Japan Bilingual Publishing Co.

Transportation Development Research
https://ojs.bilpub.com/index.php/tdr

ARTICLE

Risk Analysis of Air Cargo Discrepancies within the Air Freight Supply Chain by Utilizing Innovative Batch Mode, Blockchain, and Clustering Methods

Muhammad Salman

Computer Science Department, University of Peshawar,University Road, Peshawar 25120, KPK, Pakistan

ABSTRACT

There is no integration mechanism available between independent systems (sales and shipment), so it creates uncertainty in the air freight industry. Regarding the problems, freight forwarders or airline sales officers can make a manifest for the shippers; however, various airlines plan different routes for shipping services. But unfortunately, it provides transaction discrepancies between independent systems. The proposed methods integrate two datasets in a central repository and detect discrepancies between sales and shipment systems using novel batch mode, update, and blockchain methods. Furthermore, it provides output to clustering algorithms to highlight the risk factors or disruptions in two- or three-dimensional forms. In other words, the proposed mechanisms establish a bridge between freight forwarders and airline industries, so it minimizes the transaction discrepancies in the air freight system. Besides this, a managerial insights investigation revealed that data consistency, real-time transactions, technology, and supply chain disruptions are the most important topics. The findings indicate that the air freight industry increases the revenue and boosts the economy after implementing the proposed methods in the production environment. The proposed methods compute the theoretical and practical time complexities, which are about T (n) = n, so they are suitable for production

*CORRESPONDING AUTHOR:
Muhammad Salman, Computer Science Department, University of Peshawar, University Road, Peshawar 25120, KPK, Pakistan; Email: msalman@uop.edu.pk

ARTICLE INFO
Received: 19 September 2025 | Revised: 19 October 2025 | Accepted: 26 October 2025 | Published Online: 3 November 2025
DOI: https://doi.org/10.55121/tdr.v3i2.983

CITATION
Salman, M., 2025. Risk Analysis of Air Cargo Discrepancies within the Air Freight Supply Chain by Utilizing Innovative Batch Mode, Blockchain, and Clustering Methods. Transportation Development Research. 3(2): 72–104. DOI: https://doi.org/10.55121/tdr.v3i2.983

environments. Statistically, it provides a significant result. Finally, it resolves inconsistent issues between independent systems. It establishes a balance in the air freight supply chain industry.

Keywords: Airline and Freight Forwarders; Batch Mode and Blockchain Methods; Clustering Algorithms; Detect Discrepancies between Sales and Shipment Systems; Supply Chain Disruptions; Uncertainty in the Air Freight Industry

1. Introduction

The structure of the supply chain (SC) significantly influences factors such as price, quality, customer perception, and a company's capacity to respond to market opportunities [1,2]. A supply chain is constantly evolving, characterized by a continuous flow of products, data, and money across various phases [2]. Air cargo has become more significant and visible in supply chain planning and management as a result of the COVID-19 pandemic, geopolitical unrest, disruptions in global supply chains, and the growth of e-commerce. Air cargo has often been overlooked within the aviation industry and is frequently poorly managed in many supply chains, mainly due to its association with combination carriers. Concerns about sustainability, an impending recession, disruptions to global supply chains both physically and virtually, as well as the drone revolution, all contribute to an ever-changing environment in which one thing is for certain: there will be more uncertainty, which must be effectively managed for supply chains to continue operating [3,4]. It aims to assess whether gravity models are robust enough to predict and accurately account for significant economic shocks like the Global Financial Crisis (GFC). It suggests that US demand for air freight is highly sensitive to transport costs, competition from sea freight, and consumer spending patterns of perishable, low-value, and high-value commodities across the 19 commodity groups examined, rather than manufacturing income or factors associated with product origin, which is especially interesting given the US administration's recent protectionist rhetoric [5]. It investigates the impact of bilateral air connectivity on bilateral service trade flows. The trade data covers 'commercial', 'transport', 'travel', and 'government' services. It used the Chinese data to create a reduced-form gravity model. To address the endogeneity issue between bilateral air connectivity and service trade variables, it used an instrumental variable (IV) approach. The key results are (a) increasing the number of direct routes can significantly promote bilateral service export and import trades; (b) average route-level traffic density has only marginal positive effects; (c) improving air connectivity would enlarge China's overall service trade deficit because transport and travel service imports are promoted more than their exports; (d) 'commercial' service exports can be stimulated more than imports, making a capable commercial service trade surplus by improving bilateral air connectivity [6,7].

The freight industry is going through tremendous growth in volume thanks to the technological revolution in e-commerce and global trade liberalization. An increase in volume also corresponds to an increase in fraud cases involving false shipment declarations and smuggling. The adoption of traditional fraud detection strategies by customs and shipment companies will become unfeasible as the volume increases significantly. Routine inspections are not nearly as effective as data-driven fraud detection. The problem with data-driven detection frequently depends on the availability of data and the different fraud techniques that criminals employ to perpetrate fraud involving shipments [8]. The shipping domain's most pressing problem is obtaining effective real-time results from a sizable dataset. Since efficiency increases with dataset size, researchers must design a solution architecture that can process an ideal volume of the target dataset at a speed that allows for real-time execution. Accuracy and real-time performance are crucial in the express shipping industry because a shipment's life cycle only lasts three to five days, depending on its weight and location [8]. When fraudulent transactions are made to appear like legitimate data, they become true negative cases. This is known as overlapping data. The problem with the skewed distribution is that the ratio of fraudulent cases is very low, which might not be enough to train algorithms for a supervised classification-based fraud detection system (FDS). Data quality problems also need to be studied, as this factor directly impacts the efficiency of fraud detection [8]. Other factors that affect fraud in the shipping domain include shipment weight, payment method, and customer profile, which were not assessed because of insufficient real production data. The parameters mentioned can be considered for achieving more accurate results in future research [8]. It explains an integrated modeling method for improving flight operations at major commercial airports [9].

In Demand-Driven Services (DDS), a service loop is a fundamental structure that encompasses the service worker, the service providers, and the associated service targets. The role of the service workers is to transport either people or parcels from the providers to the designated target locations. Various planning tasks within DDS can be categorized into two distinct stages: (1) Dispatching, which involves creating service loops based on demand and supply distributions, and (2) Routing, which determines the specific serving orders within the established loops. Generating high-quality strategies in both stages is important for developing DDS, but it faces several challenges [10]. International trade serves as a crucial explanatory variable for air freight demand, while air freight also facilitates international trade. However, their roles differ significantly and depend on whether they are viewed as dependent or explanatory variables, as well as the level of economic development in the countries being analyzed. Since the onset of the COVID-19 pandemic, air cargo has been significantly negatively affected, but it appears to have had a negligible effect on international business [11]. The COVID-19 pandemic impacted nearly every aspect of airport operations, highlighting the significance of airport agility. Due to its importance, we must understand the role of agility at airports before and during unexpected disruptions [12]. It also explains the growing connection between air freight and (cross-border) e-commerce, highlighting how the demand for fast and reliable international shipment has reshaped logistics networks, freight operations, and aviation approaches. It examines the evolving roles of key aviation stakeholders, including freight forwarders, airports, integrators, airlines, and e-commerce marketplaces, as they adapt to the rapid growth of e-commerce-driven air cargo [13]. Cross-border e-commerce (CBEC) refers to online transactions that occur between sellers and consumers in different countries. Despite the growing volatility in international trade, the CBEC market is experiencing continued growth. Therefore, it is essential to understand how companies navigate logistics uncertainties. (1) External uncertainties enhance logistics flexibility, while internal uncertainties do not have a significant direct effect; (2) internal uncertainties negatively impact logistics information system (LIS) utilization, whereas external uncertainties have a slightly positive relationship; and (3) LIS utilization mediates the negative relationship between internal uncertainty and flexibility [14]. It presented that challenges related to technology, government policies, supplier and stakeholder engagement, and organizational and management support have a significant correlation with sustainable logistics practices [15]. Global supply chains are facing increasing disruptions from security-related risks, including cargo theft, illicit trade, document forgery, and cyber attacks—challenges that pose serious threats to sustainable development, particularly in vulnerable and emerging economies [16]. The critical importance of cargo forgery prompted the federal government to collect data to aid in prevention and response efforts. The analysis focuses on cargo theft data reports submitted to the FBI from 2013 to 2021, aiming to highlight the significant discrepancies between these reports and the criteria the FBI uses to classify valid cargo thefts [17]. Some important documents in the shipping industry are still transmitted on paper, such as airway bills. If the freight industry continues to use these methods to transmit information, it will consume a lot of workforce and time, which will cause many disruptions. In other words, the disappearance and deletion of data will be the main reason for the backward development of the shipping industry [18]. However, analysis of logistic and economic indicators requires having efficient IT tools. Effective management and analysis of the supply chain must occur in real time. This necessitates implementing specific organizational changes and utilizing modern tools and IT systems designed for the rapid processing of large data volumes, as well as for transparent visualization of that data [19].

1.1. Blockchain and Air Freight

Although the airline industry has experienced significant growth over the past few decades, it is still facing numerous challenges and constantly spreading due to economic recessions, technological advancements, volatility in oil prices, changing customer preferences, and global events (e.g., the COVID-19 pandemic), which have impacted demand for air transport and disrupted operations [20]. In other words, the COVID-19 pandemic has resulted in restrictions on the transportation and movement of goods and materials, particularly on routes that pass through restricted or containment areas [21]. Shipment transportation is vital to the global economy, linking manufacturers, suppliers, and consumers across various geographic regions. However, the industry encounters several challenges, including information asymmetry, a demand for greater transparency, and processes that are often inefficient and costly. To address these challenges, it provides an overview of the current and potential applications of blockchain technology in the shipment transportation industry. Various blockchain pilot projects have been launched globally, focusing on different use cases such as document management, container tracking, alternative payment methods, and enhancing supply chain visibility for stakeholders [22]. The air cargo supply chain aims to facilitate rapid shipment movement through coordinated efforts among various stakeholders. However, the lack of transparency throughout the end-to-end supply chain, along with complex digital connectivity, restricts effective peer-to-peer communication in the fragmented air cargo industry. Blockchain technology can improve communication transparency by streamlining the flow of messages across multiple stages through the use of smart contracts. Additionally, it can serve as an intermediary to simplify digital connections among stakeholders in the air cargo industry. Despite the advantages of using blockchain for air freight messaging, its status as an emerging technology raises technical concerns. Issues related to performance, privacy, transparency, and interoperability can hinder the practical adoption of blockchain systems [23]. Moreover, blockchain ensures data authenticity, real-time tracking systems, and predictive maintenance models in aircraft maintenance and safety transactions. Blockchain technology leverages traditional decentralized ledgers to reduce operational inefficiencies, enhance security, and foster trust among aviation stakeholders. Its implementation in supply chain management further promotes transparency, prevents counterfeiting, and facilitates seamless collaboration across the industry. Overall, blockchain is transforming the aviation and airline sectors by improving transparency, privacy, and efficiency [24]. Blockchain provides a secure and tamper-proof method for recording transactions, effectively reducing fraud and enhancing data transparency by 20% [25]. The combination of digital twin technology and blockchain fosters collaboration among various stakeholders in the supply chain network. This integration can lead to increased productivity, improved visibility within the supply chain, and a reduction in mistakes and fraud [26]. Blockchain technology represents a digital advancement that has significantly contributed to the growth of the transportation sector in recent decades by enhancing transparency, improving privacy, and increasing operational efficiency. However, like many emerging technologies, the adoption of blockchain has encountered several challenges. These include high implementation costs, insufficient understanding among both the public and professionals, resistance to change, and scalability issues. The findings from this study will aid individuals involved in cargo operations, transportation, and road management by reducing fraud, enhancing transparency, fostering greater trust among employees and the public, and improving operational efficiency [27]. However, several challenges emerge with the growing adoption of blockchain technology. These challenges encompass technical issues, including concerns about interoperability among various blockchain systems and difficulties related to scalability for larger and more intricate supply chain networks [28]. Taking the case of TradeLens, it explores both the potential and the limitations of blockchain adoption. Despite its technological maturity, TradeLens encountered significant challenges, including stakeholder resistance, expensive integration, lack of standardization, and governance issues—ultimately resulting in its closure [29]. It introduces a layered structure for incorporating blockchain into air freight shipping and logistics to increase its effectiveness and resistance to security attacks [30]. In contrast, blockchain is a decentralized mechanism that operates on a peer-to-peer network [21,30]. It is composed of chain blocks, each of which records a series of transactions while keeping track of previous blocks and their hashing. Convergence brings many benefits to transportation and logistics, including smart pay, improved fault tolerance, real-time knowledge sharing, privacy, and protection. It can be expanded in the future by developing a system-level model to assess the scalability and interoperability of various platforms used in smart transport and logistics using more real-world case studies [30].

1.2. Clustering Algorithms and Transaction Discrepancies

Clustering is one of the most common techniques to detect inconsistencies in various domains. Researchers use the different clustering algorithms to detect hidden facts about the business. This section will discuss a few clustering research papers that are related to detecting discrepancies in different businesses. The frequency and complexity of fraudulent operations are rising in today's increasingly digital financial scene, posing serious risks and costs for customers as well as financial institutions. This research suggests a machine learning-based K-means clustering technique to improve the precision and effectiveness of financial fraud detection in order to successfully address this problem. Researchers can quickly detect possible fraud by identifying unusual patterns and behaviors by clustering large amounts of financial transaction data. Machine learning-based systems offer greater flexibility and precision in detection while better adapting to ever-evolving fraud schemes and patterns as compared to traditional rule-based detection methods. In addition, K-means clustering facilitates targeted monitoring and prevention activities in high-risk areas, which effectively lessens the impact of fraud on the financial system as a whole. This helps financial institutions allocate resources more efficiently. As a result, the machine learning-based K-means clustering method, which aims to create a more dependable and safe transaction environment for the finance industry, has a lot of potential applications in the field of financial fraud detection [31]. The three main steps form the foundation of a novel clustering technique. Using a heuristic mechanism, it retrieves data from the database and creates thorough logical partitions in the main memory. All of the main memory's logical partitions are transformed into a text file. Additionally, it uses a Python program to display uniform data in three dimensions across all clusters [32]. It provides a few advantages as compared to k-means and orthogonal partitioning clustering techniques. First of all, it is a methodical way to handle inconsistent data and work with mixed data types and outliers. Secondly, it takes one-time database scanning when it splits data using attributes. Finally, it visualizes data in a three-dimensional format [32].

1.3. Air Freight Supply & Chain Problem and Description of the Methods

Many parties are involved in air freight throughout the whole supply chain, including in billing, costing, and other processes that require a lot of manual labor and are prone to errors [33]. It reveals significant structural changes and vulnerabilities in the global ICT supply chain through a thorough analysis of trade dependencies and patterns of technological specialization [34]. The air freight system has a couple of subsystems, such as sales and shipment. Due to the nature of the work, these modules (sales and shipment) work independently. Furthermore, the freight forwarders provide sales services to the airline industry through legal agreements, or customers directly visit the airline cargo office and make the cargo manifest, and the airlines arrange the cargo shipment services depending on the load. It is too difficult to manage the consistency between modules (shipment and sales). The sales process makes the manifest for booking, but the shipment process provides delivery services with the help of different airlines and routes. E.g., once the cargo sales department or freight forwarder makes a manifest concerning attributes ("airwaybill", "origin", "destination", "currency_code", "gross_amount", and "weight"), the cargo shipment department will provide the delivery services using different routes, but the airline industry must weigh cargo before lifting it into the aircraft or it will weigh the whole aircraft. Here is the research question: How do you detect discrepancies between the sales and shipping systems? It creates many disruptions in the freight industry, such as yield disruption, weight disruption, route disruption, freight forwarder disruption, cargo space disruption, and demand and supply disruption. The novel batch and blockchain methods can resolve the above-mentioned problems. A novel batch method detects transaction discrepancies based on the difference in values of attributes ("gross_amount" and "weight") if the attribute ("airwaybill") and its value match. Similarly, a novel blockchain detects transaction discrepancies based on the variation between records. Hence, it integrates two independent systems into a central repository, which diminishes all the above-mentioned disruptions.

2. Proposed Methods

There are two ways to detect discrepancies in the air freight business, such as novel batch mode and blockchain methods. This section will track discrepancies between shipment and sales domains. The proposed system introduces abstract and mathematical models along with algorithms. The proposed methods provide an outcome to clustering techniques to visualize the legitimacy and discrepancies in a graphical format. Acknowledge that detected discrepancies could be due to various risk factors, including both errors and potential fraud. Emphasize that the methods provide a starting point for investigation rather than definitive proof of fraudulent activity. In other words, all accounts are entered in the general ledger at the end of the financial month/year, so it is a very significant step to resolve revenue and weight conflicts between independent systems. Otherwise, it presents asset misappropriation in a financial report. In addition to this, it is essential to introduce a transport policy that will establish synchronization of revenue and weight between freight forwarders and airlines.

2.1. Novel Batch Mode Method

This section will detect transaction discrepancies using a novel batch-mode method. Furthermore, it proposes abstract and mathematical models along with algorithms and theoretical time complexity. In other words, the freight discrepancies issue creates abstract and mathematical models, and a novel batch mode algorithm is developed using a mathematical model approach along with time complexity. Finally, **Appendix A**, **Table A1** shows the specific attributes associated with the air freight domain.

2.1.1. Abstract Model and Novel Batch Mode Method

Figure 1 shows the abstract model for detecting discrepancies in the air freight business, where the sales and shipment systems work independently. Furthermore, the sales system makes a manifest with the concerned attributes ("airwaybill", "origin", "destination", "weight", and "gross_amount"). However, because of the cargo space issue and the flight diversion problem, the shipment system no longer needs to follow up with the sales system. Regarding transform-1, it extracts the sales data and stores it in the air freight central system. However, the transform-2 extracts the data from the shipment system and loads it into the air freight central system. In other words, the transform-3 process executes the batch mode algorithm, detects discrepancies between sales and shipment datasets if matched, and inserts records into the table (”match”). Otherwise, it will insert discrimination records into the table ("unmatched"). It indicates that there is no match between the datasets, yet various records are present in both. Additionally, when any update query is executed on the tables ("match" and "unmatched"), audit logs documenting transaction discrepancies will be created through a database trigger. Moreover, all inconsistencies are visualized in bar charts using Oracle Miner. Hence, if the attribute ("airwaybill") validates the records between sales and shipment datasets, it will detect discrepancies based on the attributes ("gross_amount" and "weight").

2.1.2. Mathematical Model and Novel Batch Mode

Figure 2 presents the mathematical model for detecting discrepancies between two supersets using a novel batch mode method. The mathematical model derives from the abstract model, as mentioned in **Figure 1.** There are two supersets (A and B), where A shows the sales system as mentioned in the transform-1 step, and B presents the shipment system as mentioned in the transform-2 step. All elements in a superset (A) must exist in a superset (B), and vice versa. If the conditions are not met, discrepancies will be detected." In other words, if superset (A) is a minus superset (B), as mentioned in the transform-3 step, it would present zero items as a result. Otherwise, it would show discrepancies in a comma-delimited file. For example, the light portion of the superset (A) presents the transform-1 step, and the yellow-orange part of the superset (B) shows the transform-2 step in the central repository. Furthermore, the difference between two supersets (A and B) presents inconsistencies in the air freight industry, as mentioned in the transform-3 step. Finally, it suggested that both supersets (A and B) acquire the same set of items to reduce discrepancies. Hence, it detects inconsistent records between two supersets (A and B).

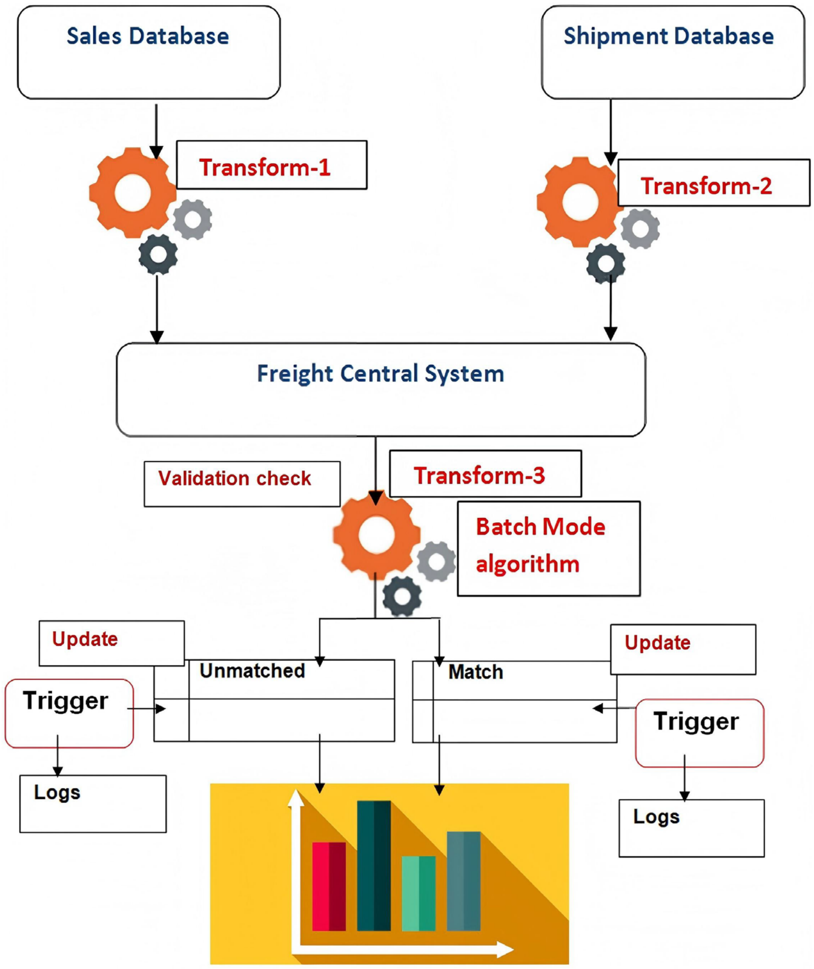
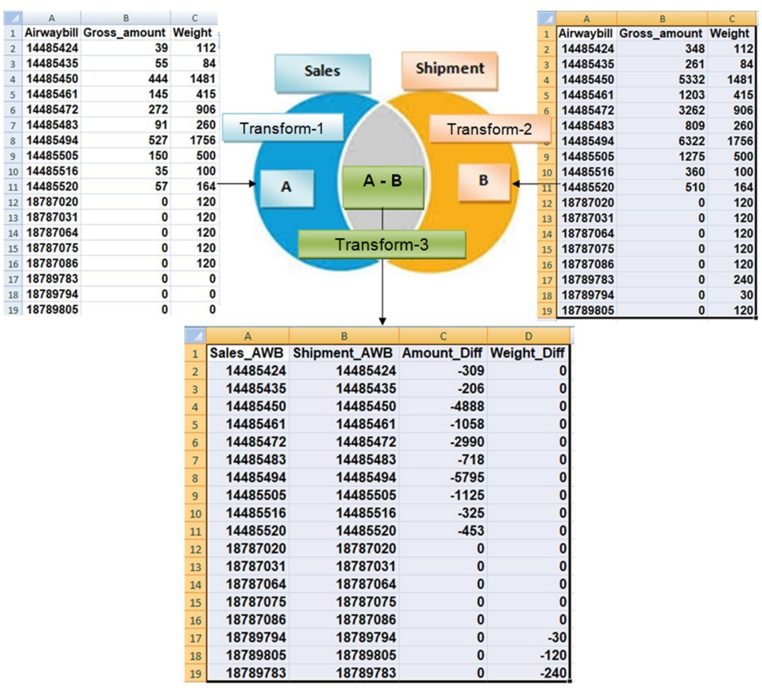


Figure 1. Abstract model and novel batch mode method.



Airwaybill	Gross_amount	Weight
14485424	39	112
14485435	55	84
14485450	444	1481
14485461	145	415
14485472	272	906
14485483	91	260
14485494	527	1756
14485505	150	500
14485516	35	100
14485520	57	164
18787020	0	120
18787031	0	120
18787064	0	120
18787075	0	120
18787086	0	120
18789783	0	0
18789794	0	0
18789805	0	0

Airwaybill	Gross_amount	Weight
14485424	348	112
14485435	261	84
14485450	5332	1481
14485461	1203	415
14485472	3262	906
14485483	809	260
14485494	6322	1756
14485505	1275	500
14485516	360	100
14485520	510	164
18787020	0	120
18787031	0	120
18787064	0	120
18787075	0	120
18787086	0	120
18789783	0	240
18789794	0	30
18789805	0	120

Sales_AWB	Shipment_AWB	Amount_Diff	Weight_Diff
14485424	14485424	-309	0
14485435	14485435	-206	0
14485450	14485450	-4888	0
14485461	14485461	-1058	0
14485472	14485472	-2990	0
14485483	14485483	-718	0
14485494	14485494	-5795	0
14485505	14485505	-1125	0
14485516	14485516	-325	0
14485520	14485520	-453	0
18787020	18787020	0	0
18787031	18787031	0	0
18787064	18787064	0	0
18787075	18787075	0	0
18787086	18787086	0	0
18789794	18789794	0	-30
18789805	18789805	0	-120
18789783	18789783	0	-240

Figure 2. Mathematical model and novel batch mode method.

2.1.3. Novel Batch Mode Algorithm

Algorithm 1 detects discrepancies between sales and shipment systems, whether datasets are matched or not. The proposed **Algorithm 1** is developed based on the abstract and mathematical models as mentioned in **Figures 1** and **2**. It takes a few steps. Firstly, if sales and shipment systems held inconsistent information regarding the attribute ("airwaybill") and its match values, it will compute revenue and weight discrepancies. As a result, it stores the data into a recordset (rs_1). However, if sales and shipment systems hold inconsistent information regarding the attribute ("airwaybills") and its mismatched values, it will compute revenue and weight discrepancies. Consequently, it stores the data into a recordset (rs_2). To start the iteration process, it fetches and inserts all features and transaction discrepancies into the table ("match"). Similarly, it fetches and inserts all unmatched airwaybill details into the table ("unmatched"). Moreover, it commits the changes and closes the loop once each iteration process is complete. Finally, it identifies revenue and weight discrepancies, regardless of whether the datasets are matched. Thus, this approach is one of the most sophisticated ways to enhance freight profit by detecting discrepancies between two independent systems.

Algorithm 1. Novel batch mode algorithm.

```
procedure Novel batch mode algorithm
rs1 ←extract all features and discrepancies from database { if sales.awb=shipment.awb}
rs2 ←extract all features and discrepancies from database { if sales.awb= shipment.awb}
begin
for r in rs1 loop
insert into match(all features and discrepancies )
values (r.feature1,r.feature2,r.rev_discrepancy,r.wgt_discrepancy);
commit
end loop
for rr in rs2 loop
insert into unmatched(all features and discrepancies )
values (rr.feature1,rr.feature2,rr.rev_discrepancy,rr.wgt_discrepancy);
commit
end loop
end procedure
```

2.1.4. Novel Batch Mode Algorithm and Theoretical Time Complexity

Regarding theoretical time complexity, it uses the first loop to identify the transaction discrepancies between sales and shipment and stores the outcome into the table ("match"). Similarly, the second loop fetches and inserts the transaction discrepancies into the table ("unmatched"). So, the time complexity of this algorithm is $T(n) = \theta(2n)$. Hence, it is also called $T(n) = \theta(n)$.

2.2. Novel Batch Mode Update Algorithm

Algorithm 2 updates the attributes ("amount_boolean" and "weight_boolean") and their values. It performs a few steps. Firstly, it creates a record set (rs_1) from the table ("match"). Secondly, it defines a record set (rs_2) from the table ("unmatched"). To start the iteration process, it fetches the records from the record set (rs_1). Additionally, if attributes ("rev_discrepancy" or "wgt_discrepancy") held the zero value, it will update the features ("amount_boolean" and "weight_boolean") and their values to zero. Otherwise, it updates the features ("amount_boolean" and "weight_boolean") and their values with one. To start the second iteration process, it fetches and updates the features ("amount_boolean" and "weight_boolean") and their values with one because it is related to unmatched records. Moreover, it performs the commit and closes the loop after completing each iteration process. Finally, it provides input to clustering algorithms for developing data clusters to present legitimacy and discrepancies in a two- or three-dimensional format. So, the proposed **Algorithm 2** utilizes one loop and an if-else statement, resulting in a total of (n + 1) executions. At each iteration of the loop, it performs an operation with a time complexity of O(1). Additionally, if the if-else statement returns a numeric value, it also has a time complexity of O(1). Therefore, the overall time complexity of this algorithm is represented as $T(n) = \theta(n)$. Hence, it provides scalability and interoperability with clustering algorithms.

Algorithm 2. Novel batch mode update algorithm.

```
procedure Novel batch mode update algorithm
rs1 ←extractall features and discrepancies from match
rs2 ←extractall features and discrepancies from unmatched
r in rs1 loop
if { r.rev_discrepancy=0 } then
update match
set r.amount_boolean = 0
else
update match
set r.amount_boolean = 1
end if
if { r.wgt_discrepancy=0 }
update match
set r.weight_boolean = 0
else
update match
set r.weight_boolean = 1
end if
end loop
rr in rs2 loop
if { rr.rev_discrepancy= 0 } then
update unmatched
set rr.amount_boolean = 1
end if
if { rr.wgt_discrepancy= 0 }
update unmatched
set rr.weight_boolean = 1
end if
end loop
commit
end procedure
```

2.3. Novel Blockchain Method

This section will detect discrepancies using a novel blockchain technology. Furthermore, it introduces abstract and mathematical models along with algorithms and theoretical time complexity. In other words, the freight discrepancies problem generates abstract and mathematical models; a novel blockchain algorithm is then developed based on these models, incorporating considerations of time complexity.

2.3.1. Abstract Model and Novel Blockchain Algorithm

Figure 3 shows the details of an abstract model to detect discrepancies using a novel blockchain algorithm. The sales and shipment systems manifest the cargo according to the attributes ("airwaybill", "origin", "destination", "gross_amount", "revenue", and "weight"). However, the shipment system provides delivery services with the help of different airlines around the globe. Once the airline cargo sales department or freight forwarders complete the booking process, shipping companies will provide delivery services across the globe after manifesting the lifted cargo. The sales and shipping systems provide air cargo delivery services independently, so there is a need to integrate both systems into a central repository and analyze the managerial insights of the freight business along with air cargo disruptions. One of the most advanced methods for detecting discrepancies between independent systems is a blockchain algorithm. In this process, the blockchain system receives data from the shipment and sales systems through CSV files. It identifies discrepancies whenever there are mismatches between the two systems. The blockchain assigns unique hash keys when tracking discrepancies between sales and shipment modules. In other words, if there is any variation between sales and shipment systems, it would detect discrepancies and record them in the output file along with a new hash key and airwaybill details. Otherwise, it would not note records in a comma-delimited file. Finally, it is suggested that a novel blockchain algorithm can diminish all air cargo demand and supply disruptions. Hence, this strategy detects discrepancies with unique hash keys.

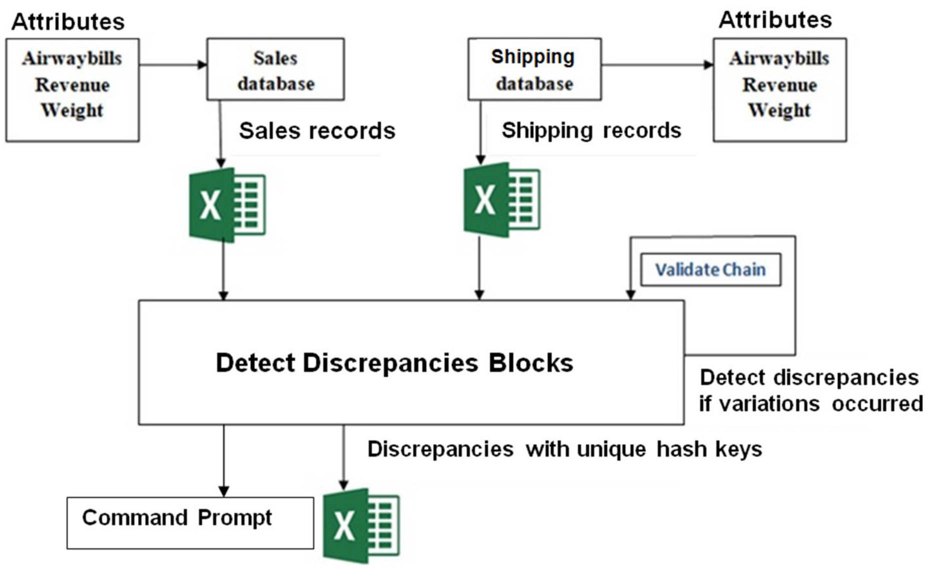
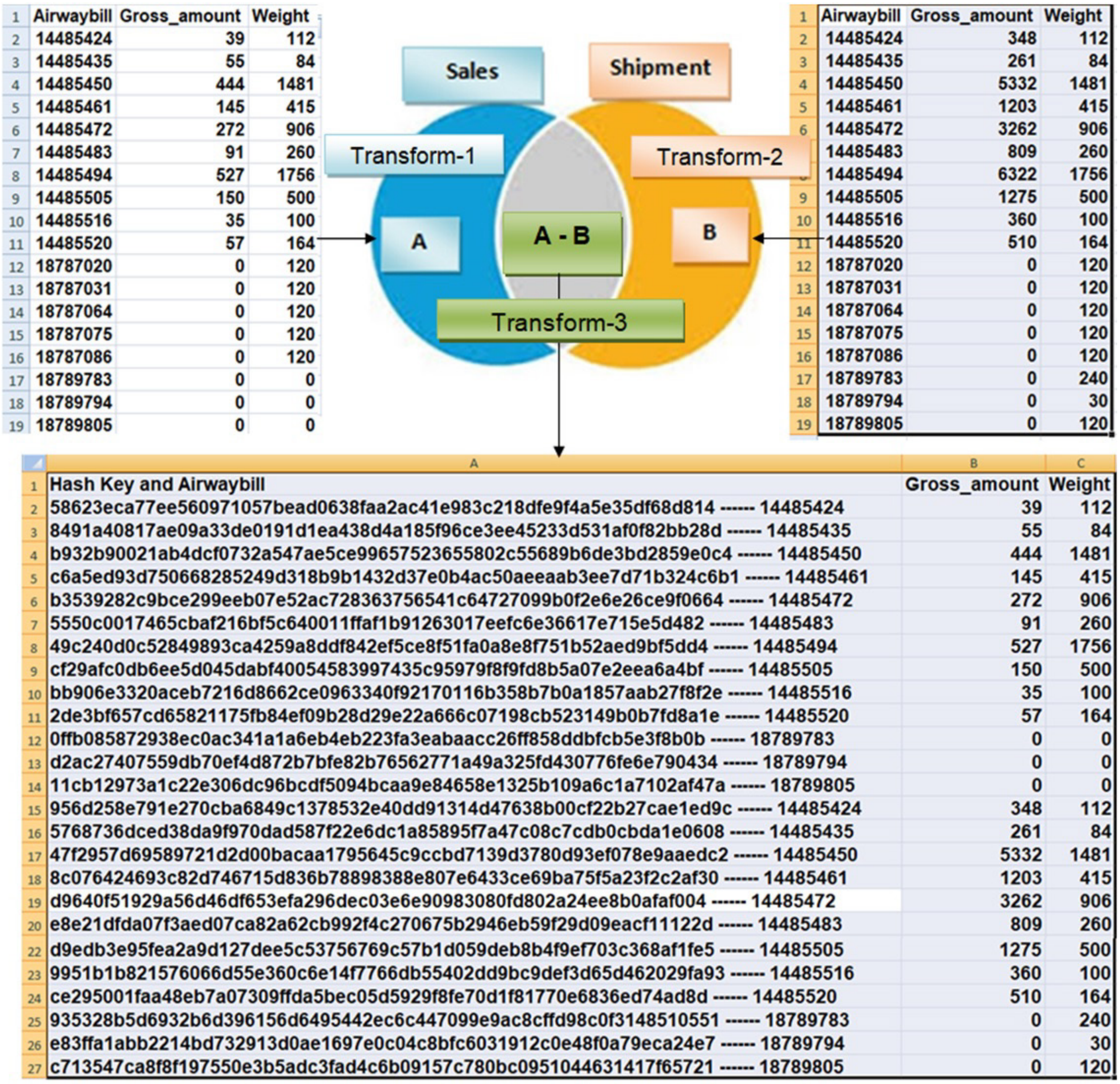


Figure 3. Abstract model and novel blockchain algorithm.

2.3.2. Mathematical Model and Novel Blockchain Algorithm

Figure 4 shows the mathematical model for detecting discrepancies between two supersets using the blockchain method. The mathematical model derives from the abstract model, as mentioned in **Figure 3**. There are two supersets (A and B), where A shows the sales system as mentioned in the transform-1 step, and B presents the shipment system as mentioned in the transform-2 step. All elements in a superset (A) must exist in a superset (B), and vice versa. Otherwise, it notes discrepancies in a comma-separated file. In other words, if superset (A) is minus superset (B), as mentioned in the transform-3 step, it would require displaying zero items with no hash key value as a result. Otherwise, it would note discrepancies in a comma-delimited file or display records on the command prompt with a new hash key. For example, the light blue portion of the superset (A) presents the transform-1 step, and the yellow-orange part of the superset (B) shows the transform-2 step in the central repository. Furthermore, the difference between the two supersets (A and B) presents transaction discrepancies in the air freight industry, as mentioned in the transform-3 step. Finally, it suggested that both supersets (A and B) acquire the same set of items to enhance freight profit. Hence, it detects discrepancies between two independent systems.



	Airwaybill	Gross_amount	Weight
2	14485424	39	112
3	14485435	55	84
4	14485450	444	1481
5	14485461	145	415
6	14485472	272	906
7	14485483	91	260
8	14485494	527	1756
9	14485505	150	500
10	14485516	35	100
11	14485520	57	164
12	18787020	0	120
13	18787031	0	120
14	18787064	0	120
15	18787075	0	120
16	18787086	0	120
17	18789783	0	0
18	18789794	0	0
19	18789805	0	0

	Airwaybill	Gross_amount	Weight
2	14485424	348	112
3	14485435	261	84
4	14485450	5332	1481
5	14485461	1203	415
6	14485472	3262	906
	14485483	809	260
	14485494	6322	1756
9	14485505	1275	500
10	14485516	360	100
11	14485520	510	164
12	18787020	0	120
13	18787031	0	120
14	18787064	0	120
15	18787075	0	120
16	18787086	0	120
17	18789783	0	240
18	18789794	0	30
19	18789805	0	120

	Hash Key and Airwaybill	Gross_amount	Weight
2	58623eca77ee560971057bead0638faa2ac41e983c218dfe9f4a5e35df68d814 ------ 14485424	39	112
3	8491a40817ae09a33de0191d1ea438d4a185f96ce3ee45233d531af0f82bb28d ------ 14485435	55	84
4	b932b90021ab4dcf0732a547ae5ce99657523655802c55689b6de3bd2859e0c4 ------ 14485450	444	1481
5	c6a5ed93d750668285249d318b9b1432d37e0b4ac50aeeaab3ee7d71b324c6b1 ------ 14485461	145	415
6	b3539282c9bce299eeb07e52ac728363756541c64727099b0f2e6e26ce9f0664 ------ 14485472	272	906
7	5550c0017465cbaf216bf5c640011ffaf1b91263017eefc6e36617e715e5d482 ------ 14485483	91	260
8	49c240d0c52849893ca4259a8ddf842ef5ce8f51fa0a8e8f751b52aed9bf5dd4 ------ 14485494	527	1756
9	cf29afc0db6ee5d045dabf40054583997435c95979f8f9fd8b5a07e2eea6a4bf ------ 14485505	150	500
10	bb906e3320aceb7216d8662ce0963340f92170116b358b7b0a1857aab27f8f2e ------ 14485516	35	100
11	2de3bf657cd65821175fb84ef09b28d29e22a666c07198cb523149b0b7fd8a1e ------ 14485520	57	164
12	0ffb085872938ec0ac341a1a6eb4eb223fa3eabaacc26ff858ddbfcb5e3f8b0b ------ 18789783	0	0
13	d2ac27407559db70ef4d872b7bfe82b76562771a49a325fd430776fe6e790434 ------ 18789794	0	0
14	11cb12973a1c22e306dc96bcdf5094bcaa9e84658e1325b109a6c1a7102af47a ------ 18789805	0	0
15	956d258e791e270cba6849c1378532e40dd91314d47638b00cf22b27cae1ed9c ------ 14485424	348	112
16	5768736dced38da9f970dad587f22e6dc1a85895f7a47c08c7cdb0cbda1e0608 ------ 14485435	261	84
17	47f2957d69589721d2d00bacaa1795645c9ccbd7139d3780d93ef078e9aaedc2 ------ 14485450	5332	1481
18	8c076424693c82d746715d836b78898388e807e6433ce69ba75f5a23f2c2af30 ------ 14485461	1203	415
19	d9640f51929a56d46df653efa296dec03e6e90983080fd802a24ee8b0afaf004 ------ 14485472	3262	906
20	e8e21dfda07f3aed07ca82a62cb992f4c270675b2946eb59f29d09eacf11122d ------ 14485483	809	260
22	d9edb3e95fea2a9d127dee5c53756769c57b1d059deb8b4f9ef703c368af1fe5 ------ 14485505	1275	500
23	9951b1b821576066d55e360c6e14f7766db55402dd9bc9def3d65d462029fa93 ------ 14485516	360	100
24	ce295001faa48eb7a07309ffda5bec05d5929f8fe70d1f81770e6836ed74ad8d ------ 14485520	510	164
25	935328b5d6932b6d396156d6495442ec6c447099e9ac8cffd98c0f3148510551 ------ 18789783	0	240
26	e83ffa1abb2214bd732913d0ae1697e0c04c8bfc6031912c0e48f0a79eca24e7 ------ 18789794	0	30
27	c713547ca8f8f197550e3b5adc3fad4c6b09157c780bc0951044631417f65721 ------ 18789805	0	120

Figure 4. Mathematical model and novel blockchain algorithm.

2.3.3. Novel Blockchain Algorithm

Algorithm 3 detects discrepancies in the air freight system using the novel blockchain method. The proposed **Algorithm 3** is developed based on the abstract and mathematical models as mentioned in **Figures 3** and **4**. It takes a few steps to detect discrepancies between sales and shipment datasets. Firstly, it imports the hash library to generate hash keys for the data block after initializing the instance method. Secondly, init is an instance method that initializes when creating an object. Furthermore, it assigns the data to the variable ("block_data") after subtracting the variable ("transactions") and its values from the attribute ("previous_hash_block"). Similarly, it creates the hash keys for the feature ("block_data") and stores them in the variable ("block_hash"). Moreover, it makes all comma-delimited files in a list format and reads the sales dataset file as $file_1$ and the shipment dataset file as $file_2$. Besides this, it opens one comma-delimited file for writing as an output file. During the loop execution process, it fetches records from comma-delimited files, such as $file_1$. If the statement validates the records in $file_2$ and returns true, it would call the instance method (init) along with functions (block_hash and block_data). In other words, if the condition returns true, it will create a class object, which will automatically call the init method and assign the hash keys to the data block. Otherwise, it will not note records in a comma-separated file. Finally, it displays all transaction discrepancies on a Python prompt and notes all discrimination information in a comma-delimited file using the Python write function. Hence, it assigns unique hash keys to both datasets if they are mismatched.

Algorithm 3. Novel blockchain algorithm.

```
import hashlib
class detect_discrepancies_blocks:
def init (previous_block_hash, transactions)
block_data ← { transactions }- { previous_block_hash }
block_hash ← block_data creates hash keys
file1 ← ['Shipment.csv']
file2 ← ['Sales.csv']
fname← open(output.csv,w)
for row in file1
if { not row in file2 }
block1← detect_discrepancies_blocks(firstblock,[row,row])
block2← detect_discrepancies_blocks(block1.block_hash,[row, row])
fname← write(block2.block_hash +− −−+row)
print("Block1data : block1.block_data")
print("Block1hash : block1.block_hash")
end if
file1, file2 ← file2, file1
end loop
```

2.3.4. Novel Blockchain Algorithm and Theoretical Time Complexity

The theoretical time complexity involves only one loop and a single if statement. The loop will run (n + 1) times, performing the O (1) operation at each iteration of the loop, and if statement will return the numeric value O (1). Therefore, the time complexity of this algorithm is expressed as T (n) = θ (n).

2.4. Proposed Methods and Challenges

There are two major independent systems in the air freight industry, such as sales and shipment. Due to the nature of the work, the air freight industry operates the business with the help of freight forwarders and other airlines, but unfortunately, there is no system available to maintain checks and balance between airlines and freight forwarders. Furthermore, if these systems independently operated in the air cargo industry, it would cause many disruptions, such as yield disruption, weight disruption, route disruption, freight forwarder disruption, cargo space disruption, and demand and supply disruption. The research methodologies can be used to overcome freight disruptions. Regarding blockchain technology, it is very easy to implement the proposed blockchain algorithm in the air freight industry because it detects document discrepancies between datasets. Every month, the air freight industry can collect the sales dataset from freight forwarders based on business agreements as well as the shipment dataset from the shipping department. It is required to define the data process layout between freight forwarders and the shipping department. Consequently, if there is any variation between datasets, the blockchain algorithm assigns new hash keys as it detects document discrepancies in the air cargo industry. There is no problem implementing the blockchain algorithm because it detects document discrepancies based on prior chains. In other words, if there is a lack of information from freight forwarders and shipping departments, it may be a big challenge for the air freight industry. The novel batch mode method is one of the decentralized approaches to detecting document discrepancies between datasets, like blockchain. First of all, it integrates sales and shipment datasets into a central repository to detect document discrepancies. If freight forwarders or shipping departments do not provide the complete details of datasets, the air freight industry will face a big barrier to overcoming the above-mentioned disruptions. Finally, the air cargo industry defines the data collection process layout between freight forwarders and shipping departments and introduces business transparency via proposed methods. Hence, it is one of the best ways to detect document discrepancies.

2.5. Proposed Methods and Practical Implications

Figure 5 shows the data pipeline process between freight forwarders and the air cargo industry. Stakeholders can provide services at various airports and enhance the freight business for airlines. The proposed methods can resolve document discrepancies between freight forwarders and the air freight industry. Therefore, the air freight industry needs to introduce a data pipeline between freight forwarders and the shipment department after completing the sales period. In the data pipeline process, all freight forwarders send sales datasets to the air freight industry or data repository via one of the sophisticated channels (FTP, email, or database link). Similarly, the shipment department sends the shipping datasets to the air freight industry or data repository through one of the sophisticated channels (FTP, email, or database link). Moreover, the proposed algorithms detect document discrepancies between sales and shipment systems after integrating the datasets. In other words, the air freight industry can share the results of imbalanced transactions with freight forwarders to resolve this issue on a priority basis.

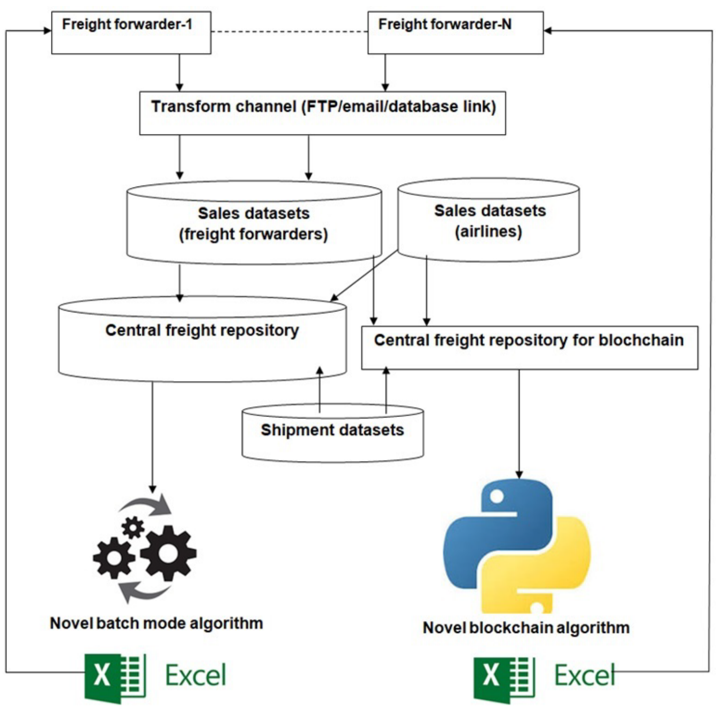


Figure 5. Proposed methods and practical implications.

3. Results and Discussion

The system configuration regarding hardware is Windows 10, with 8GB of RAM and a 3.19 GHz Core i5 processor. Furthermore, the proposed system software configuration includes Oracle 19C, Graph, PL/SQL, SQL Developer, Python, and Oracle Miner. Moreover, this section will explain the results of discrepancies using the novel batch mode algorithm and blockchain technology. It uses two freight datasets: (i) freight dataset 1 contains 169,859 records; (ii) freight dataset 2 contains 254,162. With regard to the batch mode method, it presents the discrepancies between datasets (shipment and sales), whether the datasets match or not. If any variation exists between attributes and their values, a novel blockchain algorithm will detect and note discrepancies in a comma-separated file. Finally, it discusses the benefits and threats of the evaluation.

3.1. Freight Dataset 1

It contains two subsets, which are known as sales and shipment datasets. It is also called matching datasets, but has different revenue and weight values between sales and shipment datasets. The shipment dataset contains 169,859 rows. It shows a list of attributes ("airwaybill_no", "origin", "destination", "gross_amount", "origin_currency", "weight", "bill_airline", "agreement_type", "shipping_type", "commodity_code", "incidental_charges", and "commission"). Similarly, the sales dataset contains 169,859 records. It shows a list of attributes ("airwaybill_no", "origin", "destination", "gross_amount", "weight", "agent_code", "commodity_code", and "commission").

3.2. Freight Dataset 2

It contains two subsets. It is also called matched and unmatched datasets between sales and shipment systems. The shipment dataset contains 254,162 rows. It shows a list of attributes ("airwaybill_no", "origin", "destination", "gross_amount", "origin_currency", "weight", "bill_airline", "agreement_type", "shipping_type", "commodity_code", "incidental_charges", and "commission"). Similarly, the sales dataset contains 254,162 records. It shows a list of attributes ("airwaybill_no", "origin", "destination", "gross_amount", "weight", "agent_code", "commodity_code", and "commission").

3.3. Novel Batch Mode Algorithms and Independent Systems

Figure 6 shows the execution steps of a novel batch mode algorithm. There are a few steps involved in this process to detect discrepancies and present them in a graphic format. First of all, two independent systems (sales and shipment) load the data into the central freight repository using the Oracle external table method, as mentioned in the transform-1 and transform-2 steps. Furthermore, **Algorithm 1** extracts shipment and sales data from the central freight repository, fetches records, and detects revenue and weight discrepancies based on the attribute ("airwaybill") and its value as mentioned in the transform-3 step. In other words, **Algorithm 2** updates the tables ("match" and "unmatched") regarding attributes ("amount_boolean" and "weight_boolean") to provide input to clustering methods. Moreover, all transaction discrepancies are inserted into the tables ("match" and "unmatched"). Finally, the SQL query extracts the data from the tables ("match" and "unmatched") and connects it with a graph to visualize the discrepancies in a bar chart using the Oracle Miner tool.

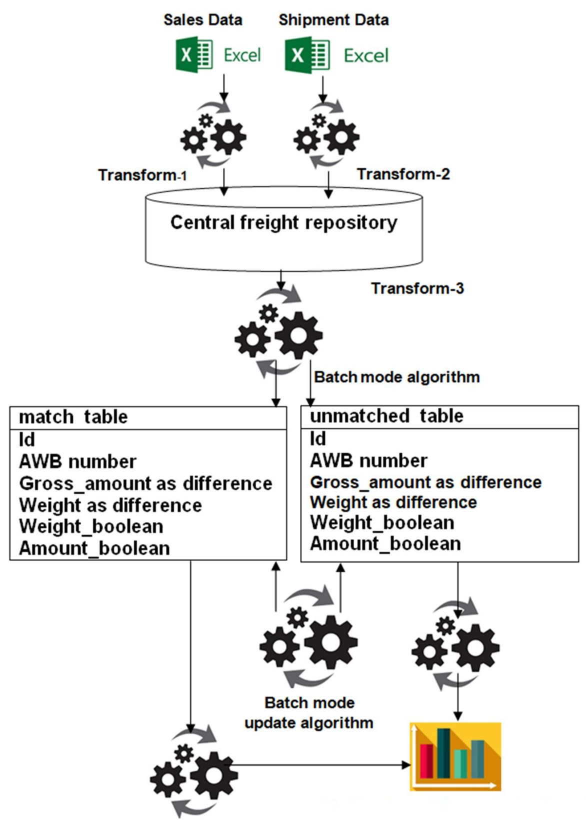


Figure 6. Novel batch mode algorithms and independent systems.

3.4. Novel Batch Mode Algorithms and Freight Dataset 1

The batch mode method is more secure and efficient than direct SQL because there is a chance of SQL injection. During the execution process, it creates the record set in main memory, where it holds the sales and shipping records. Each dataset contains 169,859 rows. Furthermore, the fetch process iterates over records between sales and shipping datasets. It validates the record based on the attribute ("airwaybill"). If the condition returns true, it will provide the revenue difference between sales and shipment domains. In this case, if the result variable holds zero values, it will provide legitimate transactions. Otherwise, it detects revenue discrepancies. Similarly, it fetches another record. Consequently, if datasets were matched between sales and shipment datasets, it would detect revenue and weight discrepancies. Concerning the weight attribute, the batch mode algorithm provides 28,714 discrepancies out of 169,859 records. Similarly, the batch mode algorithm provides 59,314 discrepancies out of 169,859 records that are revenue inconsistencies. It visualizes discrepancies in a bar graph via Oracle Miner. Firstly, it defines a pipeline to extract discrepancies from the database and provide them to the graph operator via Oracle Miner after completing the proposed **Algorithm 1** process. Secondly, the graph operator applies the average function to the features ("gross_amount_diff', "weight_diff'), where it aggregates the list of airwaybills for visualizing. Otherwise, it cannot visualize the groups regarding the feature ("airwaybill"). However, it requires developing data clusters to identify the legitimacy and discrepancies via clustering algorithms.

3.4.1. Novel Batch Mode Algorithms and Revenue Discrepancies for Freight Dataset 1

Figure 7 shows discrepancies in graphical format regarding attributes ("gross_amount" and "airwaybill") using Oracle Miner. It visualizes the discrepancies along with the list of airwaybills in a bar chart after completing the batch mode execution process. For example, if the attribute ("airwaybill") held the values (less than or equal to 52,997,581.5), it would detect revenue discrepancies between 0 and -2.0×10^4. If the attribute ("airwaybill") held the values (52,997,581.5 to 55,784,248), it would detect revenue discrepancies between 0 and -4.2×10^4. Similarly, if the attribute ("airwaybill") held the values (72,504,247 to 75,290,913.5), it would detect revenue discrepancies between 0 and -1.1×10^5. Hence, the proposed system validates the records between sales and shipping datasets based on the attribute ("airwaybill") and its values.

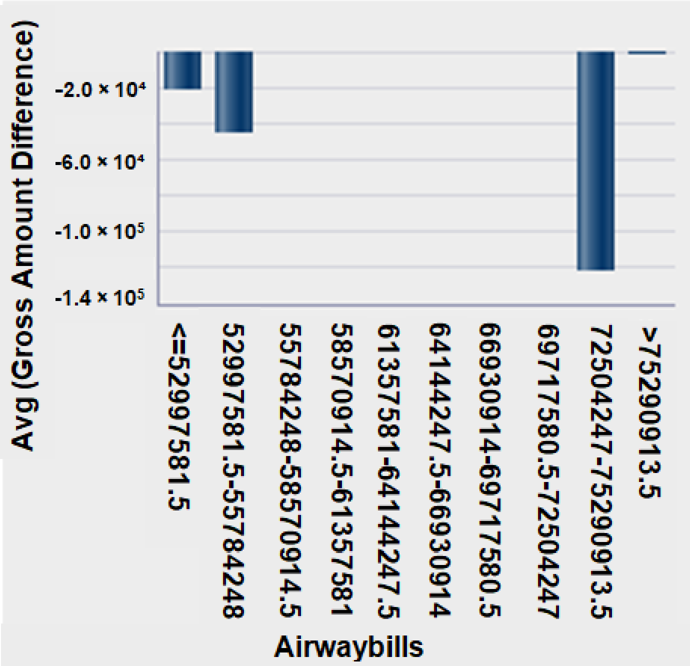


Figure 7. Novel batch mode algorithms and revenue discrepancies for freight dataset 1.

3.4.2. Novel Batch Mode Algorithms and Weight Discrepancies for Freight Dataset 1

Figure 8 shows discrepancies in the bar chart regarding attributes ("weight" and "airwaybill") after executing the proposed **Algorithm 1**. Furthermore, it visualizes a list of airwaybills along with weight discrepancies in graphical format. For example, if the shipment attribute ("airwaybill") held values between 52,997,581.5 and 55,784,248, it would detect weight discrepancies between 0 and -1.5×10^6. Similarly, if the attribute ("airwaybill") held values greater than or equal to 75,290,913.5, it would detect weight discrepancies between 0 and 1.0×10^6. Finally, positive and negative weight inconsistencies can mislead the financial report. Hence, the novel batch mode algorithm validates the bulk of records based on the attribute ("airwaybill") and its values.

3.5. Novel Batch Mode Algorithms and Freight Dataset 2

The dataset contains 254,162 records. It presents matched and unmatched records between sales and shipment systems. The novel batch mode algorithm detects revenue and weight discrepancies, whether datasets are matched or not. Consequently, the proposed **Algorithm 1** provides 105,899 out of 254,162 records as weight inconsistencies. Similarly, the batch mode algorithm provides 125,687 out of 254,162 records that are revenue inconsistencies. Once proposed **Algorithm 1** completed the execution process, it defined a pipeline to extract discrepancies from the database and provided them to the graph operator via Oracle Miner. Additionally, the graph operator applies the average function to the features ("gross_amount_diff", "weight_diff"), where it aggregates the list of airwaybills for visualizing. Otherwise, it cannot visualize the records. However, it requires developing data clusters to identify the legitimacy and discrepancies via clustering algorithms.

3.5.1. Novel Batch Mode Algorithms and Revenue Discrepancies for Freight Dataset 2

Figure 9 shows discrepancies in graphical format regarding attributes ("gross_amount" and "airwaybill") using Oracle Miner. It visualizes the list of airwaybills in a bar chart as revenue discrepancies after completing the proposed **Algorithm 1** execution process. For example, if the attribute ("airwaybill") held the values from 34,349,299 to 41,637,460, it would detect revenue discrepancies between 0 and -1.0×10^5. If the attribute ("airwaybill") held the values between 48925621 and 56,213,782, it would detect revenue discrepancies between 0 and -5.0×10^4. Similarly, if the attribute ("airwaybill") held the values from 63,501,943 to 70,790,104, it would detect revenue discrepancies between 0 and -1.5×10^4. Finally, if the attribute ("airwaybill") held values greater than 70,790,104, it would detect revenue discrepancies between 0 and -2.8×10^4. Hence, the proposed system validates the records between sales and shipping datasets based on the attribute ("airwaybill") and its values.

3.5.2. Novel Batch Mode Algorithms and Weight Discrepancies for Freight Dataset 2

Figure 10 shows weight discrepancies in the bar chart regarding attributes ("weight" and "airwaybill") after executing the proposed **Algorithm 1**. Furthermore, it visualizes a list of airwaybills in graphical format as weight discrepancies. For example, if the attribute ("airwaybill") held the values (less than or equal to 12,484,816), it would detect weight discrepancies between 0 and -1.0×10^6. If the attribute ("airwaybill") held the values from 34,349,299 to 41,637,460, it would detect weight discrepancies between 0 and -1.01×10^6. Similarly, if the attribute ("airwaybill") held values between 48,925,621 and 56,213,782 or greater than 70,790,104, it would detect weight discrepancies between 0 and -6.0×10^5. Finally, if the attribute ("airwaybill") held values (63,501,943–70,790,104), it would detect weight discrepancies between 0 and -5.0×10^5. Finally, negative weight inconsistencies can mislead the financial report. Hence, the novel batch mode algorithm validates the records based on the attribute ("airwaybill") and its values.

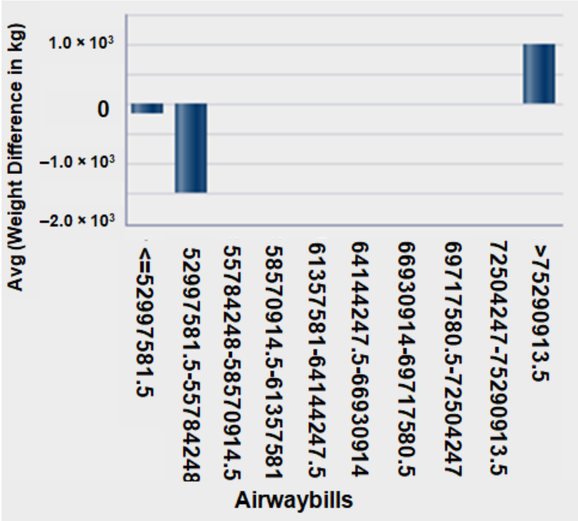


Figure 8. Novel batch mode algorithms and weight discrepancies for freight dataset 1.

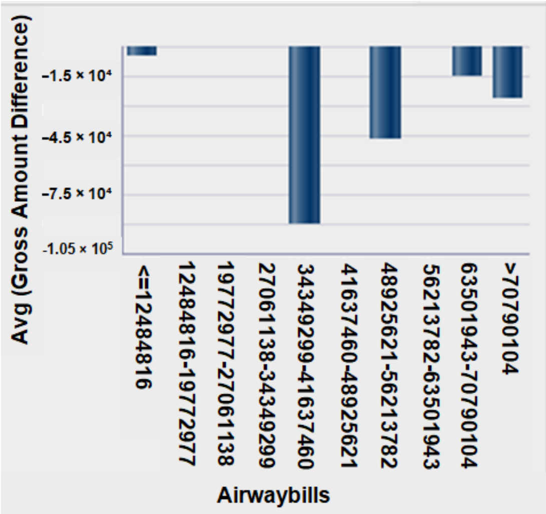


Figure 9. Novel batch mode algorithms and revenue discrepancies for freight dataset 2.

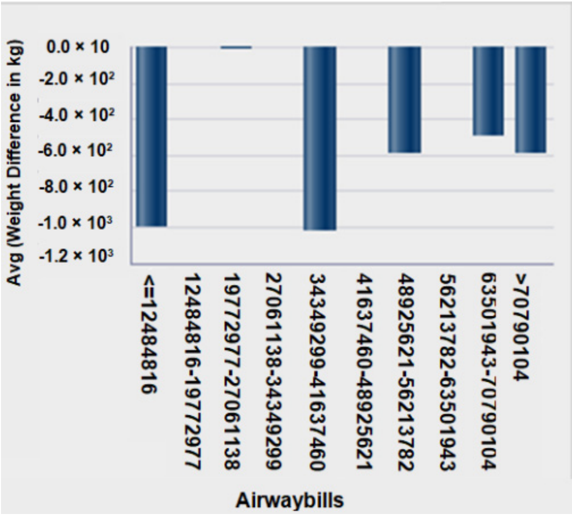


Figure 10. Novel batch mode algorithms and weight discrepancies for freight dataset 2.

3.6. Novel Batch Mode Algorithms and Real-Time Complexity

Table 1 shows the details of the real-time complexity of a novel batch mode algorithm to detect discrepancies in the air freight business. Consequently, it takes 6, 18, 40, 65, and 110 s to complete the processing of 2.0×10^4, 4.0×10^4, 6.0×10^4, 8.0×10^4, and 1.0×10^4 records, respectively. E.g., it extracts the records based on the joining condition between sales and shipment tables, so it processes 2.0×10^4 records within 6 s. Similarly, it executes the rest of the records and determines the revenue and weight discrepancies in seconds. Finally, it presents significant processing time, so it is suitable for a production environment.

3.7. Novel Blockchain Algorithm and Freight Datasets

Figure 11 shows the execution steps of the novel blockchain algorithm. First of all, the object-oriented class contains the data and function members, such as init, previous_block_hash, transactions, previous block_data, and block_hash. Furthermore, it calls the above-mentioned function members when the class creates the object after validating the if statements. Moreover, it reads the comma-delimited files, which contain sales and shipment datasets. It opens a comma-separated file ("fname") for writing output. During the fetching process, it iterates through the records of the sales and shipment datasets. If there is any variation between two datasets, it would create a class object that automatically calls function members and assigns the hash keys to the data block. The process creates transactions with unique hash keys and displays any discrepancies between them on a Python prompt. Additionally, it notes all discrepancies in a comma-delimited file. However, if there is no variation between datasets, it will not note the record in a comma-delimited file. Finally, the proposed **Algorithm 3** validates the records based on the variations that occurred between datasets.

Table 1. Novel batch mode algorithm and real-time complexity.

Number of Records	Time in Seconds
20,000	6
40,000	10
60,000	14
80,000	18
100,000	20

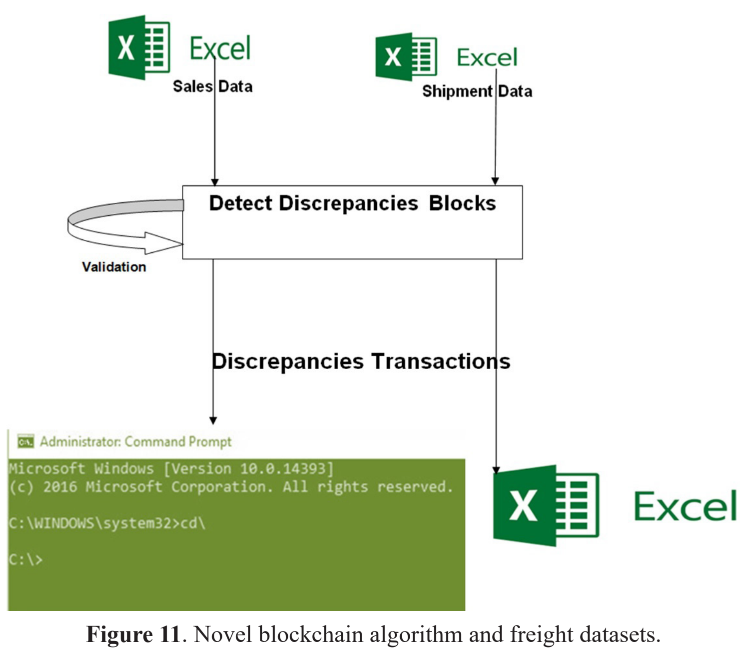


Figure 11. Novel blockchain algorithm and freight datasets.

3.7.1. Novel Blockchain Method and Revenue and Weight Discrepancies

Figure 12 shows experimental results for detecting revenue and weight discrepancies in the novel blockchain algorithm. It shows the details of block data along with hash key values. It detects the inconsistencies based on the variation in attributes ("gross_amount" and "weight") and their values. In other words, if the proposed **Algorithm 3** detected inconsistencies between datasets, it would create the transaction discrepancies with new hash keys. Otherwise, there is no new hash key assigned to records if all fields are matched.


```
In [7]: runfile('E:/main.py', wdir='E:')
Block 1 data: 14485424,39,112
 - 14485424,39,112
 - firstblock
Block 1 hash: e56c2e4a2eb3e921241c568b660ffaa99a3802b2f8e635f7c8ab872369e5e819
Block 1 data: 14485435,55,84
 - 14485435,55,84
 - firstblock
Block 1 hash: 6747e0466d44bbc2f084f409f974112f8043dc662f9cddcfd820a776e60f2559
Block 1 data: 14485450,444,1481
 - 14485450,444,1481
 - firstblock
Block 1 hash: 091addacb482c922fbfcf40d236b5b62d72e317fbff4948be411e7c9bd208384
Block 1 data: 14485461,145,415
 - 14485461,145,415
 - firstblock
Block 1 hash: 06989b8d85c3f96157cb9774c0b6d2b19dc1edac40f8c16483b846ecabee91be
Block 1 data: 14485472,272,906
 - 14485472,272,906
 - firstblock
Block 1 hash: d9080f2f1868d533b72b2264b6a457e1ad7f94124db437fbcbaaf2aefd9cb5b6
Block 1 data: 14485483,91,260
 - 14485483,91,260
 - firstblock
Block 1 hash: 1624f9ec73e51717375ba14b663fb7307a75f5b2831f8947c7e9630c4657defe
Block 1 data: 14485494,527,1756
 - 14485494,527,1756
 - firstblock
Block 1 hash: a2576ae852f516805c89b558f193c42ff3ba79f9835605a732f72ac7783851c4
Block 1 data: 14485505,150,500
 - 14485505,150,500
 - firstblock
Block 1 hash: c665019eaa61795c0644699fc4cbc92cc9b1fc8e522708541bec745b632a4815
Block 1 data: 14485516,35,100
 - 14485516,35,100
 - firstblock
Block 1 hash: 3be482712734e27ad057b9631540071d5f1c4cd05b41b5fe5930533c6ec97611
Block 1 data: 14485520,57,164
 - 14485520,57,164
 - firstblock
Block 1 hash: 45f28bcc65eeb073ece2d40ad9bd0d2fc1132f885952ed192622c6e524a4a983
Block 1 data: 14485424,348,112
 - 14485424,348,112
 - firstblock
Block 1 hash: 7721e7ee28c9a89d78318c5b61793ccdae8884e48bf504dabffd901c707630fd
Block 1 data: 14485435,261,84
 - 14485435,261,84
 - firstblock
Block 1 hash: d2195bae47c21a9996c53a7e8a8aac747b6bada22076a2a06ee60ed8934333cd
Block 1 data: 14485450,5332,1481
 - 14485450,5332,1481
 - firstblock
Block 1 hash: a1e8467a89ae38d4e53302d6209836fa1551083e592074ea1a17a33611262835
Block 1 data: 14485461,1203,415
 - 14485461,1203,415
 - firstblock
Block 1 hash: a24fb336c412769f20ccc92295a62953f2be0d16247764dc6985035102c2fe83
Block 1 data: 14485472,3262,906
 - 14485472,3262,906
 - firstblock
Block 1 hash: 588bc958f733bb7616118e8c85bab8763692d2f10d35b6a60a2c68d35f81d877
Block 1 data: 14485483,809,260
 - 14485483,809,260
 - firstblock
```

Figure 12. Novel blockchain method and revenue and weight discrepancies.

For example, the attributes (“airwaybill”, “gross_amount”, and “weight”) hold the values (14,485,424, 39, 112) and (5,160,186, 348, 112) in **Figure 12**, respectively. So the proposed **Algorithm 3** assigns the two hash keys, which start with the values “58623eca77” and “956d258e79”). Similarly, it presents the attributes (“airwaybill”, “revenue”, and “weight”) and their values (14,485,472, 272, 906) and (14,485,472, 3262, 906) in **Figure 12**, respectively. So it creates the transactions with different hash keys in block data. Finally, this strategy detects revenue and weight discrepancies with the help of different hash key values. However, if there is no variation between records, it would not assign a hash key to the transaction. For example, the attributes (“airwaybill”, “gross_amount”, and “weight”) hold the values (18,787,086, 0, 120) and (18,787,086, 0, 120) as mentioned in the transform-1 and transform-2 steps in **Figure 4**, respectively. If there is no variation between records, it cannot be noted as a transaction in a comma-separated file. Hence, each iteration validates the transaction based on the variation in the attributes (“gross_amount” and weight) and their values.

3.7.2. Novel Blockchain Algorithm and Real-Time Complexity

Table 2 shows the details of the real-time complexity of the proposed **Algorithm 3** when detecting discrepancies between datasets. Furthermore, it takes 1, 4, 9, 15, and 26 s to complete the processing of 2.0×10^4, 4.0×10^4, 6.0×10^4, 8.0×10^4, and 1.0×10^5 records, respectively. E.g., it uses two files ($file_1$ (sales_data) and $file_2$ (shipment_data)), and each file contains 10K records, so it executes an aggregated 2.0×10^4 records to determine the transaction discrepancies with respect to time in 1 s. Similarly, it executes the rest of the record sets and identifies the revenue and weight discrepancies on the Python prompt. Finally, it has the most significant time complexity, so it performs well when integrating with other modules.

3.8. Clustering Methods

Clustering techniques are used to make various groups and identify data anomalies. However, clustering methods could not identify the data anomalies between two independent systems. Consequently, the proposed methods introduce two categories (0 as legitimate and 1 as discrepancies) and visualize the records in two- and three-dimensional forms. On the other hand, classification methods (support vector machine, Naive Bayesian, and random forest) are used for predicting the business, but these techniques cannot trace the risk factors between two or more datasets. Finally, **Table A1** displays the specific attributes associated with the air freight domain.

3.8.1. Proposed Methods and K-Means Clustering with Freight Dataset 1

Figure 13 shows the comparison between legitimacy and discrepancies using the K-means algorithm. Furthermore, the K-means algorithm develops two clusters after completing the execution of the proposed methods (**Algorithm 1** and **Algorithm 2**). Moreover, it develops data clusters based on legitimacy and discrepancies in the air freight system after providing 169,859 records. The attribute (“amount_boolean”) contains two values, which are zero and 1. Consequently, it presents 65% genuine transactions and 35% inaccurate records between sales and shipment regarding the attribute (“amount_boolean”). Finally, it presents a significant difference between legitimacy and discrepancies in the air freight system. Hence, it creates freight disruptions regarding the supply and demand model.

Figure 14 shows the comparison between legitimate and discrepant transactions using the K-means algorithm. Furthermore, the K-means algorithm develops two data clusters based on the attribute (“weight_boolean”), which is held at zero or one value. Moreover, it uses 169,859 records to create data clusters between legitimate and inaccurate records. The attribute (“weight_boolean”) contains two values, which are zero and 1. As a result, it reveals that there are 83% real transactions and 17% discrepancy records between sales and shipments. Finally, if the size of datasets increases between sales and shipment systems, it will increase the inaccurate transactions. Hence, it presents weighty disruptions in the air freight industry.

Table 2. Novel blockchain algorithm and real-time complexity.

Number of Records	Time in Seconds
20,000	1
40,000	4
60,000	9
80,000	15
100,000	26

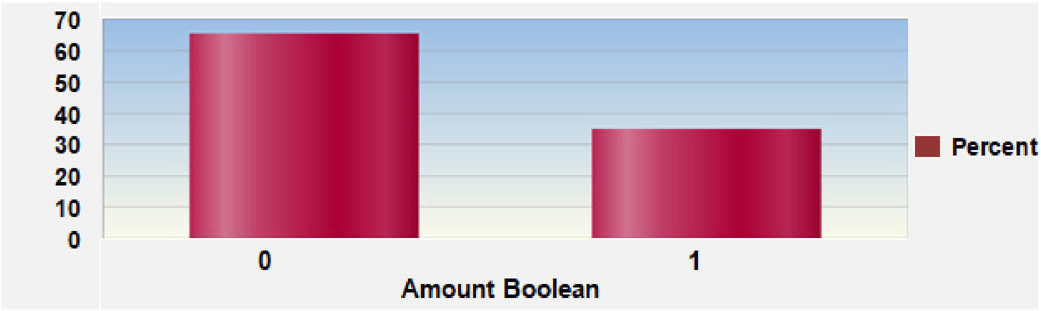


Figure 13. Revenue comparison between legitimate and discrepant clusters for freight dataset 1.

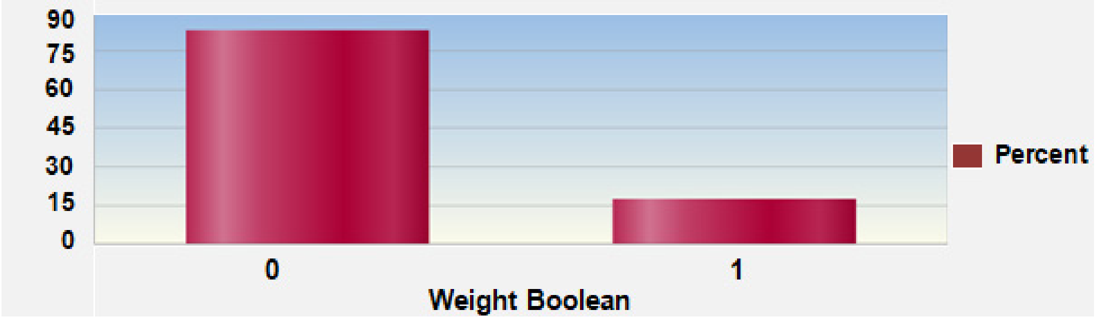


Figure 14. Weight comparison between legitimate and discrepant clusters for freight dataset 1.

3.8.2. Proposed Methods and Novel Clustering Method with Freight Dataset 1

Figure 15 shows the difference between legitimacy and discrepancies regarding attributes ("amount_boolean" and "weight_boolean") in the air freight system using a novel clustering method with consistent data in a three-dimensional format after executing the proposed algorithms (**Algorithm 1** and **Algorithm 2**). Furthermore, it uses 169,852 records to develop data clusters in a three-dimensional format using three attributes ("gross_amount", weight_boolean, and amount_boolean). Moreover, the red color shows inconsistent records, and the green color presents legitimate transactions regarding the attribute ("amount_boolean"). For example, if the attribute ("amount_boolean") held the value ("1"), it would provide the maximum value of lost revenue, which is less than 7.0 × 106. Similarly, if the attribute ("weight_boolean") held the value ("1"), it would provide the maximum value of lost weight, which is less than 5.0 × 106. Finally, it addresses a significant problem in the air freight industry. Hence, it presents the profit and loss of freight business regarding attributes ("amount_boolean" and "weight_boolean") and real differences in their values than the k-means algorithm.

3.8.3. Proposed Methods and K-Means Clustering with Freight Dataset 2

Figure 16 illustrates the comparison between legitimate and discrepant transactions in the air freight system using the K-means algorithm. Furthermore, the K-means algorithm develops two clusters. Moreover, it uses 254,162 records to develop data clusters based on the attribute ("amount_boolean"). Consequently, it presents 55% legitimate transactions and 45% discrepant records between sales and shipment regarding the attribute ("amount_boolean"). Finally, it presents a significant difference between legitimacy and discrepancies in the air cargo business. Hence, it converts profit into loss.

Figure 17 shows the comparison between real and inaccurate transactions using the K-means algorithm. Furthermore, the K-means algorithm develops two data clusters based on the attribute ("weight_boolean") after completing the execution of the proposed methods (**Algorithm 1** and **Algorithm 2**). Moreover, it uses 254,162 records to identify the legitimate and discrepancy records in a clustering format. The attribute ("weight_boolean") contains two values, which are zero for legitimacy and 1 for discrepant. The analysis reveals that 62% of the transactions are legitimate, while 38% are inaccurate records related to sales and shipment datasets. Finally, if the size of datasets increases between sales and shipment systems, it will increase the volume of discrepancies in the air freight system. For example, if we compared **Figure 14** and **Figure 17**, they would present significant weight differences, which are about 17% and 38%, respectively. Therefore, if we execute the proposed methods on a quarterly basis, it would provide a significant weight difference between sales and shipment datasets. Hence, the proposed methods provide hidden facts about the freight business.

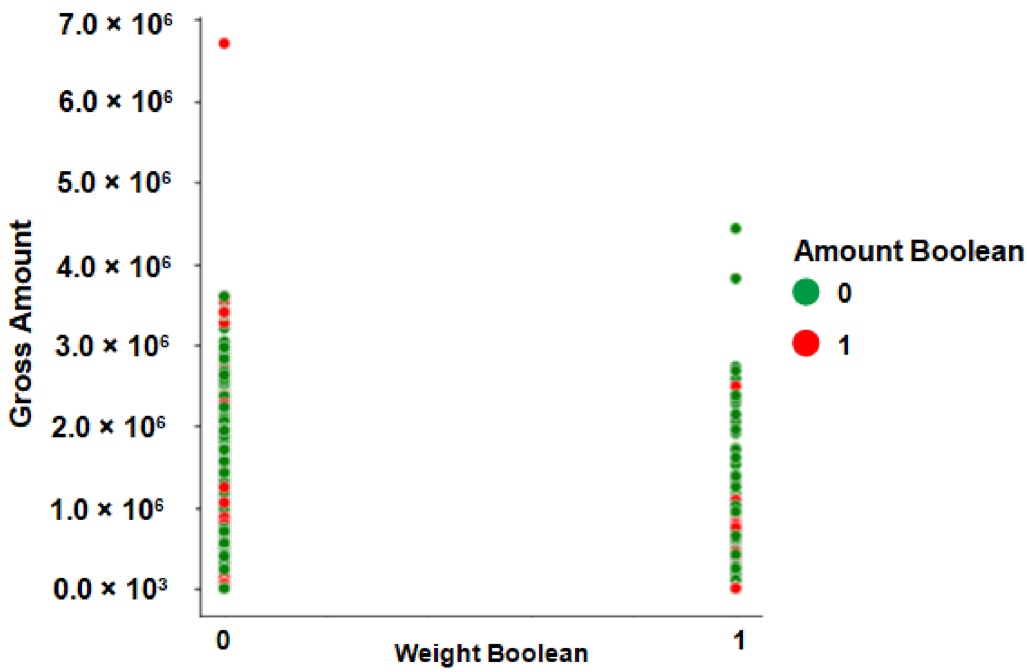


Figure 15. Data cluster between revenue and weight with real and discrepant transactions for freight dataset 1.

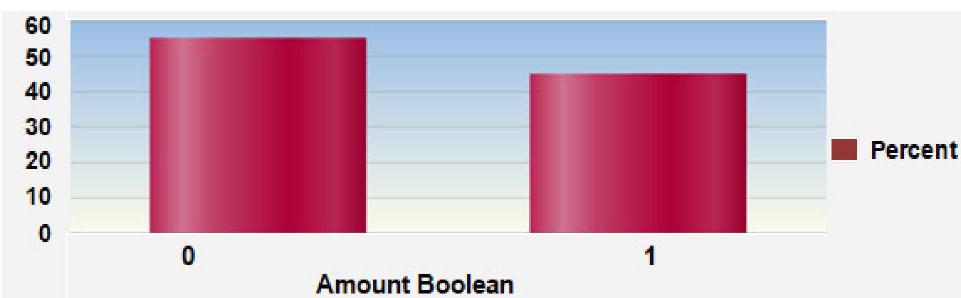


Figure 16. Amount cluster between legitimate and discrepant transactions for freight dataset 2.

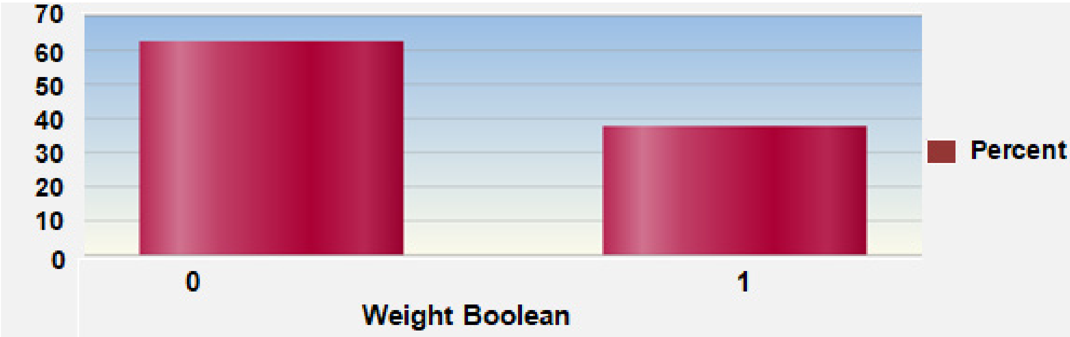


Figure 17. Weight cluster between legitimate and discrepant transactions for freight dataset 2.

3.8.4. Proposed Methods and Novel Clustering Method with Freight Dataset 2

Figure 18 shows the weight and revenue difference in a three-dimensional format using a novel clustering method with consistent data after completing the execution of the proposed methods (**Algorithm 1** and **Algorithm 2**). Furthermore, it uses 254,162 records to develop data clusters in a three-dimensional format using three attributes ("gross_amount", weight_boolean, and amount_boolean). Similarly, the red color shows transaction discrepancies, and the green color presents real records regarding the attribute ("amount_boolean"). For example, if the attribute ("amount_boolean") held the value ("1"), it would provide the maximum value of lost revenue, which is less than 7.0 × 106. Similarly, if the attribute ("weight_boolean") held the value ("1"), it would provide the maximum value of lost weight, which is less than 5000K. Finally, it is suggested that the proposed methods can overcome air freight losses via novel clustering method. Hence, it presents the profit and loss of freight business between sales and shipment datasets, whether the attributes ("amount_boolean" and "weight_boolean") and their values are matched or not.

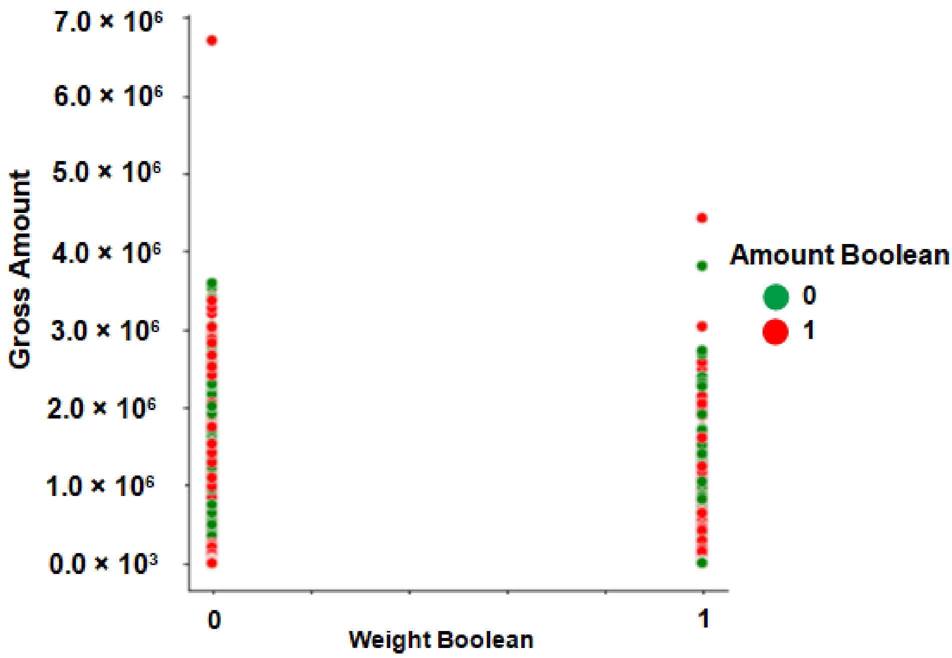


Figure 18. Data cluster between revenue and weight with legitimate and discrepant transactions for freight dataset 2.

3.9. Managerial Insights

There is no real-time transaction between shipment and sales systems, as they independently operate in the freight industry. Regarding the sales system, freight forwarders and airlines run it, but freight forwarders are the ones who make the majority of the sales revenue. However, because of issues with cargo space, different airlines handle the shipment process. This section provides managerial insights derived from the findings of related studies and proposed methods. Additionally, it identifies best practices pertinent to those topics. The present investigation revealed that data consistency, real-time transactions, technology, and supply chain disruptions are the most important topics. Consequently, the following insights can be utilized from the analysis results to identify discrepancies between the sales and shipment systems and to prevent future disruptions.

3.9.1. Data Consistency

There is no consistency between the two independent systems, as mentioned in the figures (**Figures 7**–**10** and **12**). Therefore, it poses a financial problem for airlines due to the mismanagement of data between sales and shipment systems. Gross amount and weight should be considered important attributes for the air cargo business. Therefore, air cargo managers should have a plan to maintain data consistency in sales and shipment carrier systems regarding the above-mentioned attributes to enhance the profit of airlines. Otherwise, it could magnify the effects of the air cargo industry and cause additional negative consequences.

3.9.2. Real-Time Transactions

There is no real-time transaction between sales and shipment systems due to the nature of the air cargo business. In other words, the booking process is completed by freight forwarders or airline offices, but there is no guarantee that the shipment process follows the same route as mentioned in the manifest due to aircraft space problems and flight diversion. Therefore, multiple airlines provide carrier services for handling the cargo to its destination. Moreover, both systems manifest the cargo using important attributes ("gross_amount" and "weight"), so it is a very significant step to establish real-time communication between both systems. However, if it is not possible to provide real-time transactions between independent systems, the air cargo manager should develop a batch mode or decentralized strategy between sales and shipment systems after integrating the freight data into the central repository. Finally, it requires following up on the process of **Figure 5** to resolve the inconsistencies between independent systems.

3.9.3. Technologies

Various innovative technologies, such as blockchain and Oracle PL/SQL, are valuable tools that could be used in detecting discrepancies in the occurrence of data inconsistency between independent systems. For example, blockchain can be used to detect transaction discrepancies between sales and shipment datasets and enhance the transparency and traceability of air cargo businesses. Similarly, Oracle PL/SQL can be used to detect air cargo discrepancies regarding the difference values of attributes ("gross amount" and "weight") as a batch mode method, where numbers of records can be validated by the attribute ("airwaybill").

3.10. Proposed Methods and Scalability

There are many parameters involved in detecting document discrepancies in the air freight industry, as mentioned in the future work section. For example, the novel proposed methods use two parameters ("gross_amount" and "weight") to detect discrepancies in the air cargo business. We can consider other parameters for scalability in both batch processing and blockchain implementation, such as the attributes ("incidental_charges", "commission", partial_shipment and "agreement"). In other words, proposed algorithms are not limited to the air freight industry, so we can apply the proposed methods in other domains to detect business discrepancies. Firstly, the credit card system is based on two independent modules: (i) customers utilize the credit card to buy items from the supermarket as per transaction; (ii) clients can pay the credit card bills within forty days. Therefore, we can apply the proposed algorithms to detect discrepancies in the credit card system using credit card transactions and customer bill datasets. In some big organizations, they provide medical services to employees, where doctors prescribe the medicines and pharmacists issue the medicines to patients or employees. Therefore, we can apply these algorithms in the health system to detect discrepancies between doctors who prescribe medicines and pharmacists who issue the medicines. In a supermarket, the owner buys the list of items from companies, and customers purchase the items from the supermarket, so there is a need to apply proposed algorithms between the buyer and sales entities to publish the profit and loss report. Regarding insurance companies, they introduce various insurance policies, such as vehicle and health insurance policies. In the case of claims, if the vehicle breaks down, customers will claim the breakdown cover. However, it is not certain whether it is an actual vehicle accident or not, so it needs to contact the police office for accident verification before completing the customer's claim report. In this situation, we have two independent systems: (i) the police accident system and (ii) the insurance system. The proposed methods can detect insurance claim discrepancies between police accidents and insurance systems. The degree verification process has two major systems: (i) universities issue the degrees to students; (ii) the Higher Education Commission verifies the students' degrees. Therefore, if there is a mismatch between universities' records and the higher education commission system, the proposed algorithms can detect document discrepancies. Finally, the proposed methods detect discrepancies between independent systems after changing the system's parameters. Hence, the proposed methods provide the scalability.

4. Intrinsic and Extrinsic Threats to Validation and Proposed Systems

Figure 19 shows a comprehensive analysis of the air freight system. It presents the benefits, demerits, threats, and goals of the freight industry. The proposed systems have provided a few benefits. Firstly, a novel batch mode method validates each transaction between sales and shipment datasets based on the attribute ("airwaybill") and detects transaction discrepancies in the air freight system. Secondly, if any variation exists between records, a novel blockchain method will detect discrepancies between datasets (sales and shipment). Furthermore, it diminishes inconsistencies between two independent datasets. Moreover, it removes demand and supply disruptions and enhances freight profit. Besides this, it establishes a balance between sales and shipping modules. In other words, it takes the attribute ("gross_amount") instead of the revenue field for processing because it does not deal with origin or destination currency codes, or it does not need to apply a conversion rate to the attribute ("gross_amount") when detecting the discrepancies between datasets (shipment and sales). Concerning the goals, it integrates two independent systems into a central repository, and additionally, it detects discrepancies between subsystems. It tracks disruptions of important attributes ("yield" and "weight") for the air cargo industry. Regarding the threats, it needs to introduce real-time transactions between two independent systems simultaneously. The blockchain technology develops different chains to identify the document forgery, but it is a decentralized approach. Finally, the proposed methods take significant time to complete the execution of records, so they provide the time complexity, which is about T (n) = θ (n), but they execute the million records within seconds.

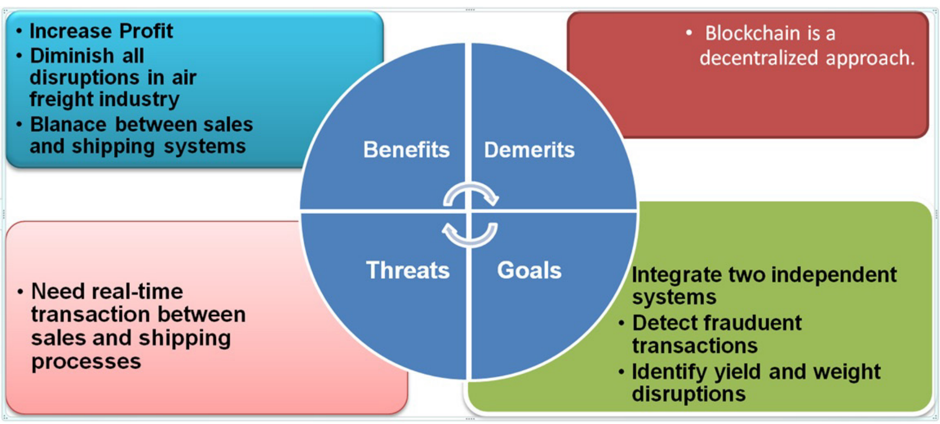


Figure 19. BDTG analysis of the air freight system.

5. Comparison between Existing Discrepancies Mechanisms and Proposed Methods

Table 3 shows the comparison between existing and proposed methods. It provided the solution for detecting discrepancies in the shipping industry via a supervised classification model using location parameters. However, other factors that affect discrepancies in the shipping domain include shipment weight, payment method, and customer profile, which were not assessed in this article because of insufficient legitimate production data [8]. This study presents a Contextual Anomaly Detection (CAD) framework designed to identify and analyze various types of anomalies in logistics operational processes [35]. It takes into account specific contexts, including branches, locations (sender and receiver cities), times, job types, and units [35]. It emphasizes the essential role of CAD in precisely detecting unusual patterns and enhancing the efficiency of operational workflows [35]. It explained the theoretical research and addressed the issues in air cargo operations. However, it is required to resolve real-world air freight problems for airlines [22]. It discussed the supply and demand disruptions for sales datasets. Concerning the demand dataset, the major findings indicate a demand imbalance in the majority of the 110 region-pair combinations studied. However, the supplied dataset shows a large imbalance in freighter capacity, a smaller imbalance in integrator capacity, and a small imbalance in wide-body belly capacity [36]. The suggested approach utilizes the IATA ONE Record standard to provide Latvian air cargo supply chain stakeholders with a clear, step-by-step plan aimed at enhancing data integration and operational efficiency. Implementing this concept is expected to decrease the reliance on manual supply chain processes and lower operating expenses, while also improving supply chain visibility and enhancing service quality. The model of integrating IATA One Record can aid industry stakeholders in evaluating the viability of similar digitalization strategies [37]. The stochastic frontier approach reduces disparities between the passenger and cargo transportation industries [38]. It was discovered that blockchain reduces both process- and document-related corruption in the air freight industry. This model is used to detect document discrepancies using blockchain, which shows the complex interplay between identities, institutional actors, technical and other resources, and practices [39]. It investigated and summarized the global practice of using blockchain technology to manage information and financial flows in the air cargo industry. This strategy made the best use of available airport resources to reduce service delays. It explains the primary issue of creating transparency in information exchange among all air cargo transportation participants [40]. It investigated the structure of the Air Cargo Handling Process (ACHP), as well as the safety and environmental aspects of the process, using a variety of scientific cognition strategies. This study proposes a blockchain-based ecosystem designed to support a sustainable and service-oriented last-mile delivery model. It includes four key stakeholders: customers, retailers, fleet carriers, and optimization specialists (miners). The model utilizes blockchain features such as transparency, traceability, agile planning, and secure claim management during the delivery process. It also discusses reverse logistics for product recycling, which enhances environmental sustainability [41]. This research can improve the process's safety and environmental aspects, which are critical for process functionality, as well as the quality level of service that meets customer requirements and the process's sustainability [42]. This model improves the customer's satisfaction and the service level. However, the air cargo handling process did not incorporate the layer between sales and shipment systems. It proposed a layered structure for incorporating blockchain into air freight shipping and logistics to increase its effectiveness and resistance to security attacks. This approach improved the service level between shipping and logistics rather than other modules of the air cargo industry, like the sales module [30]. Financial institutions understand the air freight industry as vulnerable due to discrepancies and disruptions. To address these problems, this study introduced a blockchain-based platform for air cargo financing (ACFB), which would make the supply chain more transparent, trustworthy, and efficient for all stakeholders and financial institutions [43]. It presents a novel conceptual framework that incorporates drone-based last-mile delivery into the orchestration architecture of fifth-party logistics (5PL) systems. By utilizing a multi-agent digital twin model, it combines various technologies, including the Internet of Things (IoT) for real-time tracking, AI-driven metaheuristics (such as Adaptive Large Neighborhood

Search, Particle Swarm Optimization, and Non-dominated Sorting Genetic Algorithm II) for optimizing routes and hubs, and blockchain technology to ensure compliance with service level agreements (SLAs) [44].

Table 3. Comparison between existing discrepancies mechanisms and proposed methods.

S.no.	Air Freight Discrepancies Mechanism	Purpose	Scope
1	Classification [8]	It provided the solution for detecting discrepancies in the shipping industry [8].	Based on the location parameter [8].
2	Contextual Anomaly Detection (CAD) [35]	It analyzes various types of anomalies in logistics [35].	It detects unusual patterns [35].
3	Cargo supply and demand model [36]	It shows an imbalance between demand and supply processes [36].	It improves supply and demand processes [36].
4	IATAone record [37]	Enhancing supply chain visibility and service quality [37].	Integrating IATA One Record can aid industry stakeholders in evaluating the viability of similar digitalization strategies [37].
5	Stochastic frontier approach [38]	Combination carriers with cargo-dedicated aircraft (CDA) face challenges in their air cargo business, including reduced activities, rate, and yield, economies of scale, density, and productivity [38].	It reduces disparities between the passenger and cargo transportation industries [38].
6	Blokchain [39]	Both process- and document-related corruption on a global scale across continents and economies, bringing to light improper gift giving, kickbacks, or inappropriate social activities [39].	Corruption [39]
7	Blokchain [40]	This strategy made the best use of available airport resources to reduce service delays. It explains the primary issue of creating transparency in information exchange among all air cargo transportation participants [40].	It improves airport services [40].
8	Blockchain features [41]	It provides transparent, agile planning and secure claim management during the delivery process [41].	It is limited to the delivery process and reserve logistics [41].
9	Air Cargo Handling Process [42]	The Air Cargo Handling Process (ACHP) is critical for any airport or cargo handling agent that offers cargo handling services [42].	Process safety and environmental aspects [42].
10	Blockchain [30]	This study proposes a layered structure for incorporating blockchain into air freight shipping and logistics to increase its effectiveness and resistance to security attacks. It makes the supply chain more transparent, trustworthy, and efficient for all stakeholders and financial institutions [30].	Smart shipping and logistics [30].
11	Blockchain-based platform for air cargo financing (ACFB) [43]	It makes the supply chain more transparent, trustworthy, and efficient for all stakeholders and financial institutions [43].	It enhances supply chain [43].
12	Novel conceptual framework [44]	It incorporates drone-based last-mile delivery into the orchestration architecture of fifth-party logistics (5PL) systems [44].	It improves the performance using digital twin and other technologies [44].

The above-mentioned strategies detect discrepancies between freight documents and imbalances between sales datasets, improve the service model between shipping and logistics, and bring transparency between airport services via blockchain. However, prior researchers did not discuss the freight disruptions between sales and shipment systems. Concerning proposed methods, these algorithms detect discrepancies based on the important attributes ("gross_amount" and "weight") in the air freight industry. Firstly, a novel batch mode method loads independent systems' data into a central repository. Furthermore, if the attribute ("airwaybill") and its value match between the sales and shipping datasets, it will detect discrepancies between two independent systems. As a result of this, it provides a difference between parameters ("gross_amount" and "weight"), as mentioned in **Figures 7** and **8**. If we compared this strategy with previous literature, it would provide hidden facts about two independent systems (sales and shipment). As a result of this, it diminishes freight disruptions and enhances profit for air cargo companies. Similarly, if there is any variation between datasets, the proposed blockchain algorithm assigns new hash keys because it keeps the record of the previous hash keys, as mentioned in **Figure 12**. If we compared the proposed blockchain method with the aforementioned algorithms, it would detect document discrepancies between two independent systems (shipping and sales). Consequently, the proposed methods diminish all the types of disruptions in the freight industry, such as supply and demand disruption, freight forwarder disruption, route disruption, and cargo space disruption. The outcome of this research is to enhance the freight yield as well as present the actual loaded weight of the aircraft during the shipping process.

6. Supply & Chain Disruptions and Freight System

Figure 20 shows the causes and effects of the air freight system. Furthermore, if there is an inconsistency between freight forwarders and the airline shipping process, it raises a few causes and effects of demand and supply processes, such as profit disruption, weight disruption, supply chain collapse, and economic decline. It requires improving air cargo processes, whether the air cargo manager pursues a centralized or decentralized approach to the freight business. Moreover, air cargo airlines make a manifest of parcels using important attributes ("airwaybill", "gross_amount", currency_code, and "weight") from sales data, so this strategy can facilitate synchronization between sales and shipping systems. However, it may be possible that freight forwarders present the wrong values of the manifest in the airwaybill during the sales process, so the airline industry should measure the weight of cargo along with revenue before lifting the aircraft. Therefore, the Air Cargo Manager recommends the decentralized approach, where it loads data into a central repository and detects discrepancies using proposed methods. In other words, if the air cargo manager fails to implement the recommended strategy, discrepancies in the attributes ('gross_amount' and 'weight') and their values between two independent systems could lead to significant issues and heightened freight losses for the airline industry. For example, if the shipper manifests the cargo and the attributes ("gross_amount" and "weight") and their values are less than the shipping booking, it would provide a financial loss to the airline industry. Consequently, the cargo manifest pays other airlines more than its revenue due to the inconsistency between independent systems.

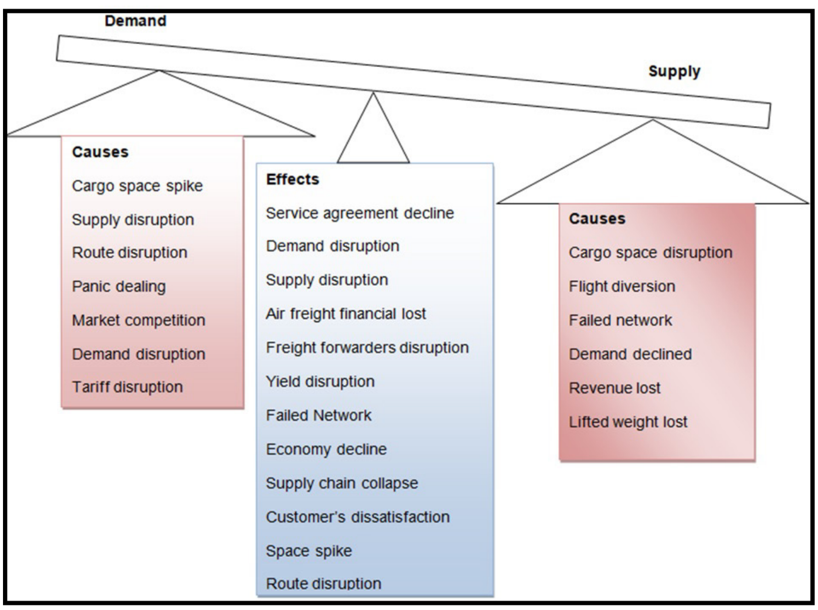


Figure 20. Supply & chain disruptions and freight system.

To encapsulate the principal managerial discoveries derived from the examination of literature and corroborating analyses, the creation of algorithms utilizing blockchain technology and Oracle PL/SQL could equip air cargo managers and legislators for potential supply chain disruptions in the future.

7. Findings and Recommendations

The major findings indicate an imbalance between sales and shipping datasets based on attributes ("gross amount" and "weight") and their values, whether the data are matched or not. It uses two freight datasets to detect discrepancies in graphical format as well as develop the data clusters. It presents airline losses and creates a serious problem with freight forwarders and other stakeholders. For example, it presents the maximum revenue lost for specific airwaybills (72,504,247–75,290,913.5), as mentioned in **Figure 7**. Similarly, it shows the list of airwaybills (less than 12,484,816, 34,349,299–41,637,460, 48,925,621–56,213,782, 63,501,943–70,790,104, and greater than 70,790,104) to lost revenue (-5.0×10^3, -9.0×10^4, -5.0×10^4, -1.5×10^4, and -4.0×10^4), respectively, as mentioned in **Figure 9**. With regard to weight discrepancies, it indicates the positive and negative weight discrepancies (-1.5×10^6 and 1.0×10^6) according to specific airwaybills from 52,997,581.5–55,784,248 and greater than 75,290,913.5, respectively, as mentioned in **Figure 8**. Furthermore, it shows specific airwaybills (less than 12,484,816, 34,349,299–41,637,460, 48,925,621–56,213,782, 63,501,943–70,790,104, and greater than 70,790,104) to lost weight (-1.0×10^6, -1.01×10^6, -6.0×10^5, -5.0×10^5, and -6.0×10^5), respectively, as mentioned in **Figure 10**. The findings indicate that it presents a problem for a few freight forwarders who work with airlines. The proposed methods take significant time to complete the processing of bulk records, as shown in tables (**Tables 1** and **2**). Hence, it is suggested that the air cargo manager should identify the freight forwarders who are tampering with the sales dataset as well as track the shipping process via specific airwaybills.

8. Paired Samples Statistics Test and Freight Datasets

Figure 21 shows the paired samples statistics test after providing the freight dataset. Furthermore, it makes two pairs ({"airwaybill", "weight_bol"} and {"airwaybill" and "amount_bol"}) using the IBM SPSS tool before taking the *t*-test. With regard to the null hypothesis (H_0), there is no difference between attributes. However, H_1, there is a difference between attributes. The evidence proved that there is a significant revenue difference between the sales and shipment datasets because the *p*-value for the attribute ("amount_bol") is $0.000 < 0.05$. Therefore, we accept the null hypothesis and reject the alternative hypothesis. The *p*-value for the attribute ("weight_bol") is approximately 0.000, which is less than 0.05. Based on the evidence, it accepts the null hypothesis and rejects the alternative hypothesis H_1 regarding the attribute ("weight"). Hence, it provides business transparency.

9. Future Work

The air freight system deals with many stakeholders, such as freight forwarders, other airlines, and IATA (International Air Transport Association). However, there is no proper integration between systems, which affects revenue and the economy. Certain crucial parameters must be identified and compared with those of other subsystems within the freight business. For example, freight forwarders take a commission for each cargo from the airline industry. Therefore, the system will need to detect transaction discrepancies based on the "commission" attribute. Another attribute (incidental_charges) is used to collect the payment at the destination for the shipment type ("charges collected"), but unfortunately, the airline industry could not get control of incidental charges due to a lack of integration between the shipping system and airport services across the world. Consequently, it requires collecting the data from domestic and international airports and integrating it into the air freight system. Similarly, when the airline lifts a large volume of cargo in the form of pieces, it is known as a partial shipment. As a result, this partial shipment is transported by several airlines since they lift the partial weight and deliver it to the destination after consolidating there. There is no check and balance between airlines for partial shipments. Finally, it is suggested that the air freight business grow more in the future after resolving the above-mentioned problems.

Paired Samples Statistics

	Variable	Mean	N	Std. Deviation	Std. Error Mean
Pair 1	AIRWAYBILL	59,070,271.59	21,896	12,546,950.04	84,792.175
	AMOUNT_BOL	0.86	21,896	0.351	0.002
Pair 2	AIRWAYBILL	59,070,271.59	21,896	12,546,950.04	84,792.175
	WEIGHT_BOL	0.85	21,896	0.355	0.002

Paired Samples Correlations

	Variables	N	Correlation	Sig.
Pair 1	AIRWAYBILL & AMOUNT_BOL	21,896	−0.103	0.000
Pair 2	AIRWAYBILL & WEIGHT_BOL	21,896	−0.064	0.000

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
	Variables	Mean	Std. Deviation	Std. Error Mean	Lower	Upper	*t*	df	Sig. (2-tailed)
Pair 1	AIRWAYBILL - AMOUNT_BOL	59,070,270.74	12,546,950.08	84,792.175	58,904,071.94	59,236,469.54	696.648	21,895	0
Pair 2	AIRWAYBILL - WEIGHT_BOL	59,070,270.74	12,546,950.07	84,792.175	58,904,071.95	59,236,469.54	696.648	21,895	0

Figure 21. Paired samples statistics test and freight dataset.

10. Conclusions

Highlighted views can identify the air freight problems, such as weight and revenue discrepancies. The current infrastructure is complex and involves multiple stakeholders, necessitating the introduction of an integration module. This module would connect independently developed systems to enhance managerial insight. For example, the air cargo system is mainly divided into two major modules (sales and shipment); each module operates solely, so it creates inconsistencies between systems and affects the other modules, such as the general ledger and ERP. With regard to the risk factor, there is no real-time transaction between sales and shipment systems, so it brings too many disruptions to the freight industry, such as yield disruption, weight disruption, supply chain disruption, route disruption, and freight forwarder disruption. The existing methods cannot resolve the inconsistency issue between sales and shipment systems, as mentioned in **Table 3**. However, the proposed novel batch mode method detects transaction discrepancies based on important parameters ("airwaybill", "gross_amount", and "weight") and their values. This strategy integrates the freight data into a central repository and detects discrepancies in a graphical format using Oracle Miner. It presents the problem for specific airwaybills, as mentioned in the figures (**Figures 7–10**). In addition to misleading records, the air cargo manager could take necessary actions against a list of discrepant airwaybills. The aforementioned views can identify discrepancies in the air freight industry through the innovative use of a blockchain algorithm. It is one of the most sophisticated ways to detect record discrepancies between datasets (shipment and sales) concerning any variation that may exist in attributes ("gross_amount" and "weight") and their values. The blockchain technology establishes a chain between previous and current transactions and assigns the new hash keys when detecting the variation between datasets. It provides a list of transaction discrepancies on a Python prompt and in a comma-delimited file without the user's interference. Blockchain technology is more suitable for sales and shipping modules due to the prior and current transactions dealt with by chains. Consequently, the proposed methods can diminish all types of disruptions in the freight industry and enhance yield. The proposed novel batch mode update method provides input to existing clustering algorithms for developing data clusters in two- or three-dimensional format. The findings indicate the legitimacy and discrepancies in the air freight industry. Statistically, it provides a significant result. Regarding future work, it needs to choose other attributes ("incidental_charges", "agreement", partial shipment, and "agent commission") to detect discrepancies in the air freight business. Finally, it is suggested that air cargo managers can increase the air cargo yield after implementing the proposed techniques in the production environment. Hence, the proposed methods provide scalability, reliability, and transparency in various domains.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The author declares no conflict of interest.

Appendix A

Table A1 shows the details of all variables that are related to the air freight industry.

Table A1. All variable details.

S.no.	Variables	Scope
1	Airwaybills	It shows complete revenue, but it does not relate to any specific currency.
2	Gross_amount	It shows complete revenue, but it does not relate to any specific currency.
3	Weight	It presents the weight regarding to true itineraries.
4	Rev_discrepancies	It shows the revenue discrepancies.
5	Wgt_discrepancies	It shows the weight discrepancies.
6	Amount_boolean	It presents the legitimate and inaccurate transactions regarding attribute ("Gross_amount").
7	Weight_boolean	It presents the legitimate and inaccurate transactions regarding attribute ("Weight").

References

[1] Matthews, R., Rutherford, B.N., Edmondson, D., et al., 2022. Uncertainty in Industrial Markets: The COVID-19 Pandemic. Industrial Marketing Management. 102, 364–376. DOI: https://doi.org/10.1016/j.indmarman.2022.02.006

[2] Detwal, P.K., Soni, G., Jakhar, S.K., et al., 2023. Machine Learning-Based Technique for Predicting Vendor Incoterm (Contract) in Global Omnichannel Pharmaceutical Supply Chain. Journal of Business Research. 158, 113688. DOI: https://doi.org/10.1016/j.jbusres.2023.113688

[3] Merkert, R., 2023. Air Cargo and Supply Chain Management. In: Sarkis, J. (Ed.). The Palgrave Handbook of Supply Chain Management. Palgrave Macmillan: Cham, Switzerland. pp. 1–18. DOI: https://doi.org/10.1007/978-3-030-89822-9_90-1

[4] Shaban, I.A., Chan, F.T.S., Chung, S.H., 2021. A Novel Model to Manage Air Cargo Disruptions Caused by Global Catastrophes such as COVID-19. Journal of Air Transport Management. 95, 102086. DOI: https://doi.org/10.1016/j.jairtraman.2021.102086

[5] Alexander, D.W., Merkert, R., 2021. Applications of Gravity Models to Evaluate and Forecast US International Air Freight Markets Post-GFC. Transport Policy. 104, 52–62. DOI: https://doi.org/10.1016/j.tranpol.2020.04.004

[6] Zhang, L., Hou, M., Liu, Y., et al., 2022. Measuring Beijing's International Air Connectivity and Suggestions for Improvement Post COVID-19. Transport Policy. 116, 132–143. DOI: https://doi.org/10.1016/j.tranpol.2021.11.015

[7] Oum, T.H., Wu, X., Wang, K., 2024. Impact of Air Connectivity on Bilateral Service Export and Import Trade: The Case of China. Transport Policy. 148, 219–233. DOI: https://doi.org/10.1016/j.tran-

pol.2024.01.015
[8] Subramaniam, G., Mahmoud, M.A., 2021. Fraud Detection in Shipping Industry Using K-NN Algorithm. International Journal of Advanced Computer Science and Applications. 12(4). DOI: https://doi.org/10.14569/IJACSA.2021.0120460
[9] Smith, L.D., Bilir, C., 2023. Modelling Airline Operations at Major Commercial Airports for Strategic Decision Support. Transportation Development Research. 1, 242–256. DOI: https://doi.org/10.55121/tdr.v1i1.101
[10] Zong, Z., Feng, T., Wang, J., et al., 2025. Deep Reinforcement Learning for Demand-Driven Services in Logistics and Transportation Systems: A Survey. ACM Transactions on Knowledge Discovery from Data. 19(4), 89. DOI: https://doi.org/10.1145/3708325
[11] Nguyen, Q.H., Nguyen, I.V., 2025. The Relationship between Air Freight and International Trade. Journal of Southeast Asian Economies. 42(1), 70–92. DOI: https://www.jstor.org/stable/27378053
[12] Hong, S.-J., Kim, W., Hiatt, B., 2025. Examining Airport Agility at Air Cargo Hub Airports. Journal of Air Transport Management. 122, 102710. DOI: https://doi.org/10.1016/j.jairtraman.2024.102710
[13] Michielsen, S., Gevaers, R., Dewulf, W., 2025. A Historical Overview and Analysis of E-Commerce's Milestones and Its Growing Connection with Air Transport. Journal of Shipping and Trade. 10(14). DOI: https://doi.org/10.1186/s41072-025-00203-5
[14] Chung, S., Kim, H., Choi, D., 2025. Impacts of Internal and External Uncertainties on Logistics Service Flexibility in Cross-Border E-Commerce Logistics: Evidence from South Korea. Systems. 13(12), 1082. DOI: https://doi.org/10.3390/systems13121082
[15] Ong, C.S., Ahmad, W.N.K.W., 2025. The Relationship between Challenges Faced by Logistics Industry in Johor and the Implementation of Sustainable Logistics Practices. Research in Management of Technology and Business. 6(2), 33–45. Available from: https://penerbit.uthm.edu.my/periodicals/index.php/rmtb/article/view/20062
[16] Andrejić, M., Pajić, V., 2025. Strengthening the Sustainability and Resilience of Global Supply Chains: An Integrated Risk Assessment Framework for International Logistics Security. Opportunities and Challenges in Sustainability. 4(1), 56–69. DOI: https://doi.org/10.56578/ocs040105
[17] Wu, J., 2025. The Evolution of Supply Chains: From Simple to Complex, from Linear to Network. Springer Series in Supply Chain Management. 26, 3–27. DOI: https://doi.org/10.1007/978-981-96-3228-2_1
[18] Yang, Y.-C., Chen, Y.-C., 2025. Risk Management of Transaction Security Using Blockchain Technology in the International Logistics Industry. Sustainable Futures. 10, 101537. DOI: https://doi.org/10.1016/j.sftr.2025.101537
[19] Gajdos, A., Gajdos, A., 2025. Trends in Logistics – Future Plans. Annales Universitatis Mariae Curie-Sklodowska, Sectio H – Oeconomia. 59(2), 57–92. DOI: https://doi.org/10.17951/h.2025.59.2.57-72
[20] Rashid, A.A., See, K.F., Yu, M.-M., 2024. Measuring Airline Efficiency Using a Dynamic Network Data Envelopment Analysis in the Presence of Innovation Capital. Technological Forecasting and Social Change. 206, 123457. DOI: https://doi.org/10.1016/j.techfore.2024.123457
[21] Raj, A., Mukherjee, A.A., Jabbour, A.B.L.S., et al., 2022. Supply Chain Management during and post COVID-19 Pandemic: Mitigation Strategies and Practical Lessons Learned. Journal of Business Research. 142, 1125–1139. DOI: https://doi.org/10.1016/j.jbusres.2022.01.037
[22] Estima, J., da Cunha, P.R., Barata, J., 2025. Blockchain in Shipment Transportation. In Handbook of Blockchain Technology. Edward Elgar Publishing: Cheltenham, UK. pp. 261–276. DOI: https://doi.org/10.4337/9781803922805.00029
[23] Lau, C.W., Liu, J., Ma, X., et al., 2024. Performance Analysis of a Blockchain-Based Messaging System Implementation for Air Cargo Supply Chains. International Journal of Production Research. 63(14), 5428–5451. DOI: https://doi.org/10.1080/00207543.2024.2407903
[24] Haq, I., Nandal, V., Uppal, H., 2025. Blockchain Applications in Aviation: Securing Transactions, Streamlining Operations, and Improving Passenger Experience. Information Systems Engineering and Management. 57(14), 131–154. DOI: https://doi.org/10.1007/978-3-031-95341-5_7
[25] Barua, B., Kaiser, M.S., 2025. A Next-Generation Approach to Airline Reservations: Integrating Cloud Microservices with AI and Blockchain for Enhanced Operational Performance. IET Blockchain. 5(1), e70020. DOI: https://doi.org/10.1049/blc2.70020
[26] Javaid, M., Haleem, A., 2025. Role of Digital Twin and Blockchain in Logistics and Supply Chain Management. In: Nguyen, T.A. (Ed.). Digital Twin and Blockchain for Sensor Networks in Smart Cities. Elsevier: Amsterdam, Netherlands. pp. 243–264. DOI: https://doi.org/10.1016/B978-0-443-30076-9.00012-1
[27] Zohourian, M.A., Pamidimukkala, A., Kermanshachi, S., 2025. Application of Blockchain Technol-

ogy in the Transportation Industry. In Proceedings of the International Conference on Transportation and Development, Glendale, AZ, USA, June 8–11, 2025; pp. 261–276. DOI: https://doi.org/10.1061/9780784486191.033
[28] Sarkis, J., Bai, C., Wang, Y., 2025. Multistakeholder Production and Logistics Management Through Blockchain Technology Applications. International Journal of Production Research. 63(14), 5031–5042. DOI: https://doi.org/10.1080/00207543.2025.2511567
[29] Zineb, K.I., Hrimech, H., Lachgar, M., 2025. Blockchain in Logistics: The Case of TradeLens and Its Impact on Supply Chain Efficiency. In Proceedings of the 16th International Conference on Logistics and Supply Chain Management (LOGISTIQUA), Casablanca, Morocco, 28–30 May 2025; pp. 1–7. DOI: https://doi.org/10.1109/LOGISTIQUA66323.2025.11122825
[30] Sahoo, R., Bhowmick, B., Tiwari, M.K., 2021. Smart Integration of Blockchain in Air Cargo Handling for Profit Maximization. IFIP Advances in Information and Communication Technology. 631, 107–114. DOI: https://doi.org/10.1007/978-3-030-85902-2_12
[31] Huang, Z., Zheng, H., Li, C., et al., 2024. Application of Machine Learning-Based K-Means Clustering for Financial Fraud Detection. Academic Journal of Science and Technology. 10(1), 33–39. DOI: https://doi.org/10.54097/74414c90
[32] Salman, M., 2023. A Novel Clustering Method With Consistent Data in a Three-Dimensional Graphical Format Over Existing Clustering Mechanisms. Information Sciences. 649, 119634. DOI: https://doi.org/10.1016/j.ins.2023.119634
[33] Vaghani, A., Sood, K., Yu, S., 2022. Security and QoS Issues in Blockchain-Enabled Next-Generation Smart Logistic Networks: A Tutorial. Blockchain: Research and Applications. 3(3), 100082. DOI: https://doi.org/10.1016/j.bcra.2022.100082
[34] Yang, W., Wen, J., Wang, F., et al., 2025. Trade Dependency and Technological Specialization in the ICT Supply Chain: Structural Dynamics and Strategic Autonomy in Major Economies. Telecommunications Policy. 49(9), 103037. DOI: https://doi.org/10.1016/j.telpol.2025.103037
[35] Çay, A., Bayram, B., Gökyılmaz, G., et al., 2025. Contextual Anomaly Detection in Logistics and Supply Chain Management. In Proceedings of the 2025 10th International Conference on Machine Learning Technologies (ICMLT), Helsinki, Finland, 23–25 May 2025; pp. 427–432. DOI: https://doi.org/10.1109/ICMLT65785.2025.11193370
[36] Van Bockstaele, V., Buyle, S., Dewulf, W., 2023. Solving the Mystery of Discrepancies and Double Counting in Air Cargo through Demand and Supply Big Data Analysis. Journal of the Air Transport Research Society. 1(1), 81–100. DOI: https://doi.org/10.59521/6A961EF46EB809C5
[37] Jurgelane-Kaldava, I., Effenberger, W.V., Batenko, A., et al., 2025. Digitalization of Air Cargo Supply Chains: A Case Study of Latvia. Systems. 13(6), 468. DOI: https://doi.org/10.3390/systems13060468
[38] Hong, S.-J., Kim, H.W., Niranjan, S., 2023. Challenges to the Air Cargo Business of Combination Carriers: Analysis of Two Major Korean Airlines. Journal of Air Transport Management. 108, 102360. DOI: https://doi.org/10.1016/j.jairtraman.2023.102360
[39] Sarker, S., Henningsson, S., Jensen, T., 2021. The Use of Blockchain as a Resource for Combating Corruption in Global Shipping: An Interpretive Case Study. Journal of Management Information Systems. 38(2), 338–373. DOI: https://doi.org/10.1080/07421222.2021.1912919
[40] Poleshkina, I., 2021. Blockchain in Air Cargo: Challenges of New World. MATEC Web of Conferences. 341, 00021. DOI: https://doi.org/10.1051/matecconf/202134100021
[41] Mofatteh, M.Y., Valilai, O.F., 2025. A Blockchain Ecosystem Model for a Sustainable Last-Mile Delivery Operation Management. Transportation Engineering. 21, 100379. DOI: https://doi.org/10.1016/j.treng.2025.100379
[42] Drljača, M., Štimac, I., Vidović, A., et al., 2020. Sustainability of the Air Cargo Handling Process in the Context of Safety and Environmental Aspects. Journal of Advanced Transportation. 2020(1), 1232846. DOI: https://doi.org/10.1155/2020/1232846
[43] Yadav, P., Bhosale, R., Sahoo, R., et al., 2022. Advances in Air Cargo Financing Using a Consortium Blockchain. IFAC-PapersOnLine. 55(10), 737–742. DOI: https://doi.org/10.1016/j.ifacol.2022.09.497
[44] Lakhwani, T.S., 2025. Integrating 5PL Frameworks with Drone-Based Last-Mile Delivery: A Model for Future-Ready Logistics. Transportation Development Research. 3(1), 27–46. DOI: https://doi.org/10.55121/tdr.v3i1.449