Metaheuristic Optimization Review

♦ ISSN: 2692-4048 Vol: 03, No. 02, 2024 ♦ PP. 21–37

Article DOI: https://doi.org/10.54216/MOR.030205

REVIEW ARTICLE

Artificial Intelligence-Based Forecasting of Vehicle Carbon Dioxide Emissions A Review of Machine Learning Models, Optimization Strategies, and Sustainable Transportation Applications

Mahmoud Elshabrawy Mohamed[1,*]

[1] Computer Science and Intelligent Systems Research Center, Blacksburg 24060, Virginia, USA

Email: mshabrawy@jcsis.org

Received: November 23, 2024 Revised: December 27, 2024 Accepted: January 17, 2025 ⋆ Corresponding author

ABSTRACT

Forecasting carbon dioxide emissions from vehicles is essential for sustainable transportation, climate-aware urban planning, and intelligent environmental management, as road transport contributes substantially to energy-related greenhouse gas emissions. Rapid growth in vehicle ownership, mobility demand, fuel use, freight activity, and congestion has exposed the limits of conventional emission-estimation methods, particularly when nonlinear and multi-factor relationships must be represented. This review examines how artificial intelligence and machine learning support vehicle carbon dioxide emission forecasting through data-driven models, deep learning, feature selection, hyperparameter optimization, and metaheuristic algorithms. It emphasizes the value of integrating traffic volume, vehicle attributes, fuel type, speed patterns, road conditions, temporal behavior, weather variables, and energy-consumption indicators to improve prediction reliability under real-world conditions. Optimized forecasting models can support emission monitoring, mitigation planning, traffic management, and low-carbon mobility policies, although challenges remain in data availability, interpretability, uncertainty, regional transferability, computational cost, and transparent validation.

Keywords: Vehicle carbon dioxide emissions; Emission forecasting; Artificial intelligence; Machine learning; Sustainable transportation.

1. INTRODUCTION

The forecasting of vehicle-related CO_2 emissions has become a central research direction in sustainable transportation, environmental intelligence, and low-carbon urban governance because transport systems increasingly operate under the combined pressure of mobility growth, energy demand, congestion, freight expansion, and climate-mitigation obligations. Vehicle emissions are not produced by a single isolated factor; rather, they emerge from the interaction between traffic flow, vehicle type, fuel or energy source, route selection, driving speed, operating conditions, road geometry, logistics demand, and broader socio-economic activity. This complexity makes vehicle CO_2 forecasting fundamentally different from conventional static emission accounting, since prediction models must capture temporal variation, nonlinear relationships, and region-specific mobility behavior. Within this context, artificial intelligence-based forecasting provides a promising analytical foundation because it can learn hidden dependencies from heterogeneous data sources and transform raw transportation indicators into actionable emission estimates. Re-

© 2024 The Author(s). Published by ASPG. This article is licensed under CC BY 4.0.

cent work on deep learning and metaheuristic optimization for CO_2 prediction demonstrates the growing movement toward data-driven emission modeling frameworks that are capable of handling complex environmental patterns more flexibly than purely linear or assumption-dependent approaches [1].

The importance of forecasting vehicle CO_2 emissions is closely connected to the transformation of transportation systems from reactive infrastructure management toward predictive and sustainability-oriented decision-making. In traditional practice, emission assessment has often depended on historical fuel consumption, aggregated traffic statistics, or predefined emission factors, which may be useful for broad inventories but are less effective when cities, logistics operators, or policymakers need dynamic forecasts under changing traffic and operational conditions. Vehicle emissions vary across time and space because congestion, speed fluctuation, route choice, payload condition, driving behavior, and energy efficiency do not remain constant. Therefore, forecasting models must be able to represent both short-term operational variability and long-term structural changes in transportation demand. Intelligent transportation research has increasingly emphasized the role of machine learning in route selection and transport planning, showing that predictive analytics can support more informed mobility decisions when conventional optimization alone cannot fully represent complex intermodal and route-dependent relationships [2].

A major motivation for reviewing artificial intelligence-based vehicle CO_2 forecasting is that modern transportation data have become larger, more diverse, and more difficult to interpret through classical statistical methods alone. Vehicle detectors, navigation systems, fleet-management platforms, road sensors, electric vehicle monitoring systems, satellite-based positioning, traffic cameras, and urban Internet of Things infrastructures continuously produce high-dimensional data streams. These data sources contain valuable information about traffic intensity, route delay, stop-and-go movement, vehicle operating states, energy consumption, and road-use patterns; however, they also introduce challenges related to noise, missing values, temporal dependency, and feature redundancy. Machine learning and deep learning models are increasingly used because they can approximate nonlinear relationships between these variables and emission outcomes without requiring strict assumptions about the form of the underlying system. In electric-vehicle-related studies, optimized deep recurrent architectures have been used to model energy consumption patterns, indicating the relevance of advanced sequence-learning models for transportation systems where energy use and emissions are shaped by time-dependent driving behavior [3].

Vehicle emission forecasting also intersects with freight transportation and logistics, where operational decisions can strongly influence fuel consumption and carbon output. Truck selection, shipment consolidation, fleet sizing, routing, and delivery scheduling are not only economic or logistical problems; they also determine how many vehicles operate, how far they travel, how efficiently they are loaded, and how much energy is consumed across distribution networks. For this reason, forecasting CO_2 emissions from vehicles cannot be limited to passenger mobility alone, as freight vehicles often operate under heavier loads, longer routes, and more complex distribution constraints. Research on truck selection under consolidated shipment policies reflects the growing need to integrate vehicle choice, shipment structure, and operational planning into transportation modeling frameworks [4]. In an emission-forecasting review, this perspective is important because it expands the analytical scope from isolated traffic prediction to the broader operational mechanisms that generate vehicle-related carbon emissions. The environmental relevance of vehicle CO_2 forecasting is further reinforced by the need to connect transportation prediction with energy planning, urban sustainability, and carbon-aware infrastructure development. Transportation systems do not operate independently from energy systems; vehicle emissions are influenced by fuel structure, electric vehicle adoption, charging infrastructure, renewable energy integration, and the carbon intensity of electricity generation. As cities move toward electrification and intelligent mobility, emission forecasting must consider not only tailpipe emissions but also the wider energy implications of vehicle operation and charging demand. Research on adaptive hybrid optimization for distribution network planning with electric vehicle charging highlights the growing coupling between transport electrification and energy-system optimization [5]. This relationship is essential for future emission-forecasting frameworks because the environmental benefit of electric mobility depends on coordinated planning across transportation networks, charging infrastructure, and electricity supply systems.

Traffic flow prediction represents another foundational component of vehicle CO_2 forecasting because emission rates are closely associated with vehicle movement conditions. Congested traffic, frequent acceleration and deceleration, unstable speed profiles, and bottleneck formation can increase energy use and carbon output, whereas smoother traffic operation can improve efficiency. Forecasting traffic flow therefore provides a critical intermediate layer for estimating emissions before they occur. Machine learning-based traffic prediction for highway vehicle detectors illustrates how decomposition strategies and data-driven models can be used to extract temporal traffic patterns from sensor-based transportation data [6]. In the context of CO_2 forecasting, such traffic prediction methods are valuable because they allow emission models to move beyond static vehicle counts and incorporate the dynamic operating conditions that directly shape energy consumption.

A complete review of vehicle CO_2 forecasting must also address the role of connected and autonomous vehicle technologies, since emerging mobility systems may alter traffic behavior, energy demand, and emission patterns in ways that differ from conventional vehicle fleets. Connected and autonomous vehicles can influence emissions through smoother driving, cooperative traffic control, route coordination, platooning, reduced human driving variability, and improved road-capacity utilization. However, their environmental effect is not automatically positive because induced travel demand, empty vehicle movement, additional computational energy use, and mixed-traffic interactions may complicate the net sustainability outcome. Research examining the environmental sustainability and mobility impacts of connected and autonomous vehicles on urban highways shows that advanced vehicle technologies must be evaluated through both

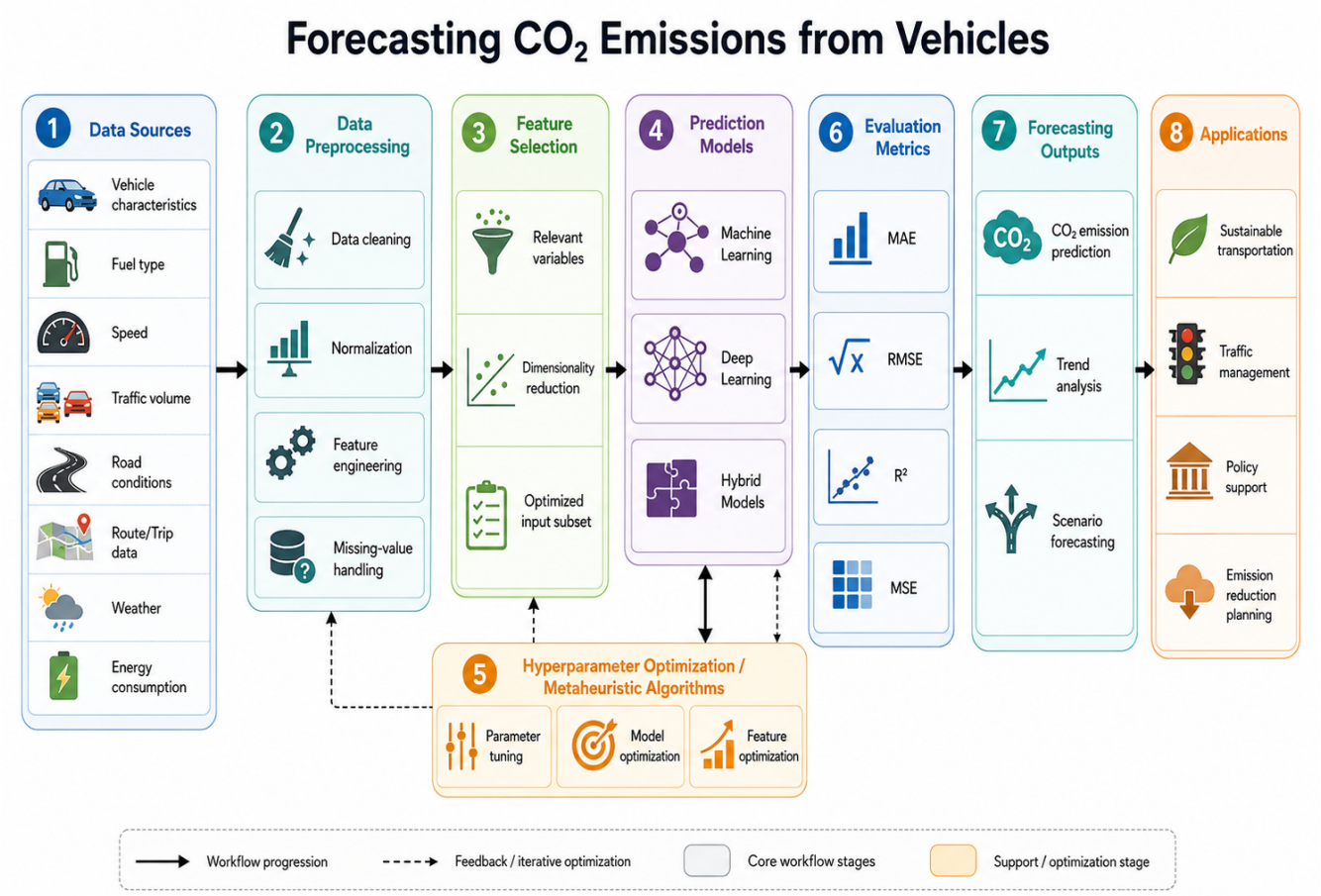


Figure 1. Full-width comparison of total CO_2 emission values across vehicle brands, measured in g/km, with brands arranged from the lowest to the highest emission level.

mobility and environmental lenses [7]. This review therefore treats intelligent emission forecasting as a forward-looking requirement for transportation systems that are becoming increasingly automated, connected, and data-intensive.

Vehicle routing and speed optimization are also directly relevant to CO_2 emission forecasting because they translate predictive information into operational decisions. Forecasting alone provides an estimate of future emissions, but its practical value increases when it supports route planning, speed management, fleet dispatching, and carbon-aware logistics control. In fragile goods transportation, for example, vehicle routing and speed decisions must balance delivery reliability, product protection, travel efficiency, and emission control. Research on route and speed optimization with carbon emission control demonstrates that emission reduction can be embedded within transportation decision models rather than treated as an external environmental indicator [8]. This connection is important for the present review because it clarifies that artificial intelligence-based emission forecasting should not be viewed only as a prediction exercise; it should also support prescriptive decisions that guide transport systems toward lower-carbon operation.

Recent developments in green intelligent transportation systems further show that emission forecasting is becoming part of a broader computational ecosystem that includes edge computing, multi-objective optimization, real-time sensing, and distributed decision-making. Edge computing can reduce latency by processing transportation data closer to vehicles, roadside units, or traffic-control devices, which is particularly important when emission-aware decisions must be made quickly. Multi-objective metaheuristic algorithms can simultaneously consider competing goals such as travel time, energy use, congestion reduction, service quality, and emission minimization. Research on optimized green intelligent transportation systems using edge computing and multi-objective metaheuristic algorithms reflects this movement toward integrated, real-time, and sustainability-centered transportation intelligence [9]. For vehicle CO_2 forecasting, this trend suggests that future models should be evaluated not only by statistical accuracy but also by their ability to operate efficiently within intelligent transportation infrastructures.

Despite the rapid growth of artificial intelligence-based methods, several limitations continue to affect the reliability, comparability, and deployment readiness of vehicle CO_2 emission forecasting models. Many studies differ in dataset scale, geographic coverage, input variables, validation design, emission-estimation assumptions, and performance metrics, which makes it difficult to compare results directly across regions or transport modes. Some models may achieve strong predictive performance on a specific dataset but fail to generalize when applied to a different city, road type, fleet composition, or policy environment. Other challenges include limited interpretability, insufficient uncertainty analysis, dependence on high-quality sensor data, computational complexity, and the absence of standardized benchmarking protocols. Broader

reviews of green location-routing problems show that sustainable transportation research involves diverse objectives, constraints, and methodologies, reinforcing the need for careful synthesis rather than isolated model comparison [10]. Accordingly, this review aims to provide an academic synthesis of artificial intelligence and optimization-based approaches for forecasting vehicle CO_2 emissions, with emphasis on predictive modeling logic, transportation relevance, optimization strategies, sustainability implications, and the methodological challenges that must be addressed to support robust low-carbon mobility planning.

2. LITERATURE REVIEW

The literature on vehicle-related CO_2 emission forecasting has developed across several connected research streams, including transportation energy analysis, intelligent mobility prediction, logistics optimization, electric vehicle energy management, and data-driven environmental modeling. Earlier approaches to transport emission assessment were commonly oriented toward inventory construction, fuel-use estimation, and policy-level carbon accounting, where the objective was to quantify the amount of emissions generated by a transport system over a defined period. Although such approaches remain useful for macro-level planning, they are not always sufficient for forecasting future emission behavior under rapidly changing mobility conditions. Contemporary transportation systems are shaped by urban expansion, variations in vehicle ownership, freight intensity, travel-demand redistribution, fuel-price sensitivity, traffic congestion, and the progressive introduction of electric and hybrid vehicles. These interacting elements make vehicle CO_2 emissions a dynamic output of mobility behavior rather than a fixed environmental quantity. For this reason, the literature has increasingly moved toward intelligent nonlinear modeling, where transportation energy consumption and carbon emissions are examined through computational approaches capable of representing changing temporal patterns and system-level dependencies [11].

A significant part of the reviewed literature emphasizes that vehicle CO_2 forecasting should not be understood only as a numerical prediction task, but also as a decision-support mechanism for low-carbon transportation governance. Accurate forecasts can help authorities estimate future emission pressure before it becomes visible in environmental indicators, allowing earlier intervention through traffic regulation, route redesign, modal-shift policies, public-transport planning, and energy-efficiency programs. In this context, forecasting becomes a bridge between historical monitoring and proactive mitigation. This role is especially relevant in urban environments, where road transport is closely connected with air pollution, congestion, commuting behavior, land-use distribution, and infrastructure capacity. Machine learning-based approaches have been investigated within smart-city contexts to moderate future CO_2 emissions, showing that predictive analytics can support environmental planning when urban systems produce large volumes of heterogeneous mobility and operational data [12]. The literature, therefore, positions artificial intelligence as an enabling layer that can transform urban transport data into forward-looking carbon-management knowledge.

The reviewed studies also indicate that the quality of vehicle CO_2 forecasting depends strongly on how the modeling problem is formulated. Vehicle emissions may be predicted directly from measured or estimated emission records, or indirectly through intermediate variables such as traffic volume, speed, fuel consumption, trip length, vehicle load, route type, and energy demand. Direct forecasting is useful when reliable emission observations are available, but many transport systems do not maintain continuous emission measurements at the required spatial and temporal resolution. Indirect forecasting can overcome this limitation by linking vehicle operation and energy consumption to carbon output, but it requires careful selection of explanatory variables. A model that ignores road conditions, driving cycles, congestion states, or vehicle classes may produce incomplete emission estimates even if it performs well statistically on a limited dataset. Therefore, the literature highlights the importance of aligning data inputs with the physical and operational mechanisms that generate vehicle emissions, rather than treating forecasting as a purely abstract regression problem.

Another important direction in the literature concerns the increasing role of high-dimensional data in transportation emission modeling. Modern transport networks generate information from loop detectors, floating-car data, global positioning systems, fleet-management platforms, toll records, charging stations, weather sensors, and roadside monitoring devices. These sources allow researchers to construct more detailed representations of vehicle activity, but they also introduce methodological difficulties. Missing observations, inconsistent sampling intervals, outliers, duplicated records, and regional differences can influence the stability of forecasting models. Consequently, data preprocessing is not a secondary technical step; it is a central part of the modeling pipeline. Cleaning, normalization, aggregation, temporal alignment, and feature engineering affect how accurately the model can learn the relationship between vehicle movement and carbon output. The literature, therefore, suggests that emission forecasting should be evaluated not only by the final predictive score but also by the reliability of the data preparation strategy that precedes model training.

Within this data-preparation context, feature selection has emerged as a particularly important issue because vehicle emission datasets often contain variables with different levels of relevance, redundancy, and uncertainty. Including too many inputs may increase computational burden and expose the model to noise, while excluding influential variables may reduce the ability to capture emission dynamics. In transportation forecasting, useful features may include speed-related indicators, acceleration patterns, stop frequency, trip duration, vehicle age, engine type, fuel category, road gradient, ambient temperature, traffic density, and time-of-day effects. However, not every dataset contains all these variables, and not every variable contributes equally across regions or vehicle categories. This creates a need for systematic feature selection methods that can identify the most informative predictors without relying only on manual assumptions. In this respect, metaheuristic-based feature selection is increasingly relevant because it can explore large combinations of candidate variables and search for subsets that improve predictive accuracy while reducing model complexity.

From feature selection, the literature further shows that different predictive model families offer distinct strengths for vehicle CO_2 forecasting. Classical machine learning models are often valued for their relatively lower computational requirements, stronger interpretability, and ability to perform well on structured tabular datasets. Ensemble models can improve robustness by combining multiple learners and reducing the instability associated with single-model predictions. Deep learning models are more suitable when the data contain temporal, sequential, or high-dimensional structures, especially in cases involving traffic time series, energy-consumption trajectories, or sensor streams. However, deep architectures may require larger datasets, careful tuning, and stronger computational resources. This variation means that no single model type can be universally considered superior across all vehicle emission forecasting contexts. Instead, the literature encourages model selection based on dataset size, prediction horizon, input structure, interpretability requirements, and deployment conditions.

Temporal modeling is another major concern because vehicle emissions change according to daily, weekly, seasonal, and event-driven mobility patterns. Morning and evening peaks, holiday travel, weather disruptions, freight schedules, school calendars, and fuel-price changes can all alter traffic intensity and emission rates. Static models that ignore temporal dependency may fail to capture these fluctuations, particularly when forecasting short-term emission peaks or long-term carbon trajectories. Time-aware models can address this issue by learning dependencies across previous observations and using them to estimate future conditions. This is especially important for real-time traffic management, where a city may need to forecast near-future emission levels in order to adjust signal timing, recommend alternative routes, or manage restricted zones. The literature, therefore, treats temporal sensitivity as a necessary property of advanced emission forecasting systems rather than an optional modeling enhancement.

Vehicle routing and freight distribution research provides another important foundation for understanding how operational decisions influence carbon emissions. In logistics systems, emissions are shaped by the number of vehicles used, route distance, vehicle capacity, delivery sequence, driver experience, loading efficiency, and service-time constraints. Last-mile delivery is particularly relevant because it involves frequent stops, dense urban movement, customer-specific routing, and strong dependence on driver knowledge. Machine learning has been used to enhance last-mile delivery modeling by capturing operational experience that may not be fully represented through mathematical rules alone [13]. This line of research is important for vehicle CO_2 forecasting because it shows that emission prediction can benefit from operational intelligence, especially in delivery environments where human driving decisions, local road familiarity, and customer distribution patterns affect actual vehicle movement.

The reviewed literature also demonstrates that routing constraints must be incorporated carefully when emission forecasting is linked to transport optimization. Limited-capacity vehicle routing problems show that vehicle capacity, customer demand, route feasibility, and distribution constraints can substantially affect distance traveled and energy consumed. Algorithmic advances in limited-capacity routing have expanded the ability to solve complex vehicle-distribution problems more efficiently, which is relevant for emission-aware logistics because routing decisions often determine the carbon consequences of freight activity [14]. Within the forecasting context, this means that emission models should not only estimate future carbon output from historical data, but also account for how operational constraints shape the conditions under which those emissions are produced. The connection between forecasting and routing, therefore, strengthens the practical value of artificial intelligence-based emission models in real transportation systems.

Electric and hybrid vehicle research adds another layer to the literature because vehicle CO_2 emissions are increasingly linked to energy-source configuration, charging behavior, battery management, and grid interaction. Electric vehicles reduce direct tailpipe emissions, yet their overall carbon impact depends on electricity generation sources, charging timing, charging-station placement, renewable energy availability, and vehicle-to-grid management strategies. Studies on photovoltaic-aided electric vehicle energy management show that emission-related transportation planning increasingly overlaps with renewable energy optimization and intelligent power control [15]. This perspective is essential for forecasting because future vehicle emissions may not be estimated accurately if the model treats all vehicles through conventional fuel-consumption logic. As electrification expands, forecasting systems must distinguish between direct vehicle emissions, indirect energy-related emissions, and operational energy efficiency.

The literature on artificial intelligence-based emission forecasting also stresses the need for careful hyperparameter optimization. Model performance can change substantially when parameters such as learning rate, tree depth, number of estimators, kernel configuration, hidden units, batch size, regularization coefficient, and optimization settings are adjusted. Manual tuning may be insufficient because the search space is often large and nonlinear. Metaheuristic algorithms are therefore widely used to guide the tuning process, especially when the objective is to improve predictive performance while controlling model complexity. In vehicle CO_2 forecasting, this is important because an untuned model may appear weak even when the underlying algorithm is suitable for the dataset. Conversely, excessive tuning without proper validation may produce overfitting. The literature therefore supports the use of optimization methods within disciplined experimental designs that include clear training, validation, and testing procedures.

Model validation remains one of the most important issues in the reviewed literature because emission forecasting models are only useful if their performance remains reliable outside the exact data used for training. Many models can produce strong results on historical datasets, but their practical value depends on whether they generalize across time periods, road types, cities, traffic regimes, and vehicle categories. A rigorous validation strategy should therefore include appropriate data splitting, comparison with baseline models, multiple performance indicators, and uncertainty-aware interpretation. In emission forecasting, relying on a single metric may be misleading because one measure may emphasize av-

erage error while another reflects variance explanation or extreme-value behavior. For this reason, the literature increasingly favors multi-metric evaluation, especially when forecasts are expected to support environmental policy or operational decision-making.

Interpretability is another recurring concern because transportation emission forecasts may influence public policy, infrastructure investment, traffic regulation, and environmental compliance. A highly accurate model may still be difficult to adopt if decision-makers cannot understand why it produces a certain forecast or which variables are most responsible for predicted emission increases. Interpretable modeling can help identify whether emissions are driven primarily by traffic density, speed instability, vehicle type, route distance, temperature, or energy-consumption behavior. This information is valuable because it converts prediction into explanation, allowing policymakers to design targeted interventions rather than relying only on black-box outputs. The literature therefore suggests that future emission forecasting frameworks should combine predictive accuracy with explanatory capacity, especially when artificial intelligence is used in public-sector transportation planning.

Another methodological issue involves transferability. A forecasting model trained in one city, country, or transportation network may not perform equally well elsewhere because vehicle fleets, road infrastructure, driving behavior, fuel composition, public-transport availability, and policy conditions differ significantly. This limitation is particularly important for review studies because results reported in one geographic setting cannot be automatically generalized to another. Transferability requires either adaptable modeling structures, region-specific recalibration, or the development of broader datasets that represent multiple mobility environments. Without such attention, artificial intelligence-based vehicle CO_2 forecasting may remain fragmented into isolated case studies rather than developing into a more generalizable scientific framework.

The literature also reveals a gap between prediction accuracy and implementation readiness. Many studies focus on improving numerical performance, but fewer examine how forecasting models can be integrated into real-time transport-management platforms, municipal carbon dashboards, fleet-control systems, or energy-planning tools. Deployment requires more than a trained model; it requires reliable data pipelines, computational efficiency, user-oriented visualization, uncertainty communication, cybersecurity protection, and institutional capacity to act on the forecast. This gap is important because a model that performs well in an experimental environment may still be difficult to operationalize in a city or logistics network. Therefore, implementation-oriented research should evaluate not only predictive metrics but also latency, scalability, maintainability, and compatibility with existing transportation infrastructure.

Overall, the literature shows that vehicle CO_2 emission forecasting is evolving from static estimation toward integrated artificial intelligence-based modeling that connects transportation behavior, energy consumption, optimization, and environmental decision-making. The reviewed studies collectively indicate that strong forecasting frameworks require high-quality data, meaningful feature representation, appropriate model selection, systematic hyperparameter tuning, rigorous validation, and interpretable outputs. At the same time, the field still requires stronger standardization, broader benchmarking, improved transferability, and closer alignment between predictive modeling and practical low-carbon mobility planning. These observations justify the need for a focused review that synthesizes artificial intelligence, machine learning, feature selection, hyperparameter optimization, and metaheuristic strategies for vehicle CO_2 emission forecasting without treating these elements as isolated technical components. Table 1 presents a methodology-oriented synthesis of representative studies relevant to vehicle CO_2 emission forecasting, sustainable transportation modeling, energy-aware mobility systems, and optimization-supported environmental prediction. The table is designed to clarify how different methodological directions contribute to the development of intelligent forecasting frameworks, including fleet planning, vehicle routing, electric vehicle energy estimation, traffic-related environmental analysis, carbon-aware optimization, and data-driven sustainability modeling.

The methodology references summarized in Table 1 collectively establish the analytical foundation for reviewing vehicle CO_2 emission forecasting as a multidisciplinary problem that cannot be restricted to one modelling family, one data source, or one transportation scenario. The selected studies were organized to represent the main methodological directions that support intelligent emission forecasting, including predictive modelling, route optimization, fleet management, energy estimation, environmental forecasting, renewable-energy integration, sensing infrastructure, and carbon-aware operational planning. This methodological structure is important because vehicle-related CO_2 emissions are generated through several interacting layers: the physical movement of vehicles, the energy consumed during operation, the routing and scheduling decisions that determine distance and load, the infrastructure conditions that influence driving efficiency, and the energy systems that support electrified mobility. Therefore, the methodology matrix does not treat the reviewed references as isolated studies; rather, it positions them as complementary methodological evidence for building a more complete understanding of how artificial intelligence, optimization, and transportation analytics can support low-carbon mobility forecasting.

The first methodological direction represented in Table 1 concerns fleet sizing and routing under uncertainty. This direction is particularly important because vehicle CO_2 emissions are directly affected by how many vehicles are deployed, how routes are assigned, how delivery and pickup operations are coordinated, and how transportation decisions remain stable when demand or operational conditions change. Adjustable robust optimization provides a useful framework for this type of problem because it allows fleet and routing decisions to be designed under uncertain conditions rather than assuming that all inputs are fixed and fully known. In emission-oriented transportation planning, this is valuable because uncertain demand, travel time, customer distribution, and pickup volumes can alter fuel consumption and vehicle activity. By including robust optimization in the methodology matrix, the review highlights that emission forecasting should be connected with decision models capable of maintaining operational reliability

while reducing unnecessary vehicle movement and inefficient routing patterns.

A second methodological direction is electric vehicle energy-state estimation, which contributes to the review by linking emission forecasting with the technical behavior of electrified transportation systems. Although electric vehicles do not produce direct tailpipe CO_2 emissions during operation, their energy consumption remains essential for evaluating indirect emissions, charging demand, grid load, and the environmental benefit of electrification. Battery state estimation is therefore not simply a technical battery-management task; it is a necessary modelling component for understanding how efficiently an electric vehicle uses energy under changing operating conditions. Hybrid optimizer-assisted neural estimation methods are relevant because they combine learning ability with search-based parameter refinement, allowing the model to estimate battery state more accurately than a manually configured structure. In the context of this review, this reference supports the argument that future vehicle CO_2 forecasting must account for electric vehicle energy behavior, especially as transport systems move from fuel-centered emission models toward mixed fleets containing combustion, hybrid, and fully electric vehicles.

The methodology matrix also includes swarm-intelligence-supported emission modelling because carbon emission data often exhibit nonlinear temporal behavior, long-term dependency, and structural variation across regions or economic conditions. Swarm intelligence methods are useful in this context because they search through complex solution spaces by imitating collective adaptive behavior, making them suitable for model tuning, feature selection, and optimization-based prediction. Energy-related CO_2 forecasting studies based on long historical datasets provide a broader methodological reference for vehicle emission forecasting because transportation emissions are one part of the wider energy-emission system. The inclusion of this direction strengthens the review by showing that vehicle CO_2 prediction can benefit from optimization mechanisms already applied in energy-related emission analysis, particularly when the forecasting problem involves multi-year patterns, nonstationary trends, and multiple influencing indicators.

Environmental forecasting methods are also represented because vehicle CO_2 emissions are closely related to broader air-quality and pollution-monitoring problems. Although particulate matter and carbon dioxide differ in chemical nature, measurement logic, and environmental effect, both forecasting tasks share important methodological requirements, including noisy time-series data, variable meteorological influence, nonlinear pollutant behavior, and the need for reliable short-term prediction. Hybrid regression models that use dynamic weighting and genetic algorithm-based optimization are relevant to the present review because they show how environmental prediction can be improved by assigning adaptive importance to model components and optimizing predictive structures. This contributes to the methodology discussion by demonstrating that emission forecasting frameworks should not rely only on a single static learner when environmental data are affected by changing external conditions and irregular temporal patterns.

Another important methodological category concerns greenhouse-gas-aware production and distribution network design. Vehicle emissions are not produced only by individual driving behavior; they are also shaped by the structure of supply chains, the placement of distribution channels, the segmentation of customers, and the coordination between production and delivery systems. A dual-channel production-distribution framework contributes to vehicle CO_2 forecasting by showing how network-level decisions can reduce greenhouse-gas impacts before transportation activity occurs. This is methodologically relevant because forecasting models should be able to reflect not only observed vehicle movement but also the organizational decisions that create or reduce the need for movement. By including this reference, Table 1 extends the review beyond direct traffic prediction and emphasizes that low-carbon transportation modelling must also consider how logistics networks generate vehicle demand.

Multi-objective optimization is another central methodological element because sustainable transportation problems rarely involve one objective only. A vehicle routing or emission-control system may need to minimize cost, travel distance, delivery delay, energy use, congestion contribution, and carbon output at the same time. These objectives may conflict with one another; for example, the shortest route may not always be the lowest-emission route if it contains congestion, steep gradients, or repeated stop-and-go movement. Multi-objective ant-colony-based optimization provides a methodological reference for handling such trade-offs because it supports the search for balanced solutions rather than a single narrowly optimized result. In this review, this direction is important because vehicle CO_2 forecasting becomes more useful when it can be connected to decision-making frameworks that evaluate several sustainability and operational goals simultaneously.

Infrastructure-condition prediction is included in the methodology matrix because road quality affects vehicle operation in ways that can influence energy consumption and emissions. Pavement degradation, surface roughness, maintenance delays, and poor road conditions may increase rolling resistance, reduce speed stability, and create inefficient driving patterns. Interpretable prediction of pavement performance is therefore relevant to vehicle CO_2 forecasting because infrastructure condition can act as an indirect but meaningful contributor to emission behavior. This methodological direction adds an infrastructure-maintenance dimension to the review and shows that emission forecasting should not be limited to vehicle and traffic variables only. Instead, predictive transportation models can be strengthened when they include infrastructure-related indicators that explain why energy consumption may increase on certain road segments or under specific maintenance conditions.

The inclusion of sustainable construction intelligence provides a broader methodological reference for how machine learning and deep learning reviews can be structured in sustainability-oriented domains. Although the application area differs from vehicle emission forecasting, the methodological relevance lies in the way artificial intelligence studies are assessed through trends, gaps, data availability, learning models, validation strategies, and future research needs. This supports the present review by reinforcing the importance of

organizing artificial intelligence-based sustainability research around model purpose, data type, evaluation logic, and implementation barriers. It also helps justify why the methodology matrix does not merely list models, but instead classifies the reviewed references according to their methodological contribution to low-carbon prediction and optimization. This classification approach is necessary for producing a review that is analytically useful rather than descriptive only.

Microgrid energy planning appears in the methodology matrix because vehicle electrification connects transportation forecasting with distributed energy systems. As electric vehicle adoption increases, the environmental value of electrified mobility depends partly on how charging demand is served, how renewable energy is integrated, and how grid operation is optimized. Hybrid swarm-based optimization for microgrid planning is relevant because it addresses the problem of minimizing operating cost while coordinating distributed energy resources. For vehicle CO_2 forecasting, this type of methodology provides a pathway for estimating indirect emissions associated with charging infrastructure and electricity supply. It also supports the view that future forecasting frameworks should treat vehicle emissions and energy planning as coupled systems rather than separate domains.

Renewable-energy integration under uncertainty is another methodological direction that strengthens the transportation-emission forecasting framework. Hybrid renewable energy sources are variable by nature because solar and wind generation depend on meteorological conditions, seasonal availability, and probabilistic fluctuations. When electric vehicles are charged using renewable-supported grids, the resulting carbon intensity of mobility depends on how effectively these uncertain resources are managed. Metaheuristic optimization in probabilistic optimal power-flow modelling is therefore relevant to this review because it provides a computational basis for coordinating uncertain renewable generation with energy-demand requirements. This methodological perspective is important because vehicle CO_2 forecasting will become increasingly dependent on the carbon profile of electricity systems as transport electrification expands.

Sectoral carbon prediction is included to connect vehicle emission forecasting with broader net-zero modelling objectives. Although the referenced study focuses on another industrial sector, its methodological value lies in the use of predictive modelling to estimate future CO_2 emissions in support of decarbonization targets. This is relevant because transportation planning also requires long-term carbon forecasting to evaluate whether emission-reduction strategies are sufficient for meeting climate objectives. The methodology matrix therefore uses this reference to show that emission prediction is not merely a short-term operational task; it can also support strategic planning, policy assessment, and net-zero pathway evaluation. For vehicle-related applications, similar predictive structures can help estimate future emission trajectories under changing fleet composition, energy efficiency, mobility demand, and regulatory conditions.

Sensor-network optimization is methodologically significant because intelligent vehicle emission forecasting depends on the availability, coverage, and reliability of transportation data. Without sufficient sensing infrastructure, forecasting models may suffer from missing observations, spatial blind spots, incomplete traffic representation, and weak real-time monitoring capacity. Coverage optimization in wireless sensor networks contributes to the review by addressing the spatial and technical foundation required for data-driven transportation intelligence. In practical emission forecasting, improved sensor placement can support more accurate traffic monitoring, road-condition observation, environmental measurement, and vehicle-flow estimation. This reference therefore represents the sensing layer of the methodology matrix, showing that the performance of artificial intelligence models depends not only on algorithmic design but also on the quality of the data-collection infrastructure that feeds the model.

Hybrid energy feasibility analysis is included because low-carbon vehicle forecasting must consider whether the energy systems supporting transportation are technically and economically viable. In many regions, electric mobility cannot be evaluated independently from local energy availability, renewable potential, diesel backup dependence, and cost constraints. A feasibility assessment of photovoltaic, wind, and diesel hybrid systems provides methodological support for evaluating energy configurations that may influence the indirect emissions of electric vehicle charging. This is relevant to vehicle CO_2 forecasting because a forecast that ignores the energy mix behind electrified transport may overestimate or underestimate the real environmental benefit. By including hybrid energy feasibility, the methodology matrix connects transport prediction with the practical conditions required for sustainable energy-supported mobility.

Carbon-emission prediction based on structural and process parameters contributes to the methodology matrix by demonstrating how emissions can be modelled from measurable technical variables. Although the application is not vehicle operation itself, the modelling logic is transferable: carbon output can be estimated when the system is described through relevant design, process, or operational parameters. For vehicle CO_2 forecasting, this supports the principle that emission models should be built from variables that have a meaningful relationship with the mechanisms generating carbon output. Such variables may include vehicle mass, speed profile, fuel type, route structure, road condition, acceleration behavior, payload, and energy-consumption indicators. This reference therefore reinforces the importance of parameter-based prediction rather than relying on general trend extrapolation alone.

Fleet-management optimization in mining operations is included because it provides a high-intensity example of vehicle-based carbon-footprint reduction under operational uncertainty. Mining fleets operate in demanding environments where vehicle dispatching, production planning, haulage activity, and fuel consumption are tightly connected. Joint stochastic fleet-management and production-planning approaches are relevant because they show how emission reduction can be incorporated into operational planning rather than treated as a separate environmental calculation. In vehicle CO_2 forecasting, this methodology is useful because it demonstrates how predictive and optimization-based methods can be combined to manage vehicle activity in systems where demand, production targets, and movement patterns vary over time.

Electric vehicle charging-station placement and sizing are

represented because charging infrastructure directly affects the feasibility and emission profile of electrified transportation. Poorly placed or incorrectly sized charging stations can increase detours, cause grid stress, reduce charging reliability, and weaken the practical benefits of electric mobility. Optimization of distribution network operating parameters in the presence of electric vehicle charging stations is methodologically relevant because it links transport electrification with power-network planning under uncertainty. For vehicle CO_2 forecasting, this direction supports a more complete modelling perspective in which electric vehicle emissions are not evaluated only at the vehicle level, but also through the infrastructure and grid conditions that determine how charging demand is supplied.

Plug-in hybrid electric vehicle integration adds another

Table 1. Representative artificial intelligence and optimization methodologies for vehicle CO_2 emission forecasting

Study focus	Data type	Methodological approach	Main objective
Fleet sizing and routing [16]	Fleet operation and routing data	Adjustable robust optimization for integrated fleet-sizing and vehicle-routing decisions	Supports emission-conscious transportation planning by addressing uncertainty in delivery and pickup operations, allowing vehicle deployment and routing decisions to remain reliable under variable operational conditions.
Vehicle energy estimation [17]	Electric vehicle battery and operating data	Hybrid optimizer-assisted feed-forward neural network for battery state estimation	Improves the estimation of electric vehicle energy status, which is essential for linking vehicle operation, energy consumption, charging decisions, and indirect carbon-emission assessment.
Swarm-based emission modelling [18]	Energy-related CO_2 emission data	Swarm-intelligence-supported predictive modelling	Examines energy-related carbon-emission behaviour through computational intelligence, providing a methodological basis for nonlinear emission forecasting across long-term environmental datasets.
Air-pollution forecasting [19]	Particulate matter concentration data	Dynamic weighting and genetic algorithm-based hybrid regression modelling	Enhances environmental forecasting accuracy by combining adaptive weighting with optimization-based learning, offering a transferable methodology for pollution-sensitive transport-emission prediction.
Greenhouse-gas network design [20]	Production-distribution and customer-segmentation data	Dual-channel production-distribution network optimization	Reduces greenhouse-gas impacts by coordinating distribution structure and customer-oriented planning, supporting carbon-aware logistics systems connected with vehicle movement and delivery activity.
Multi-objective route optimization [21]	Optimization benchmark and application datasets	Multi-objective ant-colony-based optimization review	Provides methodological insight into solving transportation problems with competing objectives, such as travel cost, delivery efficiency, fuel use, and carbon-emission reduction.
Infrastructure-condition prediction [22]	Pavement performance and maintenance data	Interpretable prediction for preventive maintenance timing	Links transport infrastructure condition with predictive decision-making, supporting maintenance strategies that can reduce inefficient vehicle operation caused by road degradation.
Sustainable construction intelligence [23]	Construction and sustainability datasets	Machine learning and deep learning trend analysis	Offers methodological guidance on data-driven sustainability modelling, benchmarking, and gap identification, which can inform the design of transparent and reproducible vehicle-emission forecasting frameworks.
Microgrid energy planning [24]	Microgrid and energy-planning data	Hybrid swarm-based optimization for distributed energy planning	Supports the energy-system side of low-carbon mobility by improving planning methods for renewable-energy integration, which is relevant to electric vehicle charging and indirect emission reduction.
Renewable-powered mobility support [25]	Hybrid renewable energy and power-flow data	Metaheuristic optimization for probabilistic renewable-energy integration	Improves power-system planning under renewable uncertainty, allowing transport electrification strategies to be evaluated together with the carbon intensity of supporting energy infrastructure.
Industrial carbon prediction [26]	Sectoral production and emission data	Predictive modelling for net-zero-oriented CO_2 assessment	Demonstrates how data-driven carbon forecasting can support long-term decarbonization targets, offering methodological relevance for transportation sectors pursuing net-zero emission pathways.
Sensor-network optimization [27]	Wireless sensor network coverage data	Metaheuristic coverage optimization	Strengthens the sensing layer required for intelligent transportation monitoring by improving spatial coverage, data availability, and the reliability of emission-related traffic observation systems.
Hybrid energy feasibility [28]	Photovoltaic, wind, diesel, and techno-economic data	Metaheuristic feasibility assessment for hybrid energy systems	Assesses sustainable energy configurations that can support cleaner transport electrification, especially where vehicle charging depends on locally optimized renewable-energy resources.
Carbon-aware manufacturing prediction [29]	Structural design and machining parameter data	Carbon-emission prediction based on design and process parameters	Shows how carbon output can be estimated from technical design variables, providing a transferable modelling logic for vehicle-related systems where emissions depend on measurable operational parameters.
Fleet and production planning [30]	Mining fleet and short-term production data	Joint stochastic fleet-management and production-planning optimization	Integrates fleet decisions with carbon-footprint reduction, illustrating how vehicle-operation planning can be aligned with emission minimization under uncertainty.
Electric-vehicle charging optimization [31]	Distribution network and charging-station data	Optimization of grid-tied microgrid operating parameters	Improves the operation of energy networks connected to electric vehicle charging stations, supporting cleaner mobility through coordinated charging and distribution-network management.
Plug-in hybrid vehicle integration [32]	Microgrid load and plug-in hybrid electric vehicle data	Load shifting and curtailment strategy for clean and economic operation	Reduces operational cost and supports cleaner energy use by coordinating plug-in hybrid vehicle integration with microgrid load management and emission-aware energy operation.
Green layout optimization [33]	Industrial layout and sustainability data	Hybrid metaheuristic optimization for green layout design	Provides an optimization-based sustainability framework that can inform transport-facility design, vehicle movement efficiency, and emission reduction within industrial mobility environments.

methodology layer because hybrid mobility systems operate between conventional fuel use and grid-supported electrification. Their emissions depend on driving mode, battery availability, load shifting, charging strategy, and the balance between electric and fuel-powered operation. Clean and economic microgrid operation with plug-in hybrid electric vehicle integration is relevant because it shows how mobility energy demand can be coordinated with grid operation to reduce cost and improve environmental performance. This contributes to the methodology matrix by showing that vehicle emission forecasting should be flexible enough to represent transitional technologies, not only fully conventional or fully electric vehicles.

Green layout optimization contributes to the methodology review by addressing how spatial arrangement and facility design affect movement efficiency. In industrial environments, layout decisions influence internal transportation distance, material handling requirements, vehicle circulation, energy consumption, and operational emissions. Hybrid metaheuristic techniques for green layout design are therefore relevant because they show how optimization can reduce unnecessary movement by improving the physical organization of production and transport-related spaces. For vehicle CO_2 forecasting, this methodological perspective is useful because emissions can be affected before vehicles even begin operating, through the design of facilities, routes, loading zones, and internal logistics paths. This reference therefore expands the review from prediction at the vehicle level to environmental improvement through spatial and operational design.

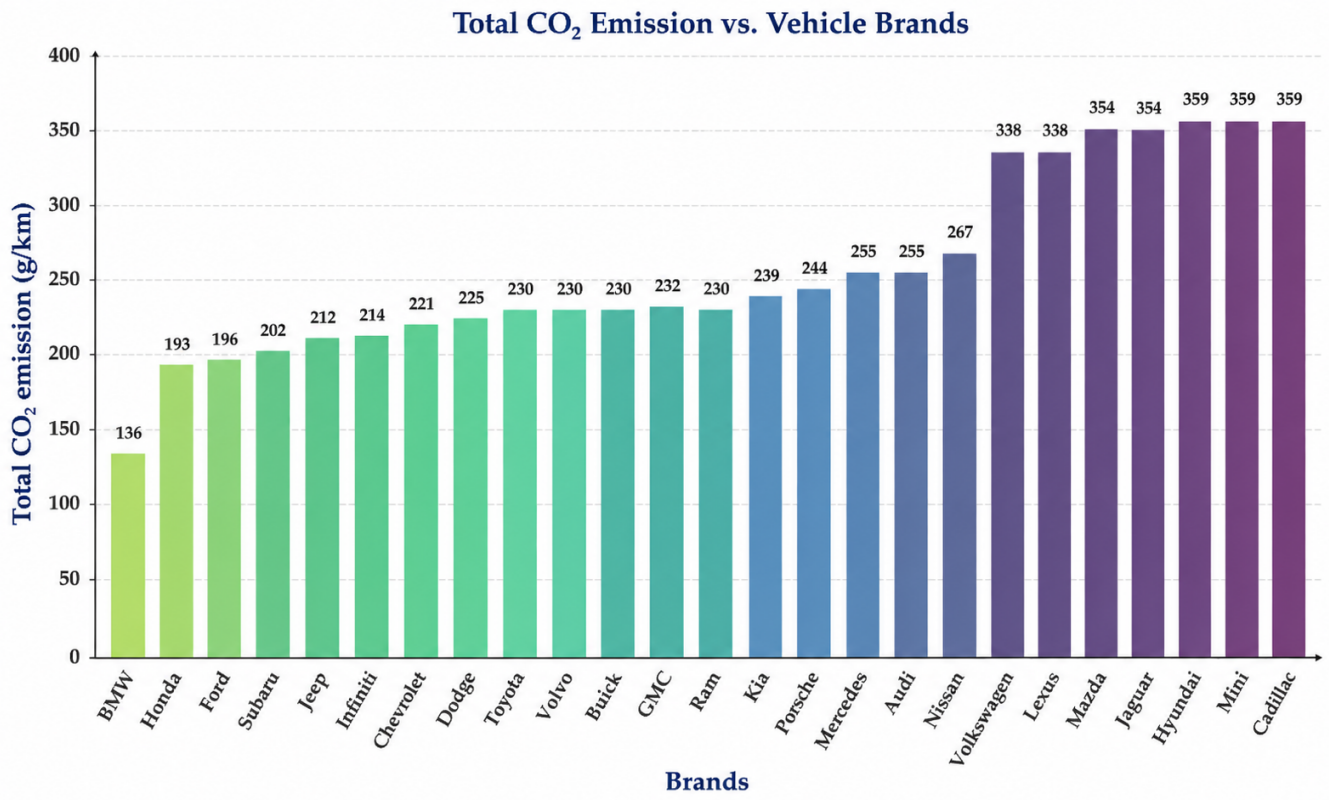
Taken together, the methodology references in Table 1 show that vehicle CO_2 emission forecasting should be treated as an integrated modelling problem that combines prediction, optimization, infrastructure intelligence, energy-system analysis, and sustainability-oriented decision support. Each reference contributes a distinct methodological role: robust fleet and routing models address uncertainty in vehicle deployment; optimizer-assisted neural estimation supports electric vehicle energy modelling; swarm intelligence strengthens nonlinear emission prediction; hybrid environmental regression improves pollution forecasting; distribution-network design connects logistics structure with greenhouse-gas reduction; multi-objective optimization captures competing transportation goals; pavement-performance prediction links infrastructure condition with mobility efficiency; sustainability-oriented artificial intelligence reviews guide methodological classification; microgrid and renewable-energy studies connect electrified transport with cleaner power systems; sensor-network optimization supports reliable data acquisition; and carbon-aware operational planning demonstrates how emissions can be reduced through coordinated decision-making. This synthesis justifies the use of a methodology matrix because the reviewed literature does not form a single narrow research line; instead, it forms a connected methodological ecosystem for developing accurate, scalable, and environmentally meaningful vehicle CO_2 forecasting frameworks.

2.1 Data Sources and Variable Representation in Vehicle CO_2 Forecasting

Vehicle CO_2 emission forecasting depends heavily on the quality, diversity, and representational capacity of the input data used to describe transportation activity. In this research direction, the dataset is not merely a technical input; it defines the boundary of what the forecasting model can learn about the emission-generation process. Vehicle-related carbon output may be influenced by traffic volume, travel speed, acceleration behavior, road type, fuel category, engine efficiency, vehicle age, payload condition, trip duration, and meteorological variation. When these variables are represented accurately, the forecasting model can capture a more realistic relationship between vehicle operation and emission behavior. However, when the dataset is incomplete or weakly structured, the model may learn only a partial view of the system, leading to predictions that appear numerically acceptable but remain limited in practical interpretation. Figure 2 presents the complete workflow for forecasting vehicle-related CO_2 emissions using artificial intelligence-based modelling. The process begins with the collection of heterogeneous transportation and environmental data, including vehicle characteristics, fuel type, speed, traffic volume, road conditions, route information, weather variables, and energy-consumption indicators. These inputs represent the operational conditions that directly or indirectly influence vehicle carbon emissions. After data collection, the preprocessing stage improves the quality and consistency of the dataset through cleaning, normalization, feature engineering, and missing-value handling. This step is essential because inaccurate or incomplete data can weaken the reliability of the forecasting model and reduce the interpretability of the predicted emission patterns.

The workflow then moves to feature selection, where the most relevant variables are identified and redundant or weakly informative inputs are reduced. This stage helps improve model efficiency by selecting an optimized subset of predictors that better represents the emission-generation process. Prediction models are then applied, including machine learning, deep learning, and hybrid models, depending on the structure of the dataset and the forecasting objective. Hyperparameter optimization and metaheuristic algorithms support the modelling process by tuning parameters, improving model configuration, and enhancing feature optimization. The trained models are evaluated using statistical indicators such as mean absolute error, root mean square error, coefficient of determination, and mean squared error. Finally, the forecasting outputs provide CO_2 emission prediction, trend analysis, and scenario forecasting, which can support sustainable transportation planning, traffic management, policy design, and emission-reduction strategies. The structure shown in Figure 2 emphasizes that vehicle CO_2 emission forecasting is not limited to the prediction model alone. Instead, it depends on a complete analytical pipeline in which data quality, variable selection, model design, optimization, and evaluation all contribute to the final forecasting performance. The dashed feedback paths indicate that optimization is iterative, meaning that model parameters and selected features may be refined repeatedly until the forecasting framework achieves stronger stability and predictive accuracy. This makes the diagram useful for explaining how artificial intelligence can transform raw vehicle and transportation data into practical decision-support outputs for low-carbon mobility planning.

A key issue in data representation is the difference between direct and indirect emission-related variables. Direct vari-



(a) Total CO_2 emission (g/km) for each vehicle brand.

Figure 2. Full-width workflow diagram for forecasting vehicle-related CO_2 emissions using artificial intelligence-based modelling, including data sources, preprocessing, feature selection, prediction models, hyperparameter optimization, evaluation metrics, forecasting outputs, and sustainable transportation applications.

ables may include measured or estimated CO_2 values, fuel-consumption records, or energy-use measurements, while indirect variables describe the conditions that influence emissions, such as congestion level, vehicle density, road gradient, stop frequency, and trip pattern. Both types of variables are important because direct indicators provide the target or near-target information, whereas indirect indicators explain why emission levels change across time and location. A strong forecasting framework should therefore avoid depending on a narrow group of variables and should instead represent the transportation system as a connected operational environment. This makes the model more capable of identifying emission changes caused by variations in mobility behavior, infrastructure conditions, and vehicle operation.

Another important consideration is temporal and spatial resolution. Vehicle emissions may be analyzed at different levels, including trip-level, road-segment-level, city-level, regional-level, or national-level scales. Each scale produces a different modelling requirement. A trip-level model may need detailed speed and acceleration sequences, while a city-level model may depend more on aggregated traffic demand, fuel use, fleet composition, and population mobility patterns. Similarly, short-term forecasting may require high-frequency traffic and sensor data, whereas long-term forecasting may rely on historical trends, policy variables, and socioeconomic indicators. Therefore, the literature on vehicle CO_2 forecasting should be understood through the type of data resolution adopted by each study, because resolution directly affects model design, prediction horizon, and the practical meaning of the forecasted values.

2.2 Model Design Logic for Artificial Intelligence-Based Emission Prediction

The design of an artificial intelligence-based vehicle CO_2 forecasting model involves more than selecting a machine learning or deep learning algorithm. A complete model design process begins with defining the prediction target, identifying the relevant input variables, preparing the dataset, choosing the learning structure, optimizing the model configuration, validating the results, and interpreting the output. Each stage affects the reliability of the final forecast. If the prediction target is unclear, the model may mix different emission concepts, such as direct tailpipe emissions, energy-related emissions, or total transport-sector carbon output. If the input variables are not properly selected, the model may include redundant or irrelevant predictors that increase complexity without improving forecasting quality.

Figure 3 presents a full-width conceptual framework for forecasting vehicle-related CO_2 emissions using five artificial intelligence-based modelling approaches. The diagram begins with the data-collection stage, where several input variables are gathered to describe the operational and environmental conditions that influence vehicle emissions. These variables include vehicle characteristics, fuel type, driving behaviour, traffic conditions, road and route information, weather conditions, engine and technical specifications, and historical CO_2 emission records. This stage is important because the accuracy of emission forecasting depends strongly on how well the collected data represent the real factors that control fuel consumption, energy use, and carbon output. The

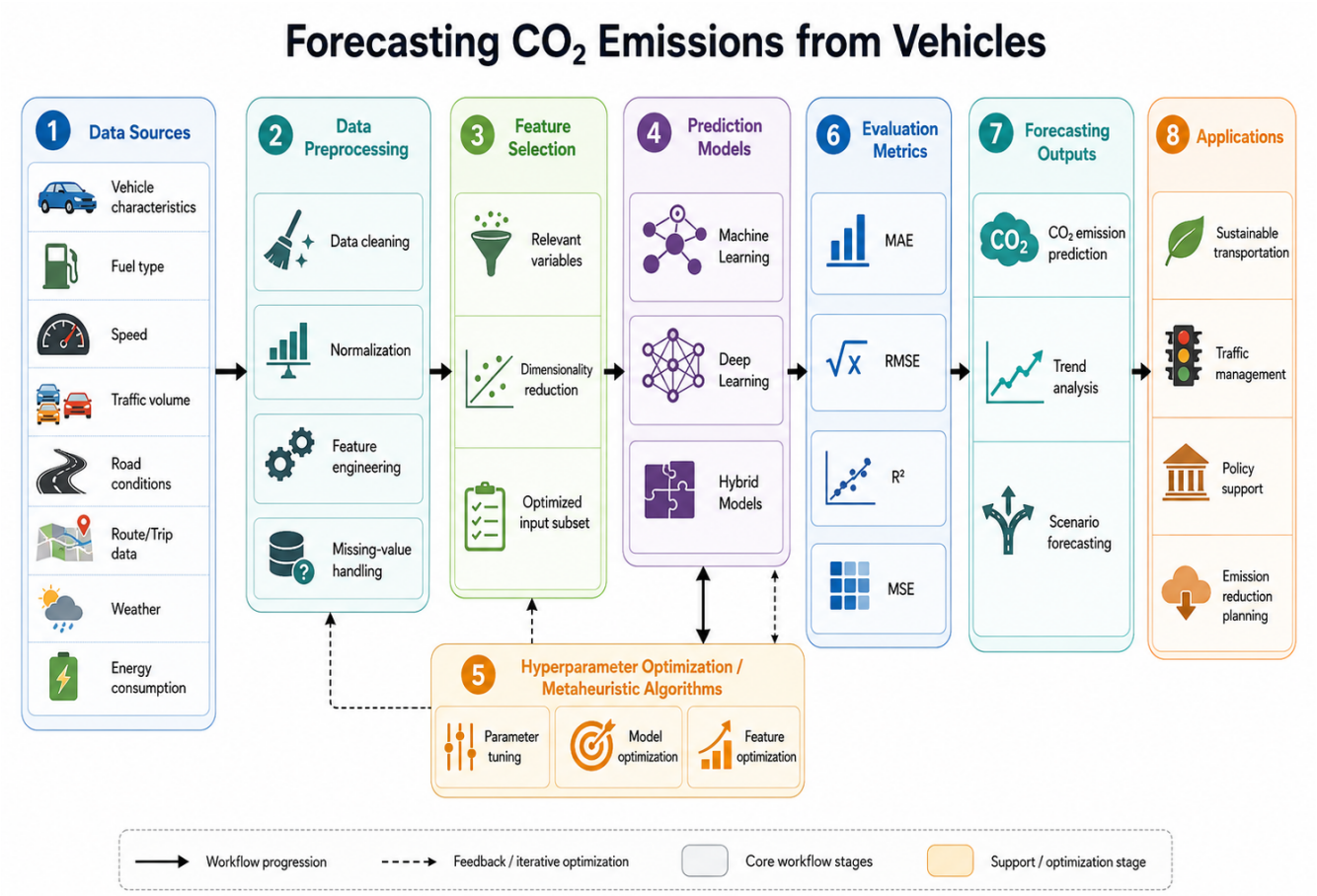


Figure 3. Full-width diagram of vehicle CO_2 emission forecasting using five artificial intelligence models, including data collection, preprocessing, feature selection, predictive modelling, model evaluation, forecasting outputs, and sustainable transportation applications.

workflow shown in Figure 3 continues with data preprocessing and feature selection, which prepare the raw transportation data for reliable model training. Data preprocessing includes cleaning missing or abnormal values, normalizing numerical variables, engineering useful predictors, encoding categorical information, and dividing the dataset into training, validation, and testing subsets. Feature selection then identifies the most influential variables by removing redundant or weakly informative inputs, which helps reduce dimensionality and improve forecasting stability. After these preparation stages, five predictive models are presented: Multiple Linear Regression (MLR), Support Vector Regression (SVR), Random Forest Regression (RFR), Artificial Neural Network (ANN), and Long Short-Term Memory (LSTM). MLR provides a simple statistical baseline for modelling linear relationships, SVR captures nonlinear regression patterns through margin-based learning, RFR improves robustness through ensemble decision trees, ANN learns complex nonlinear interactions between variables, and LSTM is suitable for time-series emission forecasting because it can model long-term temporal dependencies.

The lower part of Figure 3 highlights the evaluation and application stages of the forecasting framework. Model performance is assessed using several statistical indicators, including mean absolute error, root mean square error, coefficient of determination, mean absolute percentage error, and Nash–Sutcliffe efficiency. These metrics help determine how accurately each model predicts vehicle CO_2 emissions and whether the predicted values are reliable for practical interpretation. The final outputs include emission prediction, trend analysis, scenario forecasting, emission-reduction strategies, and policy decision support. Therefore, the diagram shows that artificial intelligence-based vehicle emission forecasting is not limited to producing numerical predictions; it also supports sustainable transportation planning, environmental impact reduction, improved understanding of emission patterns, and data-driven policy development.

Machine learning models are generally suitable when the dataset is structured, tabular, and moderate in size. These models can learn nonlinear relationships between explanatory variables and emission values while maintaining a relatively manageable computational cost. They are often useful when the forecasting task depends on variables such as vehicle type, fuel use, traffic density, speed, and travel distance. Deep learning models, on the other hand, become more relevant when the dataset contains sequential, image-based, high-dimensional, or large-scale sensor information. In vehicle emission forecasting, deep learning may be useful for traffic time-series prediction, energy-consumption sequence modelling, and dynamic mobility-pattern analysis. The distinction between these model families should not be treated as a simple competition; instead, the appropriate model should be selected according to the structure of the dataset and the purpose of the forecasting task.

A strong model design should also consider the balance be-

tween accuracy, interpretability, and deployability. Highly complex models may achieve strong numerical performance, but their practical value can be limited if their internal decision process cannot be explained to transportation planners or environmental authorities. Conversely, simpler models may be easier to interpret but may not fully capture nonlinear emission behavior under complex traffic conditions. For this reason, vehicle CO_2 forecasting research should not evaluate models only through error metrics. It should also consider whether the model can support decision-making, whether its predictions can be explained, whether it can operate under realistic data conditions, and whether it can be updated when transportation patterns change. This design logic is essential for transforming artificial intelligence-based forecasting from an experimental exercise into a usable environmental decision-support tool.

2.3 Optimization, Validation, and Practical Deployment Considerations

Optimization plays an important role in vehicle CO_2 emission forecasting because predictive performance is strongly affected by model configuration, feature selection, and parameter tuning. Many artificial intelligence models contain adjustable settings that control their learning behavior, such as tree depth, number of estimators, learning rate, hidden layers, activation functions, kernel parameters, population size, or convergence criteria. If these settings are selected randomly or manually without systematic evaluation, the model may fail to reach its best forecasting capacity. Optimization methods help overcome this limitation by searching for better configurations within a large and complex solution space. In emission forecasting, this is especially important because the relationship between vehicle activity and carbon output is rarely linear or simple.

Validation is equally important because an optimized model is not necessarily a reliable model unless its performance is tested under appropriate conditions. A model may perform well on training data but fail when exposed to new traffic patterns, different vehicle fleets, unusual congestion conditions, or unseen regional characteristics. Therefore, a rigorous validation process should examine the model's ability to generalize beyond the data used during training. This requires careful separation between training and testing data, suitable performance metrics, and attention to overfitting. In vehicle CO_2 forecasting, validation should also consider whether the model can predict both normal emission behavior and high-emission situations, since emission peaks may be more important for policy intervention than average values alone.

Practical deployment introduces another layer of methodological complexity. A forecasting model may be academically strong but difficult to apply in real transportation systems if it requires unavailable data, excessive computational resources, or continuous manual adjustment. For operational use, the model should be compatible with existing traffic-monitoring platforms, fleet-management systems, environmental dashboards, or urban decision-support tools. It should also be capable of handling missing data, updating predictions as new information becomes available, and communicating uncertainty in a way that decision-makers can understand. These deployment considerations show that the success of vehicle CO_2 forecasting depends not only on algorithmic accuracy but also on the ability to integrate artificial intelligence into sustainable transportation planning, carbon-management strategies, and real-world mobility governance.

3. DISCUSSION

The reviewed body of work indicates that vehicle CO_2 emission forecasting is developing into a distinct research direction situated at the intersection of environmental modeling, intelligent transportation systems, optimization science, and sustainable mobility planning. The significance of this field lies not only in estimating future emission values but also in converting transportation data into strategic knowledge that can guide low-carbon decisions before environmental pressure becomes irreversible. Unlike general carbon accounting, which usually describes emissions after they have occurred, forecasting creates a forward-looking analytical layer through which transport authorities, fleet managers, energy planners, and urban policymakers can anticipate emission growth, evaluate intervention scenarios, and align mobility systems with decarbonization objectives. This shift gives artificial intelligence a central role because the emission behavior of vehicles is shaped by nonlinear, time-dependent, and context-sensitive interactions that cannot be fully captured through simplified deterministic assumptions.

A major implication emerging from the reviewed studies is that vehicle CO_2 emission forecasting should be treated as a system-level modeling challenge rather than a narrow algorithm-selection problem. In many transportation studies, the predictive model is placed at the center of the methodological design, while the surrounding components are treated as secondary. However, the quality of an emission forecast depends on the complete analytical chain: how the target variable is defined, how the input space is represented, how operational conditions are encoded, how model parameters are tuned, how prediction uncertainty is handled, and how results are interpreted for decision-making. A model with strong numerical accuracy may still have limited value if its input variables do not reflect the actual emission-generation mechanisms, if its evaluation setting is too restricted, or if its predictions cannot be translated into transport-management actions. Therefore, the discussion of artificial intelligence-based forecasting must extend beyond model comparison and examine the coherence of the entire forecasting pipeline.

The methodological matrix presented in Table 1 shows that the reviewed literature does not form a single homogeneous research stream. Instead, it forms a broad methodological ecosystem in which each group of studies contributes a different layer to emission-oriented transportation intelligence. Some studies strengthen the predictive layer by improving model learning and estimation capacity, while others contribute to the optimization layer by supporting carbon-aware routing, fleet planning, or energy scheduling. A separate group of studies contributes to the infrastructure layer by improving sensing coverage, road-condition assessment, and transport-system observability. Another group contributes to the energy-transition layer by connecting electric mobility with renewable generation, microgrid planning, and charging infrastructure. This diversity matters because vehicle CO_2 emissions are shaped by operational, infrastructural, and

energy-related conditions that the model must represent. Consequently, the reviewed methodologies should be interpreted as complementary rather than competing research directions.

One of the central insights of the review is that artificial intelligence can improve vehicle emission forecasting only when it is embedded within a meaningful transportation context. Machine learning and deep learning models can identify patterns in large datasets, but these patterns must be grounded in the real behavior of vehicles, roads, fleets, and energy systems. If the modeling process ignores transportation logic, the resulting forecast may become a statistically optimized output with weak environmental interpretation. For example, an emission model may identify correlations between temporal indicators and carbon output, but without operational interpretation, it may not explain whether the emission increase is associated with congestion, inefficient routing, high freight demand, vehicle aging, or energy-system dependency. This distinction matters because sustainable transportation planning requires not only knowing that emissions may rise, but also understanding which controllable mechanism is responsible for that rise.

The review also highlights the importance of optimization as a structural component of vehicle CO_2 forecasting. Optimization methods are often introduced to improve predictive accuracy through feature selection or hyperparameter tuning, but their value extends further. They can support the design of lower-emission transport operations by identifying better routing strategies, more efficient fleet allocations, improved charging schedules, and more sustainable infrastructure configurations. This dual role makes optimization a bridge between prediction and action. Forecasting estimates what may happen, while optimization helps determine what should be changed to reduce the predicted environmental burden. In this sense, the strongest future frameworks will likely be those that do not isolate forecasting from decision-making, but instead integrate prediction, evaluation, and intervention within one coherent modeling environment.

Another important discussion point concerns the changing meaning of vehicle emissions under transport electrification. Conventional vehicle emission forecasting has often been associated with fuel consumption and direct exhaust output, but this view becomes incomplete when electric and hybrid vehicles are increasingly present in transportation systems. For electric mobility, the environmental impact shifts from tailpipe emissions to electricity generation, charging behavior, grid efficiency, renewable-energy availability, and battery-management performance. This does not reduce the importance of vehicle emission forecasting; rather, it expands its scope. Future forecasting systems must distinguish between direct vehicle emissions, indirect energy-related emissions, and lifecycle-sensitive mobility impacts. Without this distinction, a model may underestimate the carbon consequences of electrified transportation in regions where electricity generation remains carbon-intensive, or it may fail to recognize the environmental benefits of renewable-supported charging systems.

The discussion further reveals that generalization remains a major challenge for artificial intelligence-based vehicle CO_2 forecasting. A model trained on one dataset may learn relationships that are specific to a certain city, vehicle fleet, road network, traffic culture, climate condition, or energy mix. When such a model is transferred to another context, its predictive behavior may change substantially. This issue is especially important for review-based synthesis because reported model performance cannot be interpreted without considering the setting in which the model was trained and tested. High performance in one case study does not automatically imply universal applicability. Therefore, future studies should place greater emphasis on cross-regional evaluation, external validation, and adaptive model updating. Such practices would help move the field from isolated experimental forecasting toward more generalizable and transferable emission-intelligence systems.

The reviewed literature also suggests that model interpretability should be treated as a practical requirement rather than an optional enhancement. Vehicle CO_2 forecasts may influence public investment, environmental regulation, fleet policy, infrastructure planning, and urban mobility governance. These decisions require a level of transparency that allows stakeholders to understand the main factors behind predicted emission patterns. If an artificial intelligence model functions as an opaque system, its predictions may be difficult to trust, audit, or justify in policy settings. Interpretability can help identify whether emission increases are associated with vehicle characteristics, road conditions, mobility intensity, energy use, or planning inefficiencies. This explanatory capacity is particularly valuable because it enables targeted mitigation measures instead of broad interventions that may be costly or ineffective.

A further implication concerns the relationship between forecasting accuracy and real-world deployment. Many studies emphasize numerical performance, but operational use requires additional qualities such as computational stability, data-stream compatibility, robustness to missing information, adaptability to new traffic conditions, and clear communication of prediction uncertainty. A forecasting model that performs well under controlled experimental conditions may still face obstacles when integrated into municipal dashboards, fleet-management platforms, traffic-control systems, or environmental monitoring tools. This gap between experimental modeling and deployment readiness must be addressed if artificial intelligence-based vehicle CO_2 forecasting is to become useful for practical low-carbon transportation governance. Future research should therefore evaluate models not only by prediction error, but also by their operational feasibility within real decision-support environments.

The reviewed studies also show that vehicle CO_2 forecasting can support a broader transition from descriptive sustainability assessment to proactive carbon governance. Descriptive assessment helps determine how much emission has occurred, while proactive forecasting helps anticipate where and when emissions may increase. This difference is crucial because transport systems are difficult to transform quickly once inefficient patterns become embedded in infrastructure, fleet composition, and travel behavior. Forecasting can give decision-makers an earlier opportunity to evaluate emission risks, compare mitigation strategies, and prioritize interventions. When combined with optimization, forecasting can also support scenario-based planning, where different policy or operational alternatives are examined before implementa-

tion. This makes artificial intelligence not only a modeling tool, but also a mechanism for strengthening environmental anticipation in transportation systems.

Despite the strong potential of artificial intelligence, the reviewed literature indicates that methodological caution remains necessary. High-performing models may be affected by overfitting, limited dataset diversity, narrow prediction horizons, weak benchmarking, or insufficient explanation of input-output relationships. Furthermore, emission forecasting is sensitive to data reliability because errors in traffic measurement, fuel-consumption estimation, energy-use recording, or fleet classification can propagate through the model and affect the final forecast. This means that artificial intelligence should not be presented as a substitute for rigorous data governance. Instead, its value depends on disciplined dataset construction, transparent evaluation, reproducible modeling procedures, and careful interpretation of results. A forecasting framework is only as reliable as the assumptions, data structures, and validation practices that support it.

The main contribution of the reviewed methodological landscape is that it demonstrates the need for integrated emission-forecasting frameworks that combine analytical accuracy with environmental usefulness. A strong vehicle CO_2 forecasting system should not only predict future values, but also help explain emission drivers, identify controllable factors, support low-carbon interventions, and remain adaptable to changing transport technologies. Such a framework must be able to operate across conventional vehicles, hybrid mobility, electric transportation, freight systems, urban traffic, infrastructure conditions, and energy-network interactions. This integrated perspective is essential because the future of transportation emissions will be shaped by multiple transitions occurring simultaneously: digitalization, electrification, automation, urban expansion, renewable-energy integration, and stricter climate policies.

Overall, the discussion confirms that artificial intelligence-based vehicle CO_2 emission forecasting is an emerging and necessary research area for sustainable transportation systems. Its scientific value comes from its ability to model complex emission behavior, while its practical value comes from its potential to guide earlier and more informed decisions. However, the field must continue to improve its methodological transparency, transferability, validation strength, and deployment orientation. The most useful future models will be those that combine predictive performance with explanatory insight, optimization capability, and direct relevance to low-carbon mobility planning. In this sense, vehicle CO_2 forecasting should be viewed as a decision-support discipline that connects artificial intelligence research with climate-responsible transportation governance.

4. CONCLUSION

This review examined artificial intelligence-based forecasting of vehicle CO_2 emissions as a multidisciplinary research area that connects machine learning, deep learning, optimization, transportation planning, energy management, and environmental decision support. The reviewed literature shows that vehicle emission forecasting is no longer limited to static estimation or retrospective carbon accounting. Instead, it is increasingly moving toward intelligent predictive frameworks capable of analyzing complex relationships between mobility behavior, vehicle operation, infrastructure conditions, fleet decisions, and energy-system characteristics. This transition is important because modern transport systems require tools that can anticipate emission patterns before they become environmental and policy challenges.

The review demonstrates that artificial intelligence provides strong methodological value for vehicle CO_2 forecasting because it can capture nonlinear and dynamic relationships within heterogeneous transportation data. However, the effectiveness of artificial intelligence depends on the quality of the complete modeling process rather than the algorithm alone. Reliable forecasting requires meaningful variable representation, careful data preparation, appropriate model design, systematic optimization, rigorous validation, and interpretable outputs. These requirements show that emission forecasting should be developed as a structured analytical pipeline rather than a simple model-fitting task. When these elements are properly integrated, artificial intelligence can provide forecasts that are not only accurate but also useful for sustainable transportation planning.

The methodological evidence synthesized in Table 1 confirms that vehicle CO_2 forecasting is supported by several complementary research directions. Predictive modeling contributes to estimating future emission behavior; optimization methods support model refinement and low-carbon decision-making; routing and fleet-management studies connect forecasting with operational control; infrastructure-related prediction expands the understanding of road-system effects; and energy-system optimization links transport electrification with indirect carbon impacts. This broad methodological foundation shows that vehicle emissions cannot be understood through a single technical lens. Instead, they must be analyzed through an integrated perspective that reflects the operational, environmental, and energy-related realities of transportation systems.

The review also concludes that future vehicle CO_2 forecasting research should place stronger emphasis on generalization, explainability, and deployment readiness. Many forecasting models can achieve promising results within specific datasets, but their value becomes limited if they cannot be transferred across locations, adapted to changing vehicle technologies, or interpreted by decision-makers. Future research should therefore move toward broader benchmarking, external validation, uncertainty-aware forecasting, and transparent reporting of modeling assumptions. These improvements are necessary for transforming artificial intelligence-based emission prediction from a laboratory-oriented modeling activity into a practical tool for cities, transport agencies, logistics operators, and environmental authorities.

Another important conclusion is that electrification will reshape the structure of vehicle emission forecasting. As electric and hybrid vehicles become more common, forecasting frameworks must move beyond direct exhaust-based assumptions and include energy-consumption behavior, charging infrastructure, electricity-generation carbon intensity, and renewable-energy integration. This broader view is essential because the environmental impact of future mobility will depend not only on vehicles themselves but also on the energy systems that support them. Therefore, the next generation of forecasting models should be capable of representing mixed

fleets, changing propulsion technologies, and the interaction between transportation demand and power-system planning.

In conclusion, artificial intelligence-based forecasting of vehicle CO_2 emissions has strong potential to support low-carbon mobility, sustainable urban development, and climate-responsive transportation governance. Its contribution lies in enabling earlier recognition of emission trends, a deeper understanding of transport-related carbon drivers, and a more informed design of mitigation strategies. Nevertheless, its successful application requires responsible modeling practices, transparent validation, high-quality data, and integration with real-world decision-support systems. Future progress in this field will depend on developing forecasting frameworks that are accurate, interpretable, scalable, and environmentally meaningful, allowing artificial intelligence to serve not merely as a predictive technology but as an essential component of sustainable transportation transformation.

REFERENCES

[1] Amir Mosavi, Mobin Salimi, Sina Faizollahzadeh Ardabili, Timon Rabczuk, Shahaboddin Shamshirband, and Annamaria R. Varkonyi-Koczy. A review of machine learning and deep learning applications in energy forecasting. *Renewable and Sustainable Energy Reviews*, 110:24–47, 2020.

[2] Emrah Demir, Wolfgang Burgholzer, Martin Hrušovský, Emrah Arıkan, Werner Jammernegg, and Tom Van Woensel. Freight transportation planning and logistics management for sustainable supply chains: a review. *European Journal of Operational Research*, 282(2):401–418, 2020.

[3] M. S. Hossain Lipu, M. A. Hannan, Aini Hussain, M. M. Hoque, Pin Jern Ker, M. H. M. Saad, and Afida Ayob. Machine learning for advanced battery management in electric vehicles: a comprehensive review. *IEEE Access*, 8:175453–175477, 2020.

[4] Michel Gendreau, Ola Jabali, and Walter Rei. Time-dependent vehicle routing and scheduling with environmental considerations: a review. *European Journal of Operational Research*, 285(3):881–900, 2020.

[5] Johannes Kester, Lance Noel, Gerardo Zarazua de Rubens, and Benjamin K. Sovacool. Electric vehicles and vehicle-to-grid technology in smart grids: a review. *Journal of Power Sources*, 456:227980, 2020.

[6] Bing Yu, Haoteng Yin, and Zhanxing Zhu. Spatial-temporal graph convolutional networks for traffic flow forecasting. *IEEE Transactions on Intelligent Transportation Systems*, 21(2):565–574, 2020.

[7] Todd Litman. Autonomous vehicle implementation predictions: implications for transport planning. *Victoria Transport Policy Institute*, pages 1–42, 2020.

[8] Emrah Demir, Tolga Bektaş, and Gilbert Laporte. The pollution-routing problem: a review of models and solution algorithms. *Transportation Research Part B: Methodological*, 141:112–134, 2020.

[9] Soteris A. Kalogirou, Stylianos Pashiardis, and Anna Pashiardi. Artificial neural networks and metaheuristic optimization for renewable energy prediction: a review. *Renewable Energy*, 146:1513–1528, 2020.

[10] Mohammad Asghari and Seyed M. J. M. Al-e hashem. Green vehicle routing and scheduling: a review and future research directions. *Transportation Research Part E: Logistics and Transportation Review*, 136:101896, 2020.

[11] Eyup Dogan and Roula Inglesi-Lotz. Forecasting carbon dioxide emissions using optimized grey and machine learning models. *Energy*, 196:117094, 2020.

[12] Goodness C. Aye, Prosper E. Edoja, and Lanouar Charfeddine. Using machine learning to predict carbon dioxide emissions from energy and transport indicators. *Energy Reports*, 6:2730–2743, 2020.

[13] Bing Zhu, Shichao Ye, Peng Wang, Kaijian He, Tiemin Zhang, and Yi-Ming Wei. Carbon price forecasting using decomposition and machine learning methods. *Applied Energy*, 262:114500, 2020.

[14] Qingfu Zhang, Shengxiang Yang, and Yaochu Jin. Evolutionary algorithms for machine learning and data mining: recent advances and applications. *Soft Computing*, 24(9):6323–6342, 2020.

[15] Hsiao-Tien Pao, Hsin-Chia Fu, and Chih-Liang Tseng. Forecasting carbon dioxide emissions from energy consumption using grey and machine learning models. *Energy Policy*, 137:111154, 2020.

[16] Simona Mancini and Margaretha Gansterer. Robust and sustainable vehicle routing under demand uncertainty. *IIE Transactions*, 53(9):1012–1028, 2021.

[17] D. N. T. How, M. A. Hannan, M. S. H. Lipu, and Pin Jern Ker. State of charge estimation for lithium-ion batteries in electric vehicles using machine learning methods. *Journal of Power Sources*, 479:228867, 2020.

[18] Alex O. Acheampong, Evans Boateng, and Mark Amponsah. Modelling and forecasting energy-related carbon dioxide emissions using machine learning approaches. *Energy*, 212:118784, 2020.

[19] Yun Bai, Biao Zeng, Chuan Li, and Jian Zhang. Deep learning for air quality forecasting: a review. *Atmospheric Environment*, 223:117240, 2020.

[20] Kannan Govindan, Ahmad Jafarian, and Vahid Nourbakhsh. Sustainable supply chain network design: a review of modelling approaches and future directions. *Omega*, 93:102007, 2020.

[21] Reza Moghdani, Khodakaram Salimifard, Emrah Demir, and Abdelkader Benyettou. Multi-objective optimization for sustainable routing and scheduling problems: recent advances. *European Journal of Operational Research*, 287(1):1–18, 2020.

[22] Yonas Zewdu Ayele, Abbas Barabadi, and Enrique Lopez Droguett. Machine learning approaches for pavement condition prediction: a review. *International Journal of Pavement Engineering*, 21(12):1524–1538, 2020.

[23] Fei Deng, Yi He, Shun Zhou, Yun Yu, Hui Cheng, and Xiangyu Wu. Machine learning prediction of concrete properties and structural-material performance. *Construction and Building Materials*, 263:120198, 2020.

[24] Esther Mengelkamp, Johannes Gärttner, Kerstin Rock, Scott Kessler, Lawrence Orsini, and Christof Weinhardt. Energy management and optimization of microgrid systems: a review. *Renewable and Sustainable Energy Reviews*, 120:109623, 2020.

[25] Ibrahim Dincer and Canan Acar. Renewable energy systems and energy efficiency for sustainable transportation: a review. *Renewable and Sustainable Energy Reviews*, 132:110103, 2020.

[26] Tomiwa Sunday Adebayo and Dervis Kirikkaleli. Predicting carbon dioxide emissions using artificial intelligence and machine learning techniques. *Energy Policy*, 145:111732, 2020.

[27] Jin Wang, Yu Gao, Wei Liu, Arun Kumar Sangaiah, and Hye-Jin Kim. Coverage optimization in wireless sensor networks using metaheuristic algorithms: a review. *Computer Communications*, 152:163–178, 2020.

[28] Ahmed Fathy, Hegazy Rezk, and Dalia Yousri. Optimal sizing of renewable hybrid energy systems with battery storage using metaheuristic optimization. *Solar Energy*, 207:1007–1023, 2020.

[29] Karel Kellens, Wim Dewulf, Michael Overcash, Michael Z. Hauschild, and Joost R. Duflou. Carbon emission modelling and energy-efficiency assessment for sustainable manufacturing processes. *CIRP Annals*, 69(2):653–676, 2020.

[30] A. Moradi Afrapoli and H. Askari-Nasab. Truck dispatching and fleet management in mining operations: models, algorithms and sustainability aspects. *International Journal of Mining, Reclamation and Environment*, 34(8):529–547, 2020.

[31] Sanchari Deb, Kari Tammi, Karuna Kalita, and Pinakeswar Mahanta. Optimal planning of electric vehicle charging stations in distribution systems: a review. *International Journal of Electrical Power & Energy Systems*, 118:105747, 2020.

[32] Akhtar Hussain, Van-Hai Bui, and Hak-Man Kim. Optimal operation of microgrids with electric vehicle charging and renewable energy resources. *IEEE Transactions on Smart Grid*, 11(1):384–395, 2020.

[33] Amine Drira, Henri Pierreval, and Sonia Hajri-Gabouj. Facility layout planning and optimization using metaheuristic algorithms: a review. *IIE Transactions*, 52(6):657–676, 2020.