

Automating Package Type Selection in Oil and Gas Inbound Logistics: A Process Automation Framework for Cost Accuracy

Hassan M. Alhashem^{#1}, Adel A. Almarhoon^{#2}, Hani M. Aldhahri^{#3}

[#]Materials Logistics Department, Saudi Arabian Oil Company (Saudi Aramco)
Dhahran 31311, Kingdom of Saudi Arabia

¹hassan.alhashem@aramco.com

²adel.marhoon@aramco.com

³hani.dhahri@aramco.com

Received: Jul 27, 2026, **Revised:** Aug 18, 2026, **Accepted:** Aug 27, 2026, **Published Online:** Aug 29, 2026

Reviewers: Anonymous Peer Review

Citation: Alhashem, H. M., Almarhoon, A. A., Aldhahri, H. M. (2026). Automating Package Type Selection in Oil and Gas Inbound Logistics: A Process Automation Framework for Cost Accuracy. *International Journal of Supply Chain Management*, 15(4), 1-6, <https://doi.org/10.59160/ijscm.v15i4.6417>

Abstract— This paper addresses package type selection in oil and gas inbound logistics, a routine manual decision that directly determines the freight cost invoiced by the third-party carrier delivering shipments from suppliers and from the organization's own warehouses. Each selection maps to a package indicator carrying a contractually agreed carrier rate, so an inaccurate selection produces an inaccurate invoice. In a large oil and gas producing organization handling over one million packages annually, this decision depends on human judgment and is inherently error-exposed at scale, while manual verification covers less than five percent of delivered volume, leaving most selections unvalidated. Drawing on operational data covering millions of packages delivered across six years, this paper makes two contributions. It establishes package type selection as a distinct and previously unaddressed source of recurring cost leakage, a decision that no existing material classification standard governs and that the logistics literature has not treated as a problem. It then proposes a process automation framework that closes this gap without depending on labelled training data, beginning with a rule-based validation that checks each selection against the material's recorded dimensions and specifications and automates the cost calculation before shipment delivery. The framework is designed to evolve through machine-learning and master-data-driven phases toward fully autonomous, cost-accurate package selection, offering a transferable approach for any inbound logistics operation in which package type drives freight cost.

Keywords — Inbound logistics; Package type selection; Process automation; Cost accuracy; Oil and gas; Supply chain.

1. Introduction

Whenever an internal user publishes a purchase request for any needed material, a sequenced chain of business events is initiated. It starts with the procurement role: issuing a request for quotation and publishing a purchase order to the material suppliers. Once the order reaches the material issuer, whether an external supplier receiving a purchase order (PO) or an internal warehouse fulfilling a transfer order (TO), its role is to prepare the ordered materials, palletized and strapped, and then publish an Advance

Shipping Notice (ASN), which is communicated immediately to the 3PL for delivery.

This sequence also applies to Out-of-Kingdom (OOK) shipments, where materials first arrive at a port of entry an airport or seaport before the 3PL carrier takes responsibility for the onward delivery to the destination. In addition, there is a Vendor-to-Distribution-Center (VTD) incoterm, in which suppliers deliver materials directly to the carrier's hubs, where shipments are consolidated before being dispatched for final delivery.

With over 1million packages per year, a risk arises from wrongly selected package types, due to human error or supplier knowledge gaps. The package types are built as in-house formulas in the back end that calculate the RPE (Rate Pallet Equivalent). The formula is the material's length times its width, divided by the size of a standard pallet (120 × 120 cm). The resulting RPE value is linked to a cost, and this cost goes onto the monthly invoice. Therefore, if an inaccurate package type is selected, the invoice amount will be inaccurate as well.

This inaccuracy directly impacts the delivery cost contractually agreed between the oil and gas producing organization and the third-party logistics (3PL) carrier. To mitigate this, the organization introduced a manual verification practice two years ago, in which personnel across the material sorting areas physically inspect randomly selected packages multiple times daily. This manual process involves extensive back-and-forth communication through emails to make the necessary adjustments and ensure accurate payment. While valuable, this approach verifies only a small fraction of total volume less than 5 percent annually while leaving the majority of package selections unvalidated. It also does not cover the many packages delivered directly to thousands of internal-user destinations without passing through the material service centers.

For decades, package type selection in inbound logistics has relied on manual judgment, with issuers assigning

package categories based on experience and rule-of-thumb interpretation of package descriptions. While functional, this approach is inherently vulnerable to inconsistency and human error, particularly at scale and across fragmented systems. Prior efforts to improve logistics decision-making have explored automation and data-driven techniques [1], yet practical, industry-ready tools for package type selection have remained limited. More recently, advances in process automation, the availability of large operational datasets, and the maturity of rule-based and machine learning techniques have created a new opportunity to address this problem directly. Process automation, in particular, offers an immediately deployable means of standardizing package type selection and ensuring cost accuracy, while machine learning presents a longer-term path toward fully autonomous verification.

The main purpose of this paper is to demonstrate the potential of automating package type selection to eliminate selection errors and ensure accurate cost settlement across millions of packages. It proposes a feasible, dynamic framework that mitigates the cost impact of package selection, while positioning machine learning (ML) as the next stage of the framework's evolution.

2. Literature Review

2.1 Package Type Selection, Classification & Standardization

Package type selection in industrial and inbound logistics has historically relied solely on the judgment of the material issuer — whether originating from the supplier side or from within the organization. Given the scale of over one million unique materials, including sub-parts, that characterize oil and gas supply chains, and the diversity of selection references historically derived from available packaging options — boxes, cartons, and packaging materials stocked across organization warehouses — it is a near-certain outcome that no unified selection logic will prevail. This selection misalignment is further widened by the divergent priorities of two distinct professional communities: the logistics and supply chain discipline, which is primarily oriented toward cost effectiveness, and the oil and gas production domain, which is driven by operational continuity and asset integrity. Where these communities intersect in the issuer role, differences in industry background, technical expertise, and informal practice compound the inconsistency, leaving package type assignment inherently vulnerable to error at scale.

A direct consequence of the priorities misalignment between the logistics and supply chain discipline and the oil and gas production domain is the emergence of the common and largely unquestioned practice resulting in the selection of the most commonly assigned package type for a given material — a pattern driven by familiarity and repetition rather than technical accuracy. This default behavior reinforces, rather than resolves, the absence of a unified selection logic across the entities involved.

Due to the absence of a governing classification standard specific to package type assignment — unlike material classification frameworks such as UNSPSC and eCl@ss,

which address what a material is but not how it should be packaged — selection remains an informal, issuer-dependent judgment call. The cost impact of this gap is direct and measurable: where package type determines the freight rate agreed with the 3PL provider, an inaccurate selection translates immediately into cost miscalculation, invoice discrepancies, and the need for manual correction downstream.

2.2 Automation, AI/ML, and 3PL Cost Accuracy in Logistics

As Robotic Process Automation (RPA) technologies are fundamentally designed to eliminate high-volume, repetitive, rule-based tasks, their application can be leveraged at scale within the supply chain operations of oil and gas organizations [2]. The inbound logistics process, as it currently stands, relies on manual human intervention at a critical decision point — the selection of package type, whether a box, carton, or alternative packaging form — a task that is both repetitive in nature and directly consequential to freight cost settlement. The availability of structured operational data and predefined field requirements across logistical events — from purchase order issuance through ASN publication — presents a tangible opportunity to apply rule-based automation to this decision point, standardizing package type assignment and eliminating the cost inefficiencies that manual selection introduces [3]. Recent operational evidence supports this: an RPA model developed for order handling in supply chain management deployed software bots to capture and verify order data, reducing manual workload and the risk of human error in precisely the kind of repetitive, entry-dependent step that package type selection represents [9].

Before delving into the technological enablers of package selection automation, a key distinction must be clarified: the difference between a machine-learning-driven approach to package selection and a rule-based systematic approach. In the machine-learning paradigm, a supervised computer vision model would operate at the point where a supplier uploads a shipment image during Advance Shipping Notice (ASN) publication — using Optical Character Recognition (OCR) to identify the materials on the pallet and a trained classifier to determine whether the selected package type is accurate. By contrast, the rule-based approach leverages existing structured data fields to validate the selection, answering a narrower but immediately actionable question: given the dimensions and attributes entered, is this package type a logical fit? Recent reviews of the field position rule-based automation and machine learning as successive stages of supply chain intelligent automation rather than as competing alternatives [16].

Within the logistics industry, companies typically operate on thin margins, which leaves little room for error and makes capturing any opportunity for cost avoidance — if not direct cost saving all the more important [3]. In this context, the package type selection made by the material issuer has a direct and recurring cost impact, because freight charges are based on volumetric — or dimensional — weight, which is calculated from the declared package dimensions. This weight is neither the same across package

types nor fixed over time [4]. The problem is made worse by a mismatch in incentives. In the operating model of a large oil and gas organization served by thousands of supplier selections, the freight cost is paid by the organization and settled to the contracted 3PL carrier under the terms of the agreement. As a result, when a wrong selection leads to a lower charge, the carrier has no reason to flag or correct it — and when it leads to a higher charge, the extra cost falls on the organization, not on the party that made the selection. This gap between who makes the decision and who bears the cost is exactly what allows package selection errors to go unnoticed and uncorrected at scale.

2.3 Research Gap

Established material classification frameworks such as UNSPSC and eCl@ss define tens of thousands of product categories and provide standardized property definitions for describing individual items [5]. In materials-intensive industries they underpin spare-parts classification and cataloguing, supporting consistent identification of materials across systems and the operational continuity and asset reliability that depend on it. These frameworks, however, are concerned with classifying what a material is, its identity, function, and technical specification, and not with how that material should be packaged for transport.

Pivoting to the inbound logistics of these materials, package type selection as a distinct decision remains largely unaddressed in the literature. While adjacent topics — spare-parts classification, ERP implementation, and digital asset transformation — are well represented, coverage of package type selection specific to oil and gas inbound logistics is sparse to absent [6].

This paper has reviewed the available literature on material packaging classification within the logistics industry alongside the related material classification frameworks. It is evident that no globally recognized standard exists for materials packaging that links directly to the established international material classification systems. This absence — of both a governing standard and a structured body of research on automating the package selection decision — defines the gap this paper addresses, positioning the proposed process automation framework as a first structured contribution to closing it.

3. Current Practice and Cost Impact

The cost attached to every delivered package depends on a single size-based value. The current mechanism module transforms each material's length and width into a Rate Pallet Equivalent (RPE) value, applying the same length-by-width calculation uniformly across every material regardless of its wider technical specifications — including height and weight (the formula introduced in Section 1). In the operating setup of the oil and gas organization, the full span of RPE values is divided into ranges, each referenced by a package indicator, and the delivery cost linked to each indicator is fixed under the contractual agreement between the organization and the 3PL carrier.

The package type selection from the source drives this calculation of the package delivery cost. When identifying a material for the 3PL carrier to collect and deliver, the issuer must choose a package type from a long selection list; each package-type grouping maps to a package indicator, which in turn determines the applicable RPE range and its agreed cost. The accuracy of the entire cost chain therefore depends on this one manual choice — which is exposed to error for the reasons set out in Section 2.1, and is compounded when several materials are consolidated onto a single pallet, forcing the issuer to commit to one package type for a mixed load.

The literature on manual processes indicates that no amount of control can make a human-dependent selection consistent, because every selection ultimately rests on human judgment. Human reliability analysis of complex manual assembly identifies selection errors as a distinct and recurring error mode, arising wherever an operator must choose the correct item from among similar alternatives, with the probability of incorrect installation reaching 0.62 for the most error-prone part examined [7]. The scale of that exposure is quantifiable at volume: across 139,467 classification decisions made by 86 trained reviewers, the observed error rate was 10.76 percent, roughly one decision in nine, even though each decision was a binary judgment far simpler than package type selection [8].

These inaccurate selections are not captured during delivery, and there is no evidence-based method to validate them afterwards either: when administrators begin processing the invoices for the 3PL-delivered packages, they work from system records alone, with no images or technical specifications validated systematically.

In a single year, the manual verification effort described in Section 1 checked less than 5% of all packages and corrected some 2,500 wrong selections, roughly 0.25% of annual volume. If the other 95% of packages carry errors at the same rate, close to 50,000 wrong selections go uncaught each year, each one feeding an undetected cost error into the 3PL invoice.

4. Proposed Automated Package Type Selection Framework

The proposed framework ensures that the correct package type is selected for each material, so that the records reaching the 3PL invoice are accurate. Additionally, correct selection carries an operational benefit: an appropriately selected package reduces the likelihood of material damage in transit, which is costly beyond the materials themselves, since damaged materials trigger a lengthy claims process and delay materials needed for oil and gas production, in the worst-case exposing operations to the risk of a shutdown. The selection decision has been approached elsewhere as a machine-learning problem, using product attributes to assign each item the package type that best balances shipment cost against the risk of damage [4]. Using historical delivery records, supervised models have been trained to predict a suitable package size

for each new order and present that recommendation to the issuer in advance [10], while random-forest models have predicted an optimal packaging configuration from product features such as weight and volume, with reported reductions in packaging cost [11]. A reinforcement-learning framework for box sizing has since been adopted across 28 fulfilment centers of a large e-commerce platform, with estimated savings of USD 7.1 million annually [14]. Together, these studies establish package selection as a learnable decision that can be inferred from the attributes already recorded for each item though they rely on assembling large, labelled training datasets.

The framework proposed here targets this decision directly. Rather than depend on the labelled training data those approaches require, it begins with a direct, rule-based validation that is deployable immediately and is designed to evolve across three phases of progressively greater autonomy, as illustrated in Figure 1.

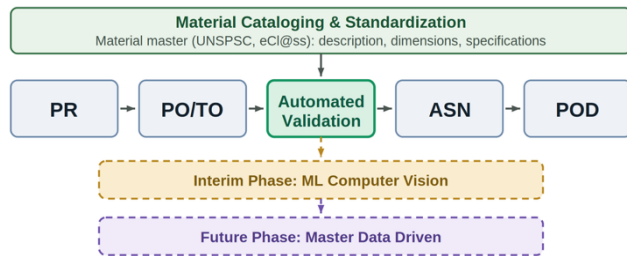


Figure 1. The automated package validation framework and its phased evolution

4.1 Current Phase

The current package selection phase relies on the data entered by the issuer. When a material is requested, the framework leverages its recorded attributes to determine the appropriate package indicator: the material description, dimensions, and technical specifications carried on the material record. Against these, it applies a set of rule-based checks that map a material's attributes to the appropriate package type, drawing on the reference tables that link each package grouping to a package indicator.

The package indicator determines the Rate Pallet Equivalent (RPE), computed from the material's length and width against the standard pallet footprint. A validated indicator therefore yields the correct RPE, and with it the correct delivery cost, before the Advance Shipping Notice is published.

4.2 Interim Phase

In the Interim phase, the issuer uploads an image of the prepared package as the Advance Shipping Notice is created, and a machine-learning computer-vision model validates the selection against the physical package. At the point of request, the system displays the material's cataloging status: where a requested material has no image or incomplete details on record, the issuer is prompted to attach them, and the upload becomes part of the material's record for subsequent shipments. Trained classifiers

applied to parcel images have been shown to identify package characteristics and detect discrepancies at high accuracy [12], [13]. This adds a layer the rule-based check cannot provide: rather than trusting the entered fields, it verifies the selection against what has physically been packed, catching cases where the recorded data itself is inaccurate or the assembled package does not match its declaration. It reduces the manual effort required of the issuer while ensuring that both the package type and the resulting cost are accurate.

4.3 Future Phase

In the future phase, the issuer enters only the material number. The system retrieves the material's description, dimensions, technical specifications, and reference images directly from the master data and determines the appropriate package type autonomously, without manual data entry, image upload, or selection.

The framework makes this state visible to the issuer: at the point of request, the system displays the material's cataloging status. Where a material is fully cataloged, the material number alone suffices and no further input is requested. Where a material is requested for the first time, or its master record lacks dimensions, specifications, or images, the issuer is shown that the material is not yet cataloged and is prompted to supply the missing details. That one-time input enriches the master record, so that subsequent orders of the same material qualify for fully autonomous handling. The framework's coverage therefore expands material by material, with each first-time request completing the record for every request that follows reducing the selection to a single input, the material number, and resulting in an accurate package type selection and a correct delivery cost reflected in the 3PL invoice.

5. Implementation Considerations

Deploying the framework in any materials-intensive oil and gas organization begins with full utilization of the material master, whose records provide both its foundation and the historical package-type associations that, once validated against each material's recorded dimensions and specifications, form the rule-based checks.

Enforcing mandatory data entry: The validation is only as reliable as the data it checks against: incomplete or inaccurate inputs cannot yield correct outputs. Studies of logistics operations confirm this at the operational level, showing that inaccurate or incomplete product master data measurably degrades process performance, with data errors translating directly into delays and downstream inefficiencies [15]. The dimension and specification fields that already exist must therefore be enforced as mandatory at the point of package selection. As noted in Section 3, these fields are presently available but not consistently completed, which is a primary reason the current process cannot self-correct; making them compulsory closes that gap and gives the rule-based checks a complete set of attributes to validate against for every material.

Package type unification: Package-type definitions and specifications vary across the oil and gas organization's systems, a variance rooted in the differing nature of operations, production output and feedstock on both the inbound and outbound fronts. Section 3 identified this variance as a structural source of inconsistent selection. Consolidating the definitions into a single, governed reference set ensures that the same material maps to the same package type and indicator regardless of the system or issuer involved a precondition for the rule-based checks to behave consistently at scale.

Cost calculation automation: The Rate Pallet Equivalent must be computed automatically within the charge-settlement flow rather than left as a manual step. Once the package indicator is validated, the system derives the corresponding RPE from the recorded dimensions and applies the contractually agreed rate, so that the figure reaching the invoice is generated from validated data rather than re-entered manually. This removes the exposure to additional error in the cost calculation after the delivery.

6. Conclusion

This paper showed that package type selection, a manual action, directly determines the invoiced delivery cost: predefined fields map each selection to a package indicator, and each indicator carries a contractually agreed carrier rate. In an inbound oil and gas logistics operation handling more than one million packages a year, manual verification was able to check less than 5% of total volume, leaving the accuracy of the remaining selections unvalidated. The first contribution of this paper is to establish this decision as a discrete source of cost leakage that recurs at scale, governed by no material classification standard and largely absent from the logistics literature. Its second contribution is a process automation framework that leverages the material master as its core data source, beginning with a rule-based validation, deployable within existing processes today and without labelled training data, that validates each selection against the recorded dimensions and automates the cost calculation in the back end. Throughout the upcoming phases, machine learning enriches this core source, as image-based validation both verifies each selection against the physical package and progressively completes the master record, enabling the future phase in which the material number alone, with quantity taken into account, is sufficient to determine the appropriate package type. The goal is inbound deliveries that are meticulously planned and flawlessly executed, giving the carrier the accurate, advance package information needed to plan trucks, forklifts, handling crews, and loading and offloading resources with far greater efficiency.

References

- [1] Babai, M. Z., Arampatzis, M., Hasni, M., Lolli, F., and Tsadiras, A., "On the use of machine learning in supply chain management: a systematic review", *IMA Journal of Management Mathematics*, Vol. 36, No. 1, pp. 21-49, 2025. <https://doi.org/10.1093/imaman/dpae029>
- [2] Flechsig, C., Anslinger, F., and Lasch, R., "Robotic Process Automation in purchasing and supply management: A multiple case study on potentials, barriers, and implementation", *Journal of Purchasing and Supply Management*, Vol. 28, No. 1, Art. 100718, 2022. <https://doi.org/10.1016/j.pursup.2021.100718>
- [3] Viale, L., and Zouari, D., "Impact of digitalization on procurement: the case of robotic process automation", *Supply Chain Forum: An International Journal*, Vol. 21, No. 3, pp. 185-195, 2020. <https://doi.org/10.1080/16258312.2020.1776089>
- [4] Gurumoorthy, K. S., Sanyal, S., and Chaoji, V., "Think Out of the Package: Recommending Package Types for E-Commerce Shipments", *Machine Learning and Knowledge Discovery in Databases: Applied Data Science and Demo Track (ECML PKDD 2020)*, Lecture Notes in Computer Science Vol. 12461, Springer, Cham, pp. 290-305, 2021. https://doi.org/10.1007/978-3-030-67670-4_18
- [5] Stolz, A., and Hepp, M., "Integrating Product Classification Standards into Schema.org: eCI@ss and UNSPSC on the Web of Data", *On the Move to Meaningful Internet Systems: OTM 2017 Workshops*, Lecture Notes in Computer Science Vol. 10697, Springer, Cham, pp. 103-113, 2018. https://doi.org/10.1007/978-3-319-73805-5_11
- [6] Khan, N., Solvang, W. D., and Yu, H., "Industrial Internet of Things (IIoT) and Other Industry 4.0 Technologies in Spare Parts Warehousing in the Oil and Gas Industry: A Systematic Literature Review", *Logistics*, Vol. 8, No. 1, Art. 16, 2024. <https://doi.org/10.3390/logistics8010016>
- [7] Torres, Y., Nadeau, S., and Landau, K., "Classification and quantification of human error in manufacturing: A case study in complex manual assembly", *Applied Sciences*, Vol. 11, No. 2, Art. 749, 2021. <https://doi.org/10.3390/app11020749>
- [8] Wang, Z., Nayfeh, T., Tetzlaff, J., O'Brien, P., and Murad, M. H., "Error rates of human reviewers during abstract screening in systematic reviews", *PLoS ONE*, Vol. 15, No. 1, Art. e0227742, 2020. <https://doi.org/10.1371/journal.pone.0227742>
- [9] Shamsuzzoha, A., and Pelkonen, S., "A robotic process automation model for order-handling optimization in supply chain management", *Supply Chain Analytics*, Vol. 9, Art. 100102, 2025. <https://doi.org/10.1016/j.sca.2025.100102>
- [10] Heininger, M., and Ortner, R., "Predicting Packaging Sizes Using Machine Learning", *Operations Research Forum*, Vol. 3, No. 3, pp. 1-14, 2022. <https://doi.org/10.1007/s43069-022-00157-5>
- [11] Zghair, H., and Konathala, R. G., "Using AI for Optimizing Packing Design and Reducing Cost in E-Commerce", *AI*, Vol. 6, No. 7, Art. 146, 2025. <https://doi.org/10.3390/ai6070146>
- [12] Lu, Z., Dai, N., Hu, X., Xu, K., and Yuan, Y., "Research on high-speed classification and location algorithm for logistics parcels based on a monocular camera", *Scientific Reports*, Vol. 14, Art. 15901, 2024. <https://doi.org/10.1038/s41598-024-66941-x>

- [13] Chaudhary, K., and Singh, A., "Parcel Damage Classification using Computer Vision: A Deep Learning Approach for Shipment Quality Assessment", 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), IEEE, Gwalior, India, pp. 1-6, 2024. <https://doi.org/10.1109/IATMSI60426.2024.10503363>
- [14] Kandula, S., Roy, D., and Akartunalı, K., "A machine learning approach to solve the e-commerce box-sizing problem", *Production and Operations Management*, Vol. 34, No. 6, pp. 1326-1345, 2025. <https://doi.org/10.1177/10591478241282249>
- [15] Božić, D., Živičnjak, M., Stanković, R., and Ignjatić, A., "Impact of the Product Master Data Quality on the Logistics Process Performance", *Logistics*, Vol. 8, No. 2, Art. 43, 2024. <https://doi.org/10.3390/logistics8020043>
- [16] Shamsuddoha, M., Khan, E. A., Chowdhury, M. M. H., and Nasir, T., "Revolutionizing supply chains: Unleashing the power of AI-driven intelligent automation and real-time information flow", *Information*, Vol. 16, No. 1, Art. 26, 2025. <https://doi.org/10.3390/info16010026>