A Literature Review on Analytical Methods in Construction Master Scheduling to Generate a Competitive Advantage

Svenja Lauble, svenja.lauble@kit.edu
Institute of Technology and Management in Construction, Karlsruhe Institute of Technology (KIT), Germany

Philipp Zielke, philipp.zielke@kit.edu
Institute of Technology and Management in Construction, Karlsruhe Institute of Technology (KIT), Germany

Shervin Haghsheno, shervin.haghsheno@kit.edu
Institute of Technology and Management in Construction, Karlsruhe Institute of Technology (KIT), Germany

Abstract
In the early phases of construction projects, initial client-set durations often face challenges due to evolving project dynamics. Despite extensive research on accurate planning, practical case studies indicate shortcomings in early-phase predictions, with analytical planning methods underutilized despite technological advancements.
This study investigates analytical methodologies used in research for master scheduling in construction projects, aiming to identify gaps and potentials for further research. A systematic literature review of 386 relevant sources categorized methodologies into statistical analyses, extrapolations, predictive techniques, and optimization methods. While statistical analyses are prevalent, predictive methods and optimizations are less explored, indicating a need to adopt advanced analytical methods for robust master scheduling models in construction projects. Addressing comprehensible methods at an advanced analytical level is crucial for enhancing effectiveness and applicability in this context.

Keywords: Master scheduling, analytics, artificial intelligence, literature review

1 Introduction

Master scheduling is crucial for the success of construction projects. It requires meticulous planning months or even years before construction begins. Stakeholders, such as owners, developers, and investors, use master scheduling to determine phase durations. This influences investment decisions and project outcomes throughout the design, construction, and operation stages. However, it is often the case that detailed information about the building and material quantities, as well as the construction companies involved, is unavailable during this initial phase. Therefore, it is imperative to conduct a comprehensive analysis of project features and characteristics (Walker 1995). Incomplete analyses or misweighted expressions can lead to unreliable milestones, which can hinder competitive advantage within the industry.

Research has focused on using analytical methods to predict project phase durations by analyzing relevant features. These methods enhance scheduling accuracy and support decision-making. Implementing analytical planning methodologies generates and applies knowledge cost-effectively, providing significant economic benefits. According to North's (2016) knowledge staircase, differentiating between signs, data, information, knowledge, and competence is crucial. Applied knowledge can become a unique competitive advantage. Thus, analytical planning methods help companies extract relevant information from large data volumes, gain valuable insights, and strengthen their market position.

In recent years, there has been a surge in digital integration, leveraging advancements in computing power and solutions such as the Internet of Things (IoT). Thorough analysis and connection of project data with external sources can unlock significant economic benefits.

However, the integration of analytical, data-driven methods is still limited (Erdis 2013), and there is a lack of literature exploring the competitive advantage of existing methods. Kanoglu (2003) proposes a classification of methods based on expert experience, parameters, and discrete simulations, while Erdis (2013) discusses data-driven methodologies such as multivariate regression and data mining. However, clear boundaries between these categories can pose challenges, highlighting the need for accessible knowledge-enhancing systems and further research. This paper explores analytical methods that support master scheduling and their contribution to a company's competitive advantage. This competitive advantage arises because in analytical methods the available data is used to identify even weak correlations, which helps the company make processes more efficient, save time and costs, and increase the quality of processes and products. A data-based analysis model delivers objective results that are free of distortions, reproducible, and adaptable in the event of deviations thanks to the application of rules. The paper therefore excludes forecasting methods that rely on participant experiences.

The following sections provide a detailed account of the systematic methodology (Section 2), content analysis and categorization of publications (Section 3), insight into current research and competitive advantage levels (Section 4), and identification of challenges and recommendations for future research (Section 5).

2 Methods and Materials

A systematic literature search was conducted to identify relevant literature on analytical methods in master scheduling of construction projects in German- and English-speaking countries. The search aimed to represent the current state of research and involved three phases (see Figure 1).

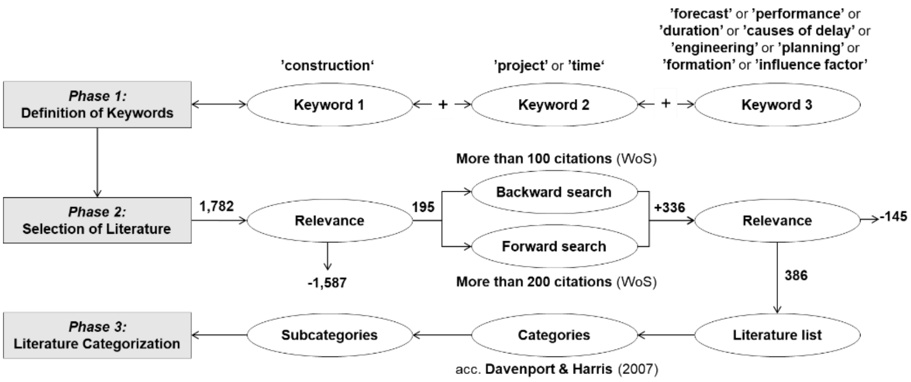

Figure 1. Overview of the systematic literature search

2.1 Definition of Keywords

The search terms used to identify analytical methods for predicting project phase durations and assessing planned/actual deviations included a combination of keywords such as 'construction,' 'project,' or 'time,' and terms like 'forecast,' 'performance,' or 'duration.' The search was conducted across construction engineering and project management journals, resulting in 1,782 publications.

2.2 Selection of Literature

The publications were identified and classified as relevant or non-relevant based on title evaluation. Titles that did not reference the construction industry were excluded. Abstracts were then evaluated for relevance, excluding those not related to construction scheduling or inaccessible. This process yielded 195 relevant publications. Furthermore, a forward and backward search was conducted using the software "research rabbit", which extended the list by 336 sources. Backward publications were included if they had more than 100 citations and forward citations were included if they had more than 200 citations. Non-relevant literature was again excluded (145 publications). The final literature list comprised 386 sources.

2.3 Literature Categorization

The purpose of categorizing literature is to provide an overview of analytical methods and their competitive advantages. Davenport and Harris (2007) categorized analysis methods into the following:

1. *Statistical correlations*: Determine the relationships between individual characteristics and outcome variables, clarifying their significance and robustness.
2. *Extrapolations*: Use stochastic methods to express mathematical relationships and answer the question, 'What will happen if the trend continues?'
3. *Predictive methods*: Analyze patterns in datasets, including unexpected events, to predict future outcomes.
4. *Optimization*: Determine optimal results based on previous stages to influence decision-making. This level offers the highest competitive advantage by answering the question, 'What should I do to achieve the best outcome for the project?'

The contributions of the literature to these stages were evaluated based on forecast accuracy. The evaluation's depth and the necessity for further research are determined by the number of literature sources used and the specific analytical method employed.

This methodology ensured a comprehensive collection of relevant sources while categorizing them to analyze analytical methods and their competitive advantage. This streamlined approach facilitated efficient literature selection and analysis, contributing to the research objectives. The literature list is not to be regarded as complete, but rather shows tendencies. The following section presents the analytical levels and introduces literature sources for a content analysis.

3 Classification of the Literature in a Taxonomy

When allocating the 386 publications to the four categories, it was found that most of the publications use methods of statistical analyses (229). The number of publications decreases with each successive step towards competitive advantage (Figure 2). Further on, ten publications appear in multiple categories. Therefore, the total number of publications in the four categories is 396 instead of the initial 386. These publications specifically compare predictive methods with less intelligent methods, statistical analyses, and extrapolations. Some publications also compare within the categories the prediction accuracy of different methods.

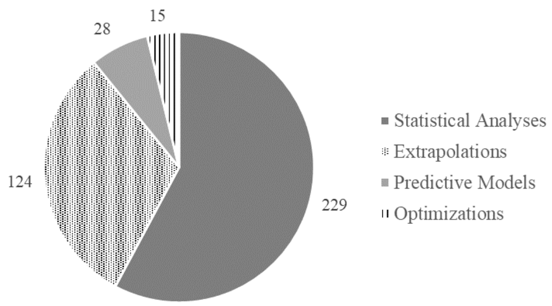

Figure 2. Number of publications identified in the categories by Davenport and Harris (2007)

3.1 Statistical Analyses

Statistical analyses, such as the Relative Importance Index (RII), correlations, and sensitivity analyses, are frequently used in master scheduling literature to identify influential characteristics that affect construction project durations.

In studies with over 100 citations, we analyzed the application of the RII, which retrospectively weights and ranks influential characteristics based on construction project samples. Respondents typically rate characteristics on a scale, and the RII is calculated by summing these ratings and dividing by the highest weight multiplied by the number of respondents (Chan & Kumaraswamy 1997).

Studies that implement the RII span various regions and focus on diverse project types, such as buildings, housing, water supply, and roads (Ogunlana et al 1996, Frimpong et al 2003, Kaliba et al 2009). Interviews with stakeholders, such as clients, project managers, planners, and contractors, have revealed common causes of time slippage. These include late changes, inadequate site management, and funding problems (Elinwa & Mangvwat 2001, Chan & Kumaraswamy 1997).

The RII identifies patterns instead of making direct predictions, providing guidance for subsequent projects. However, these findings may not be universally applicable due to regional or content-specific focuses. Nevertheless, they inform project planning and implementation, highlighting areas for attention early in construction projects. Overall, statistical analyses such as the RII provide valuable insights into understanding and addressing factors that influence construction project durations.

Additional diagnostic methods, such as correlations and sensitivity analyses, are considered statistical analyses. In Walker's (1995) study, individual factors were correlated with construction duration, demonstrating the impact of construction management performance. Kaming et al (1997) also investigated the influence of individual factors on time delays through correlations. Sensitivity analysis was used to gauge the time's sensitivity to variations in influencing factors. For example, Kumaraswamy & Chan (1995) analyzed the sensitivity of time to cost variations. Bromilow (1969) and Kumaraswamy & Chan (1995) conducted sensitivity analyses and identified a correlation between time and cost. These methods often precede extrapolations based on quantitative evaluations.

3.2 Extrapolations

Stochastic methods, which are based on statistics and probability theory, are crucial for inferring population characteristics from samples and establishing correlations between descriptive data and outcome variables. Common methods for extrapolation include multivariate regression, time series analysis, fuzzy systems, and Bayesian statistics.

Multivariate regression models are commonly used to mathematically express the relationships between project characteristics and time, allowing for the prediction of completion times for new building projects. Bromilow's model (1969) is an early example of estimating building completion time based on project cost. The formula used is $T=K*C^{B}$, where K is a constant for the project performance, T represents time in working days, C denotes the cost of the contract in millions, and B is a constant reflecting the sensitivity of time to cost.

Several models have been created to predict completion times in various regions. For example, Ireland's formula (1983) for Australian projects ($T = 219C^{0.47}$) and Chan's formulas (1999) for projects in Hong Kong ($T = 152C^{0.29}$) and Malaysia ($T = 269C^{0.32}$). However, these formulas have significant variations, resulting in significant differences in predicted completion times (Irfan et al 2011). Regressions frequently incorporate factors beyond construction costs to improve prediction accuracy, such as project and contract types.

Introduced by Lotfi Zadek in 1965, fuzzy systems represent uncertainties in construction projects by accepting intermediate values between 0 and 1. This linguistic reasoning-based approach allows for modeling complex systems that are more aligned with human thinking. Time series analysis, exemplified by the Box Jenkins method, integrates historical estimation errors to enhance regression robustness, especially with limited data sets (Zhang & Koreisha 2015). Lu and AbouRizk (2009) used the Box Jenkins method to estimate the duration of construction in large infrastructure projects.

Bayesian statistics, which assess the plausibility of statements using probability distributions, rely on prior knowledge and data. Kim and Reinschmidt (2009) demonstrated its usefulness in updating project schedules. Monte Carlo simulation, a Bayesian method, is also helpful in simulating complex problems (Albogamy et al 2003).

3.3 Predictive Models

Predictive analysis, especially through Artificial Intelligence (AI) methods, is essential in dynamic environments where discovering unknown patterns holds future significance. Unlike stochastic methods, AI methods focus on identifying patterns within complex data using versatile learning algorithms, facilitating predictions by extrapolating patterns from training data to new datasets (Bzdok et al 2018). AI methods are categorized into unsupervised and supervised learning. Unsupervised learning involves learning from input data without explicit feedback, while supervised learning involves learning relationships between input and output pairs (Kirste & Schürholz 2019).

Artificial Neural Networks (ANNs) are frequently used in supervised learning. Many studies have utilized ANNs, a machine learning technique that emulates the human brain, to predict construction project durations. Comparative analyses consistently demonstrate the superior predictive capability of ANNs over other methods. For example, Petruseva et al (2012) examined 75 construction projects and achieved a Mean Absolute Percent Error (MAPE) of 2.50 with ANN, compared to 10.36 with multivariate regression. Similar studies have also found that artificial neural networks (ANNs) outperform multivariate regression, as reported by Dissanayaka & Kumaraswamy (1999) and Chen & Huang (2006). Lam & Olalekan (2016) also noted improved prediction performance with ANN compared to the Box-Jenkins method. Additionally, ANN have been used to predict specific construction processes, such as residential construction by Naik & Kumar (2013) and shell construction activities by Golizadeh et al (2015).

AI methods are also used in ensemble techniques, which combine classification methods such as decision trees and support vector machines (SVMs) with ANN. SVMs, for example, aim to maximize the distance between data classes to achieve effective data separation, like classification expert systems. Erdis (2013) analyzed 878 and 575 public works projects to investigate the impact of new legislation on time and cost compliance in Turkish public works projects. Ensemble methods, such as decision trees, KNNs, and SVMs, are used to improve prediction accuracy in AI. This study, along with Wang et al (2012), highlights the benefits of using ensembles over individual AI techniques.

3.4 Optimization

To improve schedules, users are presented with alternative actions aimed at enhancing the current situation or eliminating specific tasks. The effects of characteristic choices on duration or costs are demonstrated using available data, with optimizations based on scenario evaluations and simulations during the predictive analysis phase.

In operational research for optimization, heuristics are important as they provide computationally feasible means to find satisfactory solutions, although not necessarily optimal ones (Nickel et al 2014). Heuristic methods focus on applying optimization techniques and decision-theoretic approaches, while AI methods employ rule-based and relational structures (Bockmayr & Radermacher 1993). Metaheuristics are a subtype of heuristics that are problem-agnostic and broadly applicable to problems with high uncertainty, such as master scheduling. They aim to enhance current solutions within a specific area based on local search principles (Nickel et al 2014). Genetic algorithms, introduced by Zheng et al (2004), iteratively select the best solutions for optimization problems. Genetic algorithms and ant algorithms have been applied in various studies, including those by El-Gafy (2007) and Zhang & Ng (2012), inspired by Darwin's theory of evolution.

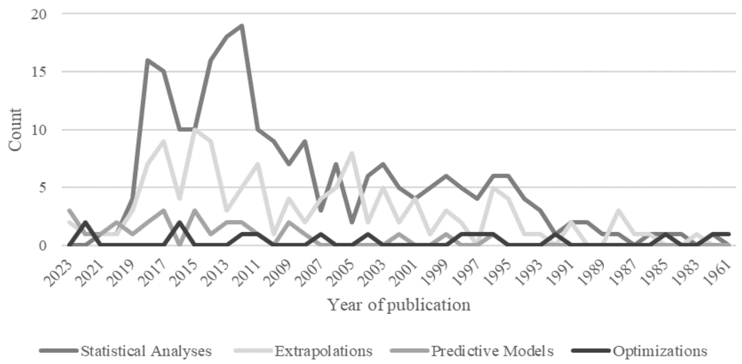

Figure 3. Number of analyzed publications in the categories by the year of publication.

In optimization problems, two metaheuristic algorithms used are tabu search and simulated annealing. Tabu search explores solution neighborhoods with predefined rules, while simulated annealing gradually decreases the acceptance of worse solutions. Jung et al (2016) used Tabu search in the construction domain, while Kirkpatrick et al (1983) introduced simulated annealing and demonstrated its use in construction scheduling in Kumar & Abdullah's (2011) work. Looking at the number of publications in each of the categories examined in relation to the year of publication, we see that statistical analysis has increased the most, while there are no clear trends, especially for predictive methods and optimization (Figure 3). The low number of analyzed publications in recent years (from 2019) could be due to the limitation of forward and backward searches to 100 and 200 citations, respectively, but also to the limited possibilities of analysis due to the Covid pandemic.

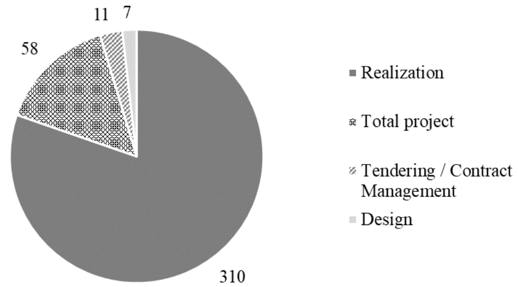

Figure 4. Focused phase of the analyzed publications

Publications can be further categorized based on the project phases they address, with a clear emphasis on the construction phase with project realization (310 publications) due to its significant variability between projects (Figure 4). Detailed documentation of construction information is a valuable source of data for analytical methods, especially for companies that rely on accurate schedules as the basis for contract documents. Overall project duration is analyzed and predicted (58 publications) by fewer authors. However, few authors explore upstream phases such as tendering and contract management (e.g., Migliaccio & Shrestha 2009) or planning (e.g., Yang & Wei 2010) using analytical methods.

4 Discussion and Conclusion

In this paper in conclusion, analytical methods for master scheduling were identified through a systematic literature review and then categorized according to their competitive advantage. For this purpose, the model of Davenport and Harris (2007) was used. The categories used are: Statistical Analysis, Extrapolations, Predictive Methods, and Optimization Methods. The categories are sequential and gain competitive advantage for companies as their level of intelligence increases.

Figure 5 displays the findings of the literature review. Along with the categories used to classify methods, we also analyzed the project phase under consideration. As the project advances, creative processes evolve into stable processes, and more information on construction processes and products becomes available and documented in each phase, resulting in an increase in information content. Also, an increase in prediction accuracy can be observed with the availability of more information and consideration of the category of competitive advantage (see Petruseva et al 2012, Dissanayaka & Kumaraswamy 1999, Chen & Huang 2006, Lam & Olalekan 2016). It is important to note that less mature methods are simpler and yield more comprehensible and explainable results. Furthermore, their limited complexity often makes them more straightforward to implement and interpret. However, algorithms created for specific datasets may lack transferability, which can impact prediction accuracy when applied to new datasets, as noted in publications on multivariate regression.

Although computing power and data volume in construction projects are steadily increasing, research in construction scheduling shows significant growth in areas with lower maturity levels, primarily focusing on project realization. In contrast, high-maturity research remains relatively constant, with cumulative totals increasing linearly over time. This disparity may be due to the construction industry's limited expertise in predictive methods and optimization, as well as a lack of explainability that impedes researchers and practitioners from effectively applying results. Additionally, factors such as a lack of pressure for change due to favorable market conditions or a shortage of digitized and structured datasets may contribute to this trend. However, it's important to note that this analysis filters out many recent studies due to a selection criterion requiring more than 100 or 200 citations. While this method eliminates less significant papers, it also excludes newer research employing innovative approaches that haven't yet accumulated

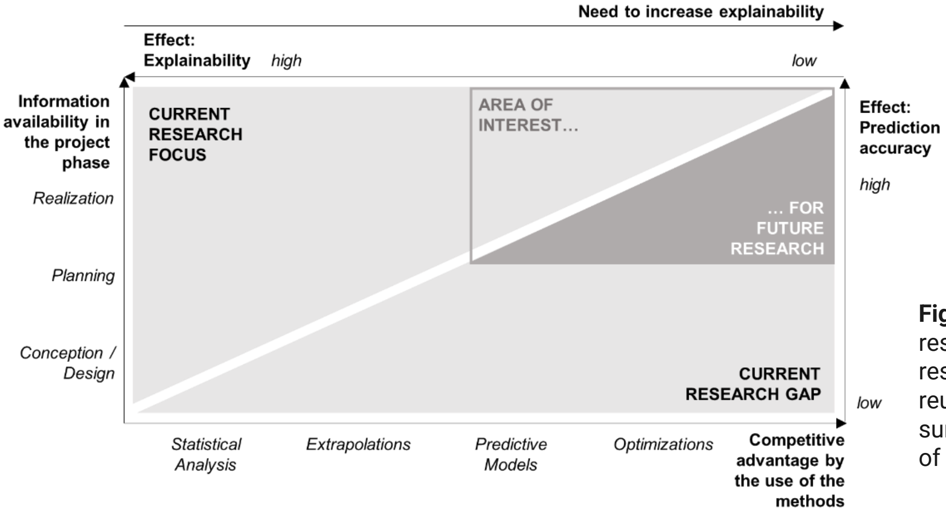

Figure 5. Current research focus and research gap in a reusable matrix summarizing the results of the literature analysis.

many citations. Still, the results remain relevant as they highlight persistent trends and challenges in construction scheduling research, such as the lack of expertise in predictive methods and the need for model explainability. The results indicate a need for further research in this area to strengthen companies' competitive advantage in the long term and facilitate practical implementation.

Recommendations for future research include:

1. Focus on high maturity methods such as predictive methods and optimizations,
2. Explore explainable approaches to master scheduling, such as decision trees,
3. Investigating human-machine collaboration in scheduling processes to support the explainability of the analysis methods, thereby increasing acceptance, and supporting learning of the methods used with expert knowledge,
4. and investigate the transferability of developed algorithms to new data sets to derive actionable project documentation.

References

Albogamy, A., Dawood, N., & Scott, D. (2014). A Risk Management Approach to Address Construction Delays from the Client Aspect. June 2014, pp. 1497–1505.

Bockmayr & Radermacher, F. J. (1993). Künstliche Intelligenz und Operations Research [Artificial intelligence and operations research]. Herzog, O., Christaller, T., and Schütt, D. (eds.) Grundlagen und Anwendungen der Künstlichen Intelligenz [Fundamentals and applications of artificial intelligence]. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 249–254.

Bromilow, FJ. (1969). Contract time performance expectations and the reality. Building forum. Vol. 1. 3, pp. 70–80.

Bzdok, D., Altman, N., & Krzywinski, M. (2018). Statistics versus machine learning. Nature Methods 15.4, pp. 233–234.

Chan, A. P.C. (1999). Modelling building durations in Hong Kong. Construction Management and Economics 17.2, pp. 189–196.

Chan, D. W.M. & Kumaraswamy, M. M. (1997). A comparative study of causes of time overruns in Hong Kong construction projects. International Journal of Project Management 15.1, pp. 55–63.

Chen, W. T., & Huang, Y. H. (2006). Approximately predicting the cost and duration of school reconstruction projects in Taiwan. Construction Management and Economics 24.12, pp. 1231–1239.

Davenport, T. H., & Harris, J. G. (2007). Competing on Analytics - The New Science of Winning. Boston, Massachusetts: Harvard Business Press.

Dissanayaka, S., & Kumaraswamy, M. (1999). Evaluation of factors affecting time and cost performance in Hong Kong building projects. Construction and Architectural Management 6.3, pp. 287–298.

El-Gafy, M. A. (2007). Construction Time-cost Trade-off Analysis Using Ant System. International Journal of Construction Education and Research 3.2, pp. 99–108.

Elinwa, A., & Mangvwat, J. (2001). Time-Overrun Factors in Nigerian Construction Industry. Journal of Construction Engineering and Management 127 (Oct. 2001).

Erdis, E. (2013). The effect of current public procurement law on duration and cost of construction projects in Turkey. Journal of Civil Engineering and Management 19 (Jan. 2013), pp. 121–135.

Flyvbjerg, B., Holm, M. S., & Buhl, S. (2002). Underestimating costs in public works projects: Error or lie? Journal of the American planning association 68.3, pp. 279–295.

Frimpong, Y., Oluwoye, J., & Crawford, L. (2003). Causes of delay and cost overruns in construction of groundwater projects in a developing country, Ghana as a case study. International Journal of Project Management 21.5, pp. 321–326.

Golizadeh, H., et al. (2015). Automated tool for predicting duration of construction activities in tropical countries. KSCE Journal of Civil Engineering 20 (Mar. 2015).

Ireland, V. (1983). The role of managerial actions in the cost, time and quality performance of high-rise commercial building projects. Construction Management and Economics 3 (Jan. 1983), pp. 59–87.

Irfan, M., et al. (2011) Planning-stage estimation of highway project duration on the basis of anticipated project cost, project type, and contract type. International Journal of Project Management, 29(1), pp. 78–92.

Jung, M., et al. (2016) Weather-Delay Simulation Model Based on Vertical Weather Profile for High-Rise Building Construction. Journal of Construction Engineering and Management, 142 (Jan.).

Kaliba, C., Muya, M., & Mumba, K. (2009) Cost escalation and schedule delays in road construction projects in Zambia. International Journal of Project Management, 27(5), pp. 522–531.

Kaming, P. F., et al. (1997) Factors influencing construction time and cost overruns on high-rise projects in Indonesia. Construction Management and Economics, 15(1), pp. 83–94.

Kanoglu, A. (2003) An integrated system for duration estimation in design/build projects and organizations. Engineering, Construction and Architectural Management, 10, pp. 272–282.

Kirkpatrick, S., Jr., & Vecchi, M. (1983) Optimization by Simulated Annealing. Science, 220, pp. 671–680.

Kirste, M., & Schürholz, M. (2019) Einleitung: Entwicklungswege zur KI. Wittpahl, V. (ed.) Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 21–35.

Kumar, V., & Abdullah, W. (2011) Optimization of uncertain construction time-cost trade-off problem using simulated annealing algorithm. pp. 489–494.

Kumaraswamy, M. M., & Chan, D. W. M. (1995) Determinants of construction duration. Construction Management and Economics, 13(3), pp. 209–217.

Lam, K., & Olalekan, O. (2016) Forecasting construction output: A comparison of artificial neural network and Box-Jenkins model. Engineering, Construction & Architectural Management, pp. 302–322.

Liu, M., et al. (2012) Predicting the Preliminary Engineering Duration of Bridge Projects. In: Construction Research Congress 2012, pp. 505–514.

Lu, Y., & AbouRizk, S. M. (2009) Automated Box–Jenkins forecasting modelling. Automation in Construction, 18(5), pp. 547–558.

Migliaccio, G. C., & Shrestha, P. P. (2009) Analysis of Design-Build Procurement Activities Durations for Highway Projects. Building a Sustainable Future.

Naik, G. M., & Kumar, M. (2013) Project Cost and duration optimization Using Soft Computing Techniques. Journal of Advanced Management Science, 1(3).

Nickel, S., Stein, O., & Waldmann, K.-H. (2014) Operations Research. Springer, Berlin.

North, K. (2016) Wissensorientierte Unternehmensführung: Wissensmanagement gestalten. [Knowledge-oriented corporate management: shaping knowledge management.] Springer, Berlin.

Ogunlana, S. O., Promkuntong, K., & Jearkjirm, V. (1996) Construction delays in a fast-growing economy: Comparing Thailand with other economies. International Journal of Project Management, 14(1), pp. 37–45.

Petruseva, S., Zujo, V., & Pancovska, V. Z. (2012) Neural Network Prediction Model for Construction Project Duration. International Journal of Engineering Research and Technology, 1 (Jan.).

Walker, D. H. T. (1995) An investigation into construction time performance. Construction Management and Economics, 13(3), pp. 263–274.

Wang, Y.-R., Yu, C.-Y., & Chan, H.-H. (2012) Predicting construction cost and schedule success using artificial neural networks ensemble and support vector machines classification models. International Journal of Project Management, 30(4), pp. 470–478.

Yang, J.-B., & Wei, P.-R. (2010) Causes of Delay in the Planning and Design Phases for Construction Projects. Journal of Architectural Engineering, 16(2), pp. 80–83.

Zadeh, L. A. (1965) Fuzzy sets. Information and control, 8(3), pp. 338–353.

Zhang, Y., & Ng, S. (2012) An ant colony system-based decision support system for construction time-cost optimization. Journal of Civil Engineering and Management, 18 (Aug.).

Zhang, Y., & Koreisha, S. (2015) Adaptive Order Determination for Constructing Time Series Forecasting Models. Communications in Statistics - Theory and Methods, 44(22), pp. 4826–4847.

Zheng, D., Ng, S., & Kumaraswamy, M. (2004) Applying a Genetic Algorithm-Based Mult objective Approach for Time-Cost Optimization. Journal of Construction Engineering and Management