FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Inventory Management in Project-Based Installation Companies: A Data-Driven Approach

Tomás Ferreira Nunes



Master in Mechanical Engineering

Supervisor: Prof. José Joaquim da Mota Machado

September 11, 2026

Inventory Management in Project-Based Installation Companies: A Data-Driven Approach

Tomás Ferreira Nunes

Master in Mechanical Engineering

September 11, 2026

Resumo

A gestão de inventário em empresas de instalações que trabalham por projeto enfrenta um desafio que os modelos clássicos não resolvem. A procura de material não é regular, ocorrendo em picos concentrados à medida que cada obra atinge a fase que a consome. Vista a partir de um armazém central que serve várias obras em simultâneo, é intermitente e irregular, um regime inadequado tanto para os modelos clássicos como para o planeamento por lista de materiais, e que os métodos de procura intermitente tratam sem explorar a sua causa conhecida. Na empresa em estudo, a PEMEL Instalações, traduz-se na ausência de uma política de stock, em ruturas recorrentes que param frentes de obra e num processo de compras reativo, com uma raiz comum, a gestão de um fluxo de material por projeto com ferramentas concebidas para operação contínua.

A dissertação propõe um modelo dinâmico de dimensionamento de inventário que explora precisamente essa causa, o portefólio de obras ativas. A partir das quantidades do orçamento de cada obra e de perfis de consumo calibrados no histórico, distribui o consumo de cada família ao longo da obra e agrega-o num stock mínimo dinâmico, que se ajusta ao portefólio e é protegido contra a variabilidade do prazo de entrega. Inclui uma classificação dos materiais por valor e padrão de procura, que delimita o seu âmbito, e uma camada não intrusiva de captura de dados por código QR e um painel de monitorização sobre o ERP existente. Segue uma metodologia de Design Science, validada retrospetivamente com dados reais.

Confirmou-se que a procura é quase universalmente intermitente ou irregular, e que os perfis de consumo são reprodutíveis num grau útil para as famílias significativas ligadas a etapas construtivas definidas. Num caso de estudo leave-one-out sobre duas dessas famílias, o modelo previu o consumo faseado muito melhor do que os métodos de Croston e de Syntetos-Boylan, reduzindo o erro entre quarenta e sessenta e cinco por cento, e uma simulação de Monte Carlo da política resultante reduziu as ruturas de cerca de uma em quatro para uma em vinte e cinco.

Os resultados são uma demonstração de viabilidade, não uma prova geral. As principais limitações são a validação ser retrospetiva e assentar em treze obras de uma especialidade, a estabilidade dos perfis ser apenas parcial e o benefício económico não ser quantificado por falta de dados de custo na empresa. A contribuição é, assim, uma alternativa baseada no conhecimento à extrapolação estatística, adequada a procura intermitente com dados escassos mas estrutura causal conhecida, e delimitada aos materiais a que se dirige.

Palavras-chave: gestão de inventário • procura intermitente • operações por projeto • portefólio de obras como motor da procura • faseamento da procura • stock de segurança dinâmico

Abstract

Inventory management in project-based installation companies faces a challenge the classical models do not resolve. Material demand is not regular but occurs in concentrated bursts, as each project reaches the phase that consumes a given material. Served from a central warehouse across several simultaneous projects, it is intermittent and irregular, a regime for which neither the classical inventory models nor bill-of-materials planning is suited, and which the methods dedicated to intermittent demand address without exploiting its known cause. In the case company, PEMEL Instalações, it manifests as the absence of any deliberate stock policy, recurrent stockouts that stall work fronts, and a reactive procurement process, deficiencies that share a single root, the management of a project-based material flow with tools conceived for continuous operation.

This dissertation proposes a dynamic inventory-sizing model that exploits precisely that cause, the active project portfolio. From the quantities in each project's budget and consumption profiles calibrated from historical records, it distributes each material family's consumption across the project lifecycle and aggregates the result into a dynamic minimum stock that adjusts to the portfolio and is buffered against the observed supplier lead-time variability. It is accompanied by a classification of materials by value and demand pattern, which delimits its scope, and by a non-intrusive layer of QR-code data capture and a monitoring dashboard operating over a closed ERP without modifying it. The work follows a Design Science methodology, validated retrospectively against real historical data.

Demand was confirmed to be almost universally intermittent or lumpy, and the consumption profiles reproducible to a useful degree for the economically significant families tied to well-defined constructive stages. On a leave-one-out case study of two such families, the model forecast the phased consumption substantially better than the Croston and Syntetos-Boylan methods, reducing the error by forty to sixty-five per cent, with a scaled error showing those methods to be no better than a naive mean, as they lack the causal information the model uses. A Monte Carlo simulation of the resulting policy reduced stockouts from roughly one requirement in four to one in twenty-five, at the cost of a modest amount of anticipated stock.

These results are a demonstration of feasibility, not a general proof. Their main limitations are that the validation is retrospective and rests on thirteen projects in a single speciality, that the profile stability is real but partial, and that the economic benefit, though favourable, is not quantified because the company does not record the necessary cost data. The contribution is thus a knowledge-driven alternative to statistical extrapolation, suited to intermittent demand with scarce data but known causal structure, and delimited to the materials it targets.

Keywords: inventory management • intermittent demand • project-based operations • project portfolio as demand driver • demand phasing • dynamic safety stock

The United Nations Sustainable Development Goals

The work connects to three United Nations Sustainable Development Goals, shown in Figure 1, in ways that follow directly from its subject, the more efficient management of material in a project-based installation company. The alignment is stated without overreach, since the contribution is a method and a prototype rather than a deployed intervention with measured outcomes.

The most direct link is to **Goal 12, responsible consumption and production**. The model sizes material to what each project requires and times its acquisition to when it is needed, rather than ordering in excess, which helps avoid the surplus that becomes waste or obsolete stock, a risk underlined by the large share of single-use items in the catalogue.

The work also aligns with **Goal 9, industry, innovation and infrastructure**, by applying a data-driven method and a lightweight digital capture layer to a traditional sector whose planning still rests on manual judgement and an under-exploited enterprise system.

A more indirect link exists to **Goal 8, decent work and economic growth**, through the operational efficiency the approach supports, since reducing the stockouts that stall work fronts contributes to a more productive use of the firm's resources, though this link is weaker as the work does not measure economic outcomes directly.



Figure 1: Sustainable Development Goals to which the work aligns, Goal 8 (decent work and economic growth), Goal 9 (industry, innovation and infrastructure), and Goal 12 (responsible consumption and production).

Acknowledgements

To my supervisor, José Machado, for his motivation, guidance, and for always pushing me to do more. Thank you for constantly challenging me to become better.

Thank you to FEUP, for being an institution that gives so much to its students and for providing me with five years of learning, challenges, and experiences that I will carry with me forever.

Thank you to PEMEL Instalações for welcoming me and giving me the opportunity to grow and flourish professionally. For more than a year now, I have been able to call it a second home, where I have had the privilege of learning, growing, and meeting people who I will carry with me into the future.

Thank you to my family. To my parents, Adriano and Maria, my siblings, Alexandre, Adriana, and Rute, my nephews, Lara, Miguel, Rodrigo, and Gonçalo, and my girlfriend, Maria. *Tudo o que sou é graças a vocês e por vocês.*

Thank you to the second family I found along this journey. To my friends Rafael, André, João, Rodrigo, Miguel, and Zé, none of this journey would have had the same meaning without you.

Finally, thank you to all the professors, colleagues, and students I have crossed paths with throughout these five years. Every person and every experience contributed, in one way or another, to the path that brought me here.

A todos, o meu mais sincero obrigado.

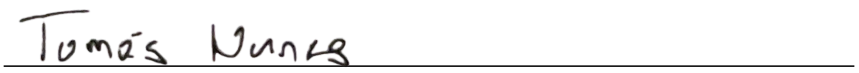
Tomás Ferreira Nunes

Declaration of honour

I declare that this dissertation is my own work and has not previously been submitted or assessed at this or any other institution. References to other authors (statements, ideas, thoughts), as well as the use of published works, strictly comply with the rules of attribution and are duly identified in the text and in the bibliographic references, in accordance with the applicable referencing standards. I am aware that the practice of plagiarism or self-plagiarism constitutes an academic offence, as established in the U.Porto Code of Ethics.

I also declare that artificial intelligence tools were used to support the drafting and revision of text, the development of data-processing scripts applied to data extracted and validated by me, and methodological reflection. All decisions, analyses, interpretations, and conclusions are my own, and all AI-assisted outputs were reviewed against the source data and corrected where necessary. I retain full responsibility for the content and integrity of this work.

Tomás Ferreira Nunes



"Success is not the key to happiness. Happiness is the key to success. If you love what you are doing, you will be successful."

Albert Schweitzer

Contents

List of Symbols xxv

1 Introduction 1
- 1.1 Context and Motivation . . . 1
- 1.2 PEMEL Instalações, S.A. . . . 2
- 1.3 Problem Statement . . . 3
- 1.4 Objectives . . . 3
- 1.5 Research Questions and Hypotheses . . . 4
- 1.6 Methodology . . . 5
- 1.7 Dissertation Structure . . . 6

2 Literature Review 9
- 2.1 Digitalization of Industrial and Organizational Processes . . . 9
- 2.2 Recent Advances and the Place of Machine Learning . . . 11
  - 2.2.1 The Rise of Data-Driven Forecasting . . . 11
  - 2.2.2 Machine Learning under Intermittent Demand and Data Scarcity . . . 11
  - 2.2.3 Positioning of the Present Work . . . 12
- 2.3 Inventory Management Fundamentals . . . 13
  - 2.3.1 Classical Reorder-Point Models for Independent Demand . . . 14
  - 2.3.2 Material Requirements Planning for Dependent Demand . . . 16
  - 2.3.3 The Assumptions Underlying the Classical Models . . . 16
- 2.4 Forecasting Demand in Intermittent and Project-Driven Contexts . . . 17
  - 2.4.1 Characterizing Intermittent and Lumpy Demand . . . 18
  - 2.4.2 Forecasting Methods for Intermittent Demand . . . 19
  - 2.4.3 The Limitation of Time-Series Methods and the Opening for a Causal Approach . . . 20
- 2.5 Inventory Management in Project-Based Environments . . . 21
  - 2.5.1 The Project-Based Production Context . . . 21
  - 2.5.2 Estimating Project Quantities . . . 21
  - 2.5.3 Temporal Distribution of Demand and the Construction S-Curve . . . 22
  - 2.5.4 Research Gap Analysis . . . 23
- 2.6 Enabling Technologies: Data Capture, Warehouse Systems, and Decision Support 23
  - 2.6.1 Data Capture and Traceability . . . 24
  - 2.6.2 Warehouse Management Systems and ERP Integration . . . 24
  - 2.6.3 Decision Support Systems . . . 25

3 Company Context and Current System **29**
3.1 Material Flow Mapping . . . 29
3.2 Warehouse Operation Description . . . 31
3.3 Order Tracking . . . 31
3.4 Operational Bottlenecks and Impact Analysis . . . 32
3.4.1 Material Shortages and Idle Labour . . . 32
3.4.2 Constraints in Reactive Order Tracking . . . 33
3.4.3 Compromised Agility and Client Relations . . . 33
3.4.4 Absence of a Performance Measurement System . . . 33
3.5 Baseline Performance . . . 34
3.6 From Diagnosis to Solution . . . 36

4 A Dynamic Phase-Based Model for Inventory Sizing in Project-Based Environments **39**
4.1 Overview of the Proposed Model . . . 39
4.2 Material Classification by Value and Demand Pattern . . . 41
4.2.1 The Value Dimension: ABC Analysis . . . 41
4.2.2 The Demand-Pattern Dimension: ADI-CV² Classification . . . 42
4.2.3 The Intersection: Scope of the Model . . . 42
4.3 Project Quantity Requirements . . . 43
4.4 Temporal Phasing of Demand . . . 44
4.4.1 Positioning Consumption on the Project Timeline . . . 44
4.4.2 Constructing the Phase Weights . . . 45
4.4.3 The Stability Premise . . . 45
4.5 Dynamic Minimum Stock Aggregation . . . 46
4.5.1 Expected Consumption over a Coverage Horizon . . . 46
4.5.2 Aggregation across the Active Portfolio . . . 47
4.5.3 The Safety Component . . . 47
4.6 Enabling Implementation Layer . . . 49
4.6.1 Order-Driven Data Capture and ERP Integration . . . 49
4.6.2 The Monitoring Dashboard as Decision-Support Layer . . . 53

5 Empirical Validation and Results **55**
5.1 Validation Methodology . . . 55
5.1.1 Data Sources and Extraction . . . 55
5.1.2 Reconstructing Family-Level Consumption . . . 56
5.1.3 Scope and Standard of the Validation . . . 59
5.2 The Value Dimension: ABC Classification Across Projects . . . 59
5.3 The Demand Regime: Testing the Lumpiness Premise (H1) . . . 62
5.4 The Stability of Phase Profiles (H2) . . . 64
5.4.1 Sensitivity to the Number of Phases . . . 68
5.5 Supplier Reliability and Lead-Time Variability . . . 69
5.6 Case Study: Forecasting Performance (H3) . . . 73
5.7 Case Study (Continued): The Inventory Trade-off (H4) . . . 76
5.7.1 Economic Estimate of the Benefit . . . 79

6 Conclusions and Future Work **83**
6.1 Conclusions . . . 83
6.2 Future Work . . . 84

References **87**

A Baseline Questionnaire **93**

B Prototype Application: QR-Based Reception Workflow **95**
B.1 Application Structure 95
B.2 The Reception Workflow 96

C Monitoring Dashboard **99**
C.1 Dashboard Pages 99
C.2 Data Model and Indicators 101

D Data-Processing Scripts **103**
D.1 ABC Classification Across Projects 103
D.2 Phase Profiles and Stability Test 107

List of Figures

1 Sustainable Development Goals to which the work aligns, Goal 8 (decent work and economic growth), Goal 9 (industry, innovation and infrastructure), and Goal 12 (responsible consumption and production). v

1.1 PEMEL Instalações logo. 2

1.2 Timeline of the dissertation, from February to September. 6

2.1 Conceptual representation of the progressive stages of digital integration, from digitization to digital transformation. 9

2.2 Qualitative positioning of forecasting and planning approaches according to the volume of historical data they require and the extent to which they exploit the causal structure of the demand. The proposed model occupies the regime of low data requirement and explicit use of the causal structure. 12

2.3 An idealized representation of how inventory levels evolve over time at a given location. Adapted from Bartholdi et al. [4]. 15

2.4 Economic Order Quantity model illustrating the reorder point, lead time, and replenishment cycle. Adapted from Jadidi et al. [17]. 15

2.5 Demand classification scheme of Syntetos, Boylan and Croston [42], defined by the Average Demand Interval (ADI) and the squared coefficient of variation (CV^2) of non-zero demand. The thresholds $ADI = 1.32$ and $CV^2 = 0.49$ delimit the smooth, erratic, intermittent, and lumpy categories. 19

2.6 Generic S-Curve representing cumulative project progress over time. Adapted from Kerzner [19]. 22

2.7 Simplified functional architecture of warehouse operations supported by a Warehouse Management System, illustrating the interaction between operational processes, centralized inven tory databases, and Enterprise Resource Planning platforms. 25

2.8 Conceptual pipeline of an end-to-end analytics process illustrating the progression from raw data sourcing to strategic decision-making. Adapted from Champa and Segall [6]. 26

3.1 Flowchart of the current materials management and logistics process, from the Internal Site Requisition (RIO) through Purchase Order (ECF), Delivery Note (GUIA), and execution. 30

4.1 Architecture of the proposed dynamic phase-based model. 40

4.2 Illustration of the dynamic minimum stock of a family over time. Projects entering at different moments would cause variations in the aggregate stock requirement (solid line) is the superposition of their phase-dependent contributions, rising when active projects coincide in a consuming phase and falling as they pass it. The temporal positioning of the projects is illustrative of the mechanism. 48
4.3 Order-driven QR reception workflow. The operator selects the ECF identified from the delivery note, scans each article against the expected order lines, associates any unrecognized code with its Enterprise Resource Planning (ERP) item, and saves a reconciled summary of ordered versus received quantities. 51
4.4 Decoupled read-only architecture. PRIMAVERA is accessed only through a one-way synchronization routine that copies ECF and article data into the local application database; Quick Response (QR) associations and stock movements are held locally and are never written back to the production ERP. 52
4.5 Overview page, showing the order-status indicators, the highlighted delayed orders, the breakdown by speciality, and the pending-by-supplier and recent-receptions tables. 54

5.1 Family sheet of a project budget, listing the thirty-two HVAC functional families with their designations and reference brands. Commercially sensitive values and identifiers are blurred. This catalogue defines the functional families used as the unit of analysis. 57
5.2 Detail sheet of a project budget. Each priced article is set against a hierarchical position and carries, in the highlighted column, the number of the functional family to which it belongs. This is the source from which the family of each article is read for the price-based reconstruction. Commercially sensitive values are blurred. 58
5.3 Proportion of projects in which each HVAC family was classified A, B or C, across the thirteen projects, ordered from the most to the least economically significant. 60
5.4 ABC class of each functional family in each project. Class A in dark blue, Class B in green, Class C in grey, and blank where the family is absent from the project. Projects with a reconstructed budget are marked with an asterisk. 61
5.5 Syntetos–Boylan classification of the installation articles in the ADI-CV^2 plane (log scale on the ADI axis). Recurring articles ($\geq$ 12 months of demand) are shown in colour by category; lower-turnover articles are shown faded. The dashed lines mark the conventional cut-offs ADI $= 1.32$ and $CV^2 = 0.49$. 63
5.6 Bias test of the reconstructed budgets. The two groups overlap broadly and the reconstructed projects do not show the reduced deviation that a bias toward artificial regularity would produce. 66
5.7 Phase-profile instability of every family across the projects, with the mean profile Q1 to Q4, the instability index and the resulting class. Families are ordered from most to least stable. 66
5.8 Intersection of value and phase stability across the HVAC projects. Each family is placed by its economic importance, measured as the share of projects in which it is Class A or B, and by the instability of its phase profile, with more stable families toward the bottom. The economically significant installation families cluster in the low-to-moderate instability band (green colour), which is the region in which the model applies. 67
5.9 Distribution of the supply lead time for current stock material. 71
5.10 Median lead time of current material by month of order. 72

5.11 Leave-one-out validation for heating copper. For each project, the phase profile calibrated on the other five projects (green) is compared against the actual profile (blue), with the mean absolute error in each project. Reconstructed-budget projects are marked with an asterisk. 74
5.12 Per-project phase profiles of the two case-study families, with the average highlighted. 74
5.13 Monte Carlo distribution of the stockout rate for refrigeration copper over ten thousand runs, under the current stock-zero policy and under the dynamic policy, with means of 24% and 4 %. The two distributions are well separated, and the dashed lines mark the means. 77
5.14 Copper lead-time distribution with the two anticipation windows, seven days for current practice and thirty for the dynamic policy. 78

6.1 Implementation roadmap for the adoption of the approach, with the phases ordered by dependency. The colour indicates the status of each, the data-capture and dashboard prototypes and the retrospectively validated model developed in this work, and the phases proposed for future work. 85

B.1 The three main areas of the prototype application. 96
B.2 Step-by-step reception workflow, from order confirmation to reconciliation. The sequence mirrors the process diagram of Figure 4.3. 97

C.1 Orders page, detailing each purchase order with its completion rate and pending value. 100
C.2 Suppliers page, with the lead-time risk quadrant and the supporting per-supplier table. 101

List of Tables

2.1 Comparison of forecasting and planning approaches by information used. The proposed model is highlighted in bold. 23

3.1 Summary of questionnaire responses ($N = 8$ operational staff). Results are reported as absolute counts, reflecting the perceptions of a small group of key informants rather than a statistically representative sample. 35
3.2 Mapping of identified problems to proposed responses. 36

4.1 Inventory management policy by material class. 43

5.1 Distribution of articles across the Syntetos-Boylan demand categories, by level of turnover. Articles are grouped according to the number of months in which demand was recorded over the observation horizon. 63
5.2 Instability of the phase profile in percentage points of each Class B installation family under four, five and six phases. 69
5.3 Supplier lead-time statistics, in days, for all orders and for the current-material population that the model governs. 70
5.4 Forecast-error metrics for the model and the benchmarks, under leave-one-out, for the two Class B families: Heaintg copper and Refrigeration copper. 75
5.5 Assumptions underlying the economic estimate. 80
5.6 Scenario estimate of the annual, organization-wide cost of the current situation and the saving the model would achieve. Frequencies are varied across scenarios, while the labour and premium rates are held at the values stated above. 80
5.7 Proforma of the data required to measure the economic benefit rather than estimate it, with the source through which each would be captured. 81

List of Acronyms

ABC	Activity-Based Classification
ADI	Average Demand Interval
BOM	Bill of Materials
DL	Deep Learning
DSS	Decision Support System
ECF	Purchase Order
EOQ	Economic Order Quantity
ERP	Enterprise Resource Planning
GUIA	Delivery Note
ML	Machine Learning
MRP	Material Requirements Planning
QR	Quick Response
RFID	Radio-Frequency Identification
RIO	Internal Site Requisition
ROP	Reorder Point
SBA	Syntetos-Boylan Approximation
TSB	Teunter-Syntetos-Babai
WMS	Warehouse Management System

List of Symbols

Symbol	Description	Domain
Indices and sets		
i	Index of a Class B material	$i \in I$
p	Index of a project	$p \in P$
ϕ	Index of a project phase (four equal quarters, Q1–Q4)	$\phi \in \{1,2,3,4\}$
t	Time (continuous)	$t \in \mathbb{R}_{\geq 0}$
active(t)	Set of projects active at time t	active$(t) \subseteq P$
$\phi(p,t)$	Set of phases project p traverses within the coverage horizon at t	$\phi(p,t) \subseteq \{1,2,3,4\}$
Proposed model (Chapter 4)		
$Q_{i,p}$	Total quantity of material i foreseen for project p (quantity map)	$\mathbb{R}_{\geq 0}$
$q_{i,p,\phi}$	Quantity of material i consumed by project p in phase ϕ	$\mathbb{R}_{\geq 0}$
$W_{i,\phi}$	Phase weight: fraction of i's total consumed in phase ϕ ($\sum_{\phi} W_{i,\phi} = 1$)	$[0,1]$
$W_{i,\phi(p,t)}$	Cumulative phase weight over the coverage horizon	$[0,1]$
$C_{i,p,t}$	Expected consumption of i by project p over the horizon	$\mathbb{R}_{\geq 0}$
$D_{i,t}$	Aggregate project-driven requirement for i at time t	$\mathbb{R}_{\geq 0}$
$\bar{D}_i$	Expected demand rate of i over the lead time	$\mathbb{R}_{\geq 0}$
SS_i	Safety (residual) component of the stock for material i	$\mathbb{R}_{\geq 0}$
$DMS_{i,t}$	Dynamic minimum stock for material i at time t	$\mathbb{R}_{\geq 0}$
Z	Service-level (safety) factor	$\mathbb{R}_{\geq 0}$
σ_L	Standard deviation of the supplier lead time	$\mathbb{R}_{\geq 0}$
Classical inventory models (Chapter 2)		
ROP	Reorder point	$\mathbb{R}_{\geq 0}$
EOQ	Economic order quantity	$\mathbb{R}_{\geq 0}$
D	Average demand rate (annual demand in the EOQ)	$\mathbb{R}_{\geq 0}$
L	Supplier lead time	$\mathbb{R}_{\geq 0}$
SS	Safety stock (classical)	$\mathbb{R}_{\geq 0}$
σ_d	Standard deviation of demand	$\mathbb{R}_{\geq 0}$
S	Fixed ordering cost per order	$\mathbb{R}_{\geq 0}$
H	Holding cost per unit per period	$\mathbb{R}_{\geq 0}$

Chapter 1

Introduction

This chapter introduces the problem and the terms in which it is approached. It sets out the context and the case company, states the problem, that a project-based installation company lacks a method to size proactively an inventory whose demand is lumpy and project-driven, and derives from it the objectives, the research question and hypotheses, and the Design Science methodology that guide the work, closing with its contributions and structure.

1.1 Context and Motivation

In recent decades, the increasing complexity of industrial operations has led organizations to seek more efficient and reliable ways to manage materials and coordinate their flow across the supply chain. In this context, inventory management plays a decisive role in ensuring operational continuity and cost control. Insufficient stock leads to shortages and project delays, while excess stock immobilizes capital and increases the risk of obsolescence. The effective sizing of inventory, determining how much of each material to hold and when, is therefore a central problem of industrial operations management.

This problem is particularly acute in project-based sectors, such as industrial installations and construction-related activities. In these environments, materials are not consumed at a steady rate but in concentrated bursts, as each project reaches the execution phase that requires a given material. The resulting demand, observed at a central warehouse serving several projects, is intermittent and irregular, a pattern known in the inventory literature as lumpy demand. This demand regime is poorly served by the classical inventory models, which assume regular and statistically stable consumption, and equally poorly served by the bill-of-materials logic of material requirements planning, whose informational requirements are rarely met in time.

However, there exists a source of information that distinguishes project-based environments from spare-parts settings for which the dedicated intermittent-demand methods were developed.

Although irregular, the demand is not random, as it is a direct and observable consequence of the active project portfolio, whose composition, scale, and schedule are known to the organization. This dissertation is motivated by the proposition that this knowledge can be exploited to anticipate material demand causally, rather than extrapolating it statistically from an irregular history, and thereby to support a more proactive sizing of inventory than current practice allows.

1.2 PEMEL Instalações, S.A.

This dissertation is developed in collaboration with PEMEL Instalações, a Portuguese company operating in industrial installations, particularly HVAC, electrical, and hydraulic systems for commercial and industrial facilities. Part of the ONIS Group, it employs around 120 people and has an annual turnover in excess of ten million euros, placing it among the medium-sized enterprises of the sector. It executes multiple projects simultaneously, each requiring a diverse set of materials, components, and equipment, and operates a central warehouse from which materials are received from suppliers, temporarily stored, and distributed to the various worksites according to project needs.



Figure 1.1: PEMEL Instalações logo.

The company's activity spans the technical installations of medium and large building projects, across the residential, commercial, public, and institutional sectors, in both new construction and refurbishment. The projects analysed in this work illustrate that range, from residential developments and office towers to schools, hospitals, and transport infrastructure, each executed as an independent contract with its own budget, schedule, and material requirements. Internally, the company is organized into the functional departments through which a project passes, estimating, which prices the work and produces the budget from which this dissertation later recovers the material families, procurement, the warehouse, production, and after-sales, supported by human resources and by safety and quality. The inventory-sizing problem addressed here sits at the intersection of estimating, procurement, and the warehouse, whose coordination determines whether material is available when a work front needs it.

This environment exhibits, in a concrete industrial setting, precisely the characteristics that make project-based inventory management difficult, high material variety, demand driven by the

progress of individual projects, and the continuous movement of materials between the warehouse and active sites. The company manages these activities through an Enterprise Resource Planning (ERP) system, complemented by manual procedure, and, as is characteristic of small and medium enterprises in the sector, without a formal, proactive inventory policy. PEMEL therefore offers a representative case in which to study, and address, the inventory-sizing problem that motivates this work.

1.3 Problem Statement

The central problem addressed by this dissertation is the absence of a method, applicable in the conditions of a medium-sized project-based enterprise, for proactively sizing the inventory of installation materials whose demand is lumpy and project-driven. What distinguishes this problem from intermittent demand in general is that the demand here is not truly unpredictable, it has a known cause, the content and schedule of the active project portfolio, which the available methods do not exploit. The problem is therefore not only that the demand is difficult, but that its difficulty is addressed by treating as random a signal whose driver is in fact known.

Methodologically, as established in Chapter 2, neither the classical reorder-point models nor material requirements planning is suited to the demand regime in question, and the dedicated intermittent-demand methods, while appropriate to the regime, forecast from the past of the demand series alone and do not exploit the project portfolio that causes it. There is thus a genuine gap between the available methods and the problem at hand. Operationally, as the diagnosis in Chapter 3 shows, this gap manifests at PEMEL as the absence of any deliberate inventory policy, in which the warehouse holds little beyond the residual surplus of completed projects and replenishment is reactive, giving rise to material shortages on one hand and immobilized capital on the other.

A further condition constrains any solution. The reliable, proactive management of inventory presupposes reliable data on consumption and supply, yet the recording of material movements remains partly manual and procurement information is accessible to only a limited number of users. Any methodological solution must therefore be accompanied by the means to capture reliable data and to make its outputs accessible to the operational staff who act on them, a requirement that shapes the solution rather than constituting a separate problem.

1.4 Objectives

The main objective of this dissertation is to develop and validate a model for the dynamic sizing of installation-material inventory in project-based environments, exploiting the active project portfolio as the driver of demand, and to provide the enabling components required to operate it within

a small enterprise.

To achieve this objective, the dissertation pursues the following specific goals:

- Characterize, through a structured diagnosis of the case company, the material and information flows and the operational deficiencies of the current reactive inventory practice.
- Establish, from the inventory and intermittent-demand literature, the methodological gap that a project-driven approach is intended to fill, and position the proposed approach relative to the recent state of the art, including data-driven and machine-learning methods.
- Develop a model that classifies materials by value and demand pattern, and that, for the materials it targets, combines the known per-project quantities with their characteristic temporal consumption profiles to compute a dynamic minimum stock across the active project portfolio.
- Validate the model and its underlying premises retrospectively, against historical consumption data, comparing its stock decisions with both the current practice and established statistical benchmarks.
- Provide the enabling layer required for operation, namely a non-intrusive data-capture mechanism and a monitoring interface, and define an implementation roadmap for adoption within the organization.

1.5 Research Questions and Hypotheses

The central research question of this dissertation is whether the demand for installation materials in a project-based environment, which the classical and intermittent-demand methods treat as an autonomous and largely unpredictable signal, can be more effectively anticipated by exploiting its known causal driver. This question is decomposed into formal hypotheses, organized in two layers that mirror the structure of the validation in Chapter 5. The first layer concerns the premises on which the proposed model depends. The second concerns its performance relative to established alternatives.

The premises of the model are testable propositions in their own right:

H1 The Class B installation materials of interest exhibit lumpy demand, confirming that the problem lies outside the regime for which the classical models are suited.

H2 The temporal consumption profile of a material family, expressed as its distribution across project phases, is sufficiently stable across projects to have predictive value. The phase at which a family is consumed is therefore a characteristic of the family rather than an idiosyncrasy of each individual project.

Conditional on these premises, the performance of the model is addressed by:

H3 A demand estimate reconstructed from the active project portfolio and the calibrated phase profiles approximates actual consumption at least as accurately as the statistical benchmark methods (Croston and the Syntetos-Boylan Approximation) applied to the same series.

H4 A dynamic minimum stock dimensioned by the proposed model, drawing on the family-specific variability of the supplier lead time, would, applied retrospectively, achieve a more favourable trade-off between service level and held inventory than the company's current reactive practice, in which material is ordered against immediate need without a planned stock.

The layers are logically ordered so that the performance hypotheses (H3, H4) are meaningful only insofar as the premise hypotheses (H1, H2) are supported, since a model built on profiles that proved unstable could not be expected to perform. Given the limited number of completed projects available for calibration and testing, the hypotheses are examined as a demonstration of feasibility supported by quantitative evidence, rather than as a claim of statistical significance. The implications of this distinction for the strength of the conclusions are discussed in Chapter 5.

1.6 Methodology

This research follows a Design Science approach, in which the central outcome is the construction of an artefact, the inventory-sizing model, and its rigorous evaluation against the problem it is intended to solve. The work is structured in four stages.

First, the current warehouse and inventory management processes of the case company are analyzed through direct observation, informal interviews with operational stakeholders, and examination of the transactional records held in the ERP. This stage maps the material and information flows, identifies the operational deficiencies, and establishes a baseline characterization of current performance.

Second, a review of the academic literature on inventory management, intermittent-demand forecasting, and project-based operations is conducted, in order to establish the methodological gap and to position the proposed approach relative to both the classical methods and the recent data-driven state of the art.

Third, the model is developed. This comprises the classification of materials by value and demand pattern, the calibration of temporal consumption profiles from historical data, and their combination with project quantities to compute a dynamic stock requirement, together with the enabling data-capture and monitoring components.

Fourth, the model is validated retrospectively. Its premises, the lumpiness of the demand and the stability of the phase profiles, are examined against the historical consumption record, and its performance is demonstrated on a case study of representative Class B families, in which a leave-one-out procedure compares the model's forecast against the Croston and Syntetos-Boylan benchmarks, and a Monte Carlo simulation compares the resulting inventory policy against the current practice. Error and inventory metrics appropriate to intermittent demand are used throughout, and the limitations imposed by the available data are made explicit.

The distribution of these stages over the work, from February to September, is shown in Figure 1.2, where the data extraction and treatment and the validation appear as the most demanding stages and overlap substantially, reflecting the effort that the reconstruction of reliable data from a fragmented ERP entailed.

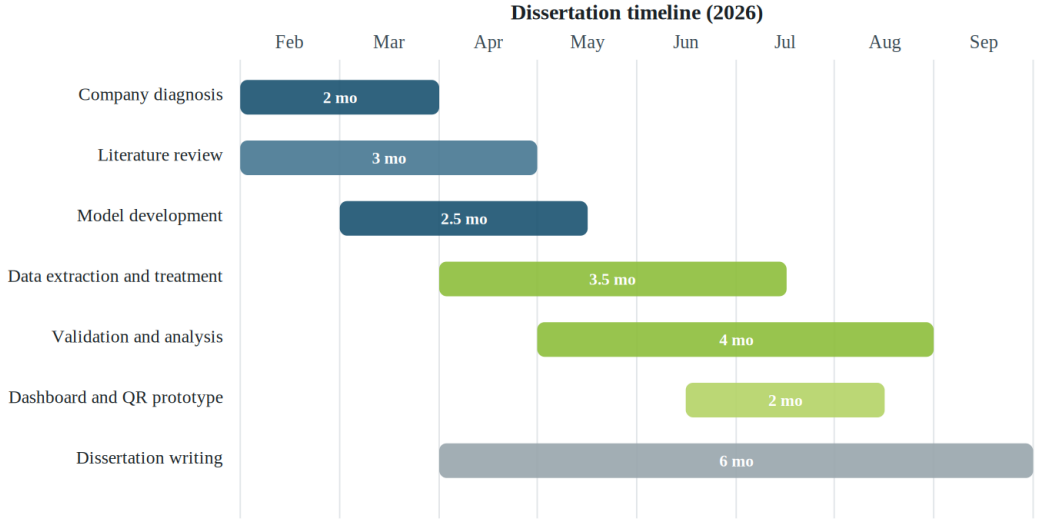


Figure 1.2: Timeline of the dissertation, from February to September.

1.7 Dissertation Structure

This dissertation is organized into six chapters. The present chapter has introduced the problem, its context at PEMEL, the objectives, and the research questions and hypotheses that guide the work.

Chapter 2 reviews the relevant literature, establishing the limitations of the classical and intermittent-demand inventory models for project-driven demand, positioning the work relative to recent data-driven and machine-learning approaches, and articulating the methodological gap that the proposed model addresses.

Chapter 3 presents the diagnosis of the case company, mapping its material and information flows, characterizing the deficiencies of the current reactive practice, and establishing the baseline against which improvement is assessed.

Chapter 4 develops the proposed model, from the classification of materials by value and demand pattern, through the temporal phasing of consumption, to the aggregation of a dynamic minimum stock across the active project portfolio, together with the enabling data-capture and monitoring layer.

Chapter 5 sets out the methodology for validating the model against historical data and reports the empirical results, from the examination of the premises to the case-study demonstration of the model's forecasting and inventory performance.

Chapter 6 draws the conclusions of the work, situates its contribution and limitations, and identifies directions for future work, including the prospective validation and the economic quantification that the present study leaves open.

Chapter 2

Literature Review

This chapter reviews the knowledge on which the work builds and establishes the gap it fills. It moves from the wider context of industrial digitalization and machine learning to the fundamentals of inventory management and the difficulty of forecasting intermittent, project-driven demand, and then to inventory management in project-based environments and the enabling technologies any solution requires. The review converges on the gap the model occupies, a regime of intermittent demand and scarce data whose causal structure is known but unexploited by existing methods.

2.1 Digitalization of Industrial and Organizational Processes

The contemporary industrial landscape is characterized by a systemic transition toward highly interconnected and data-driven operational models, in which the digitalization of organizational processes has shifted from a competitive advantage to an operational necessity [5]. For the purposes of this work, it is useful to distinguish three progressive concepts that are often conflated. *Digitization* refers to the narrow technical conversion of analog information into digital formats. *Digitalization* denotes the broader socio-technical use of digital technologies to optimize existing business processes, such as procurement, inventory management, and logistics workflows. *Digital transformation* represents the company-wide strategic renewal that leverages these digitalized processes to fundamentally alter value-creation paradigms [47, 48].



Figure 2.1: Conceptual representation of the progressive stages of digital integration, from digitization to digital transformation.

Positioning this work within that spectrum is important for scoping its contribution. The present dissertation does not pursue digital transformation as a broader strategic objective. Instead, it operates at the level of *digitalization*, exploiting the operational data that digitalized procurement and warehousing processes already generate in order to support inventory decision-making. In industrial service sectors responsible for complex HVAC, electrical, and hydraulic installations, this evolution is commonly framed within the paradigm of Industry 4.0. Frank et al. [11] distinguish between *base technologies*, such as the Internet of Things (IoT), cloud computing, and big data analytics, and *front-end technologies* that apply these capabilities to operational domains such as manufacturing and supply chain management. The base technologies provide the connectivity and information-processing capacity that, in the context of stock management, enables a transition from isolated and reactive inventory tracking toward integrated, data-driven information systems.

This transition is closely related to the concept of the Digital Supply Chain (DSC). Ivanov et al. [16] define the DSC as a highly integrated network that leverages digital technologies to improve decision-making, coordination, and operational agility across supply chain activities. Traditional supply chains frequently suffer from fragmented information flows and manual data entry, which can amplify inefficiencies such as the bullwhip effect. By digitalizing operational processes, organizations move from siloed structures toward systems in which information flows more reliably across departments.

Crucially for this work, this same digitalization produces a persistent transactional record, such as purchase orders, delivery notes, and stock movements, that constitutes the empirical basis on which data-driven inventory models can be built. Holloway [15] similarly notes that digital tools in supply chain operations, including integrated dashboards, ERP systems, and data-driven monitoring tools, can improve inventory visibility, reduce stockouts, and streamline material-management tasks. The reliability of such records depends in turn on the underlying data-capture infrastructure. Technologies such as barcode systems, Radio-Frequency Identification (RFID), and sensor-based tracking have been increasingly adopted to monitor material flows. Queiroz et al. [33] observe that, despite their infrastructural and organizational demands, these technologies improve the transparency and reliability of logistics processes. In project-based environments, where materials are continually allocated across multiple worksites, this traceability is a precondition for the accurate consumption records on which any forecasting or stock-sizing model ultimately relies.

The successful exploitation of these data, however, depends not only on technology but also on organizational capability. Warner and Wäger [50] emphasize the role of "dynamic capabilities", namely sensing opportunities for technological improvement, seizing them by integrating digital tools into core systems such as ERP platforms, and reconfiguring organizational routines accordingly, in enabling firms to adapt to digital environments. Where inventory control still relies on manual procedures or fragmented systems, the development of such capabilities is particularly relevant. Synthesizing these perspectives, digitalization within industrial organizations can be under-

stood as a multidimensional process combining technological adoption, organizational capability, and structural adaptation [7]. In the context of warehouse and logistics operations, it represents the enabling step that makes operational information accurate, accessible, and reliable enough to move from retrospective reporting toward the proactive, model-based inventory decision-making developed in the remainder of this work.

2.2 Recent Advances and the Place of Machine Learning

The methods reviewed in the preceding section, though foundational, are several decades old, and the last five years have seen intense research activity in demand forecasting and inventory management, driven largely by the availability of Machine Learning (ML) and Deep Learning (DL) techniques. A complete account of the field must therefore situate the present work relative to these developments, and explain the methodological choice that follows from that comparison.

2.2.1 The Rise of Data-Driven Forecasting

The application of ML and DL to demand forecasting has grown rapidly, with a recent critical review identifying a pronounced increase in publications between 2021 and 2024, during which the majority of the field's research output was produced [25]. The appeal of these methods lies in their capacity to learn complex, non-linear relationships and to incorporate exogenous variables, such as promotions, weather, or economic indicators, that classical time-series methods ignore [37]. Tree-based ensemble methods, in particular Random Forests and Gradient Boosting, together with deep architectures such as recurrent and attention-based networks, have become the dominant tools, and have been applied across retail, manufacturing, and, increasingly, construction-related distribution [31]. This body of work, often grouped under the headings of predictive analytics and demand sensing, represents the contemporary state of the art against which any new forecasting proposal must be positioned.

2.2.2 Machine Learning under Intermittent Demand and Data Scarcity

The relevance of these methods to the present problem is, however, qualified by two characteristics of the data available in the setting studied here. The first is the intermittent and lumpy nature of demand itself, while the second is the limited volume of historical observations available in a small enterprise executing a modest number of projects. Both conditions are precisely those under which data-hungry ML methods are known to struggle.

The recent literature is notably candid on this point. A dedicated review of ML algorithms for intermittent demand observes that, although such algorithms excel at uncovering complex patterns, sparse training datasets routinely undermine their performance, and sparsity is the defining feature of intermittent demand [45]. The same review reports that, in a controlled comparison on a widely used benchmark, Random Forests, despite their accuracy, were outperformed by the far simpler Croston method on inventory-specific metrics [45]. Comparative studies on lumpy series reach similar conclusions, finding that for the most intermittent categories the elementary baselines remain difficult to beat [21]. In short, the advantages of ML are neither automatic nor universal, but depend on data volume and demand regularity, both of which are limited in the present context.

2.2.3 Positioning of the Present Work

These observations inform the methodological stance adopted in this dissertation. The approach developed in the following chapters is deliberately not a data-hungry, purely statistical learner applied blindly to the demand series. Such an approach would be poorly matched to the available data and, on the evidence above, would not necessarily outperform the established intermittent-demand methods that already serve as benchmarks. Instead, the present work exploits a source of information that ML applied solely to demand series cannot access, namely the the structural driver of demand represented by the known content and schedule of the active project portfolio.

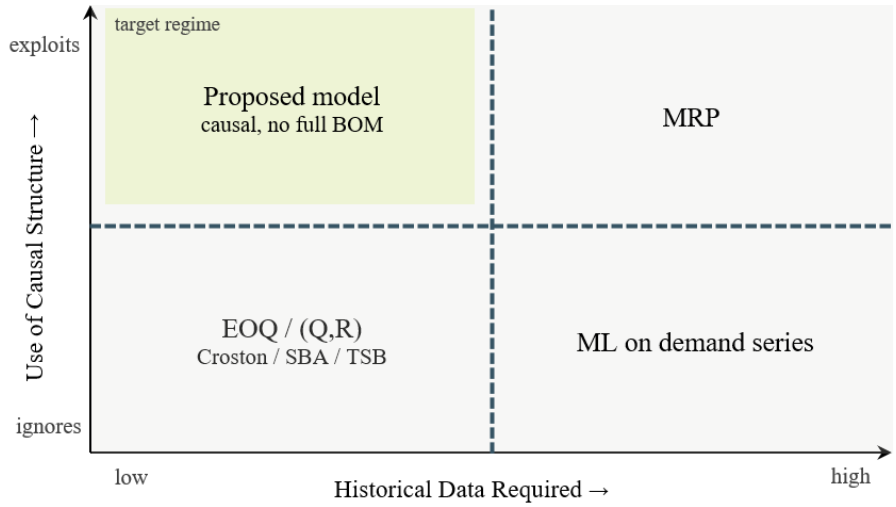


Figure 2.2: Qualitative positioning of forecasting and planning approaches according to the volume of historical data they require and the extent to which they exploit the causal structure of the demand. The proposed model occupies the regime of low data requirement and explicit use of the causal structure.

Rather than attempting to learn the demand pattern from a sparse history, it reconstructs the demand from its causes. This positions the contribution not in opposition to the data-driven paradigm

but as a complementary, knowledge-driven alternative suited to a regime of intermittent demand and scarce data, where the causal structure of the problem is known even when the data are insufficient to learn it statistically. The statistical intermittent-demand methods of Section 2.4 remain the appropriate benchmark, while the more elaborate ML methods reviewed here are considered, and set aside, on the principled grounds that the data do not support them. This positioning is summarized in Figure 2.2, which situates the approaches by the data they require and the causal information they exploit.

2.3 Inventory Management Fundamentals

Inventory management is a core operational function in industrial and logistics systems, ensuring the availability of the materials required to support production, maintenance, and project execution. In environments characterized by uncertainty in demand and supplier lead times, inventory acts as a strategic buffer that decouples interdependent processes and mitigates the risk of operational disruptions [39]. From a theoretical standpoint, its primary objective is to determine stock levels that balance service performance against holding and replenishment costs. On one hand, insufficient inventory leads to stockouts, project delays, and reduced service levels, whereas excessive stock increases capital tie-up, storage costs, and the risk of obsolescence [38]. Effective inventory management therefore requires structured decisions regarding replenishment timing, order quantities, and the allocation of resources across competing operational needs [17].

Inventory fulfils several distinct functional roles within the logistics system. *Cycle stock* supports regular demand between replenishment events, while *safety stock* is held to absorb variability in consumption and lead time. *Pipeline inventory* refers to material in transit or under procurement, while *project-specific inventory* may be temporarily dedicated to a particular operational unit or worksite [17, 38]. The coexistence of these categories increases management complexity, particularly in organizations executing several simultaneous projects. To regulate stock levels, organizations adopt inventory control policies that define *when* and *how much* to replenish. Continuous review systems monitor the inventory position on an ongoing basis and trigger a replenishment order when a predefined reorder point is reached, whereas periodic review systems assess inventory status only at fixed intervals and consequently tend to require higher safety stock to compensate for their reduced responsiveness [24]. In practice, hybrid approaches are common, combining formal control rules with operational judgment.

A more fundamental distinction in the choice of inventory policy concerns the nature of the demand experienced by the item. The inventory literature traditionally distinguishes between *independent demand*, which arises externally and is not derived from the demand for any other item, and *dependent demand*, which is determined by the requirements of a higher-level item according to a known structure [32, 38]. This distinction is decisive, because it dictates which family

of methods is appropriate. Items subject to independent demand are typically managed through the statistical, reorder-point logic described below, in which future requirements are inferred from the historical demand pattern. Items subject to dependent demand, by contrast, are more appropriately managed through Material Requirements Planning (MRP), which does not forecast consumption statistically but instead computes net material requirements by exploding the Bill of Materials (BOM) of the parent item against a production or project plan [32]. The two paradigms are predicated on contradictory assumptions. Reorder-point methods treat demand as a stochastic signal to be estimated, whereas MRP treats it as a deterministic consequence of a known plan. As will become apparent in the course of this study, the demand encountered in project-based installation work occupies a middle ground between these two extremes, a tension that this dissertation addresses directly.

2.3.1 Classical Reorder-Point Models for Independent Demand

For items managed under the independent-demand paradigm, the most widely adopted approach in industrial practice is the continuous review model, also known as the (Q,R) or (s,Q) policy [28, 38]. Under this framework, the inventory position is monitored continuously, and a replenishment order of fixed quantity Q is triggered whenever the inventory position reaches a predefined Reorder Point (ROP). The reorder point is defined as the expected demand over the supplier lead time, augmented by a safety stock component that accounts for demand variability and supply uncertainty:

$$ROP = D \cdot L + SS \tag{2.1}$$

where D is the average demand rate, L the lead time, and SS the safety stock. This formulation ensures that, on average, replenishment occurs before stock is depleted, thereby balancing service-level performance against holding costs [17].

The characteristic sawtooth evolution of the inventory level under this policy, falling as demand is served and rising with each replenishment, is illustrated in Figure 2.3.

The safety stock component itself plays a critical role in mitigating uncertainty and is conventionally derived from the desired service level and the variability of demand. A widely used formulation expresses it as:

$$SS = Z \cdot \sigma_d \cdot \sqrt{L} \tag{2.2}$$

where Z is the service-level safety factor and σ_d the standard deviation of demand. Higher target service levels require larger values of Z, increasing inventory investment while reducing the probability of a stockout [38]. It is important to note that this expression rests on the specific statistical premise that demand over the lead time can be adequately described by a normal distribution

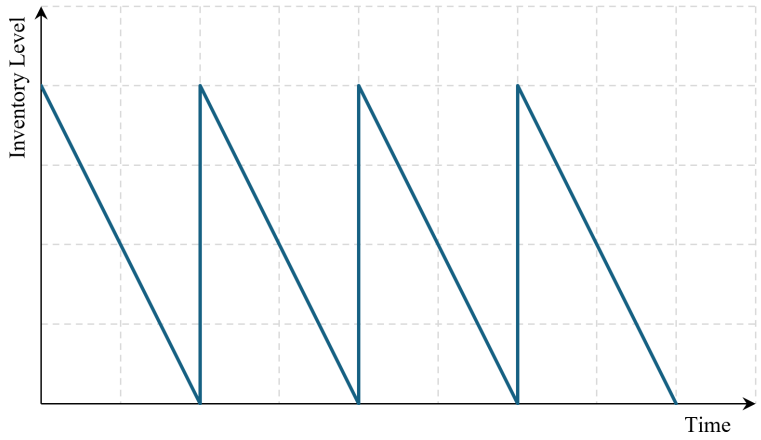


Figure 2.3: An idealized representation of how inventory levels evolve over time at a given location. Adapted from Bartholdi et al. [4].

with a stable mean and standard deviation. This premise is examined critically in Section 2.3.3.

Complementing the reorder-point logic, Economic Order Quantity (EOQ) model addresses the distinct question of *how much* to order, determining the quantity that minimizes the combined ordering and holding cost. First formalized by Harris in 1913 [13] and later popularized in industrial practice, it is among the oldest results in inventory theory and is expressed as:

$$EOQ = \sqrt{\frac{2DS}{H}} \tag{2.3}$$

where D denotes annual demand, S the fixed ordering cost, and H the holding cost per unit per period [17].

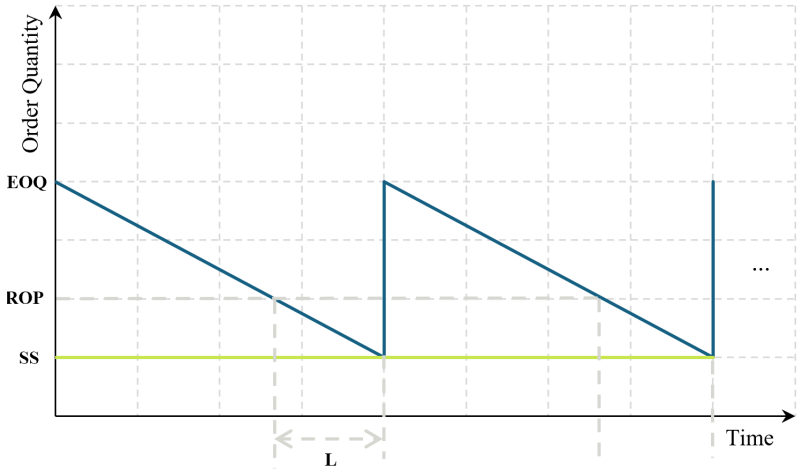


Figure 2.4: Economic Order Quantity model illustrating the reorder point, lead time, and replenishment cycle. Adapted from Jadidi et al. [17].

The EOQ remains a foundational reference for understanding the economic trade-off at the heart of every replenishment decision, even though, as discussed below, the conditions under which it was derived are rarely satisfied in practice [38]. The trade-off between ordering and holding costs that the model resolves is illustrated in Figure 2.4.

However, while mathematically robust for continuous operations, these classical models become a structural liability in project-based environments. Their reliance on mean demand and standard deviation fundamentally misinterprets lumpy consumption, inevitably leading to inflated safety stocks and immobilized capital.

2.3.2 Material Requirements Planning for Dependent Demand

Where consumption is driven by an identifiable higher-level plan rather than by an autonomous demand stream, MRP offers an alternative that does not rely on the statistical assumptions of the reorder-point models. Originating in manufacturing planning and consolidated in the work compiled by Orlicky [32], MRP starts from a master plan of end items, decomposes each into its constituent components through the BOM, and offsets the resulting gross requirements by the corresponding lead times to determine *when* and *how much* of each component must be ordered. Because requirements are computed rather than forecast, MRP avoids both the over and under stocking that arise when reorder-point logic is applied to items whose demand is, in reality, lumpy and plan-driven [32].

The relevance of this paradigm to project-based installation work is immediate as materials consumed on a worksite are, in principle, dependent on the engineering design of that specific project, much as components are dependent on the BOM of a finished product. In practice, however, the direct application of MRP in this setting is constrained by a key limitation, namely that a complete and reliable BOM is rarely available early enough in the project lifecycle to drive planning, and the effort required to maintain detailed BOMs across a large portfolio of heterogeneous projects is often prohibitive. This dissertation builds on the central insight of MRP, that project demand is structured rather than random, while seeking to relax its dependence on a fully specified BOM. The approach developed in Chapter 4 may thus be understood as occupying an intermediate position between the two classical paradigms.

2.3.3 The Assumptions Underlying the Classical Models

Although the models above provide a robust theoretical foundation, their analytical tractability is purchased at the cost of a set of restrictive assumptions. Making these assumptions explicit is essential, because it is precisely their violation that motivates the methodological direction of this

dissertation.

First, the EOQ model assumes that demand occurs at a *constant and continuous* rate and that the lead time is deterministic [13, 38]. Second, the safety stock formulation of Equation 2.2 assumes that demand is a *stationary random variable following a normal distribution*, so that its dispersion is well captured by a single standard deviation σ_d. Third, the continuous review logic of Equation 2.1 presupposes that an average demand rate D is a meaningful descriptor of consumption over the lead time, which again implies a demand process that is regular rather than sporadic. The MRP paradigm, while free of these statistical assumptions, requires the availability of an accurate, timely, and complete bill of materials. Taken together, the reorder-point models describe an idealized environment of smooth, frequent, and statistically stable demand, whereas MRP describes one of fully specified, plan-driven requirements.

In many production and distribution settings the reorder-point assumptions hold sufficiently well for the classical models to perform adequately. In project-based industrial environments, however, material consumption is typically intermittent and closely tied to the execution phases of individual projects. This is, a given installation material may exhibit long periods of zero demand interrupted by large, concentrated bursts when a project enters the corresponding phase. Under such conditions, the average rate D loses descriptive meaning, and the standard deviation σ_d, computed over a series dominated by zeros and occasional spikes, inflates the safety stock to unrealistic levels. At the same time, the BOM-driven precision required by MRP is seldom attainable at the moment procurement decisions must be made. The classical apparatus thus fails the statistical models because their assumptions of regularity are violated, and the deterministic models because their informational requirements cannot be met. Characterizing this class of demand rigorously, and reviewing the forecasting methods developed specifically for it, is the subject of the following section.

2.4 Forecasting Demand in Intermittent and Project-Driven Contexts

The preceding section concluded that the classical inventory apparatus is found to be ineffectual when confronted with installation materials from two distinct perspectives. Firstly, the statistical reorder-point models are deemed to be incompatible due to the violation of their underlying assumptions concerning the nature of demand, which is typically assumed to be normal and distributed. Secondly, the material requirements planning MRP system is found to be deficient due to the rarity of the complete bill of materials required being available in a timely manner. This section characterizes the kind of demand responsible for the first of these failures, reviews the forecasting methods developed specifically for it, and identifies the opening that the present work exploits.

2.4.1 Characterizing Intermittent and Lumpy Demand

Before any forecasting method can be selected, the demand it must address has to be characterized. The demand series of an item is simply the record of how much of it was consumed in each successive period. For an item in steady use, this series fluctuates moderately around a stable average, and the classical apparatus of Section 2.3 applies naturally. The materials consumed in project-based installation work rarely behave this way. A given item is drawn in quantity during the brief window in which an active project reaches the phase that requires it, and not at all otherwise. Its demand series is therefore dominated by zeros, interrupted by occasional, sometimes large, spikes [8]. Such a series exhibits irregularity in both *timing* and *size*. Demand occurs only sporadically, and the quantities involved may differ considerably from one occurrence to the next.

These two irregularities are captured by two distinct measures. The first, the Average Demand Interval (ADI), which addresses the mean number of periods between successive non-zero demands, computed as the total number of periods divided by the number of periods in which demand occurred. A material consumed in only three months out of twenty-four has an ADI of eight, indicating that demand arises, on average, once every eight months. The larger the ADI, the more sporadic the demand. The second measure, the *squared coefficient of variation* (CV^2) of the non-zero demands, addresses size. The coefficient of variation is the standard deviation of the demand sizes divided by their mean, a measure of variability relative to the typical magnitude, squaring it yields the CV^2. Crucially, this measure is computed using only the periods in which demand occurred, since the zeros have already been accounted for by the ADI. A low CV^2 indicates that demand, when it arises, is of a consistent size, while a high CV^2 indicates that the quantities are highly variable and therefore difficult to anticipate.

Because the two measures are independent, an item may be irregular in one dimension while regular in the other, and it is the combination that defines its demand pattern. Syntetos, Boylan and Croston [42] formalized this idea by crossing the two measures against empirically derived thresholds of $ADI = 1.32$ and $CV^2 = 0.49$, yielding the four categories shown in Figure 2.5. *Smooth* demand, in the lower-left region, is regular in both timing and size, and is the case for which the classical models were designed. *Intermittent* demand, in the lower-right, occurs sporadically but in consistent quantities each time. *Erratic* demand, in the upper-left, occurs frequently but in widely varying quantities. *Lumpy* demand, in the upper-right, is irregular in both respects at once, as it appears rarely and, when it does, in unpredictable amounts. This last category is both the most difficult to manage and the one into which the Class B installation materials of central interest to this work are expected to fall, a conjecture tested empirically in Chapter 5. Kostenko and Hyndman [23] later refined the precise boundary used to switch between forecasting methods, but the four-quadrant structure remains the standard reference for demand classification.

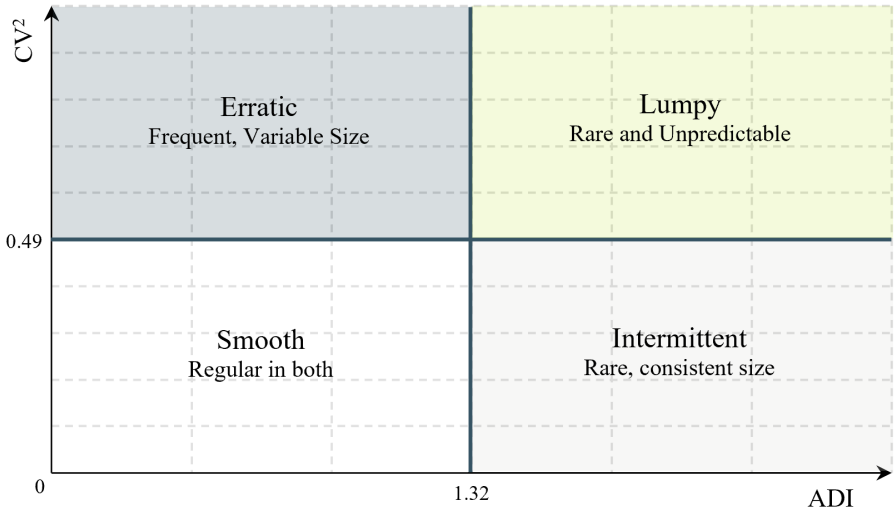


Figure 2.5: Demand classification scheme of Syntetos, Boylan and Croston [42], defined by the ADI and the squared coefficient of variation (CV^2) of non-zero demand. The thresholds $ADI = 1.32$ and $CV^2 = 0.49$ delimit the smooth, erratic, intermittent, and lumpy categories.

The value of this classification is twofold. Firstly, it replaces the subjective impression that a material's demand is "irregular" with an objective, reproducible measure computed directly from its own consumption history. This allows every item in the catalogue to be placed unambiguously in one of the four categories. Secondly, it indicates which forecasting and control method is appropriate. As the following section shows, the methods suited to smooth demand fail precisely on the lumpy items. As is shown in Chapter 4, this demand-pattern classification complements the value-based Activity-Based Classification (ABC) segmentation already used in practice. The economic importance of an item and the regularity of its demand represent distinct dimensions, and a coherent inventory policy must take into account both of these factors.

2.4.2 Forecasting Methods for Intermittent Demand

The foundational contribution to this field is due to Croston [8], who demonstrated that applying single exponential smoothing directly to an intermittent series produces biased and unresponsive estimates, and proposed instead to decompose the series into the size of demand when it occurs, and the interval between successive demands. Each component is smoothed separately, and the forecast of demand per period is obtained as the ratio of the two. This decomposition remains the conceptual basis of virtually all subsequent work on the problem [27].

Croston's original estimator was later shown to be biased. Syntetos and Boylan [41] derived a correction factor that yields an approximately unbiased estimate, a method now universally referred to as the Syntetos-Boylan Approximation (SBA) and widely regarded as the benchmark for intermittent demand forecasting. A further variant, the Teunter-Syntetos-Babai (TSB) method

[44], replaces the inter-demand interval with a demand-occurrence probability that is updated every period, making it more responsive when an item's demand is at risk of ceasing altogether, as in the case of obsolescence. The choice among Croston, SBA and TSB is itself informed by the ADI-CV^2 position of the series, with the bias-corrected methods generally preferred for the more intermittent and lumpy categories [23].

More recently, attention has turned to data-driven and machine-learning approaches. A critical review by Pinçe et al. [27] synthesizes two decades of research on spare-parts intermittent demand and notes both the persistent dominance of Croston-type methods and the growing interest in alternatives that exploit additional explanatory information. Of particular relevance to the present work, Aktepe et al. [2] applied regression and artificial-intelligence methods to demand forecasting in a construction-machinery company, demonstrating that, where covariates describing the operational context are available, methods that condition on those covariates can outperform purely time-series approaches. This points directly to the limitation addressed below.

2.4.3 The Limitation of Time-Series Methods and the Opening for a Causal Approach

A defining characteristic of the Croston family, including its SBA and TSB refinements, is its *context-agnostic*.This entails the treatment of the demand series as a self-contained statistical signal, with future demand being extrapolated from past patterns of occurrences and sizes alone. This standpoint is justifiable when the cause of demand is unobserved, as is generally the case for the spare parts for which these methods were developed, where a failure that triggers demand cannot be anticipated from the demand series itself.

In project-based installation work, however, this is not the case. The intermittency of material demand is a direct and observable consequence of the execution schedule of the active project portfolio. A spike in demand for a given material occurs because, and precisely when, a project enters the phase in which that material is required. The information that a purely time-series method must treat as hidden, namely *why* and *when* the next spike will occur, is in fact known in advance, encoded in the size and timing of the projects currently under execution. This represents a missed opportunity for the classical intermittent-demand methods and, conversely, the central opening exploited by this dissertation. Rather than extrapolating the intermittent signal statistically, the demand for installation materials can be *anticipated causally* from the characteristics of the active projects. The approach developed in Chapter 4 operationalizes this idea, while the Croston, SBA and TSB methods reviewed here provide the natural statistical benchmark against which its performance must be assessed.

2.5 Inventory Management in Project-Based Environments

The previous section established that the intermittency of installation material demand is not random but is the observable consequence of the active project schedule, and argued that this opens the way to a causal rather than a purely statistical forecast. Two established bodies of work are particularly relevant in this regard. Parametric estimating provides a basis for relating material consumption to project scale, while construction progress modelling provides a basis for distributing that consumption over the project duration. Together, these approaches provide the theoretical foundation for the model developed in Chapter 4.

2.5.1 The Project-Based Production Context

Project-based industries, including construction and the associated electromechanical, electrical, and hydraulic installation trades, are commonly characterized as engineer-to-order or make-to-order environments, in which each output is a largely unique deliverable executed against a specific contract rather than a standardized product made to stock. Material demand in these environments is *derived*, as it follows from the engineering content and execution schedule of the projects in the active portfolio rather than from an autonomous market process. This places such materials, in principle, within the dependent-demand paradigm discussed in Section 2.3.3. At the same time, projects are *heterogeneous in scale*. The material content of a large commercial facility can differ substantially from that of a small residential intervention, not only in absolute quantity but also in composition. Historical consumption therefore cannot be pooled directly across projects without some form of normalization. These characteristics create difficulties for both classical statistical models and the direct application of MRP, motivating the two techniques reviewed below.

2.5.2 Estimating Project Quantities

Where the total quantity of a material required by a project is not known in advance, it must be estimated. The established response to this problem is *parametric estimating*, in which a quantity or cost of interest is expressed as a statistical function of one or more high-level project parameters, such as floor area, capacity, or budget, calibrated on historical data from comparable projects [9, 14]. The resulting function, commonly termed a Cost Estimating Relationship and typically estimated by regression [22], is valuable precisely because it is applicable at the early project stages at which detailed design information is unavailable. In organizations that produce a detailed quantity map at the budgeting stage, however, the total requirement of each material is already known with far greater accuracy than any parametric estimate could provide, and statistical estimation of the quantity becomes unnecessary. As established in Chapter 3, this is the case at PEMEL. Parametric estimation is therefore not employed for the quantity dimension in this work as the quantity map is used directly. The technique is noted here because it remains the appropriate recourse when such documentation is absent, for instance when sizing stock against a pipeline of prospective projects

not yet awarded, a possibility revisited in the discussion of future work.

2.5.3 Temporal Distribution of Demand and the Construction S-Curve

Estimating the total quantity of a material a project will consume answers only half of the inventory question as the other half concerns *when* that consumption occurs. In project-based work this timing is highly non-uniform, because materials are drawn down according to the execution phase a project has reached (e.g. foundation and rough-in materials early, finishing components late). The standard representation of this phenomenon in construction management is the *S-curve*, which plots the cumulative value of work, cost, or resource consumption against elapsed project time. Its characteristic sigmoid shape, slow at the outset, steep through the middle phases of peak activity, and tapering towards completion, reflects the typical rhythm of project execution and is widely used for cash-flow forecasting and progress control [19].

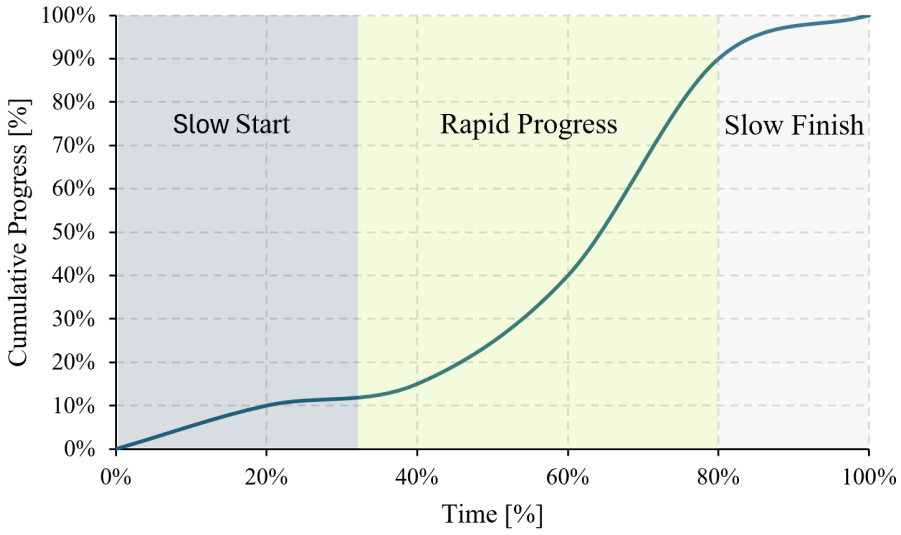


Figure 2.6: Generic S-Curve representing cumulative project progress over time. Adapted from Kerzner [19].

While the S-curve is conventionally applied to aggregate cost or progress, the same logic extends naturally to the consumption of individual materials, each of which possesses its own characteristic phasing within the project lifecycle. A material-specific delivery or consumption curve, reconstructed from the historical timing of deliveries relative to project progress, captures the phase at which a given item is required. Combined with the parametric estimate of total quantity, this temporal profile allows the expected consumption of each material to be allocated to the appropriate stage of each active project's lifecycle. The resulting material requirements can then be aggregated across projects at different stages of completion.

2.5.4 Research Gap Analysis

The preceding sections allow the gap addressed by this work to be stated not as an assertion but as a structured comparison. Each family of methods can be characterized by the information it requires and exploits, including whether it depends on a detailed BOM, whether it learns from historical demand, and whether it incorporates knowledge of the active project portfolio that drives demand. Table 2.1 summarizes this comparison.

Table 2.1: Comparison of forecasting and planning approaches by information used. The proposed model is highlighted in bold.

Approach	Needs BOM	Uses demand history	Uses active projects	Demand regime
EOQ / (Q,R)	No	Yes	No	Smooth
MRP	Yes	No	Yes	Deterministic
Croston / SBA / TSB	No	Yes	No	Intermittent
ML on demand series	No	Yes	No	Data-rich
Proposed model	**No**	**Partial**	**Yes**	**Lumpy, project-driven**

The structure of the table makes the gap explicit. The classical statistical methods and the intermittent-demand methods derived from Croston share a common blind spot as they treat demand as an autonomous signal and make no use of the project portfolio, even though, in a project-based setting, that portfolio is the direct cause of the demand. MRP does exploit the project plan, but only through a complete BOM that, as established in Section 2.3.3, is rarely available in time. Machine-learning methods applied to the demand series share the blind spot of the statistical methods and additionally require data volumes that the present setting does not provide. No approach in the reviewed literature occupies the final row that exploits the active project portfolio as a causal driver, without requiring a complete BOM, and that is tailored to lumpy, project-driven demand. It is this combination, rather than any single attribute, that constitutes the novelty of the proposed model, and the remainder of the dissertation is devoted to developing and examining it.

2.6 Enabling Technologies: Data Capture, Warehouse Systems, and Decision Support

The model outlined in the preceding sections rests the practical preconditions that the historical consumption data used to calibrate it are reliable, and that the resulting decisions are accessible to

the operational staff who act on them. Neither precondition is automatically satisfied in an environment that relies on manual record-keeping. This section reviews the technologies that satisfy them, namely automated data capture, warehouse management systems, and decision support systems. These are treated here not as the contribution of this work but as its enabling layer.

2.6.1 Data Capture and Traceability

The accuracy of any inventory model is bounded by the accuracy of the data it is built upon, and that accuracy is in turn determined by how material movements are recorded. The transition from manual counting and entry to automated identification, through barcodes, Quick Response (QR) codes, or RFID, is the foundational step in this respect [3, 36]. Barcode systems associate each item or location with a machine-readable code that can be scanned during receiving, storage, and dispatch, allowing movements to be registered quickly and with substantially fewer errors than manual entry. Istiqomah et al. [3] report corresponding gains in receiving and dispatch efficiency and in the accuracy of real-time stock records. RFID extends this principle by enabling identification without line-of-sight scanning, supporting continuous rather than transaction-triggered monitoring [43]. Beyond the current state of stock, these technologies underpin *traceability*, the ability to reconstruct the historical movement of a material through the system, which is precisely the capability on which the reconstruction of historical consumption profiles in this work depends [36, 51]. In project-based settings, where the same material is dispatched to many worksites, this historical record is the raw material from which demand patterns can later be inferred.

2.6.2 Warehouse Management Systems and ERP Integration

Where data capture addresses the reliability of individual records, the Warehouse Management System (WMS) addresses their organization. A WMS is an information system supporting the core warehouse functions of receiving, storage allocation, picking, dispatch, and stock verification, formalizing these routines so that digital records reflect physical reality more faithfully [10, 35]. Figure 2.7 illustrates how these warehouse processes are coordinated through a centralized inventory database and integrated with the ERP system.

The literature increasingly frames the adoption of such systems, together with the data-capture technologies that feed them, as part of a broader move from manual toward "smart" warehousing, with reported benefits in stock accuracy and the reduction of redundant procurement [18, 20, 46].

For small and medium enterprises in particular, the value of a WMS depends heavily on its integration with the ERP platform that holds the organization's master data, so that material availability is represented in a single, consistent source rather than in parallel and potentially divergent records [49]. This integration is, however, frequently obstructed by legacy or closed ERP systems

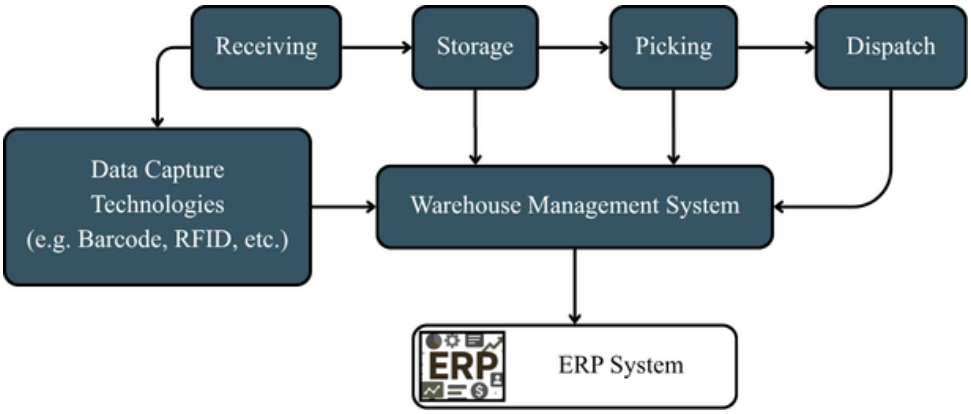


Figure 2.7: Simplified functional architecture of warehouse operations supported by a Warehouse Management System, illustrating the interaction between operational processes, centralized inven tory databases, and Enterprise Resource Planning platforms.

whose modification is costly or risky. In response, recent work has proposed lightweight, low-cost integration architectures that synchronize shop-floor data with central platforms without altering the core system [30, 34]. This line of work is directly relevant to the implementation pursued in this dissertation, which adopts a comparable read-only, non-intrusive integration with the existing ERP rather than a replacement of it.

2.6.3 Decision Support Systems

A quantitative model, however sound, produces value only when its output reaches the people who must act on it, and in a form they can use. This condition is easily overlooked, yet it is decisive in the setting of this work, where, as the diagnosis of Chapter 3 will show, the information needed to manage material is held in an ERP accessible to only a few users and remains out of reach of the site and warehouse staff who handle the material daily. The question of how analytical output becomes operational action is the domain of the Decision Support System (DSS), and it is the framework within which the model and the monitoring tools of this dissertation are best understood as a single system rather than as separate artefacts.

In the foundational framework of Sprague [40], a DSS is an interactive, computer-based system that supports decision-making on semi-structured problems through three components, a *data* component that holds the relevant information, a *model* component that processes it into a recommendation, and a *dialog* component that mediates between the system and the user. Power [29] later distinguished several types of such systems, among them *model-driven* systems, in which an analytical model is the principal engine of support, and *data-driven* systems, oriented toward the access and visualization of operational data. The value of this framing, for the present work, is that it names the parts of the solution and the relations between them, so that what might otherwise appear as two unrelated tools, a sizing model and a dashboard, are seen instead as the components

of one decision-support system with a common purpose.

Read in these terms, the solution developed in this dissertation is a model-driven DSS. The inventory model of Chapter 4 is its model component, the ERP records are its data component, and the monitoring dashboard is its dialog component, the layer through which both the model's recommendations and the underlying stock status become visible to those who lack direct ERP access. The consolidation of fragmented operational information into a single view, of the kind a dashboard provides, has been reported to improve inventory visibility and reduce stockouts [26], but its role here is more specific than general visibility. It is the channel that closes the loop between the model and operational action, without which the quantitative recommendation would remain, like the ERP data it draws on, accessible only to a few. Framed this way, the dashboard is not an end in itself but the final link in a chain that runs from reliable data capture, through structured warehouse and ERP records, to model-based decision support that reaches the operational front line. This progression directly aligns with the end-to-end analytics pipeline illustrated in Figure 2.8, which maps the necessary conceptual stages from initial data sourcing and preparation to the ultimate generation of actionable insights for decision-making.

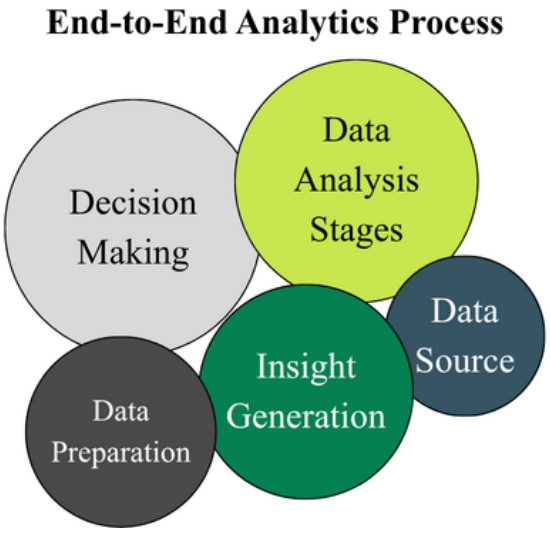


Figure 2.8: Conceptual pipeline of an end-to-end analytics process illustrating the progression from raw data sourcing to strategic decision-making. Adapted from Champa and Segall [6].

The practical realization of the aforementioned data and dialog components relies on an enabling layer driven by the paradigm of low-code development. Low-code platforms utilize intuitive visual interfaces and pre-built modular components to facilitate software creation while strictly minimizing the need for manual programming [1, 12]. In the context of a Decision Support System, these technologies act as the pragmatic medium through which complex analytical models are connected to the operational front line. They allow the interface and data capture mechanisms to be constructed and integrated accessibly, ensuring that the final system remains adaptable and

closely aligned with the actual workflows of the warehouse staff rather than becoming a rigid software artifact.

This technological approach fosters what the literature identifies as citizen development, a practice that empowers business professionals lacking formal software engineering backgrounds to design and deploy functional applications [1]. The primary advantages of this methodology include a significant acceleration of development cycles and a drastic reduction in the specialized programming competencies typically required to build custom enterprise software [12]. By abstracting the underlying code complexity into graphical environments, low-code tools allow domain experts to directly embed their operational knowledge into the system design, which concurrently simplifies the long-term maintenance and iterative improvement of the deployed tools [1, 12].

The adoption of low-code solutions is particularly critical for Small and Medium Enterprises that frequently face strict financial constraints and lack dedicated internal software development teams [6, 12]. Traditional software engineering approaches often introduce prohibitive costs and require specialized personnel that are out of reach for companies operating on tighter margins. Low-code platforms circumvent these barriers by enabling resource-constrained enterprises to build and maintain customized digital tools cost-effectively [1, 12]. Consequently, the selection of Power Apps and Power BI for the tools developed in this dissertation is a deliberate methodological choice. These platforms ensure that the resulting Decision Support System can be fully implemented, sustained, and independently modified by the company without necessitating heavy external software development investments, thereby guaranteeing the practical viability of the inventory management solution within the SME setting.

Chapter 3

Company Context and Current System

This chapter presents the diagnosis of the case company. It maps the flow of material, describes the warehouse operation and the tracking of orders, and analyses the operational bottlenecks and their impact. From this it establishes a baseline of current performance, against which improvement is later assessed, and draws the diagnosed deficiencies together into the requirements that shape the proposed solution.

3.1 Material Flow Mapping

The procurement and delivery process begins with the identification of a material requirement at the construction site. Typically, the Site Manager prepares an Internal Site Requisition (RIO), specifying the required items, their quantities, and, where applicable, a preferred supplier. This request is submitted to the purchasing department for processing.

Upon receiving the request, the purchasing department may negotiate with suppliers to obtain improved commercial conditions, particularly when this is explicitly requested by the Site Manager. A formal Purchase Order (ECF) is then issued and transmitted to the selected supplier, defining the materials, agreed quantities, unit prices, and the designated delivery location.

Depending on project requirements and logistical constraints, materials may be delivered either directly to the construction site or to the company's central warehouse for temporary storage prior to final allocation. Each delivery is accompanied by a Delivery Note (GUIA), which specifies the items supplied, their quantities, prices, and a reference to the corresponding ECF. When materials are delivered directly to site, the receiving procedure is conducted by a site supervisor or team leader, who verifies the delivered items and quantities against the delivery note and forwards photographic evidence of the documentation to the purchasing department for visual confirmation.

When delivery occurs at the central warehouse, an equivalent verification is performed by warehouse personnel.

Once the purchasing department confirms that the delivery documentation matches the purchase order, the validated information is forwarded to the accounting department for supplier payment. Where discrepancies are identified, the supplier is notified and corrective action is initiated, which may include quantity adjustments, price clarification, or the issuance of revised documentation.

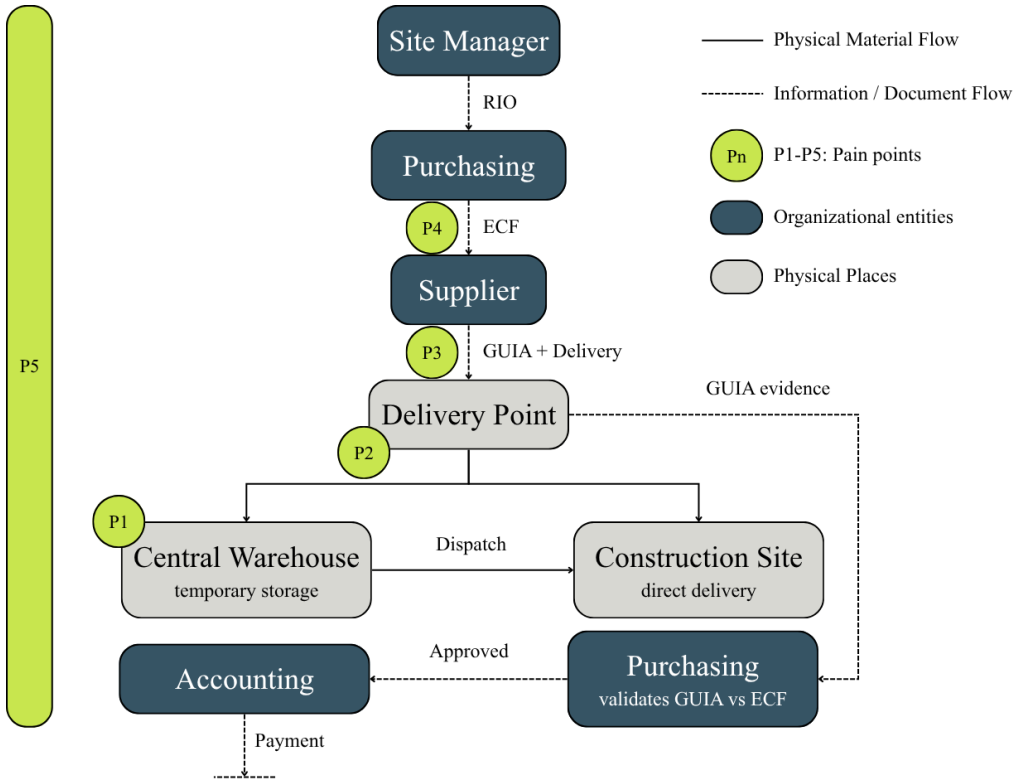


Figure 3.1: Flowchart of the current materials management and logistics process, from the RIO through ECF, GUIA, and execution.

The diagram in Figure 3.1 uses two visual conventions to expose the nature of the current process. First, two types of connector distinguish the two flows that coexist in the supply chain, with solid arrows denoting the physical movement of materials and dashed arrows denoting the flow of information and documents. Second, the node colours separate the organizational entities involved (Site Manager, Purchasing, Supplier, Accounting) from the physical locations through which materials pass (Delivery Point, Central Warehouse, Construction Site). Superimposed on this flow are five numbered markers, P1 to P5, indicating the points at which the operational deficiencies analysed in Section 3.4 arise. These are mapped to their proposed responses in Table 3.2. The figure immediately conveys that the physical flow of material proceeds in an orderly, largely linear fashion, whereas the information flow is fragmented, manual, and routed through a sequence

of disconnected validation steps. It is in this informational layer, rather than in the movement of materials itself, that most of the identified deficiencies are concentrated.

3.2 Warehouse Operation Description

The central warehouse currently operates as a temporary staging area rather than as a structured, strategic inventory facility. As reflected in the flow mapping above, materials are routed to the warehouse rather than directly to site only under circumstances when the Site Manager opts for central storage owing to the instability of construction schedules and the lack of secure, long-term storage conditions on site.

Upon arrival at the warehouse, operational procedures rely heavily on manual execution. The physical reception, validation against delivery notes, and subsequent registration of materials into the ERP system are non-automated. This reliance on manual data entry creates persistent administrative bottlenecks, increases susceptibility to human error, and severely limits real-time visibility over stored assets.

A consequence of this operational paradigm is that the company functions without a formalized, proactive inventory policy. The material kept in the warehouse is of two kinds. On one hand, a limited stock of low-value consumables, such as fixing and fastening accessories, is maintained on a continuing basis, since these items are inexpensive and in constant use. On the other, and occupying the bulk of the shelves, is an unstructured accumulation of surplus items and leftovers returned from completed projects. What is absent, in either case, is a deliberate policy governing the installation materials of intermediate value, the Class B materials at the centre of this study, whose stock is neither planned nor consistently held, but arises only residually.

3.3 Order Tracking

The final stage of the procurement cycle before physical reception is the tracking of outstanding purchase orders. In the current model, order tracking is not an automated feature of the ERP system, as it is a manual procedure executed by the site and procurement teams.

Without a centralized, automated alert mechanism, the follow-up process is fragmented. The high volume of simultaneously active projects, combined with the manual nature of the task, occasionally results in orders being lost or forgotten. This is compounded by suppliers who may themselves experience internal processing delays or fail to confirm receipt of purchase orders. In the absence of a proactive tracking mechanism, such delays typically become visible only when a Site Manager reports a material shortage, often too late to avoid disruption to the project.

Recognizing the criticality of this bottleneck, the company has recently created a dedicated role whose primary responsibility is the tracking and expediting of orders, acting as a liaison between the purchasing department and suppliers. This organizational measure is discussed further in Section 3.4.2, where its effectiveness and its limitations under the current manual infrastructure are analysed.

3.4 Operational Bottlenecks and Impact Analysis

Owing to the reactive nature of the current process and the absence of automated performance indicators within the ERP system, the quantification of supply-chain disruptions requires a dedicated diagnostic effort. To this end, an internal assessment is being conducted through structured surveys directed at Site Managers and procurement personnel, complemented where possible by transactional data extracted from the ERP. The qualitative bottlenecks are, however, clearly identifiable independently of the final metrics, and are described below together with the impact indicators that the assessment is intended to quantify.

3.4.1 Material Shortages and Idle Labour

The prevailing near-zero-stock approach assumes near-perfect synchronization between supplier deliveries and the construction schedule. While surplus materials from completed projects occasionally provide minor relief, they are insufficient to sustain the dynamic and unpredictable needs of active sites.

The absence of a structured material buffer has repeatedly exposed the vulnerability of this reactive model. Operational feedback identifies recurrent instances in which projects experienced delays due to the unavailability of basic infrastructure components, such as cabling or piping. When a supplier delay occurs, it translates directly into on-site disruption, leaving specialized labour idle. Preliminary assessment indicates that these occurrences generate a material amount of idle labour per project-month. This idle labour constitutes a hidden but substantial financial burden that may outweigh the holding costs that the zero-stock approach is intended to avoid.

The consequences of the absence of a deliberate inventory policy are felt beyond project execution itself. The after-sales service, which must respond rapidly to equipment failures, is likewise affected by the lack of readily available material, since its demand is unplanned and cannot be anticipated from the project portfolio. Although this is a relevant dimension of the problem, it is secondary to the project-driven demand that dominates the organization's material flow and is the focus of this work. After-sales demand is addressed in a separate study within the organization and is therefore not considered further here.

3.4.2 Constraints in Reactive Order Tracking

As noted in Section 3.3, the company has assigned a dedicated employee to order follow-up and supplier coordination. While this represents a deliberate step toward improving supply-chain reliability, the function remains heavily constrained by the existing manual infrastructure. In the absence of real-time digital dashboards or automated status updates within the ERP, the effectiveness of the follow-up depends on time-consuming manual communication, such as telephone calls and emails, and there is no consolidated view of all pending deliveries. Consequently, even with dedicated personnel, the tracking process remains largely reactive and struggles to keep pace with the volatility of supplier lead times.

3.4.3 Compromised Agility and Client Relations

Beyond the scope of planned projects, the lack of a dedicated stock buffer restricts the company's operational agility. PEMEL frequently receives requests for urgent, last-minute maintenance interventions from its clients, and the inability to dispatch materials immediately from a central stock delays these critical repairs.

In a sector where reliability is a primary determinant of reputation, such operational failures have a compounding negative effect. The lack of material availability not only inflates internal costs, through expedited shipping fees and idle labour, but also erodes client trust. The creation of a calculated stock buffer is therefore not merely a logistical optimization but a means of preventing project stoppages, supporting emergency interventions, and safeguarding long-term client relationships.

3.4.4 Absence of a Performance Measurement System

Underlying the operational bottlenecks described above is a more structural deficiency.The organization currently lacks any systematic system for measuring the performance of its logistics function. There are no defined key performance indicators, and no routine collection of metrics, for the warehouse, for supplier reliability, or for the procurement process. Performance is assessed reactively and anecdotally, through the reporting of incidents such as a material shortage, rather than through continuous, quantitative monitoring.

This absence has a twofold consequence. Operationally, it deprives managers of the information required to identify deteriorating supplier performance, recurring shortage patterns, or inefficient stock holding before these problems manifest as project disruptions. Decisions are therefore necessarily based on individual experience and tacit knowledge rather than on evidence. Methodologically, it is the reason that the quantification of the bottlenecks in this chapter must rely on a dedicated survey and on the manual extraction of transactional records, rather than on indicators

readily available from the ERP. The very difficulty of measuring the current state is therefore itself a symptom of the problem. An organization that measured its logistics performance routinely would not need to reconstruct that performance retrospectively. Establishing a baseline set of indicators, as the following section does, is thus not only a precondition for evaluating the proposed solution but also, in itself, a first step toward the performance-measurement capability the organization currently lacks.

3.5 Baseline Performance

To complement the qualitative diagnosis with evidence from those directly involved in the operation, a structured questionnaire was administered to the company's operational staff, covering the frequency and impact of material shortages, the responsiveness of procurement, and the accessibility of information. The full questionnaire is reproduced in Appendix A.

Eight professionals responded, spanning site direction, field technical roles, procurement, warehousing, and site administration. Given the reduced number of respondents, the results are treated as the structured perceptions of a small group of key informants rather than as a statistically representative sample, and are reported in absolute terms. The data are estimates of the professionals involved, not measurements extracted from the information system. Table 3.1 summarizes the responses, which offer a consistent picture of the operational reality described throughout this chapter.

A point of interpretation must be made regarding the frequency of shortages. When a respondent reports one or two shortages in four weeks, this reflects the experience associated with the projects that respondent follows. Since the company executes more than ten projects concurrently, a per-project frequency of this order aggregates, at the level of the organization, into a systemic incidence of the order of tens of shortage events per month. At this scale, shortages are not an occasional disturbance but a recurrent feature of the operation. Five of the six respondents in site-related roles reported shortages in their projects, and six of the eight reported that, when a shortage occurs, work is halted for half a day or more, typically affecting two to three workers and in one case as many as six. The immobilization of labour that follows from material unavailability is therefore both frequent, at the scale of the active portfolio, and materially costly when it occurs.

The reactivity of procurement and the reliability of supply emerge with equal clarity. Seven of the eight respondents reported having placed urgent orders, incurring additional cost, on more than three occasions in the preceding four weeks, and six of the eight perceived supplier deliveries as arriving later than promised frequently or almost always. This perception is consistent with the objective record of the ERP, in which a substantial majority of deliveries are received after the promised date, and lends independent support to the characterization of procurement as reactive

Table 3.1: Summary of questionnaire responses ($N = 8$ operational staff). Results are reported as absolute counts, reflecting the perceptions of a small group of key informants rather than a statistically representative sample.

Indicator	**Response ($N = 8$)**
Respondents reporting material shortages in their projects (last 4 weeks)	6 of 8 (5 of 6 among site-related roles)
Typical stoppage when a shortage occurs	6 of 8 report half a day or more
Workers halted or redeployed per shortage	Mostly 2 to 3; up to 4 to 6 in one case
Urgent orders placed (last 4 weeks, 3+ times)	7 of 8
Orders perceived as lost or forgotten in follow-up (frequently)	4 of 8
Supplier deliveries perceived as late (frequently or almost always)	6 of 8
Autonomous access to procurement information	5 of 8 report ready access; difficulty concentrated in front-line field roles
Perceived impact (mean, 1-5 scale)	Delays 3.8 Costs 3.6 Client 3.4
Materials most cited as causing stoppages	Piping, ducting, cabling, copper, fixing accessories

and of supply lead times as a significant source of uncertainty. The perceived impact of these situations was consistently moderate to high. On a five-point scale, the average perceived impact was 3.8 for project delays, 3.6 for additional costs, and 3.4 for the client relationship.

The materials most frequently identified as causing work stoppages were, without exception, the installation materials at the centre of this study, namely piping, ducting, cabling, copper, and support and fixing accessories. This convergence between the materials that the staff perceive as problematic and those that the proposed model targets confirms the study's scope from the perspective of those who experience the problem daily.

One finding refines rather than confirms the diagnosis. Autonomous access to procurement information was reported as straightforward by five of the eight respondents, which might suggest that visibility is not a general deficiency. The divergence by role, however, tells that the field technician, furthest from the administrative systems, was the one who reported depending on others for this information, while site directors and office-based staff reported ready access. The visibility deficiency thus manifests most acutely at the operational front line, where material is actually handled, rather than uniformly across the organization.

Taken together, the responses corroborate the central elements of the diagnosis, including shortages that are recurrent and disruptive across the active project portfolio, reactive procurement that is dependent on urgent orders, frequent supply delays, and a material impact on project per-

formance. The free-text responses reinforced these themes, with one respondent summarizing the situation as a combination of absent order tracking, insufficient standardization in supplier interactions, and stock levels governed by poorly defined set points. These observations provide a direct link to the solution elements developed in the following chapter.

3.6 From Diagnosis to Solution

The deficiencies identified in this chapter are heterogeneous in nature. Some stem from the absence of a quantitative basis for inventory decisions, others from the unreliability of manually captured data, and others from the limited accessibility of procurement information to operational staff. They therefore call for correspondingly different responses, not all of which fall within the scope of this dissertation. Table 3.2 maps each identified problem to its proposed response and indicates whether that response constitutes a contribution of this work or an enabling or organizational measure.

Table 3.2: Mapping of identified problems to proposed responses.

	Identified problem	Proposed response	Scope
P1	No proactive inventory policy; stock is residual only (Sec. 3.2)	Dynamic phase-based inventory model for sizing minimum stock by demand pattern (Ch. 4)	Core contribution
P2	Manual receiving; unreliable stock records (Sec. 3.2)	Automated data capture (QR code) feeding the ERP	Enabling technology
P3	Limited visibility; orders lost in follow-up (Sec. 3.3)	Procurement-monitoring dashboard over ERP data	Enabling technology
P4	Restricted ERP access for operational staff (Sec. 3.4.2)	Dashboard as accessible visualization layer	Enabling technology
P5	No systematic performance measurement (Sec. 3.4.4)	Definition of baseline KPIs; dashboard for continuous monitoring	Contribution / enabling

Although the five problems present themselves as distinct operational symptoms, they are not independent. They are the manifestations, at different points of the material flow, of a single underlying condition, that the company manages a project-based material flow with practices and

tools conceived for continuous operation. The absence of a proactive inventory policy (P1) follows because a fixed reorder rule has no meaning for a demand that arrives in the discrete pulses of a project portfolio, so stock is left to form residually. The unreliability of the stock records (P2) follows because manual capture cannot keep pace with the volume and variety of material that a changing set of projects moves through the warehouse. The loss of visibility over orders (P3, P4) follows because a procurement process organized around individual projects, rather than around a consolidated material plan, fragments the information across contracts and users. And the absence of performance measurement (P5) follows because there is no baseline stock policy against which performance could be defined in the first place. The common root is therefore not a failure of execution but a structural mismatch between the nature of the demand and the nature of the tools applied to it, which is why the problems recur despite the competence of the staff and cannot be resolved by effort alone.

This structural mismatch is precisely the one that the literature review identified at the methodological level. As established in Chapter 2, the demand regime of a project-based installation company, intermittent, driven by a known portfolio, and supported by scarce historical data, falls into a gap that none of the available families of methods addresses. The operational deficiencies diagnosed in this chapter are thus the concrete, industrial expression of that abstract gap. What appears in the literature as a vacant region in the space of methods appears at PEMEL as the daily reality of stockouts, reactive expediting, and residual stock, and the correspondence between the two is what makes the case company a representative instance of the problem rather than an idiosyncratic one.

The mapping of Table 3.2 also embodies a deliberate ordering of priorities, which the structure of the problem justifies. The inventory model is identified as the core contribution not merely because it is the most technically demanding component but because it addresses the root condition, the missing proactive policy, from which the other deficiencies partly descend, and because it is the one response that lies in the methodological gap the literature leaves open and that the company cannot obtain from existing tools or practice. The data-capture and dashboard components, by contrast, address deficiencies that are real but not novel, since the technologies to capture data reliably and to visualize it already exist and are widely applied. They are therefore treated as enabling technologies, necessary for the model to operate on trustworthy inputs and to return accessible outputs, but not themselves the scientific object of the work. The organizational measures, finally, lie outside the dissertation entirely. This ordering is not a matter of relative effort but of where the genuine contribution lies, in supplying, for a regime that the established methods do not serve, a model grounded in the causal structure of the demand.

This mapping defines the scope of the work that follows. The core scientific contribution lies in the inventory model, which provides the company with the proactive policy it currently lacks. The data-capture and dashboard components act as enabling technologies, making the model's

inputs reliable and its outputs accessible. The organizational measures already initiated by the company are noted for completeness but are not the subject of this dissertation. The following chapter develops the core contribution in detail, building on the demand characterization established in Chapter 2.

Chapter 4

A Dynamic Phase-Based Model for Inventory Sizing in Project-Based Environments

This chapter develops the model that constitutes the core contribution. After an overview, it presents the two dimensions along which materials are classified, value and demand pattern, which delimit its scope, and then the mechanism itself, from the per-project quantities read from the budgets, through the temporal phasing of consumption, to the aggregation of a dynamic minimum stock across the active portfolio. It closes with the enabling layer, the data capture and the monitoring dashboard, that supplies the model's inputs and makes its outputs accessible.

4.1 Overview of the Proposed Model

The literature reviewed in Chapter 2 established that the demand for installation materials is lumpy and project-driven, and that neither the statistical reorder-point models nor classical MRP is suited to it. The diagnosis in Chapter 3 showed the practical consequence of this gap at PEMEL, namely the absence of any proactive inventory policy. This chapter develops a model that fills the gap by exploiting the one source of information that both classical approaches neglect, namely the known size and schedule of projects currently in execution. The present section gives an overview of the model as a whole, while the sections that follow develop each component in turn. In keeping with the separation maintained throughout this work, this chapter is concerned with the design of the model. Its empirical validation is the subject of Chapter 5.

The model takes the form of a data-processing pipeline, summarized in Figure 4.1, that transforms historical records and the current project portfolio into a stock requirement that adapts continuously to operational conditions. Its inputs are of two kinds. The first is the historical record held in the ERP, comprising past material consumption, the order-to-receipt dates of ECFs and GUIAs from which supplier lead times are derived, and the characteristics of completed projects.

The second is the live state of the active project portfolio, namely which projects are currently under execution, their scale, and the phase each has reached.

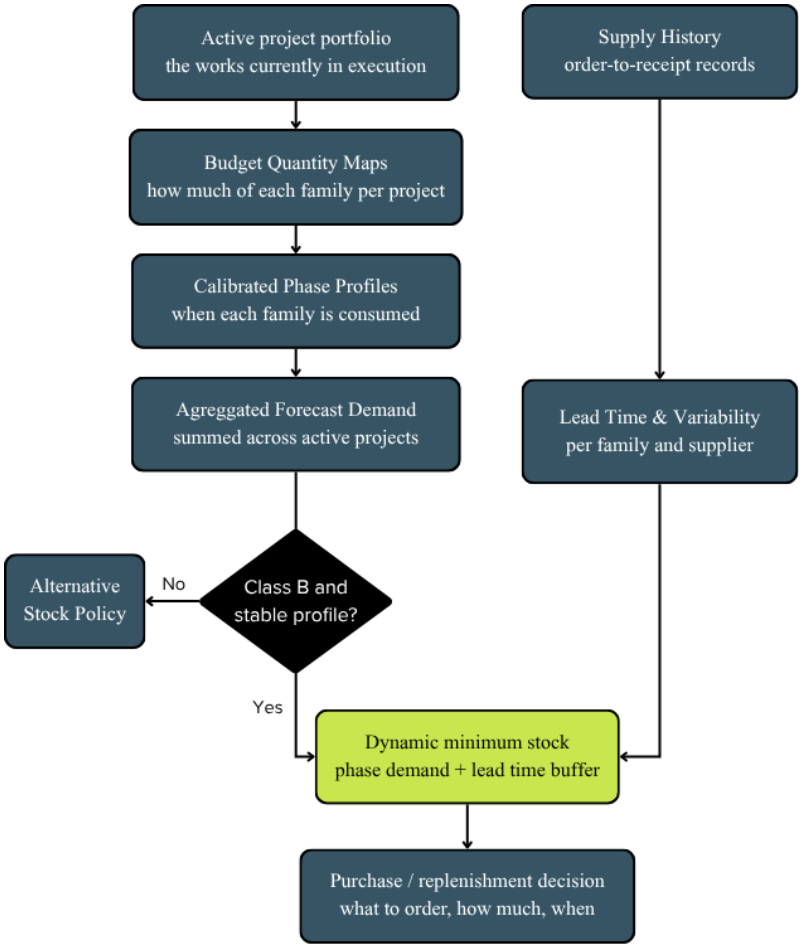


Figure 4.1: Architecture of the proposed dynamic phase-based model.

From these inputs, the model proceeds through three conceptual stages. First, a *classification* stage (Section 4.2) determines which materials are to be managed by the model, combining the value-based ABC segmentation with the demand-pattern classification of Section 2.4 to isolate the Class B installation materials for which the lumpy, project-driven dynamics are most pronounced. For each such material, the total quantity required by each active project is taken directly from that project's quantity map (Section 4.3), which answers *how much* a project needs in total, without the need for statistical estimation. The model's task is then to answer the question the quantity map leaves open, namely *when* that consumption occurs. This is the role of the *temporal phasing* stage (Section 4.4), which distributes each project's total requirement across its lifecycle according to the phase at which the material is characteristically consumed, reconstructed from historical records.

The phased per-project requirements then converge in the final *aggregation* stage (Section 4.5), which sums the phase-weighted consumption across all active projects to yield, for each material, the quantity required at the current point in time. A safety component is then added to this project-driven requirement, as discussed in Section 4.5.3. Because the model already anticipates the bulk of demand from the known quantities and their phasing, this component is dimensioned not against demand variability, as in the classical formulation, but against the residual uncertainty in supplier lead time and in the deviation of actual from budgeted consumption. The result is the dynamic minimum stock for each Class B material, a figure that rises and falls automatically as projects enter and leave the portfolio and as each progresses through its phases.

The model operates in a dynamic environment, influenced by the reliability of the input data and the subsequent utilisation of its outputs by relevant actors. These requirements are met by an enabling implementation layer, described in Section 4.6, in which automated QR-code capture improves the reliability of the consumption records that calibrate the model, and a monitoring dashboard makes its outputs accessible to operational staff. Together, the model and this enabling layer reconfigure the process diagnosed in Chapter 3, replacing its reactive, manually mediated character with one in which material requirements are anticipated and information flows reliably to those who act on it.

4.2 Material Classification by Value and Demand Pattern

An ERP of the kind operated by PEMEL holds a large and heterogeneous set of articles, ranging from major items of equipment to bulk installation materials and minor consumables. Applying a single inventory policy, or a single forecasting model, uniformly across this diversity would be both computationally wasteful and operationally misguided, since the materials differ markedly in the economic consequences of mismanaging them and in the regularity with which they are consumed. The first stage of the proposed model is therefore a classification that determines which materials warrant the phase-based treatment developed in the remainder of the chapter, and which are more appropriately managed by simpler means. This classification rests on two independent dimensions, addressed in turn below, whose intersection defines the scope of the model.

4.2.1 The Value Dimension: ABC Analysis

The first dimension concerns the economic importance of each material, and follows the classical ABC analysis derived from the Pareto principle [38]. Materials are ranked by their annual consumption value, the product of unit cost and consumption volume, and partitioned into three classes: a small number of Class A items that account for the majority of the total value, an intermediate group of Class B items of moderate individual value, and a large number of Class C items each representing a negligible fraction of the total. The rationale of the classification is that

managerial attention and inventory investment should be concentrated where they have the greatest financial effect. In the present context, Class A typically comprises major equipment such as chillers and heat pumps, Class B comprises bulk installation materials such as cabling, copper, and piping, and Class C comprises low-cost consumables such as fasteners and sealants.

4.2.2 The Demand-Pattern Dimension: ADI-CV² Classification

The value dimension alone is insufficient to determine the appropriate management method, because two materials of comparable value may exhibit entirely different demand behaviour, and it is the demand behaviour that dictates which forecasting approach can succeed. The second dimension therefore classifies each material by its demand pattern, applying the $ADI - CV^2$ scheme introduced in Section 2.4. For each material, the average demand interval and the squared coefficient of variation of its non-zero demands are computed from its historical consumption series, and the material is assigned to one of the four categories: smooth, intermittent, erratic, or lumpy. This classification is performed directly from the ERP consumption records and requires no subjective judgment, providing an objective basis on which to distinguish the materials whose demand is genuinely lumpy from those whose consumption, however valuable, is regular enough to be handled by conventional methods.

4.2.3 The Intersection: Scope of the Model

Crossing the two dimensions yields a combined classification in which each material carries both a value class and a demand-pattern category. The two are treated as independent, rather than assumed to coincide, so that the relationship between them becomes an empirical finding of the analysis rather than a premise of it. The phase-based model developed in this chapter is directed specifically at the materials that fall in the intersection of the Class B value group and the lumpy demand category, this is, those that are individually significant enough to justify active management, yet whose project-driven, intermittent consumption defeats the classical inventory models. It is for this group that the cost of mismanagement is highest in both directions, since stockouts of installation materials halt work on site, while overstocking them ties up capital and warehouse space to no purpose. Materials falling outside this intersection are managed by methods appropriate to their nature, summarized in Table 4.1.

Class A equipment, whose demand is tied to specific, identifiable project milestones, is suited to deterministic, bill-of-materials-driven planning of the MRP type, delivered just-in-time to avoid immobilizing high-value capital. Class C consumables, whose aggregate demand is high in frequency and stable in size, are adequately served by a simple static reorder-point or min-max policy, for which the low holding cost makes generous safety stock inexpensive. The phase-based model is reserved for the Class B lumpy materials, where neither of these simpler approaches applies and

Table 4.1: Inventory management policy by material class.

Attribute	Class A - Equipment	Class B - Installation materials	Class C - Consumables
Examples	Chillers, heat pumps, main panels	Cabling, cable trays, piping	Fasteners, sealants, fixings
Demand pattern	Deterministic, milestone-linked	Lumpy, phase-driven	Smooth, continuous
Forecasting basis	Engineering BOM (MRP)	Phase-based model (this work)	Historical average
Replenishment	Just-in-time to site	Dynamic project-driven min-max	Static min-max
Safety stock	None (capital tie-up unjustified)	Lead-time-based (Sec. 4.5.3)	High and static (cheap to hold)

where, consequently, the contribution of this work is concentrated.

4.3 Project Quantity Requirements

The first quantity the model requires is the total amount of each material that a given project will consume over its lifetime. In contrast to many project-based settings, where this figure must be estimated before detailed engineering, at PEMEL it is already available. When a project is budgeted, a quantity map (*mapa de quantidades*) is produced that specifies, item by item, the quantities of installation materials the project is expected to require, including the Class B materials of interest such as cabling, cable trays, and piping. The model therefore does not estimate this quantity but takes it directly from the existing project documentation, an approach that is both more accurate than any statistical estimate and consistent with the information the company already maintains.

For each active project p and each Class B material i, let $Q_{i,p}$ denote the total quantity of material i foreseen for project p, as given by that project's quantity map. This figure represents the complete requirement of the material over the entire duration of the project. It does not, however, indicate *when*, within that duration, the material is consumed, and it is this temporal dimension, absent from the quantity map, that the model must supply. The quantity map answers how much a project needs in total but does not answer how much must be available in the central warehouse at any given moment. This depends on the phase each active project has reached. Establishing that temporal distribution is the task of the following section.

It should be noted that the reliability of $Q_{i,p}$ as a representation of true consumption depends on the correspondence between the quantities foreseen in the budget and those actually consumed on site. As discussed in Chapter 3, deviations between the two arise on a project-by-project basis, owing to design changes, site conditions, and the return of surplus material. These deviations are

not systematic and cannot be predicted in advance. They are, rather, a component of the residual uncertainty that the safety stock, introduced in Section 4.5.3, is designed to absorb. For the purpose of sizing the dynamic stock requirement, the budgeted quantity is taken as the best available estimate of total project consumption.

4.4 Temporal Phasing of Demand

The quantity map provides the total quantity of each material a project will consume, but not its distribution over time. Yet it is this distribution that governs the stock requirement, since a material is not needed in the warehouse throughout a project's duration but only in the period preceding the phase that consumes it. The purpose of this section is to characterize, for each Class B material, the phase of the project lifecycle at which it is consumed, so that its known total can be allocated to the appropriate point in time. This characterization is the central modelling element of the proposed approach, and it rests on a single empirical premise that the temporal profile of a material's consumption is sufficiently stable across projects to be predictive.

4.4.1 Positioning Consumption on the Project Timeline

Projects in the company's portfolio differ widely in duration, from interventions of a few weeks to installations spanning more than a year. Elapsed calendar time in absolute terms is therefore not a meaningful basis for comparing the consumption of a material across projects, since a given month corresponds to an early stage in a long project and a late stage in a short one. To render projects comparable, the duration of each project is normalized to a relative timeline running from its start (0%) to its end (100%), and each consumption event is positioned as a fraction of the elapsed project duration at which it occurred.

For completed projects, used to calibrate the model, this elapsed fraction is computed from the recorded start and end dates of the project, against which the dates of the corresponding deliveries or stock movements are positioned. For active projects, to which the calibrated model is applied, the planned duration recorded in the project budget is used to determine the phase the project has currently reached. It should be made explicit that this measure is based on elapsed time rather than on physical progress, which is not recorded in an analyzable form. The approach therefore carries the assumption that the temporal rhythm of consumption is broadly comparable across projects. Since the purpose of the model is ultimately to anticipate *when* a material will be required, an elapsed-time basis is appropriate to that end, and the validity of the resulting profiles is assessed empirically in Chapter 5.

4.4.2 Constructing the Phase Weights

The normalized timeline of each project is partitioned into discrete phases. In this work, four phases of equal width, each spanning 25 % of the project duration, are adopted. This choice reflects a deliberate balance between resolution and robustness as a finer partition would yield a more detailed consumption profile but would distribute the limited historical record across more intervals, producing noisier and less reliable weights, whereas four phases retain enough events per interval to estimate each weight stably given the available data. The four quarters also correspond intuitively to recognizable stages of project execution, broadly associated with site preparation, structural and rough-in work, installation, and finishing, which aids the interpretation of the resulting profiles.

For a given material i and a single historical project p, the proportion of that material's total consumption falling within each phase is determined from the consumption events positioned on the project's normalized timeline. The characteristic profile of material i is then obtained by combining the profiles of all historical projects in which it was consumed. The profiles are combined as a weighted average, each project weighted by its total consumption of the material, so that projects consuming the material in larger quantities, being more representative of its consumption behaviour, exert proportionally greater influence than minor ones. This yields the *phase weight* $W_{i,\varphi}$:

$$W_{i,\varphi} = \frac{\sum_p q_{i,p,\varphi}}{\sum_p Q_{i,p}} \tag{4.1}$$

where $q_{i,p,\varphi}$ is the quantity of material i consumed by project p during phase φ, and $Q_{i,p}$ is the total quantity of material i consumed by project p. By construction the weights sum to unity across the four phases, $\sum_\varphi W_{i,\varphi} = 1$, so that $W_{i,\varphi}$ represents the fraction of any project's total requirement of material i expected to be consumed during phase φ. The discrete weights may be understood as a discretization of the cumulative S-curve of the material's consumption over the project lifecycle.

4.4.3 The Stability Premise

The use of a single characteristic profile $W_{i,\varphi}$ to represent a material across all projects rests on the premise that this profile is reasonably stable from one project to another. This premise is plausible on engineering grounds, since the sequence of construction activities, and hence the order in which materials are required, is broadly common across projects of a given type. It is not, however, assumed without verification. The degree to which the consumption profiles of a material are consistent across projects, and therefore the predictive validity of the phase weights, is an empirical question, and is examined directly in the validation of Chapter 5. Where a material's profile proves unstable, the weights carry correspondingly less information, and the residual uncertainty is absorbed, as for the quantity deviations of Section 4.3, by the safety component of the dynamic

stock. The phase weights established here are the second input, alongside the project quantities, to the aggregation developed in the following section.

4.5 Dynamic Minimum Stock Aggregation

The preceding two sections established, for each Class B material, the total quantity required by each project and the fraction of that total consumed in each phase of the project lifecycle. This section combines these elements across the active project portfolio to determine the stock that must be held at the present moment, and adds the safety component that absorbs the residual uncertainties identified along the way. The result is the dynamic minimum stock, a quantity that adjusts continuously to the composition and progress of the active portfolio.

4.5.1 Expected Consumption over a Coverage Horizon

The phase weight $W_{i,\varphi}$ describes consumption *during* a phase, whereas inventory must be available *before* the consuming phase is reached, with sufficient lead time for replenishment. The dynamic stock requirement is therefore framed in terms of a forward coverage horizon. The stock held at any moment must cover the consumption that the active projects are expected to generate over a period extending far enough ahead to accommodate both the supplier lead time and the review interval. For each active project, the phase or phases falling within this horizon, given the project's current position on its normalized timeline, determine the share of its total requirement that must be covered.

Formally, for material i and active project p at time t, let $\varphi(p,t)$ denote the set of phases the project is expected to traverse within the coverage horizon, and let $W_{i,\varphi(p,t)}$ denote the corresponding cumulative phase weight. The expected consumption of material i attributable to project p over the horizon is then the product of the project's total requirement and this weight:

$$C_{i,p,t} = Q_{i,p} \times W_{i,\varphi(p,t)} \tag{4.2}$$

where $Q_{i,p}$ is the total quantity of material i foreseen for project p, taken from its quantity map. Projects contribute to the current requirement according to their position in the lifecycle. Projects whose consuming phase is beyond the horizon contribute little or nothing, while those approaching it contribute most of their requirement. The requirement therefore shifts automatically as projects advance.

4.5.2 Aggregation across the Active Portfolio

The total project-driven requirement for material i at time t is obtained by summing the expected consumption over all projects currently active:

$$D_{i,t} = \sum_{p \in \text{active}} C_{i,p,t} = \sum_{p \in \text{active}} Q_{i,p} \times W_{i,\varphi(p,t)} \tag{4.3}$$

This aggregate captures the central insight of the model, as material demand at the central warehouse is determined by the combined phase-dependent requirements of all active projects and therefore changes as the portfolio evolves, even though each project's total consumption is fixed and known in advance.

4.5.3 The Safety Component

The aggregate $D_{i,t}$ represents the expected requirement over the coverage horizon. Because it is built from the known project quantities and their characteristic phasing, the model anticipates the bulk of demand deterministically, in contrast to the classical approach, which treats the entire demand as a stochastic quantity to be buffered. The role of the safety component is therefore confined to the uncertainty that remains after this anticipation, rather than the full variability of demand. Applying the classical safety-stock formula of Equation 2.2 to the raw, lumpy demand series would, as discussed in Section 2.3.3, inflate the buffer to unrealistic levels because the formula must absorb demand peaks that the present model instead predicts.

Residual uncertainty arises from three sources, with supplier lead-time variability being the principal one. Even when the requirement is correctly anticipated, a late delivery can cause a stockout. This source is also the most directly quantifiable, since the lead time of each order is recorded as the interval between ECF issuance and GUIA, allowing its mean and variability to be estimated per supplier or material family from the ERP. The safety component is therefore dimensioned primarily against lead-time variability, in the spirit of the classical formulation, but with lead time rather than demand as the stochastic element:

$$SS_i = Z \cdot \bar{D}_i \cdot \sigma_L \tag{4.4}$$

where Z is the service-level factor, σ_L the standard deviation of the lead time, and $\bar{D}_i$ the expected demand rate of material i over the lead time. The second and third sources of uncertainty, namely the project-by-project deviation between budgeted and actual consumption (Section 4.3) and any instability in the phase profile across projects (Section 4.4.3), are less amenable to direct parameterization with the available data. They are accommodated by the choice of service-level factor Z, a higher value of which raises the buffer to reflect a greater combined residual uncertainty, rather than by a separate closed-form term whose calibration the data would not support.

Combining the anticipated requirement with the safety component yields the dynamic minimum stock for material i at time t:

$$DMS_{i,t} = \underbrace{\sum_{p \in \text{active}} Q_{i,p} \times W_{i,\varphi(p,t)}}_{\text{anticipated requirement}} + \underbrace{SS_i}_{\text{residual buffer}} \tag{4.5}$$

This expression is the central output of the model. Its first term is deterministic and project-driven, computed from information the organization already possesses, namely the quantity maps of its active projects and their current progress, combined with the consumption profiles calibrated from history. Its second term is a comparatively small buffer against residual uncertainty, dominated by supplier reliability. The quantity $DMS_{i,t}$ is recomputed as the portfolio changes, providing a stock target that is neither the static, residual-only holding of the current state diagnosed in Chapter 3, nor the inflated buffer that a naive statistical treatment of lumpy demand would prescribe, but a requirement matched to the actual, evolving needs of the active projects.

Figure 4.2 illustrates how this behaviour would unfold for a given set of active projects, with the aggregate requirement of a family rising and falling as projects enter the portfolio and progress through their phases.

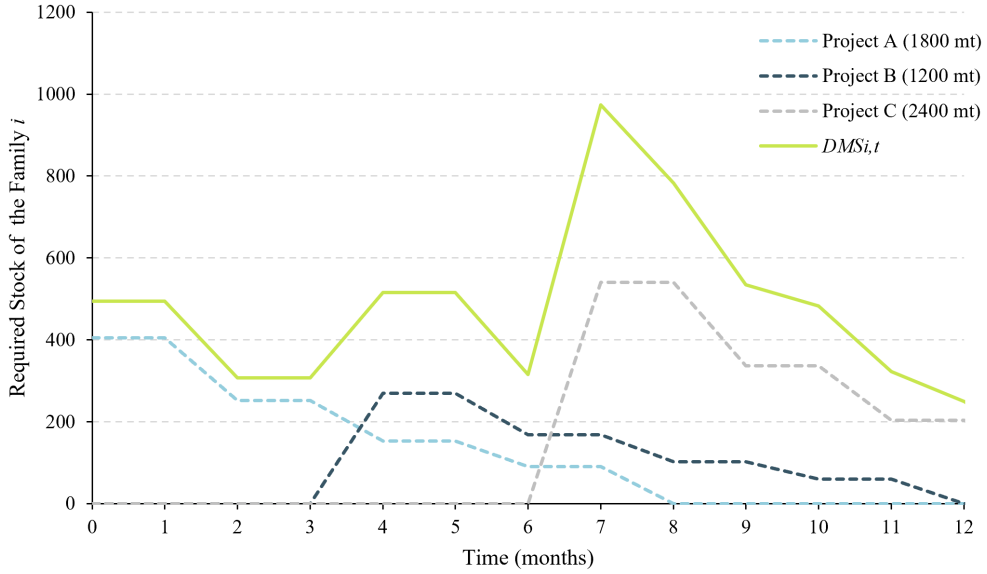


Figure 4.2: Illustration of the dynamic minimum stock of a family over time. Projects entering at different moments would cause variations in the aggregate stock requirement (solid line) is the superposition of their phase-dependent contributions, rising when active projects coincide in a consuming phase and falling as they pass it. The temporal positioning of the projects is illustrative of the mechanism.

The curve shown is a model of how the requirement would evolve for a given entry of project. In operation it is not fixed but varies continuously, as the arrival and completion of projects, or the

progress of an existing one into a consuming phase, reshape the aggregate in real time. The stock target therefore tracks the portfolio as it changes, rather than being set once and held.

4.6 Enabling Implementation Layer

The model developed in the preceding sections depends on two practical conditions that the current system, characterized in Chapter 3, does not satisfy. It requires reliable historical consumption data with which to calibrate the phase weights and the lead-time statistics, and it requires a channel through which its outputs, and the underlying stock and order status, reach the operational staff who act on them. This section describes two components implemented to meet these conditions. An automated data-capture mechanism integrated with the existing ERP, and a monitoring dashboard. They are the enabling layer that renders the model operable and, in their own right, address the data-reliability and visibility deficiencies (P2, P3, P4) identified in the diagnosis. They are presented here with emphasis on the design decisions that adapt them to the constraints of a small enterprise operating a closed, legacy ERP.

4.6.1 Order-Driven Data Capture and ERP Integration

The reliability of inventory data is determined at the point of capture. In the current process, the reception and registration of materials are manual, and it is here that the errors and delays undermining stock visibility originate. The remedy adopted is automated identification through QR-code scanning, by which articles are registered at reception, and subsequent movements recorded, without manual transcription. Its implementation is constrained by a characteristic obstacle of the small and medium-enterprise context, as the central ERP, PRIMAVERA, is a closed, proprietary system whose direct modification is costly, requires vendor intervention, and risks compromising the integrity of financial and fiscal records.

The design resolves this tension through a *decoupled, read-only architecture*. Rather than writing to or altering the ERP, a separate, lightweight application manages the operational layer of QR associations and movements in its own database, while querying the ERP strictly in read-only mode to retrieve master data. This separation is the principal design decision of the component as it allows a modern capture layer to be introduced at the warehouse floor without any intervention in the core enterprise system, obtaining the benefits of automation while guaranteeing that accounting and fiscal data remain untouched. The pattern, in which a low-cost operational layer is superimposed on an unmodified legacy platform, is recognized in the literature as an appropriate integration strategy for resource-constrained enterprises [34, 49].

A distinctive feature of the reception workflow, illustrated in Figure 4.3, is that it is driven by the purchase order rather than by the individual article. Because every delivery is accompanied by

a delivery note GUIA bearing the reference of the corresponding ECF, the operator begins by selecting that ECF in the application, whereupon the expected list of articles and ordered quantities is retrieved from the ERP in read-only mode. The operator then scans each received article in turn, and the system records the scanned quantity against the appropriate ECF line. This order-driven design turns the reception into a reconciliation of received against ordered quantities, rather than a blind identification of items, and on completion the application presents a summary of the order, listing each article with its designation, ordered quantity, and scanned quantity, before the reconciled information is saved. The functioning prototype that implements this workflow is presented in Appendix B.

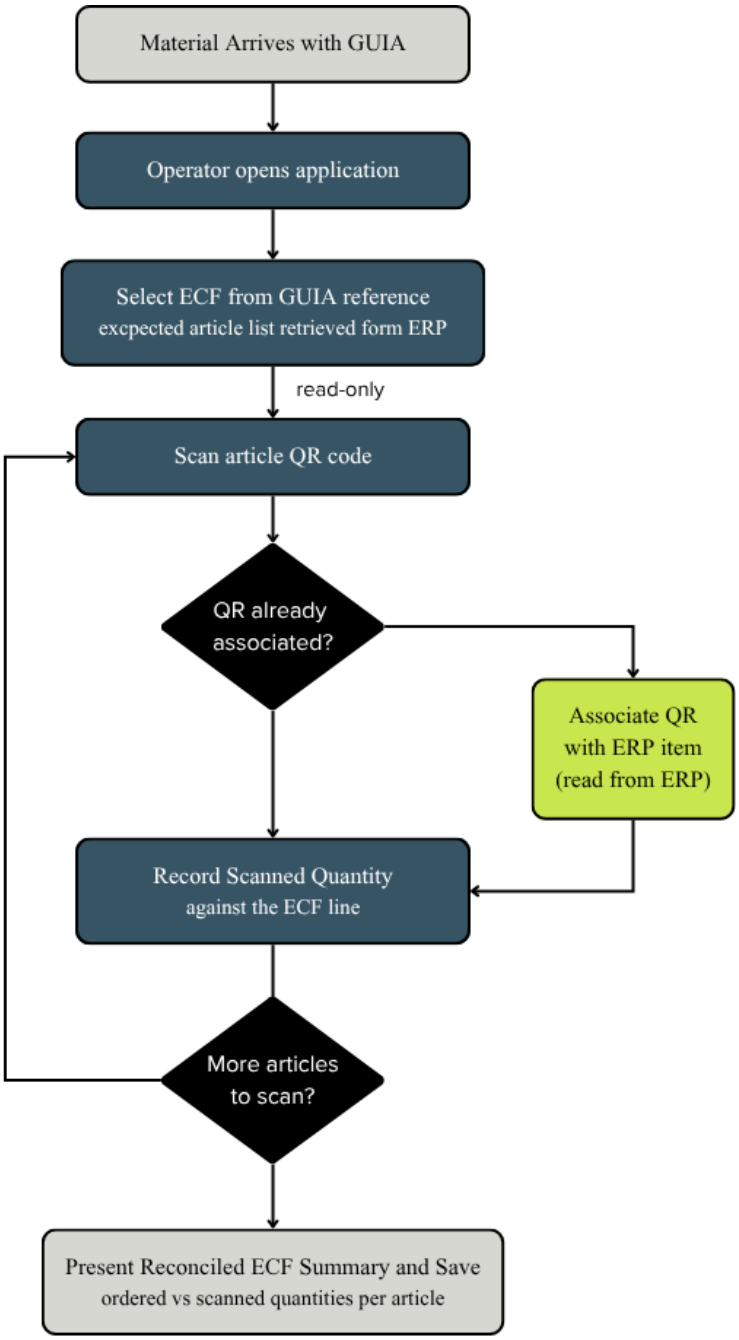


Figure 4.3: Order-driven QR reception workflow. The operator selects the ECF identified from the delivery note, scans each article against the expected order lines, associates any unrecognized code with its ERP item, and saves a reconciled summary of ordered versus received quantities.

The communication between the application and PRIMAVERA is confined to a single, one-way channel, shown in Figure 4.4. A synchronization routine reads ECF and article master data from the ERP through its standard read access, and writes them into the local database of the application, keeping the latter's copy current. All operational data generated by the application, the associations between physical QR codes and ERP items, and the records of scanned receptions and movements, reside exclusively in this local database and are never written back to the ERP.

For the purposes of this dissertation, this boundary was a deliberate scope decision. By exercising only the read capability of the production system and isolating all write operations within a separate database, the prototype could be developed and tested against real company data without any risk to the integrity of the live financial and operational records. The integration of the captured movements back into the ERP, which would be required for a full production deployment, is consequently left as an explicit item of future work rather than attempted within the scope of this work.

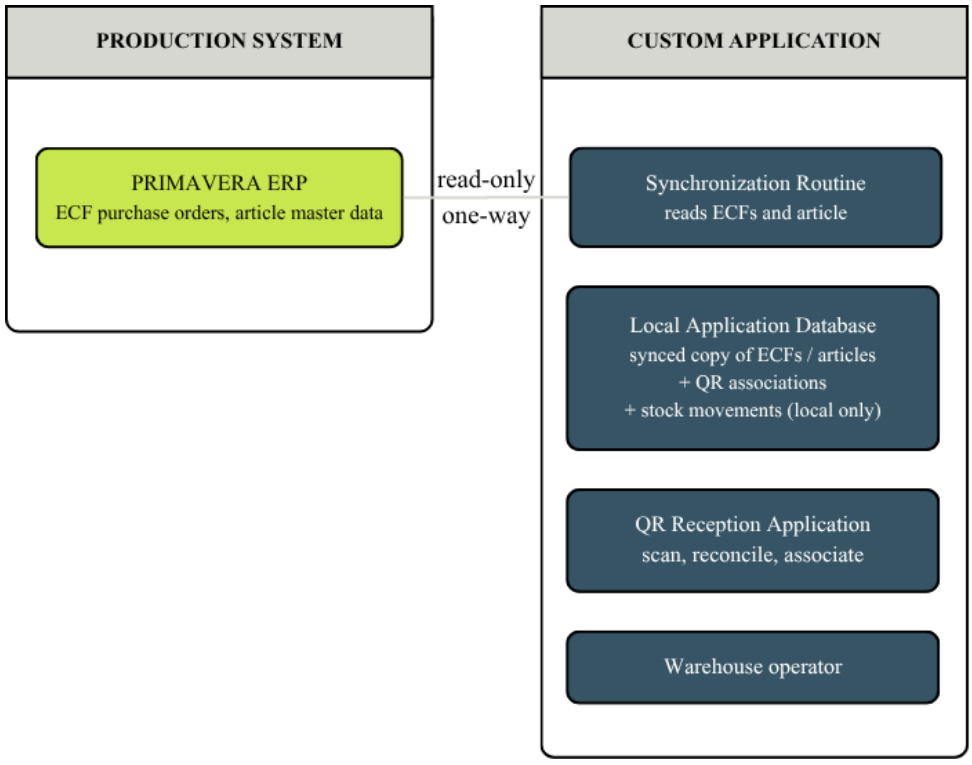


Figure 4.4: Decoupled read-only architecture. PRIMAVERA is accessed only through a one-way synchronization routine that copies ECF and article data into the local application database; QR associations and stock movements are held locally and are never written back to the production ERP.

Within this workflow, the association between a physical QR code and its ERP item arises only as an exception. Items are created in the ERP by administrative or engineering staff before physical procurement, at which point their physical QR code is unknown. When a scanned code is not yet recognized, the operator associates it, in that moment, with the corresponding ERP item retrieved from the order. Thereafter the association is stored locally and every subsequent scan of that code is recognized automatically. A one-time manual association thus yields permanent automated recognition, progressively building a reliable bridge between the physical materials and their digital records without ever writing to the ERP.

4.6.2 The Monitoring Dashboard as Decision-Support Layer

The second component of the enabling layer addresses visibility. As established in the diagnosis, procurement information resides in the ERP but is accessible only to a limited number of users, leaving site workers without an autonomous view of pending orders, expected deliveries, or order status, and reducing order follow-up to manual and reactive communication. The monitoring dashboard, developed in Power BI, remedies this not by replacing the ERP but by acting as a visualization layer over it, so that data extracted from PRIMAVERA are processed and presented in a form adapted to the operational needs of each role.

Conceptually, this component is best understood not as a reporting tool but as the *dialog component* of a decision-support system, in the terms introduced in Section 2.6. The model of this chapter constitutes the model component of that system, and the ERP data its data component, while the dashboard is the interface through which both become operationally useful. Its purpose extends along two complementary directions. The first is operational visibility, the consolidated display of pending orders, expected deliveries, and order fulfilment status that replaces fragmented manual follow-up and addresses the access deficiencies (P3, P4) identified in the diagnosis. The second is performance measurement, the quantification through indicators of the procurement and supplier functions, which addresses the structural absence of measurement (P5) identified in Section 3.4.4.

The prototype developed comprises three pages, whose full views are given in Appendix C. The first, an overview of order status, classifies every purchase order by its fulfilment state, fully received, partial, or open, and highlights those that are delayed, defined here as orders not yet fully received more than fifteen days after they were placed. This definition is deliberately grounded in the order and receipt dates, which are reliable, rather than in the promised-delivery date, which the diagnosis found to be an unpopulated default and therefore unusable. The overview resolves directly the absence of an autonomous view of order status. The second page details the orders themselves, order by order, with their degree of fulfilment and their pending value, so that the material still to arrive can be traced and prioritized. The third page characterizes the suppliers, and is the page that most directly serves the model. It positions each supplier in a plane of average lead time against lead-time variability, with the size of each point reflecting the value ordered, so that the suppliers whose deliveries are both slow and erratic, the ones that most threaten a stockout, are identified at a glance. It is the lead-time variability read from this page that feeds the safety component of the minimum stock developed in Section 4.5.3.

Figure 4.5 shows the overview page as an illustration, the full set of pages being given in Appendix C.

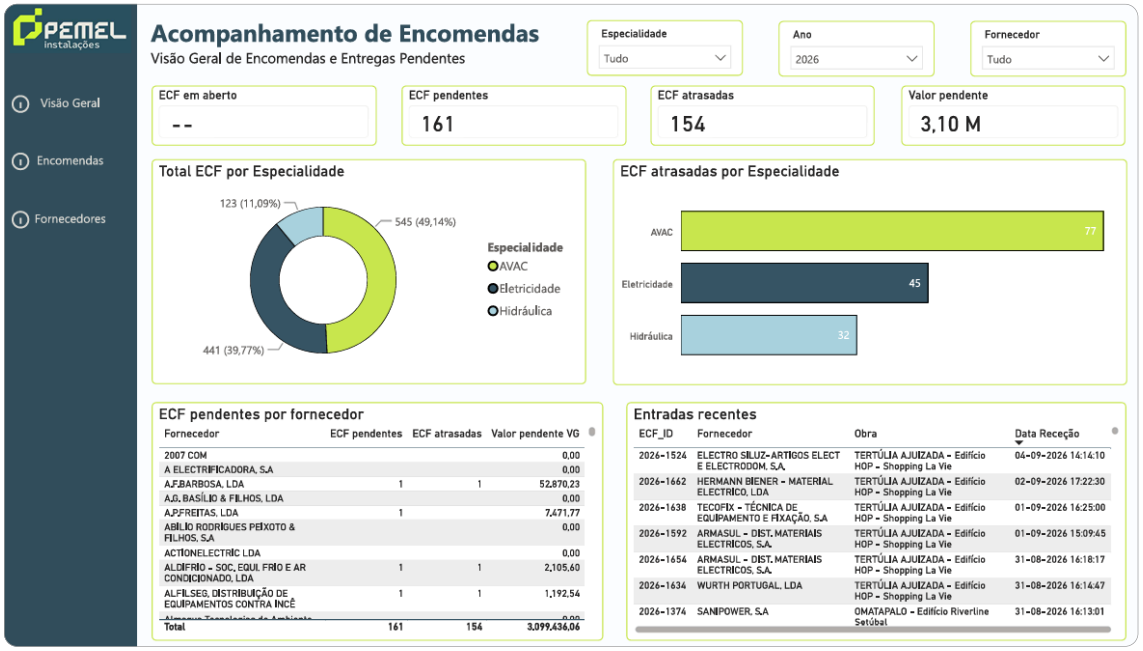


Figure 4.5: Overview page, showing the order-status indicators, the highlighted delayed orders, the breakdown by speciality, and the pending-by-supplier and recent-receptions tables.

In its present form the prototype is fed by periodic extractions from the database. Because it rests on direct SQL queries over the SQL Server instance of the PRIMAVERA system, however, its migration to a continuously refreshed layer, through a direct connection with scheduled updates, is immediate and requires no change to the architecture, the queries already written being the same in either mode. The prototype thus demonstrates the capability while leaving the step to live operation as a matter of configuration rather than of design.

The contribution of this component is therefore twofold. Methodologically, it is the channel that closes the loop between the model's output and operational action, without which a quantitative stock recommendation would remain inaccessible to those it is intended to guide. Operationally, it institutes the visibility and the measurement that the organization previously lacked, and in doing so it also makes visible, through the indicators it could not populate, the precise gaps in data capture that the remainder of the enabling layer sets out to close.

Chapter 5

Empirical Validation and Results

This chapter examines the model against the company's historical record. It sets out the validation methodology, including the reconstruction of family-level consumption on which the analysis rests, then examines the material classification and the two premises of the model, that demand is lumpy (H1) and that the phase profiles are stable enough to be predictive (H2). It characterizes the supplier lead time and its variability, and demonstrates the model on a case study, its forecasting accuracy against the classical benchmarks (H3) and the inventory trade-off it supports (H4). The results are read as a demonstration of feasibility, with the data limitations made explicit.

5.1 Validation Methodology

The empirical work reported in this chapter rests on data drawn directly from the company's Enterprise Resource Planning system. This section describes how that data was accessed, extracted and treated, and explains the reconstruction procedure that made it possible to observe material consumption at the level of functional families. The same data foundation underlies the examination of both premises in Sections 5.3 and 5.4 and the supplier analysis in Section 5.5, and it is therefore set out once, here, before the results.

5.1.1 Data Sources and Extraction

The company operates on the PRIMAVERA ERP platform, whose transactional data is held in a Microsoft SQL Server database. Access for this work was granted over an isolated copy of the production database, so that no query could affect live operations, with read-only permissions. All extraction was performed through Microsoft SQL Server Management Studio (SSMS), by means of SQL queries written to retrieve the specific tables and fields required for each analysis, which were then exported to CSV files for treatment. Working against a copy also meant that the data reflects the state of the company's records up to the moment the copy was taken, a horizon that is stated where it bears on the interpretation of a given result.

The relevant records are distributed across several modules of the database. Material consumption is recorded in the goods-issue documents that move material from the warehouse to a project. Purchasing and reception are held in the procurement module, separate from the sales and project-issue documents. Project progress is recorded in the monthly measurement statements. The article catalogue carries the family and sub-family classification together with the average and last purchase costs. Linking these modules required following the foreign-key relationships between the document headers, the document lines, the article table and the project table, which were confirmed empirically during the extraction.

The treatment of the extracted CSV files was carried out in Python, using the pandas and numpy libraries. Treatment comprised the parsing and normalization of dates and of numeric fields, the removal of placeholder and null records, the reconciliation of the several extractions into coherent datasets, and the computation of the metrics reported in the remainder of the chapter, namely the demand-classification statistics, the phase profiles, the stability indices and the supplier lead times. Where a field was found not to carry reliable information it was excluded from the analysis rather than used, and each such exclusion is stated at the point where it arises.

The two scripts that carry out this treatment, the ABC classification and the phase-profile and stability analysis, are reproduced in full in Appendix D.

5.1.2 Reconstructing Family-Level Consumption

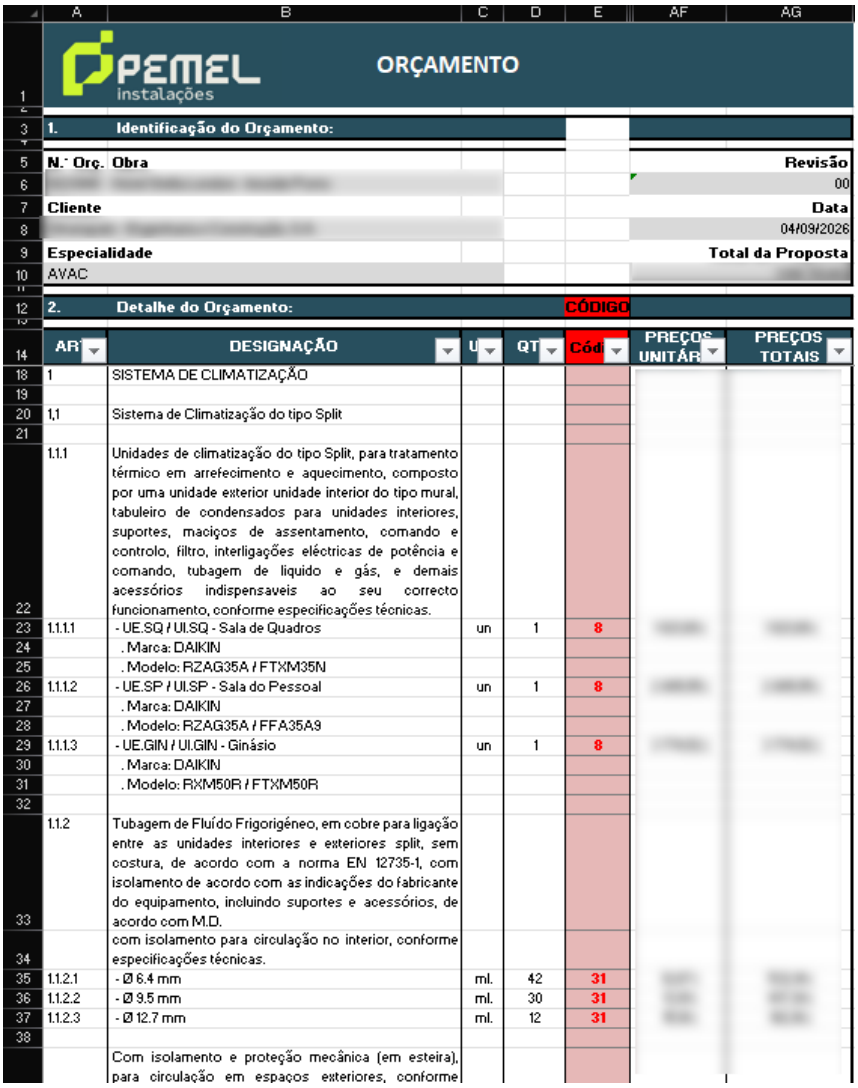
The central methodological difficulty of the validation, and the one whose solution enables the analysis of phase profiles, concerns the level at which consumption can be observed. The proposed model reasons in terms of functional families of material, the engineering groupings such as ductwork, valves or copper piping into which a budget is organized, whereas the transactional records identify only individual articles or, in the case of the measurement statements, individual budget line items. Bridging these two levels proved not to be possible through the database alone.

The company's budgets classify every article by functional family. Figure 5.1 shows the family sheet of a project budget, in which the first thirty-two HVAC families are standard across all budgets, each identified by its designation, while any additional families are budget-specific.

Figure 5.2 shows the corresponding detail sheet, in which each priced article is set against a hierarchical position and carries, in a dedicated column, the number of the family to which it belongs. This budget structure is the source from which the family of any article is known, and it exists in the spreadsheet budgets rather than in the ERP database.

	A	B	C
1	N.º Orç.	[redacted]	00
2	Obra :	[redacted]	AVAC
3	Cliente:	[redacted]	
4			
5	Cod.	Designação	Marca
6	1	Caldeiras	[redacted]
7	2	Aquecedores de Ar	
8	3	Bombas de Calor	
9	4	Ventiladores	
10	5	Unidades de Tratamento de Ar (UTA)	
11	6	Exaustores	
12	7	Unidades de Janela	
13	8	Splits	
14	9	Sistemas VRF/VRV	
15	10	Chillers	
16	11	Termostatos	
17	12	Equipamento de campo	
18	13	Sistemas de Gestão de Edifícios (BMS)	
19	14	Caminho cabos	
20	15	Tubos de PVC	
21	16	Difusores e Grelhas	
22	17	Filtros de Ar	
23	18	Válvulas	
24	19	Tubos	
25	20	Caixas	
26	21	Alimentadores	
27	22	Cabos	
28	23	Calha Rodapé	
29	24	Quadros Elétricos	
30	25	Condutas Spiro	
31	26	Condutas retangulares	
32	27	Grelhas intumescentes	
33	28	Registos de caudal	
34	29	Registos Corta Fogo	
35	30	Registos de desenfumagem	
36	31	Cobre de refrigeração	
37	32	Cobre de Aquecimento	
38	33	Cabos	
39	34	Solar Térmico	

Figure 5.1: Family sheet of a project budget, listing the thirty-two HVAC functional families with their designations and reference brands. Commercially sensitive values and identifiers are blurred. This catalogue defines the functional families used as the unit of analysis.



	A	B	C	D	E	AF	AG
1	PEMEL instalações	ORÇAMENTO					
3	1.	Identificação do Orçamento:					
5	N.º Orç.	Obra					Revisão
6							00
7	Cliente						Data
8							04/09/2026
9	Especialidade						Total da Proposta
10	AVAC						
12	2.	Detalhe do Orçamento:			CÓDIGO		
14	ART	DESIGNAÇÃO	U	QT	Cód	PREÇOS UNITÁR	PREÇOS TOTAIS
18	1	SISTEMA DE CLIMATIZAÇÃO					
19							
20	1,1	Sistema de Climatização do tipo Split					
21							
22	1.1.1	Unidades de climatização do tipo Split, para tratamento térmico em arrefecimento e aquecimento, composto por uma unidade exterior unidade interior do tipo mural, tabuleiro de condensados para unidades interiores, suportes, maciços de assentamento, comando e controlo, filtro, interligações eléctricas de potência e comando, tubagem de liquido e gás, e demais acessórios indispensaveis ao seu correcto funcionamento, conforme especificações técnicas.					
23	1.1.1.1	- UE.SQ / UI.SQ - Sala de Quadros	un	1	8		
24		. Marca: DAIKIN					
25		. Modelo: RZAG35A / FTXM35N					
26	1.1.1.2	- UE.SP / UI.SP - Sala do Pessoal	un	1	8		
27		. Marca: DAIKIN					
28		. Modelo: RZAG35A / FFA35A9					
29	1.1.1.3	- UE.GIN / UI.GIN - Ginásio	un	1	8		
30		. Marca: DAIKIN					
31		. Modelo: RXM50R / FTXM50R					
32							
33	1.1.2	Tubagem de Fluído Frigorigéneo, em cobre para ligação entre as unidades interiores e exteriores split, sem costura, de acordo com a norma EN 12735-1, com isolamento de acordo com as indicações do fabricante do equipamento, incluindo suportes e acessórios, de acordo com M.D.					
34		com isolamento para circulação no interior, conforme especificações técnicas.					
35	1.1.2.1	- Ø 6.4 mm	ml.	42	31		
36	1.1.2.2	- Ø 9.5 mm	ml.	30	31		
37	1.1.2.3	- Ø 12.7 mm	ml.	12	31		
38							
		Com isolamento e proteção mecânica (em esteira), para circulação em espaços exteriores, conforme					

Figure 5.2: Detail sheet of a project budget. Each priced article is set against a hierarchical position and carries, in the highlighted column, the number of the functional family to which it belongs. This is the source from which the family of each article is read for the price-based reconstruction. Commercially sensitive values are blurred.

The obstacle is that the line items of the measurement statements, which carry the real dated record of what was executed in each month, are not linked in the database, by any accessible key, to the article catalogue that would carry the family, nor to the hierarchical structure of the budget. The two are reconciled only within the application layer of the commercial software and not in a form that can be queried. Every attempt to join the measurement items to the budget structure through the documented identifiers failed, because the internal identifier of a measurement item does not correspond, by value or by type, to the keys of the budget-object and budget-structure tables. This fragmentation is itself a finding of some importance, since it demonstrates that the data the model requires, although present in the system, is not readily reconcilable within it, which is part of the case for the structured data-capture layer proposed in Chapter 4.

The link was ultimately established indirectly, through the unit selling price. Each measured line item carries a unit price that is very nearly unique within a project, and the same price appears against the corresponding priced article in the project budget, from which the functional family follows. Matching the measurement items to the budget articles by unit price therefore recovers, for each item consumed in each dated statement, the functional family to which it belongs. Applied to the pilot projects, this procedure achieved a match rate of 97 % to 100 %, with the few unmatched items corresponding to prices that were genuinely repeated within a project and could not be disambiguated. The correspondences obtained were verified to be physically coherent, with matched articles falling into families consistent with their descriptions. This price-based reconstruction is the mechanism by which family-level consumption is observed throughout the remainder of the chapter.

5.1.3 Scope and Standard of the Validation

Two boundaries on the validation follow from the nature of the data and are stated here so that the results of the chapter are read in the correct light. The first concerns the temporal anchor of a project. The nominal project dates recorded in the ERP were found, as discussed in Chapter 3, to be automatic placeholders in the great majority of cases rather than real start and end dates, a condition confirmed to persist in the current system. The timeline of a project is therefore taken from the sequence of its dated measurement statements, which provides a real progress curve where the nominal dates do not. The second concerns the number of projects available with a reliable timeline and a fully reconstructable consumption, which is small. In keeping with the position established in Section 1.5, the validation is accordingly conducted as a demonstration of feasibility supported by quantitative evidence rather than as a claim of statistical significance, and the strength of each result is qualified where the available sample warrants it.

5.2 The Value Dimension: ABC Classification Across Projects

The scope of the proposed model was defined in Section 4.2.3 as the intersection of the Class B value group with the lumpy demand category, and the value classes themselves were set out conceptually in Table 4.1, which assigns major equipment to Class A, bulk installation material to Class B, and low-cost consumables to Class C. This section examines that value classification empirically, in order to establish which families are in fact economically significant before the demand pattern and the phase stability of those families are examined in the sections that follow.

The ABC analysis was conducted at the level of the project rather than on the aggregate consumption of all projects combined. Aggregating the consumption into a single pool would allow one or two large projects to dominate the ranking, so that a family important in a single contract

would appear significant overall even where it is marginal in most. Classifying each project separately respects instead the project-based nature of the operation, in which each project is a distinct unit of demand, and it allows the question to be posed of whether a family is recurrently significant across projects rather than significant on one occasion.

Six of the thirteen projects draw on older budget sheets to which the functional-family classification was added by the author, a reconstruction described, and tested for bias, in Section 5.4. These projects are marked with an asterisk.

Figure 5.3 shows, for each family, the proportion of the projects in which it was classified A, B or C, ordered from the most to the least economically significant.

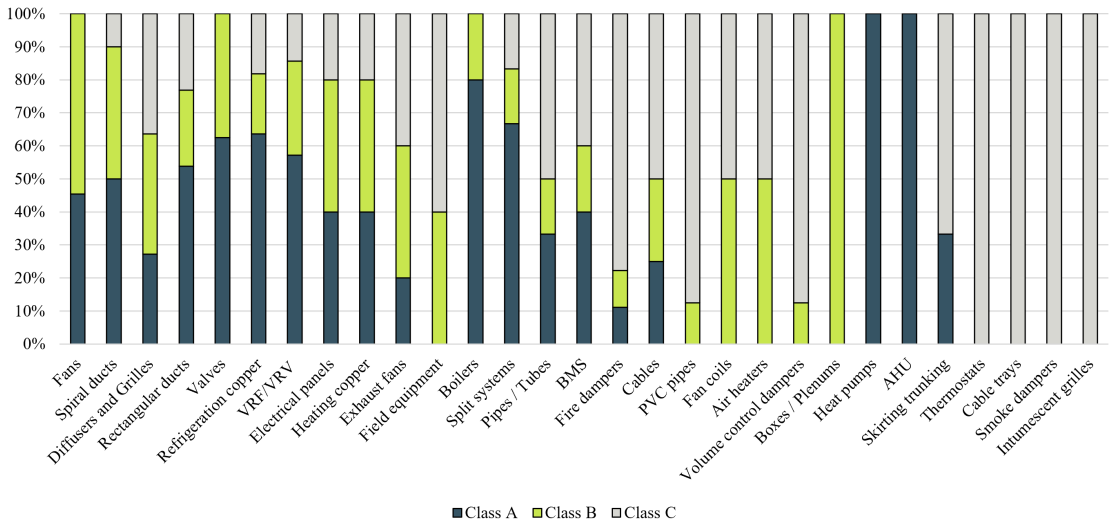


Figure 5.3: Proportion of projects in which each HVAC family was classified A, B or C, across the thirteen projects, ordered from the most to the least economically significant.

The empirical classification confirms the structure of Table 4.1 in its essentials. The major equipment families are recurrently Class A, with heat pumps, air-handling units classified A in every project in which they appear, which is consistent with their assignment to the equipment group whose demand is tied to identifiable project milestones. The low-value families are recurrently Class C, with PVC piping, cable trays and skirting trunking into Class C in the large majority of projects. Between these two ends lie the bulk installation families such as the ductwork, copper and cabling which occupy the intermediate A/B region that Table 4.1 associates with Class B.

The underlying classification, family by family and project by project, is given in Figure 5.4, where the reconstructed-budget projects are marked with an asterisk.

Family	20222007*	20222014*	20232002*	20232004	20232005*	20232011	20242003*	20242004*	20242006	20242012	20252004	20252005	20252006
Pipes / Tubes	A	A				C		C		B		C	
Valves	A	B		A	A			B	A		A	B	
Fans	A	B	B	B		A		A	A	B	B	A	B
Refrigeration copper	A	A	A	B	A	A		A	A		B	C	C
Boilers	A			A					B		A		
Rectangular ducts	B	A	B	A	C	B	A	A	C	A	A	A	C
Spiral ducts	B	A	C	A	A	B	A	A	A		B	B	B
Field equipment	B	C		C	B	C						C	
Exhaust fans	C	C			B		B					A	C
BMS	C	C		B				C			A	A	
Diffusers and Grilles	C	A	A	A	B	B	B	B	C	C	A	A	C
PVC pipes	C	C	C	C		C		C	B		C	C	
Heat pumps		A				A		A		A			A
VRF/VRV		A	A	A			C	B	B		A		
AHU		A					A			A		A	
Electrical panels		A			A	B		B			C		
Fan coils		B									C		
Fire dampers		B	C	C	C	C	C	C	C	A			
Air heaters		C		B									
Thermostats		C											
Volume control dampers		C	C	C	C	C	B	C	C	C			
Cable trays			C		C	C			C				
Smoke dampers			C										
Skirting trunking				A	C							C	
Heating copper				C	B			A	A		B		
Intumescent grilles				C		C		C	C				
Split systems					A	A		A	A			B	C
Boxes / Plenums					B								
Cables							A	C			C	B	

Figure 5.4: ABC class of each functional family in each project. Class A in dark blue, Class B in green, Class C in grey, and blank where the family is absent from the project. Projects with a reconstructed budget are marked with an asterisk.

The reading of the figure must take account of how many projects each family appears in, since a family present in only one or two projects offers weaker evidence than one present in many, and the families toward the right of the figure that show a single solid class are for the most part those that appear in few projects. Restricting attention to the families present in at least six of the thirteen projects, for which the evidence is more robust, the same three-way structure holds, with the equipment families dominated by Class A, the installation families spread across A and B, and the cheaper accessories anchored in Class C.

One feature of the intermediate group deserves comment, because it bears on the boundary between Class A and Class B. Several installation families that Table 4.1 would place in Class B, most notably the spiral and rectangular ductwork, carry a substantial Class A share across the projects. This is in part a consequence of how the budgets aggregate cost. Where a duct is budgeted together with its thermal insulation or its fire-resistant coating in a single line item, the unit value of that item rises well above that of the bare duct, and the family is displaced toward Class A. The value class of these families is therefore sensitive to a budgeting decision rather than to a stable property of the material alone, which is itself consistent with the position taken in Section 4.2.3 that the value and demand dimensions are independent and that the boundary between value classes is not sharp. It also indicates that the A/B distinction, when drawn per project, is fluid for the installation families, so that the group of economically significant installation material

is better read as the recurrent A/B region taken together than as a rigid separation of the two classes.

Taken as a whole, the per-project ABC analysis supports the value classification of Table 4.1 and identifies, with empirical backing, the installation families that are the economically significant candidates for the model. It is on these families, once their demand pattern and phase stability have been established in the sections that follow, that the intersection defining the scope of the model is drawn.

5.3 The Demand Regime: Testing the Lumpiness Premise (H1)

The first premise on which the proposed model rests is that the installation materials of interest do not follow the smooth, regular demand for which the classical reorder-point methods were devised, but rather the intermittent and lumpy demand characteristic of project-based operations (H1). This section examines that premise directly against the historical consumption record.

The test draws on the complete history of material issued from the warehouse to projects over the period covering May 2020 to September 2026, a horizon of seventy-seven months. For each article, the monthly issued quantity was reconstructed from the transactional records, and two statistics were computed following the classification scheme of Syntetos, Boylan and Croston introduced in Section 2.4. Each article was then assigned to one of the four quadrants defined by the conventional cut-offs ADI $= 1.32$ and $CV^2 = 0.49$.

Of the 6,155 distinct articles observed over the horizon, 3,766 (61 %) appeared in only a single month, or not at all in a form permitting the computation of an interval. These are single-use items, requisitioned once for a specific project and never again. This figure is itself a finding about the nature of the operation, and it is returned to below. The classification was performed on the 2,389 articles that recorded demand in at least two separate months, for which an ADI and a CV^2 are defined.

The demand for installation materials is, to an overwhelming degree, outside the regime for which the classical models are suited. Figure 5.5 plots every classified article in the ADI–CV^2 plane. The distribution is unambiguous, with 99.7 % of the articles falling to the right of the intermittency threshold and not a single article exhibiting the low-ADI, low-CV^2 signature of smooth demand in any meaningful proportion.

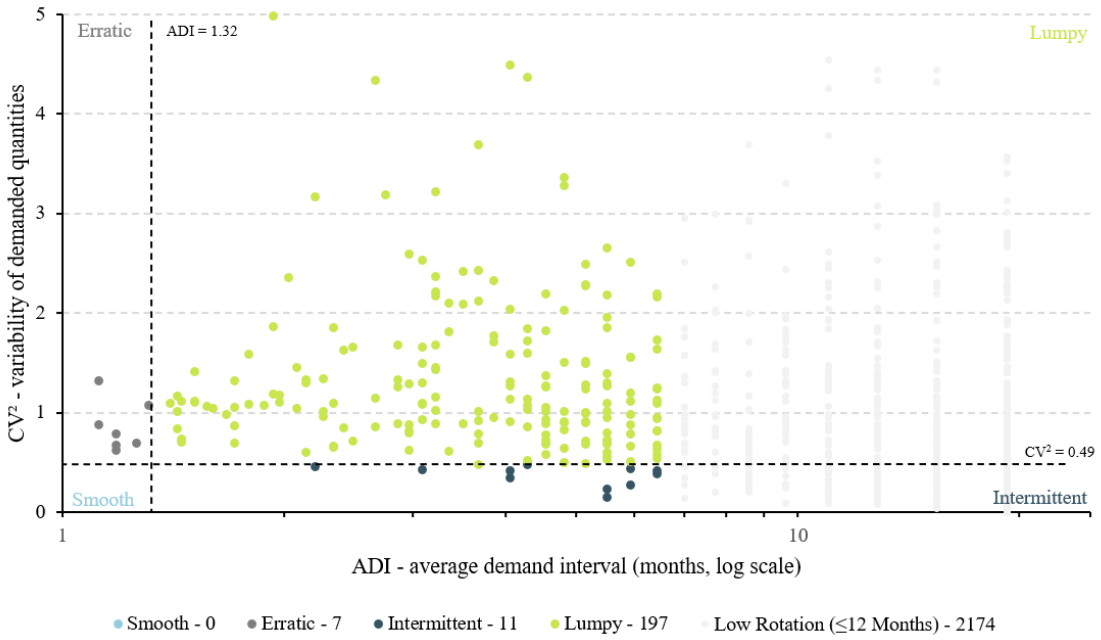


Figure 5.5: Syntetos–Boylan classification of the installation articles in the ADI-CV^2 plane (log scale on the ADI axis). Recurring articles ($\geq$ 12 months of demand) are shown in colour by category; lower-turnover articles are shown faded. The dashed lines mark the conventional cut-offs ADI $= 1.32$ and $CV^2 = 0.49$.

Table 5.1 reports the breakdown. Lumpy demand, the most demanding case in which requests are both infrequent and of unpredictable size, accounts for 54.7 % of the articles, and intermittent demand for a further 45.0 %. The median ADI is close to twenty-six months, so the typical article is consumed on average only about once every two years, against a threshold for intermittency of 1.32 months. The median CV^2 is 0.55, above the irregularity cut-off.

Table 5.1: Distribution of articles across the Syntetos-Boylan demand categories, by level of turnover. Articles are grouped according to the number of months in which demand was recorded over the observation horizon.

Article group	n	Smooth	Erratic	Intermittent	Lumpy	Median ADI
All ($\geq$ 2 mo)	2389	0.0 %	0.3 %	45.0 %	54.7 %	13.0
Recurring ($\geq$ 12 mo)	215	0.0 %	3.3 %	5.1 %	91.6 %	4.0
Highly recurring ($\geq$ 24 mo)	84	0.0 %	8.3 %	2.4 %	89.3 %	2.0

A natural objection is that this result might be an artefact of the many low turnover items. An article requisitioned in only two or three months over three years is mechanically assigned a high ADI, whether or not its demand is intrinsically lumpy. The data answer this objection directly and, in doing so, strengthen rather than weaken the premise. When the analysis is restricted to the articles that genuinely recur, namely those with demand in at least twelve of the thirty nine months,

which are the articles for which stock holding decisions are actually meaningful, the lumpy share does not fall but rises to 91.6 %. Restricting the analysis further to the highly recurring articles, defined as those with demand in twenty four or more months, corresponding to roughly monthly or more frequent demand, the lumpy share reaches 89.3 %, with a median CV^2 of 1.28.

The pattern is counter intuitive only on first inspection, as frequently consumed articles would be expected to be the most regular but, in fact, they are the most lumpy of all. The company consumes a common material such as cable or fixing hardware in almost every month, but in quantities that vary by more than an order of magnitude from one month to the next, according to which projects are active and which phase they have reached. This is precisely the demand signature that the causal, portfolio driven view of the proposed model is intended to address. The lumpiness is not noise to be smoothed, but the visible trace of an underlying driver, namely the project portfolio, that the classical methods cannot see.

The finding that 61 % of the catalogue consists of single-use items is itself evidence of the project-based nature of the operation, since the company continually handles new materials specific to each contract rather than drawing repeatedly from a stable catalogue of stock-keeping units. Also, the ADI values are computed at monthly granularity, so an ADI near unity denotes near-monthly demand rather than continuous demand, and the reported intermittency is therefore, if anything, conservative. Neither qualification alters the conclusion. Hypothesis H1 is supported without reservation, in that the demand for installation materials in the case company is almost universally intermittent or lumpy and lies outside the operating regime of the classical reorder-point models. This does not, in itself, establish that the proposed model performs well. It establishes that the problem it addresses is real, and that the methodological gap identified in Chapter 2 is a genuine one for this company.

5.4 The Stability of Phase Profiles (H2)

The second and more demanding premise is that the temporal consumption profile of a material family, its distribution across the phases of a project, is stable enough across projects to be predictive, so that the phase at which a family is consumed is a property of the family rather than an idiosyncrasy of the individual project (H2). Where H1 concerns the shape of demand in time within a single series, H2 concerns the reproducibility of a family's phase profile from one project to the next. It is this premise that would justify calibrating a profile on past projects and applying it to a new one, and it is therefore the premise on which the forecasting mechanism of the model most directly depends. This section examines it on the HVAC speciality, for which the analysis was concentrated so that the budgets could be classified by family with the accuracy that the author's domain knowledge allows.

The analysis rests on thirteen HVAC projects. For each project, the quantity of each functional family consumed in each phase of the works was obtained through the price-based reconstruction described in Section 5.1.2, and the timeline of each project was taken from the sequence of its measurement statements, for the reasons given in Section 5.1.3. The quantities measured across the sequence of statements were normalized onto the project timeline and aggregated into four equal phases from Q1 to Q4, and a family's phase profile is the share of its total consumption falling in each phase. The stability of a family was quantified as the mean, across the four phases, of the standard deviation of its phase shares between projects, an instability index expressed in percentage points, where a small value denotes a profile that reproduces closely from one project to the next. For reference, a family is termed stable below fifteen points, partially stable between fifteen and twenty-five, and unstable above twenty-five, though these thresholds are read as reference bands rather than as strict boundaries, and the instability is treated below as a continuous gradient.

Of these, six had budgets in which the classification of articles by functional family had to be reconstructed from the original quantity maps. This was necessary to obtain a broader range of projects for the analysis, since the budgets of the older projects did not classify individual articles by functional family, making a direct comparison of family-level consumption patterns impossible. Excluding them would have substantially reduced the number of observations available for testing the stability of family phase profiles and would have restricted the analysis to the more recent projects for which the company's coded classification was already available.

Because that reconstruction is a manual step, it was tested for bias before the results were interpreted. For each family, the average deviation of its per-project profiles from the family mean was computed separately for the reconstructed projects and for those classified from the company's own coded budgets, as shown in Figure 5.6.

The two groups do not diverge in a way that would inflate the apparent stability. The reconstructed projects show, if anything, a slightly smaller mean deviation than the originals, 15.2 against 17.6 percentage points, a difference of about two points in the benign direction, which is consistent with a manual classification applied by a single consistent criterion rather than with a bias toward artificial regularity, and plausibly reflecting the occasional misclassification that the company's own coded budgets contain through ordinary human error. The reconstructed projects are therefore retained in the analysis, and are marked with an asterisk wherever individual projects are shown.

Across the twenty-six families common to enough projects to be compared, five qualify as stable under the fifteen-point band, nineteen as partially stable, and only two as unstable, with a mean instability of about twenty points. The result is not that the profiles are perfectly stable, which they are not, but that the great majority of families fall in a band of low to moderate insta-

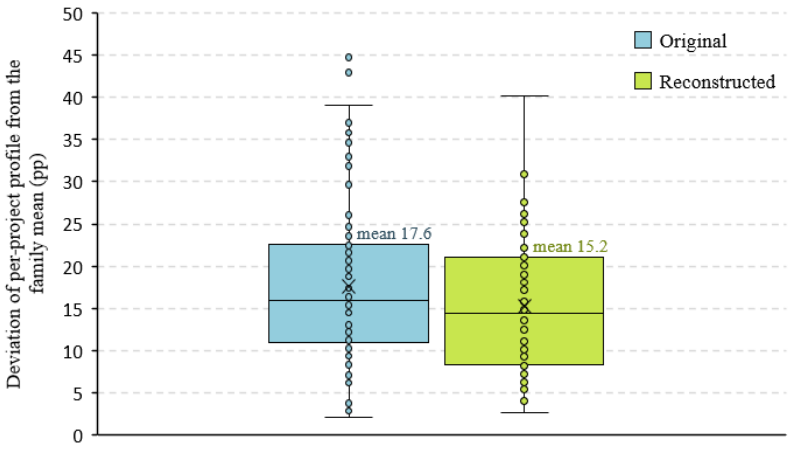


Figure 5.6: Bias test of the reconstructed budgets. The two groups overlap broadly and the reconstructed projects do not show the reduced deviation that a bias toward artificial regularity would produce.

bility rather than in the unstable extreme. The full set of families, with their mean profiles and instability indices, is reported in Figure 5.7.

Family	N° of projects	Q1	Q2	Q3	Q4	Stability [%]	Classification
Skirting trunking	3	15	8	73	4	12,3	Stable
Heating copper	6	32	32	20	17	13,7	Stable
Heat pumps	5	13	51	27	9	14,2	Stable
Intumescent grilles	4	0	11	67	22	14,8	Stable
Boilers	4	5	36	48	10	14,9	Stable
Cables	4	19	13	47	21	15,4	Partial
Field equipment	6	2	15	70	14	15,5	Partial
VRF/VRV	8	12	28	45	15	17,1	Partial
Valves	8	10	16	60	14	17,6	Partial
Spiral ducts	13	55	24	16	5	17,9	Partial
Refrigeration copper	12	53	27	14	6	18,5	Partial
Air heaters	2	36	39	20	6	19,4	Partial
Fan coils	2	25	60	11	3	20,1	Partial
Volume control dampers	9	66	20	13	1	20,7	Partial
Electrical panels	6	0	7	54	39	21	Partial
Split systems	7	18	39	37	6	21,2	Partial
Exhaust fans	6	14	29	47	10	21,3	Partial
Diffusers and Grilles	13	23	29	35	14	22,1	Partial
Fans	12	27	44	23	6	22,4	Partial
Cable trays	4	0	9	55	36	23	Partial
PVC pipes	10	41	28	24	8	23,1	Partial
Pipes / Tubes	6	22	28	34	16	24,7	Partial
AHU	4	11	58	28	3	24,7	Partial
BMS	6	3	13	46	38	24,9	Partial
Rectangular ducts	14	44	27	18	11	25,1	Unstable
Fire dampers	9	34	41	20	6	27	Unstable

Figure 5.7: Phase-profile instability of every family across the projects, with the mean profile Q1 to Q4, the instability index and the resulting class. Families are ordered from most to least stable.

Two considerations refine this into the conclusion that matters for the model. The first concern

which families are stable, and the second concern whether the stable families are the ones the model is meant to manage. The most reproducible profiles belong to families anchored to well-defined constructive stages, which is consistent with the physical reasoning behind the model, since a family whose installation is fixed within the sequence of the works will recur at the same phase, whereas a family used opportunistically across the build will not. The second consideration is the decisive one and requires crossing the stability result with the value classification established in Section 5.2. The model is directed not at every family, but at the economically significant ones, and its usefulness depends on their stability.

Figure 5.8 places every family in the plane of value against stability, with economic importance on the horizontal axis, measured as the share of projects in which the family is Class A or B, and instability on the vertical axis.

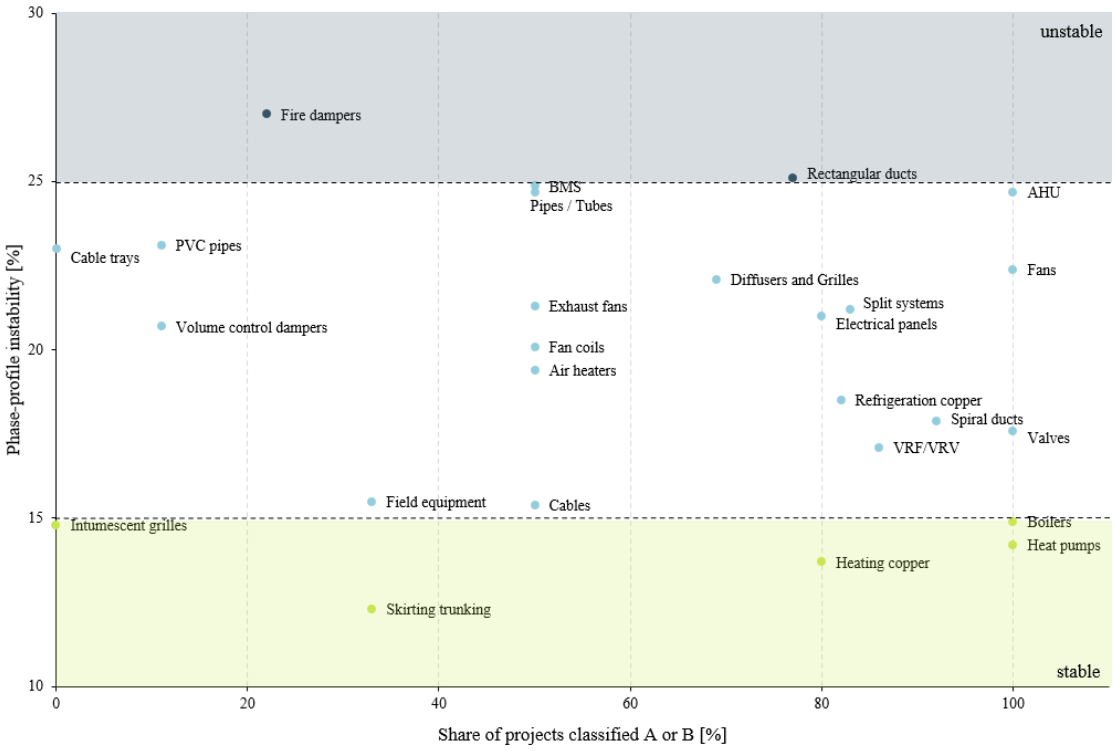


Figure 5.8: Intersection of value and phase stability across the HVAC projects. Each family is placed by its economic importance, measured as the share of projects in which it is Class A or B, and by the instability of its phase profile, with more stable families toward the bottom. The economically significant installation families cluster in the low-to-moderate instability band (green colour), which is the region in which the model applies.

The families of value concentrate in the region of low to moderate instability. Of the nineteen families that are recurrently Class A or B, three are stable under the fifteen-point band, fifteen are partially stable, and only one, the rectangular ducts, is unstable. More telling than the labels is the position on the gradient, since the core installation families of the model, among them the

refrigeration copper, the heating copper, the spiral ductwork, the valves and the VRF systems, all sit in the favourable band between roughly and eighteen points, immediately beyond the stable threshold rather than deep in the unstable region. A profile whose phase shares vary on average by some sixteen points between projects still carries substantial predictive content relative to the alternative of assuming no phase structure at all, which is the baseline against which the model must be judged.

The rectangular ductwork is the one high-value family that falls on the unstable side, at just over twenty-five points. This is coherent with the observation made in Section 5.2 that this family's classification is itself sensitive to how the budget aggregates the duct with its insulation and fire protection, a heterogeneity that plausibly carries through into a less reproducible phase profile.

Hypothesis H2 is therefore neither confirmed in general nor rejected. The evidence supports a differentiated conclusion, in that the stability of phase profiles is not a universal property of installation demand but does hold, to a degree useful for prediction, for the economically significant installation families that are the target of the model, which cluster in the low-to-moderate instability band. This delimits rather than defeats the scope of the model, and it identifies concretely the families for which the phase-based mechanism is applicable. It is on one such family that the model is put to work in the case study of Section 5.6.

5.4.1 Sensitivity to the Number of Phases

The instability index of the previous section was computed over four phases, a choice justified by their correspondence to the recognizable constructive stages of a project, namely the base infrastructure, the rough installation, the finishing and the commissioning. Because that choice is a parameter of the method rather than a property of the data, its influence on the result was examined by recomputing the instability over five and six phases as well as four. The analysis is restricted here to the Class B installation families identified in Section 5.2 as the target of the model, so that the robustness of the conclusion is tested on the families that matter for it rather than on equipment or consumables that lie outside its scope. The outcome is reported in Table 5.2.

The measured instability falls as the number of phases increases, and it does so in every one of the eight families, with the mean declining from 19.4 percentage points over four phases to 16.4 over five and 15.3 over six. Taken at face value this might seem to indicate that the profiles become more reproducible under a finer temporal resolution, and one might be tempted to adopt the larger number of phases on that basis.

In a finite-element simulation, refining the mesh improves the solution because each smaller element resolves the governing equations more accurately over its subdomain, and the computation converges toward a real physical solution that exists independently of the discretization. The

Table 5.2: Instability of the phase profile in percentage points of each Class B installation family under four, five and six phases.

Family	4 phases	5 phases	6 phases
Heating copper	13.7	10.4	10.5
Cables	15.4	13.1	14.1
Valves	17.6	14.3	14.1
Spiral ducts	17.9	15.3	13.5
Refrigeration copper	18.5	15.8	13.6
Diffusers and grilles	22.1	18.1	17.7
Pipes	24.7	21.8	18.8
Rectangular ducts	25.1	22.2	20.1
Mean	19.4	16.4	15.3

refinement of the temporal partition here has no such property. There is no underlying continuous profile toward which a finer partition converges. Increasing the number of phases merely distributes the same fixed consumption over a larger number of intervals, and because much of a family's consumption is concentrated in a few months, the additional intervals are largely empty. Two projects that both record little or nothing in the same added intervals are then scored as agreeing in those intervals, and the mean per-phase standard deviation falls mechanically. The reduction in instability is thus an artefact of dilution into empty compartments rather than a genuine gain in reproducibility. That the effect applies uniformly, reducing the measured instability even of the rectangular ductwork whose profile genuinely diverges from project to project, is itself the signature of an artefact, since a real improvement in reproducibility would discriminate between families rather than lowering all of them together.

The four-phase partition is therefore retained, not because it yields the most favourable numbers, which it does not, but because it is the conservative choice and the one with a substantive justification in the physical stages of construction. The finer partitions are reported here only to document the sensitivity of the measure and to make explicit that the stability results of the previous section are not inflated by the choice of resolution.

5.5 Supplier Reliability and Lead-Time Variability

The premises examined so far concern the demand side, the shape of consumption in time and the reproducibility of the phase profiles. The model, however, must also protect against the supply side, the time it takes for a replenishment order to arrive, and it is the variability of that time rather than its average that governs the safety component of any minimum stock. This section characterizes the supplier lead time empirically, at the level of the whole material base, both to quantify that variability for the sizing model and to test whether the supplier delays reported qualitatively in the

diagnosis are borne out by the transactional record. The lead time of the specific family used later in the case study is treated in its place, in Section 5.7, and the analysis here establishes the general picture into which that specific case fits.

The analysis draws on the complete record of purchase orders to suppliers over the period from 2020 to 2026, comprising 65,761 order lines with a valid order-to-receipt time, placed with 535 distinct suppliers. For each line the real lead time follows directly as the difference between the order date and the goods-receipt date, a quantity that is reliable because it rests on two observed events.

A measure of punctuality in the conventional sense, an on-time-in-full indicator, would compare the receipt date against a promised delivery date. That comparison was attempted and then abandoned, because the promised-delivery field in the ERP proved not to carry a genuine commitment. In 59 % of the order lines the recorded delivery date equals the order date itself, and in 87 % it falls on the order date or the following day, so the field is a system default rather than a negotiated date, and any tardiness computed against it would be an artefact of that default. The punctuality metric is therefore not reported, and the lead time from order to receipt, which does not depend on the unreliable field, is used throughout. The same limitation is revisited in the construction of the monitoring dashboard in Section 4.6.2.

Taken across all orders, the lead time is strongly right-skewed, with a median well below the mean and a long tail, so that most orders arrive quickly but a substantial minority take far longer. The full statistics are reported in Table 5.3, where orders are divided into current and made-to-order material families. This distinction is essential for interpreting the lead-time distribution, as the two populations exhibit fundamentally different replenishment characteristics.

Table 5.3: Supplier lead-time statistics, in days, for all orders and for the current-material population that the model governs.

Order population	Median	Mean	Std. dev.
All orders	9	24	42
Current material (74%)	6	12	24
Made-to-order equipment (26%)	*months, by nature of the product*		

About a quarter are placed with suppliers of made-to-order equipment, switchboards built to specification, prefabricated modules, imported units, whose lead times run to months by the nature of the product rather than through any failure of service. The remaining 74% concern current stock material, the piping, cable, ductwork and fixing hardware that the inventory model is intended to govern. For this population the lead time has a median of seven days but a mean of twelve, with a

standard deviation of 24 days, and its strongly right-skewed shape is shown in Figure 5.9.

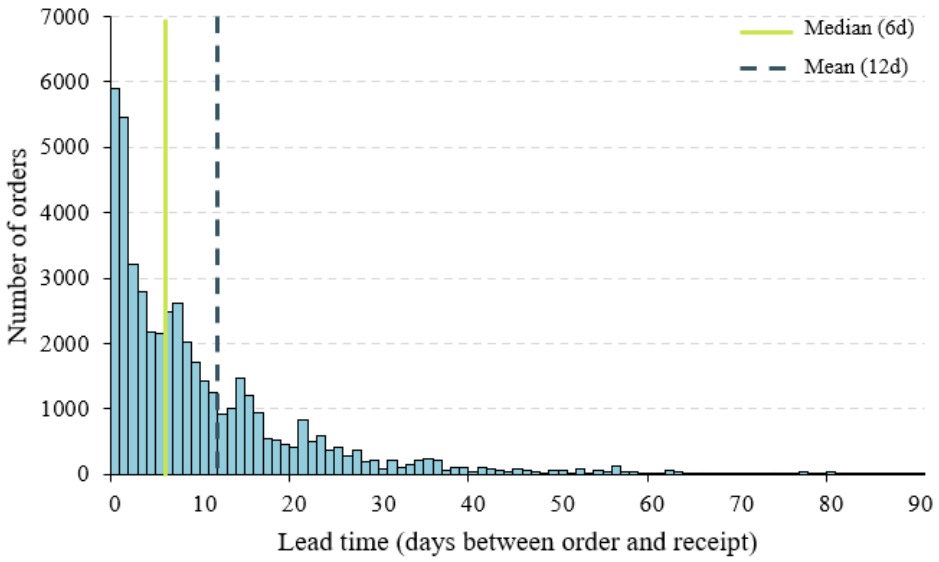


Figure 5.9: Distribution of the supply lead time for current stock material.

It is again the dispersion, not the average, that matters, and dimensioned to cover it at a 95% service level the safety component would need to protect against roughly forty days of lead-time variability beyond the mean. This provides the empirical basis, at the general level, for the lead-time term of the minimum stock developed in Section 4.5.3, grounded in the company's own supply history rather than in an assumed constant. The distinction between the two supplier populations is itself a result that reinforces the scope of the model, since current material, fast-moving and repeatedly ordered, is a sensible candidate for stock holding, whereas made-to-order equipment is better handled by project-driven procurement, in keeping with the boundary drawn around the Class B installation families.

The aggregate figure also conceals large and systematic differences between individual suppliers, which the sizing can exploit. Each supplier can be placed by its mean lead time against its lead-time variability, so that those whose deliveries are both slow and erratic are distinguished from those that are fast and consistent. This per-supplier view is realized operationally as the risk quadrant on the supplier page of the monitoring dashboard, presented in Section 4.6.2 and Appendix C, where each supplier is positioned in that plane with the size of each point reflecting the value ordered. A material sourced from a fast and consistent supplier requires far less buffer than the same material from a slow or erratic one, so the lead-time statistics computed per supplier feed a differentiated safety stock rather than a single blanket margin. This per-supplier characterization is the general counterpart of the family-specific lead time that the case study of Section 5.7 uses to size the stock of one material.

The lead time of current material also varies through the year in an operationally relevant way, as Figure 5.10 shows. Computed on the median, so that the long-lead equipment tail does not distort the reading, most of the year sits close to six or seven days, but June stands clearly apart at ten days, with August somewhat elevated. The pattern is consistent with the approach of the summer holiday period, when suppliers and manufacturers slow down or accumulate a backlog ahead of the shutdown. The practical implication is direct, since orders of current material placed in late spring, ahead of June, should allow for a longer delivery time, and anticipating replenishment before the summer would reduce the stockout risk that the extended lead time would otherwise create.

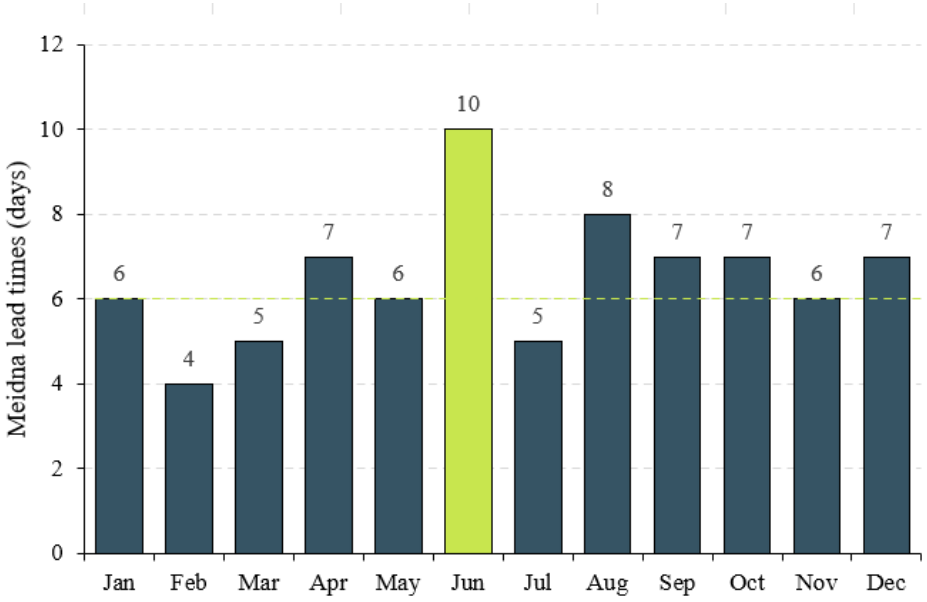


Figure 5.10: Median lead time of current material by month of order.

The purchasing record confirms, on a reliable metric, that the supply of installation material is subject to a lead time that is short on average but highly variable, that this variability differs markedly across suppliers and across the months of the year, and that the two supplier populations, current material and made-to-order equipment, must be treated separately. The supplier delays reported qualitatively in the diagnosis are therefore substantiated quantitatively, not as a failure of punctuality against a promised date, which the data do not permit assessing, but as a genuine dispersion of the realised delivery time. This dispersion is precisely what the safety component of the proposed minimum stock is designed to absorb, and the statistics established here, at the general level and per supplier, provide its empirical foundation, of which the family-specific case study that follows is a concrete instance.

5.6 Case Study: Forecasting Performance (H3)

The premises examined so far establish that the demand is lumpy (H1) and that the phase profiles of the target families are reproducible to a useful degree (H2). What remains is to test whether the model, built on those premises, actually forecasts the phased consumption of a family better than the classical methods for intermittent demand (H3). Because the number of projects is limited, this is done as a case study on individual families rather than as a fleet-wide validation, in keeping with the demonstration-of-feasibility standard adopted throughout. Two families are used, and the sequence in which they were analysed is itself part of the result.

Given the limited number of projects, a leave-one-out cross-validation procedure was adopted for this analysis. For a given family and a given target project, the phase profile is calibrated as the average of the profiles of all the other projects in which the family appears, never using the target project to predict itself. That calibrated profile is then applied to the total quantity of the family budgeted for the target project, yielding a forecast of how much is consumed in each phase, which is compared against the real phased consumption. The total is taken from the budget, under the assumption, examined below, that a project consumes what it was budgeted, so that the test isolates the quality of the phasing rather than the accuracy of the total.

The first two benchmarks are the classical methods for intermittent demand identified in Section 2.4, namely Croston's method and its Syntetos-Boylan approximation (SBA), each applied to the phased series of the target project so that it forecasts each phase from the phases already observed within that project. The third is a uniform benchmark that spreads the total equally across the four phases, representing the absence of any phase information. The comparison against the classical methods answers H3 as the literature frames it, while the comparison against the uniform benchmark isolates how much of the model's advantage comes from the phase profile itself. Forecast error is measured as the mean absolute error across the four phases.

Heating copper was chosen first because it is the most stable family by the instability index of Section 5.4, at 13.7 percentage points, and therefore the natural candidate. The model reduces the forecast error relative to Croston by 52 % and relative to SBA by 41 %, a clear advantage over the classical methods. Against the uniform benchmark, however, the reduction is a more modest 20 %, and in two of the six projects the model does not beat the uniform spread. Figure 5.11 shows the leave-one-out result project by project, with the profile calibrated on the other five projects set against the actual profile of each.

Taken alone, this is an honest but unspectacular result, and it prompted the question of whether a family analysed over a larger number of projects, even if intrinsically less stable, might yield a more convincing demonstration. This question motivated the application of the method to the refrigeration copper family, which appears in twelve projects, twice as many as heating copper,

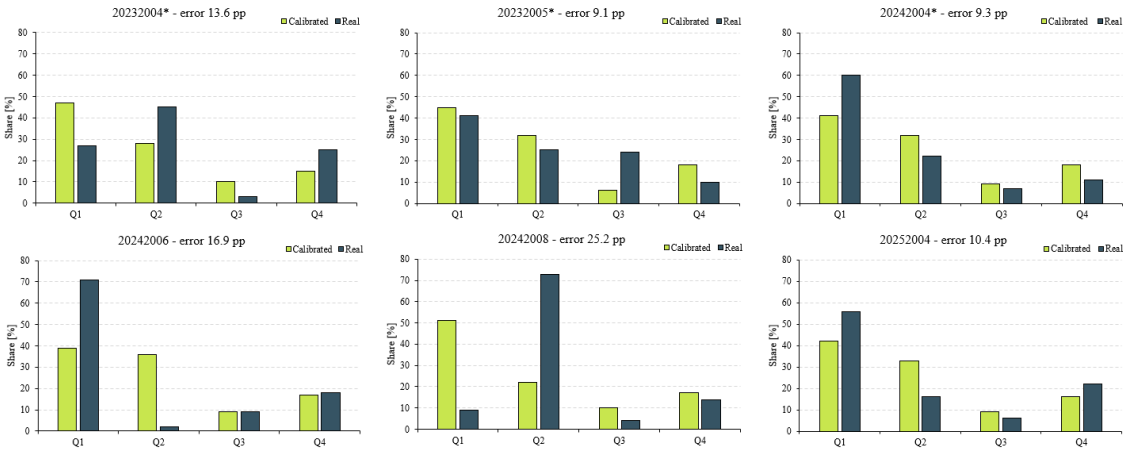


Figure 5.11: Leave-one-out validation for heating copper. For each project, the phase profile calibrated on the other five projects (green) is compared against the actual profile (blue), with the mean absolute error in each project. Reconstructed-budget projects are marked with an asterisk.

and is somewhat less stable by the index, at 18.5 percentage points. The per-project profiles of both families are shown in Figure 5.12, where the sharper, earlier concentration of refrigeration copper is apparent.

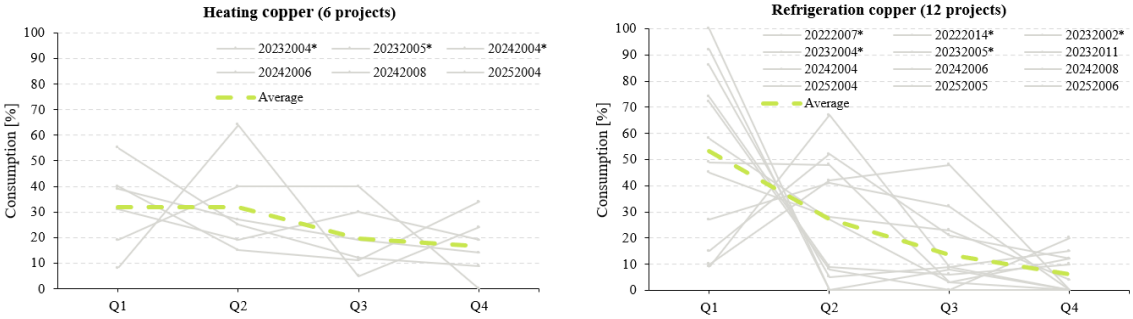


Figure 5.12: Per-project phase profiles of the two case-study families, with the average highlighted.

The comparison is reported across four forecast-error metrics in Table 5.4, the mean absolute error, the root mean squared error, the mean absolute percentage error, and the mean absolute scaled error, against the two classical methods and the uniform benchmark. The model is superior to both classical methods on every metric and in both families. The root mean squared error, which penalizes large errors more heavily than the mean absolute error, confirms that the model's advantage is not obtained at the cost of occasional large misses, since it too is lowest for the model. The mean absolute percentage error is reported for completeness but is of limited value here, because for intermittent demand with near-zero observations it is unstable and inflates, which is why all methods show high values on it. The mean absolute scaled error is the appropriate measure for this regime, as it scales the error against a naive historical mean, and it carries the most telling result. For the model it is at or just below one, whereas for Croston and SBA it exceeds one in both fami-

lies, meaning that these classical methods, lacking the causal information of the project portfolio, do not even improve on a trivial historical average. The uniform benchmark, which like the naive mean carries no phase information, sits at one by construction, and the model's advantage over it, more modest than over the classical methods, isolates the part of the model's value that comes specifically from the shape of the phase profile as distinct from the mere knowledge that a project is under way.

Table 5.4: Forecast-error metrics for the model and the benchmarks, under leave-one-out, for the two Class B families: Heaintg copper and Refrigeration copper.

Family	Method	MAE	RMSE	MAPE	MASE
Heating copper (5 proj.)	Croston	394	432	216%	1.61
	SBA	320	363	168%	1.32
	Uniform	237	290	134%	1.00
	Model	**216**	**261**	**85%**	**1.00**
Refrigeration copper (12 proj.)	Croston	1786	1866	345%	1.88
	SBA	1455	1534	271%	1.58
	Uniform	994	1158	151%	1.00
	Model	**629**	**718**	**84%**	**0.97**

The comparison of the two families carries the main message of the case study, and it is more informative than either family alone. The family with more projects gives the better result even though its instability index is higher, which shows that the number of projects available for calibration matters as much as the intrinsic variability of the family. With twelve projects, the leave-one-out profile is a far more reliable estimate of the family's characteristic phasing than it is with six, and this outweighs the modestly greater dispersion. A second factor works in the same direction. The refrigeration copper has a strongly marked profile, concentrated early in the works, with most projects placing the bulk of consumption in the first phase, which is coherent with its role as base infrastructure installed at the start. A profile with a pronounced shape departs further from the uniform spread than a flatter one, which is why the model's advantage over the uniform benchmark is larger here, 37 % against 20 %, than for the more evenly spread heating copper. The instability index alone is therefore not the sole predictor of how well the model performs, and the number of projects and the sharpness of the profile matter alongside it.

The case study supports H3 as the model, for both families, forecasted phased consumption substantially better than Croston and SBA, the reference methods for intermittent demand, with reductions in error between 41 % and 65 %. As previously stated, the classical methods forecast only from the past of the series, and in a project-based setting they are always reacting to a demand whose driver they cannot see, so that they lag the consumption of each new project. The proposed

model, by anchoring the forecast in the budget of the active project and the phase profile learned from earlier projects, incorporates the causal information that the classical methods lack. This is visible in the projects where consumption concentrates early, where Croston, having no early estimate, errs by a wide margin while the model does not.

The more modest advantage over the uniform benchmark, especially for heating copper, tempers the claim and is reported deliberately. It indicates that much of the model's value comes from knowing that a project is under way and how large it is, which already defeats the classical methods, and a smaller part from the specific shape of the phase profile. This is consistent with the differentiated conclusion of H2, in which the profiles were found reproducible only to a moderate degree, and it locates the model's contribution more in the causal anchoring of demand than in a precise deterministic profile.

It is, however, important to acknowledge that three limitations constrain the interpretation of these results. First, the test assumes that a project consumes what it was budgeted, so that the total is known without error. It presupposes that projects do not have major errors or omissions, and a project that deviated substantially from its budget would carry an error in the total that this test does not capture. Second, the demonstration rests on two families in a single speciality, and while the agreement between them is encouraging, it is not a fleet-wide validation. Third, the leave-one-out profile is calibrated on a small number of projects, and the heating-copper result in particular shows how the reliability of the calibration depends on that number. These bounds place the result where it belongs, as a demonstration that the model is feasible and outperforms the classical alternatives on the families for which it is intended, rather than as a general proof of forecasting accuracy across all materials.

5.7 Case Study (Continued): The Inventory Trade-off (H4)

Having shown that the phased consumption of refrigeration copper can be forecast (H3), the case study now turns to the policy that the forecast enables, and to whether the dynamic sizing it supports improves on the company's current practice (H4). The forecast tells the firm when the material will be needed, and the sizing question is when to order it so that it arrives in time without holding it longer than necessary. The answer depends on the lead time and, above all, on its variability, which is why the family-specific lead time established below is central to the analysis.

The lead time of refrigeration copper was isolated from the purchasing record by selecting the purchase orders whose article description identifies the material as refrigeration copper, which excludes the heating and sanitary copper. This yields 405 orders with a valid order-to-receipt time. The lead time has a mean of 7.5 days and a median of three, with a standard deviation of 13.8 days and a ninetieth percentile of twenty-one, so that most orders arrive quickly but a substantial tail

extends much further. Isolating the family was worthwhile, since this copper is both faster and less variable than the current material as a whole, whose lead time had a standard deviation of 21.7 days, and sizing against the family-specific variability avoids the over-provisioning that the aggregate figure would imply.

A Monte Carlo simulation was performed in order to compare the two policies. The current practice is a stock-zero policy in which the firm holds no copper and orders it about seven days ahead of the anticipated need, five working days. A stockout occurs whenever the realised lead time exceeds that anticipation, so that the material arrives after it was needed. The dynamic policy orders further ahead, by an amount sized to cover the lead time at the 95 % service level, which for this family is the mean plus 1.65 standard deviations, about thirty days. In each of ten thousand simulation runs, the lead time of each copper requirement is drawn by bootstrap from the 405 observed lead times, without assuming any distributional form, and the stockout rate is recorded as the share of requirements that arrive late. The number of requirements per run reflects the roughly five projects active at any time, each ordering copper at several points in its course.

Figure 5.13 shows the full Monte Carlo distributions of the stockout rate under the two policies, which are well separated, and Figure 5.14 shows the copper lead-time distribution against which the two anticipation windows are set.

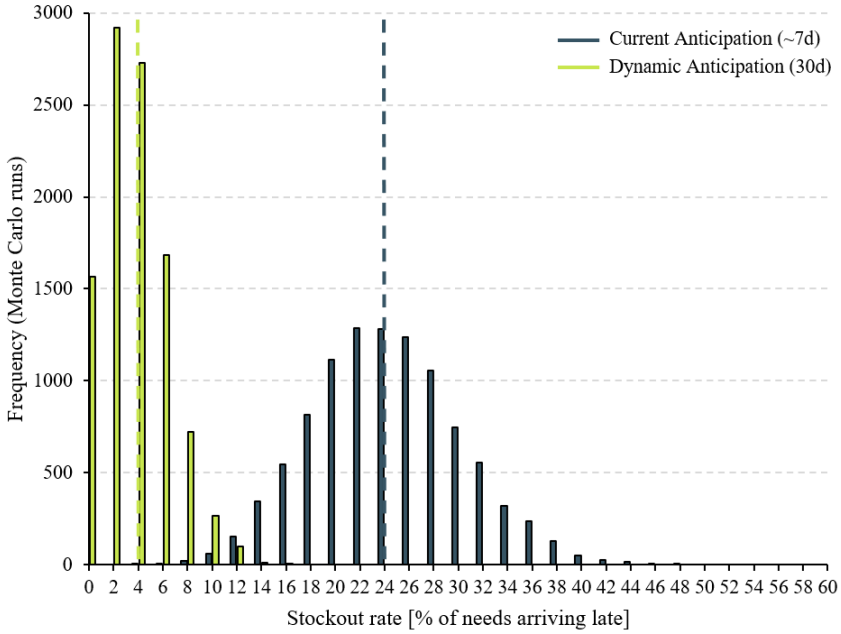


Figure 5.13: Monte Carlo distribution of the stockout rate for refrigeration copper over ten thousand runs, under the current stock-zero policy and under the dynamic policy, with means of 24% and 4 %. The two distributions are well separated, and the dashed lines mark the means.

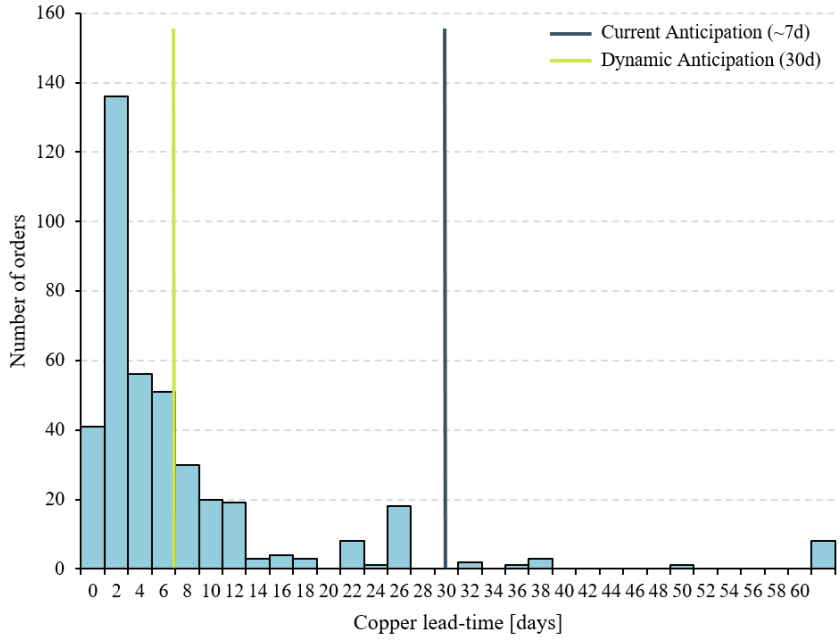


Figure 5.14: Copper lead-time distribution with the two anticipation windows, seven days for current practice and thirty for the dynamic policy.

The dynamic policy reduces the stockout rate from 24 % under current practice to 4 %, a reduction of twenty-one percentage points, or about four fifths of the stockouts eliminated. The reason is visible in in Figure 5.14. With an anticipation of seven days the firm already covers the bulk of the distribution, since the median lead time is only three days, and it is right in three quarters of cases. What it does not cover is the tail, the orders that take fifteen, twenty or more days, and it is that tail, one requirement in four, that causes the stockouts. The dynamic anticipation of thirty days reaches into the tail and covers all but the most extreme lead times, leaving a residual stockout rate of about one in twenty-five.

To cover the tail, the dynamic policy orders, on average, about twenty-three days earlier than current practice, meaning that the material is held, or is in transit and committed to the project, for that much longer before it is used. For a low-value bulk material such as copper, whose holding cost is small relative to the cost of a project stalled for want of it, this is a favourable exchange, but it is an exchange rather than a pure gain, and a firm that today operates at stock zero would be accepting a modest holding cost in return for the near-elimination of stockouts.

Two residual points qualify the result and are reported for completeness. First, the few stockouts that remain under the dynamic policy carry a larger average delay than those under current practice, because they are precisely the extreme lead times beyond the ninety-fifth percentile, so that when the dynamic policy fails it fails on the rare long tail rather than on the common case. The stockout rate, not the conditional delay, is therefore the meaningful comparison. Second, the simulation rests on an assumed current anticipation of seven days and on the number of copper

requirements implied by the typical concurrency of projects, and while neither assumption affects the mechanism, by which the dynamic policy covers the lead-time tail that the current one does not, they do condition the precise figures.

Hypothesis H4 is supported for the case-study family as dynamic sizing, fed by the H3 phase forecast and by the family-specific lead-time variability, delivers a markedly better service level than the current stock-zero practice, reducing stockouts from roughly one requirement in four to one in twenty-five, at the cost of a modest and, for this material, economically reasonable amount of anticipated stock. As with H3, this is a demonstration of feasibility on one family rather than a fleet-wide result, but it shows concretely that the model's output translates into an inventory policy with a real and quantified operational benefit.

5.7.1 Economic Estimate of the Benefit

The reduction of stockouts demonstrated in the case study carries an economic value that management is entitled to see quantified, even where the company does not yet record the data that would make an exact figure possible. What can be built from the questionnaire responses of Chapter 3, together with a small set of declared assumptions, is an order-of-magnitude estimate, and it is presented as a scenario analysis so that its sensitivity to those assumptions remains visible rather than hidden inside a single number. All the figures that follow are organization-wide, covering the Class B installation materials that the model would govern across the active portfolio, and are not per family.

The estimate rests on a small number of assumptions, collected in Table 5.5 so that they remain visible and open to challenge. The labour figure is deliberately conservative, since the interruption is charged at a single hour, the time a pair of workers takes to leave the current task and organize another, rather than as a prolonged stoppage of the work front. The urgent-order premium reflects the cost of securing immediate availability rather than the lowest price, and the carrying-cost rate is the standard reference for industrial warehousing. The proportion of stockouts the model would remove is carried over from the case study of Section 5.7.

The frequency of stockouts and urgent orders at the level of the organization is the least certain input, since the questionnaire records it per informant over a four-week window rather than for the firm as a whole, and the company runs more than ten projects at once. It is for this reason that the estimate is cast in three scenarios, reported in Table 5.6, with the conservative scenario taken as the reference so that the benefit is not overstated. Under that conservative scenario, the current cost of the problem, in idle labour and urgent-order premiums combined, is on the order of 4,300 € per year, of which the model would avoid some 3,500 by removing the greater part of the stockouts. Against this stands the carrying cost of the stock the model would introduce, of the order of 1,000 to 4,000 € per year for a Class B holding valued in the low tens of thousands, so

Table 5.5: Assumptions underlying the economic estimate.

Assumption	Value
Loaded labour cost, oficial	15 €/h
Loaded labour cost, servente	12 €/h
Workers affected per stockout	2 (1 oficial + 1 servente)
Combined stoppage cost	27 €/h
Idle time charged per stockout	1 h
Cost per stockout (idle labour)	27 €
Urgent-order premium	5%
Annual inventory carrying-cost rate	20%
Stockout reduction from the model (from H4)	∼80%

that the net position is close to break-even in the most cautious reading and clearly positive in the central and pessimistic ones, where the annual saving reaches roughly 11,000 and 22,000 €.

Table 5.6: Scenario estimate of the annual, organization-wide cost of the current situation and the saving the model would achieve. Frequencies are varied across scenarios, while the labour and premium rates are held at the values stated above.

Item (per year)	Conservative	Central	Pessimistic
Stockouts per month	6	15	25
Urgent orders per month	8	20	35
Idle-labour cost	1,944 €	4,860 €	8,100 €
Urgent-order premium	2,400	8,400	18,900
Total cost of the problem	4,344 €	13,260 €	27,000 €
Saving from the model (∼80%)	**3,475 €**	**10,608 €**	**21,600 €**
Carrying cost of added stock	1,000-4,000 €	1,000-4,000 €	1,000-4,000 €

Two things must be said about these figures, and the first is that they are modest. Even in the central scenario the saving is of the order of ten thousand euros per year, a figure that would not by itself transform the finances of a company with a turnover above ten million. This modesty is a consequence of the conservative choices made, in particular the charging of only one hour of idle labour per stockout, and it is preferable to the alternative of an inflated figure that would not survive scrutiny. The value of the result lies less in the absolute amount than in the demonstration that the model pays for itself, since the saving exceeds the carrying cost of the stock it introduces in every scenario but the most cautious, where the two are comparable.

The second observation concerns the composition of the cost, which shifts once the stoppage is charged conservatively at one hour. The dominant component is then no longer the idle labour but the premium paid on urgent orders, which in the central scenario accounts for close to two

thirds of the total. This is an informative result in its own right, since it locates the greater part of the economic damage not in the visible disruption of a halted work front but in the quieter, recurring cost of buying reactively at a premium, a cost that the anticipation the model provides is precisely suited to remove. It also indicates that the benefit would scale more than proportionally in a firm that grows or that extends the approach to its other engineering specialities, since both the frequency of urgent orders and their value would rise with the volume of work.

The estimate is deliberately an order of magnitude, because the company does not record the data that would make it exact, and the requirements to measure the benefit rigorously rather than estimate it are set out in Table 5.7. These are the same data that the enabling layer proposed in this work is designed to capture, so that instituting the measurement and deploying the enabling layer are one and the same task, and the scenario estimate can be replaced by a measured return once the layer has operated for a period.

Table 5.7: Proforma of the data required to measure the economic benefit rather than estimate it, with the source through which each would be captured.

Data to record	Source / capture mechanism
Each stockout event, with date and project	Dashboard flag on an unfulfilled requirement
Workers affected and hours lost per event	Site report, logged against the event
Loaded labour cost per worker category	Payroll and accounting
Urgent orders and the premium paid	ECF records, flagged as urgent
Value of stock held, by family and class	ERP stock records, once reliably captured
Real supplier lead time and its variability	ECF-to-reception dates in the ERP

Framed in this way the economic dimension is not left unaddressed but placed on a firm footing. The present work establishes the physical benefit and its order of magnitude under conservative assumptions, and specifies exactly what must be recorded to convert that estimate into a measured return, which is itself one of the purposes the enabling layer is meant to serve.

Chapter 6

Conclusions and Future Work

This chapter draws together the findings and situates the contribution. It states the conclusions, on the nature of the demand, the conditions under which the model applies, and its demonstrated performance, and identifies directions to extend the work, from the economic quantification and prospective validation left open to the broadening of the empirical base and the refinement of the model.

6.1 Conclusions

This dissertation set out to address a concrete and largely unstudied problem, the management of material inventory in a project-based installation company, where demand is not the smooth, stationary process that the classical inventory models assume but is instead driven by the discontinuous rhythm of a portfolio of construction projects. The work proposed a model that treats the project portfolio itself as the causal driver of demand, forecasting the phased consumption of each material family from the budgets and execution profiles of the active projects, and sizing a dynamic minimum stock that combines this forecast with the variability of the supplier lead time. Around this model it defined an enabling layer, a QR-based data-capture process and a monitoring dashboard, intended to supply and to surface the data on which the model depends.

The premises of the model were examined empirically against the company's own records. The demand for installation materials was confirmed to be, almost universally, intermittent or lumpy, with the great majority of articles falling outside the regime for which the classical reorder-point methods were devised, a result that holds across every material category and grows stronger, rather than weaker, when the analysis is restricted to the recurrently ordered articles for which stock decisions are actually meaningful. The stability of the phase profiles, on which the forecasting mechanism rests, was found to be real but partial, holding to a useful degree for a subset of families, those tied to well-defined constructive stages, which include a substantial share of the economically significant Class B installation material, while not holding for the many families whose consumption is spread opportunistically across the works. This is a differentiated rather

than a binary conclusion, and it delimits the scope of the model rather than defeating it.

On the families for which the premises hold, the model was shown to deliver. In a leave-one-out case study on two copper families, it forecast the phased consumption substantially better than the Croston and Syntetos-Boylan methods, the standard references for intermittent demand, on every error metric, with the scaled error revealing that those classical methods do not even improve on a naive historical mean because they lack the causal information of the portfolio that the model incorporates. The dynamic sizing built on this forecast, together with the family-specific lead-time variability, was shown by Monte Carlo simulation to reduce the incidence of stockouts on the case-study family from roughly one requirement in four to one in twenty-five, at the cost of a modest and economically reasonable amount of anticipated stock. The central contribution of the work, the use of the project portfolio as a causal variable in place of the extrapolation of a demand history, is therefore not only proposed but demonstrated to be effective on the subset of materials it is designed for.

The work also produced, as a by-product of its empirical grounding, a clear picture of the data limitations of the organization. The reconstruction of family-level consumption required an indirect linkage through unit price because the ERP does not expose the relation directly, the timeline of each project had to be taken from its measurement statements because the nominal project dates are placeholders, a conventional punctuality metric could not be computed because the promised-delivery date is a default, and a procurement-cycle indicator could not be built because the link between requisition and order is not registered. These are not incidental difficulties but a structural condition, and they are themselves the strongest argument for the data-capture layer that the work proposes.

6.2 Future Work

Several directions would extend and strengthen the contribution, and their place in the overall path to full operation is summarized in the implementation roadmap of Figure 6.1, whose phases are ordered by dependency, since reliable data must precede visibility, and visibility the automatic operation of the model. The first concerns the economic benefit. The present work estimated it, under conservative assumptions and in scenario form, in Section 5.7.1, but that estimate rests on assumed frequencies and costs rather than on recorded ones. Its natural extension is the measurement of the benefit in operation, drawing on the cost and event data that the enabling layer is designed to capture, which would replace the scenario estimate with a measured return and a return on the implementation, the form in which the benefit is most legible to management.

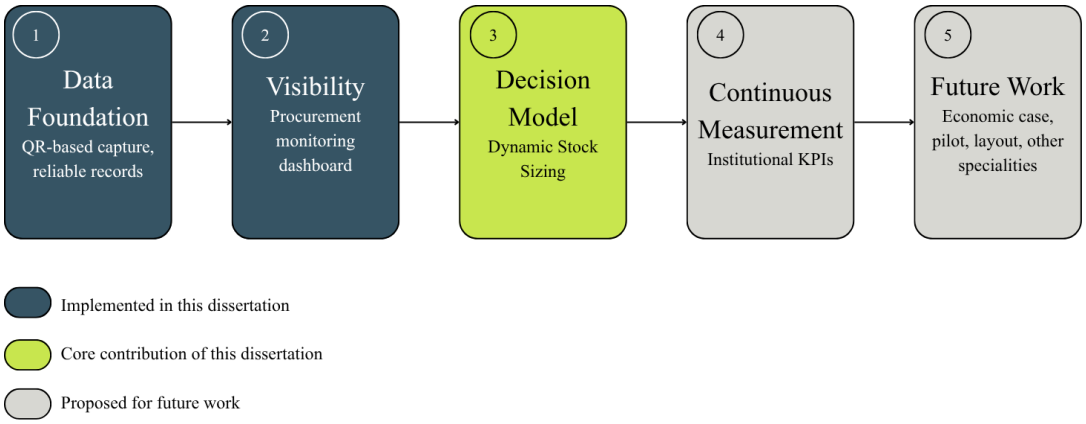


Figure 6.1: Implementation roadmap for the adoption of the approach, with the phases ordered by dependency. The colour indicates the status of each, the data-capture and dashboard prototypes and the retrospectively validated model developed in this work, and the phases proposed for future work.

A second direction is the prospective validation of the model in operation. All the validation reported here is retrospective, conducted on completed projects, and a pilot deployment over a period of several months, on one or two Class B families, measuring the real inventory and service outcomes against the current practice, would provide the operational confirmation that a retrospective study cannot. Such a pilot depends on the data-capture layer being in place, which links this direction to the deployment of the QR process and the connection of the dashboard to a live data feed.

A third direction is the broadening of the empirical base. The stability of the phase profiles was assessed over thirteen projects in a single speciality, and both a larger number of projects and the inclusion of the other engineering specialities would allow the stability, and the boundary of the model's applicability, to be established with greater statistical confidence, through confidence intervals and resampling of the phase weights rather than through the point estimates used here. The same broadening would allow the decision rule that governs whether a family is managed by the model to be calibrated on evidence rather than on reasoned thresholds.

Finally, the model itself admits refinement. The phase profile is here a simple average of past projects, and a weighting that accounts for the similarity between projects, or a treatment of the profile as a distribution rather than a point estimate, would allow the forecast to carry its own uncertainty into the stock sizing. The integration of the machine-learning methods surveyed in the literature, trained on a broader data history once the capture layer has accumulated it, is a further and longer-term possibility.

A further direction, complementary to the demand-side focus of this work, concerns the physical organization of the warehouse itself. The present dissertation addresses what to hold and when

to order it, but not where to place it. The same transactional record that drives the sizing model, the history of material movements and consumption held in the ERP, could equally inform the physical layout of the warehouse, positioning the most frequently moved materials in the most accessible locations and grouping items that are consumed together. A data-driven slotting of this kind, grounded in the movement frequencies that the ERP already records, would reduce handling time and complement the inventory policy developed here, and is a natural extension once the data-capture layer has improved the reliability of those movement records.

References

[1] Matthew Oladeji Ajimati, Noel Carroll, and Mary Maher. “Adoption of low-code and no-code development: A systematic literature review and future research agenda”. In: *Journal of Systems and Software* 222 (2025), p. 112300. ISSN: 0164-1212. DOI: https://doi.org/10.1016/j.jss.2024.112300. URL: https://www.sciencedirect.com/science/article/pii/S0164121224003443.

[2] Adnan Aktepe, Erdal Yanık, and Süleyman Ersöz. “Demand Forecasting Application with Regression and Artificial Intelligence Methods in a Construction Machinery Company”. In: *Journal of Intelligent Manufacturing* 32 (2021), pp. 1587–1604. DOI: 10.1007/s10845-020-01647-1.

[3] Nadya Amanda Istiqomah et al. “The Implementation of Barcode on Warehouse Management System for Warehouse Efficiency”. In: *Journal of Physics: Conference Series* 1573 (2020). DOI: 10.1088/1742-6596/1573/1/012038.

[4] John Bartholdi and Steven Hackman. “WAREHOUSE & DISTRIBUTION SCIENCE Release 0.94”. In: (Feb. 2011).

[5] Esther Calderon-Monge and Domingo Ribeiro-Soriano. “The role of digitalization in business and management: a systematic literature review”. In: *Review of Managerial Science* 18 (Mar. 2023), pp. 1–43. DOI: 10.1007/s11846-023-00647-8.

[6] Shalina Sultana Champa and Richard S. Segall. “Integrated Data Analytics, Business Intelligence, and Machine Learning Architecture for SMEs:” in: *International Journal of Business Analytics* 13.1 (2026). ISSN: 2334-4547. DOI: https://doi.org/10.4018/IJBAN.409371. URL: https://www.sciencedirect.com/science/article/pii/S2334454726000043.

[7] Junru Chen and Lei Shen. “A Synthetic Review on Enterprise Digital Transformation: A Bibliometric Analysis”. In: *Sustainability* 16.5 (2024). ISSN: 2071-1050. URL: https://www.mdpi.com/2071-1050/16/5/1836.

[8] J. D. Croston. “Forecasting and Stock Control for Intermittent Demands”. In: *Operational Research Quarterly* 23.3 (1972), pp. 289–303. DOI: 10.2307/3007885.

[9] Haytham H. Elmousalami. “Artificial Intelligence and Parametric Construction Cost Estimate Modeling: State-of-the-Art Review”. In: *Journal of Construction Engineering and Management* 146.1 (2020), p. 03119008. DOI: 10.1061/(ASCE)CO.1943-7862.0001678.

[10] N. Faber, René De Koster, and Ale Smidts. “Organizing warehouse management”. In: *International Journal of Operations & Production Management* 33 (Aug. 2013). DOI: 10.1108/IJOPM-12-2011-0471.

[11] Alejandro Germán Frank, Lucas Santos Dalenogare, and Néstor Fabián Ayala. "Industry 4.0 technologies: Implementation patterns in manufacturing companies". In: *International Journal of Production Economics* 210 (2019), pp. 15–26. ISSN: 0925-5273. DOI: https://doi.org/10.1016/j.ijpe.2019.01.004. URL: https://www.sciencedirect.com/science/article/pii/S0925527319300040.

[12] Dongmei Gao, Fabian Fagerholm, and Vilma Toivanen. "What does current research say about the viability of low-code development? A systematic literature review". In: *Journal of Systems and Software* 239 (2026), p. 112893. ISSN: 0164-1212. DOI: https://doi.org/10.1016/j.jss.2026.112893. URL: https://www.sciencedirect.com/science/article/pii/S0164121226001263.

[13] Ford W. Harris. "How Many Parts to Make at Once". In: *Factory, The Magazine of Management* 10.2 (1913), pp. 135–136, 152.

[14] Tarek Hegazy and Amr Ayed. "Neural Network Model for Parametric Cost Estimation of Highway Projects". In: *Journal of Construction Engineering and Management* 124.3 (1998), pp. 210–218. DOI: 10.1061/(ASCE)0733-9364(1998)124:3(210).

[15] Samuel Holloway. "Digital Transformation in Supply Chain Management: A Systematic Literature Review of Trends and Applications". In: *Preprints* (Nov. 2024). DOI: 10.20944/preprints202411.2423.v1. URL: https://doi.org/10.20944/preprints202411.2423.v1.

[16] Dmitry Ivanov, Alexander Tsipoulanidis, and Jörn Schönberger. "Digital Supply Chain, Smart Operations and Industry 4.0: A Decision-Oriented Introduction to the Creation of Value". In: Jan. 2019, pp. 481–526. ISBN: 978-3-319-94312-1. DOI: 10.1007/978-3-319-94313-8_16.

[17] Omid Jadidi, Fatemeh Firouzi, and Shahryar Sorooshian. "A closed-form solution approach for optimal reorder point in economic order quantity models with uncertain demands". In: *Decision Analytics Journal* 16 (2025), p. 100622. ISSN: 2772-6622. DOI: https://doi.org/10.1016/j.dajour.2025.100622. URL: https://www.sciencedirect.com/science/article/pii/S2772662225000785.

[18] Joakim Kembro and Andreas Norrman. "The transformation from manual to smart warehousing: An exploratory study with Swedish retailers". In: *The International Journal of Logistics Management* 33 (June 2022), pp. 107–135. DOI: 10.1108/IJLM-11-2021-0525.

[19] Harold Kerzner. *Project Management: A Systems Approach to Planning, Scheduling, and Controlling*. 12th ed. Hoboken, NJ: John Wiley & Sons, 2017. ISBN: 9781119165354.

[20] Muhammad Gufran Khan, Noor Ul Huda, and Uzair Khaleeq Uz Zaman. "Smart Warehouse Management System: Architecture, Real-Time Implementation and Prototype Design". In: *Machines* 10.2 (2022). ISSN: 2075-1702. DOI: 10.3390/machines10020150. URL: https://www.mdpi.com/2075-1702/10/2/150.

[21] Daniel Kiefer, Florian Grimm, Markus Bauer, et al. "Demand Forecasting Intermittent and Lumpy Time Series: Comparing Statistical, Machine Learning and Deep Learning Methods". In: *Proceedings of the Hawaii International Conference on System Sciences (HICSS)* (2021). DOI: 10.24251/HICSS.2021.196.

[22] Gwang-Hee Kim, Sung-Hoon An, and Kyung-In Kang. "Comparison of Construction Cost Estimating Models Based on Regression Analysis, Neural Networks, and Case-Based Reasoning". In: *Building and Environment* 39.10 (2004), pp. 1235–1242. DOI: 10.1016/j.buildenv.2004.02.013.

[23] Andrey V. Kostenko and Rob J. Hyndman. "A Note on the Categorization of Demand Patterns". In: *Journal of the Operational Research Society* 57.10 (2006), pp. 1256–1257. DOI: 10.1057/palgrave.jors.2602211.

[24] B. Mahadevan, David F. Pyke, and Moritz Fleischmann. "Periodic review, push inventory policies for remanufacturing". In: *European Journal of Operational Research* 151.3 (2003), pp. 536–551. ISSN: 0377-2217. DOI: https://doi.org/10.1016/S0377-2217(02)00627-6. URL: https://www.sciencedirect.com/science/article/pii/S0377221702006276.

[25] Miguel A. Mediavilla et al. "Machine Learning and Deep Learning Models for Demand Forecasting in Supply Chain Management: A Critical Review". In: *Applied System Innovation* 7.5 (2024), p. 93. DOI: 10.3390/asi7050093.

[26] Nhat Quang Nguyen et al. "The designing of warehouse management information system: a case study in helmet company". In: *International Journal of Services and Operations Management* 51.2 (2025), pp. 172–189.

[27] Çerag Pinçe, Laura Turrini, and Joern Meissner. "Intermittent Demand Forecasting for Spare Parts: A Critical Review". In: *Omega* 105 (2021), p. 102513. DOI: 10.1016/j.omega.2021.102513.

[28] Saeed Poormoaied and Zümbül Atan. "A continuous review policy for two complementary products with interrelated demand". In: *Computers & Industrial Engineering* 150 (2020), p. 106980. ISSN: 0360-8352. DOI: https://doi.org/10.1016/j.cie.2020.106980. URL: https://www.sciencedirect.com/science/article/pii/S0360835220306501.

[29] Daniel J. Power. *Decision Support Systems: Concepts and Resources for Managers*. Westport, CT: Quorum Books, 2002. ISBN: 9781567204971.

[30] Bobby Suryo Prakoso et al. "Optimizing a Warehouse Management System (WMS) for Small and Medium Enterprises (SMEs) Using Odoo ERP". In: *2024 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*. 2024, pp. 880–885.

[31] "Predictive Models for Inventory Optimization: A Machine Learning Application for Demand Forecasting at a Construction Supplies Distributor". In: *Future Business Journal* (2026). DOI: 10.1186/s43093-026-00807-8.

[32] Carol A. Ptak and Chad Smith. *Orlicky's Material Requirements Planning*. 3rd ed. New York: McGraw-Hill Education, 2011. ISBN: 9780071755634.

[33] Maciel M. Queiroz and Samuel Fosso Wamba. "Blockchain adoption challenges in supply chain: An empirical investigation of the main drivers in India and the USA". In: *International Journal of Information Management* 46 (2019), pp. 70–82. ISSN: 0268-4012. DOI: https://doi.org/10.1016/j.ijinfomgt.2018.11.021. URL: https://www.sciencedirect.com/science/article/pii/S0268401218309447.

[34] Lorenzo Ragazzini, Elisa Negri, and Marco Macchi. "Low-cost data integration framework for integration with simulation-as-a-service digital twins for make-to-order manufacturing SME". In: *IET Conference Proceedings* 2025 (28 2025), pp. 167–173. DOI: 10.1049/icp.2025.3684.

[35] G. Richards. *Warehouse Management: The Definitive Guide to Improving Efficiency and Minimizing Costs in the Modern Warehouse*. Kogan Page, 2021. ISBN: 9781789668414. URL: https://books.google.pt/books?id=R3RJEAAAQBAJ.

[36] Reuben Schuitemaker and Xun Xu. "Product traceability in manufacturing: A technical review". In: *Procedia CIRP* 93 (2020). 53rd CIRP Conference on Manufacturing Systems 2020, pp. 700–705. ISSN: 2212-8271. DOI: https://doi.org/10.1016/j.procir.2020.04.078. URL: https://www.sciencedirect.com/science/article/pii/S2212827120306922.

[37] Mahya Seyedan and Fereshteh Mafakheri. "Predictive Big Data Analytics for Supply Chain Demand Forecasting: Methods, Applications, and Research Opportunities". In: *Journal of Big Data* 7.1 (2020), p. 53. DOI: 10.1186/s40537-020-00329-2.

[38] Edward A. Silver, David F. Pyke, and Douglas J. Thomas. *Inventory and Production Management in Supply Chains*. CRC Press, 2016.

[39] D. Simchi-Levi, P. Kaminsky, and E. Simchi-Levi. *Managing The Supply Chain: Definitive Guide*. McGraw-Hill Education (India) Pvt Limited, 2004. ISBN: 9780070598584. URL: https://books.google.pt/books?id=KQEc_iCXUT8C.

[40] Ralph H. Sprague. "A Framework for the Development of Decision Support Systems". In: *MIS Quarterly* 4.4 (1980), pp. 1–26. DOI: 10.2307/248957.

[41] Aris A. Syntetos and John E. Boylan. "The Accuracy of Intermittent Demand Estimates". In: *International Journal of Forecasting* 21.2 (2005), pp. 303–314. DOI: 10.1016/j.ijforecast.2004.10.001.

[42] Aris A. Syntetos, John E. Boylan, and J. D. Croston. "On the Categorization of Demand Patterns". In: *Journal of the Operational Research Society* 56.5 (2005), pp. 495–503. DOI: 10.1057/palgrave.jors.2601841.

[43] Guang Hai Tang and Hui Zeng. "Visualisation technology in digital intelligent warehouse management system". In: 12.4 (2021). ISSN: 1741-847X. DOI: 10.1504/ijguc.2021.119556. URL: https://doi.org/10.1504/ijguc.2021.119556.

[44] Ruud H. Teunter, Aris A. Syntetos, and M. Zied Babai. "Intermittent Demand: Linking Forecasting to Inventory Obsolescence". In: *European Journal of Operational Research* 214.3 (2011), pp. 606–615. DOI: 10.1016/j.ejor.2011.05.018.

[45] Evangelos Theodorou, Evangelos Spiliotis, et al. "Machine Learning Algorithms in Intermittent Demand Forecasting: A Review". In: *International Journal of Production Research* (2025). DOI: 10.1080/00207543.2025.2578701.

[46] Lihle N. Tikwayo and Tebello N. D. Mathaba. "Applications of Industry 4.0 Technologies in Warehouse Management: A Systematic Literature Review". In: *Logistics* 7.2 (2023). ISSN: 2305-6290. DOI: 10.3390/logistics7020024. URL: https://www.mdpi.com/2305-6290/7/2/24.

[47] Peter C. Verhoef et al. "Digital transformation: A multidisciplinary reflection and research agenda". In: *Journal of Business Research* 122 (2021), pp. 889–901. ISSN: 0148-2963. DOI: https://doi.org/10.1016/j.jbusres.2019.09.022. URL: https://www.sciencedirect.com/science/article/pii/S0148296319305478.

[48] Gregory Vial. "Understanding digital transformation: A review and a research agenda". In: *The Journal of Strategic Information Systems* 28.2 (2019). SI: Review issue, pp. 118–144. ISSN: 0963-8687. DOI: https://doi.org/10.1016/j.jsis.2019.01.003. URL: https://www.sciencedirect.com/science/article/pii/S0963868717302196.

[49] N. V. Vodolazskaya, E. V. Kopytova, and A. A. Antonova. "Digital Technologies for Automation of Warehouse Accounting of Assembly Components". In: *Journal of Machinery Manufacture and Reliability* 54.Suppl 2 (2025), S219–S229. DOI: 10.1134/S105261882570133X. URL: https://doi.org/10.1134/S105261882570133X.

[50] Karl S.R. Warner and Maximilian Wäger. "Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal". In: *Long Range Planning* 52.3 (2019), pp. 326–349. ISSN: 0024-6301. DOI: https://doi.org/10.1016/j.lrp.2018.12.001. URL: https://www.sciencedirect.com/science/article/pii/S0024630117303710.

[51] Ray Y. Zhong, Xun Xu, and Lihui Wang. "IoT-enabled Smart Factory Visibility and Traceability Using Laser-scanners". In: *Procedia Manufacturing* 10 (2017). 45th SME North American Manufacturing Research Conference, NAMRC 45, LA, USA, pp. 1–14. ISSN: 2351-9789. DOI: https://doi.org/10.1016/j.promfg.2017.07.103. URL: https://www.sciencedirect.com/science/article/pii/S235197891730286X.

Appendix A

Baseline Questionnaire

This appendix reproduces the questionnaire administered to the operational staff of the company, whose aggregated results are presented in Section 3.5. The questionnaire was answered anonymously, and the responses are reported in aggregate form to preserve the confidentiality of the respondents.

Section A: Role and Experience

1. What is your main role?

   ☐ Site director / site supervisor

   ☐ Field technician

   ☐ Procurement / purchasing

   ☐ Warehouse

   ☐ Site administration

   ☐ Other: _____

2. How long have you worked at the company?

   ☐ Less than 1 year

   ☐ 1 to 3 years

   ☐ 3 to 10 years

   ☐ More than 10 years

Section B: Material Shortages

3. In the last 4 weeks, how often did one of your projects (or a project you followed) have to wait for material that was not available?

☐ Never ☐ 1 to 2 times ☐ 3 to 5 times ☐ More than 5 times

4. When a shortage occurs, how long, on average, is the team unable to proceed with that work?

   ☐ Less than 2 hours ☐ Half a day ☐ 1 day ☐ Several days

5. When a shortage occurs, how many workers are typically halted or redeployed to another task?

   ☐ 1 ☐ 2 to 3 ☐ 4 to 6 ☐ More than 6

6. Which materials most frequently cause these stoppages?

   _____

Section C: Procurement and Deliveries

7. In the last 4 weeks, how many times was it necessary to place an urgent order (with additional cost) because the material had been requested late?

   ☐ Never ☐ 1 to 2 ☐ 3 to 5 ☐ More than 5

8. How often does an order get lost or forgotten in the follow-up (only noticed when the material is missing)?

   ☐ Never ☐ Rarely ☐ Sometimes ☐ Frequently

9. How often do supplier deliveries arrive later than promised?

   ☐ Never ☐ Rarely ☐ Sometimes ☐ Frequently ☐ Almost always

Section D: Visibility and Impact

10. Can you, today, autonomously check the status of your orders (what was requested, what has arrived)?

    ☐ Yes, easily ☐ With difficulty ☐ No, I depend on someone else

11. In your perception, what is the impact of these situations on projects? (1 = negligible, 5 = very serious)

    Project delays: ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

    Additional costs: ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

    Client relationship: ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

12. In your opinion, what is the main cause of these problems?

    _____

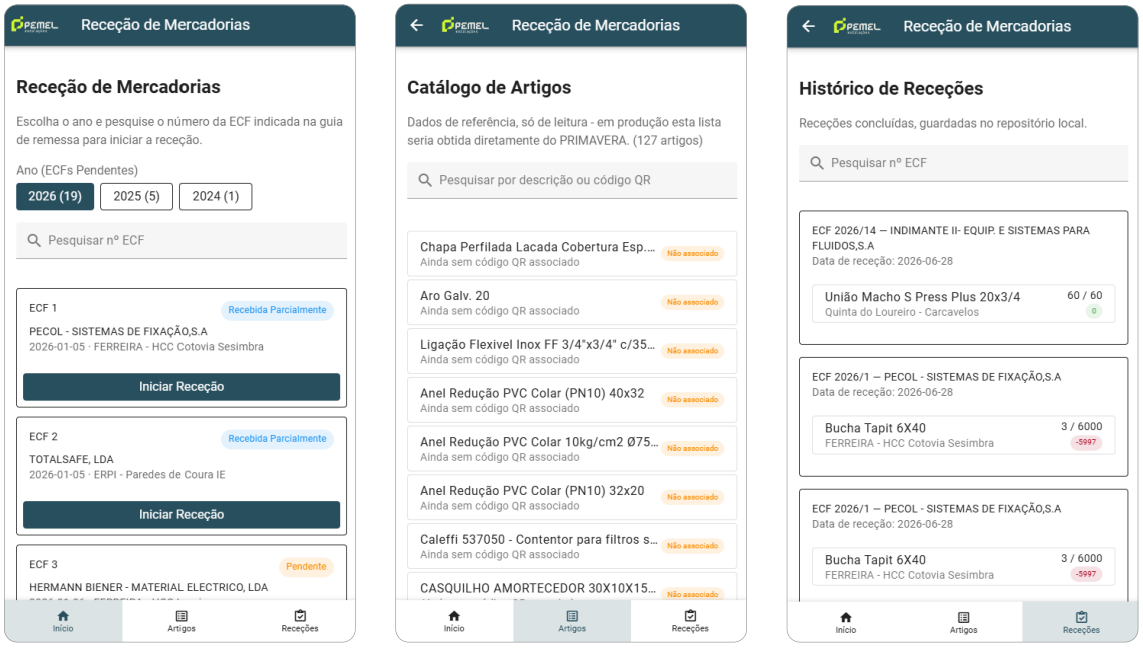
Appendix B

Prototype Application: QR-Based Reception Workflow

This appendix documents the prototype application developed to implement the order-driven QR data-capture process described in Section 4.6.1. Its purpose is not to serve as a user manual but to provide evidence that the reception workflow specified conceptually in Chapter 4 was realized as a functioning application, and to show how each stage of that workflow is materialized in the interface. The application is organized into three areas, presented in Section B.1, and implements the step-by-step reception process presented in Section B.2, which follows the same logical flow as the workflow diagram of Figure 4.3.

B.1 Application Structure

The application is organized into three main areas, accessible from the bottom navigation bar and shown in Figure B.1. The *Início* (Home) area lists the outstanding purchase orders (ECF) awaiting reception, grouped by year, from which a reception is initiated. The *Artigos* (Catalogue) area provides a read-only reference list of the articles, which in production is obtained directly from the PRIMAVERA database, indicating for each whether a QR code has yet been associated with it. The *Receções* (Receptions) area lists the completed receptions stored in the local repository, each with its reconciled quantities and reception date.



(a) Home: pending ECFs. (b) Catalogue: article reference. (c) Receptions: stored history.

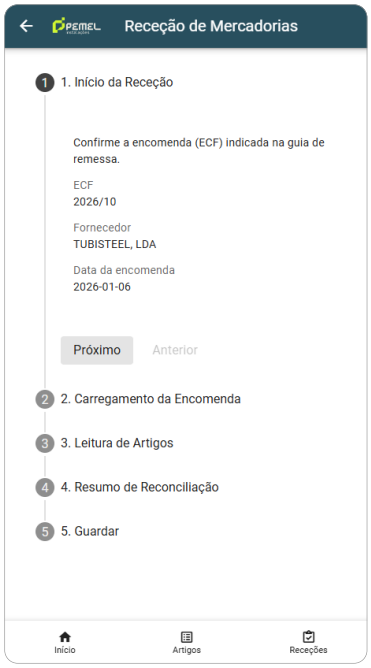
Figure B.1: The three main areas of the prototype application.

B.2 The Reception Workflow

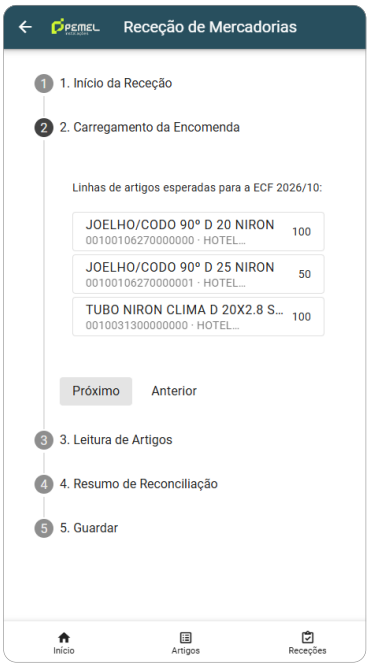
Initiating a reception from the Home area opens the step-by-step process illustrated in Figure B.2. The operator first confirms the purchase order identified from the delivery note (step 1, *Início da Receção*), whereupon the application retrieves the expected article lines and ordered quantities for that order (step 2, *Carregamento da Encomenda*). The operator then scans the QR code of each received article (step 3, *Leitura de Artigos*) and when a scanned code is not yet recognized, the first-time association procedure is invoked, prompting the operator to link the code to the corresponding article, after which the association is memorized for future scans.

On completion of scanning, the application presents the reconciliation summary (step 4, *Resumo de Reconciliação*), which lists each article with its ordered and received quantities and flags any discrepancy between them, alongside the recorded reception date. This reconciliation is the operational output of the process: beyond confirming the reception, it captures the discrepancy and timing data, ordered versus received quantities and the actual reception date, that subsequently feed the characterization of supplier reliability and lead time. The reconciled record is then stored in the local repository (step 5, *Guardar*).

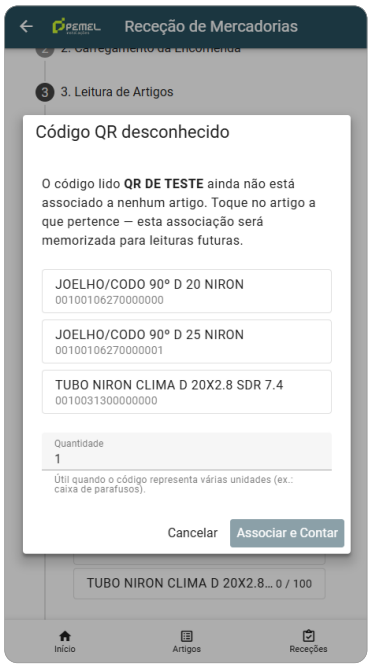
Taken together, these screens demonstrate that the reception workflow defined in Chapter 4 is fully operational in the prototype. Beyond confirming feasibility, the application establishes the practical mechanism by which reliable, structured reception data, the foundation of the data layer



(a) Step 1: confirm the ECF from the delivery note.



(b) Step 2: expected article lines are loaded.



(c) Step 3: first-time association of an unrecognized QR code.



(d) Step 4: reconciliation of received against ordered quantities.

Figure B.2: Step-by-step reception workflow, from order confirmation to reconciliation. The sequence mirrors the process diagram of Figure 4.3.

identified throughout this work, can be captured at source with minimal manual effort.

Appendix C

Monitoring Dashboard

This appendix documents the monitoring dashboard developed as the decision-support layer described in Section 4.6.2. Its purpose is not to serve as a user manual but to provide evidence that the visibility and measurement capabilities specified conceptually in Chapter 4 were realized as a functioning prototype, and to show how each page addresses a specific deficiency identified in the diagnosis. The dashboard was built in Power BI over data extracted from the PRIMAVERA SQL Server database, and is organized into three pages, presented in Section C.1. A note on the underlying data model and the indicators computed is given in Section C.2.

C.1 Dashboard Pages

The dashboard comprises three pages, each accessible from the side navigation and each directed at a distinct operational need. They are shown in Figures 4.5 to C.2.

The overview page, in Figure 4.5, gives the consolidated status of the purchase orders. Its indicators report the number of orders that are open, partial, and delayed, together with the pending value, and a dedicated element highlights the delayed orders, defined as those not yet fully received more than fifteen days after being placed. A ring chart breaks the orders down by engineering speciality, and a bar chart shows the delayed orders by speciality, so that the area under most pressure is immediately apparent. Two tables complete the page, one listing the pending orders by supplier and one showing the most recent receptions. This page resolves the absence of an autonomous, consolidated view of order status identified in the diagnosis.

The orders page, in Figure C.1, details the purchase orders individually. It reports the ordered and received quantities, the resulting completion rate, and the pending value of each order, with a progress element that makes the degree of fulfilment visible at a glance. Ordering the detail by pending value surfaces the orders with the greatest quantity of material still to arrive, which are

the ones that most warrant follow-up.

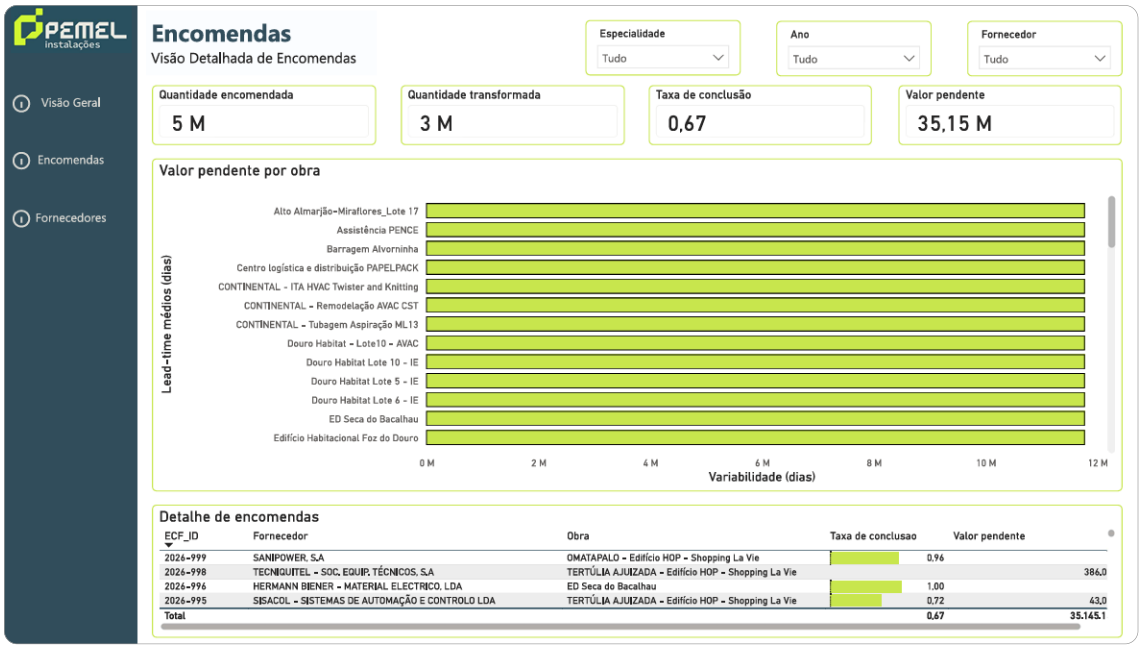


Figure C.1: Orders page, detailing each purchase order with its completion rate and pending value.

The suppliers page, in Figure C.2, characterizes the suppliers and is the page most directly connected to the model. Its central element is a risk quadrant, a scatter plot placing each supplier by its average lead time on the horizontal axis and its lead-time variability on the vertical axis, with the size of each point reflecting the value ordered. Reference lines at the means divide the plane into four regions, so that the suppliers whose deliveries are both slow and erratic, in the upper-right region, are distinguished at a glance from those that are fast and consistent. A supporting table reports, per supplier, the mean lead time, its standard deviation, the ninetieth percentile, and the value ordered. It is the lead-time variability read from this page that feeds the safety component of the minimum stock developed in this work, so that the dashboard is not only a visualization but also a source of a model input.

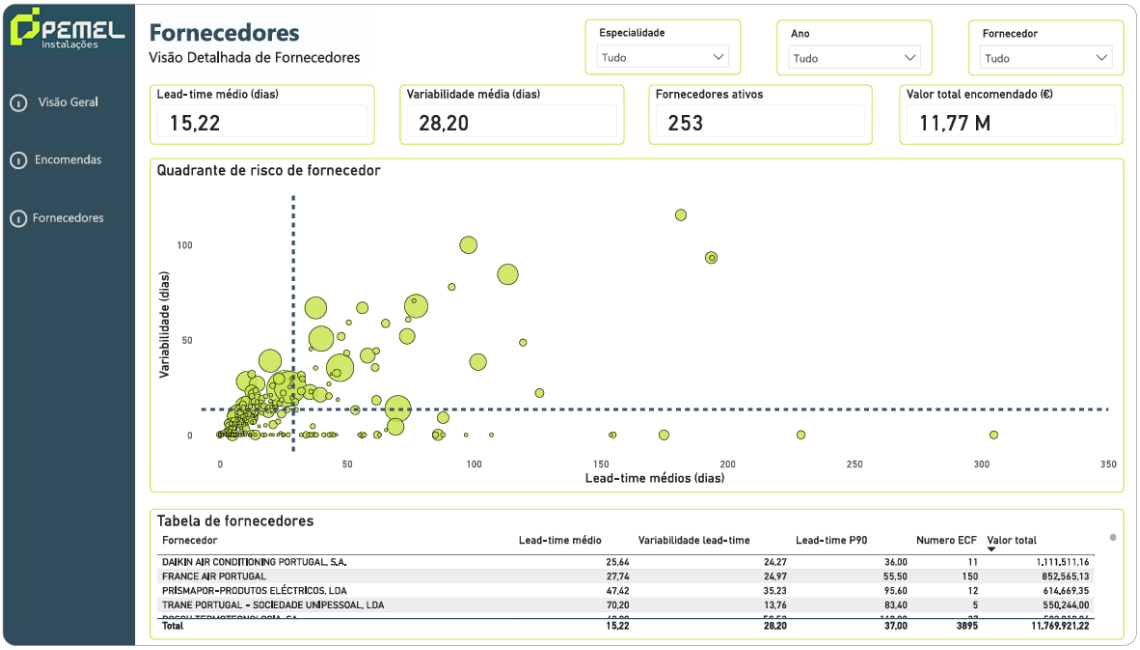


Figure C.2: Suppliers page, with the lead-time risk quadrant and the supporting per-supplier table.

C.2 Data Model and Indicators

The dashboard is built on data extracted from the PRIMAVERA SQL Server database through direct queries, exported and prepared for loading into the reporting tool. The order and reception records supply the fulfilment status and the lead time, computed as the difference between the order and receipt dates, while the purchase-order lines supply the ordered and received quantities and the value. The engineering speciality of each order is derived from the project code, whose structure encodes it. The indicators reported are those the data support reliably, the lead time and its variability, the order-fulfilment status, the pending value, and their breakdowns by supplier, speciality, and period.

Two indicators that a procurement dashboard might be expected to carry were deliberately not included, because the data do not support them. An on-time-in-full indicator would require comparing each receipt against a promised delivery date, but that date is recorded in the ERP as an unpopulated default rather than as a genuine commitment, so any punctuality computed against it would be an artefact. A procurement-cycle indicator, such as the time between an internal requisition and the resulting order, could not be constructed because the link between the two documents is not registered in the system. These indicators were therefore omitted rather than computed on unreliable data, and their absence is consistent with, and part of the evidence for, the data-capture gaps that the enabling layer of Section 4.6.1 sets out to close. In its present form the prototype is refreshed by periodic extraction, but because it rests on direct queries over the production database, migration to a scheduled, automatic refresh requires only a change of connection and no change

to the queries or the design.

Appendix D

Data-Processing Scripts

This appendix reproduces the two Python scripts used to process the data extracted from the ERP and to compute the empirical results reported in Chapter 5. Their purpose is not to serve as production software but to document, for reproducibility, the exact procedure behind the material classification and the phase-profile analysis. The scripts share the price-based reconstruction of Section 5.1.2, by which the functional family of each consumed item is recovered from its unit price in the project budget, and each reads, per project, one CSV of dated goods-issue lines together with the corresponding budget spreadsheet.

D.1 ABC Classification Across Projects

The script listed below implements the per-project ABC analysis whose results are reported in Section 5.2. For each project it links every goods-issue line to its functional family through the unit price, aggregates the consumption value by family, and assigns the A, B and C classes on the classical cumulative-value cut-offs. It then consolidates the individual projects into a family-by-project class matrix and computes, for each family, the recurrence of the A/B classification across the projects in which it appears, which is the quantity underlying the significance ranking.

```
import sys, os, glob, warnings
import pandas as pd
import numpy as np
warnings.filterwarnings('ignore')

# --- Obras fabricadas (orçamento preenchido manualmente). Marcadas com * na saí
    ↪ da.
OBRAS_FABRICADAS = {'20242004','20242003','20232005','20232002','20222014','
    ↪ 20222007'}

# --- Mapa das 32 famílias funcionais AVAC (código -> nome)
FAMILIAS = {
    1:'Caldeiras',2:'Aquecedores de Ar',3:'Bombas de Calor',4:'Ventiladores',5:'
        ↪ UTA',
```

```
    6:'Exaustores',7:'Unidades de Janela',8:'Splits',9:'VRF/VRV',10:'Chillers',
    11:'Termostatos',12:'Equipamento de campo',13:'BMS',14:'Caminho de cabos',
    15:'Tubos de PVC',16:'Difusores e Grelhas',17:'Filtros de Ar',18:'Válvulas',
    19:'Tubos',20:'Caixas',21:'Alimentadores',22:'Cabos',23:'Calha Rodapé',
    24:'Quadros Elétricos',25:'Condutas Spiro',26:'Condutas retangulares',
    27:'Grelhas intumescentes',28:'Registos de caudal',29:'Registos Corta Fogo',
    30:'Registos de desenfumagem',31:'Cobre de refrigeração',32:'Cobre de
      ↪ Aquecimento',
}

COLS_PRECO = [11, 21, 31, 28, 17, 34, 36]   # colunas de preço candidatas (auto-
  ↪ deteção)
COL_FAMILIA = 4
LINHA_INICIO = 14
SHEET_ORCAMENTO = 'Orçamento'

# cortes ABC clássicos (dentro de cada obra, sobre o valor acumulado)
ABC_CORTE_A = 0.80
ABC_CORTE_B = 0.95


def mapa_preco_familia(xlsx_path):
    """Extrai {preço: família} do orçamento, com auto-deteção da coluna de preço
      ↪ ."""
    d = pd.read_excel(xlsx_path, sheet_name=SHEET_ORCAMENTO, header=None)
    # escolher a coluna de preço que gera mais correspondências plausíveis
    melhor_col, melhor_n = None, -1
    for col in COLS_PRECO:
        if col >= d.shape[1]:
            continue
        vals = pd.to_numeric(d[col].iloc[LINHA_INICIO:], errors='coerce')
        n = ((vals > 1) & (vals < 100000)).sum()
        if n > melhor_n:
            melhor_n, melhor_col = n, col
    mapa = {}
    fam_atual = None
    for i in range(LINHA_INICIO, len(d)):
        row = d.iloc[i]
        try:
            fam = int(float(row[COL_FAMILIA]))
            if fam in FAMILIAS:
                fam_atual = fam
        except (ValueError, TypeError):
            pass
        try:
            p = round(float(row[melhor_col]), 2)
            if 1 < p < 100000 and fam_atual and p not in mapa:
                mapa[p] = FAMILIAS[fam_atual]
        except (ValueError, TypeError, IndexError):
```

```
            pass
    return mapa


def carregar_autos(csv_path):
    """Lê o CSV de autos (deteta separador, cabeçalho e vírgula decimal)."""
    melhor = None
    for sep in [';', '\t', ',']:
        try:
            tmp = pd.read_csv(csv_path, sep=sep, encoding='utf-8-sig', dtype=str
                ↪ )
            if melhor is None or tmp.shape[1] > melhor.shape[1]:
                melhor = tmp
        except Exception:
            continue
    df = melhor
    # deteta se a 1a linha é dados (sem cabeçalho)
    if not any('quant' in str(c).lower() or 'preco' in str(c).lower()
               or 'data' in str(c).lower() for c in df.columns):
        df = pd.read_csv(csv_path, sep=';', encoding='utf-8-sig', header=None,
            ↪ dtype=str)
    n = df.shape[1]
    df.columns = (['data','auto','itemid','qtd','preco'][:n]
                  if n <= 5 else list(df.columns))
    def num(s):
        s = s.astype(str).str.strip()
        return pd.to_numeric(
            s.where(~s.str.contains(',', na=False), s.str.replace('.', '', regex
                ↪ =False)
                     .str.replace(',', '.', regex=False)), errors='coerce')
    df['qtd'] = num(df['qtd'])
    df['preco'] = num(df['preco']).round(2)
    return df.dropna(subset=['qtd', 'preco'])


def abc_de_uma_obra(autos, mapa_preco):
    """Valor de consumo por família e classificação ABC dentro da obra."""
    a = autos.copy()
    a['familia'] = a['preco'].map(mapa_preco)
    a = a.dropna(subset=['familia'])
    a['valor'] = a['qtd'] * a['preco']
    val = a.groupby('familia')['valor'].sum().sort_values(ascending=False)
    total = val.sum()
    if total <= 0:
        return pd.Series(dtype=str)
    acum = val.cumsum() / total
    classe = acum.apply(lambda x: 'A' if x <= ABC_CORTE_A
                                  else ('B' if x <= ABC_CORTE_B else 'C'))
    return classe
```

```
104
105
106 def main(pasta):
107     subpastas = [d for d in sorted(glob.glob(os.path.join(pasta, '*'))) if os.
        ↪ path.isdir(d)]
108     matriz = {}   # {obra: {familia: classe}}
109     for sp in subpastas:
110         obra = os.path.basename(sp).rstrip('*')
111         csvs = glob.glob(os.path.join(sp, '*.csv'))
112         xls = glob.glob(os.path.join(sp, '*.xlsx'))
113         if not csvs or not xls:
114             print(f"[{obra}] falta CSV ou XLSX, ignorada.")
115             continue
116         autos = carregar_autos(csvs[0])
117         mapa = mapa_preco_familia(xls[0])
118         classe = abc_de_uma_obra(autos, mapa)
119         if len(classe) == 0:
120             print(f"[{obra}] sem consumo ligável, ignorada.")
121             continue
122         matriz[obra] = classe.to_dict()
123         taxa = autos['preco'].map(mapa).notna().mean()
124         print(f"[{obra}]  ligação {taxa:.0%}  |  {len(classe)} famílias "
125               f"(A:{(classe=='A').sum()} B:{(classe=='B').sum()} C:{(classe=='C
                  ↪ ').sum()})")
126
127     if not matriz:
128         print("Sem obras processadas.")
129         return
130
131     # matriz família x obra
132     df = pd.DataFrame(matriz)            # linhas=família, colunas=obra
133     df = df.reindex(sorted(df.columns), axis=1)
134     # marcar obras fabricadas no nome da coluna
135     df.columns = [f"{c}*" if c in OBRAS_FABRICADAS else c for c in df.columns]
136     df.to_csv(os.path.join(pasta, 'matriz_ABC.csv'))
137
138     # recorrência por família
139     rec = pd.DataFrame({
140         'n_obras':   df.notna().sum(axis=1),
141         'n_A':       (df == 'A').sum(axis=1),
142         'n_B':       (df == 'B').sum(axis=1),
143         'n_C':       (df == 'C').sum(axis=1),
144     })
145     rec['n_AB'] = rec['n_A'] + rec['n_B']
146     rec['pct_AB'] = (100 * rec['n_AB'] / rec['n_obras']).round(0)
147     # classe consolidada: alvo se A ou B na maioria das obras em que aparece
148     rec['alvo'] = np.where(rec['pct_AB'] >= 50, 'Alvo (A/B recorrente)', 'Não
        ↪ alvo')
149     rec = rec.sort_values(['pct_AB', 'n_AB'], ascending=False)
```

```
150     rec.to_csv(os.path.join(pasta, 'recorrencia_ABC.csv'))
151
152     print("\n=== RECORRÊNCIA DA CLASSE POR FAMÍLIA ===")
153     print(rec.to_string())
154     print(f"\nMatriz e recorrência gravadas em '{pasta}'.")
155
156
157 if __name__ == '__main__':
158     main(sys.argv[1] if len(sys.argv) > 1 else 'dados_AVAC')
```

D.2 Phase Profiles and Stability Test

The script listed below implements the phase-profiling and stability analysis reported in Section 4.4. For each project it places every consumed item on the project timeline, normalizes that timeline to the unit interval, and bins the consumption into the four equal phases Q1-Q4, so that a family's phase profile is the share of its consumption value falling in each phase. Across projects it then computes, for every family present in at least two projects, the instability index, and classifies each family against the reference bands used.

```
1  import sys, os, glob, warnings
2  import pandas as pd
3  import numpy as np
4  warnings.filterwarnings('ignore')
5
6  # ---------------------------------------------------------------------- CONFIG ----
7  # Mapa dos códigos de família do orçamento (coluna "Código" da folha Orçamento).
8  # Ajustar/estender conforme o mapa mestre da empresa (imagem das 32 famílias
     ↪ AVAC).
9  FAMILIAS = {
10     1: 'Caldeiras', 2: 'Aquecedores de Ar', 3: 'Bombas de Calor', 4: '
         ↪ Ventiladores',
11     5: 'UTA', 6: 'Exaustores', 7: 'Unidades de Janela', 8: 'Splits',
12     9: 'VRF/VRV', 10: 'Chillers', 11: 'Termostatos', 12: 'Equipamento de campo',
13     13: 'BMS', 14: 'Caminho de cabos', 15: 'Tubos de PVC', 16: 'Difusores e
         ↪ Grelhas',
14     17: 'Filtros de Ar', 18: 'Válvulas', 19: 'Tubos', 20: 'Caixas',
15     21: 'Alimentadores', 22: 'Cabos', 23: 'Calha Rodapé', 24: 'Quadros Elétricos
         ↪ ',
16     25: 'Condutas Spiro', 26: 'Condutas retangulares', 27: 'Grelhas
         ↪ intumescentes',
17     28: 'Registos de caudal', 29: 'Registos Corta Fogo', 30: 'Registos de
         ↪ desenfumagem',
18     31: 'Cobre de refrigeração', 32: 'Cobre de Aquecimento',
19 }
20
21 # Colunas do Excel (folha "Orçamento") onde procurar preços unitários.
```

```
# A leitura testa várias e usa a primeira plausível por linha.
COLS_PRECO = [11, 21, 31, 28, 17]
COL_FAMILIA = 4          # coluna "Código" (número 1-32) na folha Orçamento
LINHA_INICIO = 14        # primeira linha de dados (após cabeçalhos)
SHEET_ORCAMENTO = 'Orçamento'

N_FASES = 4              # Q1..Q4 (mudar para 3, 5, etc. se quiser outra granularidade)
LIMIAR_ESTAVEL = 15      # dif. média (pontos percentuais) abaixo da qual = estável
LIMIAR_MEDIO = 25        # entre LIMIAR_ESTAVEL e este = parcialmente estável


# ---------------------------------------------------------- EXCEL -> FAMÍLIA ---
def mapa_preco_familia(xlsx_path):
    """Extrai {preço_unitário: família} do orçamento Excel de uma obra."""
    d = pd.read_excel(xlsx_path, sheet_name=SHEET_ORCAMENTO, header=None)
    mapa = {}
    familia_actual = None
    for i in range(LINHA_INICIO, len(d)):
        row = d.iloc[i]
        # atualizar família corrente se a coluna Código tiver um número válido
        try:
            fam = int(float(row[COL_FAMILIA]))
            if fam in FAMILIAS:
                familia_actual = fam
        except (ValueError, TypeError):
            pass
        # registar preços plausíveis desta linha
        for col in COLS_PRECO:
            try:
                p = round(float(row[col]), 2)
                if 1 < p < 100000 and familia_actual and p not in mapa:
                    mapa[p] = FAMILIAS[familia_actual]
            except (ValueError, TypeError, IndexError):
                pass
    return mapa


# ---------------------------------------------------------- AUTOS -> PERFIL ----
def carregar_autos(csv_path):
    """Lê o CSV de autos exportado do SSMS (separador ; ou tab, com ou sem header)."""
    # Detetar se a primeira linha é cabeçalho ou já dados. Se a primeira célula
    # parece uma data ou um número, o ficheiro não tem cabeçalho.
    with open(csv_path, encoding='utf-8-sig') as fh:
        primeira = fh.readline()
    import re
    prim_celula = re.split(r'[;\t,]', primeira.strip())[0].strip().strip('"')
```

```
    tem_header = not (re.match(r'^\d{4}-\d{2}-\d{2}', prim_celula) or
                      re.match(r'^\d+[.,]?\d*$', prim_celula))

    # Escolher o separador que produz MAIS colunas.
    melhor = None
    for sep in [';', '\t', ',']:
        try:
            tmp = pd.read_csv(csv_path, sep=sep, encoding='utf-8-sig', dtype=str
                ↪ ,
                              header=0 if tem_header else None)
            if melhor is None or tmp.shape[1] > melhor.shape[1]:
                melhor = tmp
        except Exception:
            continue
    if melhor is None or melhor.shape[1] < 4:
        raise ValueError(f"Não consegui ler {csv_path} com >=4 colunas.")
    df = melhor

    if tem_header:
        cols = {str(c).lower().strip(): c for c in df.columns}
        def pick(*names):
            for n in names:
                if n in cols:
                    return cols[n]
            return None
        c_data = pick('dataauto', 'data')
        c_qtd  = pick('quantidade', 'amquant', 'qtd', 'quant')
        c_prec = pick('precounitario', 'precunit', 'preco')
        if c_data is None or c_qtd is None or c_prec is None:
            colnames = list(df.columns)
            c_data = c_data or colnames[0]
            c_qtd  = c_qtd  or colnames[3]
            c_prec = c_prec or colnames[4]
    else:
        # sem cabeçalho: posições fixas da query (DataAuto, Auto, ItemID,
          ↪ Quantidade, PrecoUnitario)
        colnames = list(df.columns)
        c_data, c_qtd, c_prec = colnames[0], colnames[3], colnames[4]

    df = df.rename(columns={c_data: 'data', c_qtd: 'qtd', c_prec: 'preco'})

    # Números podem vir com vírgula decimal (formato PT). Normalizar.
    def to_num(s):
        s = s.astype(str).str.replace('\u00a0', '', regex=False).str.strip()
        s = s.str.replace('.', '', regex=False).str.replace(',', '.', regex=
          ↪ False) \
             .where(s.str.contains(',', na=False), s)
        return pd.to_numeric(s, errors='coerce')

```

```
    # Datas: o SSMS exporta em ISO (ano-mês-dia). NÃO usar dayfirst.
    df['data'] = pd.to_datetime(df['data'], errors='coerce')
    df['qtd'] = to_num(df['qtd'])
    df['preco'] = to_num(df['preco']).round(2)
    df = df.dropna(subset=['data', 'qtd', 'preco'])
    if df.empty:
        raise ValueError(f"{csv_path}: 0 linhas válidas após limpeza.")
    return df


def perfil_de_fase(autos, mapa_preco):
    """Liga autos a família por preço; devolve (perfil, taxa_match,
       ↪ valor_por_familia)."""
    autos = autos.copy()
    autos['familia'] = autos['preco'].map(mapa_preco)
    taxa = autos['familia'].notna().mean()
    autos = autos.dropna(subset=['familia'])
    autos['valor'] = autos['qtd'] * autos['preco']
    # posição temporal normalizada [0,1] dentro da obra
    span = (autos['data'].max() - autos['data'].min()).days
    if span <= 0:
        return pd.DataFrame(), taxa, valor_familia
    autos['t'] = (autos['data'] - autos['data'].min()).dt.days / span
    bins = np.linspace(0, 1, N_FASES + 1)
    labels = [f'Q{i+1}' for i in range(N_FASES)]
    autos['fase'] = pd.cut(autos['t'], bins=bins, labels=labels, include_lowest=
       ↪ True)
    # perfil = % do valor de cada família em cada fase
    piv = (autos.groupby(['familia', 'fase'], observed=True)['valor'].sum()
                .unstack(fill_value=0).reindex(columns=labels, fill_value=0))
    perfil = piv.div(piv.sum(axis=1), axis=0)   # normaliza por família
    return perfil, taxa


# ---------------------------------------------------- TESTE DE ESTABILIDADE ----
def teste_estabilidade(perfis_por_obra):
    """Compara perfis de família entre obras. perfis_por_obra: {obra: DataFrame
       ↪ }."""
    labels = [f'Q{i+1}' for i in range(N_FASES)]
    # famílias presentes em >=2 obras
    contagem = {}
    for obra, perf in perfis_por_obra.items():
        for fam in perf.index:
            contagem.setdefault(fam, []).append(obra)
    comuns = {f: obras for f, obras in contagem.items() if len(obras) >= 2}

    linhas = []
    for fam, obras in sorted(comuns.items()):
```

```
159         mat = np.vstack([perfis_por_obra[o].loc[fam, labels].values for o in
                ↪ obras])
160         # instabilidade = média dos desvios-padrão por fase (em pontos
                ↪ percentuais)
161         instab = mat.std(axis=0).mean() * 100
162         perfil_medio = mat.mean(axis=0)
163         estado = ('ESTÁVEL' if instab < LIMIAR_ESTAVEL
164                   else 'PARCIAL' if instab < LIMIAR_MEDIO else 'INSTÁVEL')
165         linhas.append({
166             'familia': fam, 'n_obras': len(obras),
167             **{q: round(perfil_medio[i] * 100) for i, q in enumerate(labels)},
168             'instabilidade_pp': round(instab, 1), 'classificacao': estado,
169         })
170     return pd.DataFrame(linhas).sort_values('instabilidade_pp')
171
172
173 # ------------------------------------------------------------------- MAIN ----
174 def main(pasta_dados):
175     subpastas = [d for d in sorted(glob.glob(os.path.join(pasta_dados, '*')))
176                  if os.path.isdir(d)]
177     if not subpastas:
178         print(f"Sem subpastas de obra em '{pasta_dados}'.")
179         return
180
181     perfis_por_obra = {}
182     relatorio = ["=" * 70, "ANÁLISE DE FASEAMENTO POR FAMÍLIA (todas as famílias
            ↪ )", "=" * 70]
183
184     for sp in subpastas:
185         obra = os.path.basename(sp).rstrip('*')
186         csvs = glob.glob(os.path.join(sp, '*.csv'))
187         xlsxs = glob.glob(os.path.join(sp, '*.xlsx'))
188         if not csvs or not xlsxs:
189             relatorio.append(f"\n[{obra}] falta CSV ou XLSX, ignorada.")
190             continue
191         autos = carregar_autos(csvs[0])
192         mapa = mapa_preco_familia(xlsxs[0])
193         perfil, taxa = perfil_de_fase(autos, mapa)
194         if perfil.empty:
195             relatorio.append(f"\n[{obra}] sem span temporal utilizável.")
196             continue
197         perfis_por_obra[obra] = perfil
198         relatorio.append(f"\n[{obra}]  ligação preço->família: {taxa:.0%}  "
199                          f"({len(perfil)} famílias, {autos['data'].nunique()}
                                 ↪ autos)")
200         for fam, row in perfil.iterrows():
201             perfil_str = "  ".join(f"{q}={row[q]:.0%}" for q in perfil.columns)
202             relatorio.append(f"    {fam:24s} {perfil_str}")
203
```

```
    if len(perfis_por_obra) >= 2:
        tab = teste_estabilidade(perfis_por_obra)
        relatorio += ["\n" + "=" * 70,
                      f"TESTE DE ESTABILIDADE -- {len(perfis_por_obra)} obras (
                         ↪ todas as famílias)",
                      "=" * 70, tab.to_string(index=False)]
        n_e = (tab['classificacao'] == 'ESTÁVEL').sum()
        n_p = (tab['classificacao'] == 'PARCIAL').sum()
        relatorio.append(f"\nResumo: {n_e} estáveis, {n_p} parciais, "
                         f"{len(tab) - n_e - n_p} instáveis (de {len(tab)} famí
                            ↪ lias comuns).")
        tab.to_csv(os.path.join(pasta_dados, 'estabilidade_familias.csv'), index
           ↪ =False)
        relatorio.append("Nota: analisadas TODAS as famílias, sem filtro de
           ↪ classe. "
                         "O cruzamento com o valor (ABC) faz-se separadamente.")
    else:
        relatorio.append("\n(Só 1 obra com dados, sem teste de estabilidade.)")

    texto = "\n".join(relatorio)
    print(texto)
    with open(os.path.join(pasta_dados, 'relatorio_faseamento.txt'), 'w',
       ↪ encoding='utf-8') as f:
        f.write(texto)
    print(f"\n>> Estabilidade por família gravada em 'estabilidade_familias.csv
       ↪ '.")


if __name__ == '__main__':
    import sys
    main(sys.argv[1] if len(sys.argv) > 1 else 'dados_AVAC')
```