A Novel Concept for Steel Building Cost Estimation in Shipbuilding Using Process and Product Similarities

Alblas, G.; Pruyn, J.F.J.

DOI
10.3390/jmse12111936

Publication date
2024

Document Version
Final published version

Published in
Journal of Marine Science and Engineering

Journal of
Marine Science
and Engineering

MDPI

Article

A Novel Concept for Steel Building Cost Estimation in Shipbuilding Using Process and Product Similarities

Gerrit Alblas * and **Jeroen Pruyn**

Faculty of Mechanical, Maritime and Materials Engineering, Technical University Delft, 2628 CD Delft, The Netherlands; j.f.j.pruyn@tudelft.nl
* Correspondence: g.alblas@tudelft.nl

Abstract: Accurate cost estimates are essential for staying in business in a competitive shipbuilding industry. With new technologies and the energy transition creating an ever-changing landscape, traditional cost estimation methods based on product specifications can no longer keep pace. The need for improvement especially arises for Engineering-To-Order projects, considering that the profit margins are narrow. The use of process information in the estimation process could increase the reliability and flexibility of these estimations. This article presents a concept that utilizes graph theory to include process information in the cost estimations applied to steel buildings. This concept is specifically suited for the early stages of Engineering-To-Order projects.

Keywords: cost estimation method; graph theory; Engineering-To-Order; pre-contract; shipbuilding; steel building

Citation: Alblas, G.; Pruyn, J. A Novel Concept for Steel Building Cost Estimation in Shipbuilding Using Process and Product Similarities. *J. Mar. Sci. Eng.* **2024**, *12*, 1936. https://doi.org/10.3390/jmse12111936

Academic Editor: Md Jahir Rizvi

Received: 27 August 2024
Revised: 16 October 2024
Accepted: 18 October 2024
Published: 29 October 2024

1. Introduction

In the highly competitive shipbuilding industry, it is essential to respond quickly to a changing market. This has consequences for the way cost estimates are completed. Especially since Engineering-To-Order (ETO) products are becoming more commonplace, such a one-of-a-kind product poses challenges in cost estimation, particularly as customer involvement often continues throughout the design process. In addition, the information available at the time a shipyard must provide a quote is limited, and uncertainties within estimations and the supply chain, including owner-furnished equipment, must also be considered [1,2].

For the shipbuilding industry, accurate cost estimation is crucial for staying in business. Especially when we consider that profit margins are small. A competitiveness study [3], albeit from 2009, indicates a 2–6% profit margin, which is small compared to many other industries. The more recent Kamola-Cieślik study [4] concludes that Europe's shipbuilding industry struggles to stay active, and profit margins have not increased.

To deal with this and address the labor cost advantages of Asian countries, the European shipbuilding industry has shifted its focus toward innovation and the construction of complex, high-value ships [5]. Such ETO projects pose additional challenges for shipyards in Western Europe when it comes to cost estimating, as it is common practice to conclude fixed-price contracts based on available information developed during the pre-contract phase. In contrast, in major shipbuilding countries such as South Korea, China, and Japan, it is more common to conclude contracts for each shipbuilding phase separately.

In addition to challenging profit margins, the shipbuilding industry is now also confronted with mounting challenges related to safety regulations, cost-effectiveness, and the pressing need to address energy conservation and environmental protection [6]. All of this adds to the challenges associated with estimating costs as, for all project types, experience with such factors is limited or non-existent.

Advanced and accurate cost estimation methods necessitate comprehensive design details to deliver a well-founded and precise cost estimate. Unfortunately, such detailed information is typically unavailable during the pre-contract phase.

The literature study of Alblas et al. [7] established that the shipbuilding industry is struggling to find the right course to deal with all these mentioned challenges. They propose integrating the assembly process into the cost estimation to enhance the cost estimation accuracy for ETO projects in the pre-contract phase. Integrating the construction process means that the cost estimation will not solely depend on (basic) statistics linked to product properties, combined with expert knowledge, but also includes insights into the assembly order of the blocks. How to approach this integration is the objective of the current study. Consequently, the research question addressed is: How to incorporate, in the pre-contract phase, the product and construction process in cost estimations?

To answer this research question, we focus on the backbone of the shipbuilding process: steel building. Before delving into this, we first provide a brief overview of commonly used cost estimation methods in shipbuilding as well as in related industries.

2. Brief Literature Review

In the cost estimation methods outlined here, we focus solely on methods for estimating construction person-hours. As the costs for acquisition, overhead, design, engineering, material, and equipment are commonly included in the hourly rate used for production, this is out of the scope, as we concentrate on person-hours only.

2.1. Shipbuilding Industry

Cost estimation methods in shipbuilding that focus on estimating person-hours traditionally rely on product specifications, past performance, and expert judgment [8]. All commonly known cost estimation methods deal with these aspects in a limited number of approaches [7]. Table 1 provides an overview of the commonly used cost estimation methods in the shipbuilding industry. Based on our review, an overview is given of the advantages, drawbacks, and usability of Engineering-To-Order (ETO)-projects in the pre-contract phase for each of the methods.

Table 1. Overview of shipbuilding cost estimation methods.

Cost Estimation Methods			ETO, Pre-Contract Phase
Name	**Advantage**	**Drawback and Limitations**	**Usability**
Intuitive Method [9]	- Quick to produce - Flexible	- Susceptible to bias - Unstructured, not reproducible, untraceable - Different experts use different mechanisms	- Unreliable and poorly verifiable outcome - No traceback when design change occurs [8,10,11]
Parametric Method [12]	- Quick to produce - Repeatable and objective	- Parameters not included may be important - Logic not visible - Cannot deal with an introduction of new technologies	- False sense of security (black box) - Unsecure outcome - Traceback limited applicability [13]
Feature Based Method [14]	- Possibility of integrating CAD/CAM with cost information - Has the protentional to be automated - Link between design choices and cost consequences	- Requires large resources to implement - No consensus on what features are - Detailed design information is needed	- Not applicable, no detail design information and no past performance information available [15]

Table 1. *Cont.*

Cost Estimation Methods			ETO, Pre-Contract Phase
Name	**Advantage**	**Drawback and Limitations**	**Usability**
Case Based Reasoning [16]	- Can offer a solution rapidly - Good logical visibility - Avoid previously committed errors - Implicit storage of company knowledge	- A reliable case base is needed - Cannot deal with an introduction of new technologies - Does not handle innovative solutions	- Not useful in the absence of similarities - Cannot handle new technologies [17]
Fuzzy logic Method [18]	- Good logic visibility - Integration of the imperfection of the model	- Human expertise and know-how is needed	- Unreliable, outcome depends on variables considered - Limited traceback when design change occur - Cannot deal with new technologies or new systems, no experience [19]
Neural Network Method [20]	- Accurate estimates are possible - Can be updated and retrained	- Logic not visible (black box) - Complex - Requires a large sample database - Cannot handle innovative solutions	- Not directly applicable, a sufficiently usable example database is missing [20]

The Intuitive Method is primarily used in the shipbuilding industry for cost estimation of an ETO project in the pre-contract phase [21]. The accuracy and comprehensiveness of estimations based on expert knowledge are difficult to demonstrate, and it is often hard to reproduce and trace them back to the original cost estimation during a project. And the absence of a structured expert opinion compromises the reliability and reproducibility of estimates. These issues are particularly pronounced when estimating costs for ETO projects during the pre-contract phase [22].

2.2. Other Industries: Construction, Building, and Manufacturing

A recent review of cost estimation used for construction projects outside the shipbuilding industry was performed by Hashemi et al. [23]. While the primary focus of this analysis is machine learning techniques, it also delved into exploring alternative cost estimation methods. They showed that the lack of comprehensive data for cost estimation can be resolved by utilizing product-based generated models, as seen in the shipbuilding industry. However, no approach was identified in this study that considers the actual construction processes.

In the construction industry, Building Information Modeling (BIM) is commonly used. BIM is a process that involves creating and maintaining digital representations of the physical and functional characteristics of places. It helps architects, engineers, and other stakeholders to gain a better understanding of the project environment and make informed decisions. BIM has been around since 1970 and has been widely accepted. Its functionality has been expanded over the years, leading to the development of 5D BIM, which supports the planning and cost estimation process linked to the 3D model. However, according to Ossa Mesa [24], even with 5D BIM, cost estimation for new elements still relies on expert knowledge. The possibility of comparing costs with previously completed projects is based solely on product properties and does not consider the construction process.

The review by Miranda et al. [25] provides valuable information on the estimation methods used in the early stages of building construction. The study highlights that product-driven cost estimations are applied here, just as in the shipbuilding industry. Additionally, the article discusses the extra cost drivers that are solely based on product specifications.

The review by Ganorkar et al. [11] provides an overview of the methods utilized in the manufacturing industry. This overview reveals that, during the early stages of a project, cost estimation in the manufacturing industry relies predominantly on qualitative methods. These methods include the Intuitive Method, Case-Based Reasoning, and Regression Analytics, which are also employed in the shipbuilding industry in the early stages of a project. However, their application in manufacturing differs, as greater emphasis is placed on production processes and machining properties.

This brief review shows that methods used in the construction, building, and manufacturing industries are, just as in shipbuilding, primarily based on product specifications and past performance information. The application, however, differs. A methodical use of construction processes or construction patterns to estimate costs is only found in the manufacturing industry. In this industry, the (machine) process model for manufacturing the product is known, as with more than 1000 similar products, the inclusion and investment in process improvements pays off quickly.

In the following sections, we present a concept aimed at integrating construction processes into cost estimations to improve the accuracy, flexibility, and traceability of the estimations compared to existing methods in shipbuilding.

3. A Concept for Steel Building Cost Estimation Based on Process and Product Similarities

Before describing the concept, a short description of the information flow through the stages of the ship design and construction process is given. This is followed by paragraphs on how the available design information can be used to identify the construction process. Next, an information visualization is proposed that captures design and process information using past performance information.

During the pre-contract phase, a yard must investigate if they can build the ship on time and according to the requirements. To answer that, the yard defines a project strategy [26,27] or building strategy. This strategy is created by experts [22] using the available concept design information and knowing the possibilities of the yard and its suppliers. For ETO projects, the process of defining a building strategy, or production planning, is a task performed by experts, especially [28]. The strategy devised during the pre-contract stage incorporates a comprehensive planning approach for both interim and final products. During the contract phase, this pertains to the construction planning of the identified building blocks, as outlined in a block division plan, and the planning for assembling these blocks into a complete ship. This is referred to as the erection plan.

Traditional cost estimation methods in shipbuilding rely on expert knowledge, product specifications, and past performance information [7]. The accuracy and comprehensiveness of these estimations are difficult to demonstrate, and it is often hard to reproduce and trace them back to the original cost estimation during a project. The cost estimation concept we present here addresses these shortcomings.

In this paper, we focus on the steel building process, as it serves as the key system for comprehensive planning, with a specific focus on estimating labor hours. A substantiation for this can be found in the work of Qu et al. [29]. They enlightened the importance of the block assembly process. They argued that the hours and costs of constructing a block, respectively, account for 40~60% of the total hours and 30~50% of the cost of hull building. A steel assembly process in the basis consists of three activities (positioning, welding, and grinding). The efficiency of such activities can be influenced by factors such as connection length, weight, accessibility, position, orientation, etc.

For simplicity and clarity of the concept, we assume that the parts used in constructing steel assemblies have already been pre-processed. As such, the person-hours required for the pre-processing do not affect the proposed method. This is a simplification for clarity, not a limitation of the concept itself, as even different pre-processing methods could easily be included.

3.1. Overview of the Concept

The proposed cost estimation process, see Figure 1, is based on information that is available and accessible during the pre-contract phase. In the case of steel buildings, this information typically includes the general arrangement and a block division plan derived from it.

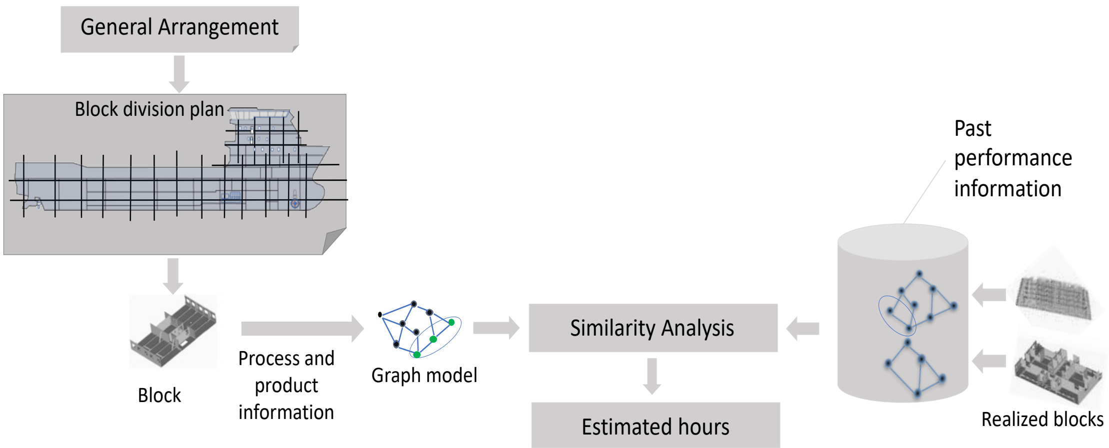


Figure 1. Estimation process overview. The block division plan demarcates the blocks to be built. Each block undergoes a defined construction process, which is then represented in a graph model. The most similar block from past performance information is determined by performing a similarity analysis, resulting in estimated person-hours. The past performance information used in the analysis includes hours spent and the applied construction process for each of the completed blocks, represented in a graphical model.

Based on the block division plan, person-hours for constructing the block are estimated. To make this estimate, it is essential to know the construction process of each block to build. One approach to defining the block construction process is to consult experts. Alternatively, the construction process can be defined using information from past performance, provided such data are already available.

The construction process information of the block to build is then translated into an information representation that captures both the block qualities and process information. A comparison is made with past performance information based on similarities to estimate the hours needed to build the block. For this purpose, the past performance information must be structured in the same information representation format. The goal of the similarity analysis is to identify the most similar block in the past performance information, which will serve as the basis for estimating the hours required to build the new block.

In the next section, we first go into detail about the available information in the pre-contract phase, followed by a description of how to represent the relation between product and process information and how to apply this in the proposed cost estimation method.

3.2. Available Information

In the pre-contract phase, design information for the steel system is contained in what is called General Arrangement (GA) documentation. A GA of a ship design encompasses all aspects of the ship's physical layout and spatial configuration to ensure optimal functionality, safety, and efficiency. The design at the level of GA represents the following elements of the ship:

- Overall dimensions;
- Hull shape;
- Bulkhead position and dimensions;
- Number and heights of decks;
- Locations and dimensions of rooms;

- Superstructure layout;
- Position and weight of main components.

Based on the GA, experts consider how the ship and part of the ship must be constructed, taking the possibilities and limitations of a yard and known sub-contractors in mind. This is illustrated in Figure 2. The construction costs estimated traditionally are, therefore, implicitly related to the construction strategy and construction schedule. How the ship is going to be constructed results in a block division plan, where blocks are large construction elements weighing between 60 and 200 metric tons and composed of several sub-assemblies. Alongside the block division plan, an erection plan is defined, outlining the sequence and methodology for assembling the constructed blocks to build the ship.

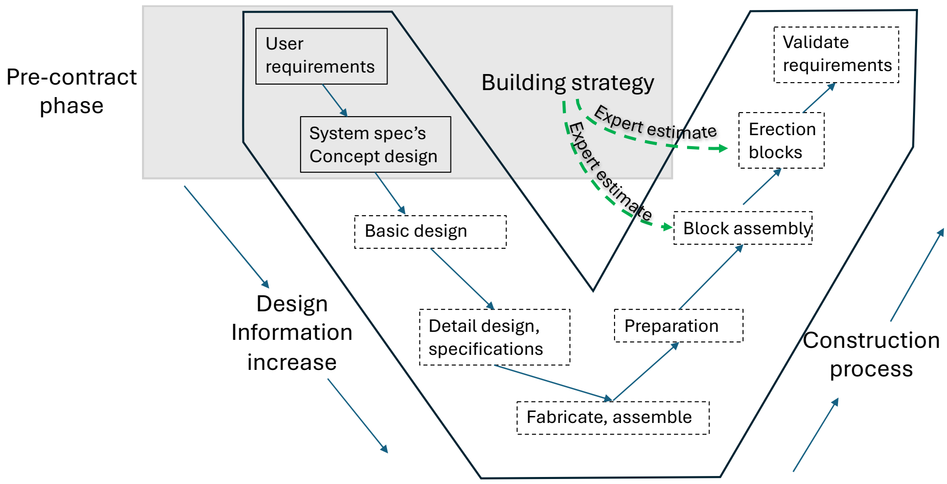


Figure 2. The evolving of design and construction information. This figure is derived from the work of Parraguez et al. [14]. The left side of the figure depicts the progression of design information until it reaches a stage suitable for part production. Conversely, the right side illustrates the development of the construction process, which begins with the fabrication of parts and their preparation for assembly into blocks. During the development of the building strategy, experts define the overall assembly methods for the blocks and outline the procedures for their erection, as illustrated by the green arrows. Once a block is assembled, the focus shifts to the erection of those blocks. At the pre-contract stage, where detailed design information is limited, experts traditionally rely on their knowledge to estimate the person-hours required for building and erecting the blocks.

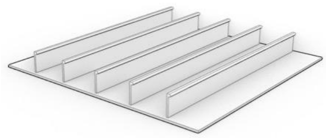
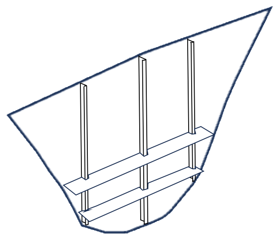
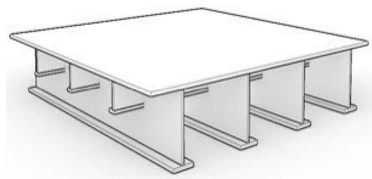
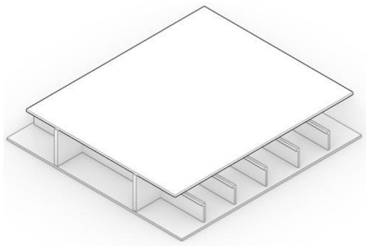
The yard's production planning is based on these plans, as the steel system forms the backbone of the entire planning process [30]. Based on the block division plan, the following information about blocks is available:

- Block type and positioning;
- (Estimated) Weight;
- Outer and inner shapes and dimensions;
- Sub-assemblies (straight panel, curved panel, bulkhead, deck, web frames);
- Sub-assembly weight and dimension;
- Three-dimensional positioning of sub-assemblies;
- Yard standards (profile type, profile distance, and directions).

As mentioned, the construction strategy, construction schedule, block distribution plan, and erection plan are determined in the pre-contractual phase. Furthermore, the subassemblies for the defined blocks are known. A list of typical sub-assemblies known in the pre-contract phase is presented in Table 2.

In conclusion, during the pre-contract phase, a layout (GA) with dimensions and potential sub-assemblies, as illustrated in Table 2, is available. Therefore, a suitable method should be capable of utilizing this limited information while evaluating product and construction process information.

Table 2. Sub-assemblies.

Sub-Assembly	Sub-Assembly Parts	Schematic Representation
Panel	Plate panels, primary and secondary stiffening structural members	
Web frame	Transverse rings of deck beams	
Bulkhead	Panels of plating, could also be stiffened with a combination of welded horizontal and vertical stiffeners	
Deck	Deck plates, girders and plate web	
Double bottom	Plates and stiffeners	

3.3. Information Representation

Considering that the construction process is known, a choice must be made on how to structure the information to be used for cost estimation purposes. Such structuring must consider the complexity of the construction process and ensure that the most relevant data points can be tracked and analyzed.

Liu et al. [21] and Iwańkowicz [31] have utilized a graph representation to optimize assembly sequences. A graph theory-based representation provides a versatile platform for detailing objects and their interconnections at the level of occurrence. As elucidated by Bunke [32] in his graph theory discourse, this representation is particularly advantageous for conducting similarity assessments. As a result, graph theory is proposed as the basis of a method to estimate production costs.

In the proposed graph model, nodes represent the steel parts, while edges between the nodes represent the activities required to connect these parts. Each edge contains information on person-hours, dimensions, and technical specifications. An illustrative example of node and edge information, specifically for a steel assembly, is given in Figure 3.

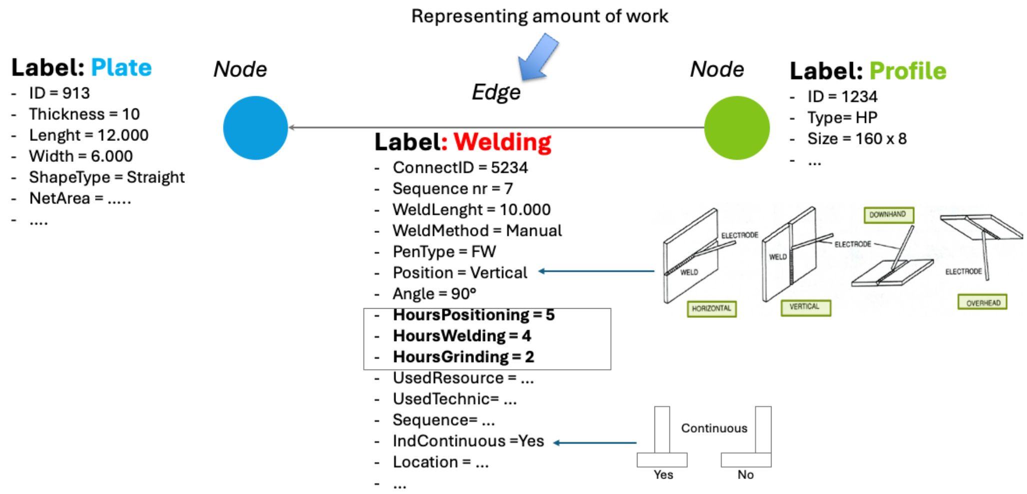


Figure 3. Graph example of a welding connection. In a graph model, each node represents entities, in this example, a "profile" and "plate". Each node has a unique identification (ID) to distinguish them from one other and properties about their dimension and shape. In this example, the edge, with label "Welding", connects the "Profile" to the "Plate". This edge carries information associated with the relationship between the two nodes. In this example, the edge includes: "Hours of work", representing the effort required to establish and complete the connection. "Welding sequence" represents the order in which a welding connection should be performed. "Welding position" indicates the orientation of the weld. "Used resources" refers to the resources used for the welding operation. "Type of welding" specifies the type of welding.

In a directed, ordered, and labeled graph, the edge sequence number and connection direction can be used to give details of the connection process. These details, expressed in information attributes, encompass specifics such as edge direction, identification, welding length, welding method, and welding type, along with allocated hours for positioning, welding, grinding, and details on resource usage. Every node is assigned a label describing the type of part or sub-assembly. A node contains identification information and information attributes containing technical information about the steel part.

Because the construction of a block is often based on assembling intermediate sub-assemblies, a mechanism is needed to represent the grouping of work [33]. Therefore, grouping information is added to the graph representation. This grouping mechanism represents a logical division of work, making it possible to define sub-assemblies, which may consist of several generic parts.

As mentioned above, a graph representation offers the flexibility of defining objects and relations between objects on an occurrence level. It provides the possibility to represent not only product information but also the way a product is built and the technology and resources used.

3.4. Example

In the previous sections, we have described the graph representation of product and process information. In this section, we further clarify this concept by using a stylized example. Our demonstration highlights how different assembly sequences for a product can be represented through the visualization of information in a directed graph.

Figure 4 contains a graph for product A, based purely on product information. It represents the specification of the several parts and that they are connected without taking the assembly process and sequence into account.

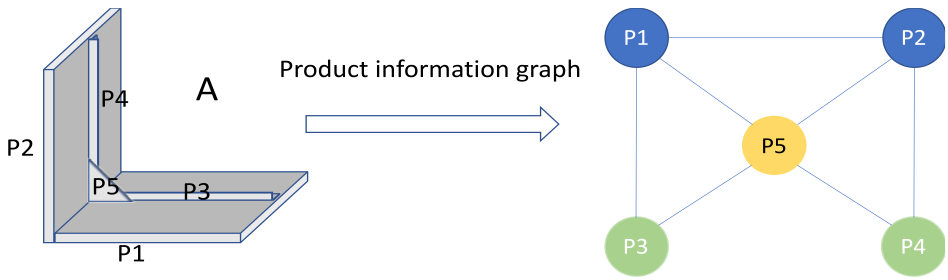


Figure 4. Product information graph. This figure illustrates for product A that plate P1 is connected to plate P2, and profile P3 is connected to plate P1 and to bracket P5. Both P1 and P2 have label "Plate". Both have attributes like length, width, plate thickness, and material type. The connection, called edge, between P1 and P2 has attributes like, for example, welding length and welding type.

There are multiple methods for constructing product A. To illustrate this, we defined two distinct assembly sequences, A1 and A2, which are presented in Figure 5.

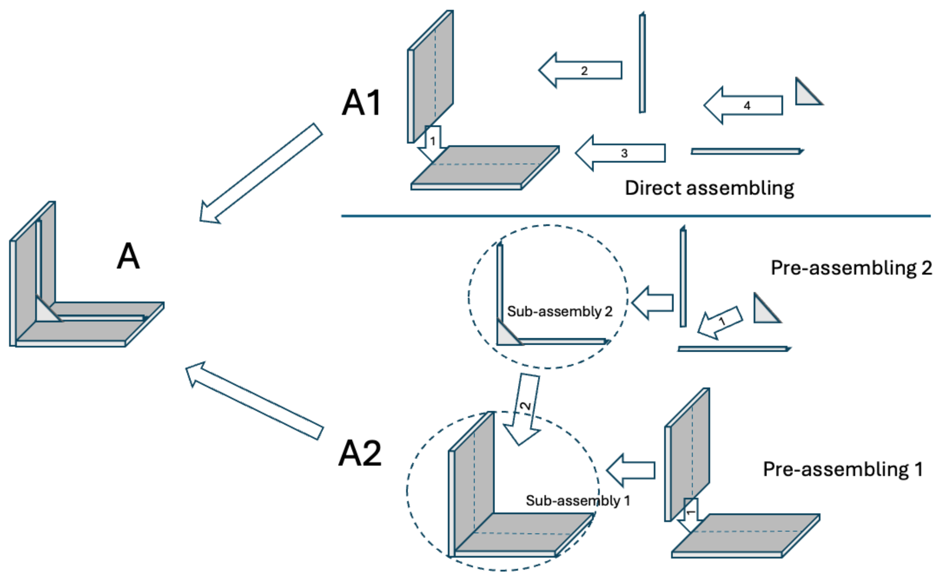


Figure 5. Two ways of assembling the same product. In the first way of assembling (A1), the two plates are first connected (1), then the profiles are each connected to a plate (2 and 3), and then the bracket is connected (4) to the profiles and the plates. In the second way of assembling (A2), parts are pre-assembled into two sub-assemblies. In the pre-assembling 1 the two plates are connected. In pre-assembling 2 the bracket is connected to the two profiles. In the next step the sub-assemblies 1 and 2 are connected (2) to construct product A.

In assembly sequence A1, the different parts are directly positioned and welded one by one, while in assembly sequence A2, the first two sub-assemblies are made and then connected to construct product A. The different assembly sequences A1 and A2 represent two different construction processes. By using a directed and labeled graph representation of the two construction processes, as illustrated in Figure 6, we capture the two ways the product can be built.

To explain the two different ways of constructing product A, a simplified list of operations is given in Tables 3 and 4. The list is simplified because operations like transport, lifting, and turning are left out. The presented person-hours in Tables 3 and 4 are not real figures; they are for illustrating purposes.

As shown in this stylized example, the two assembly sequences for building the same product result in two distinct graphs, each illustrating different work patterns. In the next section, we will discuss applying graphs in situations where not every design detail is known, as is the situation in the pre-contract phase.

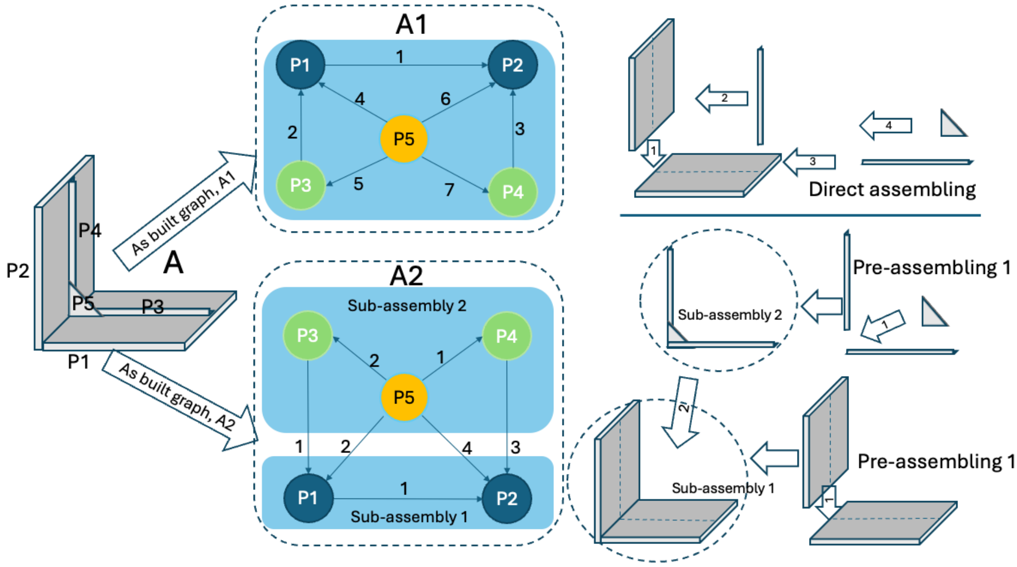


Figure 6. Two directed and ordered graphs. The two ways of constructing product A are represented in the two directed and ordered graphs, A1 and A2. In each of the presented graphs, an edge is given a direction and a number. E.g., to construct assembly A1, plate P1 is connected to plate P2 and has sequence number 1, and sequence number 2 represents the connection of profile 3 to plate P1. The graph of construction process A2 represents two levels of assembling. First, the two sub-assemblies are constructed, and then those two sub-assemblies are connected. As shown, each sub-assembly has its own sequence numbering, and the connection between the two sub-assemblies has its own sequence numbering.

Table 3. List of operations according to sequence A1 as shown in Figure 6.

Assembly Sequence A1	From Part	To Part	Activities	Person-Hours
1	Plate P1	Plate P2	Positioning, 1 × Welding, Grinding	0.5
2	Profile P3	Plate P1	Positioning, 2 × Welding, Grinding	0.8
3	Profile P4	Plate P2	Positioning, 2 × Welding, Grinding	0.8
4	Bracket P5	Plate P1	Positioning, 1 × Welding, Grinding	0.1
5	Bracket P5	Profile P3	1 × Welding, Grinding	0.05
6	Bracket P5	Plate P2	1 × Welding, Grinding	0.05
7	Bracket P5	Profile P4	1 × Welding, Grinding	0.05

Table 4. List of operations according to sequence A2 as shown in Figure 6.

Assembly Sequence A2	From Part	To Part	Activities	Person-Hours
Sequence Sub-assembly 1				
1	Plate P1	Plate P2	Positioning, 1 × Welding, Grinding	0.5
Sequence Sub-assembly 2				
1	Bracket P5	Profile P3	Positioning, 1 × Welding, Grinding	0.1
2	Bracket P5	Profile P4	Positioning, 1 × Welding, Grinding	0.1

Table 4. *Cont.*

Assembly Sequence A2	From Part	To Part	Activities	Person-Hours
Sequence Sub-assembly 1 + Sub-assembly 2				
1	Profile P3	Plate P1	Positioning, 2 × Welding, Grinding	0.8
2	Bracket P5	Plate P1	1 × Welding, Grinding	0.05
3	Profile P4	Plate P2	2 × Welding, Grinding	0.6
4	Bracket P5	Plate P2	1 × Welding, Grinding	0.05

3.5. Applying Graph Representation in Pre-Contract Phase

At the end of the pre-contract phase, costs must be estimated based on product information on a general arrangement level (or concept design), block division plan, and construction schedule. This means that when representing building block information in a graph, the objects are more generic. The objects in a graph representation of a block contain objects like panels, bottom, web-frame, bulkheads, and plate fields. See Table 2 for a list of objects that can be known in the pre-contract phase. Of those objects, not every detail is known.

During the pre-contract phase, structuring a graph model for a block revolves around determining whether the construction of a specific part is crucial for the assembly sequence of the block. E.g., a double bottom can be seen as a relevant and single object. The construction of a double bottom itself is independent of the way the double bottom is used in the assembling of the block. A double bottom is usually a standard object of which the construction hours are known from past performance information.

Figure 7 illustrates a block on the level of a concept design, together with a corresponding graph representing the construction sequence.

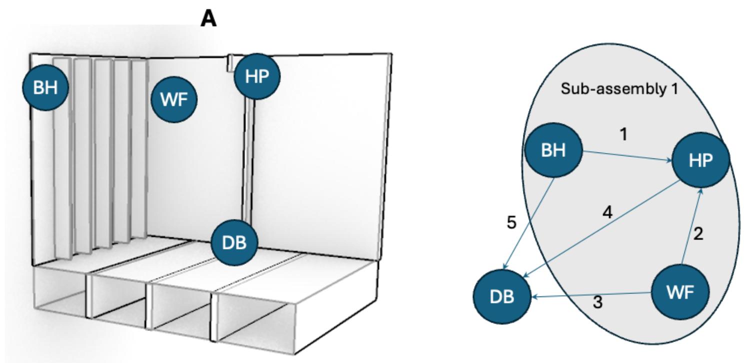


Figure 7. Block A. The block shown in this figure contains 4 sub-assemblies: BH = bulkhead, HP = hull plate, DB = double bottom, and WF = web frame. At first, the sub-assemblies BH, HP, and WF are positioned in such a way that welding can be completed underhand to create sub-assembly 1. This is done by welding BH to HP (seq.nr. 1), and then the WF is welded to HP (seq.nr. 2). After sub-assembly 1 is created in this way, it is positioned above DB to weld the connections between WF and DB (seq.nr. 3), and next the connection between HP and DB (seq.nr. 4), and next the connection between BH and DB (seq.nr 5).

The graph, as shown in Figure 7, illustrates the creation of a sub-assembly consisting of the bulkhead, hull plate, and web frame, which is then linked to the double bottom. The construction sequence shown in the graph is just one of the potential ways to assemble product A. The decision to create a distinct sub-assembly typically stems from the need to enhance the efficiency of the operations conducted and is often based on experience.

Figure 7 does not include details of transportation, lifting, or turning actions of sub-assemblies. To fully understand how the product is constructed, it is essential to consider the comprehensive activity sequence list. To illustrate this, the activity sequence list for block A, corresponding with the graph from Figure 7, is given in Table 5.

Table 5. A possible activity sequence list for block A.

Assembly Sequence A	From	To	Activities
Sequence Sub-assembly 1			
1	BH	HP	Positioning, 1 × Welding
2	WF	HP	Positioning, 2 × Welding
Sub-assembly 1 to DB			
3	WF	DB	Positioning, 2 × Welding
4	HP	DB	1 × Welding
5	BH	DB	1 × Welding

The description thus far outlines how to represent product and process information within a graph model, demonstrating its applicability in the pre-contract phase when not every detail is known yet. In the next section, we describe the use of the graph representation in the cost estimation process in the pre-contract phase.

3.6. Estimating Hours Based on Similarities

The concept introduced here, like conventional methods, leverages historical performance information. It implicitly assumes that the construction processes used in the past are relevant and transferable for estimating the costs of future projects. Therefore, estimating costs involves identifying similarities between the graphs of past projects and those planned for future construction, as illustrated in Figure 8. Identifying similarities means that a similarity analysis must be performed to estimate the hours needed to construct the products.

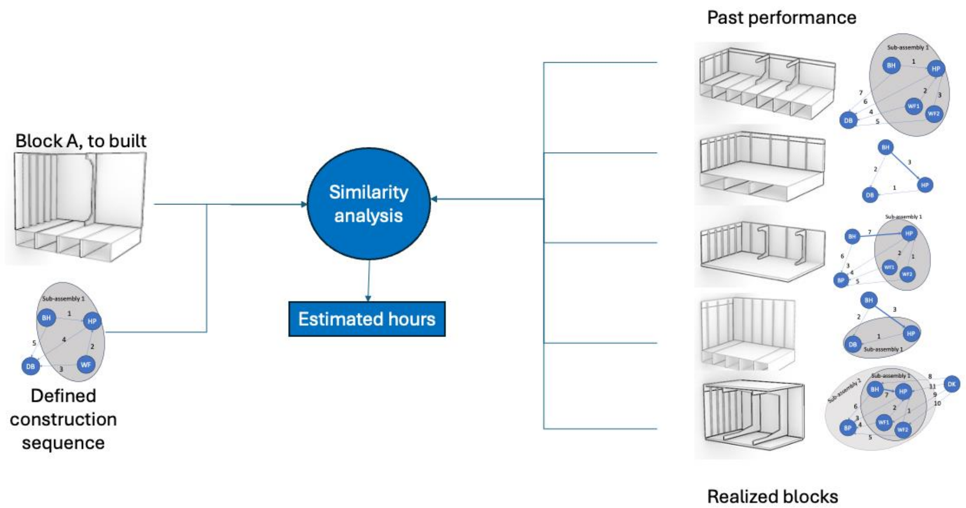


Figure 8. Similarity analysis. In this figure, block A represents the product to be built. To construct block A, a construction sequence is defined and represented in a graph model. This graph model is then compared with a set of graph models from the past performance database containing graph models of realized blocks, including construction sequences and spend person-hours. This comparison aims to find a realized block that is the most similar block to estimate the person-hours needed for the realization of block A.

To start the similarity analysis process, the construction sequence of the product to build must be defined. In our description, we take a block as determined during the pre-contact phase in the block division plan. An assembly process must be created for this block, either through a structured approach (as suggested by Zhong [33]) or through other methods. The result must be translated into a labeled, directed graph representing the product and construction information for the block under investigation. Next, the similarity analysis can be performed. Several approaches are available for conducting a graph similarity analysis. Here, we provide a brief description of two such approaches. These approaches illustrate how comparisons can be made to identify the most similar graph.

The first approach is based on error-tolerant graph matching. The core idea of error-tolerant graph matching is to calculate the cost of mapping a graph of a realized block onto the graph of the block to be built, aiming to find the realized block that minimizes the graph edit distance (GED). The GED expresses the "overall cost" of the mapping. A theoretical foundation of graph edit distance is given by Riesen [34]. GED poses a challenge as it is known to be an NP-hard problem [35,36]. NP-hard means that the solution of such a problem requires non-polynomial time to solve, making it computationally demanding and therefore hindering its practical applications.

Another approach to perform a similarity analysis is based on neural network principles. To be more specific, here, we employ graph neural networks to represent the information. Applying GNNs for similarity analysis aims to identify patterns, relationships, or clusters based on structural and characteristic similarities in the data. GNNs learn embeddings for each node in the graph by capturing both the node's feature information and its connectivity patterns with neighboring nodes. These learned node embeddings encode structural and contextual information about each node. Once node embeddings are obtained, the similarity between nodes can be computed. Ontañón [37] presents an overview of the different similarity analysis approaches using graphs and points out directions for future work.

The introduced concept focuses on cost estimation for Engineering-To-Order (ETO) in the pre-contract phase. It utilizes the available information in the pre-contract phase, where the design details are more generic. However, the method is designed to be independent of the level of information available. This means it remains applicable, even if more detailed information becomes available, like the number of stiffeners and other smaller elements. Consequently, the concept can be used effectively in subsequent phases beyond the pre-contract phase. When applied in other phases, the more detailed available design and construction process information must be included in the similarity analysis. This means that the same level of information must be made available between the block to be built and the past performance information to perform the similarity analysis. To achieve the connection between the different phases, making subsequent calculations possible, the detailing of information must be performed in a traceable and reproducible manner. The possibilities of aggregating and detailing information in a traceable and reproducible manner require further research. Implementing a uniform cost estimation method for all phases allows the return to the original cost estimate as design or construction processes take place.

4. Discussion

The presented concept makes it necessary to enrich the cost estimation information with construction process data. The development and organization of construction process information already starts in the pre-contract phase. For this, the knowledge of experts is essential, not the least, because Engineering-To-Order projects require expertise to successfully realize innovations and new requirements related to safety regulations, energy conservation, and environmental protection. Defining the construction process requires significant manual effort from domain experts. Considering that the pre-contract phase is particularly characterized by time pressure and deadlines, efficiency improvement is required for the creation of required construction process information. In the research of Iwańkowicz [31], an approach is described to reduce the manual effort of experts in defining assembly sequences. Iwańkowicz's [31] research led to the development of an intelligent hybrid sequencing method for structure assembly that uses fuzzy clustering, case-based reasoning, and evolutionary optimization.

The concept necessitates integrating past performance information, including time allocation, technology, resources utilized, and, most importantly, the specific process employed for product development. Broadly speaking, this implies that shipyards need to augment their past performance information by incorporating details on the technologies employed and the utilized construction process. For most yards, that will be a challenge, not in the least, because current available past performance information is often incomplete and not always available in well-structured and maintained databases. Improving past performance information, including process details, requires a reengineering process that involves substantial investments. An investment that can be compared to the global shift from tangible assets to intangible assets [38]. Investments that have been shown to have the potential to increase productivity are based on data-driven decisions. The concept could improve the use of collected data in shipyards, resulting in more accurate cost predictions. However, its effectiveness has yet to be demonstrated.

However, before investing in enriching past performance information, research is needed to demonstrate that the proposed cost estimation method improves the reliability and accuracy of the estimates, especially for ETO projects.

The similarity analysis is a crucial element in the described concept. When performing similarity analyses rooted in neural network principles, it is essential to identify an appropriate similarity measure before optimal similarity can be determined. To obtain the similarity measure, initial learning of the solution is required. The review of Ma et al. [39] gives an overview of the existing literature on deep graph similarity learning. The conclusion is that the literature on labeled and directed graphs for similarity learning is scarce. This implies that applying neural network principles for similarity analysis will be challenging, particularly given the limited volume of available learning data.

Structuring and enriching past performance information with construction process information opens the possibility of successfully performing optimization analyses on assembling sequence planning [40].

The principle of process patterns for cost estimates is inherently versatile and can be effectively extended to various other processes, such as piping and scaffolding. However, the successful application of these distinct processes necessitates thoughtful consideration and deliberate implementation choices.

The presented concept is a generic concept tailored to Engineering-To-Order projects. If the concept is adapted to the large-scale production environments of Asia, it should perform equally as well. However, in these environments, the production process is often fixed in more detail and compared between comparable designs, a more basic/statistical method would thus be sufficient. When considered more broadly, this concept could indicate a potential to improve operational efficiency, enable effective cost control, and support continuous improvement practices without requiring simulation or other more extensive methods of investigation.

It is essential that the introduced concept offers sufficient advantages. It is, therefore, to be demonstrated in future research whether the introduced concept will lead in practice to more accurate and reliable cost estimates.

5. Conclusions and Future Research Perspectives

The contribution of this work is a concept for the integration of the construction process in cost estimations, which shows a promising route to improve the reliability and reproducibility of estimations in the pre-contract phase for Engineering-To-Order (ETO) projects. The presented method is detail-independent. This means there is one cost estimation method for all phases of the production (ship, block, sub-section), resulting in a single cost control process and introducing the possibility of tracing back changes in design and process to the original cost estimate, and new estimations can be made.

The information representation method proposed, based on labeled directed graphs, facilitates incorporating changes in technologies and resources in a structured manner, enhancing adaptability and cost management within the project.

A conditional aspect of the introduced method is the availability of past performance information containing construction process information. To create such an information base will be challenging and asks, in most cases, a significant investment. An investment comparable to the global shift from investments from tangible to intangible assets attempts to improve business performance. This means that the proposed method would benefit from advancements in data management, particularly focusing on the precise quantification of process data. This would enable shipyards to leverage their historical performance data more effectively, leading to more accurate cost predictions.

Future research to demonstrate that the proposed method improves cost estimations for Engineering-To-Order projects in the pre-contract phase starts with how to perform the similarity analysis process and how to estimate the person-hours. The subsequent research phase involves investigating the influence of integrating the construction process into cost estimation accuracy.

Author Contributions: Conceptualization, G.A.; investigation, G.A.; methodology, and writing the original draft, G.A.; conceptualization, J.P.; supervision, J.P.; writing—review and editing, J.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Alfnes, E.; Gosling, J.; Naim, M.; Dreyer, H.C. Exploring systemic factors creating uncertainty in complex engineer-to-order supply chains: Case studies from Norwegian shipbuilding first tier suppliers. *Int. J. Prod. Econ.* **2021**, *240*, 108211. [CrossRef]
2. Bruce, G.; Bruce, G. Ship Project Strategy. In *Shipbuilding Management*; Springer: Singapore, 2021; pp. 63–74. [CrossRef]
3. ECORYS SCS Group. Study on Competitiveness of the European Shipbuilding Industry Within the Framework Contract of Sectoral Competitiveness Studies -ENTR/06/054 [pdf]. Rotterdam 2009; p. 239. Available online: https://ec.europa.eu/docsroom/documents/10506/attachments/1/translations/en/renditions/native (accessed on 29 January 2024).
4. Kamola-Cieślik, M. Changes in the global shipbuilding industry on the examples of selected states worldwide in the 21st century. *Eur. Res. Stud. J.* **2021**, *24*, 98–112. [CrossRef]
5. Gasparotti, C.; Rusu, E. An overview on the shipbuilding market in current period and forecast. *EuroEconomica* **2018**, *37*, 254–271.
6. Stanić, V.; Hadjina, M.; Fafandjel, N.; Matulja, T. Toward shipbuilding 4.0-an industry 4.0 changing the face of the shipbuilding industry. *Brodogr. Teor. I Praksa Brodogr. I Pomor. Teh.* **2018**, *69*, 111–128. [CrossRef]
7. Alblas, G.; Pruijn, J. Are current shipbuilding cost estimation methods ready for a sustainable future? A literature review of cost estimation methods and challenges. *Int. Shipbuild. Prog.* **2024**, *71*, 3–28. [CrossRef]
8. Hur, M.; Lee, S.-k.; Kim, B.; Cho, S.; Lee, D.; Lee, D. A study on the man-hour prediction system for shipbuilding. *J. Intell. Manuf.* **2015**, *26*, 1267–1279. [CrossRef]
9. Caprace, J.D. Cost Effectiveness and Complexity Assessment in Ship Design within a Concurrent Engineering and "Design for X" Framework. Ph.D. Thesis, Université de Liège, Liège, Belgium, 2010.
10. Bole, M. (Ed.) Cost assessment at concept stage design using parametrically generated production product models. In Proceedings of the International Conference on Computer Applications in Shipbuilding, Portsmouth, UK, 18–20 September 2007.
11. Evans, D.K.; Lanham, J.D.; Marsh, R. Cost estimation method selection: Matching user requirements and knowledge availability to methods. In *International Cost Engineering Council: The Singapore Institute of Surveyors and Valuers*; Systems Engineering and Estimation for Decision Support (SEEDS) Group, University of West of England: Bristol, UK, 2006.
12. Shetelig, H. Shipbuilding Cost Estimation: Parametric Approach. Master's Thesis, Institutt for Marin Teknikk, Trondheim, Norway, 2014.
13. Leal, M.; Gordo, J.M. Hull's manufacturing cost structure. *Brodogr. Teor. I Praksa Brodogr. I Pomor. Teh.* **2017**, *68*, 1–24. [CrossRef]
14. Wierda, L.S. Linking design, process planning and cost information by feature-based modelling. *J. Eng. Des.* **1991**, *2*, 3–19. [CrossRef]
15. Lin, C.-K.; Shaw, H.-J. Feature-based estimation of preliminary costs in shipbuilding. *Ocean Eng.* **2017**, *144*, 305–319. [CrossRef]

16. Kolodner, J.L. An introduction to case-based reasoning. *Artif. Intell. Rev.* **1992**, *6*, 3–34. [CrossRef]
17. Relich, M.; Pawlewski, P. A case-based reasoning approach to cost estimation of new product development. *Neurocomputing* **2018**, *272*, 40–45. [CrossRef]
18. Shehab, E.; Abdalla, H. An intelligent knowledge-based system for product cost modelling. *Int. J. Adv. Manuf. Technol.* **2002**, *19*, 49–65. [CrossRef]
19. Dixit, V.; Chaudhuri, A.; Srivastava, R.K. Assessing value of customer involvement in engineered-to-order shipbuilding projects using fuzzy set and rough set theories. *Int. J. Prod. Res.* **2019**, *57*, 6943–6962. [CrossRef]
20. Bode, J. Neural networks for cost estimation. *Cost Eng.* **1998**, *40*, 25.
21. Liu, B.; Li, R.; Wang, J.; Liu, Y.; Li, S. Subassembly Partition of Hull Block Based on Two-Dimensional PSO Algorithm. *J. Mar. Sci. Eng.* **2023**, *11*, 1006. [CrossRef]
22. Anand Alagamanna, A.; Juneja, S.S. Man-Hour Estimations in ETO: A Case Study Involving the Use of Regression to Estimate Man-Hours in an ETO Environment. Master's Thesis, Uppsala Universitet, Uppsala, Sweden, 2020.
23. Tayefeh Hashemi, S.; Ebadati, O.M.; Kaur, H. Cost estimation and prediction in construction projects: A systematic review on machine learning techniques. *SN Appl. Sci.* **2020**, *2*, 1703. [CrossRef]
24. Ossa Mesa, J.E. Cost Estimation of Construction Projects Using 5D BIM: Integrating the Cost Engineer in BIM-based Processes Through Activity Theory. Master's Thesis, Delft University of Technoloy, Delft, The Netherlands, 2021.
25. Castro Miranda, S.L.; Del Rey Castillo, E.; Gonzalez, V.; Adafin, J. Predictive Analytics for Early-Stage Construction Costs Estimation. *Buildings* **2022**, *12*, 1043. [CrossRef]
26. Bruce, G. *Shipbuilding Management*; Springer Nature: Berlin/Heidelberg, Germany, 2020. [CrossRef]
27. Goo, B.; Chung, H.; Han, S. Layered discrete event system specification for a ship production scheduling model. *Simul. Model. Pract. Theory* **2019**, *96*, 101934. [CrossRef]
28. Steinhauer, D.; Sikorra, J.N.; Haux, M.A.; Friedewald, A.; Lödding, H. Processing incomplete data for simulation-based production planning in shipbuilding. *J. Simul.* **2017**, *11*, 30–37. [CrossRef]
29. Qu, S.; Jiang, Z.; Tao, N. An integrated method for block assembly sequence planning in shipbuilding. *Int. J. Adv. Manuf. Technol.* **2013**, *69*, 1123–1135. [CrossRef]
30. Fernando, N.; TA, K.D.; Zhang, H. An artificial neural network (ANN) approach for early cost estimation of concrete bridge systems in developing countries: The case of Sri Lanka. *J. Financ. Manag. Prop. Constr.* **2024**, *29*, 23–51. [CrossRef]
31. Iwańkowicz, R. A multi-case-based assembly management method for the shipbuilding industry. *Pol. Marit. Res.* **2021**, *28*, 27–35. [CrossRef]
32. Bunke, H. Graph matching: Theoretical foundations, algorithms, and applications. In Proceedings of the Vision Interface, Montreal, QC, Canada, 14–17 May 2020; pp. 82–88.
33. Zhong, Y.; Xue, K.; Shi, D. Assembly unit partitioning for hull structure in shipbuilding. *Comput. Aided Des.* **2013**, *45*, 1630–1638. [CrossRef]
34. Riesen, K. *Structural Pattern Recognition with Graph Edit Distance*; Springer: Cham, Switzerland, 2015. [CrossRef]
35. Nair, A.; Roy, A.; Meinke, K. funcgnn: A graph neural network approach to program similarity. In Proceedings of the 14th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM), Bari, Italy, 8–9 October 2020; pp. 1–11. [CrossRef]
36. Zeng, Z.; Tung, A.K.; Wang, J.; Feng, J.; Zhou, L. Comparing stars: On approximating graph edit distance. In Proceedings of the VLDB '09, Lyon, France, 24-28 August 2009; pp. 25–36.
37. Ontañón, S. An overview of distance and similarity functions for structured data. *Artif. Intell. Rev.* **2020**, *53*, 5309–5351. [CrossRef]
38. Hazan, E.; Smit, P.S.; Woetzel, A.J.; Biljana Cvetanovski, S.; Mekala Krishnan, L.; Brian Gregg, B. *Getting Tangible About Intangible*; McKinsey Global Institute: Chicago, IL, USA, 2021; p. 40.
39. Ma, G.; Ahmed, N.K.; Willke, T.L.; Yu, P.S. Deep graph similarity learning: A survey. *Data Min. Knowl. Discov.* **2021**, *35*, 688–725. [CrossRef]
40. Shi, X.; Tian, X.; Gu, J.; Wang, G.; Zhao, D.; Ma, L. A hybrid approach of case-and rule-based reasoning to assembly sequence planning. *Int. J. Adv. Manuf. Technol.* **2023**, *127*, 221–236. [CrossRef]