

**การพัฒนาแบบจำลองแบบเรียลไทม์เพื่อเพิ่มประสิทธิภาพการควบคุม  
เรืออัตโนมัติด้วยอัลกอริทึม Twin Delayed Deep Deterministic Policy  
Gradient ในระบบการเรียนรู้เสริมกำลังเชิงลึก**

**REAL-TIME MODEL DEVELOPMENT AND PERFORMANCE  
ENHANCEMENT FOR AUTONOMOUS SHIP CONTROL USING THE  
DEEP REINFORCEMENT LEARNING TWIN DELAYED DEEP  
DETERMINISTIC POLICY GRADIENT ALGORITHM**

ธนาวุฒิ วิริยะพันธ์<sup>1</sup> สาลินี สมบูรณ์ไพศาล<sup>2</sup> ก้องเกียรติ ลีกนุช<sup>3</sup> และ ทวีศักดิ์ มั่นชวนนท์<sup>4</sup>

<sup>1,3</sup>อาจารย์, กลุ่มวิชาวิศวกรรมเครื่องกลเรือ ศูนย์ฝึกพาณิชย์นาวี กรมเจ้าท่า  
120 หมู่ 7 ซอยเทศบาล 6 ถนนสุขุมวิท อำเภอเมืองสมุทรปราการ สมุทรปราการ 10270,

<sup>1</sup>t\_wiriya@mmtc.ac.th, <sup>3</sup>Lucknuch1959@gmail.com

<sup>2</sup>อาจารย์, กลุ่มวิชาการทั่วไป ศูนย์ฝึกพาณิชย์นาวี กรมเจ้าท่า  
120 หมู่ 7 ซอยเทศบาล 6 ถนนสุขุมวิท อำเภอเมืองสมุทรปราการ สมุทรปราการ 10270,

s\_somboo@mmtc.ac.th

<sup>4</sup>อาจารย์, ภาควิชาวิศวกรรมเครื่องกล-ซ่อมบำรุงอากาศยาน คณะวิศวกรรมศาสตร์  
มหาวิทยาลัยเกษมบัณฑิต 60 ถนนร่มเกล้า แขวงมีนบุรี เขตมีนบุรี กรุงเทพมหานคร 10510,

<sup>4</sup>tmunchawanont@gmail.com

Thanawut Wiriyaphan<sup>1</sup>, Salinee Somboompaisan<sup>2</sup>, Kongkiat Luknuch<sup>3</sup>  
and Thaweesak Munchawanont<sup>4</sup>

<sup>1,3</sup>Lecturer, Marine Engineering Division, Merchant Marine Training Centre, Marine  
Department of Thailand, 120 Moo 7, Soi Thetsaban 6, Sukhumvit Road, Bang Duan  
Subdistrict, Mueang Samut Prakan District, Samut Prakan Province 10270, Thailand,

<sup>1</sup>t\_wiriya@mmtc.ac.th, <sup>3</sup>Lucknuch1959@gmail.com

<sup>2</sup>Lecturer, General Education Division, Merchant Marine Training Centre, Marine  
Department of Thailand, 120 Moo 7, Soi Thetsaban 6, Sukhumvit Road, Bang Duan  
Subdistrict, Mueang Samut Prakan District, Samut Prakan Province 10270, Thailand,

<sup>2</sup>s\_somboo@mmtc.ac.th

<sup>4</sup>Lecturer, Department of Mechanical Engineering-Aircraft Maintenance, Faculty of Engineering, Kasem Bundit University, 60 Romklao Road, Minburi District, Bangkok 10510, Thailand, tmunchawanont@gmail.com

### บทคัดย่อ

ตามข้อกำหนดขององค์การทางทะเลระหว่างประเทศ (IMO) ว่าด้วยการลดการปล่อยก๊าซเรือนกระจกและการมุ่งสู่ระบบการเดินเรืออัตโนมัติ การพัฒนาระบบควบคุมเรือที่บูรณาการประสิทธิภาพด้านการจัดการพลังงานและความปลอดภัยเข้าด้วยกันจึงเป็นสิ่งจำเป็น งานวิจัยนี้นำเสนอกรอบการทำงานแบบเรียลไทม์บนแพลตฟอร์ม MATLAB/Simulink สำหรับเรือ Ro-Pax M/S Nils Holgersson โดยพัฒนาแบบจำลองทางคณิตศาสตร์ที่ครอบคลุมพลศาสตร์การเคลื่อนที่ระบบขับเคลื่อนแบบหลายเครื่องยนต์ และแรงต้านทานจากสภาพแวดล้อมโดยอาศัยข้อมูลจากเซ็นเซอร์จำลอง นอกจากนี้ยังได้เชื่อมต่อส่วนต่อประสานโปรแกรมประยุกต์ (API) เข้ากับฐานข้อมูลสภาพอากาศทางทะเลเพื่อดึงข้อมูลแบบเรียลไทม์มาใช้ในการฝึกฝนแบบจำลอง งานวิจัยนี้ประยุกต์ใช้ระบบการเรียนรู้เสริมกำลังเชิงลึกร่วมกับอัลกอริทึม Twin Delayed Deep Deterministic Policy Gradient (TD3) เพื่อควบคุมการจัดการภาระกำลังของเครื่องจักรและการบังคับมุมหางเสือของเรือ โดยมีเป้าหมายหลักในการลดอัตราการสิ้นเปลืองเชื้อเพลิงและเพิ่มความสำเร็จในการหลีกเลี่ยงการชน ผลการจำลองพบว่า กรอบการทำงานที่นำเสนอสามารถลดการสิ้นเปลืองเชื้อเพลิงรวมได้ 1.592% (เทียบเท่า 73.56 กิโลกรัมต่อเที่ยวเดินเรือ) ในเส้นทางจาก Yachthafen Hohe Düne เมืองรอสต็อก ไปยังท่าเรือ Marina Baltica 1 เมืองลิือบะ ประเทศเยอรมนี รวมระยะทาง 43 ไมล์ทะเล ซึ่งช่วยลดการปล่อยก๊าซ CO<sub>2</sub> ได้ถึง 226.71 กิโลกรัมต่อเที่ยว นอกจากนี้ ระบบดังกล่าวยังสามารถเพิ่มอัตราความสำเร็จในการหลีกเลี่ยงการชนได้ 1.75% เมื่อเทียบกับระบบควบคุมแบบดั้งเดิม (เพิ่มขึ้นจาก 93.71% เป็น 95.46%) โดยประเมินจากการรักษาระยะห่างที่ปลอดภัยและระยะเวลาเผื่อในการหักเลี้ยวหลบสิ่งกีดขวาง ผลลัพธ์ที่ได้แสดงให้เห็นอย่างชัดเจนว่า กรอบการทำงานนี้มีศักยภาพสูงในการนำไปประยุกต์ใช้เชิงพาณิชย์ เนื่องจากสามารถบรรลุเป้าหมายทั้งด้านความปลอดภัยและประสิทธิภาพพลังงานได้พร้อมกัน ซึ่งจะเป็นแนวทางสำคัญในการพัฒนาและต่อยอดระบบควบคุมเรืออัตโนมัติในอนาคต

**คำสำคัญ:** การเดินเรืออัตโนมัติ, การเรียนรู้เสริมกำลังเชิงลึก, อัลกอริทึม TD3, การจัดการพลังงาน, การหลีกเลี่ยงการชน, ประสิทธิภาพการใช้เชื้อเพลิง

## ABSTRACT

In accordance with the International Maritime Organization (IMO) regulations concerning greenhouse gas emission reduction and the transition toward autonomous shipping, developing a ship control system that integrates energy management efficiency with safety is imperative. This research proposes a real-time framework on the MATLAB/Simulink platform for the Ro-Pax vessel M/S Nils Holgersson. A comprehensive mathematical model was developed, encompassing ship dynamics, a multi-engine propulsion system, and environmental resistance utilizing simulated sensor data. Furthermore, an Application Programming Interface (API) was integrated with a marine meteorological database to retrieve real-time weather data for model training. This study applied Deep Reinforcement Learning (DRL) via the Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm to manage the main engines' load distribution and control the rudder angle. The primary objectives were to minimize fuel consumption and enhance collision avoidance success rates. Simulation results demonstrated that the proposed framework reduced total fuel consumption by 1.592% (equivalent to 73.56 kg per voyage) along a 43-nautical-mile route from Yachthafen Hohe Düne, Rostock, to Marina Baltica 1, Lübeck, Germany. This efficiency improvement consequently mitigated CO<sub>2</sub> emissions by 226.71 kg per voyage. Additionally, the system increased the collision avoidance success rate by 1.75% compared to conventional control systems (improving from 93.71% to 95.46%). This was evaluated based on the vessel's ability to maintain a safe distance and allocate adequate time margins for evasive maneuvers. These outcomes clearly demonstrate the proposed framework's high potential for commercial implementation, successfully achieving both safety and energy efficiency targets simultaneously. This research provides a significant foundation for the future development and advancement of autonomous marine vehicle control systems.

**KEYWORDS:** Autonomous Shipping, Deep Reinforcement Learning (DRL), TD3 Algorithm, Energy Management, Collision Avoidance, Fuel Efficiency

## 1. บทนำ

ภาคการขนส่งทางทะเล ซึ่งปล่อยก๊าซเรือนกระจก (GHG) ในระดับโลกประมาณ 3% กำลังเผชิญกับช่วงเปลี่ยนผ่านที่สำคัญ [1] เพื่อตอบสนองต่อความกังวลเกี่ยวกับผลกระทบต่อสิ่งแวดล้อม องค์การทางทะเลระหว่างประเทศ (IMO) ได้ปรับปรุงเป้าหมายเชิงกลยุทธ์โดยมุ่งลด

การปล่อยก๊าซเรือนกระจกจากการขนส่งระหว่างประเทศให้เป็นศูนย์ ภายในหรือใกล้เคียงปี พ.ศ. 2593 [2] เป้าหมายดังกล่าวได้กระตุ้นให้เกิดการให้ความสำคัญต่อการเพิ่มประสิทธิภาพพลังงานและการลดการบริโภคเชื้อเพลิง ซึ่งส่งผลกระทบโดยตรงต่อต้นทุนการดำเนินงานและสภาพแวดล้อม [3] ในขณะเดียวกันอุตสาหกรรมการขนส่งทางทะเลกำลังเปลี่ยนผ่านสู่การนำระบบอัตโนมัติมาใช้งาน เพื่อตอบสนองต่อความต้องการด้านความปลอดภัยและประสิทธิภาพในยุคปัจจุบัน โดยเฉพาะเรือเดินทะเลขับเคลื่อนอัตโนมัติ (Maritime Autonomous Surface Ships: MASS) ซึ่งได้รับการนิยามโดย IMO ว่าเป็นเรือพาณิชย์ที่สามารถปฏิบัติการได้โดยมีการควบคุมจากมนุษย์น้อยที่สุด และอาศัยการทำงานของเซ็นเซอร์ ซอฟต์แวร์ และระบบสื่อสาร เพื่อนำทางหลีกเลี่ยงการชน และปฏิบัติหน้าที่ต่าง ๆ ของเรือแทน ทั้งนี้ IMO ได้จำแนก MASS ออกเป็น 4 ระดับ ดังตารางที่ 1 [4]

ตารางที่ 1 กำหนดระดับการใช้งานเรืออัตโนมัติทางทะเล (MASS)

| ระดับ | คำอธิบาย                                                | ลูกเรือบนเรือ |
|-------|---------------------------------------------------------|---------------|
| 1     | เรือที่มีกระบวนการอัตโนมัติและมีระบบสนับสนุนการตัดสินใจ | มี            |
| 2     | เรือที่ควบคุมจากระยะไกลโดยมีลูกเรืออยู่บนเรือ           | มี            |
| 3     | เรือที่ควบคุมจากระยะไกลโดยไม่มีลูกเรืออยู่บนเรือ        | ไม่มี         |
| 4     | เรือขับเคลื่อนอัตโนมัติโดยสมบูรณ์                       | ไม่มี         |

ตลาดเรืออัตโนมัติมีแนวโน้มการเติบโตอย่างต่อเนื่อง [5] โดยได้รับการสนับสนุนจากความก้าวหน้าทางด้านปัญญาประดิษฐ์ (Artificial Intelligence) ซึ่งมีความสามารถในการแก้ปัญหาและพัฒนากลยุทธ์การควบคุมผ่านกระบวนการเรียนรู้เสริมกำลังเชิงลึก (Deep Reinforcement Learning หรือ DRL) เทคโนโลยีนี้มีศักยภาพในการหาค่าที่เหมาะสมที่สุดสำหรับการเดินเรือทั้งในมิติของประสิทธิภาพและความปลอดภัย เช่น การจัดการพลังงานบนเรือ การวางแผนเส้นทางอัจฉริยะเพื่อลดการปล่อยมลพิษ และการหลีกเลี่ยงการชนเพื่อป้องกันความสูญเสียทางเศรษฐกิจและสิ่งแวดล้อม [6] อย่างไรก็ตาม การพัฒนาระบบอัจฉริยะดังกล่าวจำเป็นต้องใช้ทรัพยากรอย่างมากและมีความเสี่ยงสูงหากทำการทดสอบภายใต้สภาวะจริง [7] ดังนั้น การใช้แบบจำลองขั้นสูงจึงเป็นทางเลือกที่จำเป็น ตัวอย่างเช่น การใช้แพลตฟอร์ม MATLAB/Simulink ในการสร้างแบบจำลองแผนภาพบล็อก เพื่อออกแบบระบบหลายโดเมน ก่อนนำไปใช้งานจริงบนฮาร์ดแวร์ซึ่งช่วยประหยัดทั้งเวลาและค่าใช้จ่าย อีกทั้งยังมีความแม่นยำสูง สามารถทดสอบและปรับปรุงระบบควบคุมภายในสภาพแวดล้อมเสมือนจริงได้อย่างปลอดภัย [8] ด้วยเหตุนี้ งานวิจัยฉบับนี้จึงมุ่งเน้นที่จะพัฒนารอบการทำงานแบบเรียลไทม์บนแพลตฟอร์ม MATLAB/Simulink เพื่อสร้าง

แบบจำลองระบบควบคุมอัจฉริยะสำหรับเรือ Ro-Pax โดยประยุกต์ใช้อัลกอริทึม DRL ขั้นสูงอย่าง Twin Delayed Deep Deterministic Policy Gradient ในการควบคุมการกระจายภาระกำลังของ เครื่องจักรหลักและการบังคับหางเสือ โดยมีวัตถุประสงค์หลักเพื่อรักษาสมดุลระหว่างการลดอัตราการสิ้นเปลืองเชื้อเพลิงให้ได้สูงสุด ควบคู่ไปกับการเพิ่มอัตราความสำเร็จในการหลีกเลี่ยงการชน เพื่อเป็นแนวทางในการพัฒนาระบบเดินเรืออัตโนมัติที่ปลอดภัยและเป็นมิตรต่อสิ่งแวดล้อมต่อไป

## 2. ทบทวนวรรณกรรม

การทบทวนวรรณกรรมที่ผ่านมาพบว่า งานวิจัยด้านการหลีกเลี่ยงการชนของเรือได้รับการพัฒนาอย่างต่อเนื่อง แนวทางแรกสุดได้แก่ระบบอิงกฎเกณฑ์ เช่น ตัวควบคุมที่ใช้ตรรกศาสตร์คลุมเครือ ของ Ahmed et al [9] ซึ่งใช้เมตริกจุดเข้าใกล้ที่สุดในการประเมินความเสี่ยง แม้จะมีประสิทธิภาพในสถานการณ์ที่มีเรือหลายลำ แต่ระบบก็ยังแปรปรวนได้ง่ายจากสัญญาณรบกวน อีกทางเลือกหนึ่งคือแนวทางการหาค่าเหมาะสมที่สุดโดยอิงแบบจำลอง เช่น การควบคุมแบบทำนายผลตามแบบจำลอง ดังที่ศึกษาโดย Tengesdal et al [10] ซึ่งผนวกการประเมินความเสี่ยงจากความน่าจะเป็นและตัวกรองคาลมาน เพื่อจัดการกับความไม่แน่นอนของสัญญาณรบกวนและสถานะ อย่างไรก็ตาม วิธีการดังกล่าวต้องอาศัยการคำนวณที่แม่นยำของแบบจำลองต้นแบบ และขึ้นอยู่กับคุณภาพของการประมาณค่าเส้นทางอย่างมาก ซึ่งเป็นข้อจำกัดในการนำไปใช้งานแบบเรียลไทม์ ต่อมาจึงเกิดการปรับใช้แนวทางที่มีความยืดหยุ่นสูงกว่า นั่นคือการเรียนรู้เสริมกำลังเชิงลึก ซึ่งช่วยให้ระบบสามารถเรียนรู้นโยบายการควบคุมผ่านการโต้ตอบกับสภาพแวดล้อมได้ด้วยตนเอง [11] อย่างไรก็ตาม ความท้าทายที่สำคัญของระบบการเรียนรู้เสริมกำลังเชิงลึกคือการทำให้มั่นใจว่าระบบจะเรียนรู้พฤติกรรมที่สอดคล้องกับอนุสัญญาระหว่างประเทศว่าด้วยข้อบังคับเพื่อป้องกันการชนกันทะเล (COLREGs) อย่างสม่ำเสมอ โดยเฉพาะในสถานการณ์ที่ต้องเผชิญหน้ากับเรือหลายลำ เพื่อแก้ไขปัญหา การออกแบบฟังก์ชันรางวัล จึงกลายเป็นกลยุทธ์หลัก งานวิจัยของ Wang et al [12] ได้แสดงให้เห็นถึงการใช้ฟังก์ชันรางวัลแบบลำดับขั้น ร่วมกับอัลกอริทึม Deep Q-Network (DQN) โดยกำหนดบทลงโทษหรือลดคะแนนความสามารถของระบบตามประเภทของการเผชิญหน้ากันของเรือ (เช่น การเผชิญหน้ากันโดยตรง หรือสถานการณ์ตัดหน้า) นอกจากนี้ Xie et al [13] ได้บูรณาการฟังก์ชันรางวัลที่คำนึงถึง COLREGs เข้ากับโมดูลการประมวลผลแบบขนาน CPU-GPU เพื่อปรับปรุงการทำงานของอัลกอริทึม Proximal Policy Optimization (PPO) จนบรรลุอัตราความสำเร็จในการหลีกเลี่ยงการชนได้ถึง 98.6% ทว่าข้อด้อยของ PPO คือความเสี่ยงด้านข้อจำกัดในการสำรวจ ซึ่งอาจทำให้ระบบลู่เข้าสู่ค่าเหมาะสมที่สุดเฉพาะที่ แทนที่จะเป็นค่าเหมาะสมที่สุดสมบูรณ์ นอกเหนือจาก PPO แล้ว ยังมีการประยุกต์ใช้อัลกอริทึม DRL รูปแบบอื่น ๆ ตัวอย่างเช่น Pan et al [14] ได้ปรับปรุงอัลกอริทึม DQN โดยใช้ฟังก์ชันรางวัลที่ถ่วงน้ำหนักความสัมพันธ์เชิงพื้นที่ แม้จะสามารถบรรลุอัตราการปฏิบัติตาม

COLREGs ได้ 94.3% และลดอุบัติเหตุลงได้ 30.8% แต่การที่อัลกอริทึม DQN ต้องใช้พื้นที่การกระทำแบบไม่ต่อเนื่อง ถือเป็นข้อจำกัดสำคัญที่อาจนำไปสู่ความไม่เสถียรในการควบคุมภายใต้สภาพแวดล้อมพลวัตแบบเรียลไทม์ ในทางตรงกันข้าม Zhang and Sun [15] พบว่าอัลกอริทึม Deep Deterministic Policy Gradient (DDPG) มีอัตราการลู่เข้าที่รวดเร็วกว่าและให้ความแม่นยำสูงกว่า DQN จากการใช้พื้นที่การกระทำแบบต่อเนื่อง ในการประเมินระบบหลักเลี้ยวการชนของเรืออัตโนมัติในแม่น้ำ สอดคล้องกับงานวิจัยของ Liu et al [16] ที่แก้ไขปัญหาโดยใช้อัลกอริทึม Twin-Delayed Deep Deterministic Policy Gradient (TD3) ซึ่งพัฒนาต่อยอดมาจาก DDPG ร่วมกับฟังก์ชันรางวัลแบบหลายองค์ประกอบ ทำให้มีอัตราความสำเร็จสูงถึง 99.23% อย่างไรก็ตาม ระบบเหล่านี้ยังมีข้อจำกัดที่พึ่งพาเพียงแบบจำลองจลนศาสตร์อย่างง่าย โดยละเลยผลกระทบทางอุทกพลศาสตร์และยังขาดการบูรณาการพฤติกรรมของระบบขับเคลื่อนตลอดจนแรงภายนอกทางสิ่งแวดล้อม

ในมิติของการปรับปรุงประสิทธิภาพการใช้เชื้อเพลิง Kim [3] ได้พัฒนาระบบการวางแผนเส้นทางเดินเรือเชิงกลยุทธ์ตามสภาพอากาศโดยใช้สเปกตรัมคลื่น 2 มิติผ่านกระบวนการเขียนโปรแกรมแบบพลวัต ซึ่งช่วยลดการใช้เชื้อเพลิงได้ 6% ต่อมา Kim et al [17] ได้นำเสนอวิธีการควบคุมความเร็วร่วมกับตัวควบคุม Wave Feed Forward (WFF) ที่ชดเชยแรงต้านของคลื่นล่วงหน้า ทำให้ประหยัดเชื้อเพลิงได้ถึง 15% เมื่อพิจารณาในกลุ่มของ DRL งานวิจัยของ Moradi et al [18] พบว่าอัลกอริทึม DDPG ประสบความสำเร็จในการลดเชื้อเพลิงได้สูงสุด (6.64%) ผ่านการปรับปรุงทิศทางและความเร็วจากข้อมูลจริง เมื่อเทียบกับ PPO (1.54%) และ DQN (1.07%) ผลลัพธ์นี้ได้รับการยืนยันโดย Latinopoulos et al [19] ที่ประยุกต์ใช้ DDPG สำหรับการวางแผนเส้นทางเชิงกลยุทธ์ตามสภาพอากาศ ซึ่งลดการใช้เชื้อเพลิงได้ดีกว่า DQN ถึง 4% และดีกว่าแบบจำลองพื้นฐาน 12% สิ่งเหล่านี้ชี้ให้เห็นอย่างชัดเจนว่า DDPG มีความเหมาะสมสูงต่องานด้านการวางแผนเส้นทางนอกเหนือจากการวางแผนเส้นทาง งานวิจัยของ Alshareef and Alghanmi [20] ได้ประยุกต์ใช้ PPO ในการควบคุมเครื่องจักร ซึ่งสามารถเพิ่มประสิทธิภาพพลังงานได้ 10% ผ่านการปรับความเร็วรอบเครื่องยนต์ (RPM) และการจัดการระบบเชื้อเพลิงในสภาวะอากาศเลวร้าย ช่วยลดค่าดัชนี EEXI และ CII ลงได้อย่างมีนัยสำคัญ อย่างไรก็ตาม แม้ PPO จะแสดงศักยภาพที่ดี แต่ยังขาดการศึกษาเปรียบเทียบประสิทธิภาพกับการใช้อัลกอริทึมกลุ่ม DDPG ในการควบคุมเครื่องยนต์ นอกจากนี้ การพึ่งพาชุดข้อมูลในอดีตขนาดใหญ่ ทำให้เกิดข้อจำกัดในการประยุกต์ใช้กับข้อมูลชุดใหม่ ซึ่งเป็นอุปสรรคสำคัญต่อการปรับตัวของระบบ เมื่อต้องนำไปปฏิบัติการกับเรือลำอื่นหรือในสภาพแวดล้อมที่แตกต่างออกไป แม้ว่าการใช้อัลกอริทึม DDPG จะมีแนวโน้มที่ดี แต่ก็มักประสบปัญหาความไม่เสถียรระหว่างการฝึกฝน ดังนั้น Fujimoto et al [21] จึงได้นำเสนออัลกอริทึม TD3 เพื่อแก้ไขข้อบกพร่องดังกล่าว โดยใช้เทคนิค Clipped Double Q-learning เพื่อป้องกันปัญหาการประมาณค่าสูงเกินจริง รวมถึงใช้วิธีการอัปเดตนโยบายแบบ

หน่วงเวลา และการเพิ่มสัญญาณรบกวนในนโยบายเป้าหมาย เพื่อเพิ่มความเสถียร ซึ่งสอดคล้องกับงานวิจัยของ Hadi et al [22] ที่ใช้ TD3 ควบคุมมุมหางเสือได้อย่างปลอดภัยและประหยัดพลังงานโดยไม่ต้องมีความรู้ล่วงหน้าเกี่ยวกับสภาพแวดล้อม อย่างไรก็ตาม แกนหลักที่ทำหยาที่สุดของ DRL คือการออกแบบฟังก์ชันรางวัล เนื่องจากการใช้วิธีถ่วงน้ำหนักรวมแบบง่ายระหว่างเป้าหมายและบทลงโทษ มักนำไปสู่พฤติกรรมที่ระมัดระวังเกินไปจนสิ้นเปลืองพลังงานโดยใช่เหตุ

แม้จะมีความก้าวหน้าทางเทคโนโลยีที่กล่าวมา แต่ยังคงมีช่องว่างทางการวิจัยที่สำคัญ เนื่องจากการศึกษาที่ผ่านมาแยกส่วนการวิเคราะห์ระหว่างประสิทธิภาพการจัดการเชื้อเพลิงและความปลอดภัยในการเดินเรือออกจากกัน ส่งผลให้ระบบที่ประหยัดเชื้อเพลิงอาจเพิ่มความเสี่ยงในการชน หรือในทางกลับกัน ระบบนำทางอัจฉริยะที่ปลอดภัยอาจออกนอกเส้นทางจนสิ้นเปลืองเชื้อเพลิงเกินความจำเป็น นอกจากนี้ กรอบการทำงานแบบเรียลไทม์ที่บูรณาการตัวแปรทางสภาพแวดล้อมเชิงพลวัต (เช่น ความเร็วลม มุมตกกระทบของคลื่น อุณหภูมิ น้ำทะเล รวมถึงตำแหน่งและความเร็วของเรือเป้าหมาย) ยังไม่ได้รับการพัฒนาอย่างสมบูรณ์ อีกทั้งอัลกอริทึม TD3 ยังไม่เคยถูกนำมาประยุกต์ใช้เพื่อแก้ไขปัญหาเชิงบูรณาการลักษณะนี้ ดังนั้น งานวิจัยนี้จึงมุ่งเน้นพัฒนากรอบการทำงานสำหรับเรือ Ro-Pax M/S Nils Holgersson บนแพลตฟอร์ม MATLAB/Simulink โดยประยุกต์ใช้อัลกอริทึม TD3 เพื่อสร้างระบบควบคุมการเดินเรือแบบต่อเนื่องที่ประสานรวมการตัดสินใจด้านความปลอดภัย การหลีกเลี่ยงการชน และประสิทธิภาพการจัดการพลังงานเข้าด้วยกัน ภายใต้สภาพแวดล้อมที่จำลองข้อมูลเซ็นเซอร์แบบเรียลไทม์ เพื่อแก้ไขปัญหาความท้าทายดังกล่าวอย่างบูรณาการ

### 3. วัตถุประสงค์

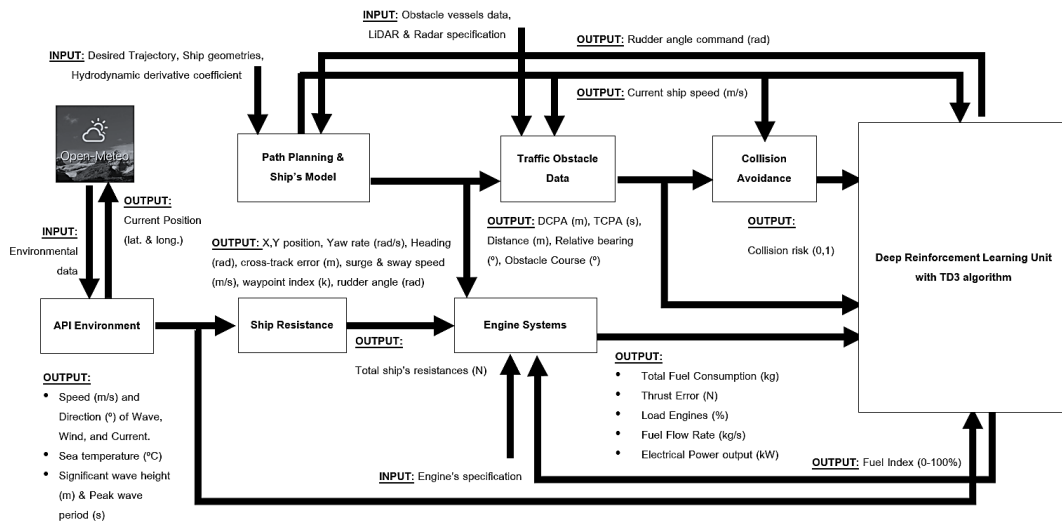
3.1 เพื่อพัฒนาและสร้างแบบจำลองทางคณิตศาสตร์ของเรือ M/S Nils Holgersson บนแพลตฟอร์ม MATLAB/Simulink ให้มีความแม่นยำตามคุณลักษณะเฉพาะจากผู้ผลิต โดยบูรณาการพลศาสตร์การเคลื่อนที่ของเรือ ระบบขับเคลื่อนแบบหลายเครื่องยนต์ ข้อมูลสภาพอากาศ และแรงต้านทานจากสภาพแวดล้อม เพื่อใช้เป็นระบบต้นแบบในการประเมินประสิทธิภาพก่อนการนำไปใช้งานจริง

3.2 เพื่อประยุกต์ใช้และประเมินประสิทธิภาพของระบบการเรียนรู้เสริมกำลังเชิงลึกผ่านอัลกอริทึม TD3 ในการสร้างนโยบายการควบคุม ที่สามารถบริหารจัดการการทำงานของเครื่องจักรเพื่อลดอัตราการสิ้นเปลืองเชื้อเพลิง ควบคู่ไปกับการบังคับทิศทางเพื่อหลีกเลี่ยงการชนได้อย่างมีประสิทธิภาพสูงสุด

#### 4. วิธีการดำเนินการวิจัย

##### 4.1 การพัฒนาแบบจำลองเรือบนแพลตฟอร์ม Simulink

กรอบการทำงานของแบบจำลองเรือบน M/S Nils Holgersson ถูกพัฒนาขึ้นบนแพลตฟอร์ม MATLAB/Simulink โดยประกอบด้วยโมดูลย่อยที่ทำงานประสานกันแบบเรียลไทม์ ดังแสดงในรูปที่ 1 โครงสร้างหลักประกอบด้วย ส่วนต่อประสานโปรแกรมประยุกต์ข้อมูลสภาพอากาศ หรือที่ดึงข้อมูลอุตุนิยมวิทยาตามจำนวนร่วมกับโมดูลพลศาสตร์การเคลื่อนที่และเส้นทาง เพื่อประเมินแรงต้านทานของเรือบ และส่งต่อไปยังระบบเครื่องจักร ในขณะเดียวกัน โมดูลข้อมูลอุปสรรคการจราจรจะคำนวณพารามิเตอร์ด้านความปลอดภัย เช่น ระยะห่างที่ใกล้ที่สุด และระยะเวลาที่เรือบจะเคลื่อนที่เข้าใกล้กันที่สุด เพื่อส่งให้โมดูลการหลีกเลี่ยงการชนประเมินความเสี่ยง ข้อมูลสถานะทางพลวัตและสภาพแวดล้อมทั้งหมดนี้ จะถูกส่งเข้าสู่หน่วยการเรียนรู้เสริมกำลังเชิงลึกที่ขับเคลื่อนด้วยอัลกอริทึม TD3 เพื่อทำหน้าที่เป็นตัวแทน ในการตัดสินใจสั่งการมุมหางเสือและดัชนีเชื้อเพลิงอย่างเหมาะสม นอกจากนี้ โครงสร้างของแบบจำลองยังถูกออกแบบมาให้สามารถสลับการทำงานเพื่อเปรียบเทียบผลลัพธ์ระหว่างระบบ DRL และระบบควบคุมแบบดั้งเดิมที่ใช้ตัวควบคุมแบบสัดส่วน-ปริพันธ์-อนุพันธ์ (PID) ได้ภายใต้เงื่อนไขเดียวกัน การเลือกระบบควบคุม PID มาเป็นเกณฑ์อ้างอิง เนื่องจากเป็นวิธีการควบคุมมาตรฐานที่ได้รับการยอมรับและใช้งานอย่างแพร่หลายในอุตสาหกรรมพาณิชย์ในปัจจุบัน โดยตัวควบคุม PID จำนวน 3 ชุดได้รับการกำหนดค่าพารามิเตอร์โดยผู้ใช้งาน เพื่อให้ได้สมรรถนะการตอบสนองที่ราบเรียบและมีค่าฟุงเกิน (Overshoot) ต่ำที่สุด พารามิเตอร์ที่ใช้กำหนดค่าไว้ที่  $K_p = 0.5$ ,  $K_i = 0.01$ ,  $K_d = 0.1$  สำหรับการควบคุมความเร็ว และ  $K_p = 1.2$ ,  $K_i = 0.05$ ,  $K_d = 0.2$  สำหรับการควบคุมหางเสือ ทั้งนี้ เพื่อให้การเปรียบเทียบประสิทธิภาพมีความเที่ยงตรงสูงสุดทั้งสองระบบจึงถูกทดสอบภายใต้สภาพแวดล้อมที่ควบคุมตัวแปรอย่างเข้มงวด ได้แก่ การใช้เส้นทางเดินเรือบเดียวกัน ข้อมูลสภาพอากาศแบบเรียลไทม์จากส่วนต่อประสานโปรแกรมประยุกต์ชุดเดียวกัน และการตั้งค่าเรือกัดขวางจำนวน 5 ลำในพิกัดตำแหน่งเริ่มต้นที่เหมือนกันทุกประการ



รูปที่ 1 แผนผังแบบจำลองเรืออัตโนมัติบนแพลตฟอร์ม MATLAB/Simulink

#### 4.2 แบบจำลองการเคลื่อนที่ของเรือและระบบขับเคลื่อน

แบบจำลองการเคลื่อนที่ของเรือถูกพัฒนาขึ้นในรูปแบบ 3 แกนอิสระบนระนาบแนวนอน โดยอ้างอิงตามหลักการของ Maneuvering Modeling Group (MMG) [23] ซึ่งครอบคลุมพลศาสตร์ของแรงและความเร่งในแนวแกนพุ่งไปข้างหน้า (Surge, X), แกนเลื่อนด้านข้าง (Sway, Y) และแกนหมุนหันเห (Yaw, N) รวมถึงพิจารณาอัตราการเปลี่ยนแปลงตำแหน่งและมุมหักเหของเรือ นอกจากนี้ เรือต้นแบบใช้ระบบขับเคลื่อนแบบดีเซล-ไฟฟ้า ซึ่งประกอบด้วยเครื่องยนต์ MAN B&W M43 series จำนวน 5 เครื่อง โดยแต่ละเครื่องยนต์ได้รับการจำลองให้เป็นระบบย่อยที่เป็นอิสระจากกัน โดยมีสมการที่เกี่ยวข้องกับระบบขับเคลื่อนดังนี้

1) ความดันประสิทธิภาพเฉลี่ยทางเบรค หรือ Brake Mean Effective Pressure (Pa) ดังสมการ

$$BMEP = \frac{2\pi T}{V_d} \quad (1)$$

เมื่อ  $T$  คือ แรงบิดทางเบรค (Nm)

$V_d$  คือ ปริมาตรกระบอกสูบรวม ( $m^3$ )

2) แรงบิดของใบจักร หรือ Propeller Load Torque (Nm) ดังสมการ

$$T_{load} = \rho \cdot K_Q \cdot n^2 \cdot D^5 \quad (2)$$

เมื่อ  $\rho$  คือ ความหนาแน่นของน้ำ ( $\text{kg/m}^3$ )

$K_Q$  คือ สัมประสิทธิ์แรงบิดของใบจักร

$n$  คือ ความเร็วรอบการหมุนของใบจักร (rev/s)

$D$  คือ เส้นผ่านศูนย์กลางของใบจักร (m)

### 4.3 แบบจำลองแรงต้านทานของเรือ

แรงต้านทานรวมที่กระทำต่อตัวเรือในแบบจำลอง ประกอบด้วยผลรวมของแรงต้านทานจากสภาพแวดล้อมและแรงต้านทานทางอุทกพลศาสตร์ ซึ่งถูกแบ่งออกเป็น 4 องค์ประกอบหลัก ดังนี้

#### 4.3.1 แรงต้านทานจากคลื่น (Wave Resistance)

ประเมินโดยใช้วิธีการของ Alexandersson [24] ดังสมการ

$$\overline{R_{AW}} = 2 \int_0^\infty R(\omega) \cdot S(\omega) d(\omega) \quad (3)$$

เมื่อ  $R(\omega)$  คือ ฟังก์ชันถ่ายโอนแรงต้านทานคลื่นเพิ่ม ( $\text{N}/(\text{m}^2/\text{s})$ )

$S(\omega)$  คือ สเปกตรัมพลังงานคลื่น ( $\text{m}^2/\text{s}$ )

$\omega$  คือ ความถี่เชิงมุมของคลื่น ( $\text{rad/s}$ )

#### 4.3.2 แรงต้านทานจากลม (Wind Resistance)

ประเมินจากผลรวมของแรงต้านที่เกิดจากลมสัมพัทธ์ (Relative Wind) และแรงต้านที่เกิดจากความเร็วของเรือเอง (Self-Induced Air Resistance) ตามวิธีการของ Park et al [25] ดังสมการ

$$R_{wind} = \frac{1}{2} \rho_A C_X(\psi_R) A_T \alpha V_R^2 - \frac{1}{2} \rho_A C_X(0) A_T \alpha V_S^2 \quad (4)$$

เมื่อ  $\rho_A$  คือ ความหนาแน่นของอากาศ ( $\text{kg/m}^3$ )

$C_X(\psi_R)$  คือ สัมประสิทธิ์แรงต้านลมตามแนวแกน (ที่มุม  $\psi_R$ )

$A_T$  คือ พื้นที่ฉายด้านหน้าเหนือแนวน้ำ ( $m^2$ )

$A$  คือ ตัวแปรโปรไฟล์ลม

$V_R$  คือ ความเร็วลมสัมพัทธ์ ( $m/s$ )

$V_S$  คือ ความเร็วของเรือผ่านน้ำ ( $m/s$ )

#### 4.3.3 แรงต้านทานตามยาวจากกระแสหน้า (Longitudinal Current Resistance)

คำนวณตามแนวลำตัวเรือในทิศทางเดินหน้าตามวิธีการของ Beji [26] ซึ่งประยุกต์ใช้พื้นฐานจากสมการ Morison ดังสมการ

$$R_C = \frac{1}{2} \rho A_{px} u_r \left| u_r \right| \left[ \left[ 1 + \frac{B}{A_{px}} \cdot \frac{\sin^2 \beta_x}{2g} u_r^2 \right] \sin^2 \beta_x + 0.072 \left( \frac{Re_x}{\sqrt{\nabla}} \right)^{-1/5} \cdot \frac{S_\omega}{A_{px}} \right] \quad (5)$$

เมื่อ  $\rho$  คือ ความหนาแน่นของน้ำ ( $kg/m^3$ )

$A_{px}$  คือ พื้นที่ฉายด้านหน้าของเรือ ( $m^2$ )

$u_r$  คือ ความเร็วสัมพัทธ์ตามยาว ( $m/s$ )

$B$  คือ ความกว้างของเรือ ( $m$ )

$g$  คือ ความเร่งเนื่องจากแรงโน้มถ่วง ( $m/s^2$ )

$\sin^2 \beta_x$  คือ ฟังก์ชันมุมรูปทรง

$Re_x$  คือ เลขเรย์โนลด์ตามยาว

$\nabla$  คือ ปริมาตรการแทนที่ของเรือ ( $m^3$ )

$S_\omega$  คือ พื้นที่ผิวเปียกของเรือ ( $m^2$ )

#### 4.3.4 แรงต้านทานในน้ำเรียบ (Calm Water Resistance)

คำนวณตามวิธีการของ Holtrop and Mennen ผ่านการประมวลผลด้วยโปรแกรม MAXSURF Resistance โดยในกรณีศึกษาของเรือลำนี้ มีค่าแรงต้านทานในน้ำเรียบอยู่ที่ประมาณ 2,150 กิโลนิวตัน ณ ความเร็วปฏิบัติการเฉลี่ยที่ 9.62 m/s (18.69 นอต)

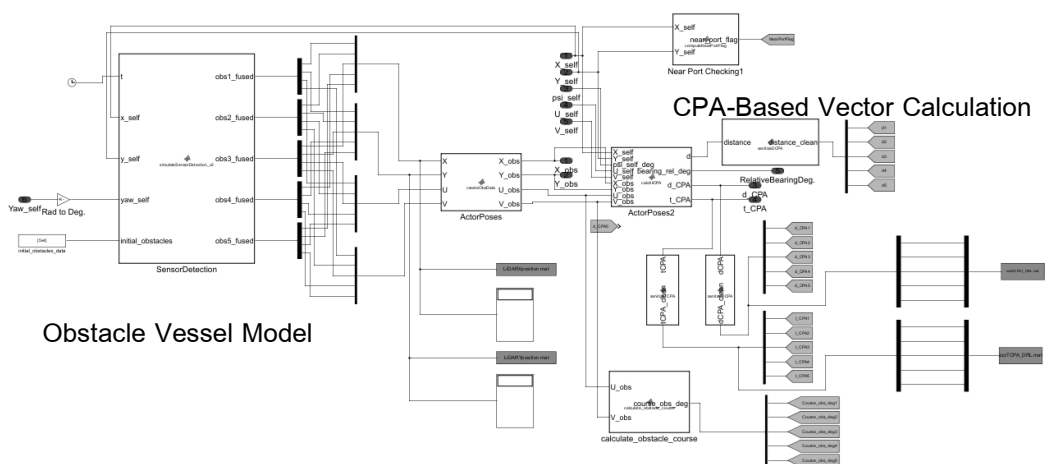
#### 4.4 ระบบการรับรู้สภาพแวดล้อมและการประเมินความเสี่ยงในการชน

ระบบการรับรู้สภาพแวดล้อมของแบบจำลองถูกพัฒนาขึ้นโดยบูรณาการข้อมูลจากเซ็นเซอร์จำลองเรดาร์และไลดาร์ เพื่อทำหน้าที่ตรวจจับและติดตามพิกัดของเรือเป้าหมาย หรือสิ่งกีดขวางในบริเวณใกล้เคียง เนื่องจากข้อมูลดิบที่ได้รับจากเซ็นเซอร์มักมีความคลาดเคลื่อนและถูกรบกวน

ด้วยสัญญาณรบกวนสภาพแวดล้อม งานวิจัยนี้จึงประยุกต์ใช้ตัวกรองคาลมาน [27] ในการทำหน้าที่เป็นตัวประมาณค่าสถานะแบบเวียนซ้ำ เพื่อกรองสัญญาณรบกวนและประมาณค่าพิกัดตำแหน่ง รวมถึงความเร็วของเรือเป้าหมายพลวัตให้มีความแม่นยำสูงสุด โดยข้อมูลสถานะของเรือเป้าหมายที่ผ่านการกรองแล้ว จะถูกนำมาคำนวณดัชนีความเสี่ยงในการชนโดยใช้หลักการประเมินจุดเข้าใกล้ที่สุด [28] ซึ่งพิจารณาจากพารามิเตอร์หลัก 2 ประการ ได้แก่ ระยะห่างที่ใกล้ที่สุด และระยะเวลาที่เรือจะเคลื่อนที่เข้าใกล้กันที่สุด โดยสามารถคำนวณได้จากพิกัดตำแหน่งและความเร็วสัมพัทธ์ และผลการคำนวณดังกล่าวจะถูกส่งต่อไปยังระบบการเรียนรู้เสริมกำลังเชิงลึก เพื่อใช้เป็นตัวแปรสถานะในการประเมินระดับความเสี่ยงและกำหนดนโยบายการหลีกเลี่ยงพลวัตต่อไป

#### 4.5 การจำลองสภาพแวดล้อมทางทะเลและสถานการณ์จราจรทางน้ำ

ระบบจำลองสภาพแวดล้อมถูกพัฒนาขึ้นเพื่อสร้างสภาวะการเดินเรือที่สมจริง โดยมีการนำเข้าตัวแปรทางอุทุนิยมวิทยาและสมุทรศาสตร์จำนวน 10 พารามิเตอร์หลัก อาทิ ความเร็วและทิศทางของคลื่น ลม กระแสน้ำ ตลอดจนอุณหภูมิของน้ำทะเล ข้อมูลเหล่านี้ถูกดึงมาใช้งานแบบเรียลไทม์ผ่านส่วนต่อประสานโปรแกรมประยุกต์จากฐานข้อมูล Open-Meteo Marine Weather [29] โดยอาศัยชุดคำสั่งภาษา Python ในการประมวลผลข้อมูลและส่งต่อไปยังแพลตฟอร์ม MATLAB/Simulink อย่างต่อเนื่อง นอกเหนือจากสภาพแวดล้อมทางธรรมชาติ แบบจำลองยังได้ออกแบบสถานการณ์จำลองการจราจรทางน้ำ เพื่อทดสอบขีดความสามารถของระบบโดยกำหนดให้มีเรือเป้าหมาย จำนวน 5 ลำ เคลื่อนที่ผ่านในพิกัดต่าง ๆ ตลอดเส้นทางเดินเรือ การจำลองสถานการณ์ดังกล่าวมีวัตถุประสงค์เพื่อประเมินประสิทธิภาพและอัตราความสำเร็จในการตัดสินใจหลีกเลี่ยงการชนแบบพลวัตของระบบควบคุม



รูปที่ 2 แบบจำลองระบบการรับรู้และการหลีกเลี่ยงการชนกัน

#### 4.6 การตั้งค่าและการฝึกฝนแบบจำลองด้วยอัลกอริทึม TD3

งานวิจัยนี้เลือกใช้อัลกอริทึม TD3 เป็นแกนหลักในการเรียนรู้ของตัวแทน เนื่องจากมีกลไกที่สามารถแก้ไขข้อบกพร่องของอัลกอริทึม กลุ่ม DDPG ตั้งเดิมได้อย่างมีประสิทธิภาพ โดยมีจุดเด่นสำคัญ 3 ประการ ได้แก่ 1) การป้องกันปัญหาการประมาณค่าสูงเกินจริง ผ่านเทคนิค Clipped Double Q-Learning ซึ่งใช้โครงข่าย Critic สองชุดประเมินค่าและเลือกค่าที่ต่ำกว่ามาใช้งาน 2) การเพิ่มความเสถียรในการฝึกฝนผ่านการอัปเดตนโยบายแบบหน่วงเวลา โดยจะอัปเดตโครงข่าย Actor ทุก ๆ 2 รอบของการอัปเดตโครงข่าย Critic เพื่อให้มั่นใจว่า Critic มีความแม่นยำเพียงพอก่อนนำไปชี้แนะ Actor และ 3) การใช้เทคนิค Target Policy Smoothing โดยการเพิ่มสัญญาณรบกวน ลงในการกระทำเป้าหมาย เพื่อลดความผันผวนและทำให้พื้นที่การเรียนรู้มีความราบเรียบขึ้น

การฝึกฝนแบบจำลองดำเนินการผ่านชุดเครื่องมือ Reinforcement Learning Toolbox บนแพลตฟอร์ม MATLAB/Simulink สถาปัตยกรรมการเรียนรู้ประกอบด้วยโครงข่ายประสาทเทียมแบบสองส่วนหลัก ได้แก่ โครงข่าย Actor ทำหน้าที่สร้างนโยบายการควบคุม และโครงข่าย Twin Critics ทำหน้าที่ประเมินคุณภาพของนโยบายผ่านค่าผลตอบแทนที่คาดหวัง (Q-Value) ในการออกแบบตัวแปรสถานะและคำสั่ง แบบจำลองจะรับข้อมูลสถานะสิ่งแวดล้อม (Observation Space,  $S \in R^{37}$ ) จำนวน 37 พารามิเตอร์ และส่งออกคำสั่งการกระทำแบบต่อเนื่อง (Continuous Action Space,  $A \in R^6$ ) จำนวน 6 ชุดคำสั่ง ซึ่งประกอบด้วยคำสั่งบังคับมุมหางเสือ 1 คำสั่ง และคำสั่งดัชนีการใช้เชื้อเพลิงสำหรับเครื่องยนต์หลักทั้ง 5 เครื่อง

ทิศทางการเรียนรู้ของแบบจำลอง TD3 ถูกกำหนดโดยฟังก์ชันรางวัลแบบหลายวัตถุประสงค์ซึ่งทำหน้าที่ประเมินและให้คะแนนการกระทำใน 4 มิติหลัก ได้แก่ มิติด้านประสิทธิภาพพลังงาน มิติด้านความปลอดภัย มิติด้านการนำทาง และมิติด้านความสำเร็จของภารกิจ นอกจากนี้ยังมีการกำหนดเงื่อนไขการสิ้นสุดรอบการฝึกไว้ 2 กรณี ได้แก่ ความล้มเหลว ซึ่งระบบจะตัดจบและลงโทษเมื่อเรือเป้าหมายมีระยะห่างน้อยกว่า 10 เมตร (เกิดการชน) และ ความสำเร็จ เมื่อเรือสามารถเดินทางถึงจุดหมายปลายทางที่กำหนดได้อย่างปลอดภัย

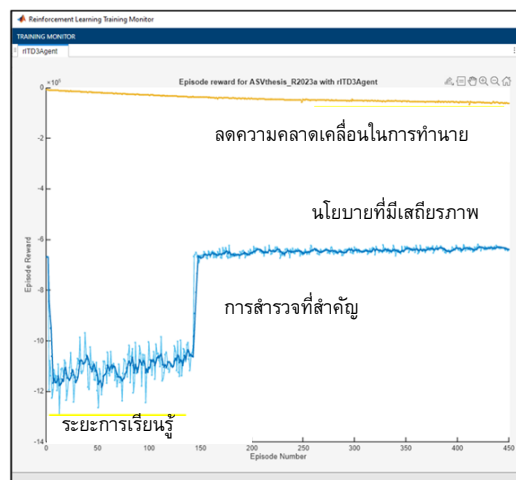
ในการทดลองนี้ การตั้งค่าจำนวนรอบการฝึกฝนสูงสุดถูกกำหนดไว้ที่ 450 รอบ โดยพิจารณาจากผลการประเมินเบื้องต้นซึ่งพบว่า พฤติกรรมการเรียนรู้ของแบบจำลองเริ่มแสดงการลู่เข้าและมีเสถียรภาพในระดับที่เพียงพอต่อการนำนโยบาย ไปใช้งานเพื่อบรรลุวัตถุประสงค์หลักได้ อย่างไรก็ตาม ในทางทฤษฎีนั้นการขยายจำนวนรอบการฝึกฝนให้สูงขึ้น (เช่น 1,000 รอบ) มีแนวโน้มที่จะช่วยเปิดโอกาสให้ตัวแทนสามารถสำรวจและค้นพบจุดเหมาะสมที่สุดสมบูรณ์ได้ดียิ่งขึ้น แต่เนื่องด้วยข้อจำกัดด้านทรัพยากรฮาร์ดแวร์สำหรับการประมวลผลที่ไม่ได้ใช้หน่วยประมวลผลกราฟิก (GPU) เฉพาะทาง ทำให้การจำลองสภาพแวดล้อมต้องใช้ระยะเวลา

ค่อนข้างสูง ดังนั้น การกำหนดรอบการฝึกฝนที่ 450 รอบจึงเป็นจุดสมดุล ที่เหมาะสมที่สุดระหว่างขีดความสามารถในการเรียนรู้ของระบบและทรัพยากรเวลาที่ใช้ในการประมวลผล

## 5. ผลการวิจัยและอภิปรายผล

### 5.1 พฤติกรรมการเรียนรู้และการลู่เข้าของตัวแทน TD3

จากข้อมูลการติดตามการฝึกฝนระบบด้วยอัลกอริทึม TD3 ตลอด 450 รอบ ดังรูปที่ 3 ผลลัพธ์ได้แสดงให้เห็นถึงพัฒนาการและพฤติกรรมการเรียนรู้เชิงลึกของตัวแทนอย่างชัดเจน ในช่วง 150 รอบแรกของการจำลอง ผลตอบแทนเฉลี่ยต่อรอบ มีความผันผวนค่อนข้างสูง โดยแกว่งตัวอยู่ที่ระดับประมาณ  $-1,100,000$  ซึ่งเป็นพฤติกรรมปกติที่บังชี้ถึงช่วงการเริ่มต้นเรียนรู้และการสำรวจสภาพแวดล้อม ของตัวแทน ตามมาด้วยช่วงเวลาที่มีความแปรปรวนเพิ่มขึ้น ซึ่งสะท้อนถึงกระบวนการสำรวจเชิงลึกที่สำคัญขณะที่ระบบกำลังพยายามก้าวข้ามจุดเหมาะสมที่สุดเฉพาะที่หลังจากผ่านพ้นช่วงการสำรวจดังกล่าว ผลตอบแทนสะสมได้เพิ่มขึ้นอย่างรวดเร็วและลู่เข้าสู่สถานะที่มีเสถียรภาพ ที่ระดับผลตอบแทนประมาณ  $-639,640$  อย่างต่อเนื่องจนถึงรอบที่ 450 ซึ่งพิสูจน์ให้เห็นว่าตัวแทนสามารถกำหนดนโยบายการควบคุม ที่มีประสิทธิภาพได้อย่างสมบูรณ์ การลู่เข้านี้ยังมีความสอดคล้องกับพฤติกรรมของค่าเมตริก Q0 (Q-value) ของแต่ละรอบการจำลอง ซึ่งปรับตัวลดลงอย่างต่อเนื่องจาก 0 ไปสู่ระดับ  $-61,196.03$  การลดลงของความคลาดเคลื่อนในการทำนายค่า Q-value ดังกล่าว เป็นตัวบ่งชี้สำคัญทางวิศวกรรมที่ยืนยันว่า โครงข่าย Critic สามารถเรียนรู้และทำนายผลลัพธ์ได้อย่างแม่นยำ ส่งผลให้เกิดการพัฒนาขีดความสามารถในการเรียนรู้ตลอดกระบวนการฝึกฝน จนบรรลุวัตถุประสงค์หลักทั้งด้านการรักษาระยะปลอดภัยและการบริหารจัดการพลังงานได้อย่างบูรณาการ



รูปที่ 3 การติดตามการฝึกฝนของระบบกระบวนการเรียนรู้เชิงลึกโดยใช้อัลกอริทึม TD3

## 5.2 การตรวจสอบความถูกต้องของแบบจำลอง

ก่อนการนำแบบจำลองไปใช้ฝึกฝนระบบปัญญาประดิษฐ์ แบบจำลองทางคณิตศาสตร์ที่พัฒนาขึ้นได้รับการตรวจสอบความถูกต้องโดยเปรียบเทียบกับคุณลักษณะทางวิศวกรรมของเครื่องจักรในเรือ M/S Nils Holgersson จากข้อมูลของผู้ผลิต ผลการจำลอง (ดังแสดงในตารางที่ 2) พบว่าค่าความคลาดเคลื่อน ของพารามิเตอร์หลัก อาทิ รอบเครื่องยนต์สูงสุด รอบใบพัด แรงดันประสิทธิผลเฉลี่ยสูงสุด และช่วงกำลังขับเคลื่อน มีค่าอยู่ในระดับต่ำระหว่าง 0.77% ถึง 5.68% ซึ่งยืนยันให้เห็นถึงความแม่นยำและศักยภาพของแบบจำลองในการประยุกต์ใช้เป็นสภาพแวดล้อมเสมือนได้อย่างมีประสิทธิภาพ

ตารางที่ 2 การเปรียบเทียบความคลาดเคลื่อนของแบบจำลอง

| ลำดับ | พารามิเตอร์                           | ข้อมูลจากผู้ผลิต | ผลลัพธ์จากแบบจำลอง | ความคลาดเคลื่อน (%) |
|-------|---------------------------------------|------------------|--------------------|---------------------|
| 1     | รอบเครื่องยนต์สูงสุด (รอบต่อนาที)     | 507              | 500                | -1.38               |
| 2     | รอบใบพัดสูงสุด (รอบต่อนาที)           |                  |                    |                     |
|       | - 8M43C                               | 155              | 153.7              | -0.84               |
|       | - 7M43C                               | 155              | 153.8              | -0.77               |
|       | - 6M43C                               | 160              | 157.4              | -1.62               |
| 3     | แรงดันประสิทธิผลเฉลี่ยสูงสุด (บาร์)   |                  |                    |                     |
|       | - 8M43C                               |                  | 25.51              | -4.63               |
|       | - 7M43C                               |                  | 25.23              | -5.68               |
|       | - 6M43C                               | 26.75            | 27.7               | 3.55                |
| 4     | ช่วงกำลังขับเคลื่อนสูงสุด (กิโลวัตต์) |                  |                    |                     |
|       | - 8M43C                               | 7,200            | 7072               | -1.78               |
|       | - 7M43C                               | 6,300            | 6119               | -2.87               |
|       | - 6M43C                               | 2,880            | 2766               | -3.95               |

## 5.3 การเปรียบเทียบสมรรถนะโดยรวม

ผลการวิเคราะห์เชิงเปรียบเทียบในตารางที่ 3 แสดงให้เห็นว่าวิธีการเรียนรู้เสริมกำลังเชิงลึกมีสมรรถนะเหนือกว่าระบบควบคุมแบบดั้งเดิมในหลายมิติ โดยเฉพาะด้านความปลอดภัยและประสิทธิภาพการใช้พลังงาน ระบบสามารถลดการสิ้นเปลืองน้ำมันเชื้อเพลิงสะสมได้ 1.592%

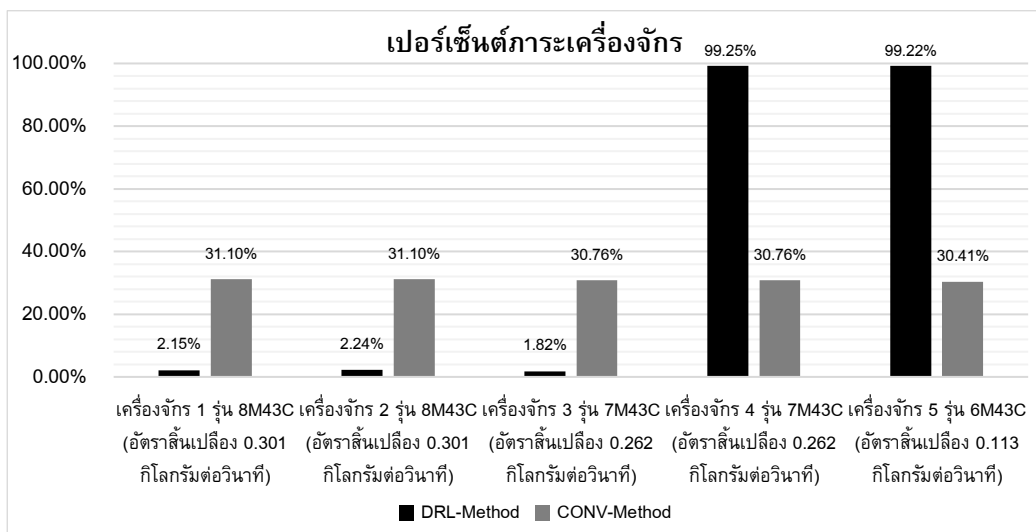
พร้อมทั้งยกระดับความสำเร็จในการหลีกเลี่ยงการชนได้ 1.75% ยิ่งไปกว่านั้น ระบบ DRL ยังมีความสามารถในการควบคุมที่แม่นยำกว่าอย่างเห็นได้ชัด โดยสามารถลดความคลาดเคลื่อนของแรงขับเคลื่อนได้อย่างมีนัยสำคัญถึง 27.12% และสามารถรักษาความเร็วให้สอดคล้องกับค่าเป้าหมายได้ดีขึ้น 4.21% พร้อมทั้งลดแรงต้านทานของเรือสะสมได้ 0.01% อย่างไรก็ดี สมรรถนะที่เพิ่มขึ้นดังกล่าวจำเป็นต้องแลกมาด้วยการเบี่ยงเบนจากเส้นทางที่กำหนด โดยระบบ DRL มีค่าความคลาดเคลื่อนสูงสุดจากเส้นทางเพิ่มขึ้น 19.64% และมีระยะเวลาเดินทางยาวนานขึ้นเล็กน้อยที่ 0.65% การค้นพบนี้สะท้อนให้เห็นถึงกลยุทธ์ของตัวแทนอัจฉริยะที่ให้น้ำหนักกับการค้นหาสมดุลระหว่างความปลอดภัยสูงสุดและการประหยัดเชื้อเพลิงเหนือกว่าการยึดมั่นในเส้นทางและเวลาเดิมอย่างเคร่งครัด ทั้งนี้ หากต้องการปรับปรุงสมรรถนะในอนาคต สามารถดำเนินการได้โดย 1) การปรับจูนน้ำหนักของฟังก์ชันรางวัล เช่น เพิ่มบทลงโทษสำหรับความล่าช้า เพิ่มรางวัลสำหรับการประหยัดเชื้อเพลิง 2) การขยายจำนวนรอบการฝึกฝนบนฮาร์ดแวร์ประมวลผลประสิทธิภาพสูง และ 3) การออกแบบฟังก์ชันรางวัลแบบลำดับขั้น เพื่อจัดลำดับความสำคัญของวัตถุประสงค์ให้ยืดหยุ่นยิ่งขึ้น

ตารางที่ 3 ผลลัพธ์ของการประยุกต์ใช้ระบบการเรียนรู้เสริมกำลังเชิงลึก (DRL)

| ลำดับ | ตัวชี้วัด                                                                         | วิธี DRL   | วิธีดั้งเดิม | การพัฒนา (%) |
|-------|-----------------------------------------------------------------------------------|------------|--------------|--------------|
| 1     | ปริมาณการใช้น้ำมันสะสม (กิโลกรัม)                                                 | 4,546.39   | 4,619.95     | 1.592        |
| 2     | ประสิทธิภาพความปลอดภัยในการหลีกเลี่ยงการชน (%)                                    | 95.46      | 93.71        | 1.75         |
| 3     | ค่าเบี่ยงเบนความเร็วเรือ (เมตร/วินาที) จากความเร็วเรือที่กำหนด (9.62 เมตร/วินาที) | 0.0091     | 0.0095       | 4.21         |
| 4     | ระยะเวลาการเดินทางเรือ (นาที)                                                     | 152.24     | 151.25       | -0.65        |
| 5     | ค่าความคลาดเคลื่อนแรงขับเคลื่อน (นิวตัน)                                          | 9.655      | 13.24        | 27.12        |
| 6     | ค่าความคลาดเคลื่อนจากเส้นทางสูงสุด (เมตร)                                         | 883.72     | 738.61       | -19.64       |
| 7     | แรงต้านทานของเรือสะสม (กิโลนิวตัน)                                                | 23,045,790 | 23,048,158   | 0.010276     |

#### 5.4 กลยุทธ์การจัดการพลังงานและการะเครื่องยนต์

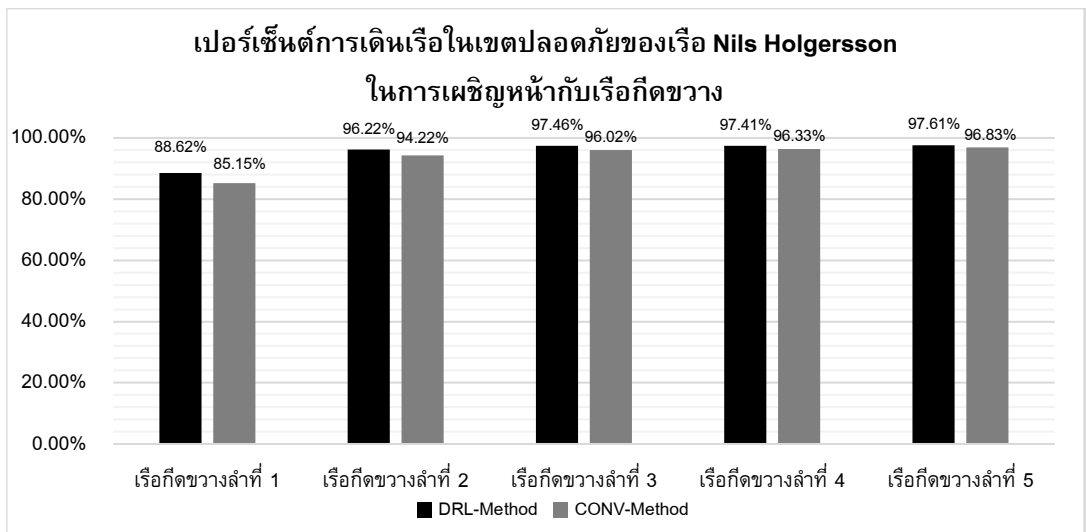
กลยุทธ์สำคัญที่ขับเคลื่อนความสำเร็จด้านพลังงานของระบบ DRL คือความสามารถในการหาค่าเหมาะสมของการกระจายภาระ ให้กับเครื่องยนต์ทั้ง 5 เครื่องที่มีคุณลักษณะแตกต่างกัน จากการวิเคราะห์ดังรูปที่ 4 พบว่าระบบควบคุมแบบดั้งเดิมมีแนวโน้มที่จะกระจายภาระงานให้เครื่องยนต์ทั้งห้าเครื่องอย่างเท่าเทียมกันที่ระดับประมาณ 30% ซึ่งตกอยู่ในช่วงที่เครื่องยนต์มีประสิทธิภาพต่ำ ในทางตรงกันข้าม ระบบ DRL ได้เรียนรู้กลยุทธ์การจัดการที่เหมาะสมกว่าแบบพอสมควร โดยตัดสินใจปิดการทำงานของเครื่องยนต์ 3 เครื่อง และสั่งการให้เครื่องยนต์ 2 เครื่องที่เหลือทำงานด้วยภาระระดับสูง (ประมาณ 99.2%) ซึ่งเป็นจุดที่เข้าใกล้บริเวณที่เครื่องยนต์มีประสิทธิภาพสูงสุด (ค่าความสิ้นเปลืองเชื้อเพลิงจำเพาะต่ำสุด) แม้ว่าการเดินเครื่องที่ 99.2% จะมีประสิทธิภาพด้อยกว่าจุดเหมาะสมที่สุดสมบูรณ์ (ที่ระดับ 85%) เล็กน้อย แต่ผลลัพธ์โดยรวมยังคงเหนือกว่าการเดินเครื่องหลายตัวที่ภาระต่ำอย่างมีนัยสำคัญ



รูปที่ 4 เปอร์เซนต์ภาระของเครื่องจักร โดยวิธีการควบคุมแต่ละแบบ

#### 5.5 สมรรถนะด้านความปลอดภัยและการหลีกเลี่ยงการชน

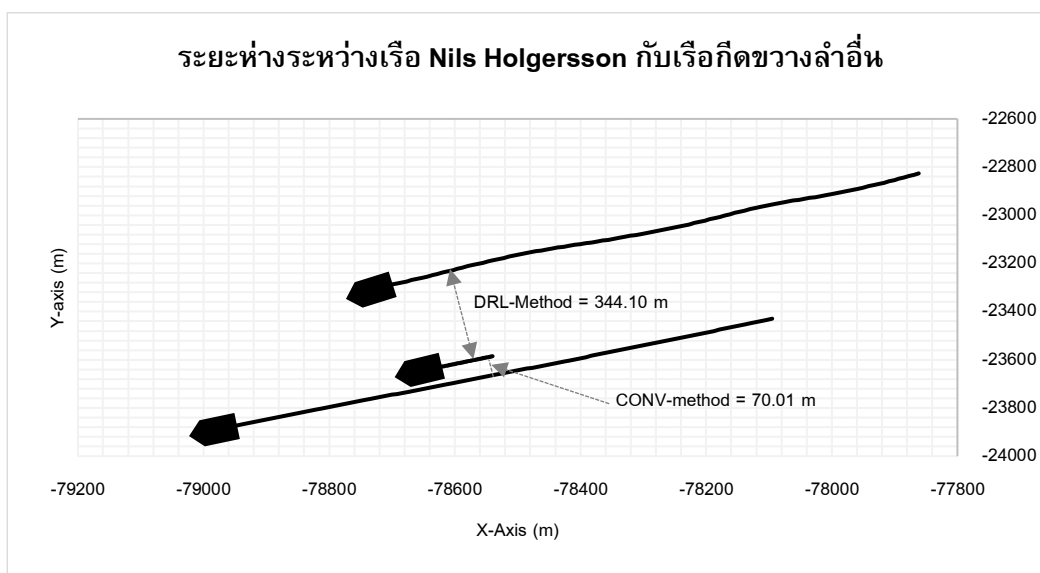
การประเมินสมรรถนะด้านความปลอดภัย พิจารณาจากสัดส่วนเวลาที่เรือสามารถรักษาสถานะอยู่นอกระศมีเขตอันตราย โดยรัศมีดังกล่าวคำนวณจากความเร็วเรือแบบพลวัตผนวกกับระยะเวลาเพื่อสำหรับการหักเลี้ยวหลบสิ่งกีดขวางซึ่งกำหนดไว้ที่ 60 วินาที ตัวอย่างเช่น หากเรือมีความเร็ว ณ ปัจจุบัน 9.62 เมตรต่อวินาที รัศมีเขตอันตรายจะถูกประเมินไว้ที่ 577.2 เมตร



**รูปที่ 5** เปอร์เซ็นต์การเดินเรือในเขตปลอดภัย โดยวิธีการควบคุมแต่ละแบบ

ผลลัพธ์การจำลองดังรูปที่ 5 แสดงให้เห็นว่า ระบบ DRL มีสมรรถนะด้านความปลอดภัยสูงกว่า โดยสามารถประคองเรือให้อยู่ในเขตปลอดภัยได้เฉลี่ย 95.46% ตลอดการเผชิญหน้าในทุกสถานการณ์ ในขณะที่วิธีการควบคุมแบบดั้งเดิมทำได้เพียง 93.71% ซึ่งหมายถึงระบบ DRL เพิ่มความปลอดภัยได้ 1.75%

ความแตกต่างนี้เกิดจากระเบียบวิธีคิดของระบบ ดังแสดงพฤติกรรมการรักษาระยะห่างในรูปที่ 6 วิธีการดั้งเดิมอาศัยกฎเกณฑ์ตายตัวแบบตอบสนอง ซึ่งจะเริ่มดำเนินการหักเลี้ยวก็ต่อเมื่อเข้าใกล้เกณฑ์ความเสี่ยง ส่งผลให้เกิดการเผชิญหน้าในระยะกระชั้นชิดเพียง 70.01 เมตร ซึ่งไม่พึงประสงค์ในการปฏิบัติงานจริง ในขณะที่ระบบ DRL ซึ่งถูกขับเคลื่อนด้วยฟังก์ชันรางวัลทางคณิตศาสตร์ จะประเมินและจดจำผลลัพธ์จากการกระทำเพื่อเพิ่มผลตอบแทนสะสม ให้ได้สูงสุด ระบบจึงแสดงพฤติกรรมเชิงรุก โดยเรียนรู้ที่จะริเริ่มการหลบหลีกอย่างราบรื่นตั้งแต่ระยะไกล ส่งผลให้สามารถรักษาระยะห่างขั้นต่ำได้กว้างขวางถึง 344.10 เมตร ซึ่งยกระดับมาตรฐานความปลอดภัยในการเดินเรือได้อย่างเด่นชัด



**รูปที่ 6 การเปรียบเทียบระยะห่างของเรือ โดยวิธีการควบคุมแต่ละแบบ**

## 6. สรุปผลการวิจัย

การศึกษานี้ประสบความสำเร็จในการพัฒนาและตรวจสอบความถูกต้องของกรอบการทำงานบนแพลตฟอร์ม MATLAB/Simulink สำหรับประเมินระบบควบคุมการเดินเรืออัตโนมัติ M/S Nils Holgersson โดยประยุกต์ใช้ระบบการเรียนรู้เสริมกำลังเชิงลึกผ่านอัลกอริทึม TD3 ผลการวิจัยพิสูจน์ให้เห็นว่าระบบสามารถจัดการกับปัญหาแบบพหุวัตถุประสงค์ ได้อย่างมีประสิทธิภาพ โดยลดการสิ้นเปลืองน้ำมันเชื้อเพลิงลงได้ 1.592% (เทียบเท่า 73.56 กิโลกรัมต่อเที่ยวเดินเรือ) ควบคู่ไปกับการยกระดับความสำเร็จในการหลีกเลี่ยงการชนเพิ่มขึ้น 1.75% เมื่อเปรียบเทียบกับระบบควบคุมแบบดั้งเดิม การค้นพบนี้สนับสนุนสมมติฐานที่ว่า ระบบปัญญาประดิษฐ์มีศักยภาพเหนือกว่าระบบที่อิงกฎเกณฑ์แบบดั้งเดิม เนื่องจากตัวแทนอัจฉริยะสามารถตัดสินใจสั่งการมูมหางเสือและดัชนีการใช้เชื้อเพลิงของเครื่องยนต์ได้อย่างยืดหยุ่น โดยเรียนรู้ที่จะยอมรับการเบี่ยงเบนจากเส้นทางเดิมในระดับที่ปลอดภัย เพื่อแลกกับผลลัพธ์ด้านความปลอดภัยและการประหยัดพลังงานที่คุ้มค่าสูงสุด ความสำเร็จนี้ยังเป็นการยืนยันว่า การจำลองสถานการณ์ที่มีความเที่ยงตรงสูง เป็นแนวทางที่มีประสิทธิผลและคุ้มค่าในการออกแบบ ทดสอบ และลดความเสี่ยงของระบบควบคุมเรืออัตโนมัติในอนาคต

ในมิติด้านสิ่งแวดล้อมและความยั่งยืน การลดปริมาณการใช้น้ำมันดีเซลสำหรับเรือเดินทะเลจำนวน 73.56 กิโลกรัมต่อเที่ยวนั้น นำไปสู่การลดการปล่อยมลพิษทางอากาศได้อย่างเป็นรูปธรรม โดยอ้างอิงจากค่าสัมประสิทธิ์การปล่อยมลพิษมาตรฐานที่ 3.082 ตันของก๊าซคาร์บอนไดออกไซด์ต่อเชื้อเพลิง 1 ตัน [30] พบว่ากรอบการทำงานนี้สามารถช่วยลดการปล่อยก๊าซเรือนกระจก ( $\text{CO}_2$ )

ได้ถึง 226.71 กิโลกรัมต่อเที่ยวเดินเรือ ผลลัพธ์ดังกล่าวนอกจากจะสร้างความคุ้มค่าทางเศรษฐกิจจากการลดต้นทุนเชื้อเพลิงแล้ว ยังตอบสนองต่อนโยบายการลดก๊าซเรือนกระจกขององค์การทางทะเลระหว่างประเทศ (IMO) ซึ่งแสดงให้เห็นถึงความรับผิดชอบต่อสังคมและสิ่งแวดล้อมสอดคล้องกับวิสัยทัศน์และแนวทางปฏิบัติด้านความยั่งยืนของอุตสาหกรรมการเดินทางเรือสมัยใหม่อย่างแท้จริง

## 7. ข้อจำกัดของงานวิจัยและข้อเสนอแนะสำหรับการวิจัยในอนาคต

การศึกษานี้ประสบความสำเร็จในการนำเสนอและตรวจสอบความถูกต้องของกรอบการทำงานบนแพลตฟอร์ม MATLAB/Simulink สำหรับการควบคุมเรืออัตโนมัติแบบบูรณาการด้วยอัลกอริทึม TD3 ผลการวิจัยยืนยันว่าระบบสามารถบรรลุวัตถุประสงค์ด้านความปลอดภัยและประสิทธิภาพพลังงานพร้อมกันได้อย่างมีประสิทธิภาพ อย่างไรก็ตาม งานวิจัยนี้ยังมีข้อจำกัดบางประการที่สามารถนำไปพัฒนาและต่อยอดได้ในอนาคต ดังนี้

1) ด้านความละเอียดของแบบจำลองพลศาสตร์: แบบจำลองการเคลื่อนที่ในงานวิจัยนี้ถูกสร้างขึ้นบนพื้นฐาน 3 แกนอิสระ (3-DOF) ในระนาบแนวนอน ซึ่งละเว้นผลกระทบของการโคลง (Roll) และการยกกระดืบหัวเรือ (Pitch) ที่อาจมีนัยสำคัญในสภาวะคลื่นลมแรง นอกจากนี้แบบจำลองเครื่องยนต์ยังอ้างอิงข้อมูลจากเซ็นเซอร์สมรรถนะของผู้ผลิต ซึ่งอาจไม่สะท้อนถึงการเสื่อมสภาพที่แท้จริงตามอายุการใช้งาน ดังนั้น งานวิจัยในอนาคตควรขยายขีดความสามารถของแบบจำลองเป็น 6 แกนอิสระ (6-DOF) และประเมินผลร่วมกับข้อมูลสมรรถนะเครื่องยนต์ที่ได้จากการปฏิบัติงานจริง

2) ด้านความซับซ้อนของสภาพแวดล้อมและการจราจร: การจำลองสถานการณ์ในงานวิจัยนี้จำกัดจำนวนเรือเป้าหมายไว้ที่ 5 ลำ ซึ่งอาจยังไม่เพียงพอต่อการจำลองสถานการณ์การจราจรที่หนาแน่น บริเวณช่องแคบหรือร่องน้ำเข้าท่าเรือ แนวทางการพัฒนาในระยะต่อไปคือการเพิ่มจำนวนเรืออุปสรรคและพัฒนาระบบจำลองการจราจรทางน้ำให้มีความพลวัตและซับซ้อนมากยิ่งขึ้น

3) ด้านความสามารถในการประยุกต์ใช้กับข้อมูลใหม่และทรัพยากรการประมวลผล: ระบบถูกฝึกฝนและประเมินผลบนเส้นทางทดสอบเฉพาะระยะทาง 43 ไมล์ทะเล ซึ่งอาจจำกัดความสามารถในการนำไปปรับใช้กับเส้นทางเดินเรืออื่นๆ นอกจากนี้ การฝึกฝนยังมีข้อจำกัดด้านทรัพยากรการคำนวณเนื่องจากไม่ได้ใช้หน่วยประมวลผลกราฟิก (GPU) เฉพาะทาง ทำให้ระบบมีความเสี่ยงที่จะเข้าสู่ค่าเหมาะสมที่สุดเฉพาะที่ การวิจัยในอนาคตจึงควรขยายขอบเขตการฝึกฝนบนฮาร์ดแวร์ประสิทธิภาพสูง ทดสอบกับเส้นทางที่หลากหลาย รวมถึงศึกษาเทคนิคการเรียนรู้แบบถ่ายทอด หรือ Transfer Learning เพื่อนำแบบจำลองไปปรับใช้กับเรือประเภทอื่น

4) ด้านความสมจริงของสถานการณ์ปฏิบัติงาน: แบบจำลองปัจจุบันยังไม่ครอบคลุมถึงอัลกอริทึมสำหรับการนำเรือเทียบท่า ซึ่งส่งผลต่อความสมบูรณ์ของโปรไฟล์การดำเนินงานจริง ตลอดจนยังไม่มีกำหนดข้อจำกัดด้านกรอบเวลาการเดินทางที่เข้มงวด ซึ่งเป็น

ตัวแปรสำคัญในอุตสาหกรรมพาณิชย์นาวี การพัฒนาระบบย่อยสำหรับการเทียบท่าและการเพิ่มเงื่อนไขด้านเวลาเข้าเป็นหนึ่งในตัวแปรของฟังก์ชันรางวัล จึงเป็นก้าวสำคัญที่จะทำให้ระบบสมบูรณ์แบบมากยิ่งขึ้น

## References

- [1] Sinay. How much does the shipping industry contribute to global CO2 emissions? [Internet]. 2023 [cited 2025 Sep 1]. Available from: <https://sinay.ai/en/how-much-does-the-shipping-industry-contribute-to-global-co2-emissions/>.
- [2] International Maritime Organization. IMO's work to cut GHG emissions from ships [Internet]. 2023 [cited 2025 Sep 1]. Available from: <https://www.imo.org/en/mediacentre/hottopics/pages/cutting-ghg-emissions.aspx>.
- [3] Kim SW. The development of route decision-making method based on tailor-made forecast 2d wave spectra due to the operation profile of the vessel. Ocean Engineering 2020;197:106907. doi: 10.1016/j.oceaneng.2019.106907.
- [4] McKie R. Maritime Autonomous Surface Ships (MASS) and SAR [Internet]. International Maritime Rescue Federation; 2023 [cited 2025 Sep 1]. Available from: <https://www.international-maritime-rescue.org/news/maritime-autonomous-surface-ships-mass-and-sar# ftn1>.
- [5] Fortune Business Insights. Autonomous ships market size, share, industry analysis & Russia-Ukraine war analysis, by autonomy, by solution, by ship type, by end user, and regional forecast, 2026-2034 [Internet]. Pune: Fortune Business Insights; 2025 [cited 2025 Oct 9]. Report No.: FBI101797. Available from: <https://www.fortunebusinessinsights.com/industry-reports/autonomous-ship-market-101797>.
- [6] Xiao G, Pan L, Lai F. Application, opportunities, and challenges of digital technologies in the decarbonizing shipping industry: a bibliometric analysis. Frontiers in Marine Science 2025;12:1523267. doi: 10.3389/fmars.2025.1523267.
- [7] Heo S, Kang M, Choi J, Yun S, Park J. A preliminary study for development of digital twin ship technology using autonomous surface vehicle. In: 20th International Conference on Ubiquitous Robots (UR); 2023 Jun 25-28; Honolulu, HI, USA. Piscataway (NJ): IEEE; 2023. p. 255-60. doi: 10.1109/UR57808.2023.10202274.

- [8] MathWorks. Simulink: simulation and model-based design [Internet]. Natick, MA: The MathWorks, Inc.; c1994-2026 [cited 2026 Jul 21]. Available from: <https://www.mathworks.com/products/simulink.html>.
- [9] Ahmed YA, Hannan MA, Oraby MY, Maimun A. COLREGs compliant fuzzy-based collision avoidance system for multiple ship encounters. *Journal of Marine Science and Engineering* 2021;9(8):790. doi: 10.3390/jmse9080790.
- [10] Tengesdal T, Brekke EF, Johansen TA. On collision risk assessment for autonomous ships using scenario-based MPC. *IFAC-PapersOnLine* 2020;53(2):14509-16. doi: 10.1016/j.ifacol.2020.12.1454.
- [11] Sutton RS, Barto AG. Reinforcement learning: An introduction. 2nd ed. Cambridge, MA: MIT Press; 2018.
- [12] Wang W, Huang L, Liu K, Wu X, Wang J. A COLREGs-compliant collision avoidance decision approach based on deep reinforcement learning. *Journal of Marine Science and Engineering* 2022;10(7):944. doi: 10.3390/jmse10070944.
- [13] Xie W, Gang L, Zhang M, Liu T, Lan Z. Optimizing multi-vessel collision avoidance decision making for autonomous surface vessels: a COLREGs-compliant deep reinforcement learning approach. *Journal of Marine Science and Engineering* 2024;12(3):372. doi: 10.3390/jmse12030372.
- [14] Pan R, Zhang W, Wang S, Kang S. Deep reinforcement learning model for Multi-Ship collision avoidance decision making design implementation and performance analysis. *Scientific Reports* 2025;15:21250. doi: 10.1038/s41598-025-05636-3.
- [15] Zhang X, Sun T. Deep reinforcement learning for collision avoidance of autonomous ships in inland rivers. *Proceedings of the Institution of Civil Engineers – Transport* 2025;178(5):343-58. doi: 10.1680/jtran.24.00068.
- [16] Liu Z, Zhou Z, Zhang M, Liu J. A twin delayed deep deterministic policy gradient method for collision avoidance of autonomous ships. *Journal of Transport Information and Safety* 2022;40(3):60-74. doi: 10.3963/j.jssn.1674-4861.2022.03.007.
- [17] Kim S, Yun S, You Y. Eco-friendly speed control algorithm development for autonomous vessel route planning. *Journal of Marine Science and Engineering* 2021;9(6):583. doi: 10.3390/jmse9060583.

- [18] Moradi MH, Brutsche M, Wenig M, Wagner U, Koch T. Marine route optimization using reinforcement learning approach to reduce fuel consumption and consequently minimize CO<sub>2</sub> emissions. *Ocean Engineering* 2022;259:111882. doi: 10.1016/j.oceaneng.2022.111882.
- [19] Latinopoulos C, Zavvos E, Kaklis D, Leemen V, Halatsis A. Marine voyage optimization and weather routing with deep reinforcement learning. *Journal of Marine Science and Engineering* 2025;13(5):902. doi: 10.3390/jmse13050902.
- [20] Alshareef MH, Alghanmi AF. Optimizing maritime energy efficiency: a machine learning approach using deep reinforcement learning for EEXI and CII compliance. *Sustainability* 2024;16(23):10534. doi: 10.3390/su162310534.
- [21] Fujimoto S, van Hoof H, Meger D. Addressing function approximation error in actor-critic methods. In: Dy J, Krause A, editors. *Proceedings of the 35th International Conference on Machine Learning (ICML)*; 2018 Jul 10-15; Stockholm, Sweden. PMLR; 2018. p. 1587-96. doi: 10.48550/arXiv.1802.09477.
- [22] Hadi B, Khosravi A, Sarhadi P. Deep reinforcement learning for adaptive path planning and control of an autonomous underwater vehicle. *Applied Ocean Research* 2022;129:103326. doi: 10.1016/j.apor.2022.103326.
- [23] Yasukawa H, Yoshimura Y. Introduction of MMG standard method for ship maneuvering predictions. *Journal of Marine Science and Technology* 2015;20:37-52. doi: 10.1007/s00773-014-0293-y.
- [24] Alexandersson M. A study of methods to predict added resistance in waves [Master's thesis]. Stockholm: KTH Centre for Naval Architecture; 2009. doi: 10.13140/RG.2.2.18796.69767.
- [25] Park H, Lee P, Kim J, Kim H, Lee H, LeeY. Study on the estimation method of wind resistance considering self-induced wind by ship advance speed. *Journal of Marine Science and Engineering* 2024;12:1694. doi: 10.3390/jmse12101694.
- [26] Beji S. Formulation of wave and current forces acting on a body and resistance of ships. *Ocean Engineering* 2020;218:108121. doi: 10.1016/j.oceaneng.2020.108121.
- [27] Welch G, Bishop G. An Introduction to the Kalman Filter [Internet]. Chapel Hill, NC: Department of Computer Science, University of North Carolina at Chapel Hill; 1995 [cited 2025 Oct 9]. Available from: <https://cis.temple.edu/~latecki/Courses/CIS750-03/Papers/KalmanFilterSIGGRAPH2001.pdf>.

- [28] Lazarowska A. Review of collision avoidance and path planning methods for ships utilizing radar remote sensing. Remote Sensing 2021;13:3265. doi: 10.3390/rs13163265.
- [29] Open-Meteo. Marine weather API [Internet]. Open-Meteo; c2022-2026 [cited 2026 Jun 8]. Available from: <https://open-meteo.com/en/docs/marine-weather-api>.
- [30] Wang H, Zhou P, Wang Z. Reviews on current carbon emission reduction technologies and projects and their feasibilities on ships. Journal of Marine Science and Application 2017;16:129-36. doi: 10.1007/s11804-017-1413-y.

### ประวัติผู้เขียนบทความ



**ธนาวุฒิ วิริยะพันธ์** อาจารย์กลุ่มวิชาวิศวกรรมเครื่องกลเรือ ศูนย์ฝึกพาณิชย์นาวี กรมเจ้าท่า 120 หมู่ 7 ซอยเทศบาล 6 ถนนสุขุมวิท ตำบลบางด้วน อำเภอเมืองสมุทรปราการ จังหวัดสมุทรปราการ 10270 โทรศัพท์ 02-7564971 E-mail: t\_wiriya@mmtc.ac.th



**ดร.สาลินี สมบูรณ์ไพศาล** อาจารย์กลุ่มวิชาการทั่วไป ศูนย์ฝึกพาณิชย์นาวี กรมเจ้าท่า 120 หมู่ 7 ซอยเทศบาล 6 ถนนสุขุมวิท ตำบลบางด้วน อำเภอเมืองสมุทรปราการ จังหวัดสมุทรปราการ 10270 โทรศัพท์ 02-7564971 E-mail: s\_somboo@mmtc.ac.th



**พล.ร.ต.ดร.ก้องเกียรติ ลีกนุช** อาจารย์ประจำหลักสูตรกลุ่มวิชาวิศวกรรมเครื่องกลเรือ ศูนย์ฝึกพาณิชย์นาวี กรมเจ้าท่า 120 หมู่ 7 ซอยเทศบาล 6 ถนนสุขุมวิท ตำบลบางด้วน อำเภอเมืองสมุทรปราการ จังหวัดสมุทรปราการ 10270 โทรศัพท์ 02-7564971 E-mail: Lucknuch1959@gmail.com



ผศ. พล.ร.ท. ทวีศักดิ์ มั่นชวนนท์ อาจารย์ประจำหลักสูตร กลุ่มวิชาเอก  
วิศวกรรมเครื่องกล-ซ่อมบำรุงอากาศยาน มหาวิทยาลัยเกษมบัณฑิต  
60 ถนนร่มเกล้า แขวงมีนบุรี เขตมีนบุรี กรุงเทพมหานคร 10510 โทรศัพท์  
02-904-2222 E-mail : tmunchawanont@gmail.com

---

**Article History:**

*Received: October 16, 2025*

*Revised: July 31, 2026*

*Accepted: August 19, 2026*