuating parking cost, UCC fees, and demand received from their customer every day.
This research used some assumptions as listed in Table 1;

| Item                              | Value                  |
|-----------------------------------|------------------------|
| Working time                      | 8 AM to 8 PM           |
| Time window                       | 60 minutes per customer|
| Capacity of the truck             | 200 parcels/truck      |
| Waiting charge (\(w_c\)) for early arrival | 1 Yen/minute          |
| Penalty charge (\(g_p\)) for late arrival | 5 Yen/minute          |

All simulations are done using MATLAB. The learning rate and the discount factor for ADP used were 0.2 and 0.6, respectively; whereas, the learning rate and the discount factor for Q-learning have been set as 0.2 and 0.8, respectively. These values were based on the results of a sensitivity analysis that has been done prior to the case study. In this study, this research performed two separate simulations using ADP and Q-learning to evaluate the impacts of UCC for sustainable city logistics using the following criteria; 1) economics’ efficiency for each learning agent, i.e., cost saving for the freight carrier, and profitability for the UCC operator; 2) environmentally friendliness.

3.1. Accuracy of the learning models

Accuracy of the outcomes obtained in the MAS-ADP and MAS-Q-learning is important for the learning models to evaluate a sustainable city logistics scheme (such as UCC in this study). The “accuracy” refers to the closeness of the gap between the expected value obtained in the MAS-ADP and MAS-Q-learning based simulations to the corresponding value experienced by the learning agent. Smaller gap between these two costs means more accurate method.

Figure 2(a) shows that the percentage gap between expected cost and the experienced cost in the ADP-based simulation is lower (39.6%) than the Q-learning based simulation (47.7%) for freight carrier A; similar pattern were obtained for freight carriers B, C, and D.

![Figure 2. Accuracy gaps (%) in ADP and Q-learning based simulations](image)
Similarly, in the case of the UCC operator, the percentage gap between the expected profit and the experienced profits in the ADP-based simulation was also lower (46.4%) than the Q-learning-based simulation (51.9%) as shown in Figure 2(b). It proved that the ADP-based learning can improve the accuracy of the simulation, thereby improving the quality of simulation.

3.2. Sustainability criteria: economics’ efficiency

3.2.1. Freight carriers
The meaning of efficiency for freight carrier is the delivery of goods at lower cost to the customers. To calculate the efficiency for freight carrier, this research compared the difference in the experienced cost for the freight carrier with UCC and without UCC. The existence of UCC provided more alternatives of goods’ delivery for freight carrier, which are direct delivery, and JDS with UCC. The MAS-ADP and MAS-Q-learning as the learning models will suggest the actions based on the reward received from the environment.

Both ADP and Q-learning-based simulation resulted in a lower delivery costs experienced by freight carriers in the case of UCC than the one without UCC case (Figure 3(a)). The delivery cost with UCC resulted from the fact that ADP-based simulation was 8.4% lower on average compared to the experienced delivery cost without UCC. Corresponding figures for the Q-learning-based simulation was 6.7%. Figure 3(b) illustrates more details with the cumulative experienced delivery cost for a freight along the simulation. It means that implementing the UCC as a sustainable city logistics policy is efficient to minimize the delivery costs for a freight carrier. Moreover, using ADP as the freight carriers’ behavior learning model is better than using the Q-learning, as the former choice can further save almost 1.7%, on average, of the total delivery costs; thereby improving the quality of the city logistics simulation i.e. ADP-based simulation provided more favorable comparison with the no-policy (without UCC) case.

![Figure 3. Experienced delivery costs for freight carriers (FC)](image)

3.2.2. UCC operator
The economic efficiency for UCC operator meant higher profits at lower UCC fee offered to freight carriers to foster more demand. To calculate the profitability, we compared the difference in the experienced profits for UCC operator under the learning environment with freight carrier and without learning. In the learning environment, we assumed the UCC operator will update the UCC fee every day by learning the reward received (profits) based on the business provided by the freight carriers. Without learning (α=0), the UCC operator will not consider the current information from the environment. Therefore, the UCC operator is assumed to offer fixed UCC fee (150 JPY/ parcel) to the freight carrier every day.
Figure 4 clearly shows that the UCC would fail dramatically if a fixed UCC fee policy is followed without learning from the environment, as the cumulative experienced profit received by the UCC operator without learning is way lower than the cumulative experienced profits resulted from learning using either of ADP or Q-learning. A dip in the cumulative experienced profit curve shows a negative profit obtained in that episode. The negative profit means the UCC failed to cover the downstream delivery cost based on the business (demand) received from the freight carriers. If the UCC operator is modelled as a learning agent, it learns from these negative reward values to adjust the UCC fee (may be able to attract more demand) and becomes profitable again. It showed the importance of the MAS-based simulations in the evaluation of the city logistics policies where a stakeholder’s action/behavior can seriously impact the other. It is important for the UCC operator to learn from the behavior of the freight carriers (refusing to join UCC due to high fees) in order to become profitable.

Moreover, Figure 5 shows that using ADP as the UCC operator’s behavior learning model is better than the Q-learning. The two simulations suggest different actions of managing the UCC fee level by the UCC operator, which resulted in the difference of the experienced profits. The UCC fee suggested by the ADP is always 7% lower, on average, than the UCC fee suggested by Q-learning (Figure 5(a)). It increases the possibility of choosing JDS with UCC by freight carrier than the direct delivery under the ADP-based learning. The impact is also evident in the profits received by UCC operator, which were 3.7% when using the UCC fee suggested by ADP (130 JPY/ parcel on the average) as compared to 2.1% when using the Q-learning (140 JPY/ parcel on the average) (Figure 5(b)).
3.3. Sustainability criteria: environmentally friendliness

We evaluated the impacts of UCC by calculating the total emissions (CO$_2$, NO$_x$, and SPM) from the delivery activities made by freight carriers and the UCC operator with and without UCC using ADP and Q-learning. The existence of UCC will reduce 36% (ADP) and 31% (Q-learning) of the total emissions as compared to the condition without UCC (Figure 6). As explained earlier, the differences of the result between these two simulations arose from the different action selection suggested by the learning model. The experienced emission level of CO$_2$ (Figure 6(a)), NO$_x$ (Figure 6(b)), and SMP (Figure 6(c)), obtained under ADP-based simulation is 5% lower than the Q-learning based simulation. It means that using ADP as the learning model for both agents is better than using the Q-learning, as the former choice can reduce 36% of the total emissions released to the environment; thereby improving the quality of the city logistics simulation, i.e. ADP-based simulation provides better comparison with the no-policy (without UCC) case.

![Figure 6. Total emissions for 24 episodes (120 days) of simulation](image)

4. Conclusions

This paper developed the ADP models to evaluate the UCC as a sustainable city logistics policy. Economic efficiency and environmentally friendliness criteria were used to evaluate the sustainability of UCC. The results showed that the implementation of UCC as a sustainable city logistics scheme is efficient by reducing 8% of the total delivery cost for the freight carrier and reducing 36% of the total emissions released to the environment. It was also showed that the use of learning agents is essential to demonstrate the successful implementation of the UCC, as it is capable to increase the experienced profit gained.

In addition, simulations should accommodate the agent’s objective. It was observed that simulation using ADP resulted in a further 1.7% less experienced cost as compared to the simulation done using Q-learning. In case of UCC operator, the ADP satisfied its objective by getting higher profits than the Q-learning based simulation. The differences of the results between these two simulations arise from the different action selection suggested by the learning model. Therefore, the accuracy of the outcomes is also very important for the learning models, especially within the uncertain environment. It was found that ADP-based simulation improved the accuracy of the expected delivery costs for the freight carrier compared to the Q-learning based simulation. The accuracy of the expected profits received by the UCC operator in the ADP-based simulation was also much better.

References

[1] Taniguchi, E., and Thompson, R. G. City Logistics I. Institute of Systems Science Research. Kyoto., Japan, 1999.

[2] Taniguchi, E., Thompson, R. G., Yamada, T. Logistics systems for sustainable cities: Visions for city logistics. Elsevier Ltd., UK, 2004.

[3] Taniguchi, E. & Tamagawa, D. Evaluating city logistics measures considering the behavior of several stakeholders. Journal of the Eastern Asia Society for Transportation Studies, Volume 6, 2005, pp. 3062-3076
[4] Browne, M., Allen, J., & Leonardi, J. Evaluating the use of an urban consolidation centre and electric vehicles in central London. IATSS Research 35, 2011, Page 1-6.

[5] Taniguchi, E., and Thompson, Russell G. City Logistics: Mapping the Future. CRC Press Taylor and Francis Group., UK, 2015.

[6] Teo, J. S. E., Taniguchi, E. & Qureshi, A. G Evaluation of Load Factor Control and Urban Freight Road Pricing Joint Schemes with Multi-agent Systems Learning Models. Procedia Social and Behavioral Sciences, vol. 125, 2014, pp. 62 – 74.

[7] Teo, J. S. E., Taniguchi, E., and Qureshi, A. G. Evaluation of Distance-Based and Cordon-Based Urban Freight Road Pricing in E-Commerce Environment with Multiagent Model. In Transportation Research Record: Journal of the Transportation Research Board, No. 2269, Transportation Research Board of the National Academies, Washington, D.C., 2012, pp. 127–134.

[8] Tamagawa, D., Taniguchi, E., and Yamada, T. Evaluating City Logistics Measures Using a Multi-agent Model. Procedia Social and Behavioral Sciences, Vol. 2, 2010, pp. 6002–6012.

[9] van Duin, J. H. R., van Kolck, A., Anand, N. and Tavasszy, L. A. Towards an agent-based modeling approach for the evaluation of dynamic usage of urban distribution centers. Procedia Social and Behavioral Sciences, Volume 39, 2012, pp. 333-348.

[10] Wangapisit, O., Taniguchi, E., Teo, J. S. E. & Qureshi, A. G. Multi-Agent systems modelling for Evaluating Joint Delivery Systems. Procedia-Social and Behavior Sciences, Volume 125C, 2014, pp. 472-483.

[11] Firdausiyah, N., Taniguchi, E., and Qureshi, A. G., (2017). Multi-agent Simulation using Adaptive Dynamic Programming in The Existence of Urban Consolidation Centers. Journal of Japan Society of Civil Engineers, Ser. D3 (Infrastructure Planning and Management), 73(5), I_835-I_846.

[12] Firdausiyah, N., Taniguchi, E., Qureshi, A. G., (2019). Modeling City Logistics using Adaptive Dynamic Programming Based Multi-Agent Simulation. Transportation Research Part E: Logistics and Transportation Review. 125, 2019, pp. 74-96.

[13] Zhang, Hua-Guang, Xin Zhang, Yan-Hong Luo, and Jun Yang. An overview of research on adaptive dynamic programming. Acta Automatica Sinica, volume 39(4), 2013, pp. 303–311.

[14] Godfrey, G. & Powell, W. B. An adaptive, dynamic programming algorithm for stochastic resource allocation problems I: Single period travel times. Transportation Science, volume 36(1), 2002, pp. 21–39.

[15] Venayagamoorthy G K, Harley R G, Wunsch D C. Dual heuristic programming excitation neuro control for generators in a multimachine power system. IEEE Transactions on Industry Applications, volume 39(2), 2003, pp. 382–394.

[16] Fagan, D., and Meier, R. Dynamic Multi-Agent Reinforcement Learning for Control Optimization. Fifth International Conference on Intelligent System, Modelling and Simulation, 2014.

[17] Karthikeyan, R., SheelaRani, B., and Renganathan, K. An Instant Path Planning Algorithm for Indoor Mobile A robot Using Adaptive Dynamic Programming and Reinforcement Learning. International Journal of Engineering and Technology, volume 6, no.2, 2014, pp. 1224-1231.

[18] Qureshi, A.G., Taniguchi, E., and Yamada, T. Hybrid insertion heuristics for vehicle routing problem with soft time windows. Journal of the Eastern Asia Society for Transportation Studies, Volume 8, 2010, pp. 827–841.

[19] Watkins, C.J.C.H and Dayyan, P. Q.Learning. Machine Learning, 8, 1992, pp. 279-292.

[20] NILIM. Qualitative appraisal index calculations used for basic unit for computation of CO2, NOx and SPM (in Japanese), 2003.