Automation in Construction 173 (2025) 106066

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

A data-driven and knowledge-based decision support system for optimized construction planning and control

Moslem Sheikhkhoshkar [a,b,*], Hind Bril El-Haouzi [c], Alexis Aubry [c], Farook Hamzeh [d], Farzad Rahimian [e,**]

[a] CNRS, CRAN, Université de Lorraine, Epinal, France
[b] University of Greenwich, London, UK
[c] CNRS, CRAN, Université de Lorraine, Nancy, France
[d] Dept. of Civil and Environmental Engineering, University of Alberta, Edmonton, Canada
[e] School of Computing, Engineering and Digital Technologies, Teesside University, United Kingdom

ARTICLE INFO

Keywords:
Decision support system
Planning and control system
Rough set theory
Knowledge repository
Mathematical model

ABSTRACT

Despite the use of various construction planning and control systems, no prior data-driven and knowledge-based system provides optimized solutions based on specific project team needs and applications. This paper presents a data-driven and knowledge-based decision support system that utilizes a knowledge database constructed from experts' experience and proposes multi-level and integrated systems for planning and control of construction projects. A mixed-method approach gathers data from industry professionals, develops a knowledge repository based on Rough Set Theory (RST), launches an inference engine using the Pyke package, and integrates these insights into a decision support system optimized by a multi-objective mathematical model. The developed system considers the functional requirements of the project team and suggests an optimized and fit-for-purpose planning and control system. To demonstrate its practicality, it applies to a real-world renovation project. This paper contributes to enhancing systematic and data-driven decision-making for planning and control systems based on expert knowledge and the specific needs of the project team.

1. Introduction

Project planning and control stands as a fundamental element of construction project management. This multi-functional domain handles a broad spectrum of decision-making challenges. These functions not only ensure that all project activities are meticulously planned, sequenced, and resourced to promote a seamless operational flow but also deal with analyzing deviations and delays, managing constraints and commitments, and fostering collaboration and communication among project team members [37,42]. To tackle these decision-making challenges, a wide variety of methods, techniques, and tools have been devised over decades. The critical path method (CPM) represents one of the earliest methods conceived for project planning, scheduling and control [29]. While it's still widely used and often required by project owners for scheduling needs, there are several concerns with this approach, including a tendency to create overly detailed schedules even when project details are uncertain, failure to encourage collaboration during the planning phase, oversight of non-critical tasks, and limitations in monitoring resource allocation [34]. Given the complexity of construction projects and the engagement of numerous stakeholders, the Architecture, Engineering and Construction (AEC) industry has recognized the need for more effective project planning and control methods. In this regard, the introduction and adoption of lean-driven planning and control methods, such as the last planner system (LPS) [8,19], location-based management system (LBMS) [40], and takt time planning (TTP) [15], represented a significant paradigm shift in project planning and control domain. Emerging building information modelling (BIM) and its integration with industry 4.0 technologies was another significant advancement in project planning and control [32,52]. While these developments have notably improved the project planning and control field, the scholars identified certain shortcomings in the independent implementation of these methods [11]. Standing alone, each system is strong in some functions but requires improvements in others [34]. Therefore, several academics have attempted to integrate these planning

* Corresponding author at: CNRS, CRAN, Université de Lorraine, Epinal, France.
** Corresponding author.
E-mail addresses: moslem.sheikhkhoshkar@univ-lorraine.fr (M. Sheikhkhoshkar), f.rahimian@tees.ac.uk (F. Rahimian).

https://doi.org/10.1016/j.autcon.2025.106066
Received 29 June 2024; Received in revised form 7 February 2025; Accepted 12 February 2025
Available online 24 February 2025

methods, aiming to establish a comprehensive approach that effectively mitigates their limitations. In this context, Olivieri, et al. [34] synthesized CPM, LBMS, and LPS to enhance the modelling of workflow dynamics. This integrated approach was intended to facilitate managerial analysis and communication regarding delays, as well as to inform decision-making processes regarding the best strategies for the critical path. Rashidi, et al. [38] focused on employing a virtual reality (VR) environment to improve 4D-BIM-based construction planning. They uncovered the potential enhancements in construction planning, particularly in spatial understanding, spatial-temporal conflict resolution, stakeholder collaboration, training and education, and safety management, through integrating these innovative technologies. Liu, et al. [30] investigated a novel experimental tool developed to examine the social mechanisms of LPS implementation by utilizing immersive virtual reality (IVR) gaming technology. Additionally, it has been proposed to utilize the LPS along with Location-Based Planning (LBP) methods to better organize work sequences across different project locations, aiming to streamline workflows further [7].

These developments go beyond just improving planning methods. Endeavors have additionally been directed towards establishing and utilizing control metrics for the continual monitoring and analysis of project performance, efficiency, and other dynamic aspects, including the efficiency of resource allocation, quality of the construction flow, constraints removals, labour productivity, and quality of the commitments [41,45]. This guarantees that projects are not solely meticulously planned but are also regularly evaluated and updated based on ongoing actual data [20].

Despite the widespread implementation of these individual and integrated planning methods and control metrics globally, the choice to prefer one planning method over another, or to integrate multiple methods, was driven by the need to address their respective shortcomings and leverage their strengths [11]. Selecting the most effective planning and control system based on the specific needs of the project team has often been overlooked in both literature and practice. Therefore, there is a need for a tool that can propose the planning and control approached based on project-specific requirements. To fill this academic and practical gap, this research aims to develop a data-driven and knowledge-based decision support system (DSS) that suggests multi-level and integrated project planning and control systems for construction projects. To achieve this aim, the research outlines the following objectives:

1- Capture and extract the knowledge and experiences of professionals in project planning and control to construct a knowledge repository
2- Develop a decision support system to use the knowledge database and suggest planning and control systems based on the project team's requirements
3- Develop and execute a mathematical model to optimize the proposed solutions by DSS

To achieve the objectives, this research captures and analyzes the knowledge and experiences of domain experts to build a knowledge repository for the DSS. An inference engine is then launched to recommend suitable planning and control systems, taking into account both the knowledge database and the project team's functional requirements. A mathematical model is subsequently developed to optimize the solutions proposed by the DSS. The practicality and usability of the system are evaluated through a case study and feedback from experts.

The paper is organized as follows: Section 2 provides the research methodology. Section 3 includes the analysis and outcomes derived from data collection efforts, extending to mathematical modelling and validation results. In section 4, the paper delves into research discussion and implications. Finally, section 5 highlights the conclusion, limitations, and avenues for future investigation.

2. Knowledge-based systems in construction management

The rapid advancement of artificial intelligence (AI) and its diverse applications have significantly enhanced decision-making processes, particularly in construction management. Given that the construction sector heavily relies on expert experience, best practices, and lessons learned, knowledge-based systems play a crucial role in capturing and preserving tacit knowledge [12]. This is especially important in mitigating the loss of critical insights due to the industry's high staff turnover. In this context, several studies have proposed knowledge-based systems to address various challenges in construction management. For instance, Dikmen, et al. [12] developed a rule-based decision support system for risk and complexity assessment in construction projects. Their mixed-method research approach involved semi-structured interviews with 18 senior project managers to explore the risk-complexity relationship and inform the knowledge framework underlying the DSS. Similarly, Okudan, et al. [33] introduced a case-based reasoning approach for a knowledge-driven risk management tool tailored to construction projects. Hwang, et al. [24] presented a knowledge-based DSS for prefabricated prefinished volumetric construction which followed a comprehensive literature review, pilot interviews with industry experts, and structured questionnaires to collect the required data and build a knowledge database. Akbari, et al. [3] employed a rough set-based fuzzy inference system to create a DSS for dynamically assessing the sustainable success of infrastructure projects.

Such advancements have also had a notable impact on construction planning and control, where systems have been developed for various functions such as schedule updating, schedule analysis, time prediction, activity duration estimation, cost estimation, and project network generation. For instance, Hendrickson, et al. [22] pioneered a knowledge-intensive expert system for generating project activity networks, cost estimates, and schedules, including defining activities, specifying precedences, selecting technologies, and estimating durations and costs. More recently, Jahr and Borrmann [25] proposed a rule-based knowledge inference system that supports semi-automated site equipment planning using data from building information models and work schedules. Additionally, Hajdasz [18] introduced an intelligent decision support tool for flexible site management in repetitive projects, while Mohamed [31] offered a knowledge-based approach for analyzing factors that influence project duration, generating both normal and productivity-adjusted schedules.

Despite the extensive developments in decision support systems and knowledge management tools within construction management—especially in project planning and control—there remains a gap in capturing expert knowledge on the applicabilities and functionalities of various planning and control systems. Addressing this gap, the present study aims to develop a data-driven and knowledge-based DSS designed specifically for the preconstruction phase, which can recommend fit-for-purpose, multi-level, and integrated planning and control systems tailored to the unique needs of construction projects.

3. Adopted research methodology

This paper's objectives were achieved through a mixed-method approach, as demonstrated in Fig. 1. This methodology encompasses three primary phases: data collection and preprocessing, expert system development, and results optimization. During the data collection phase, semi-structured interviews and surveys were conducted by domain experts to gather the requisite data for further analysis. Subsequently, in the expert system development phase, a rule-based knowledge repository was constructed. Following this, an inference engine was launched using the forward chaining method and the Pyke which is a knowledge engine in Python, aimed at recommending the most suitable planning methods, control metrics, and schedule levels based on the project team's requirements. Finally, a multi-objective mathematical model was formulated to optimize the suggestions provided by the

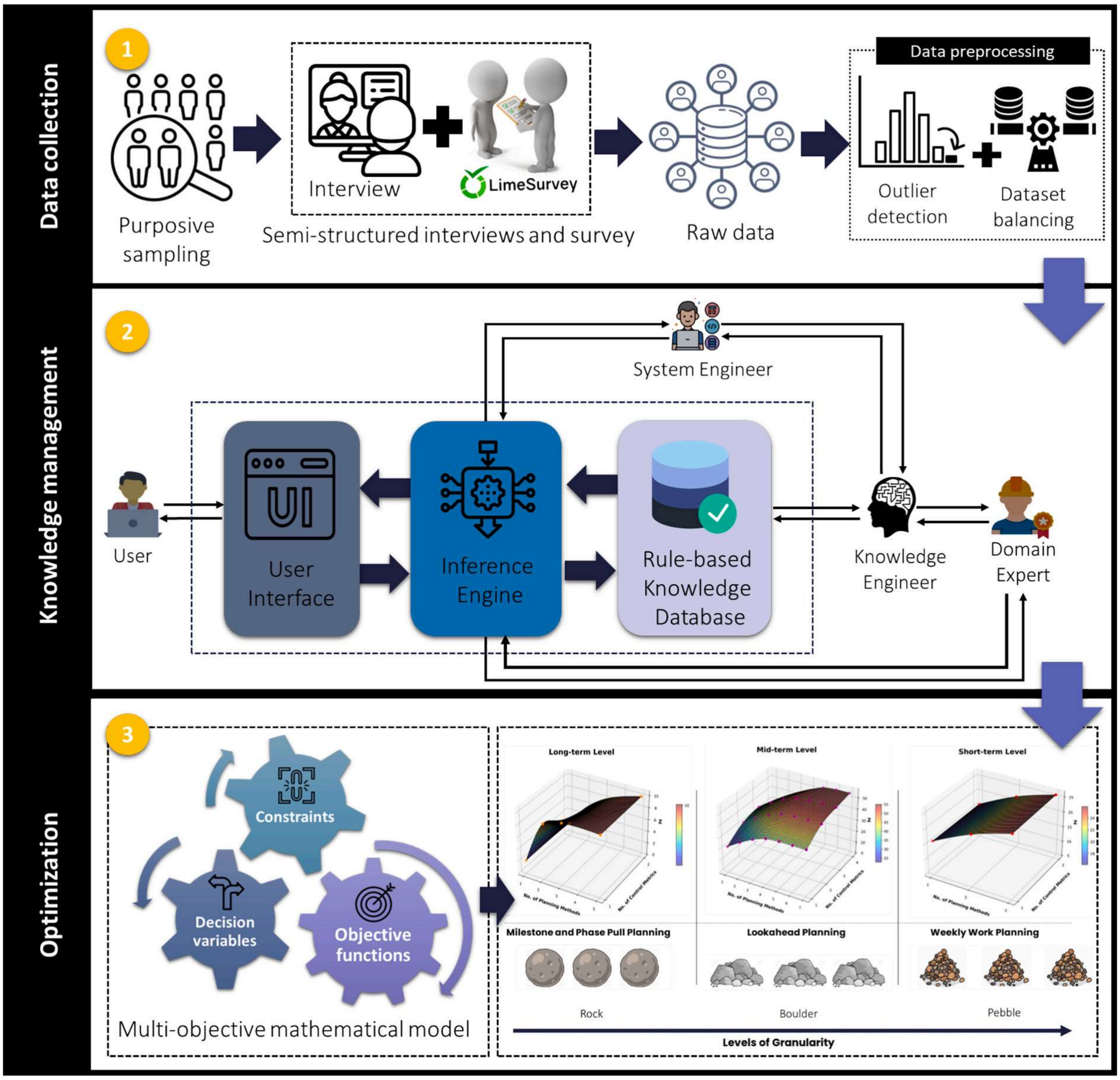


Fig. 1. Mixed-method adopted methodology.

expert system and propose an enhanced multi-level planning and control system. The following subsections provide detailed explanations for each aspect of the research methodology.

3.1. Data collection and preprocessing

3.1.1. Sampling and data collection approaches

This study used purposive sampling to select the population members to participate in the data collection process. Given the constrained availability of domain experts versed in various project planning and control systems, as well as the need to select a sample of individuals with diverse knowledge and experience in this domain, purposive sampling was deemed appropriate for the data collection process within the scope of this study. Suri [49] and Hennink, et al. [23] clarified the method's validity by highlighting its ability to allow researchers to select participants who have specific expertise relevant to the study topic.

Regarding the data collection method, a semi-structured interview was chosen as the primary approach to collect the necessary data for building a knowledge repository for project planning and control systems. In addition, a survey was designed to provide a consistent structure in the interview process, standardize the topics discussed, as well as collect the required quantitative data for building a knowledge database.

3.1.2. Data preprocessing

Following the data understanding, it was observed that the initial dataset exhibited an imbalance due to variations in expertise among participants in different planning methods. For instance, while all participants responded to the last planner system inquiries, only 50 % of participants responded to inquiries regarding the critical chain project management method. Additionally, certain responses displayed outliers, further challenging the integrity of the dataset. As a result, two preprocessing steps were deemed imperative to enhance data quality: outlier detection and addressing the imbalances inherent in the dataset.

The interquartile range (IQR) method was selected for managing the outliers. In addressing imbalanced datasets, given that this study employed purposive sampling for data collection, which inherently limits the dataset, adopting undersampling as a strategy was deemed suboptimal due to the potential data loss. Therefore, oversampling becomes the preferred approach. To do this, the Synthetic Minority Oversampling Technique (SMOTE) was employed to handle the imbalance dataset in this study. SMOTE is a powerful approach utilized for mitigating class imbalance in datasets and has demonstrated remarkable performance in a variety of applications [21]. SMOTE enhances data representation by synthesizing samples from the minority class rather than just duplicating existing records. This technique operates by identifying the nearest neighbors of a minority instance and interpolating between them to create new synthetic samples [21]. Such a procedure balances the dataset by increasing the diversity of the minority class, reducing the risk of overfitting, and enhancing model performance in scenarios where data imbalance would make accurate forecasting more difficult.

3.1.3. Reliability and consistency of the data

The consistency and reliability of the survey data were assessed using Cronbach's α coefficient method [10]. Cronbach's Alpha is a commonly used measure to assess internal consistency, which shows how effectively several survey questions evaluate the same concept. A higher Cronbach's Alpha value, typically above 0.7, indicates a higher level of reliability in the collected data. The formula for Cronbach's Alpha is:

$$\alpha = \frac{n}{n-1}\left(1 - \frac{\sum_{i=1}^{n}\sigma_i^2}{\sigma_T^2}\right) \tag{1}$$

where n is the number of survey items, σ_i^2 is the variance of each individual item, and σ_T^2 is the total variance of all items combined.

3.2. Knowledge management and expert system development

An expert system is a kind of artificial intelligence program which leverages either a predetermined set of rules or a repository of human expertise, known as a knowledge base, to replicate the decision-making ability of a human expert [1]. These systems are designed to address complex problems by reasoning a knowledge repository, typically encoded in the conditional statements form (if-then rules), rather than relying on conventional procedural programming paradigms. The main components of an expert system include 1) a knowledge database, 2) an inference engine, and 3) a user interface [2]. The outline of the implemented steps to initialize the expert system is depicted in Fig. 2 and discussed in more detail in the following subsections.

3.2.1. Building knowledge repository

A knowledge repository is a centralized system designed to collect, manage, and share valuable insights, best practices, and expert knowledge. It organizes and stores information in a structured manner, making it easily accessible for users who need it to make informed decisions [13]. This repository ensures knowledge continuity, preventing loss when team members leave, and fostering a culture of continuous learning and improvement. Within the construction sector, a knowledge repository holds significant value because the sector relies heavily on experience, best practices, and lessons learned. Such a repository enables project teams to quickly access essential information, reducing the chances of repeating past mistakes, streamlining workflows, and enhancing project outcomes. As expert insights often come with uncertainties due to subjective interpretation, incomplete data, and varying contexts, it's crucial to apply a method that effectively builds the knowledge database while addressing these uncertainties. Rough Set Theory (RST) effectively handles uncertainties by enabling data analysis with imprecise boundaries [3,28]. Thus, this method was selected to establish the knowledge database for the project planning and control system. The following subsections explore the details of the RST approach.

Rule Generation Using Rough Set Theory (RST)

Rough set theory, proposed by [35], is a key component of interpretable machine learning. It plays a vital role in artificial intelligence research, particularly in the classification, knowledge discovery, data mining, and pattern recognition domains [17]. The core principles and implementation processes of RST are outlined below:

Concept 1: Information system

In RST, a dataset is structured within an information system table wherein each column represents an attribute, such as a variable, and each row corresponds to an object or case. Formally, a dataset comprises a pair denoted as $S = (U, A)$, where U is a finite set of objects and A is a finite set of attributes. The attribute sets are split into condition attributes (C) and decision attributes (D). Thus, any information system structured as $S = (U, A = C \cup D)$ becomes a decision system, where $D = d$ and $d \notin C$ represent the decision attribute. Table 1 presents an example of an information system table that includes five functionalities of planning and control systems as condition attributes (F1 to F5), along with two planning methods, LPS and 4DBIM, as decision attributes. The table also incorporates ten objects (O1 to O10), representing expert

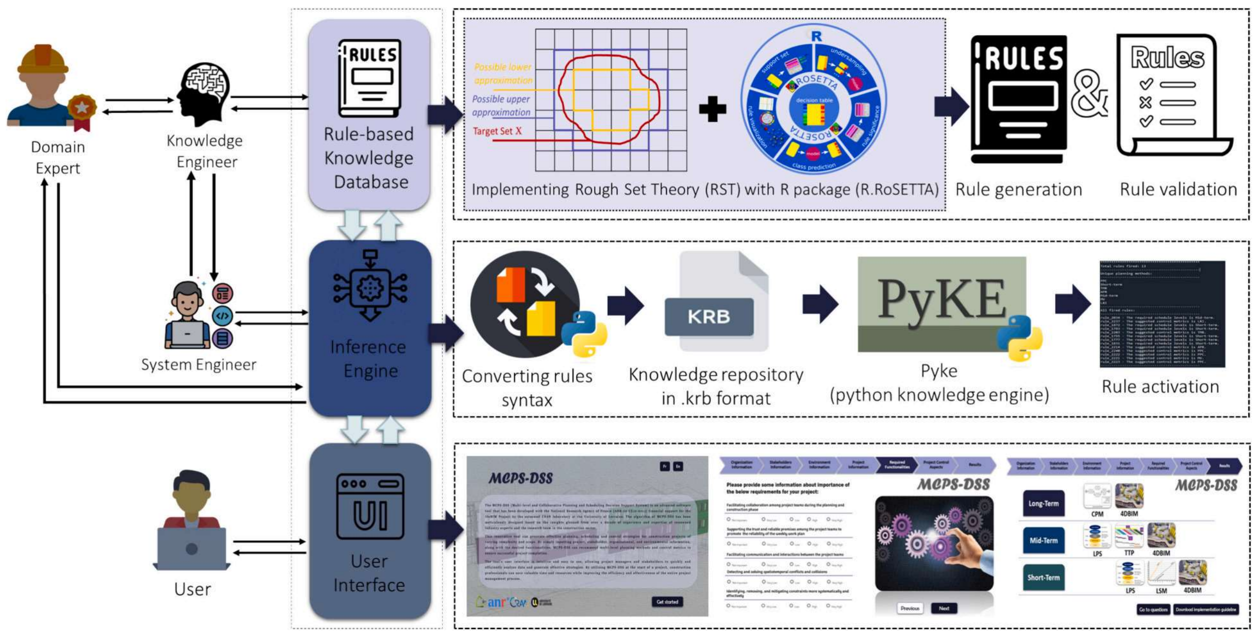


Fig. 2. Overview of the conducted steps to initialize the expert system.

Table 1
Example of an information system table.

Objects	F1	F2	F3	F4	F5	Planning methods
O1	1	0	3	4	4	4DBIM
O2	3	2	4	1	3	LPS
O3	4	1	3	0	2	LPS
O4	0	1	4	3	3	4DBIM
O5	0	1	3	3	3	LPS
O6	3	1	4	1	4	LPS
O7	4	0	3	0	3	LPS
O8	1	0	3	4	4	4DBIM
O9	3	1	4	1	4	LPS
O10	0	1	3	3	3	4DBIM

opinions on the level of support of each functionality by the planning methods. For instance, O1 indicates Expert 1's view on the level of support provided by the 4DBIM for functionalities F1 to F5. The values 0 (without support) to 4 (very high support) illustrate a Likert scale for the level of support of each functionality by the planning methods.

Concept 2: Indiscernible relation

For any subset, $B \subseteq C$, an equivalence relation denoted as $IND(B)$ is defined in Eq. (2), known as the B-indiscernibility relation.

$$IND(B) = \left\{(x, x') \in U^2 : \forall a \in B, a(x) = a(x')\right\} \quad (2)$$

If $(x, x') \in IND(B)$, then x and x' are objects that are indiscernible based on the attributes in B.

In the information system presented in Table 1, objects O6 and O9 are indiscernible concerning the recorded attributes and therefore comprise an equivalence class. Similarly, objects O5 and O10 also constitute an equivalence class. However, these objects fall into different decision classes (LPS, 4DBIM). The information system can be summarised in terms of the following equivalence classes:

Concept 3: Lower and upper approximation

Two crisp sets referred to as lower and upper approximations of a given set X concerning $IND(B)$ are defined in the approximation space.

$$\underline{B}X = \{x : IND\ (B) \subseteq X\} \quad (3)$$

$$\overline{B}X = \{x : IND\ (B) \cap X \neq \varnothing\} \quad (4)$$

$\underline{B}X$ and $\overline{B}X$ represent the B-lower and B-upper approximations of X, respectively. The B-lower approximation includes objects that are certain to be in X, whereas the B-upper approximation includes objects that may be in X. The difference is known as a boundary of X in U.

$$BN(X) = \overline{B}X - \underline{B}X \quad (5)$$

The set X is called rough if $BN(X) \neq \varnothing$ and crisp otherwise.

Considering the example, the decision classes of LPS and 4DBIM objects are rough sets because they cannot be precisely defined using a single set of equivalence classes. Instead, they are characterized by upper and lower approximations. For instance, the decision class LPS can be outlined through the equivalence classes in which all objects belong to the LPS decision class, which forms the lower approximation (i.e., equivalence classes E2, E4, E5, and E7). Alternatively, it can be described by the equivalence classes containing at least one object classified as LPS, constituting the upper approximation (i.e., E2, E3, E4, E5, and E7), as shown in Fig. 3.

Concept 4: Core and reduct of attributes

If the number of equivalent classes formed by the attribute set A is the same as that formed by $-a_i$, where $a \in A$, then a_i is redundant. Otherwise, a_i is indispensable in A. In Rough Set Theory (RST), the concept of a "reduct" is fundamental to feature engineering. A reduct refers to a minimal subset of attributes that preserves the same classification ability as the full attribute set. It is derived from the discernibility matrix using the discernibility function. By identifying and removing redundant or non-essential attributes, the reduct simplifies the dataset without compromising decision-making accuracy. This process

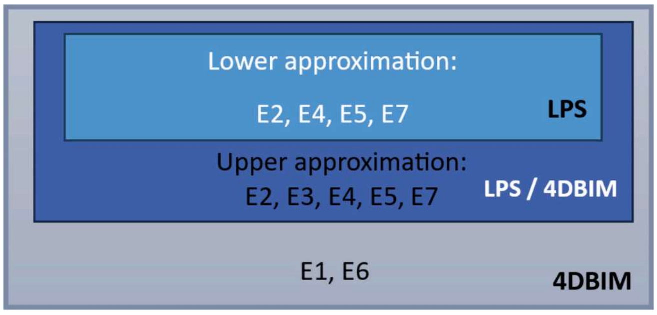


Fig. 3. Lower and upper approximation for the example.

not only streamlines the connection between input data and the conditions necessary for decisions but also enhances computational efficiency. As a feature selection technique, the reduct improves interpretability while reducing complexity. Finding a reduct, however, is an NP-hard problem, making it computationally difficult to discover all minimal reducts. To address this, algorithms like genetic reducers and Johnson reducers can be used to approximate the optimal reduct by iteratively selecting attributes that maximize the dependency degree, which reflects the classification power of the feature set. In this study, the genetic reducer was employed to compute reducts [6,17].

Another fundamental principle in RST is the concept of the "core" which is defined as the common portion of all reducts. For a given subset $B \subseteq A$, the core of B represents the set of attributes within B that are indispensable. The following equation embodies the connection between the core and reducts concepts.

$$Core(B) = \cap RED(B) \quad (6)$$

where $RED(B)$ is the set of all reducts of B.

Based on the equivalence class representation in Table 2, the discernibility function can be constructed by first developing a discernibility matrix that specifies the attributes of the different equivalence classes. The discernibility matrix for the example is shown in Table 3. It should be noted that the discernibility matrix is symmetric; for example, the entries for E1-E2 and E2-E1 are identical, so only one-half of the matrix needs to be considered. As shown in the highlighted column in Table 3, the entry for E2-E2 is empty (∅) since, naturally, the equivalence class E2 cannot be distinguished from itself. The entry for E2-E3 involves a different decision and includes attributes F1, F3, F4, and F5, for which differing values are observed between equivalence classes E2 and E3; for instance, F1 is 3 for E2, while it is 0 for E3. The entry for E2-E4 is also empty because equivalence classes resulting in the same decision do not require further discernment. The rest of the matrix is constructed similarly.

The minimal information required to discern E2 from all other objects with different decisions can now be expressed as a discernibility function:

$$f_{E2}(F1, F2, F3, F4, F5) = (F1\ \textbf{OR}\ F3\ \textbf{OR}\ F4\ \textbf{OR}\ F5)\ \textbf{AND}\ (F1\ \textbf{OR}\ F4\ \textbf{OR}\ F5)$$

To satisfy the condition for this function to be true, it is essential that at least one attribute from each E2-related entry within the discernibility

Table 2
Equivalence classes of the information system.

Equivalence classes	F1	F2	F3	F4	F5	Planning methods
E1 = {O1, O8}	1	0	3	4	4	{4DBIM}
E2 = {O6, O9}	3	1	4	1	4	{LPS}
E3 = {O5, O10}	0	1	3	3	3	{LPS, 4DBIM}
E4 = {O2}	3	2	4	1	3	{LPS}
E5 = {O3}	4	1	3	0	2	{LPS}
E6 = {O4}	0	1	4	3	3	{4DBIM}
E7 = {O7}	4	0	3	0	3	{LPS}

Table 3
Discernibility matrix for the example.

	E1	E2	E3	E4	E5	E6	E7
E1	∅						
E2	F1, F2, F3, F4	∅					
E3	F1, F2, F4, F5	F1, F3, F4, F5	∅				
E4	F1, F2, F3, F4, F5	∅	F1, F2, F3, F4	∅			
E5	F1, F2, F4, F5	∅	F1, F4, F5	∅	∅		
E6	∅	F1, F4, F5	F3	F1, F2, F4	F1, F3, F4, F5	∅	
E7	F1, F4, F5	∅	F1, F2, F4	∅	∅	F1, F2, F3, F4	∅

matrix is included. Consequently, the function can be simplified to:

$$f_{E2}(F1, F2, F3, F4, F5) = (F3 \text{ AND } F1) \text{ OR } (F3 \text{ AND } F4) \text{ OR } (F3 \text{ AND } F5)$$

which reflects the three *reducts*: $\{F3, F1\}$, $\{F3, F4\}$, and $\{F3, F5\}$.

It is important to note that identifying all reducts is an NP-complete problem [48]. Nonetheless, several approximation algorithms, such as greedy algorithms [27] and genetic algorithms [51], have been developed to facilitate the search for reducts. In this study, a genetic algorithm was employed to compute the reducts.

Concept 5: Decision rules

Rule generation is a crucial step derived from reduct computation in RST. Decision rules capture the knowledge extracted from the data and typically take the form: r : IF C THEN D. where the condition (C) is a conjunction of attribute-value pairs, like $a_1 = v_1 \wedge a_2 = v_2$, which specify the attribute values that define a particular condition. The decision (D) indicates the value associated with the decision attribute, such as $d = v_d$. These rules serve as a structured representation of knowledge, which facilitates decision-making by mapping conditions to specific outcomes.

Considering the reducts {F3, F1}, {F3, F4}, and {F3, F5} for discerning equivalence class E2 in the example, the resulting rules based on the attribute values would be:

- R1: **IF** F3 (4) AND F1 (3) **THEN** LPS
- R2: **IF** F3 (4) AND F4 (1) **THEN** LPS
- R3: **IF** F3 (4) AND F5 (4) **THEN** LPS

Rule evaluation

Rule evaluation is essential for assessing the accuracy, reliability, and usefulness of generated rules. It helps validate the rules to ensure they accurately reflect data relationships. By doing so, rule evaluation ensures that extracted knowledge is reliable and practical for informed decision-making. There are different measures for rule evaluation, including support, coverage, accuracy, and *p*-value. In this research, support, accuracy, and p-value were preferred as measures for rule evaluation. These three metrics offer comprehensive insights into the significance and applicability of rules, making coverage unnecessary as it doesn't add much beyond what others already reveal. Therefore, coverage was not included in the evaluation. Support measures how many instances meet the conditions of the rule. Left-hand side support (LHS support) counts the number of instances that satisfy the conditions in the IF part of the rule, while Right-hand side support (RHS support) measures the instances that meet the specified classes in the THEN part of the rule. The predictive efficacy of a rule is reflected in its accuracy, a metric determined through the computation of support values. More precisely, the following formula calculates a rule's accuracy:

$$accuracy(rule) = \frac{support\ (rule)_{RHS}}{support\ (rule)_{LHS}} \tag{7}$$

Moreover, the *p*-value is a measure that evaluates the statistical significance of the generated rules in the context of rule evaluation. Garbulowski, et al. [17] adopted the hypergeometric distribution to compute these *p*-values, a method that assesses the representation of rule support relative to the total number of objects. A p-value less than 0.05 is considered acceptable, indicating that the rule is statistically significant, which ensures that only meaningful rules are included in the model. Eq. (8) depicts the calculation of the *p* value.

$$P(X = r) = \frac{\binom{t_d}{r}\binom{t_0}{l-r}}{\binom{T}{l}} \tag{8}$$

where r is the RHS support of the rule, l is the LHS support of the rule. The total number of objects that align with the rule's decision class d is denoted as t_d , while t_0 indicates the number of objects belonging to decision class(es) other than the one targeted by the rule. The total number of objects within the dataset is depicted by T.

Table 4 summarizes the evaluation metrics for the rules generated from the equivalence class E2 in the example.

The first rule, R1, corresponds to the condition part of three objects in Table 1, resulting in a left-hand side support of 3. Additionally, R1 aligns with the decision part of these same three objects in Table 1, giving it a right-hand side support of 3.

Among the three objects that satisfy the IF-part of rule R1, all three also belong to the decision class specified in the THEN-part (i.e., the LPS decision class). Consequently, the rule's accuracy is 1.0. Furthermore, according to Eq. (8), the calculated *p*-value for this rule is 0.16.

RST Implementation Using R.ROSSETA

This study used the R.ROSETTA package in the R programming language to implement rough set theory. R.ROSETTA is an advanced toolkit designed to facilitate the entire spectrum of data mining and knowledge discovery processes [17]. It is an extension of the original ROSETTA system, augmenting its functionality, accessibility, and flexibility. Notably, R.ROSETTA specializes in developing and analyzing rule-based classification models, encompassing features such as data preprocessing, discretization, and reduct computation [17]. For reduct computation, a Genetic algorithm was utilized, which excels in identifying minimal attribute sets that retain essential information. R. ROSETTA's robust analytical capabilities enable the generation of effective decision rules and filters according to rigorous evaluation metrics, which offers a reliable platform for extracting insights from uncertain data.

Implementing the Inference Engine

The forward chaining approach was employed for implementing the inference engine and rule activation. Rule activation approaches include forward chaining and backward chaining [4]. Forward chaining begins with available facts and applies inference rules to derive new facts until a conclusion is reached. In contrast, backward chaining starts with a goal and works backwards to identify the supporting facts. Forward chaining was chosen for this study due to its ability to derive conclusions iteratively by applying rules based on existing data [4]. Pyke was selected for implementing the inference engine with the forward chaining approach. It is a Python knowledge engine that provides a logic

Table 4
Rule assessment results for the generated rules in the example.

ID	Rules	RHS support	LHS support	Accuracy	p-value
R1	**IF** F3 (4) AND F1 (3) **THEN** LPS	3	3	1.0 (3/3)	0.16
R2	**IF** F3 (4) AND F4 (1) **THEN** LPS	3	3	1.0 (3/3)	0.16
R3	**IF** F3 (4) AND F5 (4) **THEN** LPS	2	2	1.0 (2/2)	0.33

programming framework and supports knowledge-based inference through rule-based programming. Users are empowered to define rules and facts, making it easier to create decision-making programs that rely on logical conditions [16]. This made Pyke ideal for implementing forward chaining in this study, as it facilitated structured rule activation and data-driven reasoning processes. Pyke starts with a knowledge base, which consists of a set of facts and rules. Facts in this research are the user requirements in terms of functionalities and will be collected through the system's interface, whereas rules are conditional statements that indicate what conclusions can be inferred from what facts.

To perform forward chaining, Pyke finds rules whose "*if*" clause matches its list of known facts. When a rule's conditions are met, It activates the rule, which adds the facts in the "*then*" clause of the rule to the existing list of known facts. These newly added facts can then trigger other rules with matching "*if*" clauses, continuing the chaining process to any depth. In this way, Pyke links the "*then*" clause of one rule to the "*if*" clause of the next, progressively drawing logical inferences from the data. Pyke effectively manages the flow of rule activation, ensuring that rules are fired in a logical sequence, which allows for efficient knowledge discovery and reasoning [16].

It is worth noting that although the implemented Pyke engine initiates forward chaining and typically activates rules in a logical succession, the activation process in this study is linear and straightforward as the generated rules follow a flat structure and each rule is processed independently in a linear sequence, without triggering or relying on other rules. Rules are evaluated one by one, simplifying execution without the need for cascading activations.

Following the deployment of the inference engine, a Python script was formulated to parse the outcomes of the inference process and construct a part of the input dataset for the mathematical model and optimization purposes, which will be explained in the next sections.

User Interface

A user interface plays a crucial role in bridging the gap between end users and technical systems, enhancing user experience by providing an intuitive platform for interaction. In this study, a user interface was designed to gather user requirements effectively and facilitate communication between users, system engineers, and knowledge engineers. One of its roles is to collect project team requirements for a planning and control system. Also, by reasoning through the inference engine and optimizing via a mathematical model, the interface visually displays the results of the suggested planning methods and control metrics across three schedule levels. Fig. 4 depicts the configuration of the interface designed for the Multi-level Planning and Control System Decision Support System (MPCS-DSS). Notably, the interface comprises four primary components. The initial page outlines the key objectives of the DSS and the requisite data. Subsequently, the second component endeavors to gather project-related information, while the third segment is dedicated to collecting functional requirements, based on Sheikhkhoshkar, et al. [42], Sheikhkhoshkar, et al. [43,45] and Sheikhkhoshkar, et al. [44], from the project team for a planning and control system, serving as the primary input data for the DSS. Finally, the last component demonstrates the outcomes proposed by the DSS.

3.2.2. *Mathematical model for the results' optimization*

This section introduces a multi-objective mathematical model developed to optimize the results of DSS and suggests a multi-level and integrated project planning and control system for construction projects. The objective is to minimize the number of planning methods and control metrics at each schedule level for more practical implementation while satisfying the maximum project team's requirements. The following subsection outlines the relevant sets, indicators, parameters, decision variables, objective functions, and constraints used in the mathematical model.

Sets and Indices:

i: index for the planning method, $i \in [1..I.]$

j: index for the control metric, $j \in [1..J]$

l: index for the schedule level, $l \in [1..L]$

f: index for the functionalities, $f \in [1..F]$

Input parameters:

P = number of planning methods

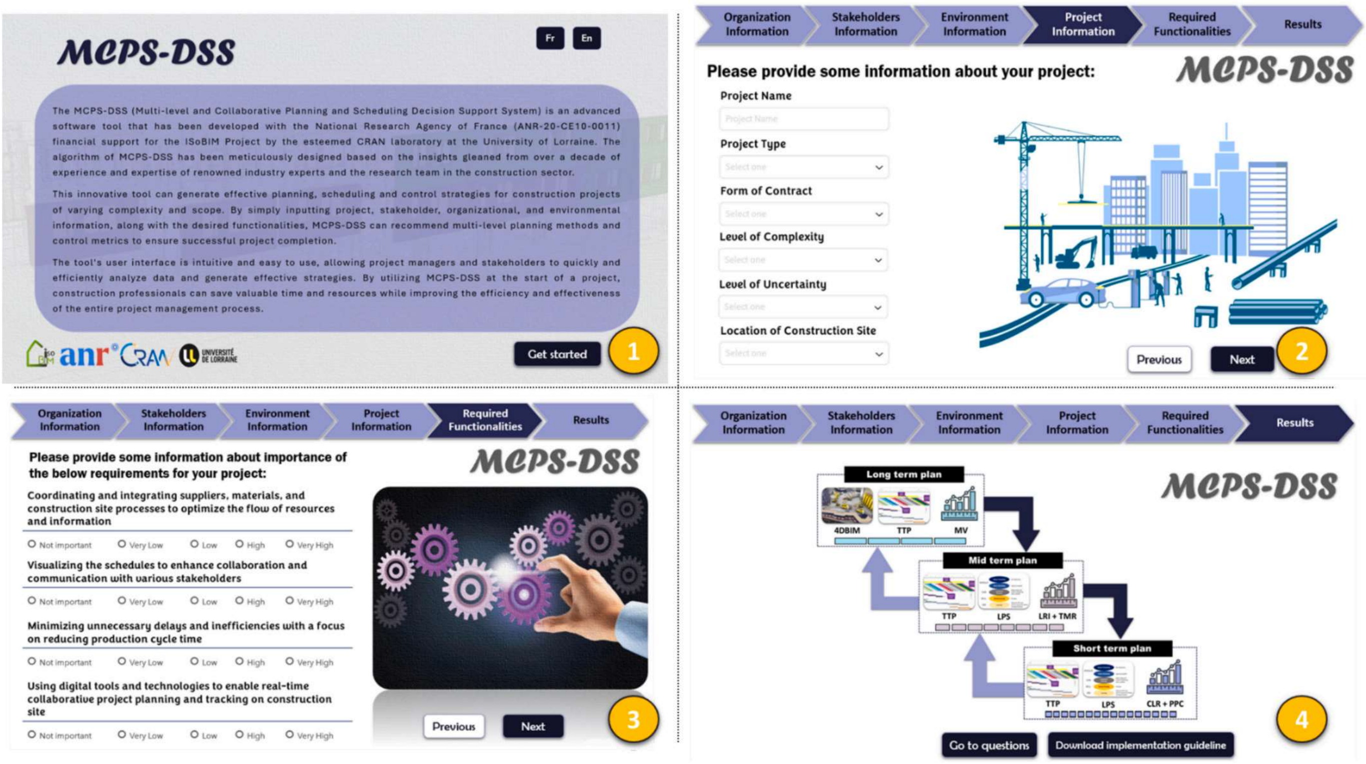


Fig. 4. Configuration of the designed interface for MPCS-DSS.

$\boldsymbol{C}$ = number of control metric
$\boldsymbol{L}$ = number of schedule levels
$\boldsymbol{F}$ = number of functionalities
$\boldsymbol{M_{fi}}$: level of support for functionality f in planning method i, $\boldsymbol{M_{fi}} = [0, 1, 2, 3, 4]$
$\boldsymbol{K_{fj}}$: level of support for functionality f in control metric j, $\boldsymbol{K_{fj}} = [0, 1, 2, 3, 4]$
$\boldsymbol{S_{fl}}$: level of support for functionality f in schedule level l, $\boldsymbol{S_{fl}} = [0, 1, 2, 3, 4]$
$\boldsymbol{A_{ij}}$: connects planning method i to control metric j. $A_{ij} = 1$ means that the planning method i mapped by control metric j.
$\boldsymbol{B_{il}}$: connects planning method i to schedule level l. $B_{il} = 1$ means that the planning method i matched with schedule level l.
$\boldsymbol{C_{jl}}$: connects control metrics j to schedule level l. $C_{jl} = 1$ means that the control metric j matched with schedule level l.
$\boldsymbol{W_i}$: A weight to evaluate the coverage of the suggested planning methods by DSS for user requirements in terms of functionalities.

$$W_i = \sum_{f=1}^{F} M_{fi} \tag{9}$$

$\boldsymbol{V_j}$: A weight to evaluate the coverage of the suggested control metrics by DSS for user requirements in terms of functionalities.

$$V_j = \sum_{f=1}^{F} K_{fj} \tag{10}$$

$\boldsymbol{U_l}$: A wight to evaluate the coverage of the suggested schedule levels by DSS for user requirements in terms of functionalities.

$$U_l = \sum_{f=1}^{F} S_{fl} \tag{11}$$

Output variables:
Binary decision variable X_i:
$\boldsymbol{X_i} = 1$ If planning method i is decided to be deployed
$\boldsymbol{X_i} = 0$ Otherwise
Binary decision variable Y_j:
$\boldsymbol{Y_j} = 1$ If control metric j is decided to be deployed
$\boldsymbol{Y_j} = 0$ Otherwise
Objective Functions:

$$Z_1 = max\left(\sum_i W_i.X_i + \sum_j V_j.Y_j\right)\Bigg) \tag{12}$$

$$Z_2 = min\left(\sum_i X_i\right) \tag{13}$$

$$Z_3 = min\left(\sum_j Y_j\right) \tag{14}$$

Eqs. (12)–(14) aim to maximize project team requirements in terms of functionalities while minimizing the number of planning methods and control metrics at each level of schedule.
Constraints:

$$\forall j \in [1..J], \sum_{i=1}^{I} (A_{ij} \times X_i) - Y_j \geq 0 \tag{15}$$

Eq. (15) ensures that if a control metric is selected, then at least one planning method that supports this control metric must also be selected.

$$\forall l \in [1..L], \sum_{i=1}^{I} (B_{il} \times X_i) \geq U_l \tag{16}$$

Eq. (16) demonstrates that deploying a particular schedule level requires the simultaneous deployment of at least one planning method corresponding to that level.

$$\forall l \in [1..L], \sum_{j=1}^{J} (C_{jl} \times Y_j) \geq U_l \tag{17}$$

Eq. (17) specifies that once a particular level of schedule has been designated for deployment, at least one corresponding control metric for that level must be deployed.

$$\forall i \in [1..I], \sum_{l=1}^{L} B_{il} \times U_l \geq X_i \tag{18}$$

Eq. (18) enforces that if a schedule level is active, at least one planning method that is applicable to this level must be selected.

$$\forall j \in [1..J], \sum_{l=1}^{L} C_{jl} \times U_l \geq Y_j \tag{19}$$

Eq. (19) enforces that if a schedule level is active, at least one control metric that is applicable to this level must be selected.

The selection of appropriate planning and control systems for projects is influenced by a range of external factors, including project type, scale, contract type, and the project team's expertise in planning and control systems. These variables contribute to the complexity of the decision-making process. To address this complexity, this study utilizes Pareto front plots to present the outcomes of the mathematical model. This approach offers project teams the flexibility to choose solutions that best meet the specific requirements of diverse project conditions. A Pareto front chart is a graphical tool used in multi-objective optimization to display a set of feasible solutions. It identifies options where no single solution is superior in all aspects, highlighting the trade-offs between competing objectives. This helps decision-makers find a balance between conflicting goals by illustrating where improvements in one objective may require sacrifices in another.

The mathematical model was executed using Python 3.11, and the Gurobi linear solver, under academic licence.

3.2.3. Solution evaluation and case project application

In this study, two approaches were adopted to assess the usability and practicality of the solution. First, a case study was conducted within the IsoBIM project framework to showcase the practicality of the decision support system and mathematical model in proposing a multi-level planning and control system for a renovation project. Second, quantitative performance measures from satisfaction surveys, as recommended by Peffers, et al. [36] were used for a general evaluation. Accordingly, a team of five experts along with the research team evaluated the usability of the proposed solution and its integration into current project planning and control processes. Through semi-structured interviews, the participants shared their functional requirements for a planning and control system using the user interface. They then assessed the DSS using a Likert scale, considering the recommendations provided by the tool. To facilitate this process, five dimensions of evaluation were considered, including ease of use, interface quality, comprehensiveness, response time, and decision quality improvement. The specifics of the evaluation criteria are detailed in Table 5. Such end-user survey for evaluating a DSS or framework was also considered in the studies by Dikmen, et al. [12], Barkokebas, et al. [9], and Sheikhkhoshkar, et al. [47].

4. Results and analysis

4.1. Data collection efforts

Using a purposive sampling approach, a targeted cohort of 45 individuals was identified for this study's data collection through a thorough review of LinkedIn profiles, peer-reviewed articles, and various online repositories. Following this identification, invitations to participate were sent via email and direct messaging channels. Of these, 23

Table 5
Details of the evaluation criteria.

Evaluation criteria	Objective
Ease of use	Assess how easy it is for new users to learn and for all users to operate the system
Comprehensiveness	Assess the extent to which the system includes all relevant data, variables, and functionalities necessary to support thorough and informed decision-making
Decision quality improvement	Determine whether and how much the DSS improves the quality of decisions compared to pre-implementation
Interface quality	Evaluate the intuitiveness, clarity, and aesthetic of the user interface.
Response time	Measure the time it takes for the system to provide outputs after inputs are given.

individuals (51 % of the identified sample) responded affirmatively. The respondents were subsequently scheduled for interviews, which included participation from both industry and academic experts. These interview sessions collectively contributed approximately 23 h of valuable insights.

To ensure data saturation and validate the sample size, numerous studies in construction management have advocated and utilized purposive sampling, typically selecting sample sizes varying from 5 to 25 [26,50,53]. Moreover, the 23 interviews were conducted with experts possessing specialized knowledge in various planning methods and their applications. As the target population of these experts is relatively narrow worldwide, the 23 interviews provide a strong and representative sample of the targeted population. The details of the participants are provided in Table 6.

This study employs a structure, depicted in Fig. 5, which synthesizes various planning methods and control metrics across different schedule levels to suggest multi-level and integrated project planning and control systems. The structure is underpinned by the functionality concept, which acts as a common principle among planning methods, control metrics, and scheduling levels to facilitate their harmonious integration. More information about defining, extracting and analyzing the functionalities is elaborated in [42,45,46].

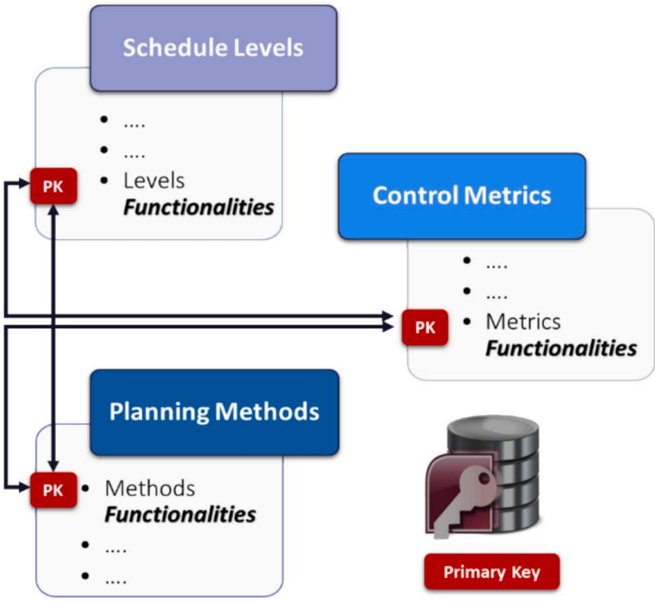


Fig. 5. Main elements of multi-level framework.

To operationalize this structure, eight commonly used planning methods in construction for different project types [42] were examined, including advanced work packaging (AWP), 4D building information modelling (4DBIM), critical chain project management (CCPM), critical path method (CPM), last planner system (LPS), location-based management system (LBMS), linear scheduling method (LSM), and takt time planning (TTP). Furthermore, the study considered nine control metrics frequently referenced in the academic literature and practical guidelines [20,45]: cost performance index (CPI), schedule performance index (SPI), milestone variance (MV), percent planned complete (PPC), required level (RL), tasks made ready (TMR), capacity to load ratio (CLR), location risk index (LRI), and task anticipated (TA). The framework also incorporates three schedule levels comprising short-term, mid-term, and long-term. Regarding the functionalities, 19 main

Table 6
Profile of Interviewees.

Expert ID	Type of Expert	Role	Organization Type	Years of Experience in Project Planning and Control
IE1	Industry experts	Department head	General contractor	11–15 years
IE2		National director of projects	General contractor	Over 15 years
IE3		Process manager	General contractor	1–5 years
IE4		Senior innovation engineer	General contractor	6–10 years
IE5		R&D manager	General contractor	11–15 years
IE6		Lean director	General contractor	Over 15 years
IE7		Scheduler	Client	Over 15 years
IE8		Project manager	Client	Over 15 years
IE9		Project manager	Client	Over 15 years
IE10		Project manager	Consultant	Over 15 years
IE11		Project manager	Consultant	Over 15 years
IE12		Superintendent	Consultant	Over 15 years
IE13		Project manager	Consultant	11–15 years
IE14		Manager - customer success	Software vendor	6–10 years
AE1	Academic experts	Professor	University	11–15 years
AE2		Professor	University	11–15 years
AE3		Professor	University	Over 15 years
AE4		Professor	University	6–10 years
AE5		Professor	University	11–15 years
AE6		Senior lecturer	University	6–10 years
AE7		Lecturer	University	6–10 years
AE8		Researcher (PhD student, post-doc)	University	6–10 years
AE9		Researcher (PhD student, post-doc)	University	6–10 years

Table 7
Incorporated functionality concepts in multi-level planning and control systems.

ID	Functionalities
F1	Collaboration management
F2	Commitment planning
F3	Communication management
F4	Conflict management
F5	Constraint management
F6	Contract and delay management
F7	Integration management
F8	Learning and knowledge sharing
F9	Risk management
F10	Process and flow management
F11	Project performance management
F12	Reliability management
F13	Resource management
F14	Root cause analysis
F15	Safety and logistic management
F16	Supply chain management
F17	Visualization
F18	Waste management
F19	Real-time site monitoring

concepts of functionalities were considered, drawing upon prior research conducted by Sheikhkhoshkar, et al. [43] illustrated in Table 7. To enhance clarity and facilitate understanding for non-expert users, the extended functionality concepts presented in Table 7 are further elaborated in Table A1, located in the appendix. It should be noted that these detailed definitions of the functionalities were incorporated into the user interface, serving as the user requirements in terms of the functionalities for a planning and control system.

Utilizing these classifications, a relational diagram was developed, as shown Fig. 6 to guide data collection and serve as a foundational element of survey design for data-gathering purposes. The survey questions were formulated to gather data aimed at identifying potential connections between planning methods and control metrics, evaluating the alignment of planning methods and control metrics across various schedule levels, and assessing the degree of support offered by each planning method, control metric, and schedule level for different functionalities.

To cover all the necessary data for the connections between the components in Fig. 6, six groups of questions needed to be addressed, including:

1. How well did each planning method align with each schedule level?
2. Which schedule level is best suited for each control metric?
3. Which planning method does each control metric align with?
4. To what extent did each planning method support a specific functionality?
5. To what extent did each control metric support a specific functionality?
6. How much did each schedule level contribute to supporting a specific functionality?

During the interview process, experts' opinions were assessed using a Likert scale ranging from 0 (no support) to 4 (very high support).

4.2. Data preprocessing efforts

Following the application of the IQR method in the data preprocessing phase, the detected outliers were carefully investigated. Given that some outliers could potentially contain valuable information or represent errors, deliberate consideration was given to appropriately address genuine outliers and anomalies before determining whether to remove or replace them. Consequently, the real outliers were replaced with the average responses for that specific record in the whole dataset and for all classes. Fig. 7 indicates an example of detected outliers for the last planner system and long-term schedule level data.

After handling outliers, SMOTE was employed to address the challenge of imbalanced data. Fig. 8 depicts the class distribution of the imbalance dataset pre- and post-SMOTE augmentation.

To assess the reliability of the collected data, Cronbach's α coefficient was computed for all classes of the dataset, including eight classes for planning methods, nine classes for control metrics, and three classes for schedule levels. Table 8 displays the computation results. Cronbach's α coefficient for all classes is more than 0.7. Therefore, the data derived from the survey is considered reliable.

4.3. Building the knowledge repository

The Genetic reducer is utilized for reduct computation in the rule generation process. A considerable number of outputs can be produced from Genetic heuristics in R.ROSETTA [17], making it challenging to interpret such models. To deal with this challenge, defining precise thresholds was imperative to identify the most significant and valid rules derived from the rule generation process. To this end, multiple criteria were set, including a p-value threshold of less than 0.05, an accuracy greater than 0.7, a minimum RHS and LHS support of five, and a maximum rule length of eight. The outcomes of the rule generation process and the construction of a knowledge repository employing Rough Set Theory (RST) with R.ROSETTA are presented in Table 9. The table contains the total number of generated rules for each category, the number of rules that meet the thresholds, as well as the mean values of

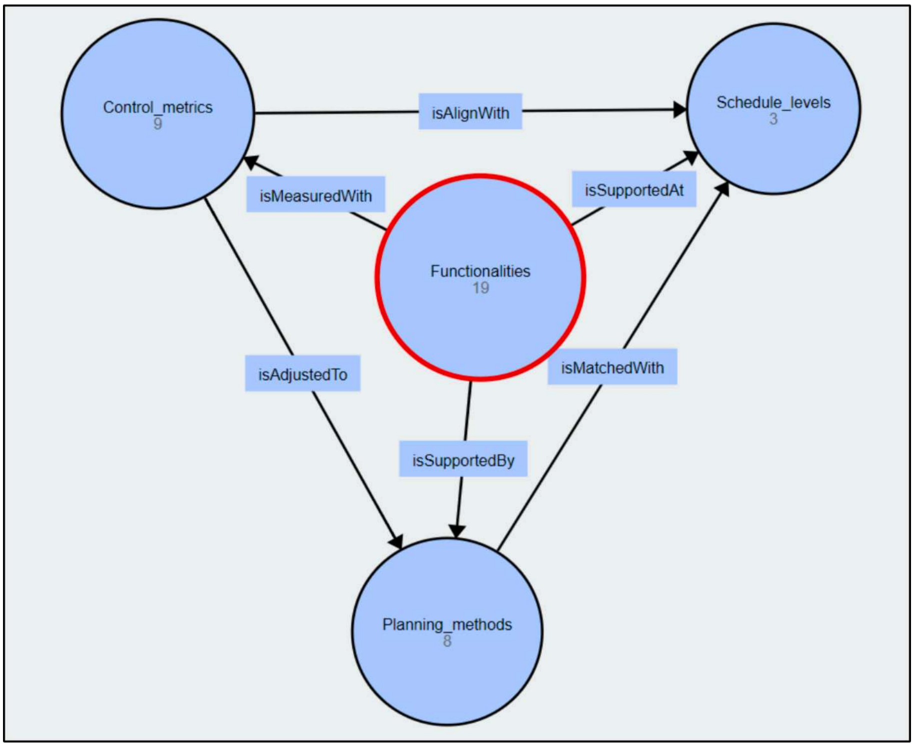


Fig. 6. Relational diagram of the main elements of the multi-level framework.

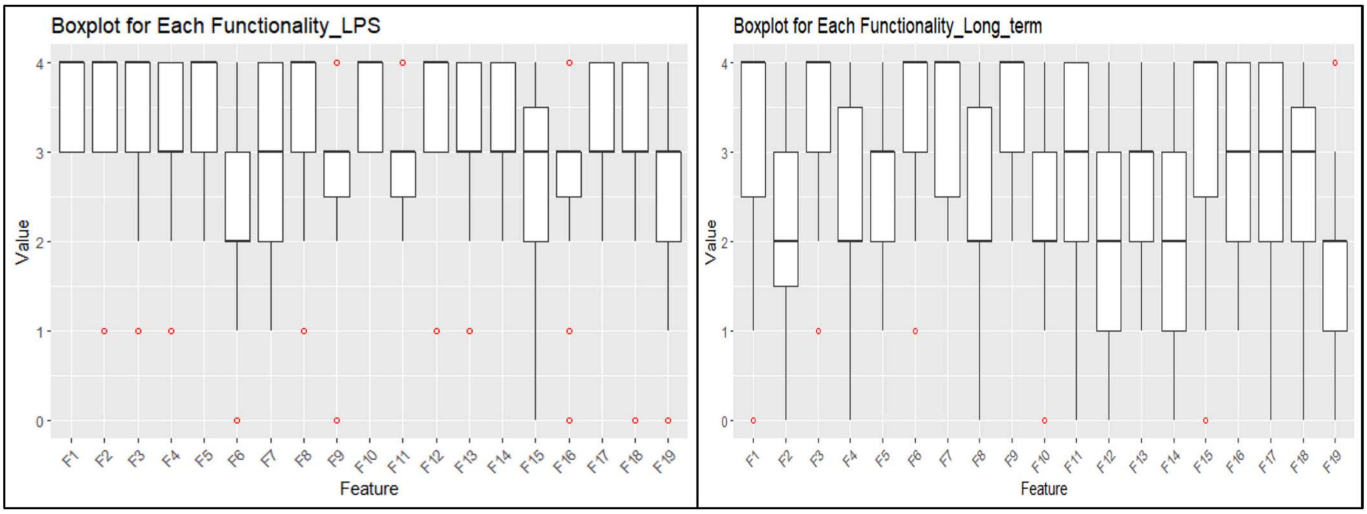


Fig. 7. Detected outliers for the last planner system (LPS) and long-term level of schedule data.

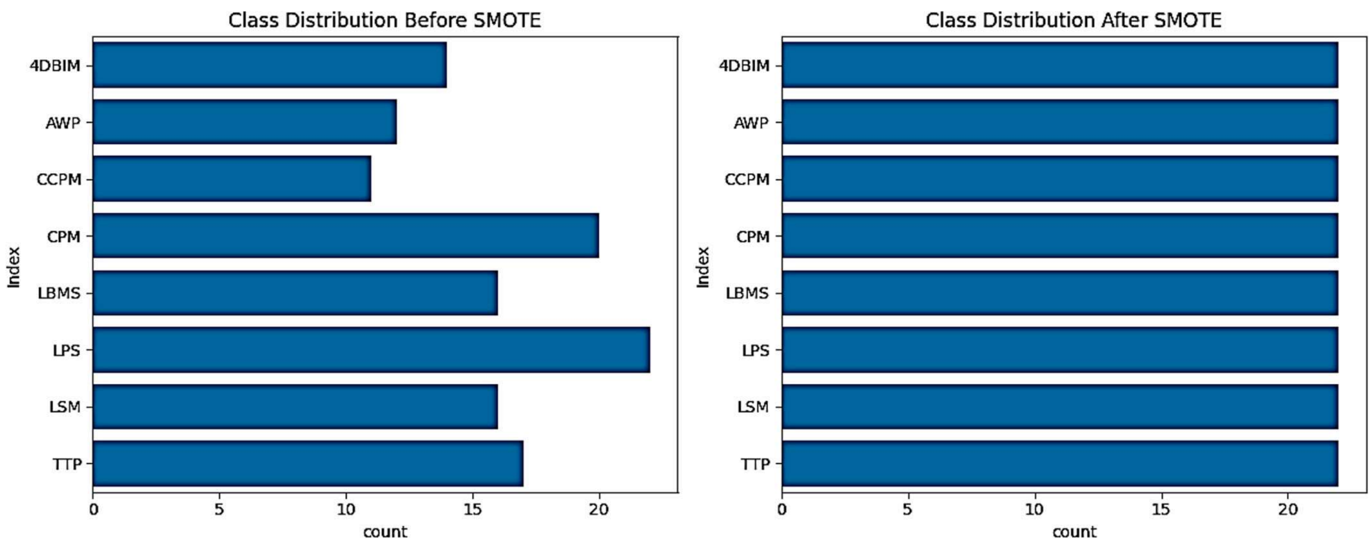


Fig. 8. Class distribution of the imbalance dataset before and after SMOTE implementation.

Table 8
Cronbach's alpha for collected data.

Planning methods		Control metrics		Schedule levels	
Class	*Cronbach's Alpha*	*Class*	*Cronbach's Alpha*	*Class*	*Cronbach's Alpha*
4DBIM	0.838	CLR	0.726	Long-term	0.924
AWP	0.841	CPI	0.814	Mid-term	0.897
CCPM	0.970	LRI	0.815	Short-term	0.824
CPM	0.935	MV	0.715		
LBMS	0.953	PPC	0.921		
LPS	0.908	RL	0.774		
LSM	0.949	SPI	0.707		
TTP	0.929	TA	0.782		
		TMR	0.879		

support, accuracy, and *p*-values associated with the generated rules.

Analysis of rule generation across planning methods, control metrics, and schedule levels reveals differentiated functionalities' support for each class and their implications for knowledge databases and decision-making. Planning methods, with a significant generation of 39,977 rules and 1753 meeting rigorous selection criteria ($P \leq 0.05$, accuracy ≥ 0.7, support ≥ 5, length ≤ 8), illustrate a multifunctional aspect, supporting a broad spectrum of functionalities within the system. In contrast, the control metrics class, generating only 463 rules with 80 meeting the thresholds, reflects its inherently focused functionality. The relatively low number of generated rules in this class is reasonable, as control metrics are designed to support specific and limited aspects of project control rather than multiple functionalities. Meanwhile, schedule levels, generating 12,962 rules with 447 qualifying the selection criteria, indicate a level of multifunctionality greater than control metrics.

The selected rules across all classes demonstrate high accuracy and validity, essential for integrating them into a knowledge repository and effective decision-making. Specifically, planning methods exhibit an accuracy of 0.895, control metrics show a remarkable accuracy of 0.962, and schedule levels maintain a strong accuracy of 0.935. These accuracy levels, coupled with very supportive *p*-values and substantial LHS and RHS support metrics, confirm the robustness and utility of the rules.

Table 10 depicts the rules information and distribution within the subclasses of the planning methods class, detailing the specific methods such as 4BIM, AWP, CCPM, CPM, LBMS, LPS, LSM, and TTP. Each subclass demonstrates a robust distribution of essential rules that

Table 9
Performance evaluation of rules for the Genetic reduction method.

Class →	Planning Methods	Control Metrics	Schedule Levels
Total number of rules	39,977	463	12,962
Selected rules statistics			
Number of rules considering thresholds ($P \leq 0.05$, accuracy ≥ 0.7, support (RHS & LHS) ≥ 5, length ≤ 8)	1753	80	447
Mean LHS support	10	12	13
Mean RHS support	9	11	12
Mean accuracy	0.895	0.962	0.935
Mean p-value	0.0065	0.00012	0.0306

Table 10
Rules information for the planning methods class.

Class →	4BIM	AWP	CCPM	CPM	LBMS	LPS	LSM	TTP
Total number of rules	5786	3899	5307	6427	5776	5259	4102	3421
Selected rules statistics								
Number of rules considering thresholds ($P \leq 0.05$, accuracy ≥ 0.7, support (RHS & LHS) ≥ 5, length ≤ 8)	120	143	197	378	249	208	285	173
Mean LHS support	10	10	9	13	9	10	9	9
Mean RHS support	9	8	8	11	8	9	8	8
Mean accuracy	0.893	0.865	0.943	0.852	0.911	0.887	0.916	0.893
Mean p-value	0.0062	0.010	0.0054	0.0013	0.0088	0.0052	0.0068	0.0083

effectively contribute to the knowledge repository, indicating a well-rounded approach to capturing diverse functional requirements. This distribution highlights that each subclass maintains a relatively consistent spread of rules that meet the selection criteria, underscoring all planning methods are included in the knowledge database.

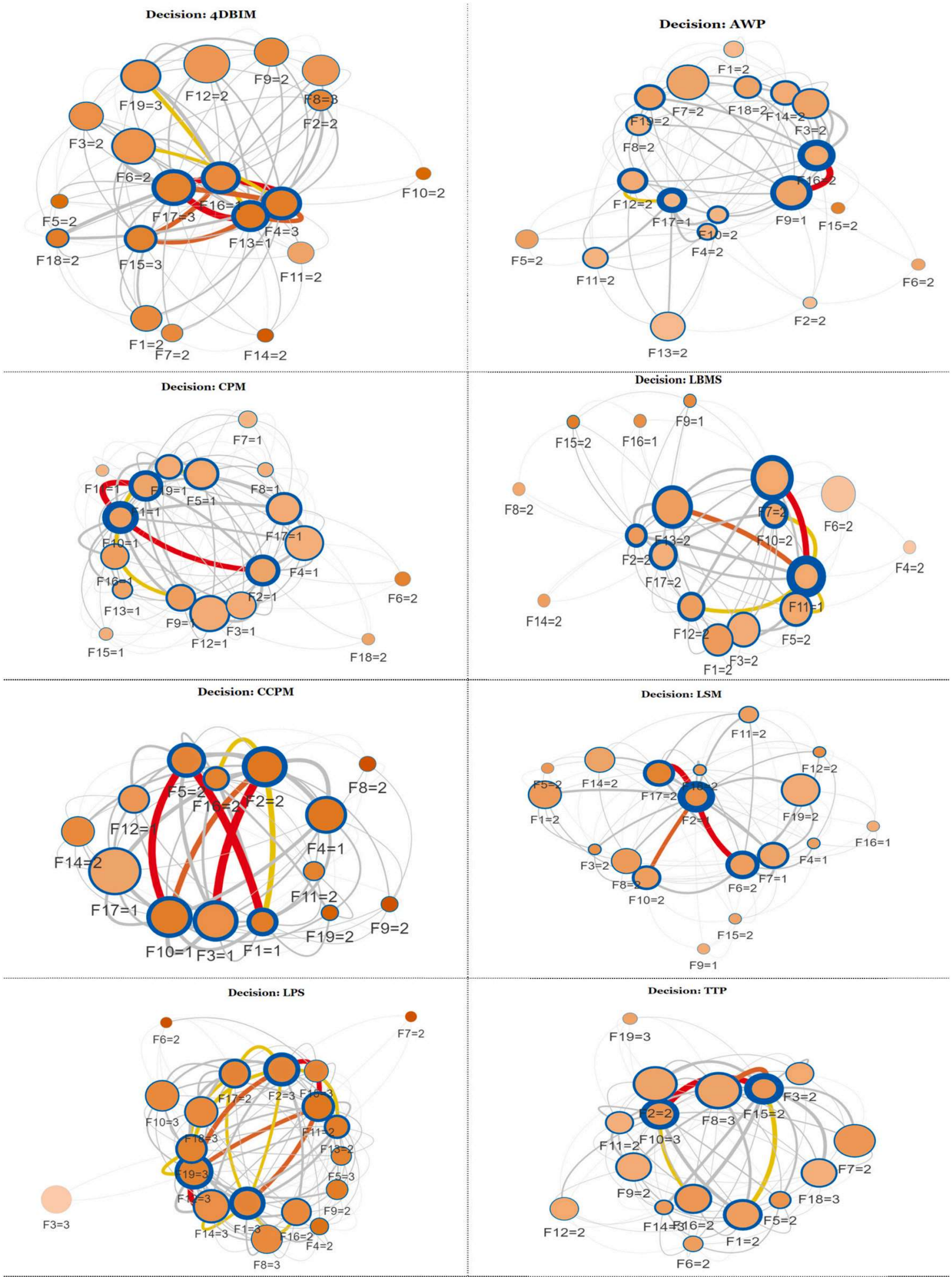
The rule distribution is also relatively robust for subclasses of schedule levels, as shown in Table 11. This distribution indicates that the existing rules in each subclass hold a relatively compatible spread of rules, ensuring that the necessary aspects of schedule levels are effectively covered.

Conversely, the control metrics class displays fewer rules, as evidenced by the limited rules meeting the stringent criteria. This shortfall is consistent with the inherent nature of control metrics, which are typically not multifunctional and are designed to measure specific aspects of projects. The limited rule count in this class illustrates a gap in the coverage of control metrics, highlighting the need for further research and development in this area to ensure that essential aspects of project control are comprehensively addressed within the knowledge repository.

Fig. 9 illustrates a visual representation of the interactions among functionalities within each subclass of the knowledge database, depicted through R.ROSETTA's rule-based model visualization feature. This holistic approach displays the entire knowledge repository as an interaction network that organizes different subclasses and their respective

Table 11
Rules information for the schedule levels class.

Class →	Long-term	Mid-term	Short-term
Total number of rules	4254	5341	3367
Selected rules statistics			
Number of rules, considering thresholds ($P \leq 0.05$, accuracy ≥ 0.7, support (RHS & LHS) ≥ 5, length ≤ 8)	107	99	241
Mean LHS support	13	12	15
Mean RHS support	12	11	13
Mean accuracy	0.939	0.944	0.921
Mean p-value	0.028	0.05	0.013



(a) Planning methods class

Fig. 9. Visual representation of knowledge repository, (a) planning methods class, (b) schedule levels class, and (c) control metrics class, (d) rule visualization guide.

functionalities. This kind of visualization not only highlights the integral network of interactions but can also be adjusted to emphasize the most relevant co-predictive functionalities and their intensity levels in each class. Additionally, Fig. 9 further supports our claim regarding the low functionalities' support by the control metrics class. It clearly shows that, compared to the planning methods and schedule levels classes, the control metrics class supports only a limited number of functionalities. This contrast underscores the multifunctional nature of the other two classes, highlighting their broader applicational scope within the knowledge database.

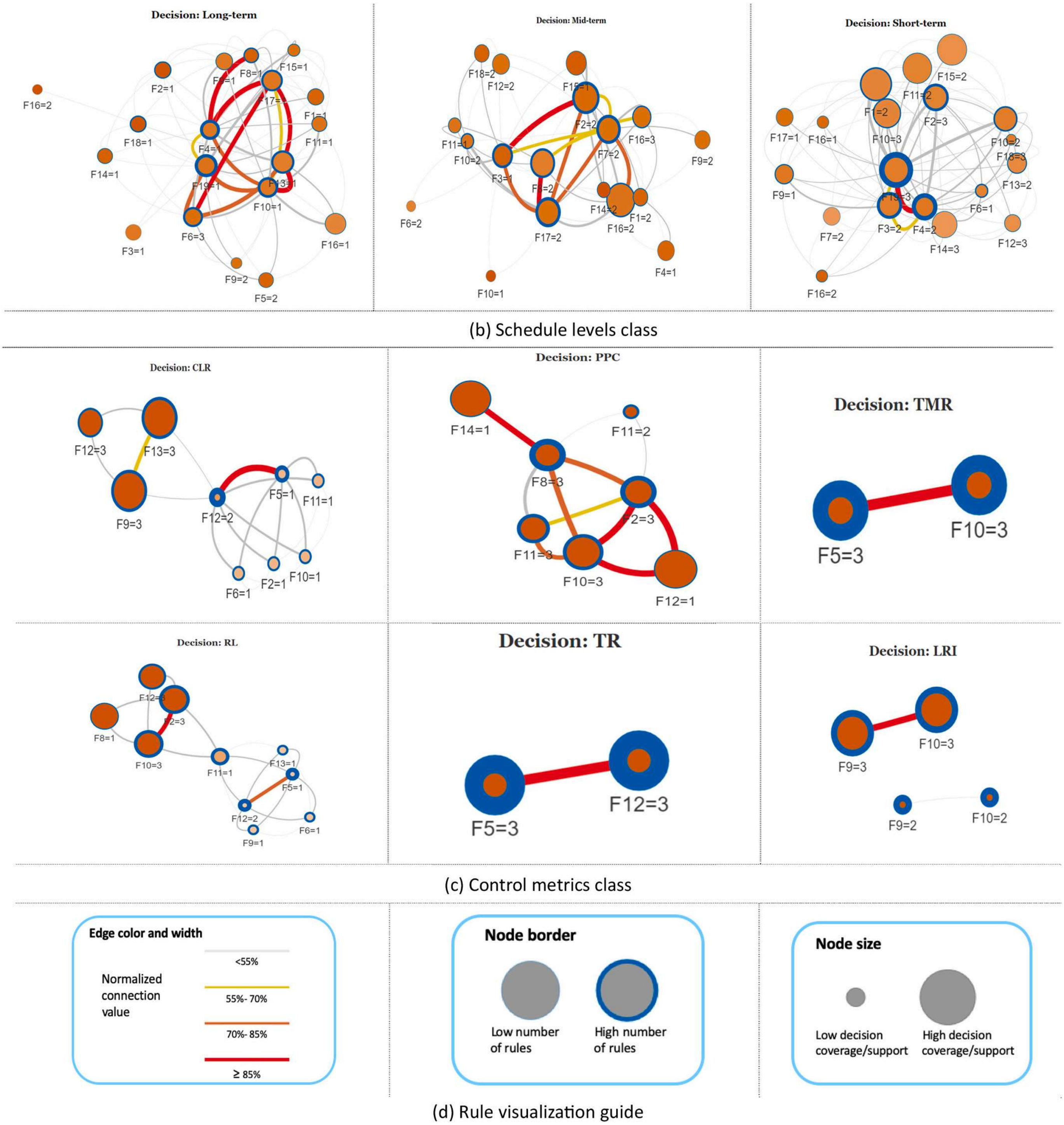


Fig. 9. (*continued*).

Following the rule generation and validation process employing RST methodology, a set of 2280 rules was formulated. These rules are stated as individual statements or combinations of statements linked by “AND” or “OR” conditions, classified into three distinct categories: planning methods, control metrics, and schedule levels. The resulting database functions as a knowledge repository for project planning and control systems. Table 12 provides a sample of selected rules within each class.

To implement the inference engine using the Pyke engine, which requires a knowledge database in its specific syntax, the syntax of the generated rules had to be converted for compatibility with Pyke. A Python script was developed to handle this conversion efficiently. An example of the converted rule syntax is illustrated in Fig. 10, demonstrating how the rules were adapted for Pyke’s framework.

The system’s inference engine utilizes a forward chaining approach, which aligns with the flat rule structure observed in the knowledge base. Since the rules are designed to directly map conditions to decisions without engaging in complex inference chains, forward chaining efficiently linearly processes these rules. The inference engine evaluates the conditions and triggers the appropriate decisions immediately, consistent with a flat structure. This straightforward reasoning process is well-suited for the system’s rule-based design.

Table 12
Filtered rules examples.

Rule class	Rule	LHS support	RHS support	Accuracy	P-value
Planning methods	IF F1(1) AND F4 (1) AND F5(1) AND F6(2) THEN CPM	14	10	0.714	0.00360
	IF F11(2) AND F12(3) AND F14 (3) AND F19(3) THEN LPS	11	11	1	3.58E-07
	IF F4(3) AND F13 (1) AND F19(3) THEN 4DBIM	12	10	0.913	0.000274
	IF F2(2) AND F8 (3) AND F10(3) AND F15(2) THEN TTP	12	11	0.916	4.02E-06
	IF F2(2) AND F7 (2) AND F11(1) AND F13(2) THEN LBMS	12	10	0.857	0.000274
	IF F7(2) AND F8 (2) AND F9(1) AND F16(2) THEN AWP	16	13	0.813	7.05E-07
	IF F2(2) AND F3 (1) AND F10(1) AND F17(1) THEN CCPM	12	12	1	2.48E-08
Control metrics	IF F2(3) AND F10 (3) AND F12(1) THEN PPC	18	18	1	3.49E-28
	IF F6(3) AND F7 (3) THEN MV	15	15	1	2.57E-22
	IF F5(3) AND F10 (1) AND F12(1) THEN TA	13	13	1	1.14E-18
	IF F5(1) AND F10 (1) AND F11(1) AND F12(2) THEN CLR	13	10	0.769	4.54E-11
	IF F5(3) AND F10 (3) THEN TMR	20	20	1	1.79E-32
Schedule levels	IF F3(2) AND F19 (3) THEN Short-term	21	19	0.905	4.44E-07
	IF F8(2) AND F16 (3) AND F17(2) THEN Mid-term	14	13	0.929	0.00448
	IF F6(3) AND F17 (1) AND F19(1) THEN Long-term	13	13	1	0.000383

4.4. Case study and inference engine results

A renovation project targeting the campuses of the University of Lorraine was considered as a case study to demonstrate the practicality of the decision support system and mathematical model for suggesting a multi-level planning and control system. This initiative is part of a significant energy renovation project launched for the IUT Nancy-Brabois campus. Over two years, the project seeks to refurbish four departments and two workshops to enhance energy efficiency and foster a more conducive learning environment. This endeavor aligns with governmental initiatives aimed at reducing carbon emissions and upgrading public building infrastructure, thereby emphasizing a commitment to sustainability and the welfare of both students and staff, while concurrently advancing academic excellence. The layout and an overview of this case study are depicted in Fig. 11. It is worth noting that this case study was selected due to the inherent planning complexity associated with renovating educational campuses, where ongoing operations must be maintained throughout the renovation process. Also, effective stakeholder collaboration was required, given the involvement of numerous parties, which necessitates precise coordination to achieve project objectives. Furthermore, the project's scale, encompassing multiple buildings across the campus, demands a multi-level and collaborative planning approach to ensure efficient management.

Various discussions with the project team revealed the importance of adopting a systematic approach for planning and control of diverse aspects such as resources, workflows, logistical considerations, uncertainties, and promoting collaboration and coordination among subcontractors, among other relevant aspects. To systematically collect the functional requirements for the case study concerning a planning and control system, which will provide input data for the DSS, the project team was asked to outline their functional requirements (based on 19 functionalities) using a Likert scale ranging from 0 (not important) to 4 (very important) via the designed interface.

Following collecting the project team functional requirements as input data for the DSS and establishing the knowledge repository, the inference engine was initiated and fired the rules that their conditions were satisfied by the requirements, as illustrated in Fig. 12. Out of the 2280 rules stored within the knowledge database, 59 rules specifically related to project planning methods, control metrics, and schedule levels were activated and fired. Based on the fired rules, 4DBIM, TTP and LPS are suggested for planning methods across all three schedule levels. Moreover, MV, PPC, RL, and CLR are proposed as the control metrics that align more with the project control requirements. This activation highlights the dynamic capability of the inference engine to selectively apply relevant rules based on the contextual demands of the project team.

The following section presents the results of the mathematical model to suggest the feasible and optimize solutions provided by the decision

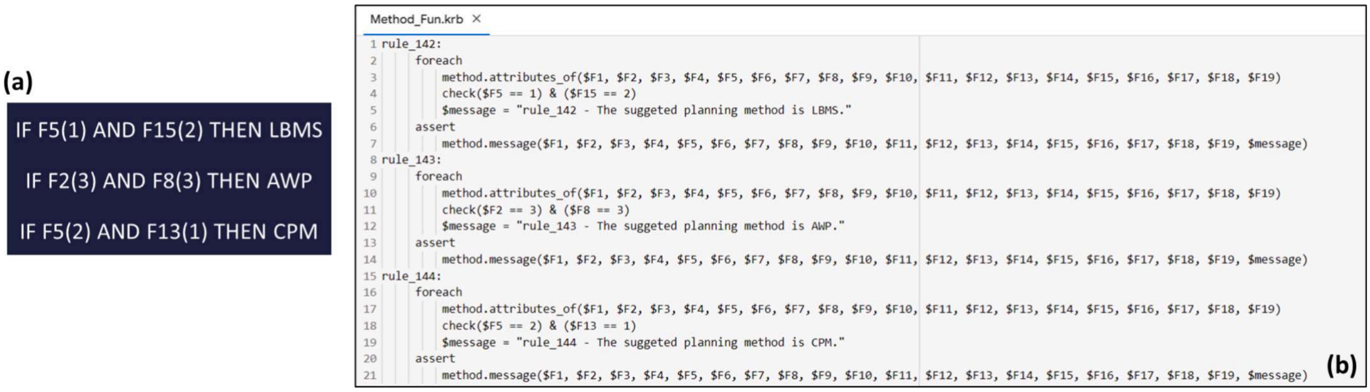


Fig. 10. (a) Examples of the generated rules' syntax through applying rough set theory using R.ROSETTA, (b) examples of the converted rules' syntax through applying a Python script, readable in Pyke engine.

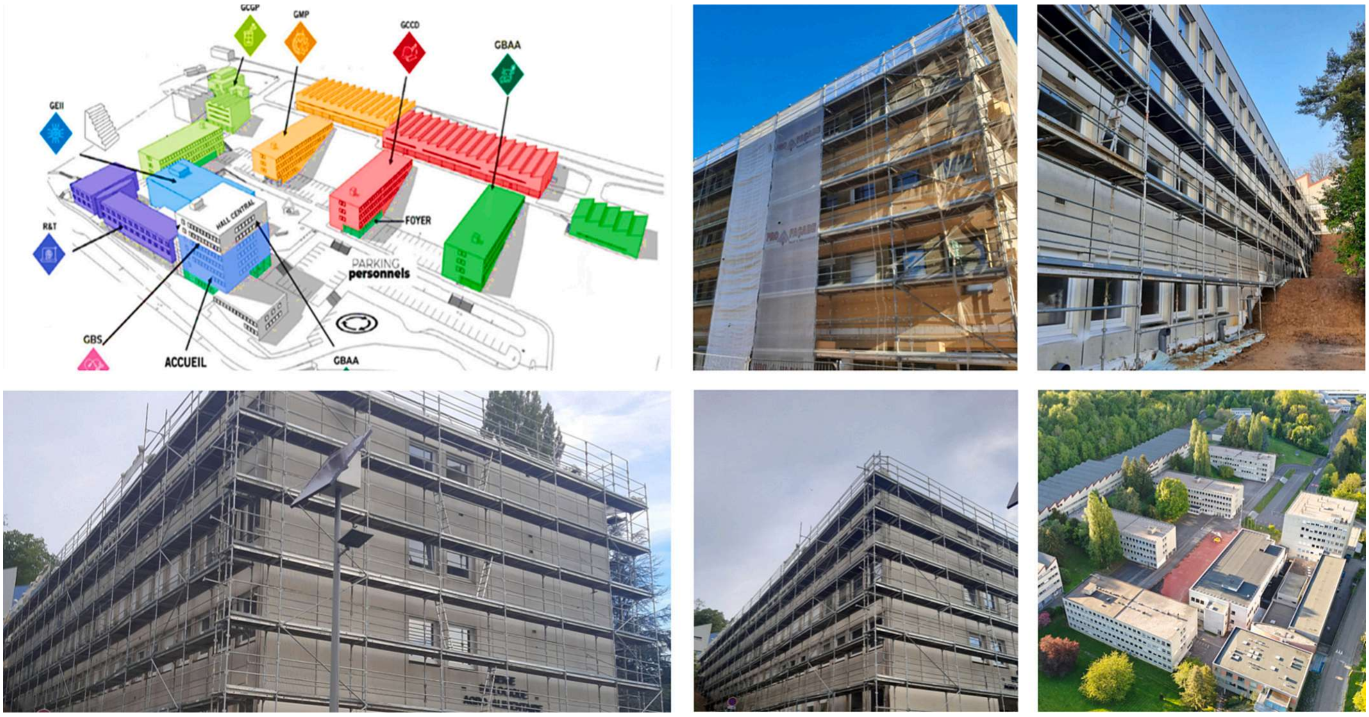


Fig. 11. Outline and layout of the case study.

support system for the case study.

4.5. Mathematical model results

Based on the collected data and the functional requirements outlined by the project team in the case study, the mathematical model was employed to propose feasible and optimized solutions provided by the DSS across various scheduling levels.

Fig. 13 illustrates the input data feasible and best-fitted scenarios, highlighting potential planning strategies and corresponding control metrics across various schedule levels for the case study. Specifically, Fig. 13 (a) displays the input data derived from parsing the information of activated rules and expert knowledge. Fig. 13 (b) outlines multiple scenarios, showcasing possible planning strategies and their respective control metrics at different schedule levels. Lastly, Fig. 13 (c) presents feasible and optimal solutions via Pareto front charts, representing the findings for the case study graphically. As previously mentioned, the project team selects the most suitable solution for each scheduling level, taking into account both project-specific factors and relevant external conditions. The project team provides the following rationale for their selection:

Scenario 1 with Z1 = 19 at the long-term level was chosen as the best fit for the case study, employing two planning methods: 4DBIM and takt time planning (TTP), along with a control metric: milestone variance (MV). This scenario was selected because it differed not much in terms of Z1 values from other scenarios while offering a more practical approach to implementation due to its reduced number of planning methods and control metrics. The use of 4DBIM provides a comprehensive visualization of the project timeline, while TTP ensures steady flow and progress. MV offers precise tracking of schedule efficiency and critical milestones, which is required for long-term windows. The scenario's advantage is its ease of use and reduced complexity, making it ideal for long-term schemes where maintaining focus and adaptability over time is crucial.

In the mid-term schedule level, takt time planning (TTP) and last planner system (LPS) were selected for planning methods, while location risk index (LRI) and task made ready (TMR) were considered as control metrics based on the project circumstances and external factors. This integration enables a more thorough and practical approach to project planning and control. TTP enhances the predictability and synchronization of work across the project, standardizing the work rhythm and facilitating efficient resource management. Complementing this, LPS involves team members in the planning process, ensuring plans are realistic and achievable through its collaboration. This adaptability is vital for mid-term planning, where project conditions can change rapidly. LRI aids in early risk identification, allowing for strategic resource allocation and effective risk mitigation. Similarly, TMR guarantees that all prerequisites for upcoming tasks are complete, thus enhancing task execution and minimizing potential delays. Together, these methods and metrics create a robust framework that significantly improves the effectiveness of lookahead planning by ensuring consistent workflow, collaborative planning, meticulous risk management, and thorough preparation.

In short-term planning, takt time planning (TTP) and last planner system (LPS) were again selected as the planning methods, complemented by percent planned complete (PPC) and capacity to load ratio (CLR) as control metrics. By aligning the work sequences with the overall project schedule, TTP optimizes workflow efficiency, preventing delays between tasks in tightly coordinated renovation projects. This approach supports the weekly work plan by establishing a clear, consistent pace that all team members can follow, ensuring tasks are completed within designated time slots. LPS augments the functionality of TTP by promoting a collaborative environment where daily and weekly work plans are developed through consensus among all project stakeholders. This system facilitates daily coordination by enabling immediate adjustments to the plan based on real-time feedback from the ground, ensuring that the project responds adaptively to any challenges or changes. This level of coordination is vital in renovation projects, where unexpected issues often arise and require quick decision-making and flexibility.

Furthermore, PPC as a control metric provides immediate feedback on the progress against the weekly plans, promoting a cycle of continuous improvement in planning accuracy and execution. CLR, on the other hand, ensures that the capacity of resources matches the demand

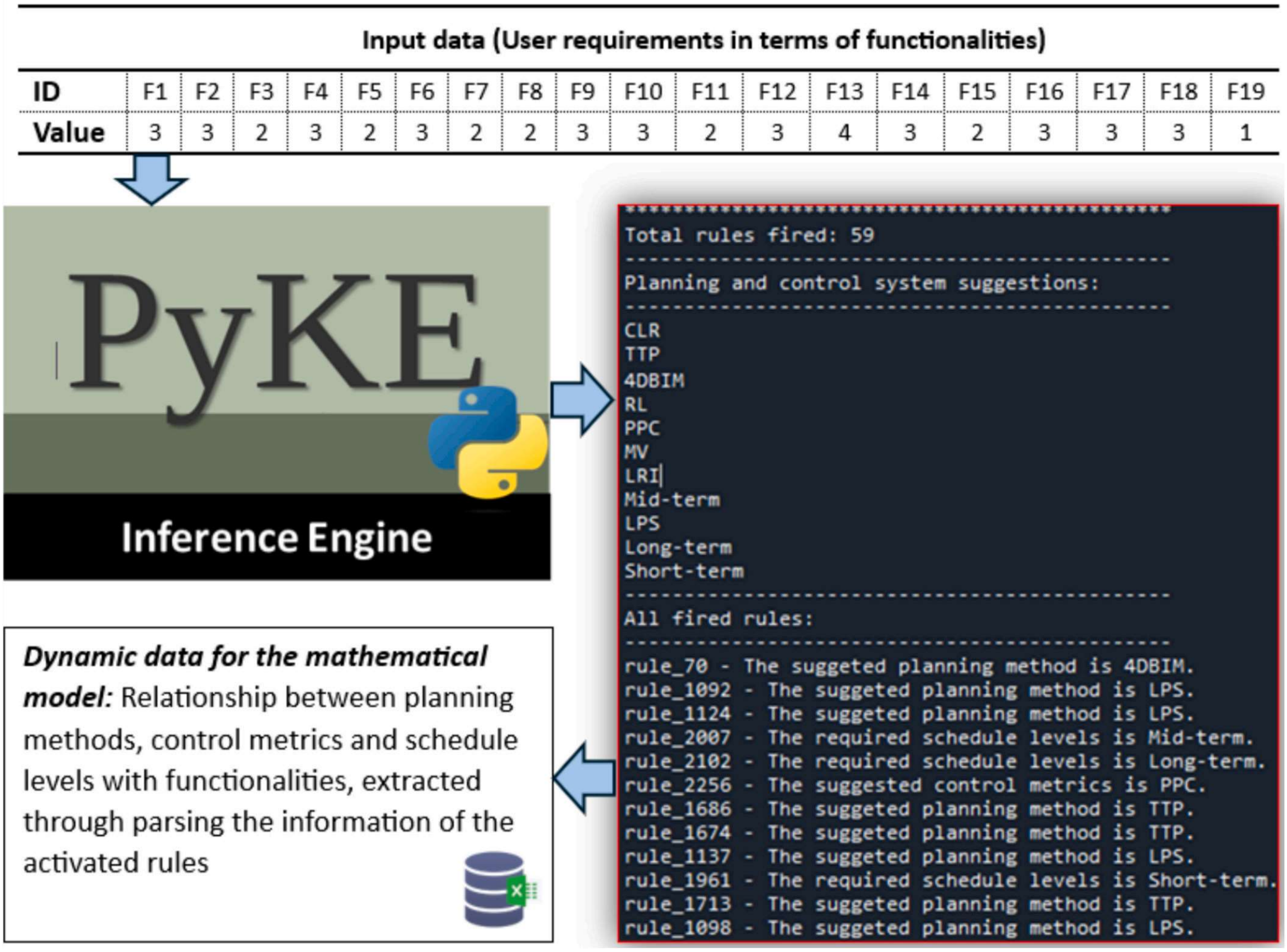


Fig. 12. Rule activation process through running the Pyke knowledge engine.

of the daily and weekly schedules, preventing overextension or underutilization of the workforce and materials. Together, TTP and LPS not only support the structuring of weekly work plans but also enhance daily coordination and operational efficiency, crucial for the dynamic and unpredictable nature of renovation projects.

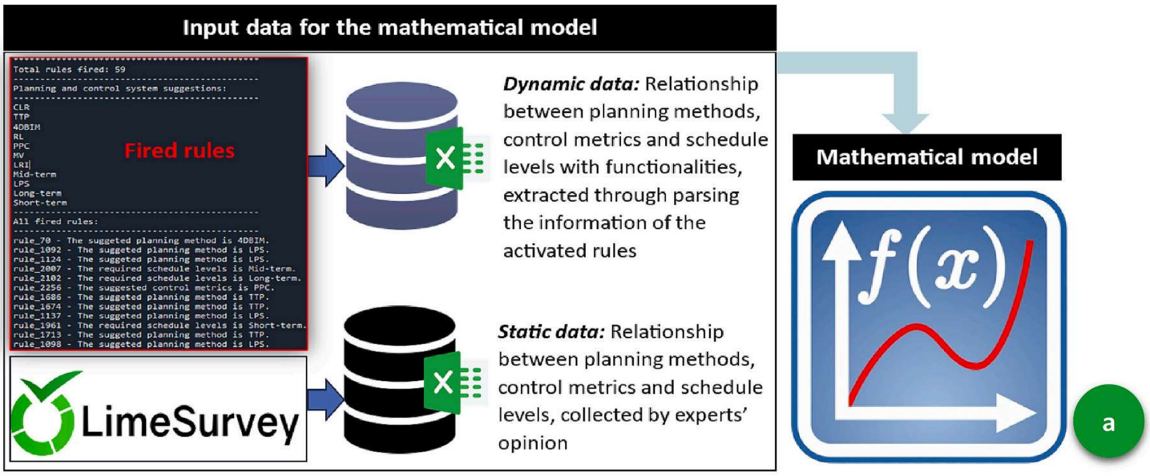
The overview of the optimized planning and control system for the case study, presented in Fig. 14, demonstrates a shift in perspective towards considering the project planning and control approach as a system. This study supports moving beyond conventional project planning methods to embrace a multi-functional system designed and implemented at various levels. By adopting this comprehensive approach, the study aims to revolutionize traditional thinking about project planning and control, emphasizing the need for a holistic and integrated system that addresses the complexities and dynamics of construction projects. This multi-level and integrated approach ensures more effective management and oversight, facilitating the achievement of project objectives through enhanced coordination and efficiency.

Regarding the engineering insights, this study adopts a data-driven and knowledge-based approach through a comprehensive and integrated process that leverages the mixed-method methodology to extract and utilize the expertise of construction professionals. By combining qualitative insights from industry experts with quantitative data analysis, the research develops a robust decision support system. This approach ensures that the practical experiences and knowledge of engineers are systematically captured and applied, leading to more accurate, reliable, and adaptive project management solutions.

To successfully implement the suggested solutions, it is critical to meticulously plan the process and workflow for integrating the proposed methods at each scheduling level, identify the responsible parties, specify the necessary reference and exchange information, and detail the required steps and activities. These essential elements for the seamless adoption of the suggested system, fall outside the scope of this research. To provide a practical guide on implementing the proposed system by DSS in practice, a methodological guideline will be developed in a subsequent step. This guideline will aim to ensure that the theoretical strategies are effectively translated into actionable, efficient practices within project management environments.

4.6. General evaluation of the proposed DSS

Although the application of the DSS to the case study has indicated its applicability, for a more comprehensive and general assessment, the DSS's overall performance was evaluated by experts using a Likert scale. This evaluation focused on several criteria, including ease of use, comprehensiveness, decision quality improvement, interface quality, and response time. As depicted in Table 13, user satisfaction across all criteria ranged between 3 and 5 out of 5. Experts generally recognized the practicality and user-friendliness of the MPCS-DSS, highlighting its consistent outputs and its role in augmenting decision-making processes within project teams concerning the selection of planning and control systems. However, limitations have been raised about the knowledge database's comprehensiveness. The knowledge repository was constru



Feasible scenarios and selected planning methods and control metrics at each schedule level for the case study

Max. Functionalities (Z1)	No. PM (Z2)	No. CM (Z3)	Planning Methods (PM)								Control Metrics (CM)								
			4DBIM	AWP	CPM	CCPM	LBMS	LSM	LPS	TTP	CLR	CPI	LRI	MV	PPC	RL	SPI	TA	TMR
Long-term level																			
19	2	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0
29	3	1	1	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0
29	4	1	1	0	1	1	0	0	0	1	0	1	0	0	0	0	0	0	0
29	5	1	1	1	1	1	0	0	0	1	0	1	0	0	0	0	0	0	0
29	6	1	1	1	1	1	0	1	0	1	0	0	0	0	0	0	1	0	0
29	6	2	1	1	1	1	0	1	0	1	0	1	0	0	0	0	1	0	0
Mid-term level																			
26	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0
45	2	2	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	1
55	3	1	1	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0
55	4	2	1	0	0	1	0	0	1	1	0	1	1	0	0	0	0	0	0
55	5	2	1	0	0	1	1	0	1	1	0	1	1	0	0	0	0	0	0
55	6	3	1	1	0	1	1	0	1	1	0	1	1	0	0	0	1	0	0
Short-term level																			
30	1	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0
36	1	2	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0
49	2	1	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0
55	2	2	0	0	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0
49	3	1	0	0	0	0	1	0	1	1	0	0	0	0	1	0	0	0	0
55	3	2	0	0	0	0	1	0	1	1	1	0	0	0	1	0	0	0	0

b

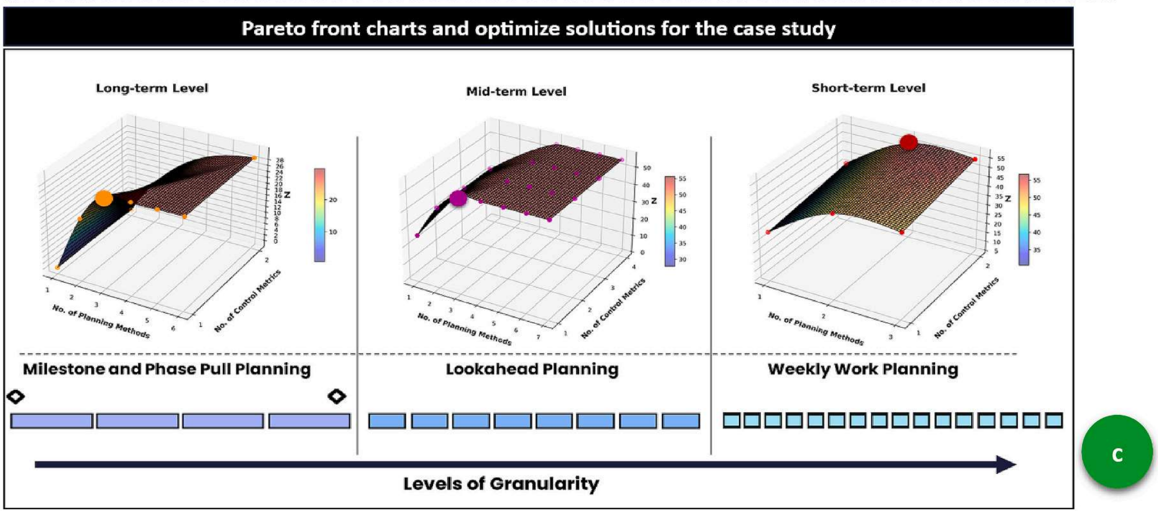


Fig. 13. (a) Input data for the mathematical model, (b) potential scenarios at each schedule level, (c) Pareto front plots of the findings.

rthermore, given the specific nature of control metrics, which cover a limited scope of functionalities, this segment of the knowledge repository contained merely 80 rules, which may have impacted its breadth. Despite these challenges, the satisfaction levels with the comprehensiveness of the DSS still surpassed the threshold of 3, which was deemed acceptable within the context of this study.

5. Research discussion and implications

This discussion tries to synthesize the key aspects and contributions

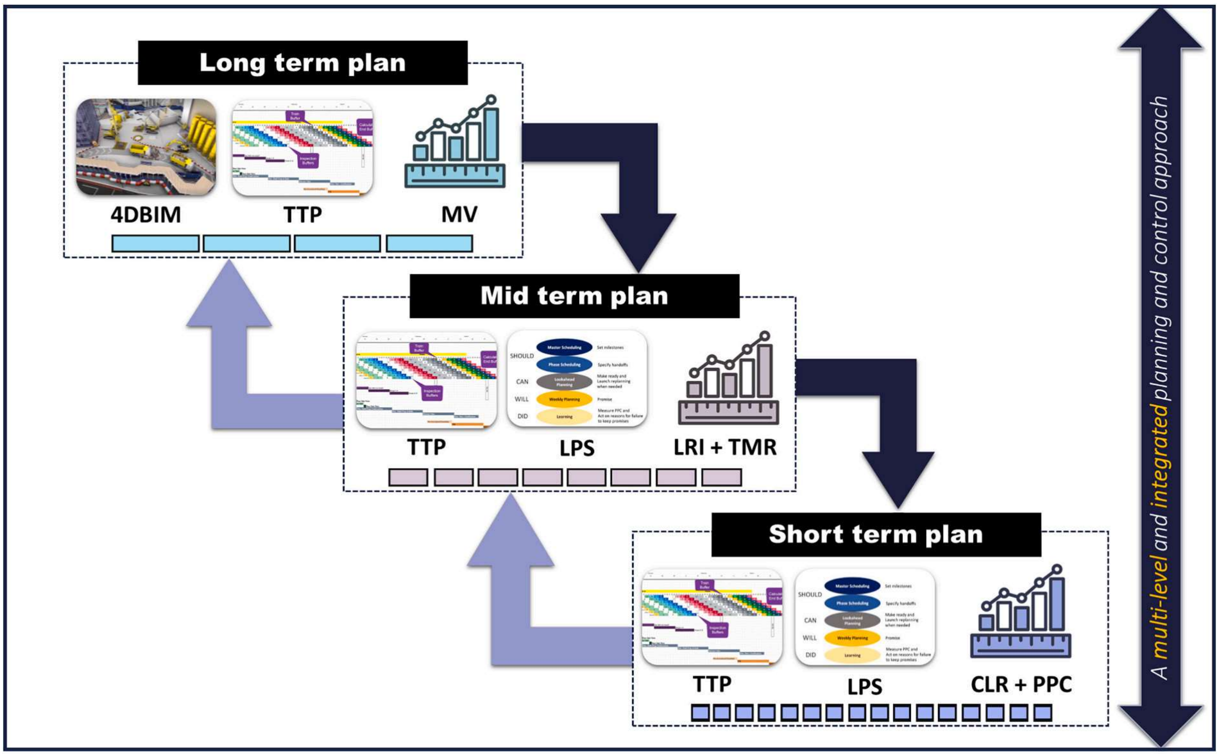


Fig. 14. Suggested multi-level and integrated solution for the case study.

Table 13
Experts' evaluations in 5-point Likert-scale.

Evaluation criteria	Likert scale					Average
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	
	1	2	3	4	5	
Ease of use						4.8
Comprehensiveness						3.4
Decision quality improvement						4
Interface quality						4.4
Response time						5

of the paper, highlighting its significance, innovations, and implications within the field of project management. This study focuses on the development of a decision support system for multi-level and integrated project planning and control systems. By bridging theoretical frameworks with practical applications, this study marks a pivotal step forward in storing and applying the experts' knowledge to propose planning and control systems for construction projects. To do so, the research methodology adopted a mixed-method approach, integrating data collection, expert system development, and result optimization. First, essential insights were gathered through semi-structured interviews and surveys conducted by domain experts. The expert system development phase involved building a rule-based knowledge repository using rough set theory (RST) and Pyke—a Python-based knowledge engine—to infer recommendations based on project requirements. The following phases involved launching an inference engine using forward chaining techniques as well as structuring and optimizing the suggestions by DSS through a mathematical model.

A key aspect of this research was the application of rough set theory (RST) to analyze the collected data systematically. RST was a crucial component in discovering and extracting the essential knowledge from experts, effectively capturing the complexities inherent in construction project planning and control systems. By structuring the data within an information system and employing concepts such as indiscernibility relations, lower and upper approximations, and attribute reducts, RST facilitated data exploration and knowledge extraction. This method not only preserved the expert insights but also ensured that the resulting decision rules were both relevant and robust. To ensure the knowledge repository's robustness, stringent criteria were established for rule selection, including *P*-value thresholds, accuracy benchmarks, and support metrics. This meticulous selection process was pivotal in retaining only the most pertinent and reliable rules, which crucially mitigated the risk of overfitting and minimized noise within the decision-making process. Consequently, this enhanced the reliability and applicability of the DSS. The high accuracy and validity of the selected rules further warranted their effectiveness in accurately representing expert knowledge and facilitating informed decision-making. Furthermore, the system's practicality was demonstrated through a case study involving a renovation project. This practical application highlighted the DSS's effectiveness in navigating complex project requirements and adapting to varied planning and control needs.

Although the developed DSS suggests a combination of advanced planning methods and control metrics as a system for project planning and control, it asks very simple questions based on tangible functional requirements that even non-experts can understand. The final solution is not only an advanced system, but it is also tailored for those who might not have extensive information and knowledge about these methods. The DSS provides a suggested system as an input for methodological guidelines. Domain experts can then use these suggested systems to develop step-by-step methodological guidelines, explaining in detail how the proposed approach by the DSS can be implemented in projects. This ensures that even non-experts can effectively utilize the DSS and implement its recommendations in a practical and comprehensible manner.

While in this research a renovation project was selected as a case study to demonstrate the effectiveness of the developed DSS, this tool can be generalized for different project types, including infrastructure projects, new construction projects, and industrial projects. The developed knowledge repository is not dependent on project types but rather on planning and control systems. By importing the functional requirements of project teams for various project types, the DSS can provide the most appropriate results. This approach ensures the system's applicability and repeatability across diverse case study projects, allowing it to be used effectively in various project management scenarios. Eventually, the system's design and functionality were evaluated on multiple dimensions, including ease of use, comprehensiveness, decision quality improvement, interface quality, and response time. The evaluation revealed generally high user satisfaction, with scores ranging from 3 to 5 on a 5-point scale. These results underscore the practicality and user-friendliness of the MPCS-DSS, particularly in its ability to deliver reliable outputs that enhance decision-making processes.

The originality of this study lies not only in developing a DSS for construction projects but also in proposing a replicable, integrated methodology for creating robust knowledge-based systems applicable across various domains. This mixed-method approach is particularly innovative in seamlessly integrating qualitative and quantitative techniques. By combining expert knowledge with advanced computational tools, the methodology establishes a tightly interconnected framework. The study's unique contribution stems from its ability to merge traditional expert systems with forward-chaining inference engines, data-driven rough set theory, and mathematical optimization models. These interconnected layers of analysis ensure that the knowledge base is not only comprehensive but also adaptable, allowing the system to evolve and provide optimized context-specific recommendations. This holistic approach offers a scalable model that other fields can adapt to develop similar decision support frameworks, thereby advancing both the theory and practice of intelligent system development.

The processes of knowledge extraction and DSS development in this research significantly contribute to the academic environment by demonstrating how rough set theory can be effectively used to handle uncertain data, a characteristic inherent to the built environment, and to harness the expertise of industry professionals. This approach is crucial in construction, a sector that heavily relies on experience, best practices, and lessons learned. Moreover, the integration of mathematical models with DSS not only showcases how theoretical applications can provide practical solutions but also suggests an optimized approach for handling complex decision-making processes in construction management. This study not only bridges the gap between theoretical research and practical application but also advances the understanding of adaptive decision-making frameworks that can cater to the dynamic nature of construction projects.

Regarding the practical implications, as several studies have shown [5,39], there is generally a low level of knowledge, understanding, and familiarity among project stakeholders regarding planning and control systems in construction. Addressing this challenge, the developed DSS simplifies the decision-making process by asking straightforward and sensible questions tailored to the project team's requirements, thus suggesting the best approach for project planning and control. This makes it a versatile tool that can be applied across various project types during the preconstruction phase to determine the most effective planning and control strategies based on team inputs. Furthermore, by presenting the results in a Pareto front plot, the DSS offers the project team multiple scenarios, providing them the flexibility to choose the most suitable option in light of specific project conditions and constraints. This adaptability enhances decision-making efficacy and promotes a more informed selection process, leading to optimized project outcomes and better alignment with strategic objectives.

While the developed DSS provides a valuable tool for project planning and control, several limitations must be acknowledged when considering its application in real-world engineering projects. Although the DSS offers multiple scenarios for integrating planning methods and control metrics at each scheduling level, the selection of the most appropriate solution is influenced by various factors, including the specific characteristics of the project, the maturity level of the organization and its stakeholders in planning and control systems, and external environmental conditions such as legal and regulatory frameworks. Due to the complexity of accounting for these variables, expert input may still be required to ensure the selection of the most suitable scenario. Additionally, implementing the multi-level and integrated system proposed by the DSS in actual construction projects would require a detailed, step-by-step methodological guideline, which falls outside the scope of this research. Future work could focus on developing such practical guidelines to facilitate the system's application. Additionally, the knowledge repository of the DSS is somewhat generalized due to the limited availability of experts with comprehensive knowledge of various planning and control systems tailored to specific project types. As such, some functionalities or rules may not be fully relevant to specific project types and would need further customization for optimal applicability.

6. Conclusion, limitations and future directions

This paper successfully demonstrated the development and practical application of a data-driven and knowledge-based decision support system for multi-level planning and control in construction projects. The study utilized rough set theory integrated with a Python-based knowledge engine, Pyke, to develop a rule-based expert system that systematically leverages experts' knowledge and builds a knowledge database for the DSS. The development of the DSS involved formulating a multi-objective mathematical model designed to enhance decision-making by evaluating various feasible solutions simultaneously. This model leveraged the outputs from the DSS—primarily the recommended planning methods and control metrics—to generate a set of optimized solutions that balance competing project requirements. The optimization component of the DSS was implemented through Pareto front plots, which are critical in multi-criteria decision-making. These plots visually represent the trade-offs between different objectives, allowing decision-makers to understand the implications of various choices and select the most appropriate strategies based on specific project needs. The case study involving a renovation project at the University of Lorraine showcased the DSS's capability to suggest a multi-level planning and control system and adapt to the specific needs of the project.

This study has several limitations that should be acknowledged. First, the data collection was limited to 23 experts, which, while sufficient for this study's scope, might not fully represent the broader range of expertise available in the construction management field. Another significant limitation is that although the DSS and mathematical model suggest a multi-level planning and control system for construction projects, there is a need for a methodological guideline to assist the project team in implementing the suggested solutions by the DSS. This aspect falls outside the scope of this study, and future research could focus on developing these methodological guidelines, which are crucial for the practical application and operational success of the DSS results in real-world settings.

Considering the scope and successful application of the DSS in this study, future research could expand in several directions. Investigating the application of the DSS in different types of construction projects, including new construction and infrastructure projects, could validate the system's adaptability and effectiveness across various contexts. Additionally, exploring alternative methods for DSS development, such as case-based systems, could offer valuable perspectives on different approaches and their respective benefits. Another promising direction involves investigating optimization algorithms beyond the current methods. Exploring alternative algorithms like metahuristic algorithms to identify the best fit could improve the robustness and effectiveness of the DSS results. Furthermore, extending the analysis to include Bayesian networks could introduce additional flexibility to the inferences produced by the system. Integrating insights from studies like Feng, et al. [14] could enrich the research and offer new avenues for enhancing the DSS's capabilities.

Future research also could explore the development of next-generation DSSs by combining rule-based expert systems with Reinforcement Learning with Human Feedback (RLHF). Such a hybrid approach would enable dynamic adjustments to rules, via penalties or rewards informed by expert insights, to prioritize the activation of contextually appropriate rules.

In conclusion, this paper marks a significant step forward in the application of decision support systems within construction management. It offers a robust framework for enhancing project outcomes by proposing multi-level and integrated planning and control systems. The implications of this research are far-reaching, promising to influence both current practices and future innovations in the construction industry. Furthermore, this framework can be applied to other case studies, demonstrating its scalability and potential for broader impact across various project management contexts.

Grant information

This research took place at the Nancy Automation Research Center (CRAN), and was made possible by funding from the National Research Agency of France under the IsoBIM project (grant number ANR-20-CE10-0011).

CRediT authorship contribution statement

Moslem Sheikhkhoshkar: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Hind Bril El-Haouzi:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition. **Alexis Aubry:** Writing – review & editing, Supervision, Methodology, Investigation, Formal analysis. **Farook Hamzeh:** Writing – review & editing, Supervision, Methodology. **Farzad Rahimian:** Writing – original draft, Supervision, Methodology, Investigation, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The research was carried out at the Nancy Automation Research Center (CRAN) and was financially supported by the National Research Agency of France for the IsoBIM project (grant number ANR-20-CE10-0011).

We acknowledge all participants for their valuable input in the data collection of this research. Participation was entirely voluntary, and respondents provided informed consent by signing a consent form before data. This study followed the standard ethical clearance protocols of the University of Lorraine, ensuring compliance with institutional guidelines. All responses were anonymized, and confidentiality was strictly maintained, with data access limited to the research team.

According to Article L. 1121-1 and Article R1121-1-II of the French Public Health Code (CSP), research that does not contribute to the development of biological or medical knowledge is classified as non-RIPH. Additionally, the Arrêté du 12 avril 2018 (Annex 1, Point 8) specifies that interviews and questionnaires that do not pose risks to participants fall outside the scope of Recherches Impliquant la Personne Humaine (RIPH). Since this study only involved low-risk, structured interviews and did not collect sensitive personal data or include vulnerable populations, it met the criteria for non-RIPH classification. Therefore, in line with these legal provisions and the University of Lorraine's ethical research guidelines, formal approval from an Institutional Review Board (IRB) or ethics committee was not required.

Additionally, in compliance with the General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679), Recital 26, anonymized data is exempt from data protection rules. This study also followed ethical research principles, including transparency, fairness, and data minimization, as outlined in GDPR Article 5 and the European Code of Conduct for Research Integrity (ALLEA, 2017).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.autcon.2025.106066.

Data availability

Data will be made available on request.

References

[1] S.O. Abioye, L.O. Oyedele, L. Akanbi, A. Ajayi, J.M.D. Delgado, M. Bilal, O. O. Akinade, A. Ahmed, Artificial intelligence in the construction industry: a review of present status, opportunities and future challenges, J. Build. Eng. 44 (2021) 103299, https://doi.org/10.1016/j.jobe.2021.103299.

[2] B.S. Adamczyk, A.L. Szejka, O.C. Júnior, Knowledge-based expert system to support the semantic interoperability in smart manufacturing, Comput. Ind. 115 (2020) 103161, https://doi.org/10.1016/j.compind.2019.103161.

[3] S. Akbari, F. Pour Rahimian, M. Sheikhkhoshkar, S. Banihashemi, M. Khanzadi, Dynamic sustainable success prediction model for infrastructure projects: a rough set based fuzzy inference system, Constr. Innov. 20 (4) (2020) 545–567, https://doi.org/10.1108/CI-04-2019-0034.

[4] A. Al-Ajlan, The comparison between forward and backward chaining, Int. J. Mach. Learn. Comput. 5 (2) (2015) 106, https://doi.org/10.7763/IJMLC.2015.V5.492.

[5] H. AlNasseri, R. Aulin, Assessing understanding of planning and scheduling theory and practice on construction projects, Eng. Manag. J. 27 (2) (2015) 58–72, https://doi.org/10.1080/10429247.2015.1035963.

[6] J.R. Anaraki, M. Eftekhari, Rough set based feature selection: A review, in: The 5th Conference on Information and Knowledge Technology, IEEE, 2013, pp. 301–306, https://doi.org/10.1109/IKT.2013.6620083.

[7] G. Ballard, I. Tommelein, Current process benchmark for the last planner system of project planning and control, 2021, 2020, https://doi.org/10.34942/P2F593.

[8] H.G. Ballard, The Last Planner System of Production Control, School of Engineering, Department of Civil Engineering, University of Birmingham, 2000. https://etheses.bham.ac.uk/id/eprint/4789.

[9] B. Barkokebas, S. Khalife, M. Al-Hussein, F. Hamzeh, A BIM-lean framework for digitalisation of premanufacturing phases in offsite construction, Eng. Constr. Archit. Manag. 28 (8) (2021) 2155–2175, https://doi.org/10.1108/ECAM-11-2020-0986.

[10] L.J. Cronbach, Coefficient alpha and the internal structure of tests, psychometrika 16 (3) (1951) 297–334, https://doi.org/10.1007/BF02310555.

[11] P. Dallasega, E. Marengo, A. Revolti, Strengths and shortcomings of methodologies for production planning and control of construction projects: a systematic literature review and future perspectives, Prod. Plan. Control 32 (4) (2021) 257–282, https://doi.org/10.1080/09537287.2020.1725170.

[12] I. Dikmen, G. Atasoy, H. Erol, H.D. Kaya, M.T. Birgonul, A decision-support tool for risk and complexity assessment and visualization in construction projects, Comput. Ind. 141 (2022) 103694, https://doi.org/10.1016/j.compind.2022.103694.

[13] N. Fatfouta, J.S. Le-Cardinal, An ontology-based knowledge management approach supporting simulation-aided design for car crash simulation in the development phase, Comput. Ind. 125 (2021) 103344, https://doi.org/10.1016/j.compind.2020.103344.

[14] Z. Feng, X. Gao, L. Wang, Bayesian network structure learning based on rough set and mutual information, TELKOMNIKA Indones. J. Electric. Eng. 12 (2) (2014) 1596–1601.

[15] A. Frandson, I.D. Tommelein, Development of a takt-time plan: a case study, construction research congress, Construct. Glob. Netw. 2014 (2014) 1646–1655, https://doi.org/10.1061/9780784413517.168.
[16] B. Frederiksen, Applying Expert System Technology to Code Reuse with Pyke, PyCon, Chicago, 2008. https://pyke.sourceforge.net/PyCon2008-paper.html.
[17] M. Garbulowski, K. Diamanti, K. Smolińska, N. Baltzer, P. Stoll, S. Bornelöv, A. Øhrn, L. Feuk, J. Komorowski, R., ROSETTA: an interpretable machine learning framework, BMC Bioinforma. 22 (2021) 1–18, https://doi.org/10.1186/s12859-021-04049-z.
[18] M. Hajdasz, Flexible management of repetitive construction processes by an intelligent support system, Expert Syst. Appl. 41 (4) (2014) 962–973, https://doi.org/10.1016/j.eswa.2013.06.063.
[19] F. Hamzeh, Improving Construction Workflow-the Role of Production Planning and Control, University of California, Berkeley, 2009, https://doi.org/10.13140/RG.2.1.1182.9286.
[20] F.R. Hamzeh, G. El Samad, S. Emdanat, Advanced metrics for construction planning, J. Constr. Eng. Manag. 145 (11) (2019) 04019063, https://doi.org/10.1061/(ASCE)CO.1943-7862.0001702.
[21] H. He, E.A. Garcia, Learning from imbalanced data, IEEE Trans. Knowl. Data Eng. 21 (9) (2009) 1263–1284, https://doi.org/10.1109/TKDE.2008.239.
[22] C. Hendrickson, C. Zozaya-Gorostiza, D. Rehak, E. Baracco-Miller, P. Lim, Expert system for construction planning, J. Comput. Civ. Eng. 1 (4) (1987) 253–269, https://doi.org/10.1061/(ASCE)0887-3801(1987)1:4(253).
[23] M.M. Hennink, B.N. Kaiser, V.C. Marconi, Code saturation versus meaning saturation: how many interviews are enough? Qual. Health Res. 27 (4) (2017) 591–608, https://doi.org/10.1177/1049732316665344.
[24] B.-G. Hwang, M. Shan, K.-Y. Looi, Knowledge-based decision support system for prefabricated prefinished volumetric construction, Autom. Constr. 94 (2018) 168–178, https://doi.org/10.1016/j.autcon.2018.06.016.
[25] K. Jahr, A. Borrmann, Semi-automated site equipment selection and configuration through formal knowledge representation and inference, Adv. Eng. Inform. 38 (2018) 488–500, https://doi.org/10.1016/j.aei.2018.08.015.
[26] J. Jepson, K. Kirytopoulos, K. London, Insights into the application of risk tools and techniques by construction project managers, Int. J. Constr. Manag. 20 (8) (2020) 848–866, https://doi.org/10.1080/15623599.2018.1494673.
[27] D.S. Johnson, Approximation algorithms for combinatorial problems, in: Proceedings of the Fifth Annual ACM Symposium on Theory of Computing, 1973, pp. 38–49, https://doi.org/10.1145/800125.804034.
[28] T. Karimi, A. Hojati, J.Y.-L. Forrest, A new methodology for sustainability measurement of banks based on rough set theory, CEJOR (2022) 1–17, https://doi.org/10.1007/s10100-020-00698-2.
[29] J.E. Kelley Jr., M.R. Walker, Critical-path planning and scheduling, in: Papers Presented at the December 1–3, 1959, Eastern Joint IRE-AIEE-ACM Computer Conference, 1959, pp. 160–173, https://doi.org/10.1145/1460299.1460318.
[30] C. Liu, V.A. González, G. Lee, G. Cabrera-Guerrero, Y. Zou, R. Davies, Integrating the last planner system and immersive virtual reality: exploring the social mechanisms produced by using LPS in projects, J. Constr. Eng. Manag. 150 (7) (2024) 04024070, https://doi.org/10.1061/JCEMD4.COENG-14430.
[31] A.N. Mohamed, Knowledge based approach for productivity adjusted construction schedule, Expert Syst. Appl. 21 (2) (2001) 87–97, https://doi.org/10.1016/S0957-4174(01)00029-X.
[32] M. Najafi, M. Sheikhkhoshkar, F. Rahimian, Innovation and lean practices for sustainable construction project management; emerging technologies, strategies and challenges, Smart Sustain. Built Environ. 13 (3) (2024) 473–478, https://doi.org/10.1108/SASBE-05-2024-406.
[33] O. Okudan, C. Budayan, I. Dikmen, A knowledge-based risk management tool for construction projects using case-based reasoning, Expert Syst. Appl. 173 (2021) 114776, https://doi.org/10.1016/j.eswa.2021.114776.
[34] H. Olivieri, O. Seppänen, A.D. Granja, Integrating Lbms, Lps and Cpm: A practical process, in: Annual Conference of the International Group for Lean Construction, National Pingtung University of Science and Technology, 2016, pp. 3–12. https://www.iglc.net/Papers/Conference/26.
[35] Z. Pawlak, Rough sets, Int. J. Comput. Inform. Sci. 11 (5) (1982) 341–356, https://doi.org/10.1007/BF01001956.
[36] K. Peffers, T. Tuunanen, M.A. Rothenberger, S. Chatterjee, A design science research methodology for information systems research, J. Manag. Inf. Syst. 24 (3) (2007) 45–77, https://doi.org/10.2753/MIS0742-1222240302.
[37] R. Pellerin, N. Perrier, A review of methods, techniques and tools for project planning and control, Int. J. Prod. Res. 57 (7) (2019) 2160–2178, https://doi.org/10.1080/00207543.2018.1524168.
[38] A. Rashidi, W.Y. Yong, D. Maxwell, Y. Fang, Construction planning through 4D BIM-based virtual reality for light steel framing building projects, Smart Sustain. Built Environ. 12 (5) (2023) 1153–1173, https://doi.org/10.1108/SASBE-06-2022-0127.
[39] S.T. Salling, C.T. Perez, S. Wandahl, Perception of project management among construction workers: A survey in Denmark, in: Proceedings of the 31st Annual Conference of the International Group for Lean Construction, 2023, pp. 882–893, https://doi.org/10.24928/2023/0124.
[40] O. Seppänen, G. Ballard, S. Pesonen, The combination of last planner system and location-based management system, Lean Construct. J. 6 (1) (2010) 43–54. https://www.researchgate.net/publication/228417459.
[41] L. Shehab, Z. Malaeb, F. Hamzeh, Harnessing LPS metrics for smarter resource allocation and project control through gamification, Lean Construct. J. (2024) 01–15. www.leanconstructionjournal.org.
[42] M. Sheikhkhoshkar, H. Bril El-Haouzi, A. Aubry, F. Hamzeh, Functionality as a key concept for integrated project planning and scheduling methods, J. Constr. Eng. Manag. 149 (7) (2023) 04023053, https://doi.org/10.1061/JCEMD4.COENG-13427.
[43] M. Sheikhkhoshkar, H. Bril El-Haouzi, A. Aubry, F. Hamzeh, F. Rahimian, From NLP to Taxonomy: Identifying and Classifying Key Functionality Concepts of Multi-level Project Planning and Control Systems, Journal of Information Technology in Construction (ITcon), 29, (Special issue: Managing the digital transformation of construction industry (CONVR 2023)). https://doi.org/10.36680/j.itcon.2024.053.
[44] M. Sheikhkhoshkar, H. Bril El-Haouzi, A. Aubry, F. Hamzeh, F. Rahimian, From process-based to technology-driven: a study on functionalities as key elements of collaborative planning methods for construction projects, Prod. Plan. Control (2024) 1–22, https://doi.org/10.1080/09537287.2024.2360581.
[45] M. Sheikhkhoshkar, H. Bril El Haouzi, A. Aubry, F. Hamzeh, An advanced exploration of functionalities as the underlying principles of construction control metrics, Smart Sustain. Built Environ. 13 (3) (2024) 644–676, https://doi.org/10.1108/SASBE-12-2023-0379.
[46] M. Sheikhkhoshkar, H.B. El-Haouzi, A. Aubry, F. Hamzeh, M. Poshdar, Analyzing the lean principles in integrated planning and scheduling methods, in: Proc., 31st Annual Conference of the International Group for Lean Construction, IGLC31, 2023, pp. 1196–1207, https://doi.org/10.24928/2023/0159.
[47] M. Sheikhkhoshkar, H.B. El-Haouzi, A. Aubry, F. Hamzeh, M. Sakikhales, An innovative integrated framework for multi-level production planning and control systems in construction, Dev. Built Environ. 100524 (2024), https://doi.org/10.1016/j.dibe.2024.100524.
[48] A. Skowron, C. Rauszer, The Discernibility Matrices and Functions in Information Systems, Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory, Springer, 1992, pp. 331–362, https://doi.org/10.1007/978-94-015-7975-9_21.
[49] H. Suri, Purposeful sampling in qualitative research synthesis, Qual. Res. J. 11 (2) (2011) 63–75, https://doi.org/10.3316/QRJ1102063.
[50] R.A. Ugulu, A. Arewa, S. Allen, Project-specific constraints influencing productivity of tradespeople in the Nigerian construction industry, Built Environ. Proj. Asset Manag. 10 (1) (2020) 94–109, https://doi.org/10.1108/BEPAM-10-2018-0124.
[51] S. Vinterbo, A. Øhrn, Minimal approximate hitting sets and rule templates, Int. J. Approx. Reason. 25 (2) (2000) 123–143, https://doi.org/10.1016/S0888-613X(00)00051-7.
[52] G. Yilmaz, L. Salter, D. McFarlane, B. Schönfuß, Low-cost (Shoestring) digital solution areas for enabling digitalisation in construction SMEs, Comput. Ind. 150 (2023) 103941, https://doi.org/10.1016/j.compind.2023.103941.
[53] S.L. Zulu, A.M. Saad, B. Gledson, Exploring leaders' perceptions of the business case for digitalisation in the construction industry, Buildings 13 (3) (2023) 701, https://doi.org/10.3390/buildings13030701.