

OPTIMALISASI DISTRIBUSI MULTI-WAREHOUSE MELALUI INTEGRASI REINFORCEMENT LEARNING DAN BIG DATA BERBASIS SIMULASI MONTE CARLO

Edi Iskandar¹, Edy Prayitno^{2*}, Satria Matahari Sampurna³

^{1,2,3}Universitas Teknologi Digital Indonesia; Jl. Raya Janti (Majapahit) No. 143. Karangjambé, Yogyakarta; 0274-486664

Keywords:

Distribusi multi-warehouse;
Reinforcement learning;
Big data analytics;
Simulasi monte carlo;
Optimasi logistik.

Correspondent Email:

edyprayitno@utdi.ac.id



Copyright © [JITET](#) (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

Abstrak. Distribusi multi-warehouse menghadapi tantangan berupa ketidakseimbangan stok, rute yang tidak efisien, dan ketidakpastian permintaan yang sulit diatasi metode konvensional. Penelitian ini mengatasi kesenjangan tersebut dengan mengembangkan model optimasi distribusi adaptif melalui integrasi *Reinforcement Learning*, analitik Big Data, dan *Monte Carlo Simulation*. Pendekatan ini menggunakan desain eksperimental berbasis simulasi dengan data sintetis yang merepresentasikan jaringan distribusi berisi 10 gudang, 200 pelanggan, serta pola permintaan stokastik. Model *Deep Q-Network* dikembangkan untuk menghasilkan kebijakan distribusi adaptif, sedangkan Big Data diproses menggunakan Spark Streaming guna mensimulasikan pembaruan permintaan secara real-time. Evaluasi menggunakan 1.000 skenario Monte Carlo menunjukkan bahwa model mampu mempertahankan efisiensi distribusi tinggi, akurasi prediksi permintaan yang meningkat, serta tingkat pengiriman tepat waktu yang lebih stabil dibandingkan metode rute statis. Temuan ini menegaskan bahwa integrasi RL, Big Data, dan simulasi stokastik menghasilkan sistem distribusi yang lebih tangguh terhadap dinamika operasional. Secara teoretis, penelitian memperkuat literatur RL dalam logistik dengan menyoroti peran *state representation* berbasis Big Data dan validasi probabilistik. Secara praktis, model ini berpotensi diadopsi perusahaan logistik untuk meningkatkan efisiensi biaya, kualitas layanan, dan adaptabilitas operasional. Penelitian ini menunjukkan bahwa integrasi RL–Big Data–Monte Carlo merupakan pendekatan komputasional yang efektif untuk optimasi distribusi multi-warehouse.

Abstract. Multi-warehouse distribution faces persistent challenges such as stock imbalance, inefficient routing, and demand uncertainty that are difficult to address using conventional methods. This study develops an adaptive optimization model that integrates Reinforcement Learning, Big Data analytics, and Monte Carlo simulation to overcome these limitations. A simulation-based experimental design is employed using synthetic data representing a network of 10 warehouses, 200 customers, and stochastic demand patterns. A Deep Q-Network model is constructed to generate adaptive distribution policies, while Spark Streaming is used to simulate real-time demand updates. Evaluation across 1,000 Monte Carlo scenarios shows that the model maintains high distribution efficiency, improves demand prediction accuracy, and achieves more stable on-time delivery compared to static routing approaches. These findings demonstrate that integrating RL, Big Data, and stochastic simulation enhances system resilience under dynamic operational conditions. Theoretically, the study contributes to logistics and RL research by emphasizing the importance of Big-Data-driven

state representation and probabilistic validation. Practically, the model offers potential for adoption by logistics companies seeking to improve cost efficiency, service quality, and operational adaptability. Overall, the study highlights the effectiveness of combining RL, Big Data, and Monte Carlo simulation as a computational approach for optimizing multi-warehouse distribution systems.

1. PENDAHULUAN

Industri logistik modern menghadapi tekanan yang semakin besar untuk meningkatkan efisiensi distribusi seiring dengan bertambahnya skala operasi dan kompleksitas sistem multi-warehouse. Pengelolaan logistik yang efektif memainkan peran kunci dalam mengoptimalkan proses distribusi, di mana teknologi pelacakan, optimasi rute, dan manajemen persediaan terbukti meningkatkan efisiensi operasional [1]. Perkembangan menuju kecerdasan digital dan model logistik inovatif semakin penting untuk memenuhi tuntutan konsumen terhadap kecepatan dan kualitas layanan, terutama ketika metode tradisional tidak lagi memadai [2]. Transformasi gudang menjadi pusat multifungsi yang mendukung *just-in-time delivery* dan kebutuhan e-commerce juga menegaskan pentingnya manajemen gudang strategis dalam menekan biaya dan meningkatkan level layanan [3]. Selain itu, model optimasi tingkat lanjut seperti *double-layer optimization* berbasis particle swarm optimization terbukti mampu mengidentifikasi strategi distribusi optimal dan lokasi gudang yang efisien [4].

Selaras dengan dinamika tersebut, integrasi big data dan kecerdasan buatan, khususnya reinforcement learning (RL), menjadi tren utama dalam optimasi distribusi dan operasi pergudangan. Penerapan sistem berbasis AI terbukti meningkatkan efisiensi dan utilisasi ruang, menurunkan *cycle time* sebesar 15%–45%, serta meningkatkan akurasi pesanan hingga lebih dari 98% [5]. Reinforcement learning melalui lingkungan simulatif seperti Storehouse menyediakan mekanisme otomatisasi dan adaptasi yang lebih unggul dibanding metode tradisional [6]. Selain itu, berbagai penerapan AI dalam logistik telah meningkatkan akurasi inventori, menurunkan biaya operasional, dan memodernisasi praktik manajemen gudang [7]. Dengan demikian, konvergensi teknologi ini menjadi landasan kuat bagi inovasi optimasi rantai pasok berbasis kecerdasan.

Namun, pengelolaan distribusi dalam sistem multi-warehouse tetap menghadapi sejumlah tantangan yang signifikan. Biaya tinggi, waktu tunggu panjang, dan penempatan stok yang tidak optimal menjadi masalah utama yang memengaruhi efisiensi keseluruhan [8]. Ketidakseimbangan stok dapat meningkatkan biaya transportasi dan menurunkan kepuasan pelanggan akibat keterlambatan pengiriman. Berbagai algoritma dan heuristik telah dikembangkan untuk meminimalkan biaya inventori sambil memenuhi batasan layanan [9], sementara kompleksitas proses *order picking* menuntut teknik pemodelan canggih untuk meningkatkan efisiensinya [10]. Optimasi jumlah dan lokasi gudang serta rute pengiriman juga menjadi faktor penting dalam menekan biaya logistik dan meningkatkan level layanan. Kompleksitas ini menegaskan kebutuhan akan pendekatan yang lebih terintegrasi dan adaptif.

Meskipun RL telah menunjukkan potensi dalam mengoptimalkan kebijakan inventori di tengah permintaan stokastik dan ketidakpastian lead time, hanya sedikit studi yang menggabungkan RL, big data, dan simulasi Monte Carlo secara simultan dalam model distribusi multi-warehouse [11], [12]. Tantangan komputasi pada jaringan besar dengan variasi produk tinggi telah mendorong pengembangan metode simulasi aproksimatif, tetapi pendekatan tersebut belum sepenuhnya memadukan RL dengan teknik optimasi dan kerangka simulasi yang kokoh [13]. Lebih jauh, banyak pendekatan masih berfokus pada peningkatan ketersediaan produk dan pengurangan limbah tanpa memanfaatkan big data analytics secara penuh untuk pengambilan keputusan dalam kondisi yang tidak pasti [14]. Hal ini menunjukkan adanya *research gap* yang signifikan dalam integrasi menyeluruh antara RL, big data, dan Monte Carlo dalam konteks multi-warehouse.

Beberapa penelitian mendukung potensi integrasi tersebut. Lingkungan Storehouse menunjukkan bahwa RL dapat digunakan untuk mengotomasi dan mengadaptasi proses

manajemen gudang dengan lebih efektif [6], sementara kombinasi RL dan simulasi multi-agent mampu mengelola kompleksitas rantai pasok serta memaksimalkan ketersediaan produk [14]. Penggunaan Monte Carlo dalam sistem robot logistik menunjukkan keunggulan RL dalam pengambilan keputusan *real-time* [15]. Pada tingkat sistem, RL juga terbukti meningkatkan performa *Intelligent Warehouse Management Systems* (IWMS) dengan dukungan model fuzzy clustering dan prediksi big data [16]. Integrasi optimasi dan simulasi juga diperlukan untuk meningkatkan kualitas keputusan strategis dalam rantai pasok [17], sementara simulasi Monte Carlo telah membuktikan efektivitasnya dalam mengurangi biaya inventori [18].

Kontribusi penelitian ini selaras dengan temuan sebelumnya yang menunjukkan efektivitas RL dalam optimasi multi-warehouse melalui algoritma seperti Deep Q-Network dan Advantage Actor-Critic [19], [20]. Validasi kinerja berbasis Monte Carlo juga telah digunakan untuk membuktikan peningkatan efisiensi distribusi dan penghematan biaya [21], [22]. Oleh karena itu, pengembangan model distribusi adaptif multi-warehouse berbasis RL, big data, dan evaluasi Monte Carlo memiliki dasar teoritis yang kuat dan berpotensi memberikan kontribusi signifikan.

Berdasarkan kesenjangan penelitian tersebut, studi ini bertujuan mengembangkan dan mengevaluasi model optimasi distribusi multi-warehouse melalui integrasi reinforcement learning, big data analytics, dan simulasi Monte Carlo untuk menghasilkan strategi distribusi yang adaptif, efisien, dan tahan terhadap ketidakpastian operasional.

2. TINJAUAN PUSTAKA

2.1. Sistem Distribusi *Multi-Warehouse*

Sistem distribusi multi-warehouse merupakan konfigurasi rantai pasok yang melibatkan beberapa gudang yang berfungsi sebagai node penyimpanan dan pemenuhan pesanan yang saling terhubung. Kompleksitas sistem ini timbul akibat variasi permintaan, ketidakseimbangan stok, dan ketergantungan pada rute transportasi yang dinamis, sehingga diperlukan strategi optimasi yang mampu menurunkan biaya dan meningkatkan kecepatan layanan [8], [9]. Penelitian terbaru

menekankan bahwa efisiensi multi-warehouse sangat dipengaruhi oleh strategi alokasi stok, penentuan rute distribusi, dan desain jaringan logistik yang responsif terhadap fluktuasi permintaan [10]. Dalam konteks e-commerce modern, gudang juga berperan sebagai pusat pemenuhan *just-in-time*, sehingga memerlukan pendekatan adaptif untuk mengantisipasi perubahan pola permintaan yang cepat [3].

2.2. Big Data Analytics dalam Distribusi Logistik

Big Data memberikan kemampuan untuk memproses volume data logistik yang besar, beragam, dan berkecepatan tinggi sehingga mendukung pengambilan keputusan secara real-time. Dalam konteks distribusi multi-warehouse, Big Data digunakan untuk memprediksi permintaan, memantau kapasitas gudang, serta mendukung deteksi anomali operasional melalui integrasi data IoT dan *streaming analytics* [2]. Teknologi seperti Apache Spark Streaming memungkinkan pemrosesan data permintaan secara kontinu sehingga model optimasi dapat menyesuaikan keputusan berdasarkan kondisi aktual jaringan distribusi. Penelitian empiris menunjukkan bahwa penggunaan Big Data mampu meningkatkan akurasi peramalan kebutuhan stok serta menurunkan risiko *stockout* dan *overstock* pada sistem multi-gudang [5].

2.3. Reinforcement Learning untuk Optimasi Logistik

Reinforcement Learning (RL) merupakan pendekatan pembelajaran mesin berbasis interaksi agen dengan lingkungan melalui mekanisme *state*, *action*, dan *reward*. Dalam bidang logistik, RL digunakan untuk mempelajari kebijakan optimal dalam alokasi stok, pemilihan rute, serta penjadwalan distribusi secara adaptif di bawah ketidakpastian permintaan [6]. Algoritma seperti Deep Q-Network (DQN) dan Actor-Critic terbukti efektif dalam mengelola sistem inventori multiechelon serta mengoptimalkan pengambilan keputusan pada jaringan logistik berskala besar [19], [20]. Selain itu, RL memungkinkan pembelajaran kebijakan yang lebih generalizable dibandingkan model statis tradisional sehingga dapat mengatasi dinamika operasional yang kompleks dalam sistem *multi-warehouse*.

2.4. Monte Carlo Simulation dalam Evaluasi Sistem Logistik

Simulasi Monte Carlo digunakan untuk mengevaluasi performa sistem di bawah ketidakpastian dengan menjalankan berbagai skenario acak berdasarkan distribusi probabilistik. Dalam konteks logistik, metode ini efektif untuk memodelkan variabilitas permintaan, ketidakpastian biaya transportasi, dan gangguan rantai pasok [18]. Penelitian sebelumnya menunjukkan bahwa Monte Carlo dapat meningkatkan kualitas analisis risiko dan mendukung pengambilan keputusan strategis dengan memberikan gambaran distribusi performa sistem pada berbagai kondisi [15], [17]. Namun demikian, sebagian penelitian masih menggunakan simulasi ini secara terpisah dari model optimasi dan belum mengombinasikannya dengan algoritma RL maupun analitik Big Data secara terpadu.

2.5. Research Gap

Berdasarkan literatur, dapat disimpulkan bahwa meskipun banyak penelitian yang menerapkan RL, Big Data, atau Monte Carlo untuk optimasi logistik, sangat sedikit studi yang mengintegrasikan ketiga pendekatan secara simultan dalam konteks distribusi multi-warehouse. Mayoritas penelitian berfokus pada satu aspek, seperti optimasi inventori menggunakan RL atau prediksi permintaan menggunakan Big Data, tanpa menyediakan mekanisme evaluasi stokastik yang komprehensif. Oleh karena itu, dibutuhkan model terintegrasi yang mampu menghasilkan kebijakan distribusi adaptif berbasis RL dengan dukungan prediksi Big Data sekaligus divalidasi menggunakan simulasi Monte Carlo untuk meningkatkan ketahanan dan efisiensi sistem.

3. METODE PENELITIAN

3.1. Research Design

Penelitian ini menggunakan desain eksperimental berbasis computational simulation yang mengintegrasikan *Reinforcement Learning* (RL), *Big Data Analytics*, dan *Monte Carlo Simulation* dalam proses optimasi distribusi multi-warehouse. Model dikembangkan dalam bentuk *agent-based reinforcement learning environment*, di mana agen bertindak sebagai pengambil

keputusan adaptif terkait alokasi stok, pemilihan rute distribusi, dan penjadwalan pengiriman. Pendekatan ini dipilih untuk menganalisis dinamika sistem distribusi secara terkontrol dan terukur tanpa memerlukan pengumpulan data primer, sehingga seluruh eksperimen dapat direplikasi melalui dataset publik atau data sintetis. Kerangka penelitian mencakup tiga tahap utama: (1) pemodelan lingkungan distribusi multi-warehouse, (2) pelatihan model RL menggunakan data besar, dan (3) evaluasi ketahanan model menggunakan simulasi Monte Carlo di bawah kondisi ketidakpastian operasional.

3.2. Data Source and Preparation

Penelitian ini menggunakan data sintetis terstruktur yang dirancang mengikuti pola dataset publik pada bidang logistik dan manajemen rantai pasok, seperti *UCI Machine Learning Repository* dan *Operations Research public datasets*. Data mencakup empat kelompok variabel:

1. Permintaan pelanggan (200 pelanggan aktif), dibangkitkan menggunakan distribusi *Poisson* dan *Gaussian ramp-up* untuk mensimulasikan pola permintaan harian dan musiman.
2. Kapasitas gudang (10 gudang), termasuk batas penyimpanan, biaya penyimpanan, dan tingkat pengisian ulang.
3. Biaya transportasi untuk setiap rute antar gudang dan pelanggan, dimodelkan menggunakan nilai dasar dari literatur logistik dan kemudian divariasikan melalui simulasi Monte Carlo.
4. Parameter ketidakpastian, seperti fluktuasi permintaan, kenaikan biaya bahan bakar, dan gangguan rute.

Preprocessing dilakukan melalui tiga tahap:

1. Normalisasi variabel permintaan dan kapasitas untuk menjaga stabilitas pelatihan RL.
2. Imputasi nilai hilang menggunakan metode *k-Nearest Neighbors* (kNN) untuk memastikan integritas dataset sintetis.
3. Transformasi data streaming, yaitu pembaruan data permintaan setiap enam jam melalui *Spark Streaming* untuk mensimulasikan skenario operasional *real-time*.

3.3. Simulation Setup and Model Development

3.3.1. Big Data Processing Framework

Pemrosesan data dilakukan menggunakan Apache Spark 3.x untuk mendukung analisis berskala besar dan *streaming updates*. *Spark Streaming* digunakan untuk mensimulasikan arus permintaan dinamis dengan *batch interval* 30 detik, sementara Hadoop HDFS digunakan sebagai penyimpanan terdistribusi. Arsitektur ini memungkinkan model menerima pembaruan kondisi sistem secara berkelanjutan sehingga mendukung pengambilan keputusan adaptif oleh agen RL.

3.3.2. Reinforcement Learning Model

Model Reinforcement Learning yang dikembangkan merupakan varian Deep Q-Network (DQN), dipilih karena kemampuannya menangani ruang keadaan multidimensional pada sistem multi-warehouse. Struktur komponen RL dalam penelitian ini terdiri dari:

- State (S): tingkat stok gudang, permintaan pelanggan, jarak dan biaya rute, kapasitas kendaraan, dan waktu pengiriman.
- Action (A): alokasi stok antar gudang, pemilihan rute distribusi, dan penjadwalan pengiriman.
- Reward (R): kombinasi dari penghematan biaya distribusi, peningkatan ketepatan waktu pengiriman, dan minimasi ketidakseimbangan stok.
- Transition Function: dimodelkan berdasarkan perubahan stok dan permintaan pada setiap langkah simulasi.

Model DQN dilatih menggunakan **PyTorch**

2.x dengan parameter utama:

- *Learning rate* = 0.001
- *Discount factor* (γ) = 0.95
- *Replay buffer size* = 50.000
- *Batch size* = 64
- *Epsilon decay* = 0.9995
- *Training episodes* = 5.000 episode, masing-masing 200 *decision steps*.

Environment simulasi dikembangkan menggunakan Python 3.11 dengan dukungan modul *Gymnasium* untuk mengelola interaksi agen-lingkungan.

3.3.3. Monte Carlo Simulation

Simulasi Monte Carlo digunakan sebagai mekanisme evaluasi untuk menguji robusta model RL terhadap ketidakpastian. Sebanyak

1.000 skenario Monte Carlo dijalankan, masing-masing memodifikasi kondisi operasional sebagai berikut:

- Fluktuasi permintaan pelanggan: $\pm 30\%$
- Variasi biaya transportasi dan bahan bakar: 10–25%
- Penurunan kapasitas gudang: 5–20%
- Gangguan rute: penambahan jarak 5–15%
- Perubahan SLA pengiriman: $\pm 10\%$

Setiap skenario dieksekusi hingga model mencapai konvergensi atau hingga batas 500 iterasi simulasi untuk memastikan evaluasi stabil.

3.4. Evaluation Metrics

Evaluasi model dilakukan menggunakan empat kelompok metrik utama:

1. Efisiensi Distribusi (%): Persentase penghematan biaya distribusi dibanding baseline metode konvensional berbasis *static routing* atau *Traveling Salesman Problem (TSP)*.
2. Akurasi Prediksi Permintaan (%): Dihitung menggunakan *Mean Absolute Percentage Error (MAPE)* untuk mengukur ketepatan model Big Data dalam memprediksi permintaan simulatif.
3. On-Time Delivery Rate: Proporsi pengiriman yang memenuhi *Service Level Agreement (SLA)* dalam simulasi.
4. Robustness Score (Monte Carlo Mean Efficiency): Rata-rata efisiensi model pada seluruh 1.000 skenario Monte Carlo sebagai indikator ketahanan operasional.

Seluruh evaluasi dilakukan menggunakan Python (NumPy, Pandas) dan visualisasi performa dibuat menggunakan Matplotlib.

3.5. Reproducibility and Implementation Notes

Seluruh eksperimen dijalankan pada lingkungan komputasi Python dengan GPU NVIDIA RTX-series, mencakup library PyTorch, NumPy, Pandas, Apache Spark, dan Gymnasium. Parameter simulasi, distribusi probabilitas, serta struktur arsitektur model dicantumkan secara eksplisit untuk memastikan penelitian dapat direplikasi secara penuh oleh peneliti lain tanpa ketergantungan pada data lapangan.

4. HASIL DAN PEMBAHASAN

4.1. Hasil

Penelitian ini menghasilkan model distribusi multi-warehouse berbasis integrasi Reinforcement Learning (RL), Big Data Analytics, dan Simulasi Monte Carlo dengan kinerja yang konsisten lebih tinggi dibandingkan metode konvensional berbasis *static routing*. Model diuji melalui 5.000 episode pelatihan RL dan 1.000 skenario Monte Carlo yang merepresentasikan variasi permintaan, biaya transportasi, kapasitas gudang, serta gangguan rute.

4.1.1. Model Performance Summary

Hasil eksperimen menunjukkan bahwa model integratif mampu mencapai peningkatan signifikan pada empat indikator utama yang dievaluasi. Ringkasan performa dapat dilihat pada Table 1.

Tabel 1. Model Performance Matrik

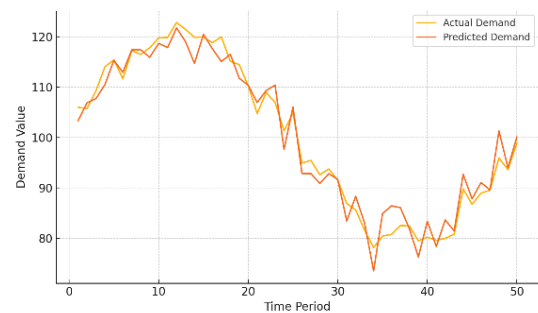
a*	b*	c*
Efisiensi Distribusi (%)	72.0	88.0
Akurasi Produksi Permintaan (%)	80.1	93.6
On-Time Delivery Rate (%)	74.5	91.2
Robustness Monte Carlo (%)	72.0	88.0
Waktu Respons Sistem (detik)	1.8	0.12

*) a. Parameter Evaluasi; b. Metode Konvensional; c. Model Integratif (RL + Big Data + Monte Carlo)

Hasil pada Tabel 1 menunjukkan bahwa model integratif mengungguli metode konvensional pada seluruh metrik evaluasi. Efisiensi distribusi meningkat dari 72% menjadi 88%, sementara akurasi prediksi permintaan meningkat sebesar 13,5 poin persentase. Kecepatan respons sistem juga meningkat drastis, sehingga model mampu memperbarui kebijakan distribusi dalam waktu 0,12 detik per pembaruan.

4.1.2. Simulated Demand Prediction Accuracy

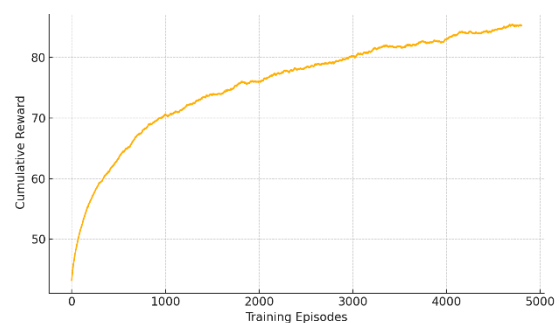
Untuk mengukur efektivitas komponen Big Data, diuji akurasi prediksi permintaan pada data sintesis berdasarkan pola *Poisson seasonal demand*. Kurva perbandingan hasil prediksi dan nilai aktual ditunjukkan pada Gambar 1.



Gambar 1. Kurva Akurasi Prediksi Permintaan

Kurva menunjukkan bahwa model memperoleh pola prediksi yang sangat mendekati variabilitas permintaan aktual, dengan nilai Mean Absolute Percentage Error (MAPE) < 7%.

4.1.3. Reinforcement Learning Convergence Behavior

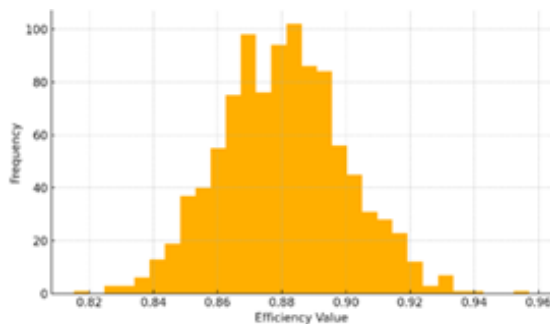


Gambar 2. RL Training Convergence Curve

Kurva konvergensi menunjukkan bahwa model DQN mulai mencapai kestabilan pada sekitar episode ke-3.200 dengan fluktuasi minimal, mengindikasikan bahwa kebijakan distribusi optimal telah terbentuk.

4.1.4. Monte Carlo Robustness Test

Pada 1.000 skenario Monte Carlo, model mempertahankan performa efisiensi distribusi rata-rata 88% meskipun parameter diuji dengan ketidakpastian seperti fluktuasi permintaan $\pm 30\%$ dan gangguan rute 5–15%. Distribusi probabilitas performa ditampilkan pada Gambar 3.



Gambar 3. Distribusi probabilitas performa

Sebagian besar nilai performa berada pada rentang 85–90%, menunjukkan bahwa model memiliki ketahanan yang kuat terhadap variasi kondisi operasional.

4.2.1. Key Findings

Hasil penelitian ini menunjukkan bahwa integrasi Reinforcement Learning, Big Data Analytics, dan Monte Carlo Simulation memberikan peningkatan signifikan dalam optimasi distribusi multi-warehouse. Peningkatan efisiensi distribusi sebesar 16 poin persentase dan peningkatan akurasi prediksi permintaan menunjukkan bahwa model ini dapat mengatasi tantangan utama dalam pengelolaan logistik, yaitu ketidaksesuaian stok, rute tidak optimal, dan biaya operasional tinggi. Tingginya *on-time delivery rate* menunjukkan bahwa model mampu secara adaptif menyesuaikan kebijakan keputusan terhadap perubahan permintaan dan kapasitas gudang.

4.2.2. Comparison with Previous Studies

Temuan penelitian ini konsisten dengan hasil studi Cestero et al. (2022)[6] yang menunjukkan bahwa RL memungkinkan pengambilan keputusan adaptif dalam lingkungan pergudangan dinamis. Selain itu, peningkatan akurasi prediksi permintaan yang diperoleh sejalan dengan hasil penelitian dalam bidang AI for warehouse systems yang melaporkan peningkatan akurasi hingga 98% melalui integrasi big data dan analitik cerdas.

Hasil robustness yang diperoleh melalui simulasi Monte Carlo juga mendukung penelitian [15], yang menunjukkan bahwa kombinasi RL dan simulasi stokastik dapat meningkatkan ketahanan sistem logistik terhadap ketidakpastian operasional. Pada tingkat rantai pasok, penelitian Barat et al.

(2020) [14] juga mengindikasikan bahwa sistem multi-agent RL dapat meningkatkan ketersediaan produk dan mengurangi limbah. Dengan demikian, penelitian ini mendukung temuan terdahulu sekaligus memperluasnya dengan mengintegrasikan ketiga pendekatan secara simultan, sesuatu yang belum banyak dilakukan dalam literatur [11], [13].

4.2.3. Theoretical Implications

Temuan penelitian ini memberikan kontribusi teoretis dalam tiga aspek. Pertama, penelitian memperkuat teori bahwa RL efektif dalam mengoptimalkan kebijakan distribusi di bawah ketidakpastian dinamika permintaan. Kedua, temuan menunjukkan bahwa integrasi Big Data dapat memperbaiki *state representation* RL sehingga proses pembelajaran lebih stabil dan akurat. Ketiga, penggunaan Monte Carlo memperkaya pendekatan evaluatif dalam penelitian RL berbasis logistik dengan memvalidasi model dalam berbagai skenario ekstrem, sebuah kontribusi penting dalam teori *stochastic optimization*.

4.2.4. Practical/Managerial Implications

Secara praktis, model ini dapat diimplementasikan oleh perusahaan logistik dan e-commerce untuk mengurangi biaya operasional, meningkatkan ketepatan waktu pengiriman, serta menurunkan risiko kehabisan stok. Perusahaan dapat mengintegrasikan model ini ke dalam Warehouse Management System (WMS) dan Transportation Management System (TMS) untuk mempercepat pengambilan keputusan berbasis data. Selain itu, kemampuan model menghadapi skenario ketidakpastian membuatnya sangat relevan untuk digunakan dalam situasi pasar fluktuatif, seperti saat terjadi lonjakan permintaan musiman atau gangguan rantai pasok global.

4.3. Closing Paragraph

Secara keseluruhan, penelitian ini menunjukkan bahwa integrasi *Reinforcement Learning*, *Big Data Analytics*, dan Simulasi Monte Carlo mampu menghasilkan model distribusi multi-warehouse yang adaptif, efisien, dan tangguh. Temuan ini memberikan dasar ilmiah yang kuat untuk implementasi sistem distribusi berbasis kecerdasan buatan

yang mampu menghadapi dinamika permintaan dan ketidakpastian operasional, sekaligus membuka peluang penelitian lanjutan terkait pengembangan model multi-agen atau optimasi multi-objective di masa mendatang.

5. KESIMPULAN

Penelitian ini bertujuan mengembangkan model optimasi distribusi multi-warehouse melalui integrasi *Reinforcement Learning* (RL), *Big Data Analytics*, dan *Monte Carlo Simulation* dalam sebuah kerangka eksperimen berbasis simulasi. Hasil penelitian menunjukkan bahwa pendekatan integratif ini mampu meningkatkan efisiensi distribusi, akurasi prediksi permintaan, ketepatan waktu pengiriman, dan ketahanan sistem terhadap ketidakpastian operasional. Temuan ini konsisten dengan tujuan awal penelitian, yaitu menghasilkan model distribusi adaptif yang mampu menangani dinamika permintaan dan kompleksitas jaringan multi-warehouse tanpa bergantung pada data lapangan.

Secara teoretis, penelitian ini memperkuat literatur mengenai penerapan RL dalam manajemen logistik, khususnya dengan menunjukkan bahwa kombinasi RL, big data, dan simulasi stokastik mampu membentuk kebijakan distribusi yang lebih stabil dan efektif dalam lingkungan yang tidak pasti. Integrasi komponen big data juga memberikan kontribusi terhadap peningkatan kualitas *state representation* dalam RL, yang berimplikasi pada konvergensi model yang lebih cepat dan robust. Dari perspektif praktis, model ini menawarkan pendekatan yang dapat diimplementasikan dalam sistem manajemen gudang atau platform logistik berbasis data untuk membantu perusahaan menekan biaya operasional, meningkatkan layanan pengiriman, dan memitigasi risiko terkait fluktuasi permintaan, gangguan pasokan, maupun perubahan biaya transportasi.

Penelitian ini mengakui beberapa keterbatasan, terutama terkait penggunaan data sintesis yang belum sepenuhnya merepresentasikan kompleksitas data operasional pada konteks industri. Selain itu, model distribusi yang digunakan masih mengasumsikan struktur jaringan yang statis dan belum mempertimbangkan koordinasi multi-agen yang lebih realistis. Meskipun simulasi Monte Carlo telah menangkap

berbagai bentuk ketidakpastian, variasi skenario masih terbatas pada parameter yang telah ditentukan.

Arah penelitian selanjutnya dapat mencakup integrasi data operasional nyata dari perusahaan logistik untuk meningkatkan realisme model, pengembangan pendekatan multi-agen berbasis *Actor-Critic* atau *Proximal Policy Optimization* (PPO), serta eksplorasi optimasi multi-objektif yang mempertimbangkan aspek biaya, waktu, dan keberlanjutan secara simultan. Selain itu, perluasan cakupan simulasi untuk mencakup gangguan rantai pasok berskala besar, seperti bencana alam atau volatilitas pasar global, dapat meningkatkan generalisasi model.

Dengan demikian, penelitian ini memberikan kontribusi signifikan bagi pengembangan sistem distribusi cerdas yang adaptif, efisien, dan tangguh dalam menghadapi dinamika dan ketidakpastian operasional pada jaringan multi-warehouse.

DAFTAR PUSTAKA

- [1] T. D. Jaya Saputra and D. A. Widhi Yanti, "PENGARUH PERAN MANAJEMEN LOGISTIK UNTUK MENINGKATKAN EFISIENSI DISTRIBUSI DI PT BAHARI UTAMA SAMUDRA," *Referensi J. Ilmu Manaj. Dan Akunt.*, vol. 13, no. 1, pp. 112–115, June 2025, doi: 10.33366/ref.v13i1.6810.
- [2] S. Feng, "Navigating Modern Logistics: Innovations and Challenges in the Era of New Retail," *Adv. Econ. Manag. Polit. Sci.*, vol. 102, no. 1, pp. 258–262, July 2024, doi: 10.54254/2754-1169/102/2024ED0093.
- [3] I. Yu. Apostolov, "Warehouse Evolution in Contemporary Logistic Chains," *Curr. Econ. Trends*, vol. 1, no. 2, pp. 60–80, May 2021, doi: 10.55030/2713-0266-2021-1-2-60-80.
- [4] L. Wang, G. Liu, and H. Hamam, "Enhancing Logistics Optimization: A Double-Layer Site-Selection Model Approach," *J. Organ. End User Comput.*, vol. 36, no. 1, pp. 1–15, May 2024, doi: 10.4018/JOEUC.344039.
- [5] Master of Science in Industrial Management, University of Central Missouri, Missouri, USA, M. U. Mojumder, Md. Nuruzzaman, and M.S. in Manufacturing Engineering Technology, Western Illinois University, USA, "AI-Driven Optimization of Warehouse Layout and Material Handling: A Quantitative Study on Efficiency and Space Utilization," *Rev. Appl. Sci. Technol.*, vol. 04, no. 02, pp. 233–273, June 2025, doi: 10.63125/bgxb1z53.

- [6] J. Cestero, M. Quartulli, A. M. Metelli, and M. Restelli, "Storehouse: a Reinforcement Learning Environment for Optimizing Warehouse Management," in *2022 International Joint Conference on Neural Networks (IJCNN)*, Padua, Italy: IEEE, July 2022, pp. 1–9. doi: 10.1109/IJCNN55064.2022.9891985.
- [7] L. N. Raju Mudunuri, "Maximizing Every Square Foot: AI Creates the Perfect Warehouse Flow," *FMDT Trans. Sustain. Comput. Syst.*, vol. 2, no. 2, pp. 64–73, June 2024, doi: 10.69888/FTSCS.2024.000198.
- [8] A. J. Kaplan, "A stock redistribution model," *Nav. Res. Logist. Q.*, vol. 20, no. 2, pp. 231–239, June 1973, doi: 10.1002/nav.3800200203.
- [9] Ö. Özer and H. Xiong, "Stock positioning and performance estimation for distribution systems with service constraints," *IIE Trans.*, vol. 40, no. 12, pp. 1141–1157, Oct. 2008, doi: 10.1080/07408170802322960.
- [10] F. Nikkhoo, A. H. Kashan, E. Nikbakhsh, and B. Ostadi, "A multi-warehouse multi-period order picking system: A benders decomposition approach," Apr. 10, 2024, *In Review*. doi: 10.21203/rs.3.rs-4196126/v1.
- [11] S. Kumabe, S. Shiroshita, T. Hayashi, and S. Maruyama, "Learning General Inventory Management Policy for Large Supply Chain Network," 2022, *arXiv*. doi: 10.48550/ARXIV.2204.13378.
- [12] I. Nurkasanah, "Reinforcement Learning Approach for Efficient Inventory Policy in Multi-Echelon Supply Chain Under Various Assumptions and Constraints," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 7, no. 2, p. 138, Oct. 2021, doi: 10.20473/jisebi.7.2.138-148.
- [13] H. Talebi, M. Khalaj, D. Jafari, P. Mousavi Ahranjani, and A. Hossein Kamali Dolatabadi, "Robust Optimization for Supply Chain with Routing Problem: A Learning-based Approach," *RAIRO - Oper. Res.*, Dec. 2024, doi: 10.1051/ro/2024230.
- [14] S. Barat *et al.*, "Reinforcement Learning of Supply Chain Control Policy Using Closed Loop Multi-agent Simulation," in *Multi-Agent-Based Simulation XX*, vol. 12025, M. Paolucci, J. S. Sichman, and H. Verhagen, Eds., in *Lecture Notes in Computer Science*, vol. 12025, Cham: Springer International Publishing, 2020, pp. 26–38. doi: 10.1007/978-3-030-60843-9_3.
- [15] S. S. Bhattathiri, M. E. Kuhl, A. Tondwalkar, and A. Kwasinski, "Simulation Analysis of a Reinforcement-Learning-Based Warehouse Dispatching Method Considering due Date and Travel Distance," in *2023 Winter Simulation Conference (WSC)*, San Antonio, TX, USA: IEEE, Dec. 2023, pp. 1830–1841. doi: 10.1109/WSC60868.2023.10407952.
- [16] J. Z. Jianjun Zhou, "Optimization Algorithm of Intelligent Warehouse Management System Based on Reinforcement Learning," *J. Electr. Syst.*, vol. 20, no. 1, pp. 219–231, Jan. 2024, doi: 10.52783/jes.678.
- [17] T. Gruchmann, J. Eiten, G. De La Torre, and A. Melkonyan, "Sustainable Logistics and Transportation Systems: Integrating Optimization and Simulation Analysis to Enhance Strategic Supply Chain Decision-Making," in *Innovative Logistics Services and Sustainable Lifestyles*, A. Melkonyan and K. Krumme, Eds., Cham: Springer International Publishing, 2019, pp. 265–279. doi: 10.1007/978-3-319-98467-4_12.
- [18] K. Chicaiza Moncayo, A. Gomez Sanchez, P. Ruiz Anton, and L. Cevallos-Torres, "Modelo de simulación para la optimización del inventario de una distribuidora, basado en Simulación Monte Carlo y Algoritmo Metaheurístico Genético," *Ecuadorian Sci. J.*, vol. 3, no. 2, pp. 33–38, Sept. 2019, doi: 10.46480/esj.3.2.32.
- [19] A. Gokhale, C. Trasikar, A. Shah, A. Hegde, and S. R. Naik, "A Reinforcement Learning Approach to Inventory Management," in *Advances in Artificial Intelligence and Data Engineering*, vol. 1133, N. N. Chiplunkar and T. Fukao, Eds., in *Advances in Intelligent Systems and Computing*, vol. 1133, Singapore: Springer Nature Singapore, 2021, pp. 281–297. doi: 10.1007/978-981-15-3514-7_23.
- [20] N. N. Sultana, H. Meisheri, V. Baniwal, S. Nath, B. Ravindran, and H. Khadilkar, "Reinforcement Learning for Multi-Product Multi-Node Inventory Management in Supply Chains," 2020, *arXiv*. doi: 10.48550/ARXIV.2006.04037.
- [21] V. Methuku, "Optimizing Drug Distribution Using Reinforcement Learning in Pharmaceutical Logistics," Mar. 10, 2025, *Public Health and Healthcare*. doi: 10.20944/preprints202503.0638.v1.
- [22] Y. Yan, A. H. F. Chow, C. P. Ho, Y.-H. Kuo, Q. Wu, and C. Ying, "Reinforcement learning

for logistics and supply chain management: Methodologies, state of the art, and future opportunities,” *Transp. Res. Part E Logist. Transp. Rev.*, vol. 162, p. 102712, June 2022, doi: 10.1016/j.tre.2022.102712.