

AI-Driven Dynamic Routing Under Uncertainty: Integrating Machine Learning and Optimization for Cost-Efficient and Sustainable Logistics Networks

Abdulazeed Baruwa¹, Ngodoo Joy Sam-Bulya^{2*}, Olaoluwa Adesanya³

¹Amazon LLC, United States

²University of Mkar, Mkar, Benue State, Nigeria

³Finance and Strategy, Exa, San Francisco, California, United States

ABSTRACT

Vehicle routing sits at the center of logistics cost and logistics emissions, yet the routes that carriers execute are usually planned on deterministic estimates of travel time and demand that are wrong by the time a vehicle leaves the depot. Machine learning now predicts travel times, demand and request arrivals with useful accuracy, and optimization methods now solve large routing problems close to optimality, but the two are often joined only loosely: forecasts are reduced to point values before they reach the optimizer, and the optimizer is rerun from scratch when conditions change. This paper reviews the literature on dynamic and stochastic vehicle routing, learning-based routing, and green routing, and organizes it around the question of how prediction, uncertainty and optimization should be connected. It classifies uncertainty sources and solution paradigms and compares classical stochastic and robust models, approximate dynamic programming and rollout, end-to-end deep reinforcement learning, and hybrid learn-and-optimize methods. On this basis it proposes the Uncertainty-Aware Learn-and-Optimize Routing framework, which passes calibrated predictive distributions rather than point forecasts to the optimizer, represents uncertainty through scenarios, budgets and chance constraints, embeds cost, emissions and tail risk in one objective, and re-plans through an event-triggered rollout or learned policy. The paper closes with a research agenda on calibration, benchmark realism, electrified fleets and governance.

Keywords: dynamic vehicle routing; stochastic optimization; machine learning; reinforcement learning; green logistics; travel-time prediction; last-mile delivery

SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology (2026);

DOI: 10.18090/samriddhi.v18i03.06

INTRODUCTION

Routing decisions determine a large share of what logistics costs and what it emits. Every day, carriers and shippers decide which vehicle serves which customer, in what order, on which road, and at what time, and the quality of those decisions shows up directly in fuel, labor, vehicle hours, missed time windows and greenhouse gas emissions (Demir et al., 2014; Savelsbergh and Van Woensel, 2016). The vehicle routing problem (VRP) that formalizes these decisions is one of the oldest and most studied problems in operations research. It began with the truck dispatching problem of Dantzig and Ramser (1959) and the savings heuristic of Clarke and Wright (1964), and six decades of work have produced exact methods and metaheuristics that solve realistic instances with thousands of customers (Laporte, 2009; Braekers et al., 2016; Vidal, 2022). In practice, however, most routes are still planned on a deterministic picture of the world, in which each travel time and each demand is a single known number.

Corresponding Author: Ngodoo Joy Sam-Bulya, University of Mkar, Mkar, Benue State, Nigeria, joyngodoo@gmail.com

How to cite this article: Baruwa, A., Sam-Bulya, N.J., Adesanya, O. (2026). AI-Driven Dynamic Routing Under Uncertainty: Integrating Machine Learning and Optimization for Cost-Efficient and Sustainable Logistics Networks. *SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology*, 18(3), 82-95.

Source of support: Nil

Conflict of interest: None

That picture is wrong in ways that matter. Travel times depend on congestion, weather, incidents and the time of day; demands are revealed only on arrival; new orders arrive after vehicles have left; and vehicles, drivers and charging stations become unavailable without notice (Gendreau et al., 1996; Pillac et al., 2013; Psaraftis et al., 2016). A route that is optimal

for mean travel times can miss most of its time windows when variance is high, and a route that ignores late requests can require costly re-dispatching (Taş et al., 2013; Ulmer et al., 2019). Global disruptions such as the COVID-19 pandemic made the fragility of deterministic plans visible at scale, as demand shifted abruptly toward home delivery and lead times became erratic (Ivanov, 2020; Dolgui and Ivanov, 2021). Applied studies in emerging economies report the same pattern at the level of individual firms, where route and throughput decisions made on fixed assumptions translate directly into logistics cost and lost profitability (Tonoyan et al., 2025; Adesanya et al., 2024). In dense African cities the problem is sharper still: dispatch riders must route deliveries through congested and poorly connected networks, and movement restrictions during the pandemic showed how quickly trip patterns and mode choices can shift (Tiamiyu and Alaka, 2020; Ajayi et al., 2020).

Two bodies of work now offer tools for this problem. The first is machine learning (ML) for prediction, which has produced accurate estimators of travel time from GPS traces and road graphs (Wang et al., 2018; Derrow-Pinion et al., 2021) and probabilistic forecasters of demand across many related series (Salinas et al., 2020; Lim et al., 2021). The second is ML for optimization, in which neural networks learn to construct routes directly or to guide the search of classical solvers (Bengio et al., 2021; Kool et al., 2019; Boggyrbayeva et al., 2024). The two streams have grown largely in parallel. Forecasts are commonly collapsed to point values before they enter an optimizer, which discards the very uncertainty that makes routing hard, while many learned routing policies are trained and evaluated on synthetic instances with deterministic costs (Hildebrandt et al., 2023). The result is a gap between what prediction knows about uncertainty and what optimization uses.

This paper reviews the literature at the intersection of dynamic routing, uncertainty and AI, and proposes a framework to close that gap. Three questions guide it. The first asks which sources of uncertainty affect routing and which ML methods predict them. The second asks how the main solution paradigms, from classical stochastic and robust optimization to deep reinforcement learning (DRL), represent that uncertainty and at what computational cost. The third asks what architecture would connect calibrated prediction, uncertainty-aware optimization and dynamic re-planning while accounting for cost and emissions together. The paper contributes a taxonomy of uncertainty sources and predictive methods (Table 1), a comparison of solution paradigms (Table 2), the Uncertainty-Aware Learn-and-Optimize Routing (ULOR) framework with its formal components (Figure 1 and Eqs. 1 to 6), and a research agenda (Table 3). The design follows the view that learning should be aligned with the downstream decision rather than trained in isolation (Elmachtoub and Grigas, 2022; Bertsimas and Kallus, 2020).

Conceptual Foundations

This section sets out the three ideas on which the review rests: the classification of routing problems by how information is

revealed, the modeling of routing as a sequential decision problem, and the extension of the routing objective from cost to emissions. Each has a mature literature. Together they define what an AI-driven routing system must represent if it is to be both cost-efficient and sustainable under uncertainty (Psaraftis et al., 2016; Powell, 2019; Bektaş and Laporte, 2011). They also fix the vocabulary used in the rest of the paper: a plan is a set of routes committed before execution, a policy is a rule that maps the current state to a decision, and uncertainty refers to any routing input whose value is not known when the decision is made (Ulmer et al., 2020).

Static, stochastic and dynamic routing

Routing problems are commonly classified along two axes: whether all information is known before planning begins, and whether plans may be revised as information arrives (Pillac et al., 2013; Psaraftis et al., 2016). A static and deterministic VRP assumes that customers, demands and travel times are known in advance. A stochastic VRP treats some of these as random variables with known distributions but still fixes the plan before execution, often with recourse actions such as a return to the depot when a vehicle runs out of capacity (Gendreau et al., 1996; Oyola et al., 2018). Laporte et al. (1992) formulated the VRP with stochastic travel times using chance constraints and recourse, and Taş et al. (2013) extended it to soft time windows with explicit service costs for late arrival. Robust formulations take a different route, protecting a plan against every realization in an uncertainty set; Bertsimas and Sim (2004) introduced budgeted uncertainty sets that control conservatism, and Gounaris et al. (2013) applied robust optimization to the capacitated VRP under demand uncertainty.

A dynamic VRP allows the plan to change during execution as new orders, travel-time updates or disruptions arrive. Early work handled this by re-solving a deterministic problem at each event, but scenario-based methods showed that sampling possible futures and choosing decisions that perform well across them gives better results. Bent and Van Hentenryck (2004) proposed a multiple scenario approach for vehicle routing with stochastic customers, and Hvattum et al. (2006) developed a sample scenario hedging heuristic for a dynamic and stochastic VRP. The trend across these studies is clear: the more information that is revealed during execution, the more value lies in policies that anticipate future information rather than in plans that react to it (Ulmer et al., 2020). Reviews of VRP software and algorithms note that commercial tools lag this literature and still emphasize static planning with manual re-dispatch (Konstantakopoulos et al., 2022).

Routing as sequential decision making under uncertainty

The modern treatment of dynamic routing models it as a Markov decision process (MDP) in which the state records vehicle positions, loads, committed routes and outstanding

requests; a decision assigns or sequences customers; and exogenous information such as new requests or realized travel times drives the transition to the next state (Ulmer et al., 2020). Powell (2019) proposed a unified framework for such problems and classified solution approaches into four families of policies: policy function approximations, cost function approximations, value function approximations and direct lookahead approximations. This classification is useful for routing because it shows that apparently different methods, such as a tuned dispatching rule, a buffered deterministic optimizer, an approximate dynamic programming (ADP) scheme and a rollout algorithm, are all instances of one decision problem that differ in how they approximate the future (Sutton and Barto, 2018).

Rollout algorithms are the lookahead method most often used in dynamic routing. A rollout evaluates each candidate decision by simulating a base heuristic forward over sampled futures and choosing the decision with the lowest expected cost. Secomandi (2000) compared neuro-dynamic programming algorithms for the VRP with stochastic demands and found that a rollout policy performed well, and Goodson et al. (2017) provided a general rollout framework for finite-horizon stochastic dynamic programs with performance guarantees relative to the base policy. Offline-online ADP combines value functions learned offline with online lookahead, and Ulmer et al. (2019) showed that this combination improves same-day delivery with stochastic requests. These methods are the natural bridge between prediction and optimization, since the quality of a rollout depends on how well the sampled futures reflect real uncertainty, and that is exactly what ML predictors can supply (Hildebrandt et al., 2023).

From cost to sustainability in routing

Cost-efficiency and sustainability are linked in routing but not identical. Fuel consumption and emissions depend on distance, but also on speed, load, road gradient and acceleration, so the shortest or fastest route is not always the cleanest (Demir et al., 2014). Bektaş and Laporte (2011) introduced the pollution-routing problem, which minimizes a combination of fuel, emissions and driver costs with speed as a decision variable, and showed that the trade-off between labor cost and fuel cost changes optimal speeds and routes. Erdoğan and Miller-Hooks (2012) formulated the green VRP for alternative-fuel vehicles that must visit refueling stations, and Schneider et al. (2014) extended routing to electric vehicles with time windows and recharging stations. Reviews of green routing confirm that emissions-aware models have become common, but they are still mostly deterministic and static (Lin et al., 2014; Demir et al., 2014).

Uncertainty and sustainability interact. Congestion raises both travel time and emissions per kilometer, so a route that is robust to congestion is often cleaner as well. Late requests that trigger extra trips add distance and emissions, so anticipating them has an environmental as well as a financial benefit (Savelsbergh and Van Woensel, 2016). Reverse-

logistics flows and humanitarian distribution add further uncertainty and further emissions when they are handled reactively (Anene et al., 2025; Anene and Clement, 2022). Predictive monitoring of emissions and infrastructure risk offers a way to measure these effects at city scale rather than per vehicle (Okojie et al., 2023). Reviews of energy efficiency in transportation logistics reach the same conclusion from the infrastructure side, arguing that routing, fleet and network decisions should be optimized together for sustainability (Opara et al., 2024). Electric fleets add a new uncertainty in energy consumption and charger availability (Schneider et al., 2014). An AI-driven routing system should therefore carry emissions and energy through the same probabilistic pipeline as time and cost, rather than computing them after the route has been chosen, a principle adopted in the framework of Section 5.

REVIEW METHOD

The review uses a structured narrative design, which suits a field that spans operations research, computer science and transportation engineering and in which the aim is conceptual integration rather than statistical pooling (Braekers et al., 2016; Konstantakopoulos et al., 2022). The corpus was assembled in four passes. The first identified anchor reviews on dynamic, stochastic and green routing and on ML for combinatorial optimization (Pillac et al., 2013; Psaraftis et al., 2016; Oyola et al., 2018; Lin et al., 2014; Bengio et al., 2021; Boggyrbayeva et al., 2024). The second traced citations backward to the foundational models and forward to methodological studies. The third added work on travel-time and demand prediction and on uncertainty quantification that routing systems depend on. The fourth added applied frameworks on logistics analytics, procurement and supply chain operations from emerging economies, which describe the operating conditions under which many routing systems must work (Okonkwo et al., 2024a; Aifuwa et al., 2020).

A study was included if it was peer reviewed or published in reviewed conference proceedings, addressed routing, the prediction of a routing input, or a method used for routing under uncertainty, and could be verified through its digital object identifier or its proceedings record. Preprints were excluded. Each study was coded on the source of uncertainty it addresses, the ML method used, the optimization paradigm, whether decisions are revised during execution, and whether emissions or energy enter the objective. The coded corpus underpins Tables 1 and 2. The design follows recent reviews of AI in operations and supply chain management that combine a structured corpus with an explicit analytical framework (Pournader et al., 2021; Rolf et al., 2023).

Machine Learning and Optimization for Routing Under Uncertainty

This section reviews the methods that an AI-driven routing system can draw on. It first considers the prediction of uncertain routing inputs, then the representation of



uncertainty in optimization, then learning-based routing policies, and finally hybrid methods that combine learning with classical search. Table 1 summarizes the main sources of uncertainty and the predictive methods that address them, and Table 2 compares the solution paradigms. The organization follows the methodological distinction of Bengio et al. (2021) between learning end to end, learning to configure an optimizer, and learning alongside an optimizer, and the policy classes of Powell (2019). Throughout, the emphasis is on what each method does with uncertainty, since that property, more than raw solution quality on deterministic benchmarks, decides whether a method is useful for dynamic routing in operating networks (Hildebrandt et al., 2023).

Predicting uncertain routing inputs

Travel time is the input on which most routing decisions depend and the one with the richest data. Wang et al. (2018) proposed DeepTTE, an end-to-end deep model that estimates the travel time of a path from raw GPS points by combining spatial convolution with recurrent layers, and showed that it outperformed segment-sum baselines. Derrow-Pinion et al. (2021) described graph neural networks deployed in Google Maps for estimated time of arrival, where road segments are grouped into supersegments and the model learns spatiotemporal interactions across them; the deployment reduced the frequency of poor arrival-time estimates in several major cities. These results show that learned travel-time models can capture interactions that historical averages miss, such as the propagation of congestion across neighboring roads. Recurrent and

gradient-boosted models remain strong baselines when data are limited (Hochreiter and Schmidhuber, 1997; Chen and Guestrin, 2016; Breiman, 2001).

Demand and request prediction draws on the forecasting literature. Probabilistic models such as DeepAR and the Temporal Fusion Transformer produce full predictive distributions across many related series and are well suited to customer-level demand (Salinas et al., 2020; Lim et al., 2021), and the M5 competition confirmed that tree ensembles and global neural models dominate large retail forecasting tasks (Makridakis et al., 2022). Applied studies report similar gains in supply chain demand forecasting, e-commerce procurement and predictive procurement planning, where better forecasts reduce both inventory inefficiency and the number of emergency deliveries (Aifuwa et al., 2020; Akin-Oluyomi et al., 2023; Okonkwo et al., 2024a). Machine learning for demand and revenue forecasting, predictive cost optimization, supply chain analytics for forecasting and bottleneck reduction, and integrated AI and predictive analytics frameworks for supply chain management extend these findings (Tonoyan et al., 2021, 2022; Akin-Oluyomi and Akhigbe, 2023b; Rukh et al., 2024), and AI-based predictive maintenance for e-commerce systems shows how the same models can anticipate failures in the digital systems that dispatch depends on (Mayo et al., 2023). The critical point for routing is that the optimizer needs the uncertainty of the forecast, not only its mean. Conformal prediction supplies distribution-free intervals with guaranteed coverage around any learned predictor (Vovk et al., 2005; Angelopoulos and Bates, 2023), which makes it a natural bridge between an ML travel-time or demand model and a chance-constrained or robust route plan.

Table 1: Sources of uncertainty in vehicle routing and the machine learning methods used to predict them

<i>Uncertainty source</i>	<i>Effect on routing</i>	<i>Predictive methods</i>	<i>Representative studies</i>	<i>Output passed to optimizer</i>
Travel time	Late arrivals, missed time windows, overtime	Deep sequence models on GPS traces; graph neural networks on road networks; gradient-boosted trees	Wang et al. (2018); Derrow-Pinion et al. (2021); Chen and Guestrin (2016)	Distribution or interval per arc and departure time
Customer demand	Capacity violations, recourse trips	Probabilistic deep forecasters; tree ensembles	Salinas et al. (2020); Lim et al. (2021); Makridakis et al. (2022)	Quantiles or scenarios per customer
Request arrival	Unserved orders, re-dispatch cost	Point-process and learned arrival-rate models	Ulmer et al. (2019); Bent and Van Hentenryck (2004)	Sampled future requests for lookahead
Energy and emissions	Range violations, unplanned charging, excess CO2	Load, speed and gradient models; learned consumption estimators	Bektaş and Laporte (2011); Schneider et al. (2014)	Expected consumption with interval
Disruption and supply	Unavailable vehicles, stock-outs, closures	Classification of disruption risk; anomaly detection	Brinrup et al. (2020); Ivanov (2020)	Scenario weights and contingency triggers

Representing uncertainty in route optimization

Classical optimization offers three ways to use predictive uncertainty. Stochastic programming with recourse optimizes the expected cost of a first-stage plan plus the cost of corrective actions, and scenario sampling makes it tractable for realistic instances (Gendreau et al., 1996; Oyola et al., 2018). Chance-constrained programming, introduced by Charnes and Cooper (1959), requires that constraints such as time windows hold with a specified probability, and Laporte et al. (1992) used it for stochastic travel times. Robust optimization protects a plan against all realizations within an uncertainty set; the budgeted sets of Bertsimas and Sim (2004) let a planner choose how many parameters may deviate at once, and Gounaris et al. (2013) showed that robust capacitated routing under demand uncertainty can be solved at useful scales. Tail-risk measures offer a fourth option: conditional value-at-risk (CVaR) penalizes the average of the worst outcomes and remains convex, so it can be optimized with sampled scenarios (Rockafellar and Uryasev, 2000).

These formulations are only as good as their inputs. A chance constraint built on a badly calibrated travel-time distribution provides false assurance, and a robust set that is too wide produces plans that are safe but expensive, while one that is too narrow offers little protection (Bertsimas and Sim, 2004; Taş et al., 2013). This is where ML adds value beyond point accuracy: learned models can produce context-specific distributions that vary by time of day, weather and road segment, and conformal calibration can size uncertainty sets so that their empirical coverage matches the target (Angelopoulos and Bates, 2023). The data-driven optimization literature supports this coupling, showing that decisions improve when learned conditional distributions are fed directly into the optimization problem rather than summarized first (Bertsimas and Kallus, 2020; Ban and Rudin, 2019).

Learning-based routing policies

The second use of ML is to learn the routing policy itself. Vinyals et al. (2015) introduced pointer networks, which output permutations of their inputs and were first applied to the traveling salesman problem. Nazari et al. (2018) trained a recurrent decoder with attention through reinforcement learning to solve the VRP, and noted that the policy could react to changes in demand without retraining. Kool et al. (2019) replaced the recurrent encoder with a transformer-style attention model trained with the REINFORCE gradient estimator and a greedy rollout baseline, obtaining near-optimal solutions on routing instances with up to 100 customers in a fraction of a second (Williams, 1992). Kwon et al. (2020) improved training by exploiting the symmetry of routing solutions through multiple starting points, and value-based methods descended from deep Q-networks have also been applied to routing and inventory problems (Mnih et al., 2015; Oroojlooyjadid et al., 2022). Reviews of these methods agree that they are fast at inference and flexible, but that most

results are on synthetic Euclidean instances of modest size (Mazyavkina et al., 2021; Bogrybayeva et al., 2024).

For dynamic and stochastic routing, the promise of learned policies is that they can make decisions in milliseconds as events arrive. Hildebrandt et al. (2023) reviewed the opportunities for reinforcement learning in stochastic dynamic vehicle routing and argued that RL is most useful where the state is complex and decisions must be made quickly, but that it should be combined with the structure of routing problems rather than replace it. Three limits recur in the learned-policy literature. The first is generalization: policies trained on one instance distribution often degrade when customer locations, fleet sizes or travel-time patterns change. The second is constraint handling: hard time windows, driver rules and vehicle capacities are difficult to guarantee through a learned policy alone. The third is evaluation: many studies use deterministic costs, which hides the very uncertainty the method is meant to handle (Gijbrecchts et al., 2022; Rolf et al., 2023). Warehouse and fleet applications in emerging markets show the same pattern, with learned path-planning and slotting models most effective when paired with explicit operational constraints (Akanbi and Sunday, 2024). A systematic review of autonomous mobile robots in micro-fulfillment centers reaches the same conclusion for fleet coordination inside the warehouse (Akanbi and Sunday, 2025).

Hybrid learn-and-optimize methods

The most effective current approaches combine learning with classical search. Metaheuristics such as adaptive large neighborhood search (ALNS) and hybrid genetic search (HGS) already achieve near-optimal solutions on large instances (Ropke and Pisinger, 2006; Vidal, 2022), and ML can improve them by learning which operators to apply, which subproblems to solve, or how to initialize the search. Li et al. (2021) trained a model to choose which regions of a large routing problem to delegate to a subsolver, reducing solution time for instances with thousands of customers while matching the quality of the full solver. Bengio et al. (2021) describe this category as ML alongside optimization and argue that it is the most promising for real deployment because the optimizer retains feasibility guarantees while learning improves efficiency.

The same logic applies to decision-focused learning. Elmachoub and Grigas (2022) showed that a predictive model trained against the loss of the downstream decision can outperform one trained on prediction error, and Bertsimas and Kallus (2020) showed how prescriptive analytics can weight historical data by its relevance to the current decision. In routing, this means that a travel-time model should be trained to minimize the cost of the routes it produces, not only the squared error of its estimates. Enterprise-scale applications add further requirements. Integrated IT and procurement systems, end-to-end visibility platforms and digital supply chain cost models supply



Table 2: Comparison of solution paradigms for dynamic vehicle routing under uncertainty

Paradigm	How uncertainty is represented	Response time to events	Feasibility guarantee	Main limitation	Representative studies
Deterministic re-optimization	Point estimates, re-solved at each event	Seconds to minutes	Strong	Ignores future information	Clarke and Wright (1964); Vidal (2022)
Stochastic programming with recourse	Known distributions, scenarios	Minutes	Strong for first stage	Distributions often assumed, not learned	Gendreau et al. (1996); Laporte et al. (1992)
Robust and chance-constrained	Uncertainty sets or probability targets	Minutes	Strong within the set	Conservatism depends on set calibration	Bertsimas and Sim (2004); Gounaris et al. (2013)
Scenario-based and rollout lookahead	Sampled futures, base heuristic	Seconds to minutes	Inherited from base heuristic	Quality depends on sampled futures	Bent and Van Hentenryck (2004); Goodson et al. (2017)
Approximate dynamic programming	Learned value functions	Milliseconds to seconds	Weak unless constrained	Feature design, training cost	Ulmer et al. (2019); Secomandi (2000)
End-to-end deep RL	Implicit in training distribution	Milliseconds	Weak	Generalization and constraints	Kool et al. (2019); Nazari et al. (2018)
Hybrid learn-and-optimize	Learned distributions feed solver; learned operators	Seconds	Strong (solver retains control)	Integration complexity	Li et al. (2021); Elmachtoub and Grigas (2022)

the data and execution channels through which routing decisions act (Okonkwo et al., 2024b, 2025; Nnabueze et al., 2021), and agentic orchestration of multi-enterprise supply chain data needs explicit process guardrails (Ogundairo, 2025). Workflow redesign for throughput, Lean Six Sigma in distribution networks, workforce optimization under staffing volatility, and leadership models that align logistics decisions with enterprise strategy determine whether an improved route can actually be executed (Nwangwu et al., 2024; Obukadata et al., 2021; Falemi et al., 2023; Akin-Oluyomi and Akhigbe, 2023a), and blockchain-based logistics records offer a verifiable log of executed deliveries (Ekwunife et al., 2024). Table 2 compares the paradigms reviewed in this section on the properties that matter for dynamic routing under uncertainty.

THE UNCERTAINTY-AWARE LEARN-AND-OPTIMIZE ROUTING FRAMEWORK

The review shows that the pieces of an uncertainty-aware routing system already exist but are rarely connected. Accurate travel-time and demand predictors, formulations that can use distributions, fast metaheuristics, and lookahead policies have each been developed and tested. What is missing is an architecture that passes calibrated uncertainty from the predictor to the optimizer, keeps cost and emissions in one objective, and revises plans as events arrive without discarding what the previous plan knew. The ULOR framework supplies this architecture. It follows the policy classification of Powell (2019) and the ML-alongside-optimization principle of Bengio et al. (2021), and it treats the governance of learned

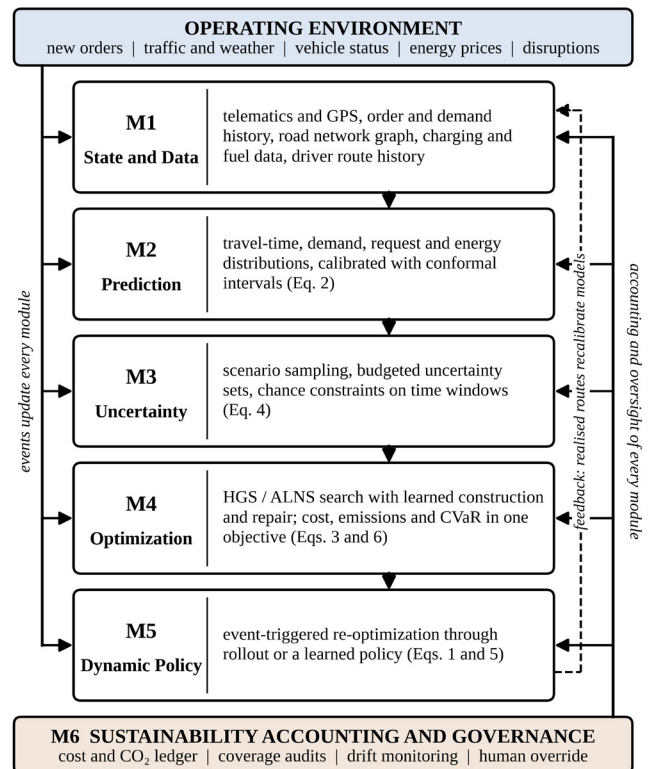


Figure 1: The Uncertainty-Aware Learn-and-Optimize Routing (ULOR) framework. Events from the operating environment update every module; calibrated predictions pass to the optimizer as distributions rather than point values; the dynamic policy re-plans on events; and executed routes return to the state and prediction modules for recalibration under a sustainability and governance layer

models as part of the system rather than an external review (NIST, 2023).

Architecture

The ULOR framework is shown in Figure 1. The state and data module (M1) maintains the current state of the network, including vehicle positions and loads, outstanding and committed orders, the road graph, energy and fuel data, and historical routes executed by drivers. The probabilistic prediction module (M2) estimates distributions of travel time, demand, request arrival and energy use, and calibrates them with conformal intervals. The uncertainty representation module (M3) converts these distributions into the form the optimizer needs, whether sampled scenarios, budgeted uncertainty sets or chance constraints. The optimization core (M4) searches for routes with a state-of-the-art metaheuristic whose construction and repair steps may be guided by learned models, and it scores routes on expected cost, emissions and tail risk together. The dynamic policy module (M5) decides when and how to re-plan as new events arrive, using a rollout over sampled futures or a learned policy. The sustainability accounting and governance module (M6) records cost and emissions for every executed route, audits the coverage of predictive intervals, monitors drift, and preserves human override.

Several design choices respond directly to the gaps identified in Section 4. Passing distributions rather than point forecasts from M2 to M3 answers the finding that most pipelines discard predictive uncertainty (Hildebrandt et al., 2023). Keeping a classical metaheuristic at the center of M4 retains feasibility for hard constraints, which answers the weakness of end-to-end learned policies (Vidal, 2022; Bogrybayeva et al., 2024). Placing emissions and energy in M2 and M4 rather than in after-the-fact reporting answers the finding that green routing models are mostly static (Demir et al., 2014). Using historical routes executed by drivers in M1 reflects the evidence from the Amazon Last Mile Routing Research Challenge that experienced drivers encode local knowledge about access, parking and sequencing that distance-based objectives miss (Merchán et al., 2024).

Formal components

The dynamic routing problem is modeled as an MDP in the sense of Ulmer et al. (2020). At decision epoch k the system is in state S_k , a policy π selects a decision $x_k = X^\pi(S_k)$, and exogenous information W_{k+1} (new requests, realized travel times, disruptions) moves the system to S_{k+1} . The objective is to find the policy that minimizes expected total cost over the horizon, starting from the initial state S_0 :

$$\min_{\pi \in \Pi} \mathbb{E} \left[\sum_{k=0}^K C(S_k, X^\pi(S_k)) \right] \quad (1)$$

The prediction module supplies, for each arc (i, j) and departure time t , a point estimate $\hat{\tau}_{ij}(t)$ from a learned

model and a conformal interval built from calibration residuals on held-out trips (Angelopoulos and Bates, 2023):

$$C_{ij}(t) = [\hat{\tau}_{ij}(t) - \hat{q}_{1-\alpha}, \hat{\tau}_{ij}(t) + \hat{q}_{1-\alpha}] \quad (2)$$

where $\hat{q}_{1-\alpha}$ is the empirical $(1-\alpha)$ quantile of absolute residuals, adjusted for sample size, so that the interval covers the realized travel time with probability of at least $1-\alpha$ under exchangeability (Vovk et al., 2005). The per-decision cost combines distance, time and emissions:

$$C(S_k, x_k) = \sum_{(i,j) \in x_k} (c_d d_{ij} + c_t \tau_{ij} + c_e e_{ij}(v_{ij}, q_{ij})) \quad (3)$$

where d_{ij} is distance, τ_{ij} travel time, e_{ij} the emissions on the arc as a function of speed v_{ij} and load q_{ij} in the manner of the pollution-routing problem, and c_d , c_t and c_e are unit costs, with c_e representing an internal or regulatory carbon price (Bektaş and Laporte, 2011). Service reliability enters as a chance constraint on each customer's arrival time A_i relative to the end of its time window l_i , for every customer i in the customer set $\mathcal{V}$:

$$\Pr(A_i \leq l_i) \geq 1 - \beta_i, \forall i \in \mathcal{V} \quad (4)$$

which is enforced through scenarios drawn from the calibrated predictive distributions of Eq. (2) (Charnes and Cooper, 1959; Laporte et al., 1992).

Dynamic policy and risk-aware objective

Re-planning in M5 uses a rollout policy that evaluates candidate decisions by simulating a base heuristic H , typically the metaheuristic of M4 run at reduced effort, over N sampled futures (Goodson et al., 2017; Secomandi, 2000):

$$X^{\text{roll}}(S_k) = \arg \min_{x \in \mathcal{X}(S_k)} \left\{ C(S_k, x) + \frac{1}{N} \sum_{n=1}^N V^H(S_{k+1}^{(n)}) \right\} \quad (5)$$

where $V^H(S_{k+1}^{(n)})$ is the cost-to-go of the base heuristic from the state reached under the n th sampled realization. A learned value function can replace or warm-start the simulation where response time is critical, following the offline-online approach of Ulmer et al. (2019). Because operators care about bad days as well as average days, the optimization core scores each plan on expected cost and on the tail of lateness and overtime losses L :

$$\min_x \mathbb{E} [C(x, \xi)] + \lambda \text{CVaR}_\gamma(L(x, \xi)) \quad (6)$$

where ξ is the joint scenario of travel times, demands and requests and $\lambda \geq 0$ sets the operator's risk appetite (Rockafellar and Uryasev, 2000). Setting $\lambda=0$ recovers expected-cost routing; increasing it produces plans that give up a little average efficiency to avoid the worst service



failures. In practice the scenario set used for Eq. (6) is the same set used for the chance constraints in Eq. (4) and the rollout in Eq. (5), so the framework needs a single well-calibrated scenario generator rather than several inconsistent ones. The weight on emissions through c_e and the weight on tail risk through λ are policy choices that should be set jointly by operations, finance and sustainability functions (Walawalkar et al., 2026).

DISCUSSION

The discussion considers the implications of the review and the framework for practice and research. It addresses four themes: the calibration and evaluation of learned components, the realism of benchmarks and the transfer of methods to operating conditions, the electrification of fleets and the measurement of sustainability benefits, and the governance and research agenda that follow. Each theme is grounded in the reviewed evidence and in current standards for the management of AI risk (NIST, 2023; Rudin, 2019). The themes are ordered from the technical to the organizational, because the value of better prediction and optimization is realized only when operators can trust, verify and govern the systems that deliver it (Konstantakopoulos et al., 2022).

Calibration and evaluation

The first implication is that routing systems should evaluate predictors by the decisions they support and by the coverage of their uncertainty estimates, not only by mean error. A travel-time model with low mean absolute error can still produce poor routes if its errors are concentrated on the arcs and times that matter most, and a chance constraint built on overconfident intervals will be violated more often than planned (Taş et al., 2013; Elmachoub and Grigas, 2022). Conformal calibration provides a check that can be run continuously: the share of realized travel times that fall within their predicted intervals should match the target, and a persistent gap signals drift that requires retraining (Angelopoulos and Bates, 2023). Explainability also matters for adoption, since dispatchers and drivers are more likely to accept a route when the reasons for a deviation from habit are visible, and post-hoc explanation methods can make the drivers of a travel-time prediction transparent (Lundberg and Lee, 2017; Rudin, 2019).

Evaluation of learned routing policies needs particular care. Many reported results compare learned policies with classical heuristics on small synthetic instances with deterministic costs, where the advantage of fast inference is visible but the cost of constraint violations and poor generalization is not (Mazyavkina et al., 2021; Bogrybayeva et al., 2024). Fair evaluation requires stochastic simulation with realistic travel-time and demand distributions, strong classical baselines such as HGS and ALNS, and reporting of feasibility rates and tail outcomes as well as average cost (Vidal, 2022; Ropke and Pisinger, 2006). Automated evaluation and release-gating practices from model governance can be

adapted to routing models, so that a new policy is promoted only when it passes a fixed test suite that includes adverse scenarios (Nwakamma et al., 2024).

Benchmark realism and transfer to operations

The second implication concerns the gap between academic instances and operating networks. Real delivery networks have asymmetric and time-dependent travel times, access restrictions, parking constraints, heterogeneous vehicles and driver preferences that standard benchmarks omit (Braekers et al., 2016; Savelsbergh and Van Woensel, 2016). The Amazon Last Mile Routing Research Challenge made a large set of real routes available and asked participants to predict the sequences that experienced drivers actually followed, which showed that learning from executed routes can capture tacit knowledge that a pure distance objective misses (Merchán et al., 2024). Such datasets are valuable because they let researchers test whether learned components improve on classical methods under realistic conditions rather than on uniform random points.

Transfer is harder still in emerging economies, where road data are sparse, addresses are informal, demand is volatile and the digital infrastructure that routing systems assume is incomplete. Applied studies from these settings show that the binding constraints are often data visibility, procurement reliability and system integration rather than algorithmic sophistication (Nnabueze et al., 2021; Okonkwo et al., 2024b; Okonkwo et al., 2025). Network analyses of intercity corridors in Nigeria show that crash clusters, emergency coverage gaps and poor access over unpaved rural roads are concentrated on specific segments, which argues for routing models that treat link reliability and safety as uncertain, location-specific inputs rather than uniform averages (Alaka et al., 2021; Tihamiyu et al., 2021; Alaka et al., 2020). Vendor management and lean practices determine whether the upstream supply that feeds a route arrives on time, so routing gains depend on procurement discipline as well (Wedraogo et al., 2025; Ike et al., 2022). Localized supply chains and cold chain distribution for temperature-sensitive goods add service constraints that routing models in these settings must respect (Anene and Clement, 2024; Aneke and Adesemoye, 2025), and GIS, drone and LiDAR programs for utility and corridor management show how the road and infrastructure data that routing requires can be collected where official maps are incomplete (Dagodzo and Ahiaekwe Patrick, 2021a, 2021b, 2023). Queueing and throughput analyses in beverage distribution and route-profitability models for delivery operations show that substantial gains are available from better use of existing data before advanced AI is introduced (Adesanya et al., 2024; Tonoyan et al., 2025). The ULOR framework is designed to degrade gracefully in such settings: when learned predictors are unavailable, M2 can fall back on historical distributions and M4 on classical search, and learned components can be added as data accumulate.

Table 3: Research agenda for AI-driven dynamic routing under uncertainty

<i>Research direction</i>	<i>Open question</i>	<i>Suggested approach</i>	<i>Anchoring studies</i>
Calibrated prediction for routing	Do calibrated intervals improve route reliability over point forecasts?	Conformal travel-time and demand models feeding chance constraints; coverage audits	Angelopoulos and Bates (2023); Laporte et al. (1992)
Decision-focused training	Should travel-time models be trained on route cost rather than prediction error?	Smart predict-then-optimize losses with routing solvers in the loop	Elmachtoub and Grigas (2022); Bertsimas and Kallus (2020)
Realistic benchmarks	Do learned policies retain their advantage on real, stochastic networks?	Evaluation on executed-route datasets with stochastic simulation and strong baselines	Merchán et al. (2024); Vidal (2022)
Hybrid scalability	How can learning scale classical search to very large dynamic instances?	Learned decomposition, operator selection and warm-started rollout	Li et al. (2021); Goodson et al. (2017)
Electrified fleets	How should energy and charger uncertainty enter dynamic routing?	Probabilistic energy models with state-of-charge chance constraints	Schneider et al. (2014); Erdoğan and Miller-Hooks (2012)
Governed deployment	How should learned routing components be validated and monitored?	Release gates, drift monitoring and fallback policies	NIST (2023); Hildebrandt et al. (2023)

Electrification and sustainability measurement

The third implication concerns fleet electrification and the measurement of environmental benefit. Electric vehicle routing adds uncertainty in energy consumption, which depends on load, speed, temperature and terrain, and in charger availability, which depends on the behavior of other users (Schneider et al., 2014). The same predictive and calibration machinery used for travel time can be applied to energy use, and chance constraints on state of charge can protect against stranded vehicles in the way Eq. (4) protects time windows. Alternative-fuel routing problems show that refueling decisions interact with route structure, so energy uncertainty cannot be treated as a separate layer added after routing (Erdoğan and Miller-Hooks, 2012; Lin et al., 2014).

Sustainability claims also need measurement. Many studies report emissions reductions computed from simplified consumption models under deterministic conditions, which may overstate benefits when real traffic and loads differ (Demir et al., 2014). The ULOR framework therefore records emissions for executed routes in M6 using the same consumption models as the optimizer and compares them with planned values, which gives a basis for verified reporting and for recalibrating the emissions model itself (Bektaş and Laporte, 2011). Where carbon pricing or reporting obligations apply, the unit cost in Eq. (3) links routing decisions directly to financial exposure, and scenario-based financial modeling can translate route-level emissions into organizational plans (Walawalkar et al., 2026).

Governance and research agenda

The fourth implication concerns governance and the research that remains. Learned routing components make decisions that affect drivers' working hours, customers' service and a firm's emissions, so their use should follow the functions of governing, mapping, measuring and managing AI risk (NIST, 2023). Practical safeguards include versioned models, documented training data, human override for dispatchers, monitoring of interval coverage and feasibility rates, and fallback to classical planning when a learned component fails (Rudin, 2019; Nwakamma et al., 2024). Reviews of algorithmic accountability and model risk management, responsible AI governance for automated decision systems, anti-discrimination compliance for automated decisions affecting workers and customers, and responsible AI auditing with human oversight offer concrete templates (Ominyi and Anichukwueze, 2023, 2024; Annan, 2025; Elebe and Anunagba, 2025). Continuous AI-enabled control, automated governance of multi-country delivery networks, standardized compliance controls for high-volume distribution networks, and performance dashboards with real-time KPI tracking give operators the means to monitor learned routing components in production (Adebayo et al., 2025; Ngonadi et al., 2025, 2023; Ezeako et al., 2024; Tonoyan et al., 2024). Table 3 sets out six research directions that follow from the review.

Across these directions, the most urgent need is for studies that evaluate complete pipelines, from raw data through prediction, uncertainty representation and optimization to executed routes, rather than individual components in isolation. Such studies would show where



the value of AI in routing actually lies, whether in better prediction, better search or better re-planning, and they would give practitioners the evidence to invest accordingly (Hildebrandt et al., 2023; Bogyrbayeva et al., 2024). Supply chain applications of reinforcement learning more broadly point to the same need for end-to-end evaluation under realistic uncertainty (Rolf et al., 2023; Toorajipour et al., 2021).

Conclusion

This review examined how artificial intelligence can make vehicle routing more cost-efficient and sustainable when travel times, demands, requests and energy use are uncertain. It found that the necessary components are well developed but poorly connected. Machine learning predicts routing inputs with increasing accuracy, optimization methods solve large routing problems close to optimality, and lookahead and learned policies can revise plans as events arrive, yet most pipelines reduce forecasts to point values, evaluate learned policies on deterministic instances, and treat emissions as an afterthought.

The Uncertainty-Aware Learn-and-Optimize Routing framework proposed here connects these components in six modules. It passes calibrated predictive distributions to the optimizer, represents uncertainty through scenarios, budgets and chance constraints, scores routes on expected cost, emissions and tail risk together, and re-plans through event-triggered rollout or learned policies while keeping a classical solver in control of feasibility. It treats sustainability accounting and model governance as parts of the system, so that environmental benefits are measured on executed routes and learned components are monitored and replaceable.

For logistics operators, the practical message is that the largest gains are likely to come from using the uncertainty that predictors already know about, rather than from replacing proven solvers with learned ones. For researchers, the priority is to evaluate complete pipelines on realistic data, to train predictors for the decisions they support, and to extend these methods to electrified fleets and to the data-scarce networks where much of the growth in logistics demand is taking place.

FUNDING

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

COMPETING INTERESTS

The authors declare no competing interests.

DATA AVAILABILITY

No new data were generated or analyzed in this study.

REFERENCES

- [1] Adebayo, A., Amadi, C., & Ijagbemi, A. O. (2025). From audit to continuous control: Implementation pathways, governance, and organizational readiness for AI-enabled compliance in critical infrastructure. *International Journal of Computer Science and Mathematical Theory*, 11(8), 88-134. <https://doi.org/10.56201/ijcsmt.vol.11.no8.2025.pg88.134>
- [2] Adesanya, O. S., Okafor, C. M., Dako, O. F., & Akinola, A. S. (2024). Value-chain automation in beverage logistics: Throughput, capacity, and cost avoidance via queueing models. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 10(4), 1112-1132. <https://doi.org/10.32628/CSEIT251134110>
- [3] Aifuwa, S. E., Oshoba, T. O., Ogbuefi, E., Ike, P. N., Nnabueze, S. B., & Olatunde-Thorpe, J. (2020). Predictive analytics models enhancing supply chain demand forecasting accuracy and reducing inventory management inefficiencies. *International Journal of Multidisciplinary Research and Growth Evaluation*, 1(3), 171-181. <https://doi.org/10.54660/IJMRGE.2020.1.3.171-181>
- [4] Ajayi, O., Alaka, O. B., & Tihamiyu, R. (2020). Essential trip making and mode substitution during COVID-19 movement restrictions in Nigerian cities. *Iconic Research and Engineering Journals*, 4(4), 442-465. <https://doi.org/10.64388/IREV414-1722673>
- [5] Akanbi, O., & Sunday, E. A. (2024). An integrated path-planning and slotting optimization model for AMR-enabled high-density warehousing. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 3226-3243. <https://doi.org/10.62225/2583049X.2024.4.6.6181>
- [6] Akanbi, O., & Sunday, E. A. (2025). A systematic review of AI-driven autonomous mobile robots (AMR) in scalable micro-fulfillment centers. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(6), 2363-2379. <https://doi.org/10.62225/2583049X.2025.5.6.6182>
- [7] Akin-Oluyomi, O. T., & Akhigbe, R. (2023a). A conceptual supply chain leadership model for aligning logistics decisions with enterprise strategy. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(1), 1725-1744. <https://doi.org/10.62225/2583049X.2023.3.1.5340>
- [8] Akin-Oluyomi, O. T., & Akhigbe, R. (2023b). A proposed supply chain analytics framework for improving forecasting accuracy and reducing bottlenecks. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(1), 1686-1703. <https://doi.org/10.62225/2583049X.2023.3.1.5338>
- [9] Akin-Oluyomi, O. T., Atima, M. E., Akinleye, O. K., & Okoruwa, P. O. (2023). Predictive analytics approaches for improving demand forecasting accuracy in e-commerce procurement systems. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6), 2173-2182. <https://doi.org/10.62225/2583049X.2023.3.6.5324>
- [10] Alaka, O. B., Ajayi, O., & Bolaji, O. (2020). Geospatial assessment of access to health facilities over unpaved rural road networks. *International Journal of Multidisciplinary Research and Growth Evaluation*, 1(5), 1199-1214. <https://doi.org/10.54660/IJMRGE.2020.1.5.1199-1214>
- [11] Alaka, O. B., Tihamiyu, R., & Bolaji, O. (2021). Spatial clustering of road traffic crash locations on intercity corridors using national road safety records. *International Journal of Multidisciplinary Research and Growth Evaluation*, 2(6), 1086-1101. <https://doi.org/10.54660/IJMRGE.2021.2.6.1086-1101>
- [12] Aneke, O. B., & Adesemoye, A. C. (2025). Advances in cold chain integrity and good distribution practice for temperature-sensitive pharmaceuticals. *International Journal of Health and Pharmaceutical Research*, 10(3), 99-137. <https://doi.org/10.56201/ijhpr.vol.10.no3.2025.pg99.137>

- [13] Anene, U. N., & Clement, T. (2022). A resilient logistics framework for humanitarian supply chains: Integrating predictive analytics, IoT, and localized distribution to strengthen emergency response systems. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 8(5), 398-424.
- [14] Anene, U. N., & Clement, T. (2024). Localized supply chain solutions for sustainable community development: A strategic model for economic revitalization and regional resilience. *International Journal of Scientific Research in Science and Technology*, 11(5), 736-776. <https://doi.org/10.32628/IJSRST52310378>
- [15] Anene, U. N., Obukadata, V. O., Uzundu, C. S., & Ibeh, A. V. (2025). Process optimization in high-volume reverse-logistics operations: A conceptual framework for category-level improvement. *Gulf Journal of Advance Business Research*, 3(12), 1670-1702. <https://doi.org/10.51594/gjabr.v3i12.221>
- [16] Angelopoulos, A. N., & Bates, S. (2023). Conformal prediction: A gentle introduction. *Foundations and Trends in Machine Learning*, 16(4), 494-591. <https://doi.org/10.1561/22000000101>
- [17] Annan, A. O. (2025). Automated decision-making and anti-discrimination compliance under U.S. law. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(2), 2522-2540. <https://doi.org/10.62225/2583049X.2025.5.2.5784>
- [18] Ban, G.-Y., & Rudin, C. (2019). The big data newsvendor: Practical insights from machine learning. *Operations Research*, 67(1), 90-108. <https://doi.org/10.1287/opre.2018.1757>
- [19] Bektaş, T., & Laporte, G. (2011). The pollution-routing problem. *Transportation Research Part B: Methodological*, 45(8), 1232-1250. <https://doi.org/10.1016/j.trb.2011.02.004>
- [20] Bengio, Y., Lodi, A., & Prouvost, A. (2021). Machine learning for combinatorial optimization: A methodological tour d'horizon. *European Journal of Operational Research*, 290(2), 405-421. <https://doi.org/10.1016/j.ejor.2020.07.063>
- [21] Bent, R. W., & Van Hentenryck, P. (2004). Scenario-based planning for partially dynamic vehicle routing with stochastic customers. *Operations Research*, 52(6), 977-987. <https://doi.org/10.1287/opre.1040.0124>
- [22] Bertsimas, D., & Kallus, N. (2020). From predictive to prescriptive analytics. *Management Science*, 66(3), 1025-1044. <https://doi.org/10.1287/mnsc.2018.3253>
- [23] Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations Research*, 52(1), 35-53. <https://doi.org/10.1287/opre.1030.0065>
- [24] Bogrybayeva, A., Meraliyev, M., Mustakhov, T., & Dauletbayev, B. (2024). Machine learning to solve vehicle routing problems: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 25(6), 4754-4772. <https://doi.org/10.1109/TITS.2023.3334976>
- [25] Braekers, K., Ramaekers, K., & Van Nieuwenhuyse, I. (2016). The vehicle routing problem: State of the art classification and review. *Computers & Industrial Engineering*, 99, 300-313. <https://doi.org/10.1016/j.cie.2015.12.007>
- [26] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- [27] Brintrup, A., Pak, J., Ratiney, D., Pearce, T., Wichmann, P., Woodall, P., & McFarlane, D. (2020). Supply chain data analytics for predicting supplier disruptions: A case study in complex asset manufacturing. *International Journal of Production Research*, 58(11), 3330-3341. <https://doi.org/10.1080/00207543.2019.1685705>
- [28] Charnes, A., & Cooper, W. W. (1959). Chance-constrained programming. *Management Science*, 6(1), 73-79. <https://doi.org/10.1287/mnsc.6.1.73>
- [29] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794). <https://doi.org/10.1145/2939672.2939785>
- [30] Clarke, G., & Wright, J. W. (1964). Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research*, 12(4), 568-581. <https://doi.org/10.1287/opre.12.4.568>
- [31] Dagodzo, D., & Ahiaeké Patrick, M. C. (2021a). A review of GIS applications in utility asset management and infrastructure planning. *Iconic Research and Engineering Journals*, 5(3), 468-492. <https://doi.org/10.64388/IREV513-1716085>
- [32] Dagodzo, D., & Ahiaeké Patrick, M. C. (2021b). An integrated framework for UAV, LiDAR, and GIS in infrastructure corridor management. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 7(5), 497-524. <https://doi.org/10.32628/CSEIT217566>
- [33] Dagodzo, D., & Ahiaeké Patrick, M. C. (2023). A framework for integrating drone operations into enterprise GIS systems for utility companies. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 9(4), 924-957. <https://doi.org/10.32628/CSEIT23564536>
- [34] Dantzig, G. B., & Ramser, J. H. (1959). The truck dispatching problem. *Management Science*, 6(1), 80-91. <https://doi.org/10.1287/mnsc.6.1.80>
- [35] Demir, E., Bektaş, T., & Laporte, G. (2014). A review of recent research on green road freight transportation. *European Journal of Operational Research*, 237(3), 775-793. <https://doi.org/10.1016/j.ejor.2013.12.033>
- [36] Derrow-Pinion, A., She, J., Wong, D., Lange, O., Hester, T., Perez, L., Nunkesser, M., Lee, S., Guo, X., Wiltshire, B., Battaglia, P. W., Gupta, V., Li, A., Xu, Z., Sanchez-Gonzalez, A., Li, Y., & Veličković, P. (2021). ETA prediction with graph neural networks in Google Maps. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management* (pp. 3767-3776). <https://doi.org/10.1145/3459637.3481916>
- [37] Dolgui, A., & Ivanov, D. (2021). Ripple effect and supply chain disruption management: New trends and research directions. *International Journal of Production Research*, 59(1), 102-109. <https://doi.org/10.1080/00207543.2021.1840148>
- [38] Ekwunife, D. I., Precious, O. T., Rasul, O. A., Akinlode, O. F., Nwokoro, T. O., & Ikpe, V. I. (2024). Using blockchain technology to maximize supply chain and logistics management in North America. *International Journal of Science and Research Archive*, 12(2), 854-863. <https://doi.org/10.30574/ijrsra.2024.12.2.1096>
- [39] Elebe, O., & Anunagba, C. O. (2025). Responsible artificial intelligence in IT auditing through evidence validation, human oversight, and regulatory compliance. *International Journal of Technology, Health and Sustainability*, 1(2), 310-328.
- [40] Elmachtoub, A. N., & Grigas, P. (2022). Smart "predict, then optimize". *Management Science*, 68(1), 9-26. <https://doi.org/10.1287/mnsc.2020.3922>
- [41] Erdoğan, S., & Miller-Hooks, E. (2012). A green vehicle routing problem. *Transportation Research Part E: Logistics and Transportation Review*, 48(1), 100-114. <https://doi.org/10.1016/j.tre.2011.08.001>
- [42] Ezeako, R., Ibeh, A. V., Alugwe, S., & Awulu, I. (2024). Performance dashboards and data-driven decision-making in operations management: A conceptual framework for leadership visibility. *International Journal of Advanced Multidisciplinary Research*



- and Studies, 4(6), 3520-3538. <https://doi.org/10.62225/2583049X.2024.4.6.6930>
- [43] Falemi, A., Akhigbe, R., & Akin-Oluyomi, O. T. (2023). A supply chain workforce optimization framework for addressing staffing volatility during rapid expansion cycles. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(1), 1745-1762. <https://doi.org/10.62225/2583049X.2023.3.1.5342>
- [44] Gendreau, M., Laporte, G., & Séguin, R. (1996). Stochastic vehicle routing. *European Journal of Operational Research*, 88(1), 3-12. [https://doi.org/10.1016/0377-2217\(95\)00050-X](https://doi.org/10.1016/0377-2217(95)00050-X)
- [45] Gijbrenchts, J., Boute, R. N., Van Mieghem, J. A., & Zhang, D. (2022). Can deep reinforcement learning improve inventory management? Performance on lost sales, dual-sourcing, and multi-echelon problems. *Manufacturing & Service Operations Management*, 24(3), 1349-1368. <https://doi.org/10.1287/msom.2021.1064>
- [46] Goodson, J. C., Thomas, B. W., & Ohlmann, J. W. (2017). A rollout algorithm framework for heuristic solutions to finite-horizon stochastic dynamic programs. *European Journal of Operational Research*, 258(1), 216-229. <https://doi.org/10.1016/j.ejor.2016.09.040>
- [47] Gounaris, C. E., Wiesemann, W., & Floudas, C. A. (2013). The robust capacitated vehicle routing problem under demand uncertainty. *Operations Research*, 61(3), 677-693. <https://doi.org/10.1287/opre.1120.1136>
- [48] Hildebrandt, F. D., Thomas, B. W., & Ulmer, M. W. (2023). Opportunities for reinforcement learning in stochastic dynamic vehicle routing. *Computers & Operations Research*, 150, 106071. <https://doi.org/10.1016/j.cor.2022.106071>
- [49] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [50] Hvattum, L. M., Løkketangen, A., & Laporte, G. (2006). Solving a dynamic and stochastic vehicle routing problem with a sample scenario hedging heuristic. *Transportation Science*, 40(4), 421-438. <https://doi.org/10.1287/trsc.1060.0166>
- [51] Ike, P. N., Ogbuefi, E., Nnabueze, S. B., Olatunde-Thorpe, J., Aifuwa, S. E., Oshoba, T. O., & Akokodari, D. (2022). Lean supply chain practices improving operational efficiency, reducing waste, and enhancing organizational competitiveness globally. *Journal of Frontiers in Multidisciplinary Research*, 3(2), 182-192. <https://doi.org/10.54660/IJFMR.2022.3.2.182-192>
- [52] Ivanov, D. (2020). Predicting the impacts of epidemic outbreaks on global supply chains: A simulation-based analysis on the coronavirus outbreak (COVID-19/SARS-CoV-2) case. *Transportation Research Part E: Logistics and Transportation Review*, 136, 101922. <https://doi.org/10.1016/j.tre.2020.101922>
- [53] Konstantakopoulos, G. D., Gayialis, S. P., & Kechagias, E. P. (2022). Vehicle routing problem and related algorithms for logistics distribution: A literature review and classification. *Operational Research*, 22(3), 2033-2062. <https://doi.org/10.1007/s12351-020-00600-7>
- [54] Kool, W., van Hoof, H., & Welling, M. (2019). Attention, learn to solve routing problems! In *International Conference on Learning Representations (ICLR 2019)*. <https://openreview.net/forum?id=ByxBFsRqYm>
- [55] Kwon, Y.-D., Choo, J., Kim, B., Yoon, I., Gwon, Y., & Min, S. (2020). POMO: Policy optimization with multiple optima for reinforcement learning. In *Advances in Neural Information Processing Systems 33* (pp. 21188-21198). <https://proceedings.neurips.cc/paper/2020/hash/f231f2107df69eab0a3862d50018a9b2-Abstract.html>
- [56] Laporte, G. (2009). Fifty years of vehicle routing. *Transportation Science*, 43(4), 408-416. <https://doi.org/10.1287/trsc.1090.0301>
- [57] Laporte, G., Louveaux, F., & Mercure, H. (1992). The vehicle routing problem with stochastic travel times. *Transportation Science*, 26(3), 161-170. <https://doi.org/10.1287/trsc.26.3.161>
- [58] Li, S., Yan, Z., & Wu, C. (2021). Learning to delegate for large-scale vehicle routing. In *Advances in Neural Information Processing Systems 34* (pp. 26198-26211). <https://proceedings.neurips.cc/paper/2021/hash/dc9fa5f217a1e57b8a6adeb065560b38-Abstract.html>
- [59] Lim, B., Arik, S. Ö., Loeff, N., & Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), 1748-1764. <https://doi.org/10.1016/j.ijforecast.2021.03.012>
- [60] Lin, C., Choy, K. L., Ho, G. T. S., Chung, S. H., & Lam, H. Y. (2014). Survey of green vehicle routing problem: Past and future trends. *Expert Systems with Applications*, 41(4), 1118-1138. <https://doi.org/10.1016/j.eswa.2013.07.107>
- [61] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems 30* (pp. 4765-4774). <https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>
- [62] Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2022). M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting*, 38(4), 1346-1364. <https://doi.org/10.1016/j.ijforecast.2021.11.013>
- [63] Mayo, W., Ogbale, J. I., Okoruwa, P. O., & Babatope, O. M. (2023). Designing an AI-predictive maintenance model for e-commerce systems using machine learning and cloud analytics. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6), 2454-2468. <https://doi.org/10.62225/2583049X.2023.3.6.5418>
- [64] Mazyavkina, N., Sviridov, S., Ivanov, S., & Burnaev, E. (2021). Reinforcement learning for combinatorial optimization: A survey. *Computers & Operations Research*, 134, 105400. <https://doi.org/10.1016/j.cor.2021.105400>
- [65] Merchán, D., Arora, J., Pachon, J., Konduri, K. C., Winkenbach, M., Parks, S., & Noszek, J. (2024). 2021 Amazon Last Mile Routing Research Challenge: Data set. *Transportation Science*, 58(1), 8-11. <https://doi.org/10.1287/trsc.2022.1173>
- [66] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529-533. <https://doi.org/10.1038/nature14236>
- [67] National Institute of Standards and Technology [NIST]. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)* (NIST AI 100-1). U.S. Department of Commerce. <https://doi.org/10.6028/NIST.AI.100-1>
- [68] Nazari, M., Oroojlooy, A., Snyder, L., & Takáč, M. (2018). Reinforcement learning for solving the vehicle routing problem. In *Advances in Neural Information Processing Systems 31* (pp. 9839-9849). <https://proceedings.neurips.cc/paper/2018/hash/9fb4651c05b2ed70fba5afe0b039a550-Abstract.html>
- [69] Ngonadi, U., Ominyi, M., Anichukwueze, C. C., & Jones, B. C. (2023). Advances in standardizing un-standardized compliance controls for high-volume non-traditional distribution networks. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6), 3178-3201. <https://doi.org/10.62225/2583049X.2023.3.6.3178>

- 49X.2023.3.6.6838
- [70] Ngonadi, U., Ominyi, M., Anichukwueze, C. C., & Jones, B. C. (2025). Advances in automated governance: Mitigating operational and systemic risks in multi-country delivery networks. *Gulf Journal of Advance Business Research*, 3(12), 1626-1669. <https://doi.org/10.51594/gjabr.v3i12.219>
- [71] Nnabueze, S. B., Ike, P. N., Olatunde-Thorpe, J., Aifuwa, S. E., Oshoba, T. O., Ogbuefi, E., & Akokodaripon, D. (2021). End-to-end visibility frameworks improving transparency, compliance, and traceability across complex global supply chain operations. *International Journal of Multidisciplinary Futuristic Development*, 2(2), 50-60. <https://doi.org/10.54660/IJMF.D.2021.2.2.50-60>
- [72] Nwakamma, S., Ojukwu, J., & Oyesiji, S. O. (2024). LLM-as-a-judge for automated model governance: A review of methods, biases, and release-gating practices. *World Journal of Innovation and Modern Technology*, 8(6), 185-216. <https://doi.org/10.56201/wjimt.v8.no6.2024.pg185.216>
- [73] Nwangwu, C. V., Onwumere, A., & Akano, O. A. (2024). Systems-level throughput optimization: A review of workflow redesign strategies to mitigate logistics bottlenecks. *International Journal of Multidisciplinary Research and Growth Evaluation*, 5(6), 1969-1980. <https://doi.org/10.54660/IJMRGE.2024.5.6.1969-1980>
- [74] Obukadata, V. O., Awulu, I., & Williams, I. (2021). A Lean Six Sigma approach to mitigating pharmacy distribution bottlenecks in developing healthcare settings. *International Journal of Multidisciplinary Research and Growth Evaluation*, 2(6), 1195-1211. <https://doi.org/10.54660/IJMRGE.2021.2.6.1195-1211>
- [75] Ogundairo, K. M. (2025). Hardening the autonomous value chain: Lean Six Sigma guardrails and multi-enterprise ETL architecture for agentic supply chain orchestration. *IJAR International Journal of Economics and Business Management*, 11(12), 441-478. <https://doi.org/10.56201/ijebm.vol.11.no12.2025.pg441.478>
- [76] Okojie, J. S., Ike, P. N., Idu, J. O. O., Nnabueze, S. B., Filani, O. M., & Ihwughwawwe, S. I. (2023). Predictive analytics models for monitoring smart city emissions and infrastructure risk in urban ESG planning. *International Journal of Multidisciplinary Futuristic Development*, 4(1), 45-57. <https://doi.org/10.54660/IJMF.D.2023.4.1.45-57>
- [77] Okonkwo, C. S., Agbabiaka, J., Mayo, W., & Okeke, O. T. (2024a). Model for predictive procurement planning to sustain operational uptime. *International Journal of Scientific Research in Humanities and Social Sciences*, 1(2), 909-928. <https://doi.org/10.32628/IJSRSSH24272>
- [78] Okonkwo, C. S., Agbabiaka, J., Mayo, W., & Okeke, O. T. (2024b). Review of digital supply chain models for cost control and operational continuity. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 2836-2846. <https://doi.org/10.62225/2583049X.2024.4.6.5496>
- [79] Okonkwo, C. S., Agbabiaka, J., Okeke, O. T., & Mayo, W. (2025). Framework for national-scale supply chain optimization through integrated IT and procurement systems. *Gulf Journal of Advance Business Research*, 3(12), 1610-1625. <https://doi.org/10.51594/gjabr.v3i12.189>
- [80] Ominyi, M., & Anichukwueze, C. C. (2023). A review of algorithmic accountability and model risk management approaches in financial services. *Journal of Accounting and Financial Management*, 9(12), 218-238. <https://doi.org/10.56201/jafm.v9.no12.2023.pg218.238>
- [81] Ominyi, M., & Anichukwueze, C. C. (2024). Conceptualizing responsible AI governance structures for automated decision systems in regulated industries. *International Journal of Social Sciences and Management Research*, 10(11), 403-430. <https://doi.org/10.56201/ijssmr.v10.no11.2024.pg.403.430>
- [82] Opara, I. S., Elumilade, R. A., Shittu, H., & Adeniji, I. O. (2024). A theoretical review of synergizing energy efficiency with transportation logistics optimization: Towards a sustainable U.S. infrastructure. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 2968-2975. <https://doi.org/10.62225/2583049X.2024.4.6.5738>
- [83] Oroojlooyjadid, A., Nazari, M. R., Snyder, L. V., & Takáč, M. (2022). A deep Q-network for the beer game: Deep reinforcement learning for inventory optimization. *Manufacturing & Service Operations Management*, 24(1), 285-304. <https://doi.org/10.1287/msom.2020.0939>
- [84] Oyola, J., Arntzen, H., & Woodruff, D. L. (2018). The stochastic vehicle routing problem, a literature review, part I: Models. *EURO Journal on Transportation and Logistics*, 7(3), 193-221. <https://doi.org/10.1007/s13676-016-0100-5>
- [85] Pillac, V., Gendreau, M., Guéret, C., & Medaglia, A. L. (2013). A review of dynamic vehicle routing problems. *European Journal of Operational Research*, 225(1), 1-11. <https://doi.org/10.1016/j.ejor.2012.08.015>
- [86] Pournader, M., Ghaderi, H., Hassanzadegan, A., & Fahimnia, B. (2021). Artificial intelligence applications in supply chain management. *International Journal of Production Economics*, 241, 108250. <https://doi.org/10.1016/j.ijpe.2021.108250>
- [87] Powell, W. B. (2019). A unified framework for stochastic optimization. *European Journal of Operational Research*, 275(3), 795-821. <https://doi.org/10.1016/j.ejor.2018.07.014>
- [88] Psaraftis, H. N., Wen, M., & Kontovas, C. A. (2016). Dynamic vehicle routing problems: Three decades and counting. *Networks*, 67(1), 3-31. <https://doi.org/10.1002/net.21628>
- [89] Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2(3), 21-41. <https://doi.org/10.21314/JOR.2000.038>
- [90] Rolf, B., Jackson, I., Müller, M., Lang, S., Reggelen, T., & Ivanov, D. (2023). A review on reinforcement learning algorithms and applications in supply chain management. *International Journal of Production Research*, 61(20), 7151-7179. <https://doi.org/10.1080/00207543.2022.2140221>
- [91] Ropke, S., & Pisinger, D. (2006). An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. *Transportation Science*, 40(4), 455-472. <https://doi.org/10.1287/trsc.1050.0135>
- [92] Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206-215. <https://doi.org/10.1038/s42256-019-0048-x>
- [93] Rukh, S., Seyi-Lande, O. B., & Oziri, S. T. (2024). An integrated framework for AI and predictive analytics in supply chain management. *International Journal of Scientific Research in Humanities and Social Sciences*, 1(1), 451-491. <https://doi.org/10.32628/IJSRSSH243671>
- [94] Salinas, D., Flunkert, V., Gasthaus, J., & Januschowski, T. (2020). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. *International Journal of Forecasting*, 36(3), 1181-1191. <https://doi.org/10.1016/j.ijforecast.2019.07.001>
- [95] Savelsbergh, M., & Van Woensel, T. (2016). 50th anniversary invited article: City logistics: Challenges and opportunities. *Transportation Science*, 50(2), 579-590. <https://doi.org/10.1287/trsc.2016.0675>



- [96] Schneider, M., Stenger, A., & Goeke, D. (2014). The electric vehicle-routing problem with time windows and recharging stations. *Transportation Science*, 48(4), 500-520. <https://doi.org/10.1287/trsc.2013.0490>
- [97] Secomandi, N. (2000). Comparing neuro-dynamic programming algorithms for the vehicle routing problem with stochastic demands. *Computers & Operations Research*, 27(11-12), 1201-1225. [https://doi.org/10.1016/S0305-0548\(99\)00146-X](https://doi.org/10.1016/S0305-0548(99)00146-X)
- [98] Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press. ISBN 978-0-262-03924-6
- [99] Taş, D., Dellaert, N., Van Woensel, T., & de Kok, A. G. (2013). Vehicle routing problem with stochastic travel times including soft time windows and service costs. *Computers & Operations Research*, 40(1), 214-224. <https://doi.org/10.1016/j.cor.2012.06.008>
- [100] Tiamiyu, R., & Alaka, O. B. (2020). Routing of motorcycle dispatch deliveries on congested and poorly connected urban networks in Lagos. *International Journal of Multidisciplinary Research and Growth Evaluation*, 1(5), 1183-1198. <https://doi.org/10.54660/IJMRGE.2020.1.5.1183-1198>
- [101] Tiamiyu, R., Alaka, O. B., & Ajayi, O. (2021). Emergency response coverage gaps on intercity highways in Kwara State: A network service-area analysis. *International Journal of Multidisciplinary Research and Growth Evaluation*, 2(6), 1102-1118. <https://doi.org/10.54660/IJMRGE.2021.2.6.1102-1118>
- [102] Tonoyan, A., Dada, O., & Ayivi-Donkor, S. S. (2021). Advances in demand forecasting: Machine learning algorithms for revenue projection and financial planning accuracy. *Gyanshauryam, International Scientific Refereed Research Journal*, 4(1), 282-317. <https://doi.org/10.32628/GISRRJ120361>
- [103] Tonoyan, A., Dada, O., & Ayivi-Donkor, S. S. (2022). Conceptual model for predictive cost optimization: Machine learning applications in business transformation analysis. *International Journal of Economics and Business Management*, 8(6), 68-102. <https://doi.org/10.56201/ijebm.v8.no6.2022.pg68.102>
- [104] Tonoyan, A., Dada, O., & Ayivi-Donkor, S. S. (2024). Real-time KPI tracking systems: A review of automated performance monitoring and data-driven decision making. *World Journal of Innovation and Modern Technology*, 8(6), 247-281. <https://doi.org/10.56201/wjimt.v8.no6.2024.pg247.281>
- [105] Tonoyan, A., Dada, O., & Ayivi-Donkor, S. S. (2025). A conceptual model for real-time route optimization: Algorithmic pathways to delivery profitability and logistics cost reduction. *World Journal of Innovation and Modern Technology*, 9(12), 308-364. <https://doi.org/10.56201/wjimt.v9.no12.2025.pg308.364>
- [106] Toorajipour, R., Sohrabpour, V., Nazarpour, A., Oghazi, P., & Fischl, M. (2021). Artificial intelligence in supply chain management: A systematic literature review. *Journal of Business Research*, 122, 502-517. <https://doi.org/10.1016/j.jbusres.2020.09.009>
- [107] Ulmer, M. W., Goodson, J. C., Mattfeld, D. C., & Hennig, M. (2019). Offline-online approximate dynamic programming for dynamic vehicle routing with stochastic requests. *Transportation Science*, 53(1), 185-202. <https://doi.org/10.1287/trsc.2017.0767>
- [108] Ulmer, M. W., Goodson, J. C., Mattfeld, D. C., & Thomas, B. W. (2020). On modeling stochastic dynamic vehicle routing problems. *EURO Journal on Transportation and Logistics*, 9(2), 100008. <https://doi.org/10.1016/j.ejtl.2020.100008>
- [109] Vidal, T. (2022). Hybrid genetic search for the CVRP: Open-source implementation and SWAP* neighborhood. *Computers & Operations Research*, 140, 105643. <https://doi.org/10.1016/j.cor.2021.105643>
- [110] Vinyals, O., Fortunato, M., & Jaitly, N. (2015). Pointer networks. In *Advances in Neural Information Processing Systems 28* (pp. 2692-2700). https://proceedings.neurips.cc/paper_files/paper/2015/hash/29921001f2f04bd3baee84a12e98098f-Abstract.html
- [111] Vovk, V., Gammernan, A., & Shafer, G. (2005). *Algorithmic learning in a random world*. Springer. <https://doi.org/10.1007/b106715>
- [112] Walawalkar, G., Oduleye, T. E., Adesuyi, M. O., & Kalu, A. (2026). Next-generation financial analytics frameworks for AI-enabled enterprises. *International Journal of Advanced Multidisciplinary Research and Studies*, 6(1), 1779-1791.
- [113] Wang, D., Zhang, J., Cao, W., Li, J., & Zheng, Y. (2018). When will you arrive? Estimating travel time based on deep neural networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1). <https://doi.org/10.1609/aaai.v32i1.11877>
- [114] Wedraogo, L., Essandoh, S., Sakyi, J. K., Ibrahim, A. K., Okafor, C. M., Babalola, A. S., & Adenuga, M. A. (2025). Analysis of vendor management practices and their effects on supply chain performance. *International Journal of Multidisciplinary Futuristic Development*, 6(1), 42-54. <https://doi.org/10.54660/IJMF.2025.6.1.42-54>
- [115] Williams, R. J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 8(3-4), 229-256. <https://doi.org/10.1007/BF00992696>