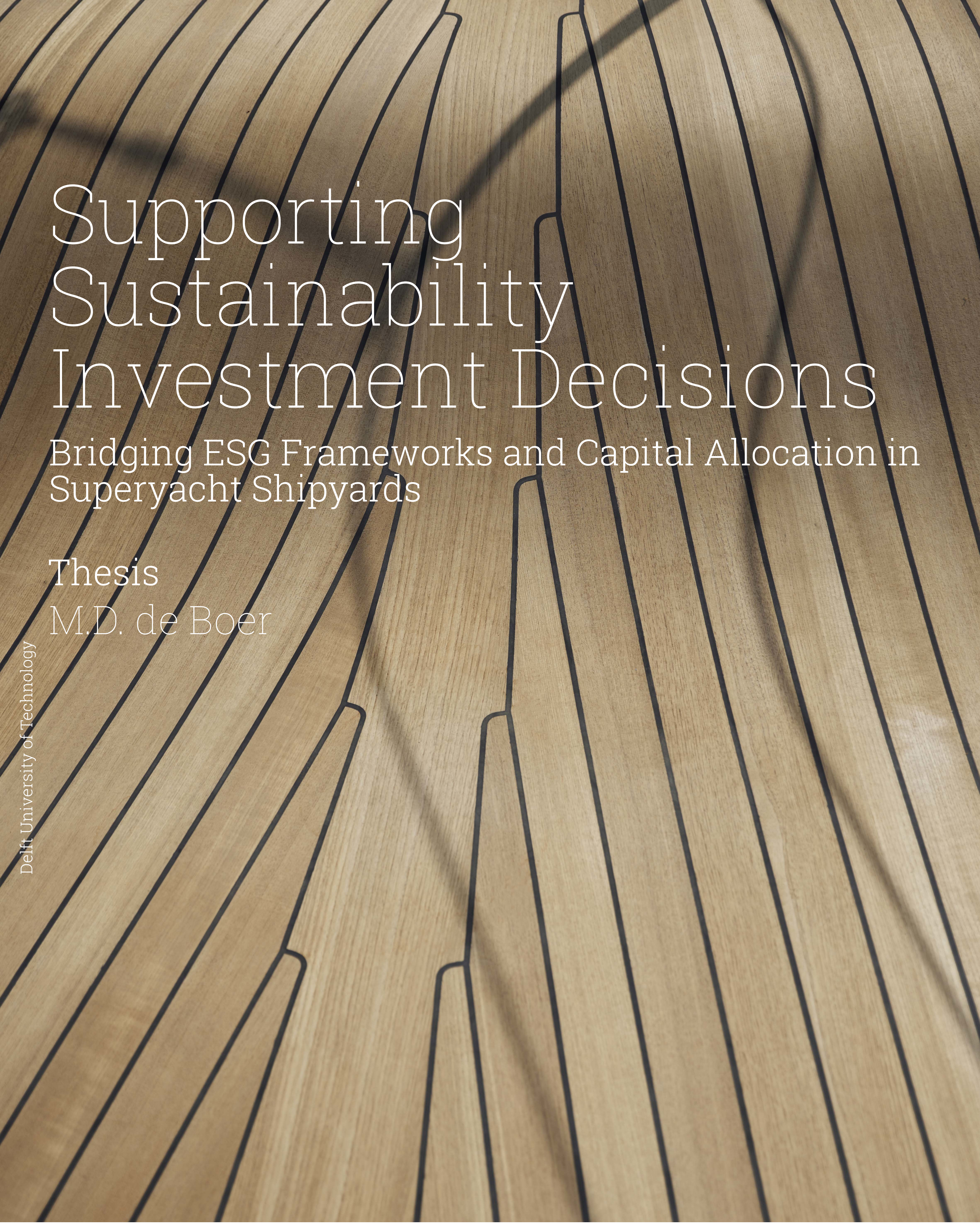


Supporting Sustainability Investment Decisions

Bridging ESG Frameworks and Capital Allocation in Superyacht Shipyards

Thesis

M.D. de Boer

Delft University of Technology

Thesis for the degree of MSc in Marine Technology in the specialisation of Maritime Operations and Management.

Supporting Sustainability Investment Decisions

by

M.D. de Boer

Performed at

Koninklijke de Vries Scheepsbouw B.V.

This thesis [MT.25/26.034.M] is classified as confidential in accordance with the general conditions for projects performed by the TU Delft.

01-07-2026

Company supervisor
Responsible supervisor: ir. C. van de Kerk

Thesis exam committee
Chair: Dr.ir. J.F.J. Pruijn
Committee member: Dr. J.M. Vleugel
Committee member: Dr. A. Napoleone
Company member: ir. C. van de Kerk

Author Details
Student number: 5076706

Cover: Feadship (2023)

TU Delft

Summary

Superyacht shipyards face growing regulatory, social, and reputational pressure to improve their Environmental, Social, and Governance (ESG) performance. In response, yards are confronted with a widening set of sustainability investment proposals that differ strongly in scale, impact pathway, and observability. Existing assessment approaches are poorly matched to this decision context. Lifecycle and valuation methods require levels of causal detail that early-stage yard-level investments cannot support, while ESG reporting frameworks describe organisational states without providing a basis for comparing discrete alternatives under scarcity. The central problem is therefore not a lack of indicators or data, but the absence of a feasible decision logic.

This thesis develops and applies a yard-level decision-making framework that supports transparent and accountable comparison of heterogeneous sustainability investments under deep uncertainty in costs, impacts, and future operating conditions. The framework combines a structured indicator basis derived from ESRS 2 and the academic literature with a System Dynamics Representation (SDR) for constructing investment consequences. Performance functions convert heterogeneous effects into comparable scores, the Analytic Hierarchy Process (AHP) elicits criterion weights, and additive Multi-Attribute Value Theory (MAVT) aggregates the results. Scenario analysis represents future-context uncertainty without assigning probabilities, minimax regret supports robust option selection, and Monte Carlo perturbation assesses framework robustness.

The framework is applied to Koninklijke de Vries Scheepsbouw. Four investments are compared across four scenarios spanning social and regulatory pressure, namely company-car electrification, cybersecurity training, yard security measures, and replacement of the final ramp with a dry dock. Under baseline aggregation, the dry dock performs best, while under minimax regret cybersecurity training is preferred. The decisive factor is therefore not the aggregation rule, but whether decision-makers regard adverse scenarios as sufficiently plausible to influence current choice. Robustness analysis shows that the recommendation is stable under variation in AHP weights and performance-function shapes, but more sensitive to scenario-conditioned impact assumptions. The SDR proved particularly valuable during elicitation by surfacing legitimacy, governance, and workforce pathways that direct assessment had initially overlooked.

The main contribution of this thesis is conceptual as well as practical. Conceptually, it reframes yard-level sustainability investment evaluation as a decision-logic problem rather than a measurement problem. The persistent gap between ESG information and capital allocation follows from a structural mismatch between backward-looking reporting logic and forward-looking investment logic. Practically, it demonstrates that a non-probabilistic, internally-valued, scenario-based framework can be operationalised in an executive shipyard setting, retaining heterogeneity across criteria without forcing premature monetisation. The framework does not eliminate uncertainty or replace managerial judgement. Rather, it disciplines deliberation by making assumptions, trade-offs, and points of disagreement explicit and contestable.

Preface

This thesis concludes my Master of Science (MSc) in Marine Technology at Delft University of Technology, in the specialisation of Maritime Operations and Management. The research was carried out at Koninklijke de Vries Scheepsbouw in Aalsmeer, where I was given the opportunity to study sustainability investment decisions in a setting that combines analytical complexity with genuine organisational stakes.

I would like to thank C. van de Kerk for her daily supervision at the yard, and for the openness with which the company supported this work. I am particularly grateful for the connections she opened within the organisation, which allowed me to conduct the case study at decision-maker level. I also thank the directors at De Vries Scheepsbouw who participated in the elicitation sessions; their willingness to engage with an unfamiliar framework shaped the thesis in ways that go beyond what the case chapter records.

I am especially grateful to J. Pruijn for his guidance throughout the project. The clarity of your feedback and the steadiness of our discussions made the process a pleasant one.

M.D. de Boer
Delft, June 2026

Contents

List of Abbreviations

AHP Analytic Hierarchy Process. 12–15, 18, 19, 23–26, 31, 32, 35–38, 43–45, 49, 50

CAPEX Capital Expenditure. 20, 31, 33, 34, 46
CEO Chief Executive Officer. 2
CFO Chief Financial Officer. 2
CLD Causal Loop Diagram. 21, 22, 30
CR Consistency Ratio. 24
CSRD Corporate Sustainability Reporting Directive. 1, 2, 8

DMDU Decision Making under Deep Uncertainty. 9, 13, 42

EBITDA Earnings Before Interest, Taxes, Depreciation and Amortisation. 20, 31, 33, 34, 42, 46
ELECTRE Élimination et Choix Traduisant la Réalité. 13
ESG Environmental, Social and Governance. 2–6, 8, 12, 13, 41, 44, 48–50
ESRS European Sustainability Reporting Standard. 2, 8, 12, 15
ETO Engineer-To-Order. 3, 5, 7, 45
ETS Emissions Trading System. 1, 2

GHG Greenhouse Gas. 20, 31, 34, 35, 38, 41

LCA Life Cycle Assessment. 3, 6, 7
LCC Life Cycle Costing. 7

MAVT Multi Attribute Value Theory. 8, 13–15, 18, 23, 24, 44, 45, 48
MCDA Multi Criteria Decision Analysis. 8, 12, 14, 15
MSc Master of Science. ii

NPV Net Present Value. 7

PROMETHEE Preference Ranking Organization Method for Enrichment Evaluations. 13

SDR System Dynamics Representation. 11, 20–22, 26, 29, 30, 38, 42–45, 48–50
SFR Stock–Flow Representation. 21, 22, 30, 31
SROI Social Return On Investment. 3

TBL Triple Bottom Line. 8

YETI Yacht Environmental Transparency Index. 7

1 Introduction

Across industrial sectors, sustainability has become increasingly important for regulation, investment, and organisational legitimacy. Environmental, social, and governance performance now shape how industrial firms are assessed by regulators, customers, employees, and the wider public (Gillan et al., 2021). These expectations are particularly relevant for sectors that are highly visible, resource-intensive, or associated with contested forms of consumption.

The superyacht sector occupies such a position. Superyachts are highly visible symbols of private wealth and are therefore easily drawn into wider societal debates on inequality, consumption, and sustainability (Hirtenlehner, 2025; Lynch et al., 2019). The relevance of this scrutiny does not depend only on the sector's aggregate environmental footprint. Rather, the visibility and symbolic character of superyachts can make the sector seem environmentally and socially problematic. This perception holds even where the underlying impacts are specific, dispersed, or hard to compare with other maritime sectors (Hirtenlehner, 2025; Visser, 2021). For superyacht builders, this creates a legitimacy challenge that extends beyond vessel performance and increasingly includes the organisational choices made in and around shipyards.

Regulatory developments reinforce this pressure through different mechanisms. Environmental regulation in Svalbard and the inclusion of large yachts in the European Union Emissions Trading System (ETS) increase the operational and commercial relevance of emissions performance (Boscawen, 2025; European Parliament and Council of the European Union, 2023). While these measures primarily target yacht operation rather than shipyard activity directly, they contribute to changing expectations around lower-impact construction, technical capability, and credible sustainability positioning. By contrast, the Corporate Sustainability Reporting Directive (CSRD) affects large shipyards more directly by requiring structured sustainability disclosure, stronger data systems, and more formal governance arrangements (European Parliament and Council of the European Union, 2022).

Social pressure further illustrates the sector's visibility. High-profile protest actions, such as the vandalism in Barcelona of a Dutch-built superyacht in 2023, show how yachts can become visible targets in broader sustainability-related critique (McCluskey, 2023). Importantly, such actions should not be interpreted only as criticism of individual vessels. They demonstrate how the symbolic visibility of superyachts can make the wider sector vulnerable to public scrutiny.

This visibility also extends to yard-related activities. In Rotterdam, a proposal to temporarily dismantle a historic bridge so a superyacht could pass drew public controversy. It showed that local communities may mobilise against activities associated with the sector (Kamerman & Verseput, 2022). The available evidence does not show that protest forced a change in logistics. Even so, the case demonstrates that superyacht production and transport decisions can become publicly contested. Such events show that social pressure is not limited to yacht operation, but may also affect the perceived acceptability of yard-level activities and decisions.

Scrutiny also extends to governance and supply chain practices. Public reporting on illegally sourced teak in superyacht supply chains has shown how materials can become focal points for critique, even when formal responsibility is vague or legally contested (Chin-A-Fo & Kuijpers, 2023). In such cases, governance risks manifest primarily as perception and legitimacy risks.

At the same time, the sector's internal sustainability focus is shifting. Attention has traditionally centred on the yacht's operational phase, but this is expected to change. As propulsion systems move towards low- or zero-carbon alternatives, construction activities account for a larger share of total life-cycle impact (Hirtenlehner, 2025). The build phase also concentrates labour conditions, subcontracting practices, safety performance, and governance responsibilities within the shipyard.

The pressures described above span environmental performance, social accountability, and governance quality. Together, they map onto the three dimensions of the Environmental, Social and Governance (ESG) framework. ESG is the dominant structure through which regulators, investors, and other stakeholders assess organisational sustainability performance (Crace & Gehman, 2023; Lombardi Netto et al., 2026). Within the European context, ESG reporting is operationalised through the European Sustainability Reporting Standard (ESRS) (European Commission, 2023). These standards define how organisations must structure and disclose sustainability information. The shift from voluntary to mandatory reporting strengthens the role of sustainability performance in strategic planning, investment behaviour, and internal governance structures. The regulatory developments driving this shift, particularly the CSRD and ETS, apply to large shipyards well beyond the superyacht sector. The challenge of translating ESG expectations into investment decisions is therefore not unique to this case. This thesis explores how ESG performance can be valued and compared at yard-level, using a superyacht shipyard as the empirical setting.

1.1. Yard-level sustainability investment

Superyacht shipyards must translate increasing regulatory requirements and societal scrutiny into concrete organisational responses, while operating under financial and organisational constraints (Ettema, 2021; Hirtenlehner, 2025). In the yachting sector, this takes the form of a growing set of sustainability-related investment proposals aimed at improving environmental performance, working conditions, and governance quality (Hojnik et al., 2020). These investments are typically motivated by regulatory anticipation, risk management, and long-term strategic positioning (Hirtenlehner, 2025; Hojnik et al., 2020). The same drivers shape internal investment proposals at the case company.

In this research, sustainability investments are defined as direct organisational or capital expenditure intended to improve ESG performance beyond short-term compliance. In a shipyard context, these range from capital-intensive measures such as insulating build halls to organisational interventions including workforce training and governance or compliance structures (Vakili et al., 2023). These investments differ in cost, scale, and in the type, timing, and observability of their impacts. While environmental effects are partly quantifiable, social and governance effects are typically indirect and inferred through proxy indicators (Hojnik et al., 2020; Young-Ferris & Roberts, 2023). Nevertheless, these integrated benefits are frequently undermined by an industry culture that treats sustainability investments as discrete operational tasks rather than holistic strategic programmes (Vakili et al., 2023).

The core decision problem addressed is the prioritisation of sustainability investments at superyacht shipyard-level. Yard management must decide which sustainability-related investments to pursue under limited financial and organisational resources and competing strategic objectives. In practice, investment proposals are consolidated at group level and prioritised in capital allocation rounds by senior management, typically the Chief Executive Officer (CEO) and Chief Financial Officer (CFO). Final selection is based mainly on intuitive judgement supported by cost estimates and simplified economic indicators, rather than systematic comparison of sustainability effects across alternatives (Kon. de Vries Scheepsbouw Personal Communication, 26-11-2025). Capital investment prioritisation is inherently a strategic management decision, involving multi-year resource commitments, competing objectives, and consequences that are difficult to reverse (Hallegatte et al., 2012; Montibeller & Franco, 2011). Within De Vries, this is reinforced by the group-level structure: proposals from subsidiaries including Makkum and Aalsmeer compete for limited capital within a common allocation process.

The scope of the decision is limited to investments over which shipyards have direct organisational control. Investments related to yacht-specific design choices and project-level supplier selection are excluded, because these decisions are largely driven by owners and designers. Although shipyards influence how such choices are engineered and manufactured in practice, this influence is indirect and project-specific, making these decisions less suitable for systematic comparison within a yard-level investment framework. By contrast, governance investments that shape supplier behaviour at organisational level, such as procurement standards, due diligence systems, and sustainability data requirements, are explicitly included. This scope therefore provides a focused basis for yard-level prioritisation, while leaving open future extension to project-level design and supply chain decisions.

1.2. Decision complexity and uncertainty

Sustainability investment decisions in pure custom superyacht shipyards violate several assumptions underlying conventional investment appraisal. While conventional shipbuilding already operates on low-volume, Engineer-To-Order (ETO) production models characterised by long cycle times and complex supply chains (Fitriadi & Mohamad Ayob, 2023; Neves et al., 2025), the superyacht sector amplifies these constraints to an extreme. Production is organised around completely unique, long-duration projects with exceptionally fragmented supply chains and highly variable processes (Cozijnsen, 2019; Ettema, 2021). As a result, sustainability impacts differ strongly between investments and cannot be meaningfully assessed using standardised, project-independent metrics.

Traditional appraisal methods such as Life Cycle Assessment (LCA) and cost models (Favi et al., 2018), Social Return On Investment (SROI) (Serrano-Cinca & Gutiérrez-Nieto, 2013), and green yard frameworks (Janson, 2016) provide valuable insights into specific impact categories. However, they rely on detailed inventories, extensive data collection across suppliers, processes, and organisational units and detailed causal modelling. In shipbuilding, such data collection is described as extremely time-consuming by Favi et al. (2018) and Heimo et al. (2024). Applying these methods to multiple early-stage investment options would require substantial time and expert effort, making them impractical for comparative prioritisation.

Beyond these methodological limitations, sustainability investment decisions are characterised by multiple forms of uncertainty. First, uncertainty exists in the causal relationships linking investments to outcomes. While the direction of impact is often understood, its magnitude is uncertain and context dependent (Marchau et al., 2019; Stanton & Roelich, 2021). Second, uncertainty exists in investment costs and organisational effort. Estimated capital expenditure and implementation timelines may deviate substantially (Flyvbjerg, 2021), particularly where organisational change and learning effects are involved (Marchau et al., 2019). Third, uncertainty exists in the future context in which sustainability investments generate value. Unlike uncertainties in impact magnitude or cost, this form of uncertainty affects the conditions under which investments are evaluated. Changes in regulation, client expectations, social acceptance of superyachts, and market conditions may alter not only how well an investment performs, but also how relevant its effects are in decision-making (Hirtenlehner, 2025; Smit, 2025). Some organisations distinguish between uncertainties treated as strategic commitments, such as internal emissions targets, and those that remain genuinely open, such as reputational pressure or customer willingness to pay for sustainability (Hirtenlehner, 2025; Smit, 2025). This distinction is also present within the case company.

Data availability further constrains early-stage evaluation. Relevant ESG data exist in fragmented form across systems and formats that make integration extremely time-consuming (Favi et al., 2018; Heimo et al., 2024). Raw ESG data are also frequently incomplete, inconsistent, and difficult to combine into a usable form for direct comparison between options (Young-Ferris & Roberts, 2023). Early-stage evaluation must therefore rely on simplified representations, proxy indicators, and structured qualitative judgement as necessary components of decision-making under conditions of incompleteness (Hojnik et al., 2020; Marchau et al., 2019).

1.3. Decision support requirements

The requirements below follow directly from the decision context outlined above. The heterogeneity of sustainability investments, the strategic and group-level nature of prioritisation, and the lack of uniform metrics require decision support that enables structured comparison without removing managerial discretion. At the same time, uncertainty, incomplete data, and limited causal knowledge require transparency and explicit treatment of assumptions.

The framework must:

- reduce reliance on intuitive judgement without eliminating managerial discretion;
- support comparison of heterogeneous sustainability investments at yard-level;
- accommodate uncertainty in investment costs, investment impacts, and future context;
- tolerate incomplete information and imprecise causal relationships;
- make assumptions explicit and adjustable;
- remain feasible for repeated use in a superyacht shipyard context;
- be scalable to accommodate additional investments and indicators over time;
- reduce the risk that investments with easily quantifiable business cases are systematically favoured over those with less quantifiable ESG effects;
- make visible how uncertainty and assumption choices affect investment comparison.

These requirements define the problem constraints that guide framework design in subsequent chapters.

1.4. Research Questions

Based on the problem described and the derived requirements, the main research question of this thesis is:

> *How does the inclusion of ESG performance expectations, uncertainty about future operating conditions, and organisational biases in capital allocation influence sustainability investment priorities in a superyacht shipyard?*

This research question is decomposed into a set of interrelated sub-questions that structure the research by addressing the elements required to construct, apply, and evaluate a decision-making framework.

The following sub-questions are addressed:

- Which ESG indicators are relevant for yard-level sustainability investment evaluation, and what gaps remain relative to existing assessment approaches?
- How can investment consequences be represented despite uncertain causal relationships and limited data availability?
- How can heterogeneous ESG and financial considerations be compared in a transparent and accountable manner?
- How does uncertainty about future operating conditions affect which investments are prioritised?
- What does each framework component contribute to making the influence of ESG expectations, uncertainty, and organisational bias on investment priorities explicit and contestable?
- What insights does application of the framework generate for sustainability investment priorities in a superyacht shipyard?

1.5. Thesis structure

Following this introduction, Chapter 2 reviews the relevant literature and derives the conceptual foundations for yard-level sustainability investment decision support. Chapter 3 presents the research design and explains how the framework is constructed, operationalised, and evaluated. Chapter 4 applies the framework to the case company, Koninklijke de Vries Scheepsbouw in Aalsmeer, and reports the implementation process, results, and case-specific analysis. Chapter 5 discusses the findings in relation to theory and practice, and Chapter 6 concludes the thesis by answering the research questions and outlining implications and recommendations.

Detailed supporting material is reported in Appendix A. This appendix contains case-specific technical material, including full mapping tables, elicitation inputs, scenario matrices, and detailed calculation outputs that support, but are not required for, the main-text argument.

2

Literature Review

This chapter establishes which ESG dimensions are relevant for yard-level sustainability investment evaluation and identifies where existing assessment approaches fall short. It reviews literature on sustainability and decision support in superyacht shipbuilding, an extreme form of ETO production where customisation, project-based working, and limited data availability shape what sustainability assessment can deliver. The chapter also identifies the analytical tasks that yard-level decision support must address, providing the foundation for framework operationalisation in Chapter 3.

2.1. Review approach and use

The literature review is structured along two complementary axes. The first axis addresses the production context and comprises literature on superyacht shipbuilding and sustainability at shipyard-level. This literature is limited in size and is predominantly descriptive, with a strong focus on vessels and environmental impacts rather than on yard-level investment choices.

The second axis addresses decision logic. It includes literature on ESG theory, sustainability indicators, organisational sustainability investment, formal decision support methods, and behavioural decision-making. These academic fields are substantially broader and were not reviewed exhaustively. Instead, they were used selectively to identify the limitations of reporting-oriented ESG systems, the organisational distortions affecting long-term investment decisions, and the classes of decision support methods that remain suitable under uncertainty.

2.1.1. Literature search strategy

The literature search was organised by theme and conducted iteratively as results were screened and search strings refined. All searches were performed using Google Scholar between 11 November and 31 December 2025. Google Scholar was selected because it captures engineering, sustainability, and decision analysis literature within a single search environment. Scopus was used as a cross-check database and added no further relevant publications.

Where appropriate, time filters were applied, most often restricting results to publications since 2021 in order to reflect recent developments in ESG regulation and sustainability research. Search results were screened first by title, then by abstract, and finally by full text where relevance remained uncertain. For most searches, screening was limited to the first five result pages because relevance declined sharply beyond that point. More specific searches with low result counts were screened further.

Beyond keyword-based search, key sources were identified through backward citation tracing. Review papers and foundational works in each thematic area provided entry points. Their reference lists were then examined to locate further relevant publications. This approach was particularly useful for the decision logic axis, where the broader interdisciplinary literature is too large to cover through keyword search alone, and helped ensure that established methodological work was incorporated alongside more recent publications.

Directly relevant literature on superyacht shipbuilding and yard-level sustainability investment was limited and fragmented, while adjacent literatures on ESG, sustainability indicators, and decision support were much larger. The second axis was therefore treated selectively rather than exhaustively.

2.1.2. Inclusion and exclusion criteria

The review is limited to custom superyacht shipyards. These are yards primarily engaged in constructing completely customised superyachts above approximately 40 metres, characterised by long build times and project-based production. The practical boundary was informed by industry expert input (Kon. de Vries Scheepsbouw Personal Communication, 22-01-2026). Studies focused on smaller semi-custom segments were treated as out of scope unless they offered uniquely transferable decision support logic.

The included decision context was required to be organisational and internal. The unit of analysis had to be a shipyard or shipyard group, and the study had to inform internal prioritisation, planning, or allocation decisions. Studies focused on external capital allocation by investors were excluded, including work framed around ESG investing in financial markets, portfolio construction, stock selection, or asset pricing.

The intervention scope was intentionally broad. Any sustainability-related intervention plausibly applicable to ships or superyachts was eligible, provided the study offered information transferable to yard-level decision-making or investment prioritisation. No restriction was imposed across environmental, social, or governance dimensions where relevance to the decision context could be established.

Studies focused on recycling, yacht design in isolation, shipping, or shipowner operational contexts were excluded. Studies that remained at the level of technical optimisation without implications for yard-level production, investment, or managerial prioritisation were also excluded. Non-peer-reviewed consultancy reports and company reports were excluded to maintain academic consistency. An exception was made for selected thesis work where it provided the only accessible literature on custom superyacht shipyards or relevant methodological insight for the sector.

A narrow exception was made for one highly cited study on ESG integration in investment analysis. Although it originates in a financial-market context, Young-Ferris and Roberts (2023) was retained because it provides a useful diagnostic account of why ESG information is difficult to incorporate into valuation-based decision frameworks. It is therefore used here as methodological evidence on the limits of ESG quantification rather than as sector-specific investment guidance.

2.1.3. Review limitations

Maritime is a niche field whose journals can be underrepresented in Scopus or Web of Science due to citation thresholds. Google Scholar's broader inclusion criteria address this, but also surface less rigorously reviewed work, requiring careful quality screening during the review. The exclusion of consultancy reports and internal industry documentation limits exposure to practitioner-developed decision tools that may be used informally in shipyards. Selected thesis work was included where peer-reviewed alternatives were unavailable, which improves sector relevance but lowers the formal standard of part of the review base. Finally, the emphasis on recent literature strengthens relevance to current ESG developments, but may favour regulatory discourse over older operational sustainability research. News articles used in Chapter 1 are treated as contextual evidence of public scrutiny and sector visibility, not as analytical evidence for framework selection.

2.2. Sustainability impacts in shipyards and shipbuilding

Sustainability impacts in shipbuilding have been studied extensively, but the literature is uneven. Environmental impacts of shipyard activity are comparatively well documented, while social and governance impacts remain less developed in research aimed at internal decision-making.

LCA studies provide detailed analyses of welding, cutting, blasting, painting, and energy use, showing that these processes contribute significantly to emissions, hazardous waste, and human toxicity during steel ship construction (Narci & Can Dogan, 2025). Broader green shipyard concepts and hybrid energy system frameworks have also been proposed (Vakili et al., 2023). However, these studies focus primarily on environmental indicators and costs, provide limited treatment of social and governance dimensions, and do not support systematic comparison of heterogeneous investments at yard-level.

A further limitation is the fragmented nature of sustainability information across shipbuilding supply chains. Heimo et al. (2024) show that incompatible naming systems, uneven supplier expertise, and limited automated data exchange hinder the repeated application of LCA and Life Cycle Costing (LCC). Even where relevant assessment methods exist in theory, these information constraints make their routine use in comparative decision-making difficult in practice.

The literature shows shipyard sustainability impacts are known and partly quantified. The dominant assessment approaches remain poorly aligned with repeated internal investment prioritisation under data scarcity. The next section examines how the superyacht production environment further shapes what can be evaluated and how.

2.3. Superyacht sustainability and the production environment

Lynch et al. (2019) frame the superyacht industry as an extreme form of luxury consumption with a disproportionate ecological footprint. This positions superyachts not merely as technical products, but as socially and environmentally contested symbols. That framing is important because it increases the strategic relevance of legitimacy, reputation, and public scrutiny for industry actors.

As shown by Hirtenlehner (2025), growing public scrutiny, regulatory attention, and stakeholder pressure have already translated into strategic responses within the superyacht ecosystem. Environmental accountability is no longer treated solely as a compliance issue, but increasingly as a condition for reputational legitimacy, long-term licence to operate, and access to skilled labour.

Despite this heightened attention, sustainability research in the superyacht domain remains predominantly vessel-focused. Dominant themes include alternative fuels, propulsion technologies, operational emissions, and marina performance (Diesveld, 2019; Hirtenlehner, 2025; Letschert, 2020). Shipyards are largely treated as background infrastructure rather than as decision-making organisations. Governance structures, workforce conditions, energy use, and material flows at yard-level receive only limited and fragmented attention (Hirtenlehner, 2025).

Studies associated with the Yacht Environmental Transparency Index (YETI) provide partial insight into production-related impacts but remain vessel-centred. Letschert (2020) confines the scope to operational emissions, while Cozijnsen (2019) extends analysis to production through a fast-track LCA but explicitly excludes yard processes. Energy use, logistics, and material efficiency at yard-level are therefore systematically underrepresented. More importantly, these approaches identify impacts, but do not explain how discrete investments should be prioritised.

The limited attention to yard-level analysis reflects structural characteristics of superyacht shipbuilding itself. As an extreme form of ETO production (Fitriadi & Mohamad Ayob, 2023; Neves et al., 2025), the sector combines high customisation, long lead times, fragmented supply chains, and craftsmanship-intensive processes. These characteristics complicate standardisation and data collection, reinforcing the tendency to assess sustainability at vessel level rather than at the level of the production processes.

Recognising sustainability impacts is one task. Evaluating them for internal investment comparison is another. In this context, the literature offers competing logics with different assumptions about information and choice.

2.4. Competing evaluation logics

The literature on sustainability assessment and the literature on investment decision-making rest on different assumptions about what information matters and how it should support organisational choice. These reflect distinct evaluation logics, as shown in Figure 2.1.

A first class of approaches conceptualises sustainability primarily through valuation and prediction. Within this logic, sustainability investments are assessed by estimating future outcomes and translating them into a common monetary metric, often through extensions of Net Present Value (NPV) analysis. This class also includes real options approaches, which attempt to capture managerial flexibility under uncertainty through option valuation.

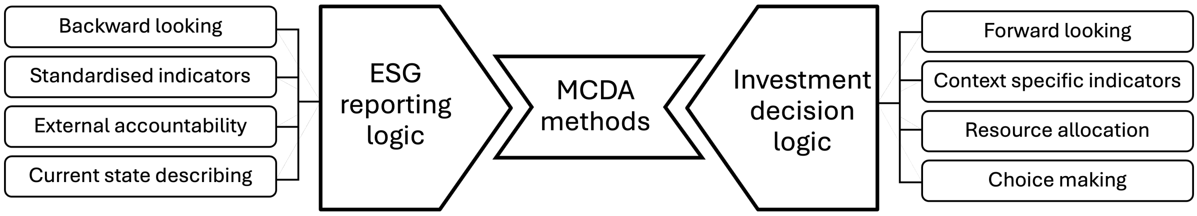


Figure 2.1: Structural discontinuity between ESG reporting logic and investment decision logic. ESG reporting is backward-looking, relies on standardised indicators, serves external accountability, and describes the current state. Investment decision logic is forward-looking, relies on context-specific indicators, supports resource allocation, and informs choice making. Multi Criteria Decision Analysis (MCDA) methods provide a bridge between the two.

Such approaches optimise for comparability and formal completeness. They assume, however, that causal relationships can be modelled and that uncertainty can be represented through defensible probability distributions (Smit, 2025). For yard-level sustainability investments, these assumptions are often not credible. Value is generated through indirect, long-term, and context-dependent mechanisms such as regulatory resilience, risk mitigation, workforce retention, and reputational effects. Under such conditions, valuation-oriented approaches become structurally inappropriate for repeated internal prioritisation.

A second class of approaches conceptualises sustainability through indicator systems structured along environmental, social, and governance dimensions. ESG and Triple Bottom Line (TBL) frameworks are designed to support transparency, accountability, and comparability across organisations (Crace & Gehman, 2023). Within the European regulatory context, this logic is now formalised through the CSRD (European Parliament and Council of the European Union, 2022), which mandates structured sustainability disclosure for large industrial firms. This requirement is operationalised through the ESRS (European Commission, 2023), providing a structured basis for identifying material sustainability topics through the principle of double materiality. While these systems are effective for describing organisational states, they do not specify how the marginal effects of discrete investments should be compared under scarcity. Reporting logic, such as that found in ESRS 2, is therefore not synonymous with choice logic.

A third class of approaches addresses trade-offs and uncertainty more directly through multi-criteria and scenario-oriented decision support. These approaches, often grouped under MCDA, preserve separation between criteria. They compare alternatives across heterogeneous objectives and plausible future conditions without requiring probabilistic ranking (Marchau et al., 2019). Methods within this class, such as additive Multi Attribute Value Theory (MAVT) or hierarchical elicitation, are designed to make value trade-offs explicit (Dyer, 2016; Montibeller & Franco, 2011). In principle, they bridge reporting logic and investment logic better than monetisation. In practice, however, many of them assume extensive data availability and long analytical cycles. In executive settings characterised by time pressure and limited organisational bandwidth, this can reduce their feasibility for routine application (Stanton & Roelich, 2021).

Some sustainability investment literature argues that impacts must be monetised or strongly aggregated to influence executive capital allocation. In the yard-level context examined here, however, the assumptions required for such valuation cannot be justified. Social, governance, and reputational effects are mediated through organisational behaviour and uncertain future conditions in ways that resist causal quantification and probabilistic modelling. Under these conditions, monetisation may obscure uncertainty rather than reduce it (Young-Ferris & Roberts, 2023).

Taken together, the literature shows that no single evaluation logic simultaneously satisfies the demands for external legitimacy, methodological completeness, organisational feasibility, and robustness under uncertainty. That tension is central to the decision problem addressed in this thesis. It also implies that framework selection cannot proceed as the adoption of a single complete method. A defensible framework must instead combine compatible components that together structure value judgements, map heterogeneous consequences onto a comparable scale, represent future uncertainty without spurious probability assignments, and evaluate alternatives across plausible futures.

2.5. Deep uncertainty and proxy-based decision support

The competing logics of the previous section share a common difficulty. None copes well with uncertainty that cannot be reduced to probabilities, which is the problem this section addresses.

Across the reviewed literature, incomplete data and uncertainty appear as defining characteristics of sustainability investment decisions rather than as residual modelling problems (Heimo et al., 2024; Marchau et al., 2019). Decision-makers therefore cannot rely on stable metrics, complete causal models, or fully quantified investment effects when comparing sustainability investments.

Within Decision Making under Deep Uncertainty (DMDU), decision support must be able to work with qualitative judgement, indirect observation, and proxy-based information. Proxy indicators provide consistent and interpretable signals when direct measurement of an effect is not feasible. Their use should therefore not be understood only as a methodological compromise. In this context, proxies are necessary because several relevant sustainability effects are indirect, delayed, or difficult to observe directly (Hojnik et al., 2020).

The literature further distinguishes between parameter uncertainty and deep uncertainty. Parameter uncertainty refers to situations where outcomes are uncertain, but the relevant model structure is broadly accepted. Deep uncertainty refers to situations where decision-makers cannot agree on the appropriate model, probability distributions, or dominant value trade-offs (Marchau et al., 2019). Walker et al. (2013) classify uncertainty along four levels, ranging from well-defined parameter uncertainty (level one) to situations in which decision-makers cannot agree on the underlying model, the relevant probability distributions, or the dominant value trade-offs (level four). The classification used in this thesis corresponds closest to this fourth level, particularly because future regulation, market expectations, social pressure, and client behaviour cannot be represented through a single agreed model or defensible probability distribution. Yard-level sustainability investment decisions are closer to this level of uncertainty than to parameter uncertainty, because the future conditions under which investments create value may change across plausible, unpredictable contexts (Hallegatte et al., 2012).

This matters because expected-value optimisation assumes that uncertain outcomes can be described through stable models and probabilities. Under deep uncertainty, that assumption is not credible. Decision support should therefore help decision-makers explore how investments perform across different plausible futures, rather than identify a single optimal investment based on one forecast (Marchau et al., 2019; Stanton & Roelich, 2021).

The literature thus suggests that uncertainty in this context is not transitional, but persistent. Better data may reduce some unknowns, particularly around costs and measurable environmental effects, but it does not remove uncertainty about future context, stakeholder expectations, or the strategic relevance of different sustainability outcomes. The framework must therefore remain usable under incomplete knowledge. This rules out method classes that rely on defensible probability assignments, including probabilistic forecasting, expected-value optimisation, and expected-utility approaches. Methods compatible with deep uncertainty must instead support reasoning across plausible futures without aggregating them into probability-weighted summaries.

2.6. Yard-level decision logic

Method suitability is necessary but not sufficient. The preceding sections show that sustainability impacts are increasingly recognised, yet remain difficult to incorporate into repeated internal investment comparison. The problem is therefore not only technical, but organisational. Even where sustainability tools are analytically sound, they may still fail to influence capital allocation in practice. The focus now shifts from analytical fit to organisational reality.

2.6.1. Capital allocation and organisational behaviour

In large industrial firms, capital allocation is not a purely analytical exercise. Decisions emerge from established routines, internal negotiations, bounded rationality, and shifting strategic priorities. Since managers must navigate conflicting goals with limited time and cognitive capacity, optimisation is only one influence among several (Stanton & Roelich, 2021). Under such conditions, organisations often satisfice rather than optimise, selecting options that are good enough relative to critical requirements

(Marchau et al., 2019).

Formal decision tools therefore function less as neutral optimisation engines than as organisational artefacts that structure deliberation and frame strategic debate (Marchau et al., 2019; Walker et al., 2013; Young-Ferris & Roberts, 2023). Their effectiveness depends on alignment with local incentives, production priorities, and prevailing strategic narratives. Empirical work in infrastructure and maritime settings shows that technically robust tools are often rejected or distorted when they conflict with operational logic (Stanton & Roelich, 2021; Vakili et al., 2023).

These findings imply that decision support in shipyards must accommodate negotiation, ambiguity, and partial information rather than suppress them through premature aggregation (Montibeller & Franco, 2011; Stanton & Roelich, 2021). These observations describe general organisational dynamics. The next subsection turns to specific behavioural mechanisms that disadvantage sustainability investments in internal capital allocation.

2.6.2. Behavioural aspects of investment decisions

The mechanisms discussed in this subsection are used to explain why sustainability-related investments may be systematically undervalued in internal capital allocation, rather than as direct empirical claims about the case company.

First, executive optimism and overconfidence shape investment choices under uncertainty. Azouzi and Anis (2012) show that optimistic executives tend to favour investments associated with visible growth narratives and short-term performance, while undervaluing initiatives whose contribution is defensive, enabling, or long-term. Sustainability investments often fall into the latter category, because their benefits frequently take the form of avoided losses, regulatory resilience, or workforce retention rather than immediate revenue generation (Young-Ferris & Roberts, 2023).

Second, strategic misrepresentation can distort project comparison. Flyvbjerg (2006) shows how projects may be framed selectively to secure approval. Sustainability investments are disadvantaged here because their benefits are often intangible and resist translation into conventional financial metrics (Young-Ferris & Roberts, 2023). By contrast, traditional capital projects more readily support optimistic narratives, contributing to what Flyvbjerg (2021) describes as a selection dynamic of "survival of the unfittest".

Third, escalation of commitment and sunk costs can crowd out deferred sustainability initiatives once resources, attention, and organisational reputation become tied to ongoing projects. This further weakens the position of sustainability-related improvements in internal competition for capital (Flyvbjerg, 2021; Smit, 2025).

The literature therefore suggests that sustainability investments are not only difficult to assess, but also structurally disadvantaged in the organisational decision environments where they must compete. However, the strength of this disadvantage may be moderated by ownership structure. In maritime contexts, founding-family ownership can foster "stewardship characteristics", where organisational objectives extend beyond maximising shareholder wealth to include long-term survival and environmental concerns (Andrikopoulos, 2026). Family-owned firms with this stewardship profile are not immune to the underlying biases, but the pressures shaping how sustainability investments compete for internal capital may differ.

These behavioural mechanisms also impose requirements on framework design. To counter optimism bias, strategic misrepresentation, and escalation effects, the framework must make value judgements explicit, expose opportunity costs across plausible futures, and support challenge and revision of elicited inputs.

2.7. Implications for framework selection

The reviewed literature yields three conclusions that directly constrain framework selection. First, yard-level sustainability impacts are real and partly understood, but current assessment approaches are poorly aligned with repeated internal investment comparison. Second, the superyacht literature remains predominantly vessel-centred, meaning that the shipyard is rarely analysed in literature as an organisation responsible for investment decisions. Third, internal capital allocation under uncertainty is shaped by bounded rationality, behavioural bias, and organisational feasibility constraints as much as by analytical sophistication.

These conclusions imply that framework selection cannot be reduced to choosing a single existing method. Existing method families address parts of the decision problem well, but none addresses all of its analytical tasks simultaneously. Framework selection is therefore treated here as the combination of compatible method components rather than the adoption of one complete decision model.

2.7.1. Framework architecture and selection criteria

Translating this requirement into framework design begins with identifying the analytical tasks the framework must perform. Four such tasks organise the remainder of this section.

The first task, **value structuring**, elicits the relative importance of the criteria against which investments are compared. The second, **value mapping and aggregation**, converts heterogeneous indicator outcomes into a common evaluative scale and combines them into one comparable score per investment. The third, **future-dependent performance**, represents how investment outcomes change across plausible external futures without assigning probabilities. The fourth, **evaluation**, compares investments across futures to identify alternatives that remain defensible under uncertainty. These four tasks form a sequential pipeline, depicted in Figure 2.2, with the output of each task feeding the next.

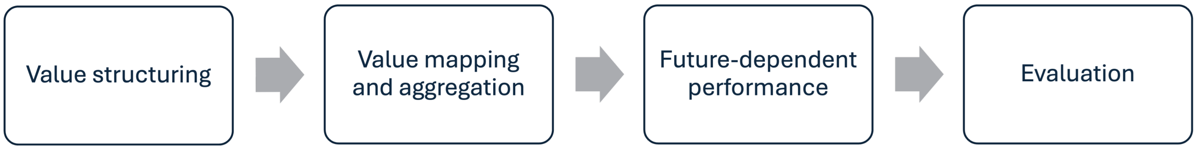


Figure 2.2: The four analytical tasks of the framework. Value structuring elicits criterion importance; value mapping and aggregation converts heterogeneous indicator outcomes into a common scale and combines them into one comparable score per investment; future-dependent performance represents how outcomes change across plausible external futures; evaluation compares investments across futures to identify alternatives that remain defensible under uncertainty.

A separate construction layer underpins this pipeline. It has two distinct elements. The first is the indicator basis, the set of proxy variables that defines what counts as a consequence. The indicator basis is established once for the framework and provides the criteria that value structuring weights and value mapping evaluates. Consequence construction is performed per investment-indicator-scenario combination, either through direct assessment or through propagation in a System Dynamics Representation (SDR). The indicator basis is further discussed in Section 2.7.2 and operationalised in Chapter 3, where the SDR is also discussed in detail.

For each task, method selection is assessed against five criteria. These criteria operationalise the framework requirements set out in Section 1.3 at the level of individual method choice, turning the problem constraints into tests that each candidate method must satisfy. Under the first criterion, *compatibility with deep uncertainty*, methods that require defensible probability assignments are excluded (Section 2.5). Under the second, *transparency of value judgements*, trade-offs and preference assumptions must be explicit and revisable rather than embedded in technical parameters (Section 2.6.2). Under the third, *organisational feasibility*, methods must remain usable under executive time pressure and limited analytical bandwidth (Section 2.6.1). Under the fourth, *behavioural-bias counterweight*, methods should support challenge of elicited inputs and expose opportunity costs that would otherwise be obscured by optimism bias and strategic misrepresentation (Section 2.6.2). Under the fifth, *robustness under variation*, chosen components must allow systematic robustness testing of their elicited inputs, with the methodological details handled in Chapter 3.

Two of the Section 1.3 requirements are properties of the framework architecture rather than tests for individual method choice. Comparing heterogeneous investments at yard-level is the function of the value mapping and aggregation task introduced above, while scalability to additional investments and indicators rests on the indicator basis and its round-specific reduction in Chapter 3. The full correspondence between the requirements and their treatment across the framework is verified in Chapter 5.

These criteria apply uniformly across the four tasks. They are used in the next subsection to justify each chosen method and the rejection of competing alternatives.

2.7.2. Defining the indicator basis for framework selection

A yard-level sustainability investment framework requires a structured indicator basis to allow heterogeneous investments to be compared in a transparent and repeatable way. In this thesis, that basis is derived from two sources, namely the ESRS 2 aligned double materiality structure as the regulatory starting point and academic literature as the analytical complement.

Double materiality offers a legitimate starting structure because it reflects sustainability topics already recognised as materially relevant at organisational level and embedded in emerging reporting practice. However, reporting-oriented structures are not sufficient on their own for investment comparison. Their primary function is to identify and disclose material topics, not to compare the marginal effects of discrete yard-level investments under resource constraints.

Academic literature is therefore used to refine and extend the reporting structure for decision support. Prior work supports distinguishing water use more explicitly in shipyard and yachting settings (Hojnik et al., 2020), treating health, safety, and workforce development as an operational domain rather than as isolated reporting items (Hojnik et al., 2020), separating economic performance (Fitriadi & Mohamad Ayob, 2023; Neves et al., 2025; Tantan & Akdağ, 2025) from investment feasibility (Koray, 2023; Vakili et al., 2023), and recognising technical maturity and implementation risk as distinct decision-relevant considerations (Koilo, 2021; Vakili et al., 2023). The resulting indicator basis is not yet an investment-ready set of decision criteria, but a defensible starting structure for framework design. Its operationalisation and reduction for decision use are explained in Chapter 3.

2.7.3. Selection of framework components

Once the indicator basis has been defined, the selection of method components for each of the four analytical tasks introduced in Section 2.7.1 can proceed. For each task, the candidate method families are compared against the selection criteria established there.

Value structuring

The first task requires eliciting the relative importance of evaluation criteria in a way that supports hierarchical structure, makes judgements traceable, and remains usable in workshop settings. The MCDA literature offers several method families for this elicitation, differing in cognitive burden, treatment of interdependencies, and explicitness with which judgement is captured.

The Analytic Hierarchy Process (AHP) elicits relative weights through pairwise comparisons (Saaty, 2008). It supports hierarchical criteria structures, makes pairwise judgements traceable, and includes a Consistency Ratio that flags logically inconsistent comparisons for review (Ishizaka, 2014; Saaty, 1977). Pairwise decomposition has been shown to perform more reliably than direct weighting approaches by breaking complex narratives into simpler relative comparisons (Ishizaka, 2014).

The Best-Worst Method reduces elicitation burden by requiring only comparisons against the most and least important criteria in the set (Velasquez & Hester, 2013). However, this reduction entails losing information from the complete set of paired comparisons, which weakens the multicriteria analysis on ESG models (Lombardi Netto et al., 2026). In a context where decision-makers must navigate trade-offs across highly heterogeneous sustainability and financial dimensions, the complete cross-comparison matrix produced by AHP generates a richer value structure than comparison against extremes alone.

The Analytic Network Process extends AHP by modelling interdependencies between criteria explicitly (Taherdoost & Madanchian, 2023). Quantifying these interlinkages between heterogeneous ESG factors is problematic, however, because the underlying causal relationships are often unclear or deeply uncertain. Forcing such interactions into the model would require modelling assumptions that the conditions in Section 2.5 cannot justify. The method is also cognitively demanding in workshop settings because the number of network comparisons grows rapidly with criterion count (Taherdoost & Madanchian, 2023).

Fuzzy extensions of AHP embed linguistic uncertainty directly into weight representation. They increase modelling complexity substantially and reduce interpretability for marginal accuracy gains (Ishizaka, 2014). Direct weighting is simpler still but offers neither hierarchical structure nor consistency control.

Against the selection criteria, standard AHP performs best. Because it requires no probability assignments, it remains compatible with deep uncertainty. Its pairwise comparisons make each judgement explicit and revisable, satisfying the transparency criterion, and it stays organisationally feasible at the criterion counts used in this framework. The Consistency Ratio acts as a behavioural-bias counterweight, surfacing logical inconsistency for participant review rather than concealing it within an aggregate weight. Finally, the reciprocal structure of pairwise judgements lends itself to the systematic perturbation used in robustness testing, addressed in Chapter 3.

Value mapping and aggregation

The second task requires translating heterogeneous indicator outcomes into a common evaluative scale and combining them into one comparable score per investment. These two operations are tightly coupled. Aggregation requires inputs on a shared scale, and the choice of mapping logic constrains which aggregation rules remain admissible. They are therefore treated as one task with two stages.

Indicator outcomes differ in unit, direction, and meaning, and cannot be combined directly. Three approaches dominate the literature. Performance functions map indicator outcomes onto a bounded evaluative scale, typically $[0, 1]$, while making thresholds, diminishing returns, and preferred directions explicit (Boix-Cots et al., 2022; Dyer, 2016; Keeney, 1988). Direct normalisation rescales raw values using methods such as min-max normalisation, which embeds linear value assumptions implicitly. Equal increments in raw performance are treated as equally valuable, even where threshold effects clearly apply (Gan et al., 2017). Monetisation converts heterogeneous outcomes into financial units, which is poorly suited to ESG effects that are indirect, delayed, or difficult to value credibly, and may force heterogeneous effects into a financial logic that favours what is easiest to price (Flyvbjerg & Bester, 2021; Young-Ferris & Roberts, 2023).

For aggregation, additive MAVT combines criterion-level scores using explicit weights, preserving the interpretability of the resulting trade-offs (Dyer, 2016; Gan et al., 2017; Velasquez & Hester, 2013). The additive form requires preferential independence between criteria and permits full compensation between them (Dyer, 2016). Outranking methods such as Preference Ranking Organization Method for Enrichment Evaluations (PROMETHEE) and Élimination et Choix Traduisant la Réalité (ELECTRE) rely on pairwise preference relations and method-specific parameters such as preference thresholds, which embed key value judgements within technical settings rather than making them explicit and tend to produce unstable rankings when performance varies across scenarios (Behzadian et al., 2010). Dominance-based approaches such as Pareto optimisation produce large efficient sets with limited decision relevance when alternatives perform unevenly across multiple criteria, as is characteristic of ESG comparisons (Behzadian et al., 2010). Related multi-objective optimisation approaches reintroduce hidden value judgements through implicit priority structures or remain unstable under uncertainty (Velasquez & Hester, 2013).

Against the same criteria, performance functions combined with additive MAVT perform best. Needing no probability assignments, the combination is compatible with deep uncertainty. It is transparent because mapping and weighting are kept separate, with within-criterion value judgements expressed in the function shape and between-criterion trade-offs expressed in the AHP weights. Feasibility follows from eliciting performance functions through a small set of templates rather than free-form specification, which limits elicitation burden (Boix-Cots et al., 2022). As a behavioural-bias counterweight, the combination exposes where value depends on threshold effects or diminishing returns rather than concealing these assumptions in normalisation. Its bounded scale and parametric function shape again lend themselves to systematic perturbation in robustness testing, addressed in Chapter 3.

Future-dependent performance

The third task requires representing how investment outcomes change across plausible external futures in a way compatible with deep uncertainty. The DMDU literature offers several approaches, differing primarily in whether they assign probabilities to future states and in the analytical infrastructure they require.

Probabilistic forecasting represents future uncertainty through probability distributions over key parameters and supports expected-value optimisation. As established in Section 2.5, defensible probability assignment is not possible under the level-four uncertainty that characterises yard-level sustainability investment, ruling out this class of method.

Scenario analysis represents future uncertainty through a small set of internally consistent, qualitatively distinct future contexts. Each scenario describes a plausible external state without assigning likelihood, and investments are evaluated separately under each scenario (Marchau et al., 2019; Montibeller & Franco, 2011). This makes assumptions about external developments explicit and revisable rather than aggregating them into probability-weighted summaries (Hallegatte et al., 2012; Zwaginga, 2026). The combination of scenario analysis with value-based MCDA methods is an established response to deep uncertainty in strategic decision contexts, with MAVT and AHP identified as the leading methods used to evaluate alternatives across scenarios (Marttunen et al., 2017).

Robust Decision-Making and Dynamic Adaptive Policy Pathways extend scenario analysis by exploring large numbers of futures computationally or by designing strategies that adapt over time (Marchau et al., 2019). These methods preserve robustness but require substantial data, computational infrastructure, and repeated stakeholder engagement that limit their feasibility in executive settings characterised by time pressure and limited analytical bandwidth (Stanton & Roelich, 2021).

Scenario analysis is likewise the strongest option for representing future-dependent performance. It assigns no probabilities, so compatibility with deep uncertainty holds by construction. Each scenario can be described in terms decision-makers can challenge and revise, which keeps the approach transparent, and its modest elicitation burden at the two-axis, four-scenario level adopted in Chapter 3 keeps it organisationally feasible. Its behavioural-bias counterweight comes from forcing each investment to be assessed under adverse as well as favourable futures, which limits the room for optimism bias and selective framing of investment narratives. Its discrete scenario structure also allows the scenario-specific impacts to be perturbed systematically in robustness testing, addressed in Chapter 3.

Evaluation

The fourth task requires comparing aggregated investment outcomes across scenarios to identify alternatives that remain defensible under uncertainty. The literature offers several criteria for this comparison, differing in whether they require probability assignments and in how they weight different types of risk exposure.

Expected value and expected utility criteria weight scenario outcomes by their probabilities. As established in Section 2.5, defensible probability assignment is not possible in the present context, ruling out these criteria.

Maximin selects the alternative with the best worst-case outcome across scenarios. It avoids probability assignment but evaluates each alternative only through its single worst absolute outcome, regardless of how that outcome compares to alternatives in the same scenario (Marchau et al., 2019). This produces overly conservative recommendations. An alternative with a moderately weak outcome in one adverse scenario is penalised identically whether competing alternatives perform similarly or much better in that same scenario. The criterion therefore conveys exposure to bad outcomes but not exposure to opportunity loss, which is the more decision-relevant quantity when alternatives are being compared rather than accepted in isolation.

Minimax regret evaluates each alternative through its maximum opportunity loss across scenarios, where regret in a given scenario is the difference between the alternative's outcome and the best alternative's outcome in that same scenario (Hallegatte et al., 2012; Stanton & Roelich, 2021). It does not require probability assignments and penalises alternatives that perform well in narrow subsets of futures while performing poorly elsewhere. By framing performance as opportunity cost relative to the best-performing alternative in each scenario, it filters out brittle strategies and identifies alternatives that remain defensible across plausible futures.

Among the evaluation criteria, minimax regret performs best. Like the preceding methods, it requires no probability assignments and so is compatible with deep uncertainty. Its regret values are interpretable as opportunity costs relative to the best alternative in each scenario, which makes it transparent, and it remains organisationally feasible because it operates directly on the aggregated value scores from the preceding tasks without additional elicitation. Its behavioural-bias counterweight is particularly direct. By making the penalty of being wrong under each plausible future visible, it reduces the room for optimism bias and strategic misrepresentation that thrive when consequences are aggregated into single-point forecasts. The discrete scenario-by-alternative structure also supports systematic robustness assessment, through perturbation of scenario-specific impacts and aggregated value scores, addressed in Chapter 3.

2.7.4. Compatibility of selected components

Table 2.1: Framework components selected through the literature review.

Task	Selected method	Framework requirement	Methodological role
Indicator basis	Structured reduction of ESRS 2 and literature indicators	Define a defensible set of potentially relevant indicators	Combines regulatory traceability with literature-based extensions for yard-level decision support
Value structuring	AHP	Make value judgements explicit and traceable	Derives relative criterion weights through pairwise comparisons with consistency checks
Value mapping and aggregation	Performance functions and additive MAVT	Convert heterogeneous outcomes into comparable scores and combine them into one value per investment	Maps indicator outcomes onto a bounded evaluative scale and combines them through weighted aggregation
Future-dependent performance	Scenario analysis	Represent future uncertainty without assigning probabilities	Evaluates investment performance across plausible external futures
Evaluation	Minimax regret	Compare alternatives across scenarios without expected values	Identifies exposure to opportunity loss relative to the best-performing alternative in each scenario

The four chosen methods are individually defensible against the selection criteria. They must also compose into a coherent framework, since the output of each task is the input of the next. This subsection establishes that the pairwise interfaces between components are methodologically validated in the literature and that the underlying assumptions remain consistent across the pipeline.

The interface between value structuring and aggregation is the most direct. Additive MAVT requires explicit weights, and AHP provides them through ratio-scale pairwise judgements that can be converted into the weights additive aggregation needs. The use of AHP weights within an additive MAVT structure is methodologically validated. Measurable multiattribute value functions can be successfully assessed using the ratio judgements provided by the AHP methodology (Dyer, 2016). In environmental MCDA reviews, this hybrid is identified as standard practice, where decision-makers structure part of a model using AHP methods for convenience, while relying on value-based methods like MAVT for transparent calculations (Huang et al., 2011).

The interface between value mapping and aggregation is the assumption of preferential independence. Performance functions express within-criterion value judgements, while additive aggregation requires that trade-offs between any subset of criteria do not depend on fixed levels of the remaining criteria. Keeping performance functions monotonic within each indicator and eliciting weights at the criterion level rather than at the indicator-pair level preserves this independence (Dyer, 2016).

The interface between aggregation and future-dependent performance separates valuation from context. Weights and performance functions remain fixed across scenarios. Only the consequences of investments change. This separation prevents shifts in external context from being confused with shifts in organisational preferences, and is the established way of integrating value-based MCDA with scenario analysis in strategic decision contexts (Marttunen et al., 2017).

The interface between future-dependent performance and evaluation is structural. Minimax regret operates directly on the matrix of aggregated value scores by scenario produced by the preceding tasks, requiring no additional elicitation or transformation.

Beyond the individual interfaces, the framework as a whole is methodologically validated in the broader MCDA literature. Integrating distinct MCDA techniques to leverage their complementary strengths is an established practice in environmental and strategic decision-making (Huang et al., 2011; Marttunen et al., 2017), and the specific combination of AHP for weight elicitation, additive MAVT for aggregation, scenario analysis for future-context modelling, and minimax regret for cross-scenario evaluation reflects an established response to deep uncertainty in strategic investment contexts (Lombardi Netto et al., 2026; Marttunen et al., 2017).

These connections are summarised in Table 2.1, which lists each task, the selected method, the framework requirement it addresses, and its methodological role in the pipeline. This selection provides the theoretical foundation for the framework operationalisation developed in Chapter 3, which specifies the procedural steps, elicitation protocols, and robustness assessment used to apply the framework in the case of Koninklijke de Vries Scheepsbouw.

3
Methodology

This chapter operationalises the framework logic derived from Chapter 2 for the comparison of sustainability investments in a superyacht shipyard. It defines the research design and analytical boundary, and specifies how investments are assessed, valued, and compared under uncertainty.

3.1. Research design and case scope

3.1.1. Research design

The thesis follows a design-oriented case study approach. Its purpose is to analyse the existing decision process and to develop and test a framework for yard-level sustainability investment prioritisation under uncertainty.

The framework is developed from the literature and then configured and applied in one industrial case. This case is used to examine whether the framework can be operationalised under realistic organisational conditions, using actual constraints on time, data, and expert availability. The case therefore functions as an empirical test of methodological feasibility and analytical usefulness.

A single case is appropriate because the objective is framework development and application rather than cross-case comparison. It provides the organisational setting, data conditions, and expert judgement needed to evaluate whether the framework remains coherent and usable when moved from conceptual design to practical use.

3.1.2. Case scope and analytical boundary

The empirical setting is a superyacht shipyard operating in a customised, project-based production environment. The analysis focuses on yard-level sustainability investments that compete in recurring group-level capital allocation rounds.

The framework is limited to investments under direct yard control, including facilities, infrastructure, governance systems, digital systems, vehicle fleets, and security measures. Vessel-specific design choices and owner-driven customisation decisions fall outside this boundary, because they are not primarily governed through yard-level capital allocation.

This boundary reflects the decision problem addressed in the thesis. The investments considered differ in scale, impact pathway, feasibility, and uncertainty, yet must still be compared within one organisational decision process. Case-specific parameter choices and participant composition are reported in Chapter 4.

3.2. Operational framework overview

This section presents how the framework components selected in Chapter 2 are combined. It follows a fixed analytical sequence with four layers. First, the evaluation basis is defined through an active indicator set, a finite set of admissible investment alternatives, and a bounded scenario set. Second, investment consequences are constructed at indicator level. Third, these consequences are converted into comparable values through elicited performance functions, elicited AHP-based criterion weights, and MAVT aggregation. Fourth, investment alternatives are compared across scenarios through minimax regret, after which framework robustness is assessed by varying uncertain elicited inputs.

This sequence is methodologically important because it separates consequence construction from value judgements and robustness evaluation. It therefore helps prevent changes in future context from being confused with changes in organisational preferences.

Figure 3.1 shows the operational framework used in this thesis.

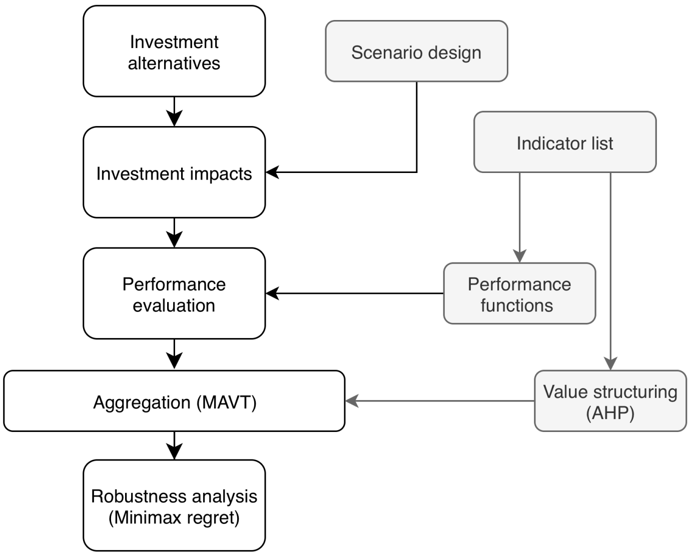


Figure 3.1: Flow chart of framework operationalisation.

3.3. Defining the evaluation basis

3.3.1. Indicator selection

Indicators define the criteria through which investment alternatives are evaluated. The indicator universe derived in Chapter 2 contains 206 candidates spanning environmental, social, governance, financial, and feasibility effects. This universe is intentionally broad, so that all potentially relevant effects can enter the framework.

The universe is reduced in two stages. The first stage is performed once and yields the framework-level active set. The second stage is performed for each investment round and yields the round-specific evaluation set.

Stage 1: Framework-level reduction (206 → 20).
The framework-level reduction proceeds in three steps, depicted in Figure 3.2.

Step 1: Investment relevance screening (206 → 65). The initial universe is screened against the decision context. Indicators that only provide contextual company information (135) or that describe targets rather than evaluative criteria (6) are excluded, leaving 65 investment-relevant indicators.

Step 2: Yard-level decision relevance screening (65 → 25). Remaining indicators are screened for whether they can in principle be influenced by investments in superyacht shipyards. Indicators that fall outside this decision boundary (40) are excluded.

Step 3: Consolidation (25 → 20). Of the 25 yard-level relevant indicators, 12 are kept as unique criteria while 13 are combined into 8 proxy indices where this improves interpretability and reduces redundancy. This yields the active framework set of 20 indicators.

Detailed explanation of indicators can be found in Appendix A.

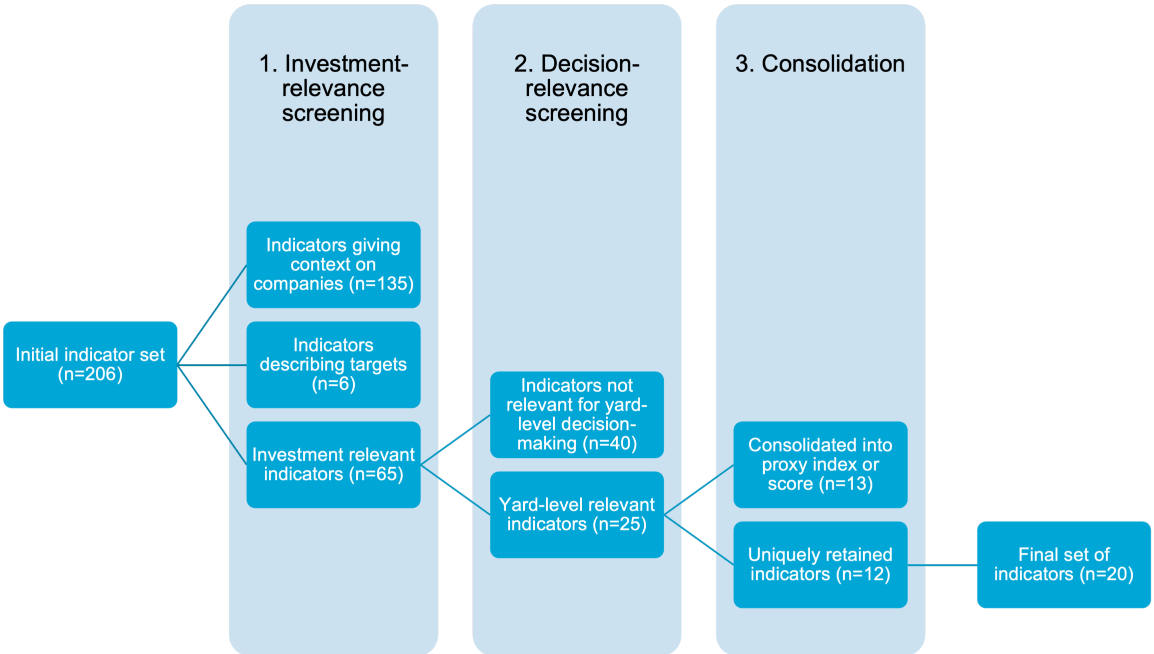


Figure 3.2: Stage 1: Three-step reduction of the initial indicator universe from 206 literature derived indicators to a framework-level active set of 20 indicators.

Stage 2: Round-specific reduction (20 → fewer).
For any given investment round, a subset of the 20 framework indicators may be non-discriminating. An indicator is inactive for a round when it is already at maximum performance, remains unchanged across all alternatives under consideration, or falls outside the practical decision boundary of that specific decision. Such indicators are set aside for the round but remain in the framework set, since an indicator that is non-discriminating in one round may become decision relevant in another. All round-level exclusions are documented together with their rationale.

The second reduction is used for feasibility of the framework. The number of indicators directly affects elicitation burden. The number of AHP pairwise comparisons grows quadratically with the number of criteria (Saaty, 2008), performance-function elicitation scales with the size of the active set, and assessment of scenario-dependent effects becomes more demanding as the indicator space expands. Indicator reduction therefore improves usability without changing the underlying conceptual scope of the framework.

The resulting active evaluation set is then used consistently throughout the analysis. Framework results should be interpreted as conditional on the selected indicator set.

3.3.2. Investment alternatives

Within the framework, investments are treated as discrete decision alternatives that can be selected by decision-makers. The framework compares a finite set of such alternatives and does not model adaptive investment pathways over time.

Each investment must be specified clearly to support consistent interpretation. At minimum, this requires a description of the intervention, its organisational scope and the primary mechanism through which it is expected to affect the yard. This information defines the causal entry point of the investment into the framework.

Only investments already deemed admissible for consideration by the company enter the formal comparison. The framework is therefore a decision support tool rather than a gatekeeping tool. It does not determine whether an investment is allowed, necessary, or impossible. Instead, it compares the relative value and robustness of alternatives that decision-makers have already accepted as worthy of

Table 3.1: Final active indicator set used for investment decision-making.

Environmental	Social	Governance	Financial	Feasibility
Scope 1 GHG emissions	Absence rate	Supplier ethical standards	EBITDA	Implementation complexity
Scope 2 GHG emissions	Employee turnover	Corruption / bribery	CAPEX	Organisational effort
Water consumption	Safety / incident rate	Human rights incidents	Financial risk	Technology maturity
Total waste generation	Workforce flexibility	Fines / penalties		
Total hazardous waste generation	Training intensity			

consideration. Company specific exclusion criteria or hard minimum requirements should therefore be applied before the formal comparison begins.

Investments are defined independently of scenarios. Scenarios alter the context in which an investment performs, not the identity of the investment itself.

3.3.3. Scenario design

Scenarios represent a small set of plausible external futures under deep uncertainty. Their purpose is to test how the relative attractiveness of investments changes across materially different operating environments (Marchau et al., 2019). One established approach to scenario construction is to identify principal axes of uncertainty and combine their extreme states into internally consistent future contexts (Marchau et al., 2019).

Scenario axes are identified through expert elicitation and analysis of the industry and company context. To qualify as a scenario axis, a driver must be (i) highly influential for the investment alternatives under consideration, (ii) meaningfully uncertain, and (iii) outside the control of the decision-maker. The first criterion ensures analytical relevance, while the third ensures that scenarios capture exogenous conditions rather than strategic choices (Marchau et al., 2019; Zwaginga, 2026). Axes must be analytically distinct, so that two axes do not reflect the same underlying uncertainty.

In this thesis, two axes are used, producing four discrete scenarios through combination of their extreme states. Two axes strike a practical balance between analytical depth and elicitation burden. Additional axes would increase the number of scenarios rapidly and reduce interpretability for decision-makers (Lempert, 2019; Stanton & Roelich, 2021). Each scenario represents a coherent and internally consistent future context and is treated as a conditional external state rather than a probabilistic forecast.

Within the framework, scenarios affect investment consequences only (Zwaginga, 2026). They do not alter value weights or performance functions. Variation across scenarios should therefore reflect changing consequences rather than changing preferences. Each scenario is ultimately translated into inputs for a causal map, following Sterman (2000), where it affects the external conditions under which investment effects propagate.

3.4. Constructing investment consequences

3.4.1. Impact assessment logic

Impact assessment specifies how the effect of each investment on each indicator is determined. The objective is to apply a consistent assessment route for every investment indicator combination so that alternatives remain comparable and intuitive judgement is made explicit.

For each investment indicator combination, the first methodological decision is whether the effect can be assessed directly or whether it must be propagated through the SDR. The direct route is used when the effect can be specified without modelling system interaction. The SDR route is used when the effect depends on indirect pathways, feedback structure, or interactions between multiple organisational variables.

After the route is selected, impacts are represented either quantitatively or qualitatively. Quantitative representation is used when the effect can be expressed through measurable data, calculations, or defensible conditional assumptions. Qualitative representation is used when the effect cannot be specified with sufficient precision and is then expressed on the common five-level scale shown in Table 3.2. Scenario-dependent impacts may be represented quantitatively only when they follow from explicit scenario assumptions together with defensible technical, operational, or accounting relations. Where longer causal chains or feedback-dominated behaviour prevent credible quantification, impacts remain ordinal.

Quantitative baselines, where available, define the reference state for all impact assessments of that indicator. When baseline values are not available, the five-level scale is used exclusively.

It is important to distinguish direct effects from net system effects. An investment may have limited direct influence on a variable and still generate a negative overall outcome through broader system dynamics in a given scenario. In such cases, the framework records the net consequence relevant for evaluation rather than the direct causal contribution in isolation.

This approach introduces methodological risks. Mixed data types may reduce consistency across assessments. Results may remain sensitive to baseline definitions and to the way qualitative and quantitative information are combined. These risks are accepted in exchange for a key benefit. The method makes intuitive judgement explicit rather than leaving it as a tacit assumption embedded in the analysis (Sterman, 2000).

3.4.2. System Dynamics Representation

The SDR is the construction route used when consequences cannot be assessed directly. It provides a shared structural representation of the decision environment against which scenarios and investments are interpreted.

The SDR is developed in three stages. First, a Causal Loop Diagram (CLD) is constructed. Second, this structure is formalised as a Stock–Flow Representation (SFR). Third, the SFR is used to propagate scenario conditions and investment interventions to relevant indicator consequences (Dall-Orsoletta et al., 2025; Sterman, 2000).

The SDR is used qualitatively under deep uncertainty and is not intended as a calibrated predictive model. Its role is to support disciplined causal reasoning by making explicit how scenario conditions and investment effects interact through feedbacks, delays, and accumulations (Stanton & Roelich, 2021; Sterman, 2000).

Causal Loop Diagram

The Causal Loop Diagram is the first formal modelling layer. Its role is to identify the causal relations and feedback loops that shape investment consequences. This matters in complex organisational settings, where relevant effects arise through indirect pathways and non-linear interactions that unaided mental models tend to underrepresent (Dall-Orsoletta et al., 2025; Marchau et al., 2019; Montibeller & Franco, 2011; Sterman, 2000).

In this thesis, the CLD is constructed in three steps. First, the organisational domains relevant to the decision problem are identified from the industry and company context, together with the key variables within each domain. Second, direct causal links between these variables are specified, each annotated with a polarity indicating whether an increase in the cause raises or lowers the effect. Third, the resulting link structure is inspected for closed loops, which are classified as reinforcing or balancing and retained only when they are relevant to at least one indicator or scenario axis. Variables and links that do not participate in a retained loop or in a direct pathway to an indicator are pruned to keep the diagram interpretable (Sterman, 2000).

Stock Flow Representation

The Stock Flow Representation formalises the CLD by distinguishing stocks, flows, and auxiliary variables (Dall-Orsoletta et al., 2025; Sterman, 2000). Unlike the CLD, which shows only the direction of influence, the SFR requires that changes in system state occur through explicitly defined inflows and outflows. Within the framework, this added structure serves causal consistency. It reduces the risk that accumulations and rate dependencies remain hidden in qualitative reasoning.

In this thesis, the SFR is constructed from the CLD in three steps. First, each variable in the retained CLD is classified as a stock, a flow, or an auxiliary variable, based on whether it represents an accumulation, a rate of change, or an intermediate quantity. Second, stocks are connected to their governing inflows and outflows, and auxiliary variables are used where the rate depends on other system states. Third, the resulting structure is checked for consistency with the loop polarities established in the CLD, so that reinforcing and balancing feedbacks identified at the causal level are preserved in the stock flow structure (Sterman, 2000).

Simulation and indicator mapping
The SFR is used as the basis for simulation. Each scenario is first translated into changes in relevant external conditions, establishing a scenario-specific baseline state. Each investment is then represented through a limited set of primary effects on model variables, which are propagated through the SFR. Restricting investments to a limited number of primary effects focuses attention on the most influential pathways and keeps the SDR usable.

The resulting system changes are translated into impacts on the retained indicators through predefined mappings. Where the propagated effect cannot be expressed credibly as a cardinal outcome, it remains ordinal or directional.

The SDR introduces methodological risks specific to this stage. Because the model remains judgement based, its structure may reflect simplifications, omissions, or causal assumptions that are more stable than the available evidence justifies. Indicator mappings in particular must be specified carefully to avoid double-counting against directly assessed effects.

The output of this stage is a set of scenario-specific and investment-specific directional or quantified effects at indicator level for those combinations requiring indirect assessment.

3.4.3. Indicator levels

Indicator levels are constructed for every investment scenario indicator combination from the preceding steps. They represent the final output of consequence construction and the immediate input for value scoring.

For directly assessed combinations, the indicator level follows from measured data, calculation, or structured expert judgement. For indirectly assessed combinations, it follows from propagation through the SDR and subsequent translation into the relevant indicator.

Indicator levels must be constructed consistently across all alternatives and scenarios, using the same baseline, interpretation rules, and direction of improvement.

For some indicators, impacts cannot be expressed directly as quantitative differences even though a meaningful quantitative baseline exists. In such cases, the framework introduces an indicator specific translation rule that maps the ordinal outcome onto a baseline-referenced level. The rule is fixed before evaluation based on the indicator's baseline scale and the plausible range of its variation, and is applied consistently across all investments and scenarios in the round. Where a baseline referenced quantitative proxy is defensible, the ordinal level is translated into a multiplier or calibrated adjustment relative to the baseline. Where this is not defensible, the outcome remains on a bounded ordinal proxy scale (Dyer, 2016; Hojnik et al., 2020).

Together, these indicator levels provide a scenario-specific description of how each investment performs on each criterion, but they are not yet directly comparable across criteria. That conversion is performed through the performance functions in the next section.

3.5. Converting consequences into comparable values

3.5.1. Performance functions

Performance functions define how performance on each indicator is valued within the framework (Keeney, 1988). Their purpose is to express how different levels of indicator performance are valued by

Table 3.2: Qualitative indicator levels used in impact assessment.

Level	Meaning
- -	Strong negative change
-	Negative change
0	No expected change
+	Positive change
+ +	Strong positive change

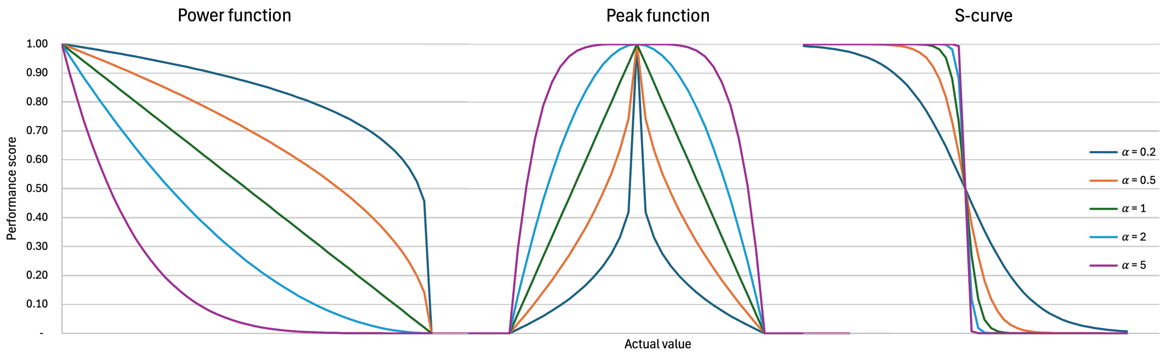


Figure 3.3: Performance function template families available for indicator valuation. Each family admits five curvature levels ($\alpha \in \{0.2, 0.5, 1, 2, 5\}$). From left to right: monotonic decreasing (invertible for increasing indicators), peak-shaped (preferred intermediate level), and S-curve (invertible). Linear valuation is the $\alpha = 1$ case of the monotonic family.

the organisation.

Each indicator is assigned one fixed performance function. These functions map indicator outcomes to dimensionless performance scores on the interval $[0, 1]$, where 1 denotes the most preferred outcome and 0 the least preferred outcome (Dyer, 2016). In MAVT terminology these are value scores. In this thesis, the terms performance function and performance score are used to distinguish value conversion (performance functions) from criterion weighting (AHP).

Performance functions are elicited through a restricted set of templates rather than through unconstrained free-form specification. This restriction applies across indicator types, including both continuously and ordinally represented indicators. It is a deliberate design choice. It limits elicitation complexity, improves consistency across indicators, and makes it easier to revise or extend the framework later when additional information becomes available (Boix-Cots et al., 2022). The template families used in this thesis are monotonic power curves, increasing and decreasing S curves, and peak shaped curves for indicators with an intermediate preferred level.

The same template principle is retained across indicator types. Indicators can enter the value function through ordered levels or continuous measurements, which preserves comparability across indicator types. Because the value function is defined continuously, quantitative data can replace ordinal inputs as they become available without requiring the value functions to be re-elicited.

For all indicators, admissible template parameters are fixed in advance. When robustness is tested, variation in template curvature is restricted to adjacent admissible parameter values so that the perturbed function remains similar to the originally elicited template choice.

3.5.2. Performance scoring

Performance scores are obtained by applying the fixed performance functions to the estimated indicator consequences of each investment. This step converts heterogeneous impact tables into a common bounded scale that can be compared across indicators.

For each investment, indicator, and scenario, the estimated consequence is entered into the corresponding performance function. Quantitative consequences are inserted directly. Qualitative consequences are first located on the common qualitative scale and, where necessary, translated into the relevant indicator specific proxy level before being scored.

The result is a scenario-specific performance score on the interval $[0, 1]$. If a scenario does not alter the estimated consequence for a given investment indicator combination, the performance score remains unchanged across scenarios.

The resulting performance scores form the direct input for criterion weighting and MAVT aggregation.

3.5.3. Value structure through AHP

The value structure is elicited using the AHP (Saaty, 2008). AHP is used to derive relative importance weights for the criteria through pairwise comparisons among elements at the same hierarchical level.

Each participant first completes individual pairwise comparison matrices. These matrices are checked separately for logical consistency. This thesis adopts Consistency Ratio (CR) < 0.10 as the target consistency threshold. Where this threshold is not met, participants are invited to review and, where appropriate, revise the relevant judgements. In expert elicitation on complex criteria, however, full compliance with $CR < 0.10$ is not always achievable in practice. Residual inconsistency is therefore not ignored, but documented explicitly and considered in the interpretation of results and, where relevant, in sensitivity analysis (Saaty, 2008).

After individual review, expert judgements are aggregated at the level of the pairwise comparison intensities using the geometric mean method. The aggregated comparison matrices are then used to derive local and global criterion weights for the evaluation.

The resulting weights remain fixed within an evaluation round and across scenario analysis. In this way, scenarios alter the estimated consequences of investments, but not the relative importance assigned to the criteria.

3.5.4. Aggregation through MAVT

MAVT aggregation combines the criterion level performance scores into one overall value score for each investment under each scenario. Aggregation is implemented as an additive weighted sum (Dyer, 2016; Gan et al., 2017):

$$V_{is} = \sum_{j=1}^{n} w_j \, p_{ijs}, \tag{3.1}$$

where V_{is} denotes the aggregated value score of investment i under scenario s, w_j denotes the AHP weight of criterion j, and p_{ijs} denotes the performance score of investment i on criterion j under scenario s.

The resulting value score remains bounded on the interval $[0, 1]$. A higher score indicates stronger overall performance under the fixed value structure of the evaluation round.

The additive form requires mutual preferential independence between criteria, meaning that trade-offs between any subset of criteria do not depend on the fixed levels of the remaining criteria. Additionally, it permits full compensation between the criteria (Dyer, 2016; Gan et al., 2017).

3.6. Robust option selection and framework robustness

Robustness is assessed at two analytically distinct levels. First, investment robustness is assessed across scenarios through minimax regret. Second, framework robustness is assessed by varying uncertain elicited inputs and observing whether the overall conclusion remains stable. These two forms of robustness should not be merged, because they answer different methodological questions.

3.6.1. Robust option selection through minimax regret

Robust option selection is based on Savage's minimax regret criterion (Hallegatte et al., 2012; Stanton & Roelich, 2021). It is applied after MAVT aggregation and before framework robustness testing.

For each scenario, regret is defined as the difference between the aggregated value of the best performing investment in that scenario and the aggregated value of the investment under consideration:

$$R_{is} = \max_{k \in I} V_{ks} - V_{is}. \tag{3.2}$$

For each investment, the maximum regret across all scenarios is then determined:

$$MR_i = \max_{s \in S} R_{is}. \tag{3.3}$$

The preferred investment is the one with the lowest maximum regret:

$$i^* = \arg\min_{i \in I} MR_i. \tag{3.4}$$

3.6.2. Framework robustness through Monte Carlo analysis

Framework robustness is assessed through Monte Carlo analysis. The purpose is to test whether the framework conclusions remain stable when uncertain elicited inputs are varied within predefined bounds.

In each Monte Carlo run, uncertain elicited inputs are perturbed, and the full framework is recalculated. This yields an updated set of aggregated value scores, scenario-specific regret values, and a revised minimax regret recommendation. The distribution of recommendations across runs indicates how strongly the final conclusion depends on the elicited inputs. Four input families may be varied, each with its own perturbation logic.

AHP pairwise comparisons.
The aggregated geometric mean pairwise comparison values are perturbed multiplicatively in log space, truncated to the Saaty scale [1/9, 9]. A new AHP weight vector is derived from the perturbed matrix in each run. Multiplicative perturbation preserves the reciprocal structure of pairwise comparisons and reflects the ratio scale on which AHP judgements are expressed.

Performance function curvature.
The curvature parameter α of the quantitative performance function templates is varied by discrete shifts to adjacent admissible template values. If the elicited value lies in the interior of the admissible set, variation is allowed to the adjacent lower and higher options. If it lies at the boundary, variation is restricted to the nearest available adjacent option. Restricting variation to adjacent templates keeps the sampled functions close to the originally elicited preference shape.

Ordinal to baseline translation.
The translation rule $T_j(\cdot)$ is fixed for the evaluation round but is varied during Monte Carlo analysis to test sensitivity to its calibration. The baseline (ordinal level 0) is held fixed as the reference state and is not perturbed.

Perturbation is applied to the positive side of the scale for each ordinal level $\ell \in \{+, ++\}$ with elicited multiplier $T_j(\ell) > 1$. The perturbation amplitude equals the distance between the elicited multiplier and the baseline. The perturbed value therefore ranges from the baseline itself up to $2T_j(\ell) - 1$, centred on $T_j(\ell)$.

The negative side is set reciprocally to preserve symmetry around the baseline:

$$T_j^{\text{pert}}(--) = 1/T_j^{\text{pert}}(++), \qquad T_j^{\text{pert}}(-) = 1/T_j^{\text{pert}}(+).$$

This ensures that the sampled translation rule does not systematically skew the indicator scale upward or downward relative to the baseline.

Scenario-specific investment impacts.
Quantitative impacts are perturbed proportionally. Qualitative impacts on the five-level scale are perturbed by at most one step up or down, clipped to the admissible range. Where ordinal outcomes are translated onto a quantitative baseline, the translation value itself is perturbed within its admissible range. Because these inputs form the immediate consequence layer of the framework, sensitivity to them is analytically expected and should not in itself be interpreted as a methodological defect.

3.6.3. Interpreting robustness results

The main output of the Monte Carlo analysis is the frequency with which each investment is selected as the minimax regret option across all runs.

High selection frequencies indicate that the framework conclusion is relatively stable under the specified input variation. Lower frequencies, or repeated switching between a small number of alternatives, indicate greater sensitivity to elicited assumptions. In this thesis, the primary criterion for framework robustness is stability of the qualitative conclusion, meaning that the overall interpretation of which investments are attractive, vulnerable, or scenario sensitive remains broadly stable even when modest numerical differences occur across runs.

It is also important to distinguish impact sensitivity from valuation sensitivity. Sensitivity to scenario-specific investment inputs reflects dependence on the consequence layer, meaning what investments are assumed to do under each scenario. Sensitivity to AHP weights, performance-function shape, or translation assumptions reflects dependence on the valuation layer, meaning how decision-makers value those consequences. The two are not interchangeable, and a robustness result against one does not imply robustness against the other. Sensitivity to the consequence layer is analytically expected, since these inputs directly determine the performance scores that propagate through the remainder of the framework.

3.7. Elicitation procedure

The framework relies on structured expert elicitation because several inputs cannot be derived mechanically from observed data alone. This applies in particular to criterion importance, performance function shape, and scenario-dependent impacts whose effects are uncertain, indirect, or only partially measurable. Expert judgement is therefore treated as a methodological necessity rather than as a fallback.

Elicitation proceeds in four stages. First, the goal, decision context, and methodological logic of the framework are introduced. Second, the screened indicator sets and investment definitions are confirmed. Third, AHP pairwise comparisons are completed individually. Fourth, performance functions and key scenario-dependent impacts are elicited using templates, baseline data, and structured discussion.

3.8. Methodological limitations

The methodology is subject to several limitations. First, indicator reduction, consequence construction, and performance scoring involve researcher judgement, even where these steps are documented and supported by expert input. Second, parts of the framework rely on subjective assessment because not all relevant dimensions can be measured directly or represented through defensible quantitative relations. This applies particularly to qualitative impacts, indirect effects, and SDR based consequence construction.

Third, the additive aggregation structure simplifies interactions between criteria by assuming preferential independence. Fourth, the five-level ordinal scale reduces some effects that are in practice more continuous or ambiguous. Fifth, judgement based inputs remain vulnerable to bias, especially where direct measurement is unavailable.

These limitations define the conditions under which the methodological application should be interpreted. Broader limitations of the framework, including transferability beyond the case context and implications for practice, are discussed in Chapter 5.

4
Case Study

This chapter applies the decision support framework developed in Chapter 3 to Koninklijke de Vries Scheepsbouw in Aalsmeer. The purpose is to assess whether the framework can provide a transparent, robust, and practically useful comparison of heterogeneous investments in a realistic superyacht shipyard context.

The chapter introduces the case context and design, explains how the generic framework was operationalised for this case, and works through Investment A, the electrification of the company car fleet, as an illustrative example. It then reports the main results and interprets them against current practice, verification, validation, and case-specific limitations. Detailed tables and supporting material are reported in Appendix A.

4.1. Case context and chapter purpose

In this case study, the framework is used to compare four investment alternatives and to make the underlying trade-offs explicit. The purpose is to provide structured decision support for a limited yet varied set of alternatives. The case concerns Koninklijke de Vries Scheepsbouw, a custom superyacht shipyard operating in a project-based, capital intensive production environment. In practice, the final investment decision remains the responsibility of the board, while yard-level management plays an important role in originating and supporting investment proposals. The chapter therefore examines both what the framework produces in this case and where the resulting conclusion remains sensitive to uncertain assumptions.

4.2. Case design

4.2.1. Elicitation setup

The case study was conducted in four steps. First, available data on the indicators and investment alternatives were collected. Second, elicitation was conducted in two sessions using these data to structure the discussion. Third, the collected case information and elicited expert judgements were combined into the final framework inputs. Fourth, these inputs were used to determine comparative performance across scenarios and to assess both investment and framework robustness.

Elicitation took place in two separate sessions with four participants. The first session consisted of one member of the Group Board of Directors (*Session 1*). The second session consisted of two yard-level Directors and the Group Sustainability Director (*Session 2*). The expert group spans three perspectives: the level making investment decisions, the level proposing them, and the sustainability function seeking to steer toward more sustainable outcomes.

Because of time constraints, elicitation was conducted in one session per group rather than through a longer multi round process. This reduced the breadth of expert input and likely lowered elicitation quality. The trade off was nevertheless accepted in order to test the framework with the actual decision-makers involved in the company context. Both sessions were recorded for traceability.

During the ninety minute sessions, participants were first introduced to the goal, purpose, and methodological logic of the framework. The indicator reduction step was then discussed. Next, the individual value structure was elicited using a prepared online questionnaire. Performance functions were elicited through templates, sketching, and structured discussion. Scenario dependent investment effects were identified during the open discussion phase, primarily in *Session 1*. Table 4.1 summarises the composition of the two sessions and indicates which elicitation activities took place in each.

After elicitation, a validation round was conducted to assess participant consistency, the plausibility of relative magnitudes, and alignment with operational knowledge.

Table 4.1: Composition of the elicitation sessions.

Session	Duration	Participants	Value structure	Performance functions	Scenario impacts
1	90 minutes	Member of the Group Board of Directors	✓	✓	✓
2	90 minutes	Two yard-level directors and the Group Sustainability Director	✓	✓	

4.2.2. Investment set

Four investments that the yard has recently implemented or is considering implementing are included in the case study. They were selected deliberately to test the framework's ability to compare alternatives with strongly different scales, impact pathways, and implementation logics. Investments that had already been implemented were included because this improved data availability and supported reflection on the realism of the elicited effects.

The four investments were selected because they differ in scale and impact pathway, and together test the main functions of the framework. **Investment A**, the electrification of the company car fleet, represents a comparatively traditional sustainability investment with a relatively direct and partly quantifiable environmental effect, comparing electric vehicles with conventional diesel replacements. **Investment B**, cybersecurity training for employees, tests whether the framework can handle an investment whose value lies mainly in uncertain risk reduction. **Investment C**, yard security improvements combining external disruption defence with internal access control and subcontractor hour tracking, provides the clearest test of strong scenario dependence. **Investment D**, replaces the final ramp with a dry dock, lifting capacity constraints on vessel size. It tests whether the framework can compare a large strategic investment with smaller sustainability-oriented alternatives without reducing the comparison to financial scale alone. Together they form a compact yet analytically broad test set, summarised in Table 4.2.

The case deliberately combines implemented and non implemented investments. This reflects a trade off between methodological control and practical realism. Because the framework is intended for decision support under incomplete information, it would not be representative to include alternatives for which no usable company information existed. Using investments that were already known within the organisation improved data availability and made the elicitation more concrete within the available time. Although some base case consequences of implemented investments are partly known retrospectively, their scenario dependent consequences remain uncertain because the future contexts represented in the case have not occurred and cannot be observed directly. Their inclusion therefore still provides a valid test of the framework's handling of uncertainty.

4.2.3. Scenario set

To evaluate the four investments under deep uncertainty, the case study evaluates the investments across four scenarios defined along two external uncertainty axes: social pressure and regulatory pressure. These axes were identified by yard experts as the most decision relevant external uncertainties for the present case (Kon. de Vries Scheepsbouw Personal Communication, 26-11-2025), and regulatory pressure is also supported in the literature as major driver of sustainability related decisions (Hirtenlehner, 2025). Social pressure captures uncertainty regarding public scrutiny, protest intensity, employer attractiveness, and broader legitimacy dynamics. Regulatory pressure captures uncertainty regarding compliance

Table 4.2: Overview of the investment alternatives.

ID	Investment	Category	Value type	Scenario sensitivity	Status
A	Car fleet electrification	Environmental	Emissions reduction	Low	Under discussion
B	Cybersecurity training	Governance	Risk mitigation	Moderate	Implemented
C	Yard security improvements	Operational	Continuity protection	High	Under discussion
D	Dry dock replacement	Strategic	Capacity and flexibility	Moderate	Implemented

obligations and the tightening or relaxation of formal institutional requirements. Other uncertainties were considered but not adopted as primary axes. Notably, yard experts did not identify market demand as the most relevant axis for the present case despite its long-term uncertainty.

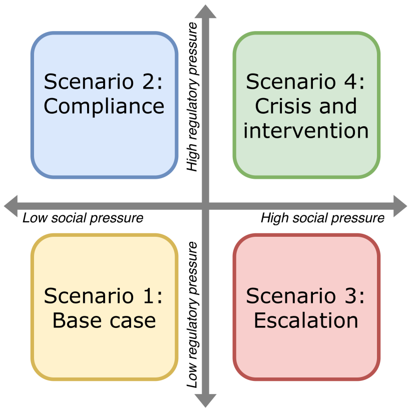


Figure 4.1: Case study scenario axes.

Scenario 1: Base case combines low social pressure with low regulatory pressure. The yard operates in a relatively stable environment with limited scrutiny, manageable workload pressure, and few disruption risks.

Scenario 2: Compliance environment combines low social pressure with high regulatory pressure. Operational legitimacy remains relatively stable, but new environmental, labour, or safety requirements increase compliance costs and reduce margins.

Scenario 3: Escalation combines high social pressure with low regulatory pressure. Public scrutiny increases, employer reputation deteriorates, labour availability declines, and disruption risks intensify through reinforcing legitimacy effects.

Scenario 4: Crisis and intervention combines high social pressure with high regulatory pressure. Scrutiny remains intense while institutional intervention adds compliance burdens and operational constraints to an already difficult environment.

These four scenarios form the fixed context for the case comparison reported in the remainder of the chapter.

4.3. Case-specific operationalisation

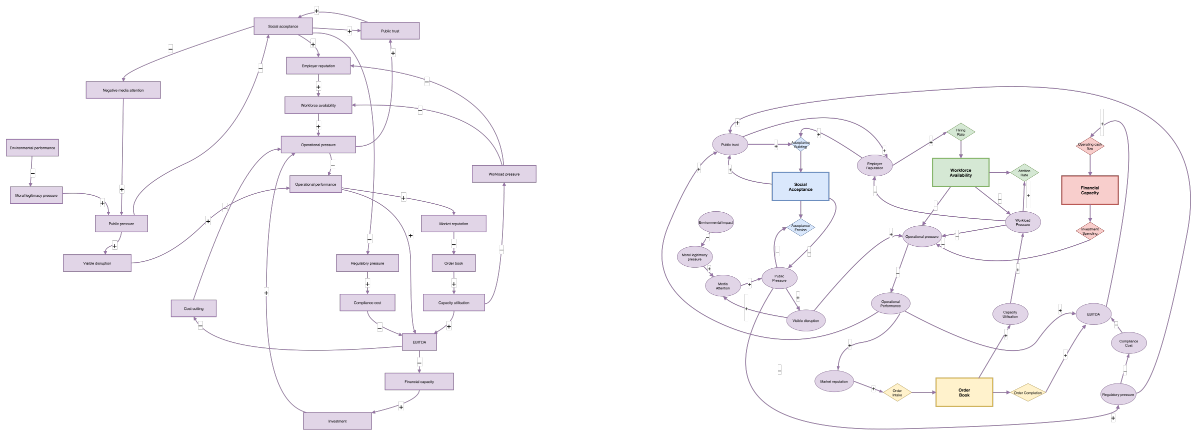
Following the investments under consideration and scenario construction, the generic framework described in Chapter 3 is now operationalised for the yard. This requires four case-specific steps: constructing an SDR that links scenario conditions to indicator consequences, reducing the indicator set to those that plausibly differentiate between the four alternatives, preparing baseline values and preliminary function anchors as inputs for elicitation, and eliciting the value structure and performance functions with the experts. The following subsections address each step in turn.

4.3.1. System dynamics representation for the case

In the case, the SDR is used to inform scenario dependent indicator levels for effects that cannot credibly be entered as direct and scenario invariant inputs. The scenario logic is informed by ongoing company discussions (Dall-Orsoletta et al., 2025). In the superyacht context, the dominant mechanism is moral legitimacy pressure (Stuart et al., 2023). Social dissatisfaction can accumulate and, when amplified by media attention, trigger visible disruption and operational pressure.

A CLD was constructed to capture the feedback structure of the case, based on company discussions, literature on legitimacy in contested industries, and internal understanding of shipyard operations. The model links four domains: social legitimacy, workforce dynamics, operational performance, and financial performance. Environmental performance is treated as an upstream driver of legitimacy pressure and social acceptance.

The CLD was then translated into the SFR to add temporal structure and dynamic consistency. The model centres on the key stocks social acceptance, workforce availability, order book, and financial capacity (Dall-Orsoletta et al., 2025; Sterman, 2000). Several system variables within the representation are structurally dependent, notably social acceptance, employer reputation, and workforce availability. The constructed CLD and SFR are shown in Figure 4.2, larger versions can be found in Appendix A and A respectively.



(a) Causal loop diagram (b) Stock–flow representation

Figure 4.2: System dynamics representations used in the case study.

After construction, the model structure and main causal pathways were reviewed by experts during elicitation. This validation focused on causal plausibility, internal consistency, and recognisability of the main feedback mechanisms within the yard context, rather than on formal behavioural calibration against historical time series. No equations were formulated. The SFR is used as a structured reasoning scaffold, through which scenario conditions and investment effects are propagated as ordinal directional changes and then mapped onto indicator outcomes via the rules reported in Appendix A. For each scenario in section 4.2.3, baseline levels are assigned to the exogenous drivers, after which the model is used to derive scenario consistent system conditions. Investment effects are then applied as interventions on selected system variables and propagated through the model. Finally, system level consequences are translated into indicator level outcomes. The formal mapping tables used for this translation are reported in Appendix A.

4.3.2. Case-specific indicator reduction

Preparation was completed before the expert elicitation sessions. Each investment was pre screened against the full indicator set using technical investment descriptions and the SFR to assess whether a plausible effect pathway existed in at least one scenario. If no plausible effect was expected for any investment in any scenario, the corresponding entries in the investment indicator scenario matrix were fixed to zero. Because such indicators do not differentiate between alternatives, they were excluded from the reduced set used in subsequent weighting and aggregation.

In this case, *water consumption, total waste generation, total hazardous waste generation, workforce flexibility, supplier ethical standards, and human rights incidents* met this exclusion criterion and were removed from the case-specific set. *Corruption and bribery* was initially screened out on the same basis, but was reintroduced after elicitation once a plausible pathway for governance oriented investments became apparent. A separate reduction had already been made in Chapter 3, where *organisational effort* was removed because it overlapped substantially with *implementation complexity*. The reintroduction of corruption and bribery shows that this step is not purely mechanical. It remains a structured judgement that can be revised when a credible impact pathway becomes visible during case application. The full reduction table is reported in Appendix A, and the per-investment screening reasoning for the other three investments is reported in Appendix A. A worked illustration of the screening logic for **Investment A** is provided in Section 4.4.1.

4.3.3. Case-specific indicator inputs

For each retained indicator j, the baseline level B_j, preferred direction, unit, and preliminary performance function anchors were prepared as inputs for expert discussion. Where company data were available, baseline and target values were prefilled. For ordinal indicators, category meanings were refined through structured discussion using the common five level scale (- -, -, *0*, +, + +).

A distinction was made between direct inputs and scenario dependent inputs. Direct inputs are those whose values do not vary across scenarios. These include *Scope 1 GHG emissions, Scope 2 GHG emissions, training intensity, CAPEX, corruption and bribery, implementation complexity, and technology maturity*. Their direct input values can be found in Appendix A. Scenario dependent inputs were derived through the SFR because their consequences vary by future context. These include *absence rate, employee turnover, safety and incident rate, fines and penalties, EBITDA*, and *Financial risk*. The expanded input table is reported in Appendix A.

All alternatives are evaluated over a common five year horizon, consistent with recurring capital allocation rounds (Kon. de Vries Scheepsbouw, Personal Communication, 03-12-2025).

4.3.4. Case elicitation and resulting value structure

The value structure was elicited through pairwise comparisons using the prepared questionnaire. Individual judgements were checked for consistency and then combined through the geometric mean method at the level of the pairwise comparison values to obtain one organisational weight profile.

Performance functions were elicited through templates, sketching, and structured discussion. For continuous indicators, the discussion focused on anchor points, preferred directions, and the shape of the value relationship. For ordinal indicators, the discussion focused on the interpretation of ordered performance levels and their mapping to bounded scores.

As discussed in Chapter 3, $CR < 0.10$ was treated as the target threshold for AHP consistency (Saaty, 2008). In the case, this threshold was not always met at the individual level. Several participants exceeded it on the dimension-level pairwise comparisons in particular, reflecting the time-constrained nature of the elicitation and the genuine difficulty of making strictly transitive judgements across heterogeneous criteria within a single ninety-minute session. The aggregated geometric-mean matrices, however, satisfied the threshold throughout, with $\max(CR) = 0.039$. This indicates that individual inconsistencies were not systematic but varied across participants and offset at the group level. Residual inconsistency is treated as part of the practical case limitation rather than as a methodological flaw, and the resulting weight profile is interpreted accordingly. The geometric mean weight profile is shown in Figure 4.3. Detailed consistency results are reported in Appendix A.

4.3.5. Case robustness setup

Robustness in the case study was assessed by varying the four input families as discussed in Chapter 3, separately, jointly, and in pairwise combinations. First, the geometric mean AHP pairwise comparison values were perturbed using low and high variation levels of 15% and 30%, after which a new set of AHP weights was derived in each run. Second, the curvature parameter α of the performance function templates was varied over adjacent admissible template options. Third, ordinal translation assumptions were varied. Fourth, the scenario-specific investment impacts entering the framework, including both ordinal and quantitative inputs, were varied. All perturbed inputs were then propagated through

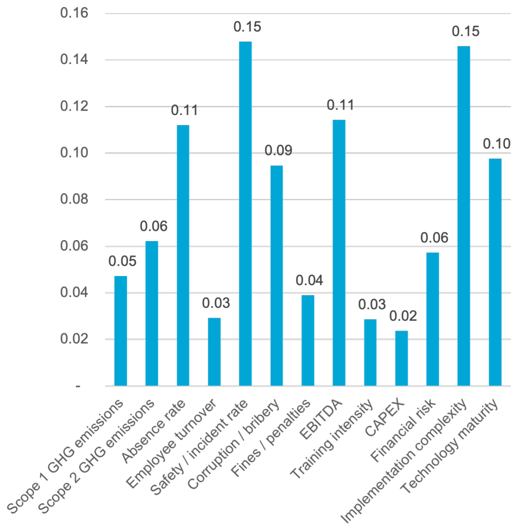


Figure 4.3: Geometric mean AHP weight profile for the case study based on the aggregated pairwise comparison judgements of the elicited experts $max(CR) = 0.039$. Detailed consistency results are reported in Appendix A.

the full framework, yielding updated performance scores, aggregated value scores, regret values per scenario and investment, and a revised minimax regret recommendation. Table 4.3 summarises the perturbation rules used in these Monte Carlo experiments, while detailed robustness outputs are reported in Appendix A.

Table 4.3: Monte Carlo perturbations used in the robustness analysis. Each experiment was run for 5,000 iterations.

Input	Perturbation	Settings	Bounds
AHP comparisons	Log multiplicative noise on aggregated pairwise comparisons	$\sigma = 0.15$ and 0.30	Truncated to Saaty scale $[1/9, 9]$
Performance function curvature	Discrete shift to adjacent admissible α levels	$\alpha \in \{0.2, 0.5, 1, 2, 5\}$	Restricted to nearest admissible level
Impact translation	Ordinal score to baseline change	Linear perturbation of ±50 percentage points applied to the baseline translation value.	Clipped to admissible range
Investment impacts (quantitative)	Uniform proportional perturbation of each quantitative impact	±10%	No additional bounds
Investment impacts (qualitative)	One step down, unchanged, or one step up	Probabilities 0.25, 0.50, 0.25	Clipped to range $[1, 5]$

4.4. Worked example: Investment A (Electrification of the company car fleet)

To make the case application transparent, this section works through **Investment A**. This investment is the clearest example of a relatively direct sustainability measure whose wider organisational value still depends on future context. The full worked example tables for the other investments are provided in Appendix A.

4.4.1. Pre-screening of indicators for Investment A

To make the indicator reduction step concrete, Table 4.4 shows the pre-screening outcome for **Investment A** across the full indicator set. Each indicator was assessed against the four scenarios to determine whether a plausible effect pathway from fleet electrification existed in at least one scenario. An indicator with no plausible pathway from **Investment A** can still be retained at the case level when another investment provides such a pathway, since the reduction operates on the investment set as a whole. The expanded screening tables for the other investments are reported in Appendix A.

Table 4.4: Pre-screening of all indicators for **Investment A** (electrification of company cars). For each indicator, the table records whether a plausible effect pathway from electrification exists in at least one scenario, and the reasoning behind that judgement. Indicators marked as *Not relevant for* ***Inv. A*** may still be retained at the case level if another investment provides a plausible pathway, since the reduction operates on the investment set as a whole.

Indicator	Verdict	Reasoning
Scope 1 GHG emissions	Retained	Direct reduction through replacement of combustion vehicles with electric equivalents.
Scope 2 GHG emissions	Retained	Increase through additional charging load on the grid.
Water consumption	Excluded	Water use limited to car washing; unaffected by drivetrain change and no pathway from other investments.
Total waste generation	Excluded	No change in waste streams expected from fleet replacement, and no pathway from other investments.
Total hazardous waste generation	Excluded	No change in hazardous waste streams expected from fleet replacement, and no pathway from other investments.
Absence rate	Not relevant for **Inv. A**	No plausible pathway from electrification; retained at case level via reputational and working-condition pathways from other investments.
Employee turnover	Not relevant for **Inv. A**	Marginal pathway through employer reputation only; retained at case level via stronger pathways from other investments.
Safety / incident rate	Not relevant for **Inv. A**	No plausible pathway from electrification; retained at case level via operational and disruption pathways from other investments.
Supplier ethical standards	Excluded	Drivetrain change does not affect supplier conduct in the relevant tiers, and no pathway from other investments.
Corruption / bribery	Not relevant for **Inv. A**	No pathway from electrification; reintroduced at case level after governance pathways from other investments became clearer during elicitation.
Human rights incidents	Excluded	No plausible pathway from electrification or any other investment in the set.
Fines / penalties	Retained	Plausible reduction in emission-zone and reporting-related fines under active regulation.
EBITDA	Retained	Plausible benefit under tightening emission regulation, since electricity is less heavily taxed than fossil fuels.
Workforce flexibility	Excluded	Fleet composition does not affect workforce capacity or skill adaptability, and no pathway from other investments.
Training intensity	Not relevant for **Inv. A**	No training component associated with vehicle replacement; retained at case level through other investments.
CAPEX	Retained	Direct upfront cost of vehicles and charging infrastructure.
Financial risk	Retained	Plausible reduction under regulation favouring electrification, through reduced depreciation and stranded-asset risk.
Organisational effort	Excluded	Implementation effort comparable to standard fleet replacement; not differentiating across investments.
Implementation complexity	Not relevant for **Inv. A**	Standard replacement project; retained at case level because it differentiates across the investments.
Technology maturity	Not relevant for **Inv. A**	Mature technology; retained at case level because it differentiates across the investments.

4.4.2. From system effects to indicator effects

The effects of **Investment A** are traced through four steps, following the structure formally defined in Chapter 3.

Investment A is the clearest example of a visible sustainability investment with mature technology and low operational disruption, as discussed in *Session 1*. Its direct effects enter the model through four system variables: public trust and employer reputation benefit from the visible sustainability signal; compliance cost decreases through alignment with commuting and transport emissions reporting requirements; and financial capacity decreases through the *CAPEX* of vehicles and charging infrastructure.

Table 4.5: Direct effects of **Investment A** on system variables.

System variable	Direct effect
Public trust	+
Employer reputation	+
Compliance cost	-
Financial capacity	-

Whether these direct effects are meaningful depends on the scenario context. The public trust and employer reputation effects are modest in magnitude and are not amplified by high social pressure: *Session 1* explicitly noted that visible environmental investments do not appease activists in escalating social conflict. The compliance cost benefit amplifies strongly whenever regulatory reporting is active, and the financial capacity drag is a constant *CAPEX* effect.

Table 4.6: Scenario-conditional amplification of the direct effects of **Investment A**.

System variable	Base	Compliance	Escalation	Crisis
Public trust	1	1	1	1
Employer reputation	1	1	1	1
Compliance cost	0	2	0	2
Financial capacity	1	1	1	1

The net effect on system variables, obtained by multiplying the direct effect vector by the scenario-conditional amplification, is shown in Table 4.7. Propagating these through the indicator mapping table in the Appendix (Table A), yields the scenario-dependent indicator outcomes in Table 4.8.

Table 4.7: Scenario-dependent effect of **Investment A** on system variables.

System variable	Base	Compliance	Escalation	Crisis
Public trust	+	+	+	+
Employer reputation	+	+	+	+
Compliance cost	0	- -	0	- -
Financial capacity	-	-	-	-

Table 4.8: Final indicator states for **Investment A** per scenario, obtained by combining the scenario baseline (Table A) with the marginal investment effect of **Investment A**. Signs follow the convention that a plus indicates a desirable state.

Indicator	Base	Compliance	Escalation	Crisis
Absence rate	0	0	-	- -
Employee turnover	+	+	-	-
Safety / incident rate	0	-	- -	- -
Corruption / bribery	0	-	-	- -
Fines / penalties	0	0	-	0
EBITDA	-	- -	- -	- -
Financial risk	-	-	- -	- -

Investment A produces modest, broadly stable benefits through employer reputation and fines-penalties, amplified where regulation is active. Its scenario sensitivity is small compared to the other alternatives, because its direct effect channels are limited to public trust, employer reputation, and compliance, none of which reverse or collapse under high pressure. The *CAPEX* drag appears as a constant small negative on *EBITDA* and *Financial risk*. The scope of the environmental benefit itself is discussed separately through the *Scope 1 GHG emissions* contribution in Section 4.4.3.

4.4.3. From indicator effects to performance scores

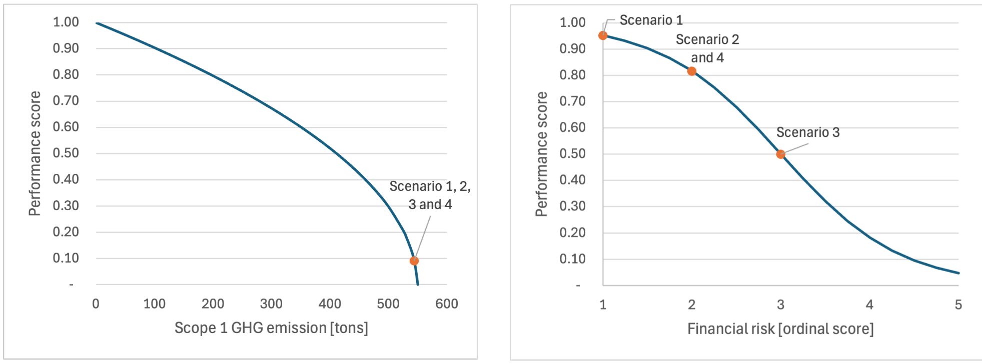
In the case study, performance functions were elicited through structured discussion and calibrated to reflect how decision-makers valued changes in indicator performance rather than the physical magnitude of change alone.

For **Investment A**, two indicators are used here to illustrate this step: *Scope 1 GHG emissions* as a direct continuous indicator, and *Financial risk* as a scenario dependent ordinal indicator. Together, these examples show how both direct and indirect investment consequences are translated into comparable performance scores.

Investment A reduces *Scope 1 GHG emissions*, by 4.50 tonnes, from 548.36 to 543.86 tonnes (see Appendix A for the underlying calculation). When this change is mapped onto the elicited performance function, it results in a low but positive performance contribution. This reflects that the investment improves environmental performance, while the magnitude of the improvement remains limited relative to the full reference range between the current state and the long-term target.

For *Financial risk*, the effect is scenario dependent and therefore cannot be represented by a single score across all futures. The same investment receives different evaluative scores depending on the scenario-specific risk level that enters the function. The performance function is non-linear, so movement towards the high risk end produces a sharper score reduction than equivalent movement within acceptable ranges.

These examples show that performance score mapping is the stage at which case-specific indicator outcomes become directly comparable across criteria. The *Scope 1 GHG emissions* example illustrates a small but positive direct score contribution that remains constant across scenarios, whereas the *Financial risk* example shows how scenario dependence enters the evaluation through differentiated mapped scores. The full worked example tables and performance function plots are reported in Appendix A.



(a) Performance function for *Scope 1 GHG emissions*. The orange marker shows the mapped performance score of **Investment A**, which is constant across all scenarios.

(b) Performance function for *Financial risk*. The orange markers show the mapped performance scores of **Investment A** for the scenario-specific *Financial risk* outcomes used in the worked example.

Figure 4.4: Illustrative performance functions for the worked example of **Investment A** (electrification of company cars). Panel (a) shows the direct *Scope 1 GHG emissions* indicator, while panel (b) shows the scenario dependent *Financial risk* indicator.

4.5. Case results

4.5.1. Baseline and scenario-specific outcomes

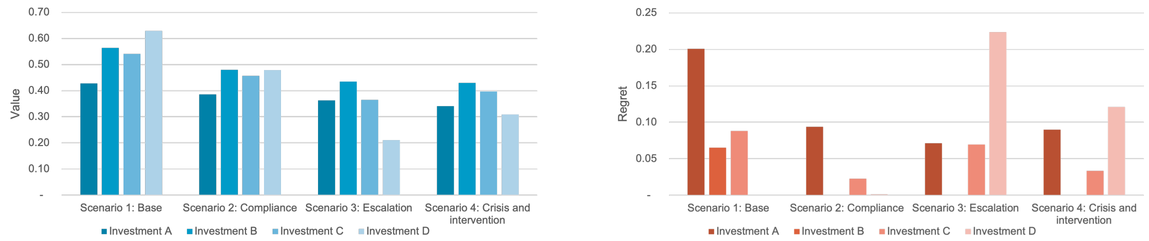
The framework generates three main types of outputs for the case study. First, it produces a baseline comparison of the four investments under fixed assumptions. Second, it shows how the relative attractiveness of the investments changes across the four scenarios. Third, it produces regret profiles that make the opportunity cost of each investment visible across alternative futures.

In substantive terms, the results show a clear difference between baseline attractiveness and robustness under uncertainty. **Investment D** (dry dock replacement) achieves the highest aggregated value score in the base case (Figure 4.5a; Table A). **Investment B** (cybersecurity training) achieves the lowest maximum regret across the four scenarios and is therefore the minimax regret choice (Figure 4.5b; Table A). The contrast between these two outcomes is analytically important because it shows that the investment with the highest upside under relatively favourable conditions is not necessarily the investment with the lowest vulnerability across uncertain futures.

The resulting value and regret per scenario and investment are shown in Figure 4.5. Detailed result tables are reported in Appendix A.

4.5.2. Framework robustness in the case

Framework robustness was assessed through the Monte Carlo experiments defined in Section 4.3.5. Each run was iterated 5,000 times. The seven runs separate into two analytical groups. Five perturb only the valuation layer: AHP under low and high variation, the performance function curvature parameter



(a) Aggregated value score of each investment across the four case study scenarios.

(b) Scenario-specific regret of each investment relative to the best performing alternative.

Figure 4.5: Baseline case study results comparing the scenario-specific value and regret profiles of the four investment alternatives. Higher value scores are preferred, while lower regret indicates greater robustness across scenarios.

α, the translation matrix, and the joint variation of AHP (high), α, and the translation matrix. One perturbs only the consequence layer, through the scenario-specific investment impacts. The seventh perturbs all four input families together. The valuation layer governs how consequences are valued. The consequence layer governs what those consequences are. Their robustness behaviour differs sharply.

Under valuation-layer perturbation, the recommendation almost does not move. **Investment B** is selected as the minimax regret option in all 5,000 runs under each of the four single-input perturbations: AHP (low), AHP (high), α, and translation matrix. It is also selected in 4,993 of 5,000 runs when the three valuation inputs are perturbed jointly (Figure 4.6). The corresponding value (panel a) and regret (panel b) violin plots in Figure 4.7 illustrate why. Although the value distributions show moderate spread within each scenario, the recommendation depends on the gaps between investments in each scenario, not on their absolute values. The bounded perturbation applied here shifts these gaps only modestly. In Scenarios 3 and 4, **Investment B** wins under fixed assumptions. The perturbation is too small to flip this result. **Investment B**'s regret in these scenarios therefore stays near zero across runs.

The picture changes once the consequence layer is perturbed. When only the scenario-specific investment impacts are varied, no investment commands a majority. **Investment B** is selected in 1,925 of 5,000 runs, **Investment D** in 1,822, **Investment C** in 1,198, and **Investment A** in only 55 (Figure 4.6). The pattern is comparable when all four input families are varied jointly: **Investment D** is selected in 2,130 runs, **Investment B** in 1,505, **Investment C** in 1,257, and **Investment A** in 108. In both runs, three of the four investments win in non-trivial fractions while **Investment A** is effectively eliminated. The modal winner shifts between **Investment B** and **Investment D** depending on whether valuation inputs are held fixed or also perturbed. The more important feature, however, is shared: once consequence assumptions are allowed to vary, no single investment dominates.

The contrast between the two layers points to a clear analytical interpretation. The recommendation is robust to disagreement about how consequences are valued, but sensitive to disagreement about what those consequences are. The implications for decision-maker discussion are taken up in Section 4.6.1.

4.6. Case interpretation and validation

4.6.1. Interpretation of the case results

The interaction between investment scale, downside exposure, and scenario dependence explains the difference between **Investment D** and **Investment B**. **Investment D** performs well when conditions remain favourable, but its larger scale also makes it more exposed when adverse conditions materialise. **Investment B**, by contrast, provides a smaller but more stable contribution across futures and is less vulnerable to severe regret. The case therefore shows that the framework does not simply reward absolute impact or financial scale, but distinguishes between high upside and robust defensibility under deep uncertainty.

Sensitivity to scenario-specific investment impacts is substantively important in this case, but should not be treated as a surprising methodological result. These impacts form the consequence layer of the case study. If they change, the aggregated value scores and regret results are expected to change as well. In this sense, impact sensitivity is mainly a transparency result. It shows how strongly the

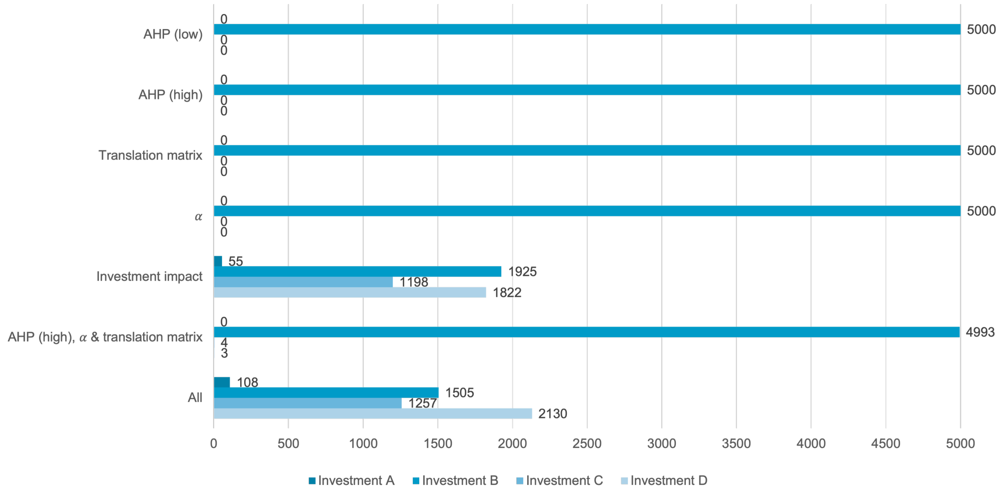
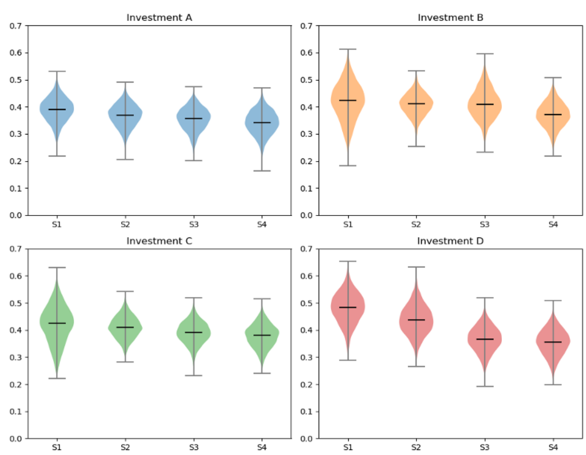
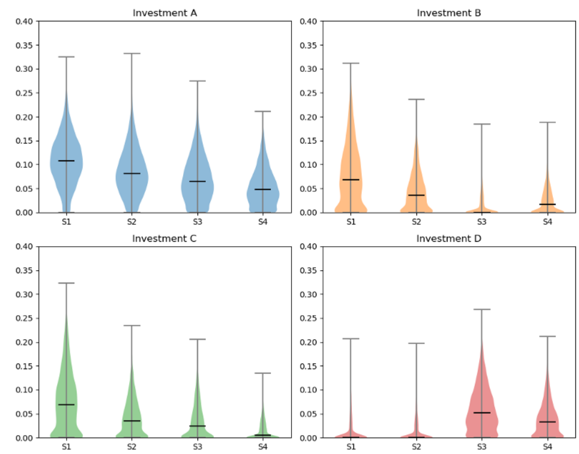


Figure 4.6: Monte Carlo robustness results for the case study, showing how the minimax regret recommendation changes under variation in the tested input families.



(a) Monte Carlo value diagram showing the distribution of aggregated value outcomes under uncertainty in AHP elicitation input, translation matrix, and performance function curve (α).



(b) Monte Carlo Regret diagram showing the distribution of aggregated value outcomes under uncertainty in AHP elicitation input, translation matrix, and performance function curve (α)

Figure 4.7: Monte Carlo diagrams showing the distribution of aggregated value and regret outcomes under input uncertainty.

recommendation depends on the assumed magnitude and direction of the scenario conditioned effects.

Variation in the AHP weights and performance function shape is analytically different. These inputs do not alter the assumed consequences of the investments, but the way those consequences are valued. The case therefore suggests that the conclusion is comparatively robust in the valuation layer, but more sensitive in the consequence layer.

This distinction also clarifies the practical meaning of the results for decision-makers. Strong dependence on impact assumptions suggests that further effort should be directed primarily towards improving the quality, traceability, and verifiability of the consequence estimates. Strong dependence on AHP weights or performance functions, by contrast, would suggest that the main source of instability lies in unresolved differences in managerial preferences rather than in the assumed case mechanics.

Taken together, these results indicate that the framework does not eliminate uncertainty, but helps to locate it. In this case, that is valuable in itself because it shows whether disagreement should be directed primarily at the expected effects of the investments or at the way those effects are valued in the final comparison.

The low score of **Investment A** on *Scope 1 GHG emissions* is partly explained by the scale of the reference target used in the performance function. When a long-term strategic target is used directly as the upper anchor, modest improvements from individual investments may translate into only very small score differences. Future applications could therefore test whether intermediate reference levels improve discriminatory power while remaining consistent with the underlying strategic ambition.

4.6.2. Comparison with current practice

Compared with current practice, the framework adds value most clearly by compelling explicit discussion of assumptions that would otherwise remain implicit. In that sense, it functions as a check on intuition rather than a replacement for it. The framework improves transparency by making trade-offs, scenario assumptions, and judgement based inputs visible and contestable. At the same time, the case also shows that this transparency comes at the cost of complexity. For first time users, the framework is data intensive, and not all assumptions are equally easy to challenge immediately.

In practice, this limitation is partly mitigated by the fact that not all components need to be re elicited in every investment round. In particular, the value structure and performance functions are likely to remain relatively stable, allowing future applications to focus more strongly on scenario dependent impacts and comparative discussion of the investment alternatives. This suggests that the framework is most useful for strategically relevant investment rounds, where the value of structured discussion justifies the time required from senior decision-makers, while very small investments may not warrant the same level of analytical effort. This trade-off should be made deliberately, since selective application of the framework could itself introduce bias.

4.6.3. Verification and validation

Verification and validation address both technical correctness and practical interpretability. Verification in this case consisted of checking completeness of the investment–indicator–scenario matrix, sign consistency between direct and SDR based consequence construction, correct derivation of AHP weights from the aggregated pairwise comparison matrices, boundedness of performance scores, and internal consistency of the MAVT aggregation, regret calculations, and Monte Carlo propagation of perturbed inputs through the full framework.

Validation addressed whether the case-specific causal structure, value structure, and resulting output patterns were recognisable and interpretable to knowledgeable participants. The SDR structure and main feedback pathways were reviewed with experts during elicitation for plausibility and internal consistency. The resulting rankings, trade-offs, and regret patterns were then assessed against yard knowledge and current decision practice. In this sense, validation in the case was not a claim that the framework identifies objectively correct investments, but that it produces outputs that are intelligible, challengeable, and practically usable within the organisational context.

4.6.4. Case-specific limitations

Three limitations are particularly relevant in this case study. First, the high social pressure scenarios rely on wider and less certain impact pathways, which increases the range of plausible effects. Second, the limited availability of experts constrained the depth of elicitation and therefore reduced transparency and reproducibility relative to a more extensive multi round process. Third, the elicitor's relatively limited company specific experience may have reduced understanding of complex investment effects. In combination with the limited elicitation time, this may have skewed parts of the case toward easy to interpret causal effects.

4.7. Case conclusion

This chapter has shown how the framework can be operationalised in a realistic superyacht shipyard setting. The case combines heterogeneous investments, incomplete information, and deep uncertainty regarding the future operating context. The worked example of the electrification of company cars illustrates why even relatively direct sustainability investments require explicit treatment of scenario dependent effects once broader legitimacy, workforce, and financial consequences are taken into account.

Taken together, the case study suggests that the framework provides useful decision support in this setting by producing a transparent comparative ranking by showing the difference in perceived value between heterogeneous investments. The case also shows that the framework does not eliminate judgement, but improves the quality and traceability of that judgement by making key assumptions explicit. Its main practical limitation is analytical complexity, especially during first time application. Even so, the results indicate that the framework is capable of supporting a more structured and contestable investment discussion than current practice alone. These case findings raise three questions. What the divergence between baseline and regret means for the framework, how the results sit against the literature, and what they reveal about the methodological choices behind them. The next chapter takes up each in turn.

5 Discussion

This chapter discusses what the case results mean for the framework developed in this thesis. It interprets the findings, relates them to the literature reviewed in Chapter 2, and reflects on the methodological choices set out in Chapter 3. On that basis, it identifies the main implications of the work and the limitations and boundary conditions that qualify its use. The research questions are answered separately in Chapter 6.

5.1. Interpretation of the case findings

5.1.1. Scenario buy-in as the decisive factor

The framework yields two different recommendations. Under baseline value aggregation, **Investment D**, the dry dock replacement, performs best. Under minimax regret, **Investment B**, cybersecurity training, performs best. The decisive issue is therefore not the numerical output alone, but whether decision-makers regard the adverse scenarios as plausible enough to matter for present choice.

This became clear during elicitation. In one of the sessions, the high social pressure escalation (Scenario 3) was not considered a realistic future. Under that interpretation, the scenario carries little decision weight, and **Investment D** remains preferred because it performs strongly in the base case and compliance-oriented scenario. The other session regarded legitimacy pressure as much more credible. Under that view, the regret profile of **Investment D** becomes material, and a more robust option such as **Investment B** becomes more attractive. The remaining alternatives sit between these extremes. **Investment A** contributes only modestly under every scenario and neither dominates nor collapses, while **Investment C** offers a narrower but more stable profile than **Investment D** without matching **Investment B**'s low maximum regret.

This matters because the framework does not resolve disagreement about the future, nor is it meant to. Its role is narrower and more useful. It makes the disagreement explicit and locates it in a specific analytical layer. In this case, the disagreement sits mainly in scenario plausibility, not in indicator valuation or impact estimation. The recommendation is therefore conditional. The framework narrows the decision to a more precise managerial question. Which futures are plausible enough to shape current investment choice?

5.1.2. Where the sensitivity lies

The Monte Carlo results in Section 4.5.2 support a clear distinction between the valuation layer and the consequence layer of the framework. When the AHP pairwise comparisons and performance function curvature parameters were perturbed within plausible bounds, the minimax regret recommendation remained stable. However, when the scenario-specific impact inputs were perturbed, the recommendation frequently collapsed. In these cases, no single investment emerged as a clear robust winner, indicating that the decision outcome is primarily sensitive to the estimated magnitude of scenario effects.

This pattern is structurally expected. Impact inputs form the immediate consequence layer of the framework. Changes in them therefore propagate directly into aggregated value scores and regret results. Sensitivity to impact inputs should not be read as a flaw in itself. In the present case it is better understood as a transparency result. It shows how strongly the recommendation depends on assumptions about the magnitude and direction of scenario-conditioned effects.

The practical consequence follows directly. If sensitivity concentrates in the consequence layer, analytical effort should focus on improving the quality, traceability, and contestability of impact estimates. If sensitivity were instead concentrated in the valuation layer, the main source of instability would be unresolved differences in managerial preferences. In the present case, the first situation applies. The main fragility lies in the consequences attributed to investments under different futures, not in the value structure used to compare them.

5.1.3. Perceived value versus physical contribution

One finding from the case concerns a subtler risk than initially anticipated. Small numerical changes did not produce disproportionate score shifts. The more relevant issue was different. The perceived value of visible environmental improvements may exceed their actual physical contribution.

Investment A, the electrification of the company car fleet, illustrates this. Its reduction in *Scope 1 GHG emissions* is small relative to yard-wide emissions, and the framework reflects this accordingly. The investment receives only a trivial environmental contribution in the aggregated score (Table A). Within the organisation, however, car fleet electrification tends to be discussed as a meaningful sustainability step. The difference is not extreme in this case, but it points to an important dynamic. Without an explicit reference point, the perceived value of visible sustainability measures can drift beyond their measurable environmental contribution.

This dynamic is not produced by the framework itself, and the framework cannot eliminate it. It can, however, make it visible. By producing transparent indicator-level scores, the framework gives decision-makers a reference point against which internal perceptions can be checked. In that sense, this observation reinforces rather than weakens the rationale for structured evaluation.

Two technical responses appear possible. First, performance function anchors could be calibrated against intermediate reference levels rather than long-term targets, so that small absolute improvements yield smaller score differences. Second, environmental indicators could be scored relative to the most-reducing alternative in the option set, so that each investment is valued against what is actually achievable in the decision round. The second approach would increase data requirements and requires the full option set to be fixed before scoring. Its advantage is that it ties perceived value more closely to comparative physical impact rather than to stated ambition. Both approaches are consistent with the template discipline recommended by Boix-Cots et al. (2022) and merit further testing.

Having interpreted what the case shows on its own terms, the next step is to ask how far it confirms, qualifies, or extends what the literature already claims.

5.2. Findings in light of the literature

5.2.1. Internal value rather than monetary value

Young-Ferris and Roberts (2023) argue that ESG information resists integration into valuation-based decision frameworks because social, governance, and reputational effects cannot be assigned defensible monetary values. The case supports that diagnosis. **Investment B** contributes mainly through risk mitigation and organisational legitimacy. It has almost no cash-flow signature that a valuation approach could capture without highly contestable assumptions.

The framework developed in this thesis responds by replacing monetary value with internal organisational value. Performance functions express how the organisation values different levels of indicator performance, and AHP elicitation structures the trade-offs between them. This is not merely a technical workaround. More fundamentally, it aligns with the underlying argument of Young-Ferris and Roberts (2023). Monetisation imports precisely the bias that value scoring is intended to avoid. Retaining heterogeneity across indicator types, rather than collapsing them into a single financial metric, is therefore not a concession. It is the point of the framework.

The same logic that resists monetisation also bears on how investments are selected, which is where Flyvbjerg's account becomes relevant.

5.2.2. Confirming and extending Flyvbjerg

Flyvbjerg (2021) argues that capital allocation systematically favours investments with articulable upside and visible growth narratives, producing what he describes as a selection dynamic of survival of the unfittest. The case confirms this pattern at the baseline level of analysis. **Investment D** is regarded within the company as one of the strongest options, partly because the larger dock enables construction of larger vessels and thus strengthens EBITDA potential. This assessment reflects current practice and was incorporated into the analysis rather than treated as an error to be corrected.

Under minimax regret, however, **Investment B** performs best. This does not contradict Flyvbjerg (2021). It shows what happens when the same alternatives are evaluated through a decision logic explicitly designed to resist narrow upside orientation. The framework does not overturn the baseline assessment. It exposes the asymmetry between upside-oriented valuation and robustness-oriented evaluation and makes that asymmetry contestable within the same organisational process. Its contribution is therefore not to replace existing capital allocation logic, but to supplement it with a perspective that is structurally less vulnerable to the bias Flyvbjerg describes.

Resisting that bias depends on methods that remain usable in practice. That feasibility concern is the one raised by the deep-uncertainty literature.

5.2.3. Deep uncertainty in practice

Marchau et al. (2019) and Stanton and Roelich (2021) note that DMDU methods are theoretically well developed but often difficult to apply in practice because they assume extensive data availability, long analytical cycles, and repeated stakeholder engagement. The case confirms that these feasibility constraints are real. Elicitation had to be compressed into two ninety-minute sessions, expert availability was limited, and several assumptions could only be checked once within the time available.

At the same time, the case suggests that these constraints are not fatal. Scenario-based comparison without probability assignment produced outputs that current practice, which relies mainly on intuition supported by cost estimates, does not generate. The framework did not eliminate uncertainty, nor was it intended to. It produced a structured comparison of heterogeneous investments and made visible where the result is robust and where it is vulnerable. In that sense, the case provides modest empirical support for the practical feasibility of DMDU methods in executive settings, provided their analytical ambition is calibrated to the time and expert resources actually available.

Feasibility is one constraint on those methods. Another is whether they surface the right pathways at all, which is where explicit causal representation earns its place.

5.2.4. System dynamics as a means of surfacing hidden feedbacks

Sterman (2000) argues that unaided mental models systematically underrepresent delays, accumulations, and feedback structure. Explicit causal representation is therefore not only an aid during elicitation. The elicitation sessions produced three concrete cases in which the SDR surfaced pathways that direct assessment had missed.

The first concerns **Investment B**. Initially, *Session 1* did not identify a legitimacy dimension for cybersecurity training and treated it as a purely governance-oriented risk reduction measure. While reasoning through the consequences of a data leak using the causal structure, the participant recognised it as social pressure waiting to happen (*Session 1*). A leak of data would erode internal legitimacy, increase absence as disgruntled employees stayed away, and trigger regulatory attention. Alternatively, a leak of customer data could increase outside social pressure. The feedback structure was not visible under direct assessment but became immediate once the causal pathway was made explicit.

The second concerns **Investment C**. In *Session 1*, direct assessment dismissed yard security as unnecessary because Dutch trespassing law already excludes activists from the yard. In *Session 2*, causal reasoning surfaced two neglected loops. First, the absence of visible security established a cultural baseline in which minor infractions, such as employees taking small items, were normalised, creating slow drift towards broader governance failure. Second, subcontractors might be using borrowed badges to bring in untrained or undocumented workers, a pattern that could eventually trigger intervention by the Labour inspection. Both pathways reframe yard security from an operational nuisance into a governance investment.

The third concerns **Investment D**. The dry dock replacement was initially motivated by operational need. The existing slipway was ageing, could not handle the heavier vessels being sold, and was rendered unsafe by its stairs (*Session 1*). Reasoning through the scenario structure revealed an unintended positive effect. Heavy insulation and the removal of windows sharply reduced disturbance to local residents. A strategic investment thus also became a legitimacy investment. *Session 1* described this as an increase in awareness of social pressure, a connection that had not previously been made to this investment.

In none of these cases did the SDR generate new empirical information. What it did was make existing organisational knowledge explicit, structured, and contestable by embedding it in a causal representation (Sterman, 2000).

5.3. Methodological reflections

The literature clarifies what the findings mean. It says less about the choices that produced them, which the following reflections on method address directly.

5.3.1. Divergence hidden by aggregation

Aggregating the AHP pairwise comparisons through the geometric mean produced an organisational weight profile that no individual participant actually held. Each of the four participants had a different dominant indicator: *safety and incident rate*, *EBITDA*, *corruption and bribery*, and *technology maturity*. The aggregated profile balances these positions, but it also conceals their divergence.

This is not a methodological error. Geometric mean aggregation is defensible and consistent with established AHP practice (Saaty, 2008). It does, however, compress genuine disagreement into a composite that does not directly represent any single participant. Inspection of the individual results (Appendix A) showed that three of the four participants would have selected the same investment as the aggregated result under baseline aggregation, and three of the four under minimax regret. The aggregated result therefore aligned with a majority on both criteria. More importantly, the shift from baseline to minimax regret was not an artefact of aggregation. Two of the four participants individually moved from a different baseline winner to the aggregated minimax-regret choice once regret was considered. A different aggregation rule, such as role-weighted or subset-based aggregation, could nevertheless have produced a different headline result. None of these rules is methodologically wrong, but each implies a different view of whose judgement should count. The implications of this choice for repeated use are discussed in Section 5.4.2.

5.3.2. Elicitation depth, bias, and time

A practical challenge in this research was finding an acceptable balance between elicitation depth, bias management, and limited time. Greater depth can improve nuance and reduce the risk that relevant considerations remain implicit, but it also increases session burden and may introduce fatigue or inconsistency. This trade-off mattered especially because the framework was operationalised by a company outsider. Beyond the initial case setup, the elicitation sessions were the only moments in which expert input could be incorporated directly into the framework application.

As a result, choices about indicator selection, interpretation of effects, and other framework settings had to be checked within a narrow time window. The main compensating strategy was to document assumptions and choices as explicitly as possible, so that they remained transparent and open to later review. This mitigates, but does not remove, the risk that some elicited inputs may reflect time pressure rather than considered judgement.

5.3.3. The role of a preselection session

For further development of the framework, a preselection session would likely improve the quality of the main elicitation. Such a session would serve three purposes. First, it would allow discussion of the scenario axes before participants are asked to assess impacts. Second, it would explain the methodological logic of the framework, including the role of AHP, the performance function templates, and the interpretation of minimax regret. Third, it would allow the SDR to be reviewed and revised against participants' own understanding of the yard.

The purpose is not to pre-empt the main elicitation. It is to allow participants to arrive with considered rather than improvised positions. Directors would have time to reflect on what they value, what they regard as irrelevant, and which causal relations they recognise. The subsequent discussion of scenario-specific impacts could then be substantive rather than mainly orienting. The trade-off is obvious. It demands more time from senior participants. That cost may be justified for strategically significant investment rounds, but not for smaller allocations.

5.4. Implications

5.4.1. Implications for theory

The main theoretical implication of this work is that sustainability investment evaluation is primarily a decision-logic problem rather than a measurement problem. The persistent gap between ESG reporting and capital allocation is not mainly caused by poor data quality or insufficient indicator coverage. It stems from a structural incompatibility between backward-looking, externally accountable ESG logic and forward-looking, internally deliberative investment logic.

The case suggests that this gap can be narrowed by a framework deliberately designed for the latter context. The combination of non-probabilistic scenario analysis, internal value rather than monetary value, and explicit rather than implicit valuation structures allows heterogeneous sustainability investments to enter the same decision process without forcing premature monetisation. In that sense, the contribution is not simply another ESG measurement architecture. It is an argument that the central problem lies in how sustainability information is translated into organisational choice.

5.4.2. Implications for practice

For practice, the framework appears best suited to strategically significant capital allocation rounds in which the analytical effort is justified by the value of structured deliberation. For very small investments, the time required from senior decision-makers may outweigh the benefit of formal comparison, and intuitive prioritisation may remain acceptable.

The case also indicates that the framework's components differ in necessity. The AHP value structure, performance functions, scenario set, and structured elicitation are necessary, since without them the framework loses its valuation basis, its uncertainty treatment, or its organisational grounding. The SDR sits in a supporting role rather than a structural one. It is not strictly necessary if consequence inputs can be determined through other means, but in the present case it surfaced indirect pathways that direct assessment had missed. MAVT aggregation and minimax regret are stable analytical operations once the inputs are set, and Monte Carlo robustness analysis is most valuable at framework setup or after substantial change.

A further practical implication is that the main product of the framework is not only the ranking it produces. The discussion required to generate the inputs, and the process that surfaces implicit assumptions about scenarios, causal structure, and organisational value, are at least as important as the final score. The framework should therefore not be treated as a verdict machine. Its primary practical value lies in disciplining deliberation.

Translated into the case yard's next investment round, the framework can build directly on the existing setup and fit into the annual capital allocation cycle. The valuation layer elicited in this thesis can be reused, since organisational preferences are unlikely to have shifted materially. The main effort should sit with the consequence layer, where the analysis showed real sensitivity and where the inputs feed directly into the framework's output. A broader candidate set of around ten investments would test the framework's discriminating power at realistic scale. To strengthen the deliberative quality observed in this thesis, both board-level and yard-level directors should be present in the same room, together with the sustainability function, provided the board values the inclusion of these views. Two half-day sessions are recommended. The first covers scenario buy-in and impact discussion, the second covers results discussion. The second session is equally important, since its purpose is to challenge the resulting recommendation rather than to ratify it. Framework outputs should not be presented as the decision itself. A separate meeting should determine which conclusions are carried into the final investment choices.

Three operational requirements support repeated use. First, candidate investments should enter the comparison through board approval, preferably as the complete list of proposed investments rather than a pre-selected subset. Second, framework continuity requires a designated custodian who inherits the elicitor's role. This person would maintain the elicited weights, performance functions, and the SDR, and apply lightweight updates as new investments enter. Third, the choice of AHP aggregation rule embeds a view of whose judgement should count and in what proportion, and a different rule could produce a different headline result. The case company should therefore set this choice deliberately rather than adopt it by default, and framework outputs should be read as conditional on that choice.

A separate operational point concerns the selection rule itself. Because capital allocation typically funds multiple investments rather than a single winner, the minimax regret rule will need adaptation in practice. Either investment candidates are partitioned into categories with one robust selection per category or the top investments by lowest maximum regret are reported as a recommended bundle.

5.4.3. Meeting the framework requirements

Section 1.3 set out nine requirements that a yard-level sustainability investment framework must satisfy. The case application provides the evidence needed to judge how far the framework meets them. Table 5.1 maps each requirement to the component that addresses it and to the supporting case evidence. Seven are met, through a combination of design and demonstrated case performance. Two are only partially met, because the framework was applied once to a subset of investments rather than to a full capital allocation round. Feasibility for repeated use was shown for a single application, not across successive rounds. Scalability has been argued from the reduction logic, but not tested beyond the four investments and thirteen criteria of the case. Both are revisited in Section 5.5.

Table 5.1: Treatment of the Section 1.3 requirements, with case evidence.

Requirement	Treatment	Status	Evidence
Reduce intuition reliance without removing discretion	AHP weights and performance functions	Met	Sec. 4.6.2
Compare heterogeneous investments at yard-level	Performance functions and MAVT	Met	Sec. 4.5
Accommodate uncertainty in costs, impacts, and context	Scenario analysis	Met	Sec. 4.2.3
Tolerate incomplete information and imprecise causality	SDR and ordinal proxies	Met	Sec. 4.3.1
Make assumptions explicit and adjustable	Documented elicited inputs	Met	Sec. 4.3
Remain feasible for repeated use	Template-based elicitation	Partially met	Sec. 4.2.1
Scale to more investments and indicators	Indicator basis and round-specific reduction	Partially met	Sec. 3.3
Reduce favouring of easily quantifiable cases	Minimax regret and non-monetised scoring	Met	Sec. 5.1.3
Make uncertainty and assumption effects visible	Monte Carlo perturbation	Met	Sec. 4.5.2

5.5. Limitations and boundary conditions

The framework, as distinct from the specific case application discussed in Section 4.6.4, has several limitations and boundary conditions that qualify its use.

First, it has been applied to a single case. Transferability beyond the custom superyacht shipyard has not been empirically tested. The framework leans on characteristics shared across ETO production more broadly. These are heterogeneous investment portfolios, fragmented data environments, long project horizons, and exposure to external legitimacy pressure (Heimo et al., 2024; Neves et al., 2025). Its value is highest where deep regulatory uncertainty and legitimacy risk coexist. ETO firms in stable regulatory environments, or without meaningful legitimacy exposure, would likely find the SDR and minimax regret logic disproportionate to the decision at hand, and a lighter MAVT or AHP application would suffice. This expectation remains provisional until tested in further cases.

Second, the additive MAVT aggregation assumes preferential independence between criteria. In the present context that assumption is only approximate. Two concerns emerged in the case. The indicators *Implementation complexity* and *Organisational effort* overlapped sufficiently that one was removed at framework level. More fundamentally, *CAPEX*, *EBITDA*, and *Financial risk* are not parallel goods. *CAPEX* is an input cost incurred to generate *EBITDA*, and higher *Financial risk* is often accepted in expectation of higher *EBITDA*. Additive aggregation treats these as separable and compensable, which may produce internally inconsistent valuations when participants implicitly condition one criterion on another. The limitations of additive value functions in such settings are discussed by Dyer (2016) and Gan et al. (2017).

Third, the scenarios are static. They represent alternative external futures, but do not model adaptive investment pathways across successive allocation rounds. Path dependence between rounds, sequencing effects, and the possibility that organisational preferences evolve in response to outcomes are therefore not captured.

Fourth, the five-level ordinal scale used for qualitative indicators deliberately privileges interpretability over resolution. This makes the framework easier to use in executive settings, but it also coarsens outcomes that are in practice more continuous. Small but decision-relevant differences may therefore disappear in the translation to the common scale.

Fifth, the framework is sensitive to the skill of the elicitor. Its outputs depend on the elicitor's ability to surface implicit assumptions without imposing them. Too little guidance produces superficial discussion and missed feedbacks. Too much guidance makes the output partly a reflection of the elicitor's own framing. This is not a minor operational issue. It is a structural boundary condition for transfer to organisations without an embedded or methodologically experienced facilitator.

Sixth, the Monte Carlo robustness analysis tests stability under elicitation noise, not under fundamental disagreement about the value structure. The Saaty-scale truncation and adjacent-template restrictions prevent extreme reweighting. A decision-maker holding a radically different value structure than the aggregated elicited profile is therefore outside the tested perturbation space. The stability result should be read as robustness under modest preference variation, not as robustness under arbitrary preference disagreement.

Seventh, the framework handles non-linear reversals only through manual scenario-specific input. In the case application, Investment D's operational capacity flips from a strategic asset to a potential liability under scenario 3. While this is captured in the impact estimates, the additive structure does not inherently model such threshold effects. The framework remains a static evaluation tool rather than a dynamic real-options model.

Eighth, the framework was applied once, to four investments and thirteen weighted criteria, rather than to a full capital allocation round repeated across cycles. The number of AHP pairwise comparisons grows quadratically with the criterion count, and performance-function and impact elicitation scale with the size of the active set, so the elicitation burden rises as the comparison widens. The round-specific reduction in Chapter 3 is designed to contain this. Even so, both scalability to a larger candidate set and feasibility across successive rounds remain argued from the design rather than demonstrated. A live round of around ten investments, repeated in a later cycle, would be needed to test both.

Taken together, these limitations do not invalidate the framework, but they define the conditions under which its conclusions should be read. They also point to the directions where further development is most needed, which the next section develops.

5.6. Future research

The limitations identified above, together with the broader interpretation of the case findings, point to several directions for further development. The most immediate next step is application in a live capital allocation round with a larger set of competing investments. This would test whether the framework scales beyond the limited set used in the case, whether the elicitation process remains feasible under actual organisational time pressure, and whether the deliberative effects observed in this thesis translate into measurable changes in allocation outcomes.

A second priority is methodological refinement of the elicitation process. The case suggests that a preselection session could improve the quality of the main workshops by allowing participants to reflect in advance on scenario plausibility, causal structure, and value trade-offs. This should be tested directly. A comparison between applications with and without such a step would show whether better preparation materially improves consistency and depth, or whether the observed compression is mainly a structural consequence of limited executive attention.

A third priority concerns calibration. The case indicated that visible environmental measures can acquire more perceived importance than their physical effect warrants. Future work should therefore test alternative anchoring strategies for performance functions, including intermediate reference levels and comparative scoring against the most reducing alternative in the option set. This would clarify whether environmental value scores can be made more closely proportional to actual contribution without reducing interpretability.

A fourth direction is extension towards adaptive decision making over time. The framework developed here compares static alternatives under static scenarios. It does not capture how investment sequences might be adjusted across successive allocation rounds as external conditions change. Extending the framework towards adaptive pathways would make it possible to examine not only which investment is preferable now, but also when later investments become attractive under changing conditions.

Fifth, the application of minimax regret logic to multi-investment portfolios requires further investigation. Standard capital allocation involves selecting a bundle of investments rather than a single winner. Testing whether category-based selection or "Top-N" robust variants of the regret rule provide useful selection heuristics would bridge the gap between individual project evaluation and annual portfolio construction.

Finally, cross-case application is needed. Testing the framework in other custom superyacht shipyards would show which elements are genuinely sector specific and which generalise. Application in other capital-intensive, project-based sectors facing uncertainty about regulation and legitimacy would go further still. Such work is necessary before broader claims about transferability can be made with confidence.

6
Conclusion

This thesis examined how sustainability investments can be compared at yard-level in a decision context marked by heterogeneous effects, incomplete data, and deep uncertainty. The conclusion draws together the main findings. It answers the research questions, states the contribution of the thesis, and outlines the most important directions for further research. Detailed interpretation, implications, and limitations are discussed in Chapter 5.

6.1. Answers to the sub-questions

SQ1: Relevant indicators and gaps relative to existing assessment approaches

Relevant indicators for yard-level sustainability investment evaluation span environmental, social, governance, financial, and feasibility dimensions. No single existing framework provides an investment-ready set for this purpose. The appropriate basis is therefore a broad indicator universe, grounded in reporting logic (European Commission, 2023) and literature (Hojnik et al., 2020). This universe is then reduced to an active set for the specific decision context. The gap relative to existing approaches is not primarily one of measurement coverage. Reporting oriented approaches identify material topics but do not show how discrete investments should be compared under scarcity. More data-intensive approaches such as lifecycle assessment and valuation-based methods require levels of causal detail and measurement that are difficult to justify for early-stage, heterogeneous yard-level investments. The central gap is therefore the absence of a feasible way to compare investment alternatives under uncertainty, not a shortage of measurement categories. In this thesis, the relevant indicator set was not treated as fixed or universal. It was treated as something to be derived systematically and reduced transparently for each application.

SQ2: Representing expected effects under uncertain causal relationships and limited data

Expected effects can be represented through a combination of direct assessment and System Dynamics Representation. Direct assessment is suitable where an investment affects an indicator through a short and interpretable pathway. The SDR is suitable where effects depend on indirect pathways, feedbacks, delays, or accumulations that are difficult to assess reliably through unaided judgement alone (Sterman, 2000). Together, these two routes make it possible to represent expected consequences without forcing a level of precision that the available data and causal knowledge cannot support.

SQ3: Transparent and robust comparison of heterogeneous considerations

Heterogeneous ESG and financial considerations can be compared transparently by converting them into a common evaluative space without monetising them. In this thesis, performance functions translated indicator outcomes into value scores, AHP elicitation made the relative weights between criteria explicit (Saaty, 2008), and additive MAVT aggregated the results into an overall assessment (Dyer, 2016). This allowed environmental, social, governance, financial, and feasibility effects to remain analytically distinct while still being compared within one decision structure. Transparency followed because the value judgements were explicit, contestable, and revisable.

SQ4: Accounting for uncertainty about future regulatory, social, and market developments

Uncertainty about future conditions can be incorporated through a bounded scenario set and a non-probabilistic robustness rule (Marchau et al., 2019; Walker et al., 2013). In this thesis, scenarios represented plausible external futures rather than forecasts with assigned likelihoods, and minimax regret was used to compare investments across them (Hallegatte et al., 2012; Stanton & Roelich, 2021). This made it possible to distinguish options that perform well in one expected future from options that remain defensible across several plausible futures. Framework robustness was then assessed separately through Monte Carlo perturbation of elicited inputs. This separation is important, because uncertainty about future conditions and uncertainty about framework inputs are analytically different problems.

SQ5: Framework component contributions to making the three influences explicit

Each framework component contributes to making one or more of the three influences explicit and contestable. Together, the indicator selection and reduction process, performance functions, AHP value structure, and structured elicitation make ESG expectations explicit. They force the organisation to specify which dimensions matter, how they are valued, and how trade-offs are weighted. The scenario set and the minimax regret rule together make uncertainty about future operating conditions explicit, structuring plausible alternative futures and selecting investments on the basis of robustness rather than upside under any single future. Together, the minimax regret rule, the SDR, and the explicit value structure counter organisational bias in capital allocation. They resist upside-oriented selection, surface indirect pathways that direct assessment had overlooked, and make implicit preferences contestable. Monte Carlo robustness analysis sits across all three by showing how sensitive the recommendation is to the elicited inputs that operationalise them. The framework's contribution therefore lies in distributing these three influences across distinct analytical layers rather than collapsing them into a single judgement step.

SQ6: Insights from application in a superyacht shipyard

Application in the case study showed that the framework can produce a transparent comparison of heterogeneous investments in a realistic executive setting. More importantly, it surfaced insights that current practice did not make explicit. Scenario plausibility emerged as a decisive factor in the final recommendation. The SDR surfaced legitimacy and governance pathways that direct expert assessment had not made explicit. The case also showed where the main sensitivity of the framework lies. In this application, the decisive uncertainty was concentrated in the consequence layer rather than in the valuation layer. At the same time, the case made clear that the framework requires time, preparation, and disciplined elicitation. It is therefore most suitable for strategically significant investment rounds rather than routine small-scale decisions.

6.2. Answer to the main research question

Three influences shape investment priorities at the yard in distinct ways. They are ESG performance expectations, uncertainty about future operating conditions, and organisational biases. ESG performance expectations expand the criterion set beyond financial considerations, requiring environmental, social, and governance effects to enter the comparison alongside cost and capacity. Uncertainty about future operating conditions makes scenario plausibility the decisive factor. The same case data produced different recommendations depending on whether adverse scenarios were treated as realistic. The dry dock replacement was preferred under stable futures, whereas robustness-oriented evaluation favoured cybersecurity training. Organisational bias in capital allocation favours investments with articulable upside and quantifiable returns (Flyvbjerg, 2021). As a result, visible strategic investments are systematically advantaged over robust but less narrative-friendly alternatives.

The framework developed in this thesis does not resolve these three influences, but makes their effects on investment priorities visible and contestable. ESG expectations are made explicit through the active indicator set, performance functions, and AHP elicitation. Uncertainty about future operating conditions is structured through bounded scenarios and minimax regret rather than through probabilistic forecasting. Organisational bias is countered by separating value aggregation from robust selection, so that upside-oriented preferences cannot dominate the recommendation without being challenged by regret across adverse futures. The case study shows that this approach is feasible in a superyacht shipyard context. The framework therefore offers a way to make these three influences traceable in actual investment decisions rather than allowing them to operate silently. This holds provided the framework is used as a tool for disciplined deliberation rather than as an automatic decision rule.

6.3. Contribution

The main contribution of this thesis is conceptual as well as practical. Conceptually, it shows that yard-level sustainability investment evaluation should be treated primarily as a decision-logic problem rather than a measurement problem. Specifically, it identifies how three influences shape investment priorities and shows how their effects can be made explicit through a structured decision process. These influences are ESG expectations, uncertainty about future operating conditions, and organisational bias in capital allocation. The central difficulty is not that organisations lack categories, indicators, or reporting frameworks. It is that these do not by themselves provide a credible way to compare discrete investment alternatives under deep uncertainty. By shifting the focus from measurement completeness to decision structure, the thesis offers a clearer basis for understanding why ESG information often remains weakly connected to internal capital allocation.

Practically, the thesis develops and applies a framework that combines explicit value structuring, internal value scoring, scenario-based comparison, and robustness analysis in a form that can be used in an executive shipyard setting. This is the central practical contribution. The framework does not attempt to monetise all relevant effects, nor does it rely on probability assignments that the case cannot justify. Instead, it retains heterogeneity, makes trade-offs explicit, and allows decision-makers to see how recommendations change across plausible futures. The case study shows that this approach is not only methodologically coherent, but also operationally usable within the limits of a real organisational setting.

A further contribution lies in the role of the System Dynamics Representation. In this thesis, the SDR did not serve as a predictive model. Its value was analytical. It helped surface indirect pathways, hidden feedbacks, and delayed consequences that direct assessment alone did not reliably capture. This mattered especially for legitimacy- and governance-related effects, which proved difficult to articulate until they were embedded in an explicit causal structure.

These contributions should be read within the limits discussed in Chapter 5, especially the single case setting, the approximate nature of additive aggregation, the static scenario structure, and the dependence of the framework on elicitation quality. These limits qualify the scope of the claims, but they do not remove the central finding that a structured, non-probabilistic, yard-level sustainability investment framework is both possible and useful.

References

Andrikopoulos, A. (2026). ESG in the shipping industry: A literature review. *Corporate Governance: The International Journal of Business in Society*, *26*(1), 199–232. https://doi.org/10.1108/CG-05-2024-0264

Azouzi, M. A., & Anis, J. (2012). CEO emotional bias and investment decision, bayesian network method. *Management Science Letters*, *2*(4), 1259–1278. https://doi.org/10.5267/j.msl.2012.02.012

Behzadian, M., Kazemzadeh, R., Albadvi, A., & Aghdasi, M. (2010). PROMETHEE: A comprehensive literature review on methodologies and applications. *European Journal of Operational Research*, *200*(1), 198–215. https://doi.org/10.1016/j.ejor.2009.01.021

Boix-Cots, D., Pardo-Bosch, F., Blanco, A., Aguado, A., & Pujadas, P. (2022). A systematic review on MIVES: A sustainability-oriented multi-criteria decision-making method. *Building and Environment*, *223*, 109515. https://doi.org/10.1016/j.buildenv.2022.109515

Boscawen, G. (2025, January 15). *New svalbard environmental regulations introduced: What does it mean for yachts?* [Boat international]. Retrieved November 27, 2025, from https://www.boatinternational.com/destinations/new-svalbard-environmental-regulations

Chin-A-Fo, H., & Kuijpers, K. (2023). Superyacht builders close their eyes to illegal teak trade. *NRC*. Retrieved January 22, 2026, from https://www.nrc.nl/nieuws/2023/03/03/superyacht-builders-look-away-at-illegal-teak-trade-a4158553

Cozijnsen, E. J. (2019, December 17). *The footprint of yacht production* [Master Thesis]. Delft University of Technology.

Crace, L., & Gehman, J. (2023). What really explains ESG performance? disentangling the asymmetrical drivers of the triple bottom line. *Organization & Environment*, *36*(1), 150–178. https://doi.org/10.1177/10860266221079408

Dall-Orsoletta, A., Verrier, B., Uriona-Maldonado, M., Dranka, G. G., & Ferreira, P. (2025). How does social acceptance affect transition minerals production in europe? a system dynamics approach and case study in portugal. *The Extractive Industries and Society*, *22*, 101625. https://doi.org/10.1016/j.exis.2025.101625

Diesveld, B. H. M. (2019). *The fuel cell powered superyacht* [Master Thesis]. Delft University of Technology.

Dyer, J. S. (2016). Multiattribute utility theory (MAUT). In S. Greco, M. Ehrgott, & J. R. Figueira (Eds.), *Multiple criteria decision analysis* (pp. 285–314, Vol. 233). Springer New York. https://doi.org/10.1007/978-1-4939-3094-4_8

Ettema, B. (2021, October 18). *Circular yacht production: Mapping environmental impacts of business processes in yacht building and indicating innovation potentials* [Master Thesis]. Delft University of Technology.

Commission Delegated Regulation (EU) 2023/2772 of 31 July 2023 supplementing Directive 2013/34/EU of the European Parliament and of the Council as regards sustainability reporting standards (2023, July 31). Retrieved February 19, 2026, from http://data.europa.eu/eli/reg_del/2023/2772/oj

Directive (EU) 2022/2464 of the European Parliament and of the Council of 14 December 2022 amending Regulation (EU) No 537/2014, Directive 2004/109/EC, Directive 2006/43/EC and Directive 2013/34/EU, as regards corporate sustainability reporting (2022, December 14). Retrieved May 22, 2026, from http://data.europa.eu/eli/dir/2022/2464/oj

Directive (EU) 2023/959 of the European Parliament and of the Council of 10 May 2023 amending Directive 2003/87/EC establishing a system for greenhouse gas emission allowance trading within the Union and Decision (EU) 2015/1814 (2023, May 10). Retrieved May 22, 2026, from http://data.europa.eu/eli/dir/2023/959/oj

Favi, C., Campi, F., Germani, M., & Manieri, S. (2018). Using design information to create a data framework and tool for life cycle analysis of complex maritime vessels. *Journal of Cleaner Production*, *192*, 887–905. https://doi.org/10.1016/j.jclepro.2018.04.263

Feadship. (2023, July 31). *(copyright) obsidian: Feadship's first delivery of 2023 raises the bar on carbon...* Retrieved January 20, 2026, from https://www.feadship.nl/pressroom/obsidian-feadships-first-delivery-of-2023-raises-the-bar-on-carbon-reduction

Fitriadi, F., & Mohamad Ayob, A. F. (2023). Continuous improvement strategy in traditional shipyard industry: A holistic approach towards sustainability. *Journal of Optimization in Industrial Engineering*. https://doi.org/10.22094/joie.2023.1982419.2052

Flyvbjerg, B. (2006). From nobel prize to project management: Getting risks right. *Project Management Journal*, *37*(3), 5–15. https://doi.org/10.1177/875697280603700302

Flyvbjerg, B. (2021). Top ten behavioral biases in project management: An overview. *Project Management Journal*, *52*(6), 531–546. https://doi.org/10.1177/87569728211049046

Flyvbjerg, B., & Bester, D. W. (2021). The cost-benefit fallacy: Why cost-benefit analysis is broken and how to fix it. *Journal of Benefit-Cost Analysis*, *12*(3), 395–419. https://doi.org/10.1017/bca.2021.9

Gan, X., Fernandez, I. C., Guo, J., Wilson, M., Zhao, Y., Zhou, B., & Wu, J. (2017). When to use what: Methods for weighting and aggregating sustainability indicators. *Ecological Indicators*, *81*, 491–502. https://doi.org/10.1016/j.ecolind.2017.05.068

Gillan, S. L., Koch, A., & Starks, L. T. (2021). Firms and social responsibility: A review of ESG and CSR research in corporate finance. *Journal of Corporate Finance*, *66*, 101889. https://doi.org/10.1016/j.jcorpfin.2021.101889

Hallegatte, S., Shah, A., Lempert, R. J., Brown, C., & Gill, S. (2012, September). *Investment decision making under deep uncertainty - application to climate change*. World Bank, Washington, DC. https://doi.org/10.1596/1813-9450-6193

Heimo, O., Vainio-kaila, T., Kinnunen, K., Hänninen, S., Helle, S., Majaniemi, S., Jokinen, L., & Lehtonen, T. (2024). The flow of sustainability information through interorganisational shipbuilding ecosystem. *146*. https://doi.org/10.54941/ahfe1005156

Hirtenlehner, J. J. (2025, January 3). *Rocking the boat: Alt-fuels and the super yacht industry* [Master Thesis]. Universidade Católica Portuguesa.

Hojnik, J., Biloslavo, R., Cicero, L., & Cagnina, M. R. (2020). Sustainability indicators for the yachting industry: Empirical conceptualization. *Journal of Cleaner Production*, *249*, 119368. https://doi.org/10.1016/j.jclepro.2019.119368

Huang, I. B., Keisler, J., & Linkov, I. (2011). Multi-criteria decision analysis in environmental sciences: Ten years of applications and trends. *Science of The Total Environment*, *409*(19), 3578–3594. https://doi.org/10.1016/j.scitotenv.2011.06.022

Ishizaka, A. (2014). Comparison of fuzzy logic, AHP, FAHP and hybrid fuzzy AHP for new supplier selection and its performance analysis. *International Journal of Integrated Supply Management*, *9*(1), 1. https://doi.org/10.1504/IJISM.2014.064353

Janson, D. I. (2016, November). *The development of a green shipyard concept* [Master Thesis]. Delft University of Technology.

Kamerman, S., & Verseput, S. (2022). Een spoeddebat en rotte eieren: Bezos' jacht veroorzaakt onrust. *NRC*. Retrieved November 27, 2025, from https://www.nrc.nl/nieuws/2022/02/04/een-spoeddebat-en-rotte-eieren-bezos-jacht-veroorzaakt-onrust-a4084424

Keeney, R. L. (1988). Building models of values. *European Journal of Operational Research*, *37*(2), 149–157. https://doi.org/10.1016/0377-2217(88)90324-4

Koilo, V. (2021). Evaluation of r&d activities in the maritime industry: Managing sustainability transitions through business model. *Problems and Perspectives in Management*, *19*(3), 230–246. https://doi.org/10.21511/ppm.19(3).2021.20

Koray, M. (2023). Prioritizing shipyard conversion requirements regarding green ship and green shipyard concept. *Transactions on Maritime Science*, *12*(2). https://doi.org/10.7225/toms.v12.n02.w02

Lempert, R. J. (2019). Robust decision making (RDM). In V. A. W. J. Marchau, W. E. Walker, P. J. T. M. Bloemen, & S. W. Popper (Eds.), *Decision making under deep uncertainty* (pp. 23–51). Springer International Publishing. https://doi.org/10.1007/978-3-030-05252-2_2

Letschert, M. E. (2020, March). *A process design towards the yacht environmental transparency index* [Master Thesis]. Delft University of Technology.

Lombardi Netto, A., Salomon, V. A. P., & Ortiz-Barrios, M. A. (2026). An environmental, social, and governance (ESG) model for multiple criteria analysis of investments. *International Transactions in Operational Research*, *33*(1), 489–506. https://doi.org/10.1111/itor.70026

Lynch, M. J., Long, M. A., Stretesky, P. B., & Barrett, K. L. (2019). Measuring the ecological impact of the wealthy: Excessive consumption, ecological disorganization, green crime, and justice. *Social Currents*, *6*(4), 377–395. https://doi.org/10.1177/2329496519847491

Marchau, V. A. W. J., Walker, W. E., Bloemen, P. J. T. M., & Popper, S. W. (Eds.). (2019). *Decision making under deep uncertainty: From theory to practice*. Springer International Publishing. https://doi.org/10.1007/978-3-030-05252-2

Marttunen, M., Lienert, J., & Belton, V. (2017). Structuring problems for multi-criteria decision analysis in practice: A literature review of method combinations. *European Journal of Operational Research*, *263*(1), 1–17. https://doi.org/10.1016/j.ejor.2017.04.041

McCluskey, M. (2023, July 17). *Spanish activists vandalize superyacht in ibiza believed to belong to billionaire walmart heiress* [CNN]. Retrieved November 25, 2025, from https://www.cnn.com/2023/07/17/europe/activists-vandalize-yacht-ibiza-climate-intl

Montibeller, G., & Franco, L. A. (2011). Raising the bar: Strategic multi-criteria decision analysis. *Journal of the Operational Research Society*, *62*(5), 855–867. https://doi.org/10.1057/jors.2009.178

Narci, A. O., & Can Dogan, E. (2025). Sustainable shipbuilding: Enhancing environmental performance through life cycle assessment and process optimization. *Process Safety and Environmental Protection*, *202*, 107712. https://doi.org/10.1016/j.psep.2025.107712

Neves, A., Erikstad, S. O., & Godina, R. (2025). Enhancing efficiency and sustainability in shipbuilding: Insights from lean implementation and challenges. *Journal of Marine Science and Technology*, *30*(3), 575–616. https://doi.org/10.1007/s00773-025-01078-8

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, *15*(3), 234–281. https://doi.org/10.1016/0022-2496(77)90033-5

Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, *1*(1), 83. https://doi.org/10.1504/IJSSCI.2008.017590

Serrano-Cinca, C., & Gutiérrez-Nieto, B. (2013). A decision support system for financial and social investment. *Applied Economics*, *45*(28), 4060–4070. https://doi.org/10.1080/00036846.2012.748180

Smit, H. (2025). Sustainability real options. *California Management Review*, *67*(3), 55–85. https://doi.org/10.1177/00081256251331264

Stanton, M. C. B., & Roelich, K. (2021). Decision making under deep uncertainties: A review of the applicability of methods in practice. *Technological Forecasting and Social Change*, *171*, 120939. https://doi.org/10.1016/j.techfore.2021.120939

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world* (International student edition). Irwin, McGraw-Hill.

Stuart, A., Bond, A., Franco, A. M., Baker, J., Gerrard, C., Danino, V., & Jones, K. (2023). Conceptualising social licence to operate. *Resources Policy*, *85*, 103962. https://doi.org/10.1016/j.resourpol.2023.103962

Taherdoost, H., & Madanchian, M. (2023). Analytic network process (ANP) method: A comprehensive review of applications, advantages, and limitations. *Journal of Data Science and Intelligent Systems*, *1*(1), 12–18. https://doi.org/10.47852/bonviewJDSIS3202885

Tantan, M., & Akdağ, H. C. (2025). An assessment of sustainability indicators for a shipyard, utilizing the analytic hierarchy process (AHP). *Proceedings of the International Conference on Industrial Engineering and Operations Management*. https://doi.org/10.46254/AF06.20250069

Vakili, S., Schönborn, A., & Ölçer, A. I. (2023). The road to zero emission shipbuilding industry: A systematic and transdisciplinary approach to modern multi-energy shipyards. *Energy Conversion and Management: X*, *18*, 100365. https://doi.org/10.1016/j.ecmx.2023.100365

Velasquez, M., & Hester, P. T. (2013). An analysis of multi-criteria decision making methods. *International Journal of Operations Research*, *10*(2), 56–66.

Visser, B. C. W. (2021, August 16). *The plug-in hybrid electric superyacht* [Master Thesis]. Delft University of Technology.

Walker, W. E., Lempert, R. J., & Kwakkel, J. H. (2013). Deep uncertainty. In M. C. Fu & S. I. Gass (Eds.), *Encyclopedia of operations research and management science* (pp. 395–402). Springer US. https://doi.org/10.1007/978-1-4419-1153-7_1140

Young-Ferris, A., & Roberts, J. (2023). 'looking for something that isn't there': A case study of an early attempt at ESG integration in investment decision making. *European Accounting Review*, *32*(3), 717–744. https://doi.org/10.1080/09638180.2021.2000458

Zwaginga, J. (2026). *Enabling ship changeability; a lifecycle approach to the maritime energy transition* [Doctoral Thesis]. Delft University of Technology. https://doi.org/10.4233/UUID:D0BB8816-22D9-459F-816C-C1A434C50297

A

Case study

Appendix A contains confidential, case-specific data from Koninklijke de Vries Scheepsbouw and has been redacted in this public version. The full version is available to the examination committee.

Supporting Sustainability Investment Decisions: Bridging ESG Frameworks and Capital Allocation in Superyacht Shipyards

M.D. de Boer[a] (Master Student), J.F.J. Pruyn[a] (Supervisor)

[a]*Faculty of Mechanical Engineering, Department of Marine Technology, Delft University of Technology , Mekelweg 2, 2628 CD Delft, The Netherlands*

ARTICLE INFO

Keywords:
Sustainability investment
multi-criteria decision analysis
deep uncertainty
ESG
capital allocation
superyacht shipyard

ABSTRACT

The gap between sustainability reporting and capital allocation in capital-intensive, project-based industries is commonly diagnosed as a measurement problem. This paper argues that, for yard-level investment decisions in custom superyacht shipbuilding, it is instead a decision-logic problem. Reporting standards produce backward-looking information for external accountability. Investment choice instead requires forward-looking deliberation under deep uncertainty about regulation, legitimacy, and client expectations. Existing valuation, indicator-based, and multi-criteria approaches each address part of this problem but none is sufficient alone. The paper develops a modular non-probabilistic framework with four components. A European Sustainability Reporting Standards-grounded indicator basis and a System Dynamics Representation handle indirect and feedback-mediated consequences. AHP-weighted Multi-Attribute Value Theory aggregates non-monetised value, and minimax regret across bounded scenarios supports cross-context comparison. Monte Carlo perturbation of elicited inputs tests framework robustness. The framework is applied to four investments at a custom superyacht shipyard, selected to span scale, impact pathway, and scenario sensitivity. The application demonstrates three results. First, baseline-attractive and robustness-attractive investments diverge. A large strategic investment wins under additive aggregation but carries the highest maximum regret, while a small governance investment minimises regret across scenarios. Second, the System Dynamics Representation surfaces legitimacy and governance pathways that direct expert assessment systematically overlooks. Third, sensitivity concentrates in the consequence layer rather than the valuation layer, locating productive disagreement in empirical rather than normative questions. The framework does not eliminate uncertainty. It disciplines deliberation by making assumptions, trade-offs, and points of disagreement explicit and contestable.

1. Introduction

Across capital-intensive, project-based industries, firms face a growing set of sustainability investment proposals that resist comparison through conventional appraisal. Environmental, Social and Governance (ESG) considerations increasingly shape how these firms are financed, regulated, and judged by their stakeholders (Andrikopoulos 2026; Gillan, Koch, and Starks 2021). These proposals are heterogeneous. Their effects span environmental, social, governance, financial, and feasibility dimensions. Their causal pathways are partly indirect. Their future value depends on conditions that cannot be forecast with defensible probabilities.

This pressure is increasingly codified in regulation. In the European Union, the Corporate Sustainability Reporting Directive (CSRD) (European Parliament and Council of the European Union 2022) and the European Sustainability Reporting Standards (ESRS) (European Commission 2023) mandate detailed sustainability disclosure. Sectoral measures such as the extension of the EU Emissions Trading System (EU ETS) to maritime transport (European Parliament and Council of the European Union 2023) translate broad commitments into concrete obligations. These frameworks produce information optimised for backward-looking, externally accountable disclosure. That information does not automatically support forward-looking, internally deliberative investment choice.

ORCID(s): 0000-0002-4496-4544 (J.F.J. Pruyn)

In these industries, the gap between sustainability reporting and capital allocation is fundamentally a decision-logic problem. The mismatch is sharpened by what Walker, Lempert, and Kwakkel (2013) term level-four deep uncertainty, in which decision-makers cannot agree on the underlying causal model, the relevant probability distributions, or the dominant value trade-offs. Expected-value optimisation and monetised valuation are therefore structurally inappropriate (Marchau et al. 2019). Flyvbjerg (2021) additionally documents systematic biases in organisational decision environments that disadvantage investments whose benefits are defensive, indirect, or long-term. This is precisely the profile of most sustainability investments.

These conditions appear in an especially concentrated form in custom superyacht shipbuilding, the setting used here for validation. Sectoral visibility translates regulatory and social pressure directly into yard-level legitimacy risk (Hirtenlehner 2025). This pressure ranges from tightening environmental regulation to high-profile protest actions (McCluskey 2023). In current practice, proposals are decided by senior management on the basis of cost estimates and intuitive judgement, without systematic comparison of sustainability effects across alternatives. This tends to favour investments with short-term, easily quantifiable returns over those whose value is defensive, indirect, or long-term.

This gap is addressed by developing and demonstrating a yard-level decision support framework explicitly designed for the decision-logic problem rather than the measurement problem. The framework combines four components.

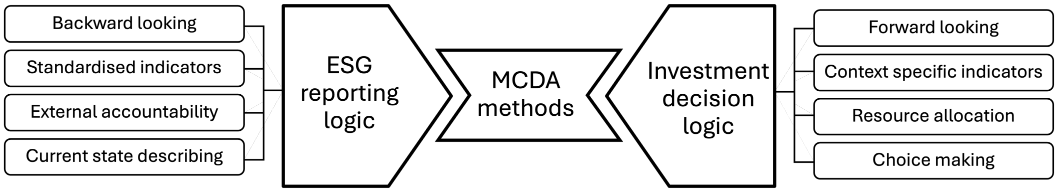


Figure 1: Structural discontinuity between ESG reporting logic and investment decision logic. ESG reporting is backward-looking, relies on standardised indicators, serves external accountability, and describes the current state. Investment decision logic is forward-looking, relies on context-specific indicators, supports resource allocation, and informs choice making. MCDA methods provide a bridge between the two.

An ESRS-grounded indicator basis and a System Dynamics Representation (SDR) construct indirect and feedback-mediated consequences. AHP-weighted Multi-Attribute Value Theory (MAVT) aggregates non-monetised value, and minimax regret across non-probabilistic scenarios supports robust cross-context comparison. Monte Carlo perturbation of elicited inputs then tests framework robustness.

The paper makes three contributions. First, it reframes the ESG–capital-allocation gap, for the class of decision contexts examined here, as a decision-logic problem rather than a measurement problem. Second, it develops a modular non-probabilistic framework to address this gap. Within it, the SDR surfaces indirect, feedback-mediated pathways around legitimacy and governance that unaided expert judgement routinely overlooks. Third, a case application at a Dutch pure-custom superyacht shipyard shows that the framework produces decision-relevant outputs absent from current practice. These include a divergence between baseline-attractive and robustness-attractive investments and a clear separation between the valuation layer, which is stable under perturbation, and the consequence layer, which is decisive.

2. Background and positioning

The literature relevant to yard-level sustainability investment decisions spans three distinct evaluation logics, each structurally mismatched to the decision context. Figure 1 summarises the discontinuity between reporting-oriented ESG logic and forward-looking investment decision logic, and the bridging function of MCDA methods.

2.1. Valuation and prediction.

A first class of approaches conceptualises sustainability investments through valuation and prediction. These typically extend Net Present Value (NPV) analysis or real options reasoning to translate future outcomes into a common monetary metric (Smit 2025). These approaches assume that causal relationships can be modelled and that uncertainty can be represented through defensible probability distributions. For sustainability investments at organisational level, value is generated through indirect, long-term, context-dependent mechanisms such as regulatory resilience, risk mitigation, workforce retention, and reputational effects. Translating these into cash-flow terms requires assumptions the decision context cannot supply. Young-Ferris and Roberts (2023) document the resulting integration failure empirically. ESG information resists incorporation into valuation-based frameworks not because the information is poor, but because monetisation imports the very bias it is meant to resolve, systematically favouring effects that are easiest to price.

2.2. ESG reporting and indicator systems.

A second class of approaches structures sustainability information through ESG indicator systems, increasingly codified in mandatory reporting standards such as the ESRS (European Commission 2023). These frameworks are designed for transparency, accountability, and cross-organisational comparability (Crace and Gehman 2023). They identify and disclose material topics, but they do not specify how the marginal effects of discrete investments should be compared under scarcity. A richer reporting standard does not by itself produce a sharper decision rule.

2.3. Multi-criteria and scenario-based methods.

A third class of approaches addresses trade-offs and uncertainty more directly through Multi-Criteria Decision Analysis (MCDA) and scenario-based methods. These preserve dimensional separation, compare alternatives across heterogeneous objectives, and accommodate plausible future conditions without requiring probabilistic ranking (Marchau et al. 2019). In principle, they fit sustainability investment problems closely. In practice, applications often assume extensive data availability, long analytical cycles, and repeated stakeholder engagement. This limits feasibility in executive settings characterised by time pressure and limited organisational bandwidth (Stanton and Roelich 2021).

2.4. Deep uncertainty as the binding constraint.

The decision context examined here corresponds to the level-four deep uncertainty in the Walker, Lempert, and Kwakkel (2013) classification. Future regulatory trajectories, social legitimacy dynamics, and client expectations are not parameter uncertainties susceptible to better data. They are conditional contexts in which different investments generate different value (Hallegatte et al. 2012). This rules out expected-value optimisation and probability-weighted scenario analysis on principled rather than pragmatic grounds. It reframes proxy indicators and structured qualitative judgement as necessary components of the decision process (Marchau et al. 2019).

2.5. Organisational decision environments.

Even where analytically sound methods exist, internal capital allocation is shaped by bounded rationality, behavioural bias, and organisational feasibility constraints (Stanton and Roelich 2021). Flyvbjerg (2021) documents a selection dynamic, "survival of the unfittest", in which projects with articulable upside and visible growth narratives systematically outcompete projects with defensive, indirect, or long-term benefits. Sustainability investments fall predominantly in the latter category. A framework designed for this context must therefore do more than accommodate qualitative information. It must structurally resist the upside orientation that monetised valuation embeds.

2.6. Method selection as combination of components.

Three constraints jointly imply that no single existing method family is sufficient. Defensible monetisation is unavailable, reporting logic does not support choice, and organisational decision environments bias selection systematically. The decision problem combines four analytically distinct tasks. These are representing causal effects under partial knowledge, structuring trade-offs across heterogeneous criteria, comparing alternatives across plausible futures, and resisting the bias that favours easily articulated upside. Method selection is therefore best treated as the combination of compatible components rather than the adoption of a single complete model. The framework presented in Section 3 reflects this position, with each component selected for its fit with a specific analytical task.

3. Framework

This section specifies the framework in general form, independent of any particular application. The case in Section 4 is one application of it. Each of its four components addresses one of the analytical tasks identified in Section 2, and the components are presented in the order in which they are applied.

3.1. Architecture

The framework follows a fixed analytical sequence with four layers. First, the evaluation basis is defined through an active indicator set, a finite set of admissible investment alternatives, and a bounded scenario set. Second, investment consequences are constructed at indicator level, either through direct assessment or through propagation in an SDR. Third, these consequences are converted into comparable values through elicited performance functions, Analytic Hierarchy Process (AHP)-derived criterion weights, and additive MAVT aggregation. Fourth, alternatives are compared across scenarios through minimax regret, after which framework robustness is evaluated by varying uncertain elicited inputs. Figure 2 shows the full sequence.

The separation of consequence construction, value conversion, and robustness evaluation prevents changes in future context from being confused with changes in organisational preferences.

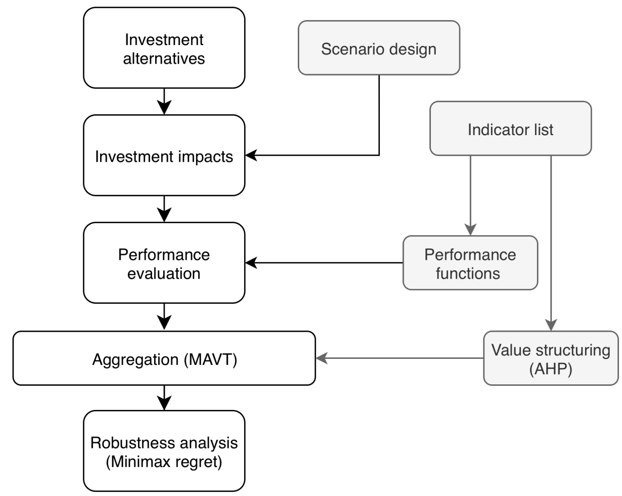


Figure 2: Operational sequence of the framework.

3.2. Indicator basis and reduction

Indicators define the criteria through which investments are evaluated. The framework derives its indicator basis from two strands. ESRS material topics provide a regulatory starting structure (European Commission 2023). The academic literature on yard-level sustainability (Hojnik et al. 2020; Janson 2016; Vakili, Schönborn, and Ölçer 2023) extends and refines the reporting structure for decision-relevant comparison. The result is an indicator universe spanning environmental, social, governance, financial, and feasibility dimensions.

Reduction occurs in two stages. The framework-level stage, performed once, screens the universe for investment relevance and structural decision relevance, then consolidates redundant indicators into proxy indices where this improves interpretability. This yields the active framework set. The round-specific stage, performed for each investment round, sets aside indicators that do not discriminate between the alternatives under consideration, either because they are already at maximum performance, remain unchanged across alternatives, or fall outside the practical decision boundary of that round. Indicators excluded in one round remain in the framework set and may become decision-relevant in another. Reduction is documented at both stages so that exclusions remain auditable.

3.3. Consequence construction

For each combination of investment, indicator, and scenario, a determination is made whether the effect can be assessed directly or must be propagated through an SDR. Direct assessment applies when the effect can be specified without modelling system interaction, for example through measurable data or defensible conditional assumptions. The SDR route applies when the effect depends on indirect pathways, feedback structure, or interactions between organisational variables that unaided judgement systematically underrepresents (Sterman 2000; Dall-Orsoletta et al. 2025).

Effects are represented quantitatively where they can be expressed through measurable data or calculation, and qualitatively otherwise, using a common five-level scale ($--, -, 0, +, ++$). Quantitative baselines, where available, define the reference state. The framework distinguishes direct effects from net system effects. An investment may have limited direct influence on a variable yet generate a meaningful net consequence through the SDR in a given scenario, or vice versa.

The SDR is constructed in three stages. A Causal Loop Diagram (CLD) identifies the variables and feedback structure relevant to the decision context. Each link is annotated with a polarity (whether an increase in the cause raises or lowers the effect), and each closed loop is classified as reinforcing or balancing depending on whether it amplifies or dampens change (Sterman 2000). The diagram is then formalised as a Stock–Flow Representation (SFR) that distinguishes accumulations, rates, and auxiliary variables, ensuring that delays and rate dependencies are made explicit (Sterman 2000). Finally, scenario conditions are translated into changes in exogenous drivers and investment effects are applied as interventions on selected system variables. The propagated state is mapped onto the affected indicators through predefined translation rules.

The representation is used qualitatively under deep uncertainty and is not intended as a calibrated predictive model. Its analytical role is to surface causal pathways, particularly indirect, delayed, or feedback-mediated effects that direct expert assessment routinely misses.

3.4. Value conversion

Performance functions translate indicator outcomes into dimensionless performance scores on the interval $[0, 1]$, where 1 denotes the most preferred outcome (Dyer 2016). Each indicator is assigned one fixed performance function, elicited from a restricted set of templates. These are monotonic power curves, increasing and decreasing S-curves, and peak-shaped curves for indicators with an intermediate preferred level. Restricting elicitation to templates limits complexity, improves consistency across indicators, and supports later revision when additional information becomes available (Boix-Cots et al. 2022). The same template families apply across continuous and ordinal indicators. This preserves comparability and lets quantitative measurements replace ordinal inputs as data quality improves, without re-eliciting the value function.

Criterion weights are derived through the AHP. Each participant completes individual pairwise comparison matrices, which are assessed against the conventional threshold $CR < 0.10$ for logical consistency (Thomas L Saaty 1977). Individual judgements are aggregated at the level of pairwise comparison intensities through the geometric mean, and the aggregated matrix is used to derive local and global criterion weights (Thomas L. Saaty 2008). Weights remain fixed within an evaluation round and across scenario analysis.

Aggregation is performed through additive MAVT (Dyer 2016):

$$V_{is} = \sum_{j=1}^{n} w_j \, p_{ijs}, \tag{1}$$

where V_{is} is the aggregated value of investment i under scenario s, w_j is the AHP weight of criterion j, and p_{ijs} is the performance score of investment i on criterion j under scenario s. The additive form requires mutual preferential independence between criteria and permits compensation between them. Both assumptions are approximate in the present decision context.

3.5. Robustness

Robustness is assessed at two analytically distinct levels. *Investment robustness* evaluates how alternatives compare across scenarios. *Framework robustness* evaluates how the recommendation responds to variation in elicited inputs.

Let I denote the set of investment alternatives and S the set of scenarios. Investment robustness uses Savage's minimax regret criterion (Hallegatte et al. 2012; Stanton and Roelich 2021; Lempert 2019). For each scenario $s \in S$, regret R_{is} is the value gap between the best-performing alternative in that scenario and investment i:

$$R_{is} = \max_{k \in I} V_{ks} - V_{is}. \tag{2}$$

The preferred investment minimises the maximum regret across scenarios, $MR_i = \max_{s \in S} R_{is}$. Minimax regret is preferred over expected-value or expected-utility approaches because it requires no probability assignments over scenarios. It is preferred over maximin because it evaluates exposure relative to the best achievable outcome in each scenario rather than absolute worst-case performance.

Four input families are perturbed. AHP pairwise comparisons are perturbed multiplicatively in log-space and truncated to the Saaty scale [1/9, 9]. Performance function curvature is varied by discrete shifts to adjacent admissible templates. The ordinal-to-baseline translation rule is varied within a defined band. Scenario-specific investment impacts are perturbed proportionally for quantitative inputs and shifted by at most one ordinal step for qualitative inputs. In each run, the full value and regret matrices are recalculated and the minimax regret recommendation is recorded. The distribution of recommendations across runs indicates how strongly the conclusion depends on elicited assumptions.

The four input families fall into two categories. AHP weights and performance function curvature (α) constitute the valuation layer, governing how consequences are valued. Translation rules and scenario-specific impacts constitute the consequence layer, governing what investment consequences are assumed. Sensitivity in the consequence layer is analytically expected, since these inputs propagate directly into all downstream comparisons.

This completes the framework's analytical sequence, from indicator basis through to robustness evaluation. The following section applies it to a real allocation decision and

Table 1
Investment alternatives in the case application.

ID	Investment	Primary value type	Status
A	Car fleet electrification	Emissions reduction	Under discussion
B	Cybersecurity training	Risk mitigation	Implemented
C	Yard security improvements	Continuity protection	Under discussion
D	Dry dock replacement	Capacity and flexibility	Implemented

tests whether it produces decision-relevant outputs under realistic constraints on time, data, and expert availability.

4. Case application

The framework is now demonstrated on a real capital allocation decision rather than illustrated abstractly. A single case cannot establish statistical generalisation. What it can establish is whether the framework is operable with actual decision-makers and what character its outputs take. The section first sets out the decision context and the four investments compared, then the scenario set that bounds future uncertainty, and finally the operational choices made in applying each component. The resulting comparison is reported in Section 5.

4.1. Setting and decision context

The framework was applied at Koninklijke de Vries Scheepsbouw, a Dutch, pure-custom superyacht shipyard operating within a group structure that consolidates capital allocation across multiple yards. Investment proposals originate at yard level, compete in recurring group-level allocation rounds, and are decided by senior management on the basis of cost estimates and intuitive judgement, without systematic comparison of sustainability effects across alternatives. The case examines whether the framework can produce decision-relevant outputs in this setting under realistic time, data, and expert-availability constraints.

Four investments were selected primarily for their heterogeneity, spanning scale, impact pathway, and scenario sensitivity (Table 1). This heterogeneity, rather than the mix of implemented and prospective investments, is what makes the set a meaningful test of the framework. Two had been implemented and two were under consideration. This mix improved data availability without compromising the test of scenario-conditioned reasoning. The future contexts represented in the case have not occurred, so they remain unobserved for implemented and prospective investments alike.

Investment A (car fleet electrification) is a relatively direct environmental measure with mature technology. **Investment B (cybersecurity training)** is a governance measure whose value lies in tail-risk mitigation rather than directly observable performance. **Investment C (yard security improvements)** is the clearest case of strong scenario dependence. **Investment D (dry dock replacement)** is a large strategic investment that enables construction of larger vessels.

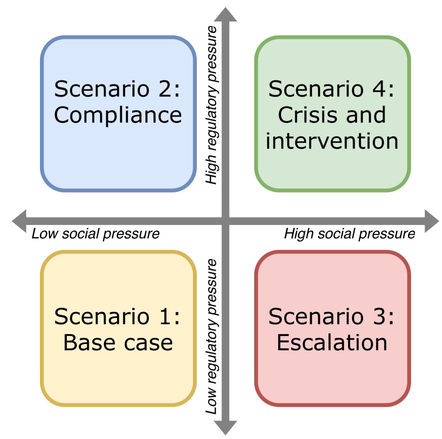


Figure 3: Two-axis scenario set used in the case application.

4.2. Scenarios

Two external uncertainty axes were identified through expert elicitation as the most decision-relevant for the present case: social pressure and regulatory pressure. These mirror the two channels through which sectoral visibility translates into yard-level legitimacy risk, as set out in Section 1. Regulatory pressure is also identified in the literature as a major driver of sustainability-related decisions in the sector (Hirtenlehner 2025). Social pressure captures uncertainty regarding public scrutiny, protest intensity, employer attractiveness, and broader legitimacy dynamics. Regulatory pressure captures uncertainty regarding compliance obligations and the tightening or relaxation of formal institutional requirements. Other axes were considered, including market demand, but participants did not regard demand variation as the most decision-relevant uncertainty. Combining the two axes at their high and low extremes yields four scenarios (Figure 3): *Base case* (low/low), *Compliance environment* (low social, high regulatory), *Escalation* (high social, low regulatory), and *Crisis and intervention* (high/high).

4.3. Operationalisation

Indicator reduction

The framework-level reduction yielded an active set of 20 indicators across environmental, social, governance, financial, and feasibility dimensions. Round-specific reduction for the case set aside seven indicators because no investment in the set produced a discriminating effect on them. Five had no plausible pathway from any alternative (*water consumption, total waste, hazardous waste, supplier ethical standards,* and *human rights incidents*). Two were set aside for other reasons, namely *workforce flexibility*, which no alternative affected, and *organisational effort*, which overlapped substantially with implementation complexity. One indicator (*corruption/bribery*) was initially screened out but reintroduced during elicitation, when participants identified a plausible governance pathway for **investments B (cybersecurity training)** and **C (yard security)**. This left a discriminating set of 13 indicators, which form the criteria weighted in Figure 4. The reintroduction illustrates that round-specific reduction remains revisable when credible impact pathways become visible during application.

System Dynamics Representation

A CLD and SFR were constructed for the case, linking four organisational domains: social legitimacy, workforce dynamics, operational performance, and financial performance (Dall-Orsoletta et al. 2025). The dominant mechanism in the superyacht context is moral legitimacy pressure (Stuart et al. 2023). Social dissatisfaction can accumulate and, when amplified by media attention, trigger visible disruption and operational pressure. Environmental performance enters as an upstream driver of legitimacy pressure rather than as an independent outcome stock. The model centres on four key stocks, namely social acceptance, workforce availability, order book, and financial capacity. It is used to derive scenario-specific baseline states and to propagate investment effects through the system. Construction was reviewed with experts during elicitation for causal plausibility and recognisability rather than for behavioural calibration against historical data.

Elicitation

Elicitation was conducted in two ninety-minute sessions with four participants spanning three perspectives: a member of the Group Board of Directors (the level deciding investments), two yard-level directors (the level proposing investments), and the Group Sustainability Director (the function steering toward sustainable outcomes). Each session covered framework introduction, indicator reduction confirmation, AHP pairwise comparisons via a structured questionnaire, performance function elicitation through templates and sketching, and structured discussion of scenario-conditioned investment effects. Compressing elicitation into single sessions per group reduced depth relative to a multi-round process, but allowed the framework to be tested with the actual decision-makers rather than with proxies. Both sessions were recorded for traceability.

AHP consistency and aggregation

Individual AHP matrices frequently exceeded the conventional threshold of $CR < 0.10$ on dimension-level comparisons. This reflects the genuine difficulty of making strictly transitive judgements across heterogeneous criteria under time-constrained elicitation. Geometric-mean aggregation across the four participants produced consistent matrices throughout, with $\max(CR) = 0.039$. The aggregated weight profile, shown in Figure 4, distributes weight across criteria with two notable concentrations: safety/incident rate

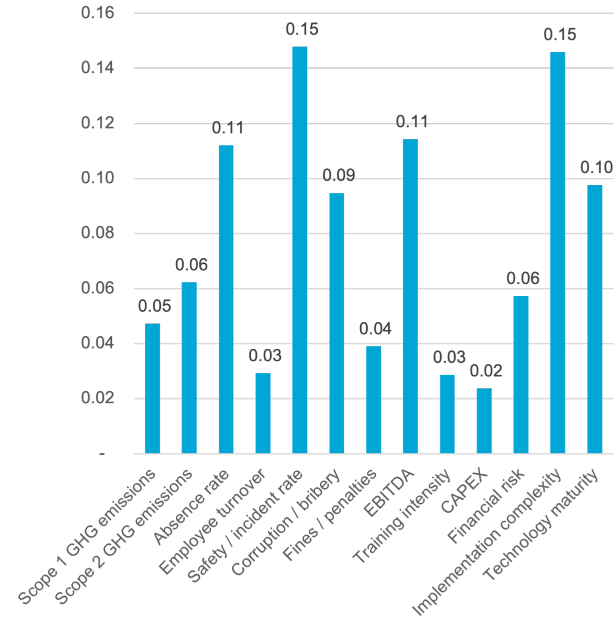


Figure 4: Geometric-mean AHP weight profile across elicited participants ($\max(CR) = 0.039$).

($w = 0.15$) and implementation complexity ($w = 0.15$). The aggregation balances substantially divergent priorities, since no individual participant held this profile.

Robustness setup

Monte Carlo robustness tests were configured following the four-input-family logic of Section 3.5. AHP pairwise comparisons were perturbed multiplicatively in log-space at two noise levels ($\sigma = 0.15$ and 0.30), truncated to the Saaty scale. Performance function curvature (α) was perturbed by discrete shifts to adjacent admissible template levels. Ordinal-to-baseline translation values were perturbed within a ± 50 percentage-point band centred on the elicited value. Scenario-specific quantitative impacts were perturbed proportionally at $\pm 10\%$; qualitative impacts were perturbed by at most one ordinal step (with probabilities 0.25, 0.50, 0.25 for down/unchanged/up). Each experiment was run for 5,000 iterations, and combined-perturbation experiments were also conducted to test interaction effects.

5. Results

The results are presented in two steps. The first establishes how the four investments compare under fixed assumptions, both in the base case and across scenarios. The second tests how far that comparison depends on the elicited inputs.

5.1. Baseline and scenario-specific outcomes

Under fixed assumptions and additive aggregation, **Investment D (dry dock replacement)** achieves the highest aggregated value in **Scenario 1 (Base case)**, driven by

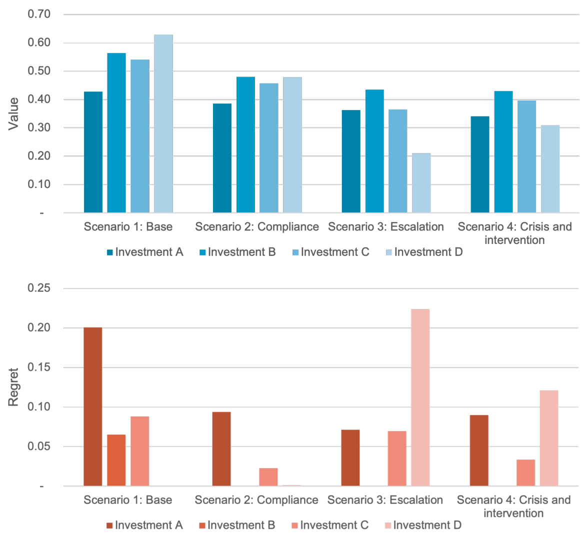


Figure 5: Aggregated value (upper) and scenario-specific regret (lower) per investment across the four scenarios. Higher value is preferred; lower regret indicates greater robustness.

strong performance on safety, EBITDA, and workforce-related indicators in conditions where capacity expansion delivers operational benefit. Under **Scenario 3 (Escalation)**, however, **Investment D**'s value drops sharply. The same capacity that supports operational performance under stable demand becomes a sunk-cost exposure when legitimacy pressure constrains workforce availability and order intake. **Investment B (cybersecurity training)** shows a flatter profile across scenarios, with consistent contributions through risk mitigation and compliance pathways but limited upside in any single context. **Investments A** and **C** occupy intermediate positions, with **Investment A** delivering modest, scenario-stable environmental benefit and **Investment C** concentrating its value in scenarios where enforcement is active.

Figure 5 reports value and regret per investment per scenario. Under minimax regret, **Investment B** is preferred ($MR_B = 0.065$). **Investment D**, despite winning under baseline aggregation in the favourable scenarios, has the highest maximum regret ($MR_D = 0.224$) due to its exposure under Escalation. The two recommendations therefore diverge, since the investment with the strongest baseline upside is not the investment with the lowest exposure across plausible futures. Both follow from the same elicited inputs, the same indicator set, and the same propagated consequences. What differs is the decision rule applied to the resulting value matrix.

Whether this divergence is an artefact of the elicited inputs or a stable feature of the case is the question the robustness analysis addresses next.

5.2. Robustness of the recommendation

Monte Carlo perturbation of elicited inputs revealed a sharp asymmetry between the valuation layer and the consequence layer. When AHP pairwise comparisons were perturbed at both tested noise levels ($\sigma = 0.15$ and $\sigma = 0.30$), **Investment B** was selected as the minimax regret recommendation in all 5,000 runs, indicating full stability of the conclusion under tested variation in elicited preferences. Variation in performance function curvature similarly left the recommendation unchanged across all runs.

Perturbation of consequence-layer inputs produced a substantially different picture. Perturbing the ordinal-to-baseline translation rules alone left **Investment B** selected in all 5,000 runs. When only scenario-specific investment impacts were varied (proportional perturbation for quantitative impacts, one-step ordinal shifts for qualitative impacts), **Investment B** was selected in 1,925 of 5,000 runs. Under combined perturbation of all four input families, the distribution flattened further: **Investment D** became the modal winner with 2,130 of 5,000 runs, followed by **Investment B** with 1,505, **Investment C** with 1,257, and **Investment A** with 108 runs (Figure 6).

This asymmetry is methodologically expected. Consequence-layer inputs feed directly into the performance scores that propagate through the entire framework, so sensitivity to them is structural rather than a sign of model misspecification. The substantive finding is the asymmetry itself. In this case, the recommendation does not depend on contested judgements about how organisational priorities should be weighed. Rather, it depends on contested assumptions about what investments will do under each scenario.

6. Discussion

The case results speak to the framework's analytical contribution along four lines, taken up in turn below.

6.1. Findings in light of the central claim

First, retaining heterogeneity across indicator types, rather than collapsing them into a single monetary metric, produced a comparison in which **Investment B** emerged as the robust recommendation. Its value lies almost entirely in tail-risk mitigation and would be largely invisible to a cash-flow-based appraisal. Second, replacing probabilistic forecasting with bounded scenario analysis allowed **Investment D**'s exposure under **Scenario 3 (Escalation)** to enter the comparison without requiring participants to commit to a probability distribution they did not hold. Third, sensitivity in the case concentrates in the consequence layer rather than the valuation layer. This demonstrates that the modular separation of value structuring, consequence construction, and robustness evaluation does analytical work, not merely organisational work.

6.2. Pathways surfaced by the system dynamics representation

In three instances during elicitation, the SDR identified causal pathways that direct expert assessment had initially missed. These pathways materially affected the case recommendation and illustrate the analytical role the representation plays under deep uncertainty.

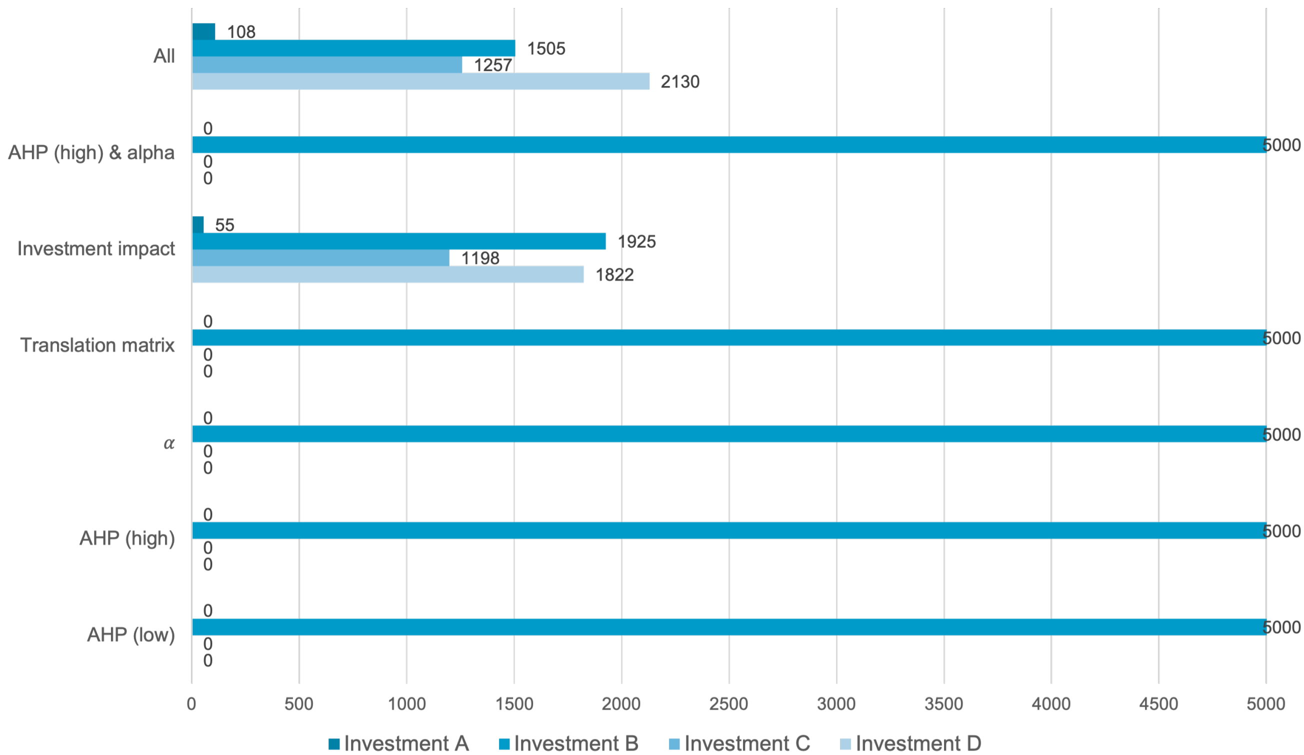


Figure 6: Monte Carlo selection frequencies across input perturbation experiments (5,000 runs each). The valuation layer (AHP, α) and the translation rules both leave the recommendation unchanged, while only the scenario-conditioned investment impacts move it.

For **Investment B**, direct assessment initially treated it as a purely governance-oriented risk reduction measure with no legitimacy dimension. Reasoning through the consequences of a hypothetical data breach surfaced what one participant described as "social pressure in the closet". A leak of employee data could erode internal legitimacy, increase absence as affected employees disengage, and trigger regulatory attention. A leak of owner data could translate into social pressure and visible disruptions. Both are reinforcing loops invisible under direct assessment but immediate once made explicit. This pathway is partly responsible for **Investment B**'s flat regret profile.

For **Investment C**, direct assessment dismissed yard security as unnecessary on the grounds that Dutch trespassing law already excluded activists from the yard. Causal reasoning surfaced two neglected loops. First, the absence of visible security established a cultural baseline in which minor governance infractions were normalised, drifting toward broader integrity failure. Second, subcontractors were observed using borrowed badges to bring in untrained or undocumented workers, a pattern that could eventually trigger labour-inspectorate intervention. Both pathways reframed yard security from operational nuisance to governance investment.

For **Investment D**, the dry dock was initially motivated solely by operational need (the existing slipway could not handle larger vessels). Reasoning through the scenario structure surfaced an unintended legitimacy effect. The new structure's heavy insulation and elimination of windows sharply reduced disturbance to local residents. This converted what had been framed as a capacity investment into a partial community-relations investment as well.

In none of these three cases did the representation generate new empirical information. It made existing organisational knowledge explicit, structured, and contestable by embedding it in a causal representation. This is the analytical role Sterman (2000) attributes to system dynamics under deep uncertainty, and it is the role the representation played in the case.

6.3. Reproducing the upside-bias dynamic

The divergence between the baseline-attractive **Investment D** and robustness-attractive **Investment B** reproduces, within a single decision process, the dynamic that Flyvbjerg (2021) identifies in capital allocation more broadly (Flyvbjerg and Bester 2021). **Investment D** has the profile of an investment that wins under upside-oriented valuation. It is large in scale, articulable in business-case terms, capacity-expanding, and supported by a clear EBITDA pathway. **Investment B** has the profile that loses. It is small in scale, defensive, and intangible in primary effect, with benefits realised mainly through avoided losses in adverse contexts.

The framework does not overturn the baseline assessment that **Investment D** is attractive under favourable conditions. It exposes the asymmetry between that assessment and the alternative's resilience under conditions the organisation cannot rule out. The contribution is therefore not a competing recommendation but a contestable second view, generated from the same data and the same elicited preferences, and applied through a decision rule designed to resist the bias that systematically disadvantages defensive investments.

6.4. Locating uncertainty

The Monte Carlo asymmetry, valuation-layer stability versus consequence-layer sensitivity, has a practical implication that survives the single-case setting. When a framework recommendation depends substantially on how consequences are estimated rather than on how they are valued, the productive disagreement is empirical rather than normative. Effort directed toward improving the traceability and challengeability of consequence estimates is well spent. Effort directed toward renegotiating criterion weights is not.

The residual question is therefore managerial: whether the high-social-pressure scenarios are regarded as material. The framework does not resolve that question. It narrows the deliberation to it.

6.5. Limitations

Several limitations qualify these claims. First, empirical evidence comes from a single case in a single sector. Transferability to other capital-intensive, project-based sectors with strong legitimacy exposure is plausible but untested. Second, the additive MAVT aggregation assumes mutual preferential independence between criteria, an assumption that is only approximate here (Dyer 2016; Gan et al. 2017). CAPEX, EBITDA, and Financial risk are not parallel goods, since CAPEX is incurred in expectation of EBITDA and higher Financial risk is often accepted in pursuit of it. Third, the case scenarios are static. The framework compares alternatives across distinct external futures but does not capture path dependence across successive allocation rounds or the possibility that organisational preferences evolve in response to outcomes. Fourth, elicitation was compressed into two ninety-minute sessions, trading depth for access to actual decision-makers. This limitation interacts directly with the Monte Carlo finding reported in Section 6.4. The framework's recommendation is most sensitive to scenario-conditioned investment impacts, which are precisely the inputs that benefit most from extended expert engagement. The value the framework can deliver is therefore bounded by the depth of consequence-layer elicitation it receives. Fifth, the framework's outputs depend on the skill of the elicitor in surfacing implicit assumptions without imposing them. This is a structural boundary condition for transfer to organisations without an embedded, methodologically experienced facilitator.

7. Conclusion

This paper has argued that the persistent gap between sustainability reporting and capital allocation is fundamentally a decision-logic problem rather than a measurement problem, and has developed a yard-level framework designed for the former. The framework integrates an ESRS-grounded indicator basis, an SDR for indirect consequence construction, AHP-weighted MAVT for non-monetised value aggregation, and minimax regret across non-probabilistic scenarios for robust cross-context comparison.

Application to a pure-custom superyacht shipyard demonstrated that the framework produces decision-relevant outputs that current practice does not generate, and that its modular architecture remains usable in an executive decision setting under realistic constraints on time and expert availability. Three outputs were decisive. The investment with the strongest baseline value, a large dry dock replacement, carried the highest maximum regret, while a small cybersecurity training investment minimised regret across scenarios. The SDR surfaced legitimacy and governance pathways that direct expert judgement had overlooked. Sensitivity concentrated in the scenario-conditioned consequence estimates rather than in the elicited value structure, which locates the productive disagreement in empirical rather than normative questions.

The framework therefore does not resolve the underlying uncertainty. It disciplines deliberation and narrows it to a single managerial question, whether the high-pressure scenarios are regarded as plausible enough to act on.

Two priorities for further work follow directly. First, cross-case application in other capital-intensive, project-based sectors operating in deep uncertainty would test whether the locating capability observed here generalises beyond the present setting. Second, extending the framework toward adaptive decision-making over successive allocation rounds would address the static-scenario limitation and connect the approach more directly to dynamic adaptive policy pathways.

References

Andrikopoulos, Andreas (Jan. 29, 2026). “ESG in the shipping industry: a literature review”. In: *Corporate Governance: The International Journal of Business in Society* 26.1, pp. 199–232. ISSN: 1472-0701, 1758-6054. DOI: 10.1108/CG-05-2024-0264. URL: https://www.emerald.com/cg/article/26/1/199/1249660/ESG-in-the-shipping-industry-a-literature-review (visited on 05/13/2026).

Boix-Cots, David et al. (Sept. 2022). “A systematic review on MIVES: A sustainability-oriented multi-criteria decision-making method”. In: *Building and Environment* 223, p. 109515. ISSN: 03601323. DOI: 10.1016/j.buildenv.2022.109515. URL: https://linkinghub.elsevier.com/retrieve/pii/S0360132322007454 (visited on 04/01/2026).

Crace, Logan and Joel Gehman (Mar. 2023). “What Really Explains ESG Performance? Disentangling the Asymmetrical Drivers of the Triple Bottom Line”. In: *Organization & Environment* 36.1, pp. 150–178. ISSN: 1086-0266, 1552-7417. DOI: 10.1177/10860266221079408. URL: https://journals.sagepub.com/doi/10.1177/10860266221079408 (visited on 11/14/2025).

Dall-Orsoletta, Alaize et al. (June 2025). “How does social acceptance affect transition minerals production in Europe? A system dynamics approach and case study in Portugal”. In: *The Extractive Industries and Society* 22, p. 101625. ISSN: 2214790X. DOI: 10.1016/j.exis.2025.101625. URL: https://linkinghub.elsevier.com/retrieve/pii/S2214790X25000140 (visited on 03/30/2026).

Dyer, James S. (2016). “Multiattribute Utility Theory (MAUT)”. In: *Multiple Criteria Decision Analysis*. Ed. by Salvatore Greco, Matthias Ehrgott, and José Rui Figueira. Vol. 233. New York, NY: Springer New York, pp. 285–314. ISBN: 978-1-4939-3094-4. DOI: 10.1007/978-1-4939-3094-4_8. URL: https://link.springer.com/10.1007/978-1-4939-3094-4_8 (visited on 01/21/2026).

European Commission (July 31, 2023). *Commission Delegated Regulation (EU) 2023/2772 of 31 July 2023 supplementing Directive 2013/34/EU of the European Parliament and of the Council as regards sustainability reporting standards*. OJ L, 2023/2772, 22.12.2023. URL: http://data.europa.eu/eli/reg_del/2023/2772/oj (visited on 02/19/2026).

European Parliament and Council of the European Union (Dec. 14, 2022). *Directive (EU) 2022/2464 of the European Parliament and of the Council of 14 December 2022 amending Regulation (EU) No 537/2014, Directive 2004/109/EC, Directive 2006/43/EC and Directive 2013/34/EU, as regards corporate sustainability reporting*. URL: http://data.europa.eu/eli/dir/2022/2464/oj (visited on 05/22/2026).

— (May 10, 2023). *Directive (EU) 2023/959 of the European Parliament and of the Council of 10 May 2023 amending Directive 2003/87/EC establishing a system for greenhouse gas emission allowance trading within the Union and Decision (EU) 2015/1814*. URL: http://data.europa.eu/eli/dir/2023/959/oj (visited on 05/22/2026).

Flyvbjerg, Bent (Dec. 2021). “Top Ten Behavioral Biases in Project Management: An Overview”. In: *Project Management Journal* 52.6, pp. 531–546. ISSN: 8756-9728, 1938-9507. DOI: 10.1177/87569728211049046. URL: https://journals.sagepub.com/doi/10.1177/87569728211049046 (visited on 01/05/2026).

Flyvbjerg, Bent and Dirk W. Bester (2021). “The Cost-Benefit Fallacy: Why Cost-Benefit Analysis Is Broken and How to Fix It”. In: *Journal of Benefit-Cost Analysis* 12.3, pp. 395–419. ISSN: 2194-5888, 2152-2812. DOI: 10.1017/bca.2021.9. URL: https://www.cambridge.org/core/product/identifier/S2194588821000099/type/journal_article (visited on 01/05/2026).

Gan, Xiaoyu et al. (Oct. 2017). “When to use what: Methods for weighting and aggregating sustainability indicators”. In: *Ecological Indicators* 81, pp. 491–502. ISSN: 1470160X. DOI: 10.1016/j.ecolind.2017.05.068. URL: https://linkinghub.elsevier.com/retrieve/pii/S1470160X17303357 (visited on 12/12/2025).

Gillan, Stuart L., Andrew Koch, and Laura T. Starks (Feb. 2021). “Firms and social responsibility: A review of ESG and CSR research in corporate finance”. In: *Journal of Corporate Finance* 66, p. 101889. ISSN: 09291199. DOI: 10.1016/j.jcorpfin.2021.101889. URL: https://linkinghub.elsevier.com/retrieve/pii/S0929119921000092 (visited on 06/11/2026).

Hallegatte, Stéphane et al. (Sept. 2012). *Investment Decision Making under Deep Uncertainty - Application to Climate Change*. World Bank, Washington, DC. DOI: 10.1596/1813-9450-6193. URL: https://hdl.handle.net/10986/12028 (visited on 12/23/2025).

Hirtenlehner, Jerome Joseph (Jan. 3, 2025). “Rocking the Boat: Alt-Fuels and the Super Yacht Industry”. Master Thesis. Universidade Católica Portuguesa.

Hojnik, Jana et al. (Mar. 2020). “Sustainability indicators for the yachting industry: Empirical conceptualization”. In: *Journal of Cleaner Production* 249, p. 119368. ISSN: 09596526. DOI: 10.1016/j.jclepro.2019.119368. URL: https://linkinghub.elsevier.com/retrieve/pii/S0959652619342386 (visited on 11/13/2025).

Janson, Daniel I (Nov. 2016). “The Development Of A Green Shipyard Concept”. Master Thesis. Delft: Delft University of Technology.

Lempert, Robert J. (2019). “Robust Decision Making (RDM)”. In: *Decision Making under Deep Uncertainty*. Ed. by Vincent A. W. J. Marchau et al. Cham: Springer International Publishing, pp. 23–51. ISBN: 978-3-030-05252-2. DOI: 10.1007/978-3-030-05252-2_2. URL: http://link.springer.com/10.1007/978-3-030-05252-2_2 (visited on 01/21/2026).

Marchau, Vincent A. W. J. et al., eds. (2019). *Decision Making under Deep Uncertainty: From Theory to Practice*. Cham: Springer International Publishing. ISBN: 978-3-030-05252-2. DOI: 10.1007/978-3-030-05252-2. URL: http://link.springer.com/10.1007/978-3-030-05252-2 (visited on 12/29/2025).

McCluskey, Mitchell (July 17, 2023). *Spanish activists vandalize superyacht in Ibiza believed to belong to billionaire Walmart heiress*. CNN. URL: https://www.cnn.com/2023/07/17/europe/activists-vandalize-yacht-ibiza-climate-intl (visited on 11/25/2025).

Saaty, Thomas L (June 1977). “A scaling method for priorities in hierarchical structures”. In: *Journal of Mathematical Psychology* 15.3, pp. 234–281. ISSN: 00222496. DOI: 10.1016/0022-2496(77)90033-5. URL: https://linkinghub.elsevier.com/retrieve/pii/0022249677900335 (visited on 05/22/2026).

— (2008). “Decision making with the analytic hierarchy process”. In: *International Journal of Services Sciences* 1.1, p. 83. ISSN: 1753-1446, 1753-1454. DOI: 10.1504/IJSSCI.2008.017590. URL: http://www.inderscience.com/link.php?id=17590 (visited on 12/09/2025).

Smit, Han (May 2025). “Sustainability Real Options”. In: *California Management Review* 67.3, pp. 55–85. ISSN: 0008-1256, 2162-8564. DOI: 10.1177/00081256251331264. URL: https://journals.sagepub.com/doi/10.1177/00081256251331264 (visited on 11/13/2025).

Stanton, Muriel C. Bonjean and Katy Roelich (Oct. 2021). “Decision making under deep uncertainties: A review of the applicability of methods in practice”. In: *Technological Forecasting and Social Change* 171, p. 120939. ISSN: 00401625. DOI: 10.1016/j.techfore.2021.120939. URL: https://linkinghub.elsevier.com/retrieve/pii/S0040162521003711 (visited on 12/29/2025).

Sterman, John D. (2000). *Business dynamics: systems thinking and modeling for a complex world*. International student edition. Boston Burr Ridge, IL Dubuque, IA Madison, WI New York San Francisco St. Louis Bangkok Bogotá Caracas Lisbon London Madrid Mexico City Milan New Delhi Seoul Singapore Sydney Taipei Toronto: Irwin, McGraw-Hill. 982 pp. ISBN: 9780071179898.

Stuart, Alice et al. (Aug. 2023). “Conceptualising social licence to operate”. In: *Resources Policy* 85, p. 103962. ISSN: 03014207. DOI: 10.1016/j.resourpol.2023.103962. URL: https://linkinghub.elsevier.com/retrieve/pii/S0301420723006736 (visited on 03/09/2026).

Vakili, Seyedvahid, Alessandro Schönborn, and Aykut I. Ölçer (Apr. 2023). “The road to zero emission shipbuilding Industry: A systematic and transdisciplinary approach to modern multi-energy shipyards”. In: *Energy Conversion and Management: X* 18, p. 100365. ISSN: 25901745. DOI: 10.1016/j.ecmx.2023.100365. URL: https://linkinghub.elsevier.com/retrieve/pii/S2590174523000211 (visited on 11/13/2025).

Walker, Warren E., Robert J. Lempert, and Jan H. Kwakkel (2013). "Deep Uncertainty". In: *Encyclopedia of Operations Research and Management Science*. Ed. by Michael C. Fu and Saul I. Gass. Boston, MA: Springer US, pp. 395–402. ISBN: 978-1-4419-1153-7. DOI: 10.1007/978-1-4419-1153-7_1140. URL: http://link.springer.com/10.1007/978-1-4419-1153-7_1140 (visited on 12/29/2025).

Young-Ferris, Anna and John Roberts (May 27, 2023). "'Looking for Something that Isn't There': A Case Study of an Early Attempt at ESG Integration in Investment Decision Making". In: *European Accounting Review* 32.3, pp. 717–744. ISSN: 0963-8180, 1468-4497. DOI: 10.1080/09638180.2021.2000458. URL: https://www.tandfonline.com/doi/full/10.1080/09638180.2021.2000458 (visited on 12/17/2025).