RAIRO-Oper. Res. 59 (2025) 549–578
https://doi.org/10.1051/ro/2024162

RAIRO Operations Research
www.rairo-ro.org

SUSTAINABLE SOLUTIONS ANALYSIS OF A BI-OBJECTIVE GREEN INVENTORY ROUTING PROBLEM WITH HETEROGENEOUS FLEET AND DIFFERENT TYPES OF FUELS

Arianne A.S. Mundim[1,*], Maristela O. Santos[1] and Reinaldo Morabito[2]

Abstract. One of the main agents responsible for global warming is greenhouse gases, especially carbon dioxide (CO_2) associated with fuel combustion. Most works in the literature address logistics transportation from an economic perspective, giving little attention to the existing trade-off with sustainability. In this work, we develop a bi-objective approach to the inventory routing problem with heterogeneous fleet, where we minimize costs while simultaneously reducing CO_2 emissions. First, we present an explicit vehicular equation developed to calculate CO_2 emissions for different types of vehicles and fuels. We demonstrate that this equation is statistically precise by conducting a study with a database in which machine learning techniques were applied to assess the predictive accuracy of CO_2 emissions. The comparison between the explicit equation and machine learning models proves its efficacy as a suitable approximation for practical applications. Then, we propose an augmented ϵ-constrained method to find the efficient Pareto frontier using a branch-and-cut method. Computational experiments were conducted on 285 instances, of which 125 were adapted from the literature, solving the augmented ϵ-constrained optimally. Result analysis indicates the ability of the approach to trade off between economy and sustainability, where, on average, lexicographic solutions show a 58% reduction in emissions and a 36% increase in costs. We conclude with a managerial analysis providing insights into the proposed approach, highlighting the advantages of using different vehicles and fuels.

Mathematics Subject Classification. 90B06, 90B50, 90C11, 90C29.

Received November 22, 2023. Accepted August 15, 2024.

1. Introduction

The Inventory Routing Problem (IRP) is one of the most extensively studied problems in the literature, beginning with the pioneering work of [15]. Hundreds of works consider the minimization of transportation and inventory costs, as highlighted in the following review papers [7, 17, 23, 24, 55]. Despite IRP being studied since the 1980's, according to [65], the number of works addressing sustainable issues in the supply chain has increased only in recent years.

Keywords. Multi-objective green inventory routing problem, carbon dioxide emission, augmented ϵ-constrained method, heterogeneous fleet, different types of fuels.

[1] University of Sao Paulo, Mathematics and Computer Sciences Dept., Sao Paulo, Brazil.

[2] Federal University of Sao Carlos, Production Engineering Dept., Sao Paulo, Brazil.

[*]Corresponding author: arianne@alumni.usp.br

One reason for the growing interest in sustainable works is due to global warming, and the situation is now at another level because "global boiling" is an expression used by the Secretary-General of the United Nations (UN) to describe the current phase of accelerated global warming and climate change.

This issue has emerged as one of the greatest challenges of our era, as emphasized by [11,37,69]. Measurements taken at the Earth's surface indicate an average increase of 0.74 °C over the past century. Projections suggest that if carbon emissions remain unchanged, average temperatures could rise by up to 3.4 °C by the end of this century, as stated by [26, 72]. One of the primary drivers of global warming is greenhouse gases, with approximately 73% of emissions composed of CO_2, according to [76]. While global carbon emissions dropped significantly (−8.8%) in the first half of 2020 due to the pandemic's effects [44], the most recent data show a strong rebound, and carbon emissions are approaching 2019 levels [67].

The increase in CO_2 emissions has raised concerns related to the topic of sustainable development. These concerns have driven a growing interest in alternative approaches to logistics since transportation plays a fundamental role in the context of sustainable development, as discussed in the studies by [28, 43]. Different aspects of gas mitigation can be observed in the works [32, 75].

Sustainable development has been incorporated into supply chain management [61] and logistics systems [19]. This is due to the significant impact that logistics and supply chain activities have on economic growth and competitiveness [46]. Additionally, it plays a crucial role in managing non-renewable resources and mitigating emissions of gases [63].

In the supply chain, the transportation sector is the largest contributor to carbon dioxide emissions [29]. The Intergovernmental Panel on Climate Change reported that transportation was responsible for 14% of greenhouse gas emissions by economic sectors in 2010 [53].

In this study, we explore the concept of sustainable IRP, also known as green IRP, which simultaneously addresses inventory management decisions, vehicle routing, and product delivery scheduling from an environmentally responsible perspective. Our goal is to minimize CO_2 emissions while maintaining economic viability.

The IRP was chosen over traditional Vehicle Routing Problem (VRP) models due to its potential for significant reductions in CO_2 emissions through integrated inventory management and delivery scheduling. This choice is particularly effective in scenarios where inventory costs are not prohibitive, allowing for strategic stockpiling that reduces frequent shipments, thereby lowering overall emissions.

Previous references, such as [3, 4, 21, 64], have dealt with these issues. However, these studies simplify the problem by combining different objectives into a single objective function, limiting the discussion of conflicting solutions in economic and environmental terms.

An overlooked area in the green IRP literature is the use of multiple fuel types and various modes of transportation, such as rail, waterway, and others. While [21] address green IRP with a heterogeneous fleet, considering only one type of fuel and using a nonlinear equation that incorporates variable speed, Xiao *et al.* [74] also consider a single type of fuel, with emissions depending on the vehicle's load state, empty or full. This gap in the literature deserves attention due to the complexity of green IRP with multiple fuel types and a heterogeneous fleet in terms of capacity, where different modes can be considered.

The formulations proposed in [21, 74] are currently used to estimate CO_2 emissions; however, in addition to considering only one type of fuel, they lack an assessment of the associated error in this estimate. Cheng *et al.* [21] use the formulation from [13], where the authors themselves state that the model needs to be periodically updated to adequately represent current vehicles in any fleet. Aware of these limitations, we present an explicit equation that can be applied to different types of vehicles and fuels, and conduct an error evaluation study using a test set of several thousand vehicles produced between 1995 and 2022. To validate the effectiveness of our approach, we compare the results obtained with widely used Machine Learning methods in the specialized literature. Our results reveal that our approach has high explanatory power and performs comparably to Machine Learning methods, highlighting its utility and reliability.

Our research aims to address the highlighted gaps by introducing a multi-objective approach and presenting an explicit vehicular equation applicable to transportation problems. Below, we outline the main contributions of this study:

(1) Gathering data and information to clearly describe a mathematical expression that estimates CO_2 emissions based on vehicle efficiency, CO_2 emitted per fuel consumption, and distance traveled. The proposed explicit vehicular equation is applicable to five different types of fuels, making it adaptable to a wide range of routing problems.
(2) Computational analysis of the explicit vehicular equation, indicating an explanatory power above 99%, ensuring an average error of 0.98%. Additionally, implementation and comparison with seven Machine Learning methods.
(3) Presentation of a sustainable bi-objective model that aims to minimize both economic and sustainability compromises for a heterogeneous fleet.
(4) Application of an exact scalarization method, called the augmented ϵ-constrained, to estimate the Pareto frontier of the proposed problem, providing adaptations of well-known instances from the literature.
(5) Analysis of computational results providing an understanding of the method's effectiveness and managerial insights on the use of multiple fuels and analysis of Pareto frontier solutions.

This paper is organized as follows. The literature review is conducted in Section 2. The presentation of the Explicit Equation for estimating CO_2 emissions is in Section 3. In Section 4, we present a mathematical modeling of the bi-objective green IRP and a solution method. In Section 5, computational experiments are presented. The conclusions of the work and prospects are outlined in Section 6.

2. Literature review

In this section, we organize the main studies that underpin the development of our work. In Section 2.1, we present the formulations employed for estimating CO_2 emissions. In Section 2.2, we provide an overview of studies investigating sustainable approaches in logistics problems. In Section 2.3, we discuss works that consider green IRP with conflicting objectives. Finally, we present a tabular review of studies related to our research.

2.1. Numerical alternatives for calculating CO_2 emissions

There are different numerical alternatives for estimating CO_2 emissions. For example, Xiao *et al.* [74] consider vehicle load, both full and empty, and this formulation is generally used for problems involving packing or multiple products, as applied in [33]. The [18, 35] formulations are most commonly used in transportation problems in the agricultural sector.

Demir *et al.* [30], Asghari and Mirzapour Al-e-hashem [10] propose various solutions to address the issue of carbon dioxide emissions. In particular, Asghari and Mirzapour Al-e-hashem [10] highlights the approach put forth by Barth *et al.* [13], which has gained widespread recognition in the scientific literature. Barth *et al.* [13] describe the emissions model for heavy diesel trucks, using a physical energy demand approach. When developing the model, Barth *et al.* [13] attempted to capture many important aspects of vehicle operation and its impact on tailpipe emissions. However, the authors saw the need for many variables, making it impossible to model all aspects with a high level of detail. Furthermore, the authors state that the model needs to be periodically updated to adequately represent current vehicles in any fleet. Future vehicle fleets will certainly include new technologies that are not represented in this version of the emission model.

Bektaş and Laporte [14] adapted the formula proposed by Barth *et al.* [13] for the pollution-routing problem, and [21], in turn, adapted it for the IRP with a heterogeneous fleet, obtaining an mixed integer problem (MIP), modeled with a non-linear objective function. To handle the non-linearity of the model, Cheng *et al.* [21] added a new variable for linearizing the velocity variable v, following [14]. The authors consider the IRP restricted to three types of diesel combustion vehicles, *i.e.*, different types of vehicles and one type of fuel.

Based on [13] and adapted from [21], in Appendix A, we present equation (A.1), which quantifies the emission of CO_2 in kilograms for a vehicle of type k, depending on the distance d (km) and the velocity v (km/h). In addition to [13], who propose equation (A.1), [14, 30] provide an explanatory summary of its derivation and mention that they initially include terms related to the fuel rate, such as engine power second by second and the total tractive power applied to the vehicle's wheels, which for a given arc (i, j) of length d consider v as the vehicle velocity variable traveling that arc.

While the equation proposed by Barth *et al.* [13] is comprehensive, its impracticality arises from the use of over thirty parameters. This complexity renders it unfeasible for practical applications in our context, where we examine fleets comprising more than three distinct vehicles and consider multiple fuels. It is essential to note that the formulation by Barth (2005) is specifically applicable to diesel and lacks scalability for larger fleets. Nonetheless, presenting this expression aims to provide readers with insight into one of the most frequently employed approaches in the existing literature.

Recent papers in machine learning and artificial intelligence have also brought studies in estimating CO_2 emissions, such as [38, 52, 71]. Machine learning, a subarea of artificial intelligence, empowers systems to learn and improve based on data-derived experience. Among the existing methods, supervised learning is notable for using labeled data sets to train algorithms, which learn to make predictions. The main goal is to efficiently map inputs to desired outputs, eventually adjusting parameters to minimize errors between predictions and actual values.

In this context, Wen *et al.* [71] proposes a Random Forest algorithm to estimate CO_2 emissions from large cities, using 272 indicators such as road characteristics, population density, and information on dirt roads to train the model. Niroomand *et al.* [52] uses Deep Learning techniques to assess the differences in CO_2 emissions among different vehicle classes and categories. Jiménez *et al.* [38] discusses the factors that dictate the difference between official and actual vehicle efficiency and CO_2 emissions, reviewing the influence of vehicle classification, vehicle features, vehicle brand, and reference year on real-world CO_2 emissions. They used a database of 650 passenger cars to explain the impact of these factors on the gap between real-world and homologation emission values.

This study aims to explore applying supervised machine learning methods to understand their functionality and performance as numerical alternatives for estimating $\mathrm{CO_2}$ emissions. To this end, we employed various commonly used supervised learning methods, Mahesh [45], to predict $\mathrm{CO_2}$ emissions. Each method offers a distinct approach to managing the complexity of the data, providing a comprehensive overview of the predictive capabilities of machine learning. A comparative study involving these methods was conducted to demonstrate their effectiveness in emission prediction, which will be presented later.

2.2. Overview of logistics problems considering CO_2 emissions

One of the most widely used ways in the literature to calculate carbon dioxide emissions is based on the expression proposed by Barth *et al.* [13], where the authors describe the emissions model for heavy diesel trucks, using a physical energy demand approach. Many research studies in logistics problems, specifically sustainable IRP, use Barth's formulation, such as [3, 21, 41, 48, 64, 66].

Other studies on sustainable logistics problems can be seen in: [20, 31, 39, 41, 73] and simpler emission formulations can be seen in: [12, 33, 42, 57, 68, 74]. Although each formulation has its particularities, they all share the same purpose of mitigating the negative impacts of human activity on the environment and promoting sustainability. Through these studies, they aim to reduce and estimate emissions, among other more conscious and sustainable practices.

Alkawaleet *et al.* [4] investigate the effect of $\mathrm{CO_2}$ emissions, considering the acquisition of carbon rights in IRP decisions, determined over a time horizon. A heterogeneous fleet is considered, but in each time period, at most one vehicle is used for product distribution to each customer. The results indicate that emission costs influence inventory and routing decisions. Fixed known values are considered for emissions.

In [68], in addition to the classic definition of IRP, CO_2 emissions are determined and added to the objective function using a carbon price. The IRP is then solved with the goal of minimizing total costs. The authors use the formulation of [13] as a set of constraints for the model. The model is applied to a case study in the petrochemical industry. The study considers a homogeneous fleet.

Soysal *et al.* [64] present a multi-period IRP model that includes truckload-dependent distribution costs, analysis of $\mathrm{CO_2}$ emissions and fuel consumption, perishability, and a service level constraint to meet uncertain demand. Considering a homogeneous fleet, the results suggest that the proposed integrated model can achieve total cost savings while meeting service level requirements.

Cheng *et al.* [21] study an IRP with the objective of minimizing the sum of inventory and routing costs, where the latter includes driver salary, vehicle fixed cost, fuel costs, and emissions, with fuel consumption determined by load, distance, speed, and vehicle characteristics. Numerical tests were conducted to quantify the benefits of using a comprehensive objective and heterogeneous vehicles. Managerial insights are also extracted from parameter analyses. However, it is limited to three types of diesel trucks, which may not be practical with large heterogeneous fleets and different fuel types. Additionally, the authors suggest a multi-objective analysis of the problem.

Alinaghian *et al.* [3] propose a mathematical model for the green IRP with time windows (GIRP-TW). The goal is to minimize fuel consumption cost, driver cost, inventory cost, and vehicle usage cost, taking into account several factors, similar to what was done in [21]. To solve the problem, three metaheuristic methods are designed, including the original and augmented Tabu Search algorithms and the Differential Evolution algorithm. The authors consider a fleet that consumes diesel.

Most of the works presented in this section address gas reduction by minimizing emissions in the model's objective function. In this article, we choose to concentrate on minimizing carbon dioxide emissions and operational costs within the context of the IRP.

2.3. Multi-objective IRP with CO_2 emissions

The literature on multi-objective methods for studying the IRP is still limited compared to the literature on single-objective papers. However, there are studies that address green IRP from a multi-objective perspective, such as [34, 58].

Franco *et al.* [34] introduce the multi-objective algorithm NISE (Non-Inferior Set Estimation), which uses the column generation technique to solve an IRP with two objectives. The authors do not describe in details the data they use to calculate CO_2 emissions.

In the article by Rahbari *et al.* [58], a bi-objective IRP is addressed, in which three vehicles distribute products from various suppliers to a retailer to meet product demand. The customer-associated demand is assumed to be time-variable and deterministic. Transshipment is considered in the model. The authors consider given values for emissions and minimize the costs of emitted CO_2. The proposed model is implemented and solved using the modeling language GAMS and its solvers.

In recent years, the multi-objective approach has received attention from researchers regarding other related problems, as seen in [49, 56, 59].

Misni *et al.* [49] consider a problem of the Location-Inventory-Routing Problem (LIRP) based on the economic order quantity model with environmental concerns. The study aims to minimize the total cost of operational facilities, inventory, and vehicle distance as the first objective and minimize the cost of CO_2 emissions as the second objective. The Multi-Objective Hybrid Simulated Annealing Harmony Search Algorithm (MOHS-SA) is applied to find the trade-off between these two objectives.

Rabbani *et al.* [56] present a multi-objective optimization model for a sustainable urban solid waste management system. The objective functions minimize: (i) the net total cost, (ii) the greenhouse gas emissions, and (iii) the total waste collection and treatment time. The model is implemented in a case study and solved using AUGMECON2 (enhanced ϵ-constrained). The authors consider that the values of CO_2 emissions come from the literature and do not provide an estimation of how to calculate this value.

Rahbari *et al.* [59] study an LIRP with a fleet of heterogeneous vehicles aiming to minimize supply chain risk, supply chain costs, and reduce greenhouse gas emissions. A meta-heuristic algorithm called MOBWO (Multi-Objective Binary Whale Optimization) is presented to solve multi-objective optimization problems. The algorithm's performance is compared with MOSA and NSGA-II.

Among the presented works, [34, 56] use exact methods to address the multi-objective IRP. Franco *et al.* [34] use NISE along with column generation, and [56] use the augmented ϵ-constrained method. Both works consider

TABLE 1. Summary of sustainable logistics problems literature.

References	Main problem	Fleet		Sustainability		M. O.	Objective		Expression type	Types of fuels				Solution method
		Ho.	He.	Constr.	Obj.		E	Sus.		D	G	E	Unk.	
Xiao *et al.* [74]	VRP	√			√		√		Linear				√	Modified S. A.
Demir *et al.* [31]	VRP	√			√			√	Non-linear	√				ALNS
Alkawaleet *et al.* [4]	IRP		√		√		√		Linear				√	MIP
Treitl *et al.* [68]	IRP	√		√	√		√		Linear	√				MIP
Koc *et al.* [41]	VRP		√		√		√		Non-linear	√				Evolutionary metaheuristic
Soysal *et al.* [64]	IRP	√			√		√		Non-linear	√				MIP
Franco *et al.* [34]	IRP		√		√	√	√	√	Linear				√	Column generation and NISE
Cheng *et al.* [21]	IRP		√		√		√	√	Non-linear	√				B&C
Rahbari *et al.* [57]	IRP		√	√	√	√	√		Linear				√	MIP and S. A.
Balamurugan *et al.* [12]	IRP	√			√			√	Linear				√	Evolutionary algorithm
Ferreira *et al.* [33]	VRP	√			√			√	Linear	√				B&C
Alinaghian *et al.* [3]	IRP		√		√		√		Non-linear	√				Augmented Tabu search and heuristics
Rabbani *et al.* [56]	LIRP		√		√		√	√	Non-linear				√	ϵ-constrained
Rahbari *et al.* [59]	LIRP		√		√	√	√	√	Linear				√	MOBWO, MOSA and NSGA II
Our research	IRP		√		√	√	√	√	Linear	√	√	√		B&C and ϵ-constrained

fixed known values for CO_2 emissions. The difference in our work is that we explore different trade-offs of green IRP compromises with different values for ϵ.

Table 1 presents a summary of these reviewed studies on green logistics and a comparison between them and our work. In it, we have the references, highlighting the main problem of the article, whether the fleet is homogeneous or heterogeneous, whether they address sustainability as a constraint and/or as an objective, whether the article considers multi-objective optimization, whether the objective(s) are economic and/or sustainable, whether the terms used to measure CO_2 emissions are linear or nonlinear, which fuels are used by the fleet, and what solution method is used. We use in the table: Ho: Homogeneous, He: Heterogeneous, Cons: Constrained, Obj: Objective, E: Economic, Sus: Sustainable, D: Diesel, G: Gasoline, E: Ethanol, Unk: Unknown. Acronyms used: VRP: Vehicle Routing Problem, M. O.: Multi-Objective. MIP: Mixed-Integer Programming, ALNS: Adaptive Large Neighborhood Search, B&C: branch-and-cut, S. A.: Simulated Annealing.

We estimate CO_2 emissions through an explicit linear equation, which is considered one of the objectives to be minimized alongside operational costs. To solve this problem, we apply the augmented ϵ-constrained method.

3. Machine learning methods and explicit vehicular equation for estimating CO_2 emissions

3.1. Machine learning methods

Machine learning empowers systems to learn and improve based on data-derived experience. Among the existing methods, supervised learning is notable for using labeled data sets to train algorithms, which learn to make predictions. The process eventually involves adjusting the model settings so that the learned function minimizes the error between the model-specific settings and the actual values in the training data. Over time, and with exposure to training data, the model becomes more effective at predicting results for new data sets, adapting to provide accurate flexibility based on learned characteristics of the data.

In this work, we analyzed annual databases from 1995 to 2022, totaling 26,146 vehicles fabricated in Canada and we employed seven supervised learning methods including Decision Trees, K-Nearest Neighbors (KNN), Lasso Regression, Random Forest, Linear Regression, Ridge Regression, and Support Vector Regression (SVR) to

TABLE 2. Different error measures for comparing machine learning methods used in emission estimation.

Method	ME	MGD	MAPE	MSE	r2
Decision tree	30,75	0,75%	0,01%	7,44	99,83%
KNN	70,50	0,90%	0,02%	13,61	99,69%
Lasso	52,95	1,15%	0,05%	29,24	99,32%
Random Forest	23,38	0,79%	0,01%	7,63	99,82%
Linear Regression	52,95	1,15%	0,05%	29,24	99,32%
Ridge	52,94	1,15%	0,05%	29,24	99,32%
SVR	108,98	0,81%	0,01%	11,09	99,74%

predict CO_2 emissions. From the entire database, we processed the data to remove duplicates, resulting in a total of $20,358$ examples. Each example in databases stores the following information: year of manufacture, brand, model, vehicle class, engine size, number of cylinders, transmission type, fuel type, city consumption, highway consumption, combined consumption, miles per gallon (MPG), and CO_2 emissions in grams per kilometer. The number of examples by fuel types is: gasoline (11,367), premium gasoline (7,790), ethanol (857), diesel (365), and natural gas (37). Some features in the model description include whether the vehicle has: four-wheel drive, all-wheel drive, is a hybrid vehicle, whether the wheelbase is short, long, or extended. The vehicle class informs whether it is a small or large SUV, compact, two-seater, pickup truck, truck, or others. The database is available on the Canadian government's website: https://open.canada.ca/data/en/dataset/98f1a129-f628-4ce4-b24d-6f16bf24dd64. The Canadian database is also available at the following link from the Github repository: https://github.com/ariannesilvamundim/RAIRO-OR.

We compared the results of the seven methods studied to estimate vehicle CO_2 emissions, as presented in Table 2. We allocated 70% of the dataset for training and tested the algorithms on the remaining 30% of the data. In this test subset, the actual value y_i is compared with the predicted value $\hat{y}_i$ to evaluate the performance of each method. The columns indicate the method used, the maximum error found (ME), the mean gamma deviation (MGD), the mean absolute percentage error (MAPE), the mean squared error (MSE), and the R-squared (r2). The description of these metrics can be found in Appendix B. These measures are widely used to assess the quality of proposed methods. We can observe that the proposed Explicit Vehicular Equation had the lowest maximum error, with ME $= 20.17$, while SVR had ME $= 108.98$. On the other hand, Decision Tree and Random Forest were the best methods in terms of MGD, MSE, and r2. For MAPE, the best methods are Decision Tree, Random Forest, and SVR.

Although using any of the tested machine learning methods to estimate CO_2 emissions is possible, given that all applied metrics resulted in very accurate predictions, we opted for a different approach. Instead of using these methods, we will employ a formula that is easy to apply. This formula is not only practical to use but also offers the flexibility to adapt the fleet in terms of fuel type and vehicle capacity. The only information required is the average efficiency (fuel consumption) of the available vehicles in the fleet. Furthermore, as the formula is statistically accurate, we will demonstrate its performance subsequently, reinforcing our decision to use it, as it can be easily integrated into the investigated mathematical model.

In Section 3.2, we introduce a method for calculating gas emissions. Next, in Section 3.3, we assess the accuracy of this approach and compare it to Machine Learning methods to measure the quality of emission estimation.

3.2. Proposed approach for green IRP

In this section, we gather data and provide a clear and explicit presentation of our methodology for estimating CO_2 emissions in kilograms, simplifying the requirement for detailed vehicle-specific information. When referring to this approach, we will term it the "Explicit Vehicular Equation". To develop it, we utilized data on CO_2 emission factors for multiple fuel types published by the Canadian government (see [51]).

During a vehicle's travel on the road, the internal combustion engine converts the energy stored in the fuel into mechanical energy to drive the vehicle's propulsion system, resulting in the emission of carbon dioxide as a byproduct of this process. According to [51], burning 1 liter of gasoline produces approximately 2.3 kg of CO_2. This means that an average vehicle, burning 2000 liters of gasoline per year, releases about 4600 kg of CO_2 into the atmosphere. From chemical studies, it is known that 1 liter of gasoline, weighing 0.75 kg, produces 2.3 kg of CO_2, as gasoline contains carbon and hydrogen atoms. During combustion, the carbon (C) in the fuel combines with the oxygen (O2) in the air to produce CO_2. The additional weight comes from oxygen.

As [70] consider, the density of gasoline is 0.75 kg/L, and the conversion factor from gasoline to CO_2 is 3.7 (C/CO_2). They emphasize that in some countries, such as Brazil, the volume of ethanol mixed with gasoline, which is approximately between 18% and 25%, should be subtracted. In this case, each liter of gasoline blended with ethanol has a maximum of 82% pure gasoline, which should be the focus of greenhouse gas emissions originating from fossil fuels in transportation. Thus, based on the study by Wang *et al.* [70], the calculation is as follows: 1 liter of gasoline $= 1\times0.82\times0.75\times3.7 =$ total kilograms of CO_2 emitted per liter. For example, for a daily commute in the city of 20 kilometers (km) with a car that has a consumption rate of, for instance, 10 km/liter, its consumption will be 2 liters. For this example, the total emissions are given by: $2\times0.82\times0.75\times3.7 = 4.55$ kg of CO_2. Now, if approximately 1 gasoline tank (50 liters) is used per week, driving 500 km/week, in a car that travels 10 km per liter of gasoline, the gas emission will be: $50 \times 0.82 \times 0.75 \times 3.7 = 114$ kg of CO_2 per week.

Emissions from a vehicle vary depending on the type of fuel due to different densities. More dense hydrocarbon fuels, such as diesel, contain more carbon and, therefore, produce more CO_2 for a given volume of fuel. Emissions from various types of transportation fuels are well-defined and known, as documented in [51]. The relationship between a vehicle's consumption (efficiency) in kilometers per liter and CO_2 emissions is known, Pinto and Oliver-Hoyo [54]. Thus, the liters of fuel consumed on a given route, based on the vehicle's average consumption, is given by:

$$\frac{\text{distance}}{\text{avg\_veh\_cons}}\,\frac{\text{(km)}}{\text{(km/L)}}, \tag{1}$$

which implies:

$$\frac{\text{distance}}{\text{avg\_veh\_cons}}\,\text{(L)}. \tag{2}$$

Equation (1) represents the ratio between the distance in kilometers and the average consumption of a vehicle, which involves the distance traveled and the amount of fuel used, in kilometers per liter. Thus, by simplifying units of measurement, Expression 2 provides how many liters of fuel are spent to cover a distance d, knowing the average consumption of the vehicle. Knowing the direct relationship between how many kilograms of CO_2 are emitted for each liter of fuel consumed, multiplying the CO_2 emission in Expression 2, we have:

$$\text{CO2\_emission (kg/L)} \times \frac{\text{distance}}{\text{avg\_veh\_cons}}\,\text{(L)}. \tag{3}$$

Thus,

$$\frac{\text{CO2\_emission} \times \text{distance}}{\text{avg\_veh\_cons}}\,\text{(kg)}. \tag{4}$$

Therefore, Expression 4 provides how many kilograms of CO_2 were emitted to cover a distance d, knowing in advance the average consumption of the vehicle for any type of fuel used. In other words, this well-defined expression accommodates a heterogeneous fleet of vehicles. That's why this equation will be used in the proposed green IRP model in this work.

TABLE 3. Different error measures for comparing results from the explicit vehicular equation in emission estimation.

Method	ME	MGD	MAPE	MSE	r2
Explicit Veh Eq. $F^k_{CO_2}$	20,17	0,98%	0,02%	12,23	99,72%

So, to obtain a sustainable approach to the IRP, it is necessary to calculate CO_2 emissions, as they are intended to be minimized. For this, a practical way to calculate how many kilograms of gas are emitted to cover a certain distance is to know the average consumption of the vehicle used, regardless of vehicle deterioration, new vehicle technologies, and other factors. The equation is given below, written for each type of vehicle/fuel k:

$$F^k_{CO_2} = \frac{f^k\left(\frac{\text{kg}}{\text{L}}\right) \times d\,(\text{km})}{cv^k\left(\frac{\text{km}}{\text{L}}\right)} = \frac{f^k \times d\left(\text{kg}\,\frac{\text{km}}{\text{L}}\right)}{cv^k\left(\frac{\text{km}}{\text{L}}\right)} = \frac{f^k \times d}{cv^k}\,(\text{kg})\,, \tag{5}$$

where,

- f^k represents the kilograms of CO_2 emitted for each liter of fuel consumed by vehicle/fuel k (kg CO_2 / L),
- d is the distance traveled by the vehicle (km),
- cv^k is the average consumption of fuel of vehicle/fuel k (km/L).

The Explicit Vehicular Equation, $F^k_{CO_2}$ given by 5, defines how many kilograms of CO_2 are emitted into the atmosphere based on the average vehicle consumption. This expression is considered generic, providing greater scalability to the problem by allowing the consideration of various fuels and types of vehicles. Additionally, it is straightforward to apply, requiring few parameters and can be adapted as needed.

In general terms, the first step is to obtain the average consumption of the vehicle/fuel, which may vary depending on the driver, vehicle condition, maintenance, and weight changes. Then, we use the following values for f^k for different types of fuels (*e.g.*, Gasoline, E10 (10% ethanol and 90% gasoline), E85 (85% ethanol and 15% gasoline), Diesel, B5 (5% biodiesel and 95% diesel), B20 (20% biodiesel and 80% diesel)): 2.29, 2.21, 1.61, 2.66, 2.65, 2.62, respectively.

3.3. Validation of the proposed formulation and comparison with machine learning methods

In this section, we evaluate the performance of the explicit vehicular equation compared to machine learning methods, using the same metrics and test data. Following the comparison, we demonstrate the validation of the formula across the entire dataset. The results of applying the Explicit Vehicular Equation to the test data, previously segregated for analysis, are presented in Table 3.

It is essential to highlight that all estimates, $F^k_{CO_2}$ and the methods, were useful in estimating CO_2 emissions. An r2 above 99% is an excellent result for an approximation. However, the simplicity and quality of the approach proposed in this work compared to Machine Learning Methods are undeniable.

The proposed equation is easy to interpret, which is a significant advantage compared to other proposed algorithms that are often treated as black boxes or are challenging to interpret, except for the decision tree algorithm. Additionally, the proposed equation is computationally efficient because it can be directly used to predict gas emissions without the need for significant data collection and preprocessing steps or hyperparameter training and optimization to generate a prediction model.

In summary, applying the data to our Explicit Vehicular Equation, we obtained an R-squared (r2) of 99.72%. An r2 above 99% means that the values predicted by the proposed equation can account for more than 99% of gas emissions. In addition, since the Explicit Vehicular Equation can be directly applied to the entire dataset, we applied it to all 20,416 instances and obtained a mean absolute error of 1.07%. For more details on the data, see Appendix B.

In this section, we adapt the Explicit Vehicular Equation $F^k_{CO_2}$, so that we can assess its accuracy on a public dataset. Our equation uses the known values of the amount of kilograms of CO_2 emitted for each liter of fuel and the average consumption to develop an expression. This expression, when multiplied by the distance traveled, results in a unique unit of measurement: kilograms of CO_2 emitted by the vehicle over the corresponding distance.

The required adaptation to apply the Explicit Vehicular Equation involved only simple algebraic manipulations, enabling the estimation of the amount of CO_2 emitted in grams per kilometer. Thus, the equation to be employed to validate the dataset is presented below:

$$F_{CO_2}\left(\frac{\mathrm{g}}{\mathrm{km}}\right) = f^k\left(\frac{kg}{L}\right) \times FCC\left(\frac{\mathrm{L}}{100\,\mathrm{km}}\right) = 10 \times f^k \times FCC\left(\frac{\mathrm{g}}{\mathrm{km}}\right),$$

where FCC represents the value in the "Fuel Consumption Comb" column of the dataset, measured in liters per 100 kilometers.

In the following, we use the described dataset to evaluate the Explicit Vehicular Equation for estimating CO_2 emissions, simultaneously comparing its effectiveness with traditional Machine Learning methods. Seven regression algorithms were implemented, [40], with the purpose of estimating CO_2 emissions: Decision Tree, KNN, Lasso, Random Forest, Linear Regression, Ridge, and SVR. A detailed description of each algorithm and parameter calibration can be found in Appendix B. It is relevant to emphasize that all employed algorithms are of the supervised type, meaning they have the ability to learn from previously provided data. For this study, we split the dataset into two distinct parts: allocating 70% of the data for the model training process, while the remaining 30% were used to test the algorithms' performance.

The database contains independent variables (or features) and the target variable (the CO_2 emissions). The features are: engine size in liters (Engine Size), the number of cylinders (Cylinders), average fuel consumption (Fuel Consumption Comb), and fuel type (Fuel Type).

To apply the Explicit Vehicular Equation to this dataset, the process is straightforward: for each entry in the "Fuel Type" variable, we use the following symbols: X for regular gasoline, Z for premium gasoline, D for diesel, E for ethanol (E85), and N for natural gas. Then, for each type of fuel, we use the corresponding values of f^k: 2.29 for regular gasoline, 2.29 for premium gasoline, 2.66 for diesel, 1.61 for ethanol, and 0 for natural gas.

Consider an example taken from the database: a 2022 Toyota Corolla that uses regular gasoline ($f^k = 2.29\,\mathrm{kg/L}$), has a fuel consumption of $FCC = 7.1\,\mathrm{L}/100\,\mathrm{km}$, and emits 165 g/km of CO_2. Simply multiply the fuel consumption (FCC) by the corresponding fuel type value to obtain the amount of CO_2 in grams per kilometer using the following equation:

$$F_{CO_2}\frac{\mathrm{g}}{\mathrm{km}} = \frac{10}{10} \times 2.29\,\frac{\mathrm{kg}}{\mathrm{L}} \times 7.1\frac{\mathrm{L}}{100\,\mathrm{km}} = 162.59\,\frac{\mathrm{g}}{\mathrm{km}}\cdot$$

Note that we included the factor of ten divided by ten for unit conversion to grams per kilometer (g/km) to ensure consistency in the dataset's unit of measurement. Thus, the estimated value is 162.59 g/km, while the actual value is 165 g/km, resulting in a difference of −1.46%.

4. Problem modeling and solution methods

In Section 4.1, we introduce the model that we developed. Subsequently, in Section 4.2, we provide the essential definitions for the multi-objective problem, along with an explanation of the augmented ϵ-constrained approach.

TABLE 4. Sets and parameters of the GIRP.

Sets:	
$\mathbb{V}$	Set of all vertices
$\mathbb{A}$	Set of edges
$\mathbb{V}'$	Set of customers
$\mathbb{K}$	Set of vehicles
τ	Set of time periods
Parameters:	
C^k	Capacity of each vehicle
c_{ij}	Transportation cost between nodes i and j
p_{ij}	Euclidean distance of the path between nodes i and j
h_i^t	Inventory holding cost at node i at the end of time period t
C_i	Storage capacity at customer i
r^t	Quantity available at the supplier in time period t
I_i^0	Initial inventory at node i
d_i^t	Customer i demand in time period t
f^k	CO_2 emitted for each liter of fuel consumed by vehicle k (kgCO_2/L)
cv^k	Average consumption of fuel of vehicle k (km/L)

4.1. Mathematical model GIRP and B&C algorithm

The proposed mathematical formulation for the green IRP, referred to as GIRP, which minimizes the emissions of kilograms of CO_2 as seen in Expression 5, is subject to constraints that arise from the IRP with heterogeneous fleet, based on [5, 24].

The problem is defined on an undirected graph $\mathbb{G} = (\mathbb{V}, \mathbb{A})$, where $\mathbb{V} = \{0, \ldots, n\}$ is the set of vertices, and $\mathbb{A} = \{(i, j), |, i, j \in \mathbb{V}, , i < j\}$ is the set of edges. Vertex 0 represents the supplier, and the remaining vertices in $\mathbb{V}' = \mathbb{V} \backslash \{0\}$ represent n customers. The supplier has a heterogeneous fleet composed of K vehicles, denoted by the set $\mathbb{K} = \{1, \ldots, K\}$, where each vehicle $k \in \mathbb{K}$ has unique characteristics, such as capacity C^k and the type of fuel it consumes (*i.e.*, each $k \in \mathbb{K}$ denotes a vehicle/fuel). Each vehicle can perform a route per time period to deliver products from the supplier to a subset of customers. A travel associated with all the edges $(i, j) \in \mathbb{A}$ have a cost c_{ij} and a distance p_{ij}. Both the supplier and the customers have unit inventory holding costs h_i^t at the end of each time period, $i \in \mathbb{V}$, and each customer has a storage capacity C_i, $i \in \mathbb{V}$. The planning horizon size is T, and in each time period $t \in \tau = \{1, \ldots, T\}$, the production/quantity of product available at the supplier is r^t. We assume that the supplier has sufficient stock to meet the total customer demand during the planning horizon, and all demand must be satisfied; in other words, backlogging is not allowed, and inventories cannot be negative. The variables I_0^0 and I_i^0 are defined as the initial inventories at the supplier and customer $i \in \mathbb{V}'$, respectively. At the beginning of the planning horizon, the decision-maker knows the demand d_i^t of each customer i for each time period t.

The model uses variables x_{ij}^{kt} equal to the number of times edge (i, j) with $i < j$ is used in the route of vehicle k in period t. Binary variables y_i^{kt} are equal to 1 if and only if node i (the supplier or a customer) is visited by vehicle k in period t. Let I_i^t be the inventory level at vertex $i \in \mathbb{V}$ at the end of time period $t \in \tau$. q_i^{kt} is the quantity of product delivered from the supplier to customer i using vehicle k in time period t.

Table 4 summarizes all the notation introduced above.

With this, the GIRP consists of determining when customers will be visited, the quantity of product that will be delivered to each customer, and which routes should be taken to make these deliveries. The objective is to minimize the total operational cost, as well as minimize the emission of CO_2 kilograms.

The MIP model for GIRP is presented below:

$$\text{Minimize} \quad F_{IRP} = \sum_{i\in\mathbb{V}} \sum_{i\in\tau} h_i^t\, I_i^t + \sum_{i\in\mathbb{V}} \sum_{j\in\mathbb{V}, i<j} \sum_{k\in\mathbb{K}} \sum_{t\in\tau} c_{ij}\, x_{ij}^{kt} \tag{6}$$

$$\text{Minimize} \quad F_{CO_2} = \frac{\sum_{i\in V} \sum_{j\in V, i<j} \sum_{k\in\mathbb{K}} \sum_{t\in T} f^k\, p_{ij}\, x_{ij}^{kt}}{\sum_{k\in\mathbb{K}} cv^k} \tag{7}$$

Subject to

$$I_0^t = I_0^{t-1} + r^t - \sum_{k\in\mathbb{K}} \sum_{i\in\mathbb{V}'} q_i^{kt}, \qquad t\in\tau, \tag{8}$$

$$I_i^t = I_i^{t-1} + \sum_{k\in\mathbb{K}} q_i^{kt} - d_i^t, \qquad i\in\mathbb{V}', t\in\tau, \tag{9}$$

$$\sum_{k\in\mathbb{K}} q_i^{kt} \le C_i - I_i^{t-1}, \qquad i\in\mathbb{V}', t\in\tau, \tag{10}$$

$$q_i^{kt} \le C_i\, y_i^{kt}, \qquad i\in\mathbb{V}', k\in\mathbb{K}, t\in\tau, \tag{11}$$

$$\sum_{i\in\mathbb{V}'} q_i^{kt} \le C^k\, y_0^{kt}, \qquad k\in\mathbb{K}, t\in\tau, \tag{12}$$

$$\sum_{j\in\mathbb{V}, i<j} x_{ij}^{kt} + \sum_{j\in\mathbb{V}, j<i} x_{ji}^{kt} = 2\, y_i^{kt}, \qquad i\in\mathbb{V},\ k\in\mathbb{K}, t\in\tau, \tag{13}$$

$$\sum_{i\in\mathbb{S}} \sum_{j\in\mathbb{S}, i<j} x_{ij}^{kt} \le \sum_{i\in\mathbb{S}} y_i^{kt} - y_g^{kt}, \qquad \mathbb{S}\subseteq\mathbb{V}', k\in\mathbb{K}, t\in\tau, \forall g\in\mathbb{S}, \tag{14}$$

$$\sum_{k\in\mathbb{K}} y_i^{kt} \le 1, \qquad i\in\mathbb{V}, t\in\tau, \tag{15}$$

$$I_i^t \ge 0, \qquad i\in\mathbb{V}, t\in\tau, \tag{16}$$

$$q_i^{kt} \ge 0, \qquad i\in\mathbb{V}', k\in\mathbb{K}, t\in\tau, \tag{17}$$

$$x_{0j}^{kt} \in \{0,1,2\}, \qquad j\in\mathbb{V}', k\in\mathbb{K}, t\in\tau, \tag{18}$$

$$x_{ij}^{kt} \in \{0,1\}, \qquad i,j\in\mathbb{V}' : i<j, k\in\mathbb{K}, t\in\tau, \tag{19}$$

$$y_i^{kt} \in \{0,1\}, \qquad i\in\mathbb{V}, k\in\mathbb{K}, t\in\tau. \tag{20}$$

The objective function 6 F_{IRP} minimizes inventory and transportation costs, while 7 F_{CO_2} minimizes the kilograms of carbon dioxide emitted, through the proposed formulation described in 5. Constraints (8) and (9) define inventory levels at the supplier and customers, respectively. Constraints (10) impose a maximum inventory level at the customers. Constraints (11) associate the quantity of products delivered with routing variables, allowing a vehicle to deliver products to a customer only if that customer is visited by that vehicle. Constraints (12) ensure that vehicle capacities are respected, while constraints (13) and 14 are degree constraints and subcycle elimination constraints, respectively. Constraints (15) state that each customer can be visited at most once in each period. Constraints (16) prevent stockouts at the supplier and customers, and Constraints (17)–(20) impose non-negativity and integrality conditions on the variables.

To solve this model exactly, we employed a basic branch-and-cut (B&C) algorithm based on [6]. Several authors in the IRP literature use exact B&C methods, as in [8, 9, 23]. Green IRP studies also employ this type of method, such as in [21, 22, 27, 60]. The GIRP and ϵ-GIRP models contain an exponential number of subcycle elimination constraints (14), so a B&C-type algorithm was applied in this work. These constraints are relaxed in the formulation and added iteratively whenever they are violated in the nodes of the branch-and-bound tree.

In the implementation, an exact separation algorithm is used to solve a series of minimum cut problems to identify the violated constraints for each vehicle and time period. At each node of the branch-and-bound tree, let $\bar{y}_i^{kt}$ and $\bar{x}_{ij}^{kt}$ be the values of the visit (y) and flow (x) variables in the solution, respectively. A graph is constructed for each vehicle k and time period t for the nodes with $\bar{y}_i^{kt} > 0$, defining the edge weights of the new graph as $\bar{x}_{ij}^{kt}$, $\forall (i, j) \in \mathbb{A}$.

Next, we will present some definitions of multi-objective optimization and describe the augmented epsilon-constrained method for GIRP.

4.2. Multi-objective optimization for GIRP and augmented e-constrained method

A multi-objective (or multicriteria) problem is characterized by the existence of a set of feasible solutions defined through a set of constraints, where the objectives are specified through an objective function. In general, in multi-objective problems, there is no solution that optimizes all objective functions simultaneously [2], therefore, they involve the optimization of a vector composed of scalar functions chosen as a way to assess the impact of feasible decisions in the problem, according to different performance indices.

The multi-objective approach seeks to find the set of optimal points for the components of the vector objective function (the vector is composed of the objective functions to be optimized), where, unlike mono-objective optimization, the solution to the problem is a set of efficient points (solutions). Each efficient solution is optimal in the sense that no improvement can be achieved in one component of the vector function without worsening at least one of the remaining components of the vector function [62]. Within the set of efficient solutions, the decision-makers will choose the one they find most satisfactory.

Here are some definitions, which can also be found in [25]:

– *Dominated solution*: A solution is dominated if and only if there exists another solution that is better in at least one criterion/objective, without being worse in any of the others.
– *Efficient solution (Non-Dominated or Pareto Optimal)*: A solution is efficient if and only if it is not dominated by any feasible solution, meaning that in multicriteria formulation, a non-dominated solution is a solution that outperforms another solution in all objectives, or it is a solution that cannot be improved in terms of one objective without worsening at least one other objective.
– *Weakly efficient solution*: A solution is weakly efficient if it can be improved in terms of one objective without worsening at least one other objective.
– *Efficient set (Pareto Set)*: It is the set of non-dominated solutions.
– *Pareto Front*: It is the image of non-dominated solutions of the Pareto set in the objective values space.

Professionals in various fields often face situations where it is necessary to consider multiple objectives and evaluate the possibility of relaxing constraints to achieve a better fit of the model to the real-world problem, Aliano Filho *et al.* [1]. For example, when planning product deliveries to meet customer demand, managing inventory, and determining vehicle routes, it can be stated that executing this plan will emit a certain amount of CO_2. Therefore, subjecting this plan to emissions as a constraint in the model, limited to acceptable emission values, is an interesting way to analyze the trade-off behavior between costs and emissions.

In this context, within the scope of the GIRP, this work conducts an analysis of solutions with the aim of capturing the trade-off between costs and emissions, through the development of a multi-objective optimization model, as suggested by [21, 66].

The ϵ-constrained method overcomes some difficulties of previous methods. Proposed in [36], this method scalarizes a multi-objective problem by selecting one function as the objective and restricting the others with specified limits. When these limits vary appropriately, efficient solutions can be obtained.

In this study, we apply the augmented ϵ-constrained method, Mavrotas [47], as it avoids weakly efficient solutions when slightly modifying the contours of the objective to be optimized. We achieve this by prioritizing the costs of IRP in the GIRP model, described in Expression 6 as F_{IRP}, and adding a term that introduces sustainability to the model, computing the emission of gas described in Expression 7 as F_{CO_2} multiplied by $\rho \approx 0$. Finally, we include the emission of CO_2 as a constraint limited to ϵ. Although the literature formulation F^k can be used in the multi-objective approach, we chose to use the gas emission estimate proposed in this work, given by equation (5), which is the F_{CO_2} of the GIRP model. With this, the model is named ϵ-GIRP:

$$\text{Minimize} \quad F_{IRP} + \rho\, F_{CO_2} \tag{21}$$

Subject to

$$F_{CO_2} \leq \epsilon, \tag{22}$$

$$\text{constraints (8)–(20).} \tag{23}$$

With $\rho \approx 0$, we slightly modify the contours of the objective to be optimized. The objective function 21 minimizes inventory and transportation costs plus the emission of kilograms of CO_2. While the set of constraints (22) ensures that the established ϵ is considered, *i.e.*, the CO_2 emission will be limited to ϵ values. Constraints (8)–(20) are described earlier. The way ϵ is defined will be described in Section 5.3.

5. Computational experiments, results, and discussions

We present the entire testing environment and data in Section 5.1, with details of the general data. And we present its results in Section 5.2. In Section 5.3, we show the results of the ϵ-GIRP model. Finally, in Section 5.4 we also discuss managerial insights from the decision-maker's perspective.

5.1. Testing environment and data for GIRP

The algorithms were implemented in the C++ programming language, and the computational experiments were conducted on a machine with an Intel Core i5-11400H @ 2.70 GHz × 12, 16 GB RAM, and Ubuntu 20.04.3 LTS as the operating system. The model was solved with IBM ILOG CPLEX 12.9, considering its default settings and a time limit of 3600 s as the stopping criterion.

The instances are inspired by Archetti *et al.* [8]. For all models, we used the following data: The number of customers: $n = 5, 10, 15, 20, 25, 30, 35, 40, 45, 50$; inventory costs at customers, $h_i = [0.1,\ 0.5]$ and at the supplier $h_0 = 0.3$ (denoted as H, representing high costs) and $h_i = [0.01,\ 0.05]$ and at the supplier $h_0 = 0.03$ (denoted as L, representing low costs); thus, they were named H3, L3, H6, L6 for $T = 3$ and 6, representing the planning horizon size; demand d_i^t is randomly generated as an integer in the range $[10, 100]$ and remains constant over time, *i.e.*, $d_i^t = d_i$; the product quantity $r^t = \sum_{i \in \mathbb{V}'} d_i$ at time t; the maximum inventory level at customer i is generated between 0 and 500; the initial inventory level at the supplier is the sum of the maximum inventory level of all customers; the initial inventory level at customer i is the maximum inventory level minus d_i. Euclidean distance is considered for p_{ij}, where the points (X_i, Y_i) and (X_j, Y_j) in the plane are obtained by automatic generation of each coordinate as an integer in the range $[0, 500]$. As done by literature, travel costs correspond to Euclidean distances rounded to the nearest integer, *i.e.*, $\lfloor p_{ij} \rceil = c_{ij}$. This totals 32 groups of instances, as [8] in their benchmark configure five variations for each group, totaling 160 instances used in this article.

For instances where the model did not find the optimal solution, we calculated the solution gap, obtained by $\frac{UB-LB}{UB}$, where UB is the upper bound and LB is the lower bound.

TABLE 5. Data used in the proposed instances for GIRP.

Type of fuel	CO_2 emissions (kg/L)	cv^k
Gasoline	2.29	8
E10 (10% ethanol + 90% gasoline)	2.21	10
E85 (85% ethanol + 15% gasoline)	1.61	12
Diesel	2.66	2
B5 (5% biodiesel + 95% diesel)	2.65	3
B20 (20% biodiesel + 80% diesel)	2.62	5

For the computational tests, we generated instances with different fleet sizes, six types of fuels, and distinct vehicle capacity proportions. Based on [51], the fuels are gasoline, E10, E85, diesel, B5, and B20, along with their respective emission factors f^k for each vehicle/fuel k, as shown in the second column of Table 5. Fleet sizes include 6, 18, and 30 vehicles, which are described below. In this table, we have data for each vehicle that makes up the fleets and instances proposed for GIRP. The table rows provide information for each vehicle/fuel k, including their respective emissions and vehicle efficiency cv^k. With this data, the table illustrates the configuration of a fleet with 6 vehicles, where each vehicle consumes a specific type of fuel. The fleet with 6 vehicles is such that each vehicle consumes a specific fuel type. The fleet with 18 vehicles includes three gasoline vehicles, seven E10 vehicles, three E85 vehicles, one diesel vehicle, one B5 vehicle, and three B20 vehicles. The largest fleet with 30 vehicles includes four gasoline vehicles, eight E10 vehicles, ten E85 vehicles, two diesel vehicles, two B5 vehicles, and four B20 vehicles.

The vehicle capacity C^k is estimated based on the capacity created by Archetti *et al.* [8], where vehicles with higher capacity are diesel, B5, B20, followed by gasoline, E10, and E85, with capacities of 90%, 85%, 70%, 25%, 20%, and 5% of the Archetti instances' capacity.

5.2. GIRP results

We conducted extensive computational tests for GIRP. In the result tables in this section, we show the solutions obtained from the model for each of the objective functions separately. This allows us to analyze how CO_2 emissions impact IRP costs and how costs influence emissions. The tests cover fleets of different sizes and configurations to analyze the model's performance and effectiveness. Once again, it should be noted that the model proposed in this work allows us to include more vehicles in the problem with different fuels.

In Tables 6, 7 and 8, each row contains the average solution of a group of five instances, totaling 160 in each table, arranged as follows: in the first column is the number n of customers; the second column contains the number p of planning periods, which are 3 and 6, accompanied by the letters H and L, representing high and low inventory costs, respectively; the third column "CO_2 o. f." shows the amount of CO_2 emitted from the GIRP model when minimizing only CO_2, GIRP (min CO_2); in the fourth column "IRP v.", the resulting cost value when minimizing CO_2 is presented; the fifth column displays the computational time in seconds, and the sixth column shows the solution gap. The last four columns present the results obtained from the GIRP model when minimizing only costs, GIRP (min transp_c + invent_c), which include: minimized costs in "IRP o. f.", the CO_2 emission value in "CO_2 v.", the time, and the gap, respectively.

Some observations about the tables should be highlighted: Each row in the tables contains the result of the average of five instances, but for some groups, this was not possible. To represent the group of instances for which the algorithm did not find feasible solutions for at least one of them, an asterisk (*) is subscripted in the tabulated values. All analyses and considerations between the models were made for the groups for which it was possible to calculate the averages of the five, *i.e.*, groups for which the algorithm found viable and/or optimal solutions within the 3600 s time limit.

Table 6. Computational results for the proposed model, for 6 vehicles.

		GIRP (min CO_2)				IRP (min transp_c + invent_c)			
n	p	CO_2 o. f.	IRP v.	Runtime	Gap %	IRP o. f.	CO_2 v.	Runtime	Gap %
5	H3	794.8	4367.2	15.1	0	3674.7	1999.4	2.5	0
15	H3	818.7	5769.1	188.1	0	4531.6	2289.8	3.4	0
20	H3	946.0	7457.8	1525.8	1	5992.3	2631.7	89.2	0
25	H3	1039.1	8639.2	2303.0	8	7409.1	2680.9	52.9	0
30	H3	1025.9	10579.1	2786.1	20	8770.0	3014.0	41.2	0
35	H3	1157.4	11943.4	3600.6	20	9267.8	3121.2	802.0	0
40	H3	1219.5	12755.4	3600.0	30	10048.9	3229.4	144.0	0
45	H3	1554.0	15217.8	3600.0	40	11119.8	3364.8	821.9	0
50	H3	2549.4*	22019.6*	3600.0*	64*	12354.9	3128.0	2333.3	1
5	H6	1407.1	6160.1	39.0	0	5041.6	3724.2	4.9	0
10	H6	2052.0	10674.3	3259.5	10	8113.6	5532.4	678.4	0
15	H6	2273.7	14563.1	3600.0	22	10637.7	5748.9	598.8	0
20	H6	2905.5	17821.6	3600.0	32	13403.1	6755.2	2459.7	1
25	H6	3495.1*	22309.2*	3600.0*	35*	15482.1	7537.5	2416.6	1
30	H6	3534.0*	27693.9*	3600.0*	46*	18498.6	7455.4	3600.0	1
5	L3	523.6	1875.9	0.9	0	1246.7	1454.4	0.4	0
10	L3	794.8	2549.1	15.5	0	1863.4	1935.9	3.4	0
15	L3	818.7	3342.3	193.3	0	2118.0	2249.7	4.9	0
20	L3	946.0	4023.1	1561.3	1	2589.5	2647.6	64.7	0
25	L3	1039.1	4124.0	2310.1	8	2940.4	2627.3	164.8	0
30	L3	998.4	4752.5	3600.0	13	3066.7	2947.9	44.2	0
35	L3	1160.5	5966.7	3600.0	21	3303.0	3353.7	1217.0	0
40	L3	1269.0	6588.7	3600.0	32	3463.3	3083.7	190.2	0
45	L3	1544.0	7671.6	3600.0	40	3666.3	2946.1	1074.8	0
50	L3	–*	–*	–*	–*	4121.9	3709.2	2395.2	3
5	L6	1407.1	4247.1	39.6	0	3143.7	3622.3	4.2	0
10	L6	2049.3	7158.8	3269.4	9	4712.5	5680.6	915.9	0
15	L6	2273.1	9011.4	3600.0	22	5402.8	5953.4	980.5	0
20	L6	2900.1	11064.2	3600.0	32	6584.5	7198.2	2979.4	2
25	L6	3512.6*	14596.7*	3600.0*	35*	7258.4	7374.6	3365.3	2
30	L6	3522.7*	15863.5*	3600.0*	46*	7635.5	7709.5	3600.0	3

In Table 6, solutions are presented for a fleet with 6 vehicles, each consuming a different type of fuel as described in Table 5. There are 160 instances in total, for which the GIRP (min transp_c + invent_c) model found solutions for all instances, while the GIRP (min CO_2) model did not find a solution for at least one instance in each of the following groups: $n = 50$ and H3; $n = 25$, 30, and H6; $n = 50$ and L3; $n = 25$, 30, and L6, within the time limit. In terms of optimality proof, the GIRP (min CO_2) model was able to prove the optimality of 35.62% of the instances compared to 79.37% for the IRP (min transp_c + invent_c) model. For this fleet of vehicles, GIRP (min CO_2) obtained an average gap of 13.88%, while the gap for the GIRP (min transp_c + invent_c) model was less than 1%. Nevertheless, the CO_2 emissions from GIRP (min CO_2) were impressive, achieving a 61% reduction. In terms of computational time, GIRP (min CO_2) was 76.7% more costly.

In Table 7, solutions are presented for a fleet of 18 vehicles, with three gasoline vehicles, seven E10 vehicles, three E85 vehicles, one diesel vehicle, one B5 vehicle, and three B20 vehicles. There are 160 instances in total, for which the GIRP (min transp_c + invent_c) model found solutions for all instances, while the GIRP (min CO_2) model did not find a solution for at least one instance in each of the four groups: $n = 50$ and H3; $n = 5$ and H6; $n = 50$ and L3; and $n = 5$ and L6, within the time limit. In terms of optimality proof, the GIRP (min CO_2) model was able to prove the optimality of 5.63%, while the GIRP (min transp_c + invent_c) model proved

TABLE 7. Computational results for the proposed model, for 18 vehicles.

		GIRP (min CO_2)				IRP (min transp_c + invent_c)			
n	p	CO_2 o. f.	IRP v.	runtime	gap %	IRP o. f.	CO_2 v.	runtime	gap %
5	H3	518.9	2834.8	854.7	0	1934.3	1386.4	1.2	0
10	H3	771.3	4994.23	3600.0	19	3674.7	2045.2	19.9	0
15	H3	806.2	6423.3	3600.0	32	4531.6	2278.6	52.5	0
20	H3	1034.7	8711.7	3600.0	39	5992.3	2546.8	1029.4	0
25	H3	1039.4	9847.9	3600.0	33	7409.1	2715.7	1936.0	0
30	H3	1302.9	12749.0	3600.0	51	8770.0	2993.3	783.8	0
35	H3	1737.2	14881.3	3600.7	55	9290.9	3141.9	2449.4	1
40	H3	6642.5	26733.2	3600.8	90	10073.4	2894.5	2448.8	1
45	H3	8426.3	32019.9	3591.3	92	11549.9	2914.0	3490.5	4
50	H3	13286.9*	43212.3*	3600.1*	94*	12653.9	3819.1	3600.3	4
5	H6	1147.1*	5899.3*	3600.0*	8*	5041.6	3533.8	30.2	0
10	H6	2030.8	12308.5	3600.0	25	8118.6	5416.6	2568.3	1
15	H6	2604.1	17055.5	3600.0	47	10638.9	5672.9	3231.9	1
20	H6	5621.2	33522.8	3600.0	67	13503.8	6737.8	3600.0	2
25	H6	10067.9	41083.4	3600.0	75	15632.6	7735.9	3600.0	2
30	H6	12983.2	47232.5	3600.1	88	35718.2	11072.1	3600.1	38
5	L3	518.9	2143.3	850.2	0	1246.7	1456.8	1.2	0
10	L3	771.3	3170.4	3600.0	20	1863.4	2014.5	23.8	0
15	L3	795.8	3961.3	3600.0	31	2118.0	2219.2	77.4	0
20	L3	1038.2	5340.0	3600.0	40	2589.5	2652.4	963.2	1
25	L3	1037.5	5314.8	3600.0	33	2940.4	2445.1	1743.7	1
30	L3	1371.0	7051.5	3600.3	52	3066.7	2961.4	581.9	0
35	L3	1718.6	8753.8	3600.0	54	3322.8	3213.4	2332.0	3
40	L3	6625.4	19623.6	3600.0	90	3463.3	3250.0	2128.6	1
45	L3	8370.7	24353.6	3600.0	92	3962.4	2760.5	3568.9	9
50	L3	13286.9*	35196.0*	3600.1*	94*	4330.3	3662.4	3600.0	9
5	L6	1424.4*	4296.4*	3600.0*	10*	3143.7	3703.4	42.6	0
10	L6	2025.3	8897.5	3600.0	25	4646.9	5103.9	2705.6	2
15	L6	2600.9	11819.4	3600.0	47	5408.8	5659.2	3546.9	2
20	L6	5617.3	26701.0	3600.0	67	6696.8	7061.3	3600.1	4
25	L6	10572.4	33972.5	3600.03	76	7558.9	7733.3	3600.0	7
30	L6	12983.2	36259.5	3600.0	88	16809.7	9673.3	3600.1	40

it for 53.75%. With this fleet of available vehicles, GIRP (min CO_2) obtained solutions with an average gap of 51%, while the gap for the GIRP (min transp_c + invent_c) model was less than 4.43%. Unexpectedly, CO_2 emissions also reduced by 5.2%. The computational time for GIRP (min CO_2) was approximately 40% more costly.

In Table 8, solutions are presented for 30 vehicles, including four gasoline vehicles, eight E10 vehicles, ten E85 vehicles, two diesel vehicles, two B5 vehicles, and four B20 vehicles. There are 160 instances in total, for which the GIRP (min transp_c + invent_c) model found solutions for all instances, while the GIRP (min CO_2) model did not find a solution for at least one instance in each of the following groups: $n = 10$ and H3; $n = 50$ and H3; $n = 25$, 30, and H6; $n = 50$ and L3; $n = 25$, and L6; and $n = 30$ and L6, within the time limit. In terms of optimality proof, the GIRP (min CO_2) model was able to prove the optimality of 5.62% of the instances, while the GIRP (min transp_c + invent_c) model proved it for 35%. Now, considering a larger fleet of vehicles, the CO_2 emission solutions of the GIRP (min CO_2) model were slightly more polluting due to a gap of over 52%, while the average gap for the GIRP (min transp_c + invent_c) model was 3%. In terms of computational time, GIRP was 33% more costly.

TABLE 8. Computational results for the proposed model, for 30 vehicles.

		GIRP (min CO_2)				IRP (min transp_c + invent_c)			
n	p	CO_2 o. f.	IRP v.	Runtime	Gap %	IRP o. f.	CO_2 v.	Runtime	Gap %
5	H3	522.4	2879.0	1655.2	1	1934.3	1436.5	3.2	0
10	H3	820.8 *	5005.3 *	3600.0 *	24*	3674.7	2156.5	139.3	0
15	H3	812.2	6582.2	3600.0	34	4531.6	2323.5	420.3	0
20	H3	1043.2	8779.6	3600.0	42	5992.3	2673.0	1560.8	1
25	H3	2079.7	15399.7	3600.0	52	7416.6	2699.8	2318.4	1
30	H3	3081.8	18991.7	3600.0	77	8852.7	2728.7	2880.8	1
35	H3	8133.2	31484.5	3600.1	90	9304.3	3159.6	3060.3	2
40	H3	8619.5	31453.2	3600.1	90	10178.6	2939.8	3600.0	2
45	H3	8621.7	32431.9	3600.1	92	11323.3	3864.4	3604.2	3
50	H3	8253.1*	37474.2*	3600.1*	90*	22690.7	7968.7	3600.0	29
5	H6	1416.5	6830.0	3600.0	10	5041.6	3712.6	319.9	0
10	H6	2033.1	11850.1	3600.0	26	8118.6	5438.5	3600.0	3
15	H6	2593.1	17761.0	3600.0	47	10656.9	5927.5	3600.0	2
20	H6	7338.4	38517.4	3600.0	77	13741.3	6684.4	3600.0	4
25	H6	11718.0*	43346.2*	3600.1*	83*	15758.7	8327.9	3600.0	3
30	H6	7696.2*	43232.8*	3601.7*	80*	43311.8	12429.3	3600.1	57
5	L3	522.4	2184.1	1683.7	1	1246.7	1397.8	2.3	0
10	L3	771.6	3094.1	3600.0	21	1863.4	1992.7	188.4	0
15	L3	816.5	4234.8	3591.3	34	2118.0	2305.5	424.9	0
20	L3	1038.1	5345.3	3600.0	41	2589.5	2730.3	1608.2	2
25	L3	2054.8	10758.4	3600.0	50	2948.2	2690.9	2943.6	4
30	L3	3106.1	13472.9	3600.0	77	3071.0	3048.0	2786.0	1
35	L3	8156.6	25150.4	3600.1	90	3332.3	3303.4	2706.6	3
40	L3	8619.5	24813.0	3600.1	90	3831.6	2864.1	3600.2	11
45	L3	8621.7	24833.0	3600.1	92	4540.8	3493.5	3602.6	20
50	L3	8253.1*	29585.5*	3600.1*	90*	12023.4	6892.6	3600.0	36
5	L6	1415.9	4933.9	3600.0	11	3143.7	3638.3	449.3	0
10	L6	2037.0	8478.2	3600.0	26	4726.6	5393.0	3600.0	5
15	L6	2824.2	13407.9	3600.3	51	5411.1	5889.1	3600.0	3
20	L6	7338.4	31625.7	3600.0	77	6958.1	7610.3	3600.0	8
25	L6	11718.0*	35391.1*	3600.1*	83*	17478.6	10524.5	3600.1	27
30	L6	7696.2*	34551.4*	3609.4*	80*	20155.9	8098.4	3600.8	43

Therefore, we conducted computational tests to evaluate the model's performance, considering various fleet configurations that differ in the number of vehicles and types of fuels used. We observed a substantial reduction in emissions of the pollutant gas by the model, aligning with the objectives of our research. However, we noted that, for fleets with more than 18 vehicles, the model faces greater challenges in achieving efficient solutions, particularly in larger instances with numerous clients.

Upon examining the obtained solutions, we emphasized the trade-off between the commitments of the GIRP: CO_2 emissions and IRP costs. When minimizing one objective function, the other function reaches its maximum value and vice versa. Additionally, for each instance where the model achieves an optimal solution for both objectives, we can see the difference between the values of each commitment when the function is minimized and when it is not. From these observations, we identified the opportunity to address the problem through multi-objective methods.

5.3. ϵ-GIRP results

In this section, we present the details of the data and computational experiments of the ϵ-constrained multi-objective method for GIRP. In Section 5.4, we present some perceptions from a managerial point of view.

To find the feasible range of ϵ values, we define the minimum and maximum values, which are denominated by lexicographic points, for each instance, as follows:

- LCO_2: Left lexicographic (lower limit of CO_2): Result obtained by solving the GIRP model (F_{CO_2}), that is, the objective function minimizes the emission of kilograms of CO_2, and
- UCO_2: Right lexicographic (upper limit of CO_2): Result of the amount of CO_2 emitted by solving the GIRP model (F_{IRP}), where we have the cheapest solution from the IRP point of view.

The GIRP model, with only the objective function F_{CO_2}, minimizes the carbon dioxide CO_2 emissions and obtains the minimum value for the interval (LCO_2). For the other extreme of the interval, we solve the GIRP model, with only the objective function F_{IRP}, which minimizes the operational costs of the IRP, and compute the gas emissions, obtaining the maximum value for the interval (UCO_2). In other words, the model is solved twice, using the B&C method, to find the interval $[LCO_2, UCO_2]$. This interval is divided into 5 parts, with each division named ϵ_i (with i ranging from 1 to 5), where a higher value of i indicates a more relaxed constraint ϵ, and a lower value indicates a more restricted space of feasible solutions. The steps to obtain the values of the 5 ϵ's are: (i) calculate the interval size $diff = UCO_2 - LCO_2$; (ii) calculate the step size $step = \frac{diff}{5}$; (iii) for each $i = 1..5$, calculate $\epsilon_i = LCO_2 + (step \times (i-1))$.

Due to the sensitivity of the input data and to validate the proposed approach for ϵ-GIRP, we generated 5 fleet configurations for each instance from [8], such as variations of the instance presented in Table 5, where we defined alternative values for vc^k and C^k for each vehicle/fuel type.

For each calculated ϵ value, the ϵ-GIRP model is executed, and thus we obtain the efficient solutions, which constitute the efficient set.

In this section, computational experiments for the augmented ϵ-constrained method for the GIRP model, called ϵ-GIRP, represented by equations (21)–(23), will be presented. This formulation uses a multicriteria approach and was developed for GIRP, where we prioritize and minimize the costs of IRP restricted to gas emissions limited by an ϵ, in addition to the other sets of constraints of the IRP model. Thus, it is possible to find intermediate solutions within a feasible interval for each value of ϵ.

Values of ϵ were obtained through initial experiments, where we sought the optimal lexicographic points within a time limit of 3600 s. Intermediate non-dominated solutions from the efficient set were found, on average, within 13 min. For instances that required more computational time, the model found the solutions in less than 50 min.

Table 9 presents the solutions for a fleet with 3 vehicles (gasoline, E85, and diesel) and $T = 3$. Each row represents the average of the five generated fleet configurations. The first column represents the instance from the literature, the second column (nC) is the number of customers *i.e.*, $n-1$, followed by the efficient solutions obtained by applying the ϵ-constrained method for 5 values of ϵ. It can be observed that, for each ϵ, the emission function increases while the cost function decreases. Calculating the average variation across all instances in the table between ϵ_1 and ϵ_5, we find that the average of the lexicographic solutions indicates a 58% reduction in emissions while costs increased by 36%.

Therefore, we can observe the behavior of the solutions as the Pareto set is obtained. Thus, we highlight the importance of multi-objective analysis in problems where the involved trade-offs are conflicting.

For a better visualization of this analysis, Figure 1 is the Pareto frontier with 5 non-dominated points for the abs3 instance with 5 customers and a fleet of 3 vehicles, presented in Table 9. The frontier is steep in the proximity of the emission minimum (left side) and becomes flatter towards the IRP cost minimum (right side). This means that small changes in CO_2 emissions lead to sharp differences in costs, while solutions with relatively low costs generate notably different emissions. This suggests that our choice of identifying three Pareto-optimal solutions between the two lexicographic solutions provides a reasonably representative sample.

The Table 10 presents the solutions for a fleet of 6 vehicles, with each vehicle consuming a type of fuel, and Table 11 presents the solutions when the fleet consists of 12 vehicles, with every two vehicles consuming a type

TABLE 9. Results of the ϵ-GIRP model, for a fleet with 3 vehicles and $T = 3$.

		ϵ_1		ϵ_2		ϵ_3		ϵ_4		ϵ_5	
	nC	CO_2 emission	IRP costs	CO_2 emission	IRP costs	CO_2 emission	IRP costs	CO_2 emission	IRP costs	CO_2 emission	IRP costs
abs1	5	386.8	2852.3	543.7	2430.1	920.6	2350.3	1110.6	2097.4	1542.6	1933.7
abs2	5	510.9	2993.0	597.1	2907.7	741.3	2816.3	887.7	2383.1	1250.2	1615.7
abs3	5	1250.1	4911.7	1472.7	4046.7	1830.3	3249.2	2062.3	3116.7	2635.2	2977.1
abs4	5	1036.1	2983.8	1216.2	2590.9	1339.3	2440.9	1602.5	2322.1	1951.6	2113.4
abs5	5	486.0	3009.9	663.2	2336.0	679.4	2303.7	919.0	2093.2	1343.8	1952.3
abs1	10	1072.1	5917.5	1235.7	5186.5	1538.3	4833.9	1658.6	4414.6	2150.6	4032.5
abs2	10	1415.1	6344.6	1698.0	5452.3	1970.3	4652.4	2435.8	4509.5	2840.8	4260.0
abs3	10	908.3	5320.9	1240.3	4432.9	1491.4	3758.4	1923.2	3727.1	2442.9	3614.3
abs4	10	1344.9	5818.6	1575.8	5191.3	1968.0	4182.2	2052.9	4076.5	2655.6	3927.8
abs5	10	1035.7	5293.6	1151.4	4902.0	1300.8	4606.4	1430.9	4375.3	1893.9	4119.1
abs1	15	975.8	6656.1	1268.2	5968.5	1577.2	5395.4	1904.5	4850.9	2297.0	4761.3
abs2	15	1222.2	6689.3	1542.8	5749.1	1772.4	5305.3	2153.0	5086.1	2556.5	4916.6
abs3	15	1134.4	7590.0	1505.8	6744.0	1956.9	6110.2	2442.9	5763.8	3019.4	5504.6
abs4	15	1053.5	6022.7	1363.0	5068.4	1670.2	4872.6	1897.3	4593.4	2374.0	4458.0
abs5	15	1346.8	6221.5	1629.7	5218.8	1921.7	4969.1	2307.9	4646.7	2691.2	4476.6

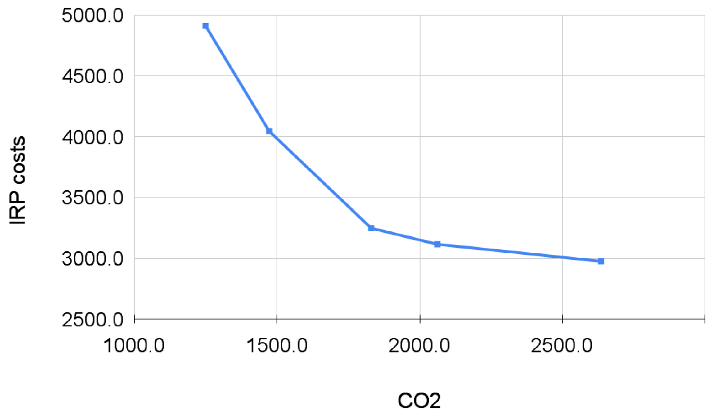

FIGURE 1. Pareto frontier with 5 non-dominated points for the abs3 instance with 5 customers and a fleet of 3 vehicles.

TABLE 10. Results of the ϵ-GIRP model, for a fleet with 6 vehicles and $T = 3$.

		ϵ_1		ϵ_2		ϵ_3		ϵ_4		ϵ_5	
	nC	CO_2 emission	IRP costs	CO_2 emission	IRP costs	CO_2 emission	IRP costs	CO_2 emission	IRP costs	CO_2 emission	IRP costs
abs1	5	383.3	2681.6	568.9	2557.7	825.4	2391.5	1007.7	2271.3	1338.3	2172.0
abs2	5	460.7	3009.8	668.3	2830.2	950.6	2618.5	1161.6	2408.2	1508.7	2259.8
abs3	5	1040.3	5980.8	1448.1	5554.8	1920.2	4706.2	2271.9	4428.7	2961.9	4210.9
abs4	5	1112.0	4331.4	1403.8	3721.3	1676.2	3362.6	2019.5	3090.9	2176.3	2940.7
abs5	5	576.3	3061.5	752.0	2809.2	940.2	2710.4	1063.5	2611.1	1376.3	2394.7

of fuel. The presented solutions are Pareto-optimal, allowing for a better analysis of the trade-off between the objectives. In these tables, the efficient solutions are also presented for each ϵ. For these fleets, we can also observe the increasing and decreasing behavior of the objective functions. When a manager has more choices of vehicles, they are assuming that there may be higher costs associated with the distribution of their products. These analysis possibilities are important for decision-makers as they provide different possible scenarios for the company. This reinforces the importance of multi-objective optimization.

TABLE 11. Results of the ϵ-GIRP model, for a fleet with 12 vehicles and $T = 3$.

		ϵ_1		ϵ_2		ϵ_3		ϵ_4		ϵ_5	
	nC	CO_2 emission	IRP costs	CO_2 emission	IRP costs	CO_2 emission	IRP costs	CO_2 emission	IRP costs	CO_2 emission	IRP costs
abs1	5	380.1	2996.7	574.5	2830.2	790.2	2534.0	1003.9	2391.3	1546.1	2133.5
abs2	5	464.2	3477.5	677.3	3063.8	890.3	2730.5	1091.2	2513.1	1672.3	2218.6
abs3	5	1177.9	6378.1	1743.4	5194.3	2460.3	4455.7	3044.0	4248.3	3375.8	4190.2
abs4	5	1190.4	4350.9	1614.2	3451.4	2065.5	3177.3	2496.4	3025.0	2578.7	3024.7
abs5	5	593.9	3386.3	818.3	2948.2	1045.1	2682.8	1298.6	2525.5	1608.0	2363.0

TABLE 12. ϵ-GIRP management analysis study with a multi-objective approach for 5 clients.

Fleet available	ϵ	CO_2	Transportation cost	Inventory cost	IRP costs	Used fleet
	ϵ_1	334.5	2398	978.2	3376.2	G / E
	ϵ_2	424.1	1921	988.7	2909.7	G / E
	ϵ_3	536.6	1815	988.7	2803.7	G / E / D
G / E85 / D	ϵ_5	837.3	1748	988.7	2736.7	G / D
	ϵ_1	801.0	3148	978.7	4126.7	G / G
	ϵ_2	1097.4	2770	965.5	3735.5	G / G / D
	ϵ_3	1097.4	2770	965.5	3735.5	G / G
	ϵ_4	1674.3	2108	973.3	3081.3	G / G
G / G / D	ϵ_5	2425.1	2087	976.2	3063.2	G / G / D
G / G / G	ϵ_1	975.0	3832	966.6	4798.6	G / G / G
	ϵ_1	459.0	2473	983.7	3456.7	G / E
	ϵ_3	465.7	2207	987.2	3194.2	G / E / E
	ϵ_4	502.7	2180	976.2	3156.2	G / E / E
G / E / E	ϵ_5	526.9	2121	965.7	3086.7	G / E
E / E / E	ϵ_1	733.6	5468	983.9	6451.9	E / E / E
	ϵ_1	537.6	2113	988.7	3101.7	G
	ϵ_4	539.7	2121	972.2	3093.2	G
G / D / D	ϵ_5	650.4	2108	972.2	3080.2	G / D
D / D / D	ϵ_1	4737.5	3562	979.2	4541.2	D / D / D

5.4. Managerial perceptions

In this section, we conducted a study that provides detailed managerial insights. In this study, we performed a more specific analysis, within the multi-objective approach, for two groups of instances, in order to obtain a broader view, with managerial perceptions of the characteristics of the various solutions identified for each instance. This analysis aims to help decision-makers better understand the trade-offs achieved between the two objectives, as well as the impact of emissions resulting from the minimization of operational costs. Thus, the selected type of fuel and vehicle are provided, along with the associated costs over the planning horizon.

For different fleet configurations, supposedly available from the supplier, we present the efficient solutions that provide which vehicles were selected for use, considering the conflicting objectives. In Tables 12 and 13, we provide these analyses. In the first column, we have the fleet available from the supplier; in the second column, labeled as ϵ, we show the ϵ values that managed to limit the CO_2 emissions while minimizing the IRP costs; the next three columns show the separate cost values, for transportation and inventory, and the total costs; in the last column, we present the vehicles used, *i.e.*, the fleet that the model suggests for each determined ϵ.

For example, for a fleet with three vehicles, one running on gasoline (v1), one on E85 (v2), and one on diesel (v3), and to serve a set of 10 customers, as shown in Table 13, the model determines, for each value of ϵ, which vehicles should be used for product delivery. When analyzing the trade-off for ϵ_1, we obtain the lowest gas emissions and the highest operational cost. It can be observed that smaller vehicles running on gasoline and ethanol are used. At time

TABLE 13. ϵ-GIRP management analysis study with a multi-objective approach for 10 clients.

Fleet available	ϵ_i	CO_2	Transportation cost	Inventory cost	IRP costs	Used fleet
	ϵ_1	435.4	2541	1786.4	4327.4	G / E
	ϵ_2	598.6	2412	1783.9	4195.9	G / E
	ϵ_3	1326.0	1897	1793.5	3690.5	E / D
	ϵ_4	1734.7	1856	1777.4	3633.4	G / D
G / E85 / D	ϵ_5	2468.5	1856	1772.9	3628.9	D
	ϵ_1	956.1	3340	1779.2	5119.2	G / G
	ϵ_3	1494.5	1954	1781.1	3735.1	G / D
G / G / D	ϵ_5	2468.5	1856	1772.9	3628.9	D
	ϵ_1	509.5	3146	1772.6	4918.6	G / E / E
G / E85 / E85	ϵ_3	554.1	2498	1782.4	4280.4	G / E
	ϵ_4	609.8	2496	1783.9	4279.9	G / E
	ϵ_5	690.4	2412	1789.8	4201.8	G
G / D / D	ϵ_3	690.4	2412	1789.8	4201.8	G
	ϵ_5	1274.3	2355	1778.6	4133.6	G / D

$t = 1$, v2 visits customers 3, 9, and 10; at time $t = 2$, v1 visits customers 4 and 6, while v2 visits customers 5, 2, 7, and 8; and at time $t = 3$, v2 is used again to visit customer 1. For ϵ_5, we have the other extreme of the Pareto frontier, where the highest gas emissions are produced, and the lowest total cost is obtained. In this case, only the diesel vehicle (v3) is used, which has the largest capacity. From the perspective of discrete-time planning, at time $t = 1$, no deliveries are made, meaning that the stored product quantity is sufficient to meet the demand for that period; at time $t = 2$, v3 visits customers 5, 3, 6, 2, 7, and 8; and at time $t = 3$, v3 visits customers 10, 9, and 4. For intermediate values of ϵ, we obtain intermediate efficient solutions and, consequently, an average total transport capacity. In such cases, the method usually opts for considering a mix of the fleet, including not only lighter and smaller vehicles that emit less CO_2, but also heavier and larger vehicles that emit more CO_2.

To visualize a solution from a managerial perspective, Figure 2 illustrates the first case from Table 13, where the available fleet consists of three heterogeneous vehicles in terms of fuel type and vehicle capacities. For each value of ϵ, the time periods, vehicles, fuels used, and the quantity of products delivered to customers are shown. In the case of eps1 $= \epsilon_1$, we have one of the lexicographic points where the CO_2 emissions are minimized, but operational costs are higher. Therefore, there are more routes, utilizing smaller trucks that consume fuels with lower emission of pollutants. In the first period, T1, the ethanol truck visits customers 3, 9, and 10, delivering 58, 32, and 80 products, respectively. In the T2 period, two routes are performed: one by the gasoline truck, visiting customers 4 and 6 and delivering 26 and 198 products, and the other route is done by the ethanol truck, visiting customers 2, 5, 7, and 8 and delivering 44, 36, 50, and 42 products. In the last period, T3, customer 1 is supplied with 63 products by the ethanol truck. The values eps2, eps3, and eps4 represent points that provide intermediate solutions compared to the extremes. In these three solutions, there is a reduction in the number of routes, and a preference for a larger capacity vehicle, which consequently emits more CO_2. In the case of the other lexicographic point, eps5 $= \epsilon_5$, we have the other extreme with the lowest total cost but the highest gas emission, representing the worst case from a sustainability perspective. It is worth noting that customer 9 receives 96 products in this solution, while in the case of eps4 $= \epsilon_4$, in the T3 period, only 39 products are delivered to the same customer due to the lower capacity of the gasoline truck. It is important to remember that the considered supply policy does not necessarily deliver products to customers until their inventory capacity limits are reached.

Based on the conducted computational experiments, it is observed that when considering the option of having a homogeneous fleet, minimizing the distance traveled is directly proportional to minimizing the CO_2 emissions, demonstrating the advantage of using a heterogeneous fleet. Furthermore, when the fleet is heterogeneous in terms of vehicle capacities and fuels used, there is an overall reduction in the involved trade-offs.

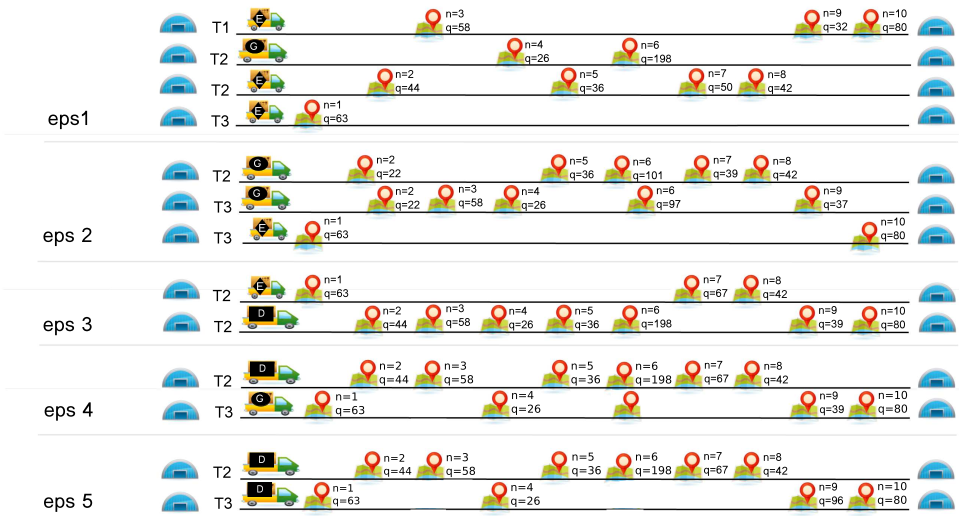

FIGURE 2. Illustration of the solution for ϵ-GIRP with 10 customers and the available fleet of G/E85/D.

6. Conclusion

The literature on logistical optimization models with a sustainable approach has grown significantly in recent years, driven by the concern for addressing the challenges in promoting sustainable development. In this regard, this work presents solutions obtained from a B&C algorithm for solving the developed models. We also considered an exact multi-objective method, the augmented ϵ-constrained, to obtain the trade-off between conflicting objectives in the problem of green IRP with a heterogeneous fleet and different fuels, using our proposed explicit vehicular equation to estimate CO_2 emissions for each vehicle/fuel.

This study aimed to acquire information, gather data, and formulate a modeling approach to minimize CO_2 emissions. The primary aim was to create a method that is both simple and practical, with broad applicability. Additionally, the study sought to investigate the economic implications of reducing carbon dioxide through multi-objective optimization.

Our gas emission formulation, which we consider to be generalizable, allows for the inclusion of various types of vehicles and different fuels, requiring only information on vehicle efficiency. To validate our proposed formulation, which numerically estimates CO_2 emissions, we evaluated the equation on a dataset with over 20,000 vehicles, achieving an error rate of less than 1%. Furthermore, we compared our equation with Machine Learning methods. All estimates proved useful for measuring CO_2 emissions, but the simplicity and quality of our proposed approach were evident compared to Machine Learning methods. Thus, the present GIRP model exhibits interesting characteristics to be applicable in practice.

For the analysis of efficient solutions resulting from the trade-off between the operational costs of the IRP and gas emission, the ϵ-augmented method was implemented for the proposed GIRP model to provide valuable insights for decision-makers and evaluate the behavior of criteria associated with green IRP. The multi-criteria study revealed possibilities of alternative and intermediate solutions for decision-makers. Furthermore, we analyze alternative scenarios that provided managerial insights. Through computational experiments, we observed that, when considering the option of a homogeneous fleet, minimizing distance traveled is directly proportional to minimizing CO_2 emissions, demonstrating the advantage of using a heterogeneous fleet. Moreover, when the fleet is heterogeneous in terms of vehicle capacities and fuels used, there is an overall reduction in the involved trade-offs.

In conclusion, addressing CO_2 emissions is of utmost importance as it contributes to sustainability and the green concept of sustainable development by reducing the emissions of a pollutant gas. This work presents results from the study of the green IRP model, which aims to minimize carbon dioxide emissions for fleets of heterogeneous vehicles with different fuel consumption. In future work, we intend to further investigate this study using heuristic and matheuristic approaches for the problem, as well as explore other methods to handle the multi-objective nature of the problem. Another interesting future research would be to apply this approach and evaluate its results in real logistics settings of green inventory routing.

Acknowledgments

The authors would like to thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for the financial support. To FAPESP (grant 2013/07375-0) for the computational resources provided by the Centro de Ciências Matemáticas Aplicadas à Indústria (CeMEAI) and to the Thematic Project FAPESP (grant 2016/01860-1).

Data availability statement

The data used in this paper is available online in a Github repository: https://github.com/ariannesilvamundim/RAIRO-OR [50].

References

[1] A. Aliano Filho, H. Oliveira Florentino, M.V. Pato, S.C. Poltroniere and J.F.S. Costa, Exact and heuristic methods to solve a bi-objective problem of sustainable cultivation. *Ann. Oper. Res.* **314** (2022) 347–376.

[2] A. Aliano Filho, W.A. Oliveira and T. Melo, Multi-objective optimization for integrated sugarcane cultivation and harvesting planning. *Eur. J. Oper. Res.* **309** (2023) 330–344.

[3] M. Alinaghian, E.B. Tirkolaee, Z.K. Dezaki, S.R. Hejazi and W. Ding, An augmented Tabu search algorithm for the green inventory-routing problem with time windows. *Swarm Evol. Comput.* **60** (2021) 100802.

[4] N. Alkawaleet, Y.-F. Hsieh and Y. Wang, Inventory routing problem with CO2 emissions consideration. In: *Logistics Operations, Supply Chain Management and Sustainability.* Springer, Cham (2014) 611–619

[5] A. Alvarez, P. Munari and R. Morabito, Iterated local search and simulated annealing algorithms for the inventory routing problem. *ITOR* **25** (2018) 1785–1809.

[6] A. Alvarez, J.-F. Cordeau, R. Jans, P. Munari and R. Morabito, Formulations, branch-and-cut and a hybrid heuristic algorithm for an inventory routing problem with perishable products. *Eur. J. Oper. Res.* **283** (2020) 511–529.

[7] C. Archetti and I. Ljubić, Comparison of formulations for the inventory routing problem. *Eur. J. Oper. Res.* **303** (2022) 997–1008.

[8] C. Archetti, L. Bertazzi, G. Laporte and M. Grazia Speranza, A branch-and-cut algorithm for a vendor-managed inventory-routing problem. *Transp. Sci.* **41** (2007) 382–391.

[9] C. Archetti, M. Grazia Speranza, M. Boccia, A. Sforza and C. Sterle, A branch-and-cut algorithm for the inventory routing problem with pickups and deliveries. *Eur. J. Oper. Res.* **282** (2020) 886–895.

[10] M. Asghari and S.M.J. Mirzapour Al-e-hashem, Green vehicle routing problem: a state-of-the-art review. *Int. J. Prod. Econ.* **231** (2021) 107899.

[11] G. Aydin, I. Karakurt and K. Aydiner, Evaluation of geologic storage options of CO2: applicability, cost, storage capacity and safety. *Energy Policy* **38** (2010) 5072–5080.

[12] T. Balamurugan, L.K. Karunamoorthy, N. Arunkumar and D. Santhosh, Optimization of inventory routing problem to minimize carbon dioxide emission. *Int. J. Simul. Model.* **17** (2018) 42–54.

[13] M. Barth, T. Younglove and G. Scora, *Development of a heavy-duty diesel modal emissions and fuel consumption model*, California PATH Research report (2005).

[14] T. Bektaş and G. Laporte, The pollution-routing problem. *Transp. Res. Part B Methodol.* **45** (2011) 1232–1250.

[15] J. Bell Walter, L.M. Dalberto, M.L. Fisher, R. Greenfield, R. Jaikumar and P. Kedia, Improving the distribution of industrial gases with an on-line computerized routing and scheduling optimizer. *Interfaces* **6** (1983) 4–23.

[16] S. Benjaafar, Y. Li and M. Daskin, Carbon footprint and the management of supply chains: insights from simple models. *IEEE Trans. Autom. Sci. Eng.* **10** (2012) 99–116.

[17] L. Bertazzi and M. Grazia Speranza, Inventory routing problems: an introduction. *EURO J. Transp. Logist.* **1** (2012) 307–326.

[18] J.E.H. Branco, D.B. Bartholomeu, P.N.A. Junior and J.V.C. Filho, Evaluation of the economic and environmental impacts from the addition of new railways to the Brazilian's transportation network: an application of a network equilibrium model, *Transp. Policy* **124** (2022) 61–69.

[19] W.-R. Bretzke, Sustainable logistics: in search of solutions for a challenging new problem. *Logist. Res.* **3** (2011) 179–189.

[20] G.P. Cachon, Retail store density and the cost of greenhouse gas emissions. *Manag. Sci.* **60** (2014) 1907–1925.

[21] C. Cheng, P. Yang, M. Qi and L.-M. Rousseau, Modeling a green inventory routing problem with a heterogeneous fleet. *Transp. Res. Part E Logist. Transp. Rev.* **97** (2017) 97–112.

[22] C. Cheng, M. Qi and L.-M. Rousseau, Fuel consumption optimization model for the multi-period inventory routing problem. *Transp. Res. Rec.* **2672** (2018) 59–69.

[23] L.C. Coelho and G. Laporte, A branch-and-cut algorithm for the multi-product multi-vehicle inventory-routing problem. *Int. J. Prod. Res.* **51** (2013) 7156–7169.

[24] L. Coelho, J.-F. Cordeau and G. Laporte, Thirty years of inventory routing. *Transp. Sci.* **48** (2013) 1–19.

[25] Y. Collette and P. Siarry, Multiobjective Optimization: Principles and Case Studies. Springer Science & Business Media, Heidelberg (2004).

[26] P.M. Cox and M.S. Williamson, Emergent constraint on equilibrium climate sensitivity from global temperature variability. *Nature* **553** (2018) 319–322.

[27] M. Darvish, C. Archetti and L.C. Coelho, Trade-offs between environmental and economic performance in production and inventory-routing problems. *Int. J. Prod. Econ.* **217** (2019) 269–280.

[28] E. Deakin, Sustainable development and sustainable transportation: strategies for economic prosperity, environmental quality, and equity. IURD Working Paper Series (2001).

[29] R. Dekker, J. Bloemhof and I. Mallidis, Operations research for green logistics–an overview of aspects, issues, contributions and challenges. *Eur. J. Oper. Res.* **219** (2012) 671–679.

[30] E. Demir, T. Bektaş and G. Laporte, A comparative analysis of several vehicle emission models for road freight transportation. *Transp. Res. Part D Transp. Environ.* **16** (2011) 347–357.

[31] E. Demir, T. Bektaş and G. Laporte, An adaptive large neighborhood search heuristic for the pollution-routing problem. *Eur. J. Oper. Res.* **223** (2012) 346–359.

[32] B.K. Dey, J. Park and H. Seok, Carbon-emission and waste reduction of a manufacturing-remanufacturing system using green technology and autonomated inspection. *RAIRO-OR* **56** (2022) 2801–2831.

[33] K.M. Ferreira, T.A. de Queiroz and F.M.B. Toledo, An exact approach for the green vehicle routing problem with two-dimensional loading constraints and split delivery. *Comput. Oper. Res.* **136** (2021) 105452.

[34] C. Franco, E.R. López-Santana and G. Méndez-Giraldo, A column generation approach for solving a green bi-objective inventory routing problem. In: *Advances in Artificial Intelligence - IBERAMIA 2016*. Springer International Publishing, Cham (2016) 101–112.

[35] F. Goodarzian, V. Kumar and P. Ghasemi, Investigating a citrus fruit supply chain network considering CO2 emissions using meta-heuristic algorithms. *Ann. Oper. Res.* (2022). DOI: 10.1007/s10479-022-05005-7.

[36] Y. Haimes, On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE Trans. Syst. Man Cybern.* **3** (1971) 296–297.

[37] J.S. Hoffman, P.U. Clark, A.C. Parnell and F. He, Regional and global sea-surface temperatures during the last interglaciation. *Science* **355** (2017) 276–279.

[38] J.L. Jiménez, J. Valido and N. Molden, The drivers behind differences between official and actual vehicle efficiency and CO2 emissions. *Transp. Res. Part D Transp. Environ.* **67** (2019) 628–641.

[39] P. Karakostas, A. Sifaleras and M.C. Georgiadis, Adaptive variable neighborhood search solution methods for the fleet size and mix pollution location-inventory-routing problem. *Expert Syst. Appl.* **153** (2020) 113444.

[40] V. Kishore Ayyadevara, Pro Machine Learning Algorithms. Apress, Berkeley, Springer (2018).

[41] C. Koc, T. Bektas, O. Jabali and G. Laporte, The fleet size and mix pollution-routing problem. *Transp. Res. Part B Methodol.* **70** (2014) 239–254.

[42] M. Kumari, P.K. De, P. Narang and N.H. Shah, Integrated optimization of inventory, replenishment, and vehicle routing for a sustainable supply chain utilizing a novel hybrid algorithm with carbon emission regulation. *Expert Syst. Appl.* **220** (2023) 119667.

[43] Y.-F. Lin, T.-Y. Chang, W.-R. Su and R.-K. Shang, IoT for environmental management and security governance: an integrated project in Taiwan. Sustainability **14** (2022) 217.

[44] Z. Liu, P. Ciais, Z. Deng, R. Lei, S.J. Davis, S. Feng, B. Zheng, D. Cui, X. Dou, B. Zhu and R. Guo, Near-real-time monitoring of global CO2 emissions reveals the effects of the COVID-19 pandemic. *Nat. commun.* **11** (2020) 1–12.

[45] B. Mahesh, Machine learning algorithms-a review. *Int. J. Sci. Res. (IJSR)* **9** (2020) 381–386.

[46] E.B. Mariano, J.A. Gobbo Jr., F. de Castro Camioto and D.A. do Nascimento Rebelatto, CO2 emissions and logistics performance: a composite index proposal. *J. Clean. Prod.* **163** (2017) 166–178.

[47] G. Mavrotas, Effective implementation of the e-constraint method in multi-objective mathematical programming problems. *Appl. Math. Comput.* **213** (2009) 455–465.

[48] G.J.L. Micheli and F. Mantella, Modelling an environmentally-extended inventory routing problem with demand uncertainty and a heterogeneous fleet under carbon control policies. *Int. J. Prod. Econ.* **204** (2018) 316–327.

[49] F. Misni, L.S. Lee and N.I. Jaini, Multi-objective hybrid harmony search-simulated annealing for location-inventory-routing problem in supply chain network design of reverse logistics with CO2 emission. *J. Phys. Conf. Ser.* **1** (2021) 012054.

[50] A.A.S. Mundim, M.O. Santos and R. Morabito, Sustainable solutions analysis of a bi-objective green inventory routing problem with heterogeneous fleet and different types of fuels. *RAIRO-OR* **59** (2025). Available at: https://github.com/ariannesilvamundim/RAIRO-OR. Accessed on: January 30, 2025.

[51] Natural Resources Canada, Learn the Facts: Fuel Consumption and CO2. Government of Canada (2014).

[52] N. Niroomand, C. Bach and M. Elser, Segment-based CO2 emission evaluations from passenger cars based on deep learning techniques. *IEEE Access* **9** (2021) 166314–166327.

[53] R.K. Pachauri and L.A. Meyer, IPCC, 2014: Climate Change 2014. In: *Intergovernmental Panel on Climate Change* (2014).

[54] G. Pinto and M.T. Oliver-Hoyo, Using the relationship between vehicle fuel consumption and CO2 emissions to illustrate chemical principles. *J. Chem. Educ.* **85** (2008) 218.

[55] L. Qin, L. Miao, Q. Ruan and Y. Zhang, A local search method for periodic inventory routing problem. *Expert Syst. Appl.* **41** (2014) 765–778.

[56] M. Rabbani, K.R. Mokarrari and N. Akbarian-saravi, A multi-objective location inventory routing problem with pricing decisions in a sustainable waste management system. *Sustaina. Cities Soc.* **75** (2021) 103319.

[57] M. Rahbari, B. Naderi and M. Mohammadi, Modelling and solving the inventory routing problem with CO2 emissions consideration and transshipment option. *Environ. Process.* **5** (2018) 649–665.

[58] M. Rahbari, B. Naderi and M. Mohammadi, Modeling a multi-objective green inventory-routing problem with transshipment option. In: *14th International Industrial Engineering Conference (IIEC 2018)*. Tehran, Iran (2018).

[59] M. Rahbari, A.A. Khamseh, Y. Sadati-Keneti and M.J. Jafari, A risk-based green location-inventory-routing problem for hazardous materials: NSGA II, MOSA, and multi-objective black widow optimization. *Environ. Dev. Sustain.* **24** (2022) 2804–2840.

[60] C.M. Schenekemberg, T.A. Guimarães, A.A. Chaves and L.C. Coelho, A three-front parallel branch-and-cut algorithm for production and inventory routing problems. *Transp. Sci.* **58** (2023) 687–707.

[61] S. Seuring, A review of modeling approaches for sustainable supply chain management. *Dec. Support Syst.* **54** (2013) 1513–1520.

[62] Y. Shi, Multiple Criteria and Multiple Constraint Levels Linear Programming: Concepts, Techniques and Applications. World Scientific Publishing Company, Singapore (2001)

[63] R. Smokers, L. Tavasszy, M. Chen and E. Guis, Options for Competitive and Sustainable Logistics. Emerald Group Publishing Limited (2014)

[64] M. Soysal, J.M. Bloemhof-Ruwaard, R. Haijema and J.G.A.J. van der Vorst, Modeling an inventory routing problem for perishable products with environmental considerations and demand uncertainty. *Int. J. Prod. Econ.* **164** (2015) 118–133.

[65] M. Soysal, M. Çimen, S. Belbağ and E. Toğrul, A review on sustainable inventory routing. *Comput. Ind. Eng.* **132** (2019) 395–411.

[66] H.M. Stellingwerf, G. Laporte, F.C.A.M. Cruijssen, A. Kanellopoulos and J.M. Bloemhof, Quantifying the environmental and economic benefits of cooperation: a case study in temperature-controlled food logistics. *Transp. Res. Part D Transp. Environ.* **65** (2018) 178–193.

[67] J. Tollefson, COVID curbed carbon emissions in 2020 - but not by much. *Nature* **589** (2021) 343.

[68] S. Treitl, P.C. Nolz, W. Jammernegg, Incorporating environmental aspects in an inventory routing problem. A case study from the petrochemical industry, *Flex. Serv. Manuf. J.* **26** (2014) 143–169.

[69] Q. Wang and F. Zhang, The effects of trade openness on decoupling carbon emissions from economic growth–evidence from 182 countries. *J. Clean. Prod.* **279** (2021) 123838.

[70] C. Wang, W. Cai, X. Lu and J. Chen, CO2 mitigation scenarios in China's road transport sector. *Energy Convers. Manag.* **48** (2007) 2110–2118.

[71] Y. Wen, R. Wu, Z. Zhou, S. Zhang, S. Yang, T.J. Wallington, W. Shen, Q. Tan, Y. Deng and Y. Wu, A data-driven method of traffic emissions mapping with land use random forest models. *Appl. Energy* **305** (2022) 117916.

[72] R. Wu and Z. Xie, Identifying the impacts of income inequality on CO2 emissions: empirical evidences from OECD countries and non-OECD countries. *J. Clean. Prod.* **277** (2020) 123858.

[73] W. Wu, W. Zhou, Y. Lin, Y. Xie and W. Jin, A hybrid metaheuristic algorithm for location inventory routing problem with time windows and fuel consumption. *Expert Syst. Appl.* **166** (2021) 114034.

[74] Y. Xiao, Q. Zhao, I. Kaku and Y. Xu, Development of a fuel consumption optimization model for the capacitated vehicle routing problem. *Comput. Oper. Res.* **39** (2012) 1419–1431.

[75] L. Xu, C. Wang, Z. Miao and J. Chen, Governmental subsidy policies and supply chain decisions with carbon emission limit and consumer's environmental awareness. *RAIRO-OR* **53** (2019) 1675–1689.

[76] D. Zhang, R. He, S. Li and Z. Wang, A multimodal logistics service network design with time windows and environmental concerns. *PLOS ONE* **12** (2017) e0185001.

Appendix A.

Based on [13] and adapted from [21], we present equation (A.1), which quantifies the emission of CO_2 in kilograms for a vehicle of type k traveling a distance d (km) at a speed v (km/h) and consuming:

$$\lambda\left(\frac{K^k\, N^k\, V^k\, d}{v} + M^k\, \gamma^k\, \alpha\, d + \beta^k\, \gamma^k\, dv^2\right)\sigma. \tag{A.1}$$

As for the parameters, $\lambda = \epsilon/(\kappa\,\psi)$, $\gamma^k = 1/(1000\, n_{tf}^k\, \eta)$, $\alpha = \tau + g\,\mathrm{sen}(\theta) + g\, C_r \cos(\theta)$, $\beta^k = 0.5\, C_d^k\, \rho\, A^k$. $M^k = w^k + Q^k$, representing the total weight of the fully loaded vehicle, which is the sum of the curb weight and the payload, respectively. Tables B.1 and B.2 contain information about the parameters needed for the mentioned formulation defined by F^k in equation (A.1). Table B.1 provides information about common vehicle characteristics such as air density, unit fuel cost, gravitational constant, with distance d and velocity v as variables, and Table B.2 contains specific vehicle parameters like curb weight, engine friction factor, engine speed, engine displacement, frontal surface area, among others.

Appendix B.

In this appendix, we briefly present the seven regression algorithms used in this work to estimate CO_2 emissions. We also present the error metrics for comparing the algorithms.

Used machine learning algorithms

All algorithms are supervised learning methods, Mahesh [45], which means they can learn from data. In this work, we separated a training dataset that was used to create the models. This dataset contains seven independent variables (or features) and the target variable (the CO_2 emissions). The features include: X_{ES}, representing the engine size in liters (Engine Size); X_C, representing the number of cylinders (Cylinders); X_{FC}, representing the average fuel consumption (Fuel Consumption Comb); and the fuel type is represented by four binary variables (X_D for diesel, X_G for gasoline, X_{GP} for premium gasoline, and X_E for ethanol). For all cases, $X_D + X_G + X_{GP} + X_E = 1$. Next, we present all the methods, Kishore Ayyadevara [40], along with some details about parameter calibration.

Decision tree

Table B.1. Definition of common vehicle parameters.

Notation	Description	Typical values
ϵ	Fuel-to-air mass ratio	1
g	Gravitational constant (m/s^2)	9.81
ρ	Air density (kg/m^3)	1.2041
C_r	Coefficient of rolling resistance	0.01
η	Efficiency parameter for diesel engines	0.45
σ	CO_2 Emitted by unit fuel consumption (kg/L)	2.669
κ	Heating value of a typical diesel fuel (kJ/g)	44
v	Speed (m/s)	–
ψ	Conversion factor (g/s to L/s)	737
θ	Road angle	0
τ	Acceleration (m/s^2)	0
d	Travel distance (km)	–

Source. Adapted from [21].

Table B.2. Definition of specific vehicle parameters.

Notation	Description	Light vehicle	Medium vehicle	Heavy vehicle
w^k	Curb weight (kg)	4672	6328	13154
Q^k	Maximum payload (kg)	2585	5080	17236
K^k	Engine friction factor (kJ/rev/L)	0.25	0.2	0.15
N^k	Engine speed (rev/s)	39	33	30.2
V^k	Engine displacement (L)	2.77	5	6.66
C_d^k	Coefficient of aerodynamics drag	0.6	0.6	0.7
A^k	Frontal surface area (m^2)	9	9	9.8
n_{tf}^k	Vehicle drive train efficiency	0.4	0.45	0.5

Source. Adapted from [21].

A decision tree is an algorithm that constructs a regression model in the structure of a tree. This is one of the most commonly used algorithms in Machine Learning, and its basic idea is to divide a complex problem into a set of simple decisions. We can formally define it as a directed acyclic graph formed by a set of leaf nodes and a set of split nodes. The split nodes divide the input data into two sets, and this conditional test will always be based on some input feature. On the other hand, the leaf nodes return the value of the variable of interest as the average of the examples contained within them. To find the best configuration for our decision tree, we test the maximum depth of the tree for heights ranging from 1 to 20. To measure the quality of the splits and the responses, we consider the mean squared error, mean squared error with Friedman's improvement score, and mean absolute error. The best configuration is a tree with a height of 20, and the splitting function is the mean squared error function.

Random Forest

The major advantage of using a Random Forest is the fact that it reduces the overfitting of decision trees. Typically, a decision tree can overfit to the training data and fail to generalize the knowledge to new instances. A Random Forest is an ensemble, which means it is a collection of decision trees where each tree is trained on a random sample with replacement from the original dataset. For the construction of each node in a tree, a randomly selected subset of features is considered. Finally, the decision consists of an average of the results from all the trees. This way, the Random Forest produces uncorrelated trees, and the average reduces the variability in predictions for unseen data during training. We tested trees with depths ranging from 1 to 10, with a number of features ranging from 1 to the total number of features. The number of trees considered in the predictor was: 50, 100, 200, 500, 750, 1000. The best configuration was with a height of 9, 100 trees, and all features as input.

K-Nearest Neighbors

The K-Nearest Neighbors (KNN) is a distance-based algorithm. It stores all the training data in an n-dimensional space and simply calculates, for each new instance, which are the K closest data points in the training set and returns the average of those cases. Despite its simplicity, depending on the problem, it is a very useful algorithm, and its main weakness is the need to store and calculate the distance to all the points in the training set. In this work, we used the Euclidean distance and tested all values of K from 1 to 1000 instances. The best result was achieved with $K = 4$.

Linear regression

A linear regression is an algorithm that adjusts a set of coefficients associated with features to minimize the sum of squared errors between the expected values (dependent variable) and the values predicted by the linear approximation. The better the fit, the smaller the residual error in the data. The great advantage of linear regression is its easy interpretability.

Lasso e Ridge

Lasso (Least Absolute Shrinkage and Selection Operator) and Ridge are widely used regularization techniques in linear regression to deal with overfitting problems and improve the model's generalization ability. Both techniques add a regularization term to the objective function of the model, which controls the model's complexity by penalizing larger regression coefficients.

Lasso: In Lasso regression, the regularization term is based on the L1 norm of the regression coefficients. The objective function of Lasso regression is defined as:

$$\text{minimize } \frac{1}{2}\text{RSS} + \alpha\|w\|_1,$$

where:

- RSS is the residual sum of squares;
- α is the hyperparameter that controls the strength of regularization;
- w are the regression coefficients.

Ridge: In Ridge regression, the regularization term is based on the L2 norm of the regression coefficients. The objective function of Ridge regression is defined as:

$$\text{minimize } \frac{1}{2}\text{RSS} + \alpha\|w\|_2^2,$$

where:

- RSS is the residual sum of squares;
- α is the hyperparameter that controls the strength of regularization;
- w are the regression coefficients.

The main difference between them lies in how they penalize the coefficients. Lasso uses the L1 norm, leading some coefficients to zero and creating a more sparse model. Ridge uses the L2 norm, reducing the magnitude of the coefficients without completely nullifying them. Lasso also has the property of automatically selecting the most relevant features, while Ridge does not. However, Lasso is more sensitive to small changes in the data, making feature selection less stable. Ridge is more stable and has a closed-form analytical solution, while Lasso requires an iterative procedure to find the solution. In summary, Lasso is useful for automatic feature selection and creating sparser models, while Ridge is suitable for reducing the magnitude of coefficients while maintaining stability. For the Lasso technique, the following values were used for the hyperparameter Alpha: we generated a list of 100 equally spaced values between 5^9 and 5^{-11}. As for the Ridge technique, the hyperparameter Lambda: we generated 100 equally spaced values in the range from 5^9 to 5^{-3}.

Epsilon-Support Vector Regression SVR

Epsilon-Support Vector Regression (SVR) is a regression model that uses the Support Vector Machines (SVM) technique to make predictions in regression problems. One of the distinctive features of SVR is the ability to use kernels, which are data transformations into a higher-dimensional space. These transformations allow SVR to find an approximately linear relationship between the features and the value of the target variable in the new space. The goal of SVR is to find a function $h(x)$ that minimizes the difference between the predictions and the actual values, within a tolerance margin defined by the epsilon hyperparameter. The objective function of SVR can be expressed as:

$$\min_{w,b,\xi,\xi^*} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*)$$

subject to:

$$y_i - w^T\Phi(x_i) - b \leq \epsilon + \xi_i$$
$$w^T\Phi(x_i) + b - y_i \leq \epsilon + \xi_i^*$$
$$\xi_i, \xi_i^* \geq 0, i = 1, \ldots, n.$$

where:

– w is the weight vector that defines the regression function;
– b is the bias term;
– ξ_i and ξ_i^* are slack variables;
– C is a hyperparameter that controls the penalty for margin violations;
– ϵ is the width of the tolerance margin.

The values of the hyperparameters C used were 0.01, 0.1, 0.4, 5, 10, 20, 30, 40, and 50. The adopted value for epsilon was 0.2. We used the Radial basis function kernel.

Used error metrics

The 5 error measures used and presented here are widely used to assess the quality of the proposed methods. They are: maximum error (ME), mean gamma deviation (MGD), mean absolute percentage error (MAPE), mean squared error (MSE), and R-squared (r2). Below is a description of each measure.
Maximum error (ME): also known as maximum absolute error, it is the difference between the actual value and the predicted value that exhibits the highest absolute deviation compared to the other points. In other words, it is the largest error observed among all predictions made by the model.
Mean gamma deviation (MGD): also known as mean absolute error, it is a measure that evaluates the overall performance of the model by calculating the average of the absolute deviations between the predicted values and the actual values. It is a measure of how close or far the predictions are from the actual values, regardless of their direction.

Both ME and MGD are error measures, and the smaller their values, the better the model's performance. These metrics are commonly used to assess the accuracy and precision of prediction models, such as regression models, which are used in this article.
Mean absolute percentage error (MAPE): This is another interesting metric to use, typically used in management reports because the error is measured as a percentage. This allows for comparisons between model percentage errors across products. MAPE is calculated as follows:

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\%$$

where:

– n is the number of observations.
– y_i is the actual value.
– $\hat{y}_i$ is the predicted value.

Mean squared error (MSE): is commonly used to assess the accuracy of models and gives greater weight to larger errors. This is because, when calculated, each error is squared individually, and then the average of these squared errors is calculated. MSE is calculated as follows:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

where:

– n is the number of observations;
– y_i is the actual value;
– $\hat{y}_i$ is the predicted value.

R-squared (r2): also known as the coefficient of determination, is a measure that indicates the proportion of the variance in the dependent variable that can be explained by the model. R-squared ranges from 0 to 1, where 0 indicates that the model does not explain any variability, and 1 indicates that the model explains all the variability in the data. It is calculated as follows:

$$\text{r2} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

where:

– n is the number of observations;
– y_i is the actual value;
– $\hat{y}_i$ is the predicted value;
– $\bar{y}$ is the mean of the actual values.