

# **Relative Performance-Based Reinforcement Learning for Large-Scale Dynamic and Stochastic Electric Vehicle Routing**

**Emine Gülşen Torunoğlu Aslan**

Graduate School of Natural and Applied Sciences  
Gazi University, Ankara, Türkiye  
[egulsen.aslan@gazi.edu.tr](mailto:egulsen.aslan@gazi.edu.tr)

**Gonca Yıldırım and Selçuk Kürşat İşleyen**

Department of Industrial Engineering  
Department, Faculty of Engineering  
Gazi University, Ankara, Türkiye  
[goncayildirim@gazi.edu.tr](mailto:goncayildirim@gazi.edu.tr), [isleyens@gazi.edu.tr](mailto:isleyens@gazi.edu.tr)

## **Abstract**

The Dynamic and Stochastic Electric Vehicle Routing Problem (DS-EVRP) involves routing decisions under both dynamic and stochastic uncertainty. However, most DS-EVRP studies rely on deterministic assumptions or treat these uncertainties separately due to the computational limitations of traditional optimization methods, limiting their applicability in real-world electric vehicle (EV) operations, where uncertainty is inherently more pronounced than in conventional vehicle routing. While reinforcement learning (RL) has emerged as a promising alternative, its application to DS-EVRP remains limited, particularly in settings that jointly consider dynamic and stochastic elements. Moreover, the impact of reward design on learning performance remains underexplored. In our prior work, we compared two RL approaches based on tabular Monte Carlo control with different reward designs against a classical optimization method. The results demonstrated that both RL approaches outperformed the classical method, and that reward design plays a critical role in this performance: the proposed relative performance-based reward consistently outperformed the conventional reward scheme, yielding improved learning effectiveness and solution quality in dynamic and stochastic decision environments. Building on these findings, the present study extends the proposed reward mechanism to larger-scale DS-EVRP instances within a sequential decision-making framework, where tabular RL approaches face significant limitations due to rapid growth of the state-action space. To address this, we transition from tabular RL to function approximation via deep RL and integrate the proposed reward design within actor-critic architectures. Initial experiments indicate that this combination provides a promising direction for achieving robust, adaptive, and scalable routing decisions, directly relevant to real-world EV operations.

## **Keywords**

Dynamic and Stochastic Electric Vehicle Routing Problem (DS-EVRP), Reinforcement Learning, Relative Performance-Based Reward

## **Biographies**

**Emine Gülşen Torunoğlu Aslan** is a PhD candidate within the Graduate School of Natural and Applied Sciences at Gazi University, Türkiye. Her research focuses on reinforcement learning approaches to dynamic and stochastic vehicle routing problems within the field of supply chain and logistics management.

**Gonca Yıldırım** is an Assistant Professor of Industrial Engineering at Gazi University in Türkiye. She holds a PhD from the University of Florida. Her research interests include optimization applications, inventory planning, and control. She has published in the International Journal of Production Economics and the Journal of the Operational Research Society, among others.

**Selçuk Kürşat İşleyen** is a full professor in the Department of Industrial Engineering of Gazi University, Türkiye. He graduated from the Faculty of Engineering at Gazi University in 2001. He received M.Sc. and Ph.D. degrees in Industrial Engineering from Gazi University in 2004 and 2008, respectively. His research mainly focuses on optimization problems, logistics, and assembly line balancing problems.