



Knowledge Extraction and Fuzzy Expert System Modeling for IoT-Enabled Asset Management in Construction and Infrastructure Projects: Insights from a Practitioner Survey

Sameer Jain^{1,*} 

¹ NICMAR Business School (NBS), NICMAR University, Pune, Maharashtra, India

ARTICLE INFO

Article history:

Received 25 August 2026

Received in revised form 1 October 2026

Accepted 9 October 2026

Available online 17 October 2026

Keywords:

IoT; asset management; Fuzzy expert system; Knowledge extraction; Construction and infrastructure; Wang–Mendel rule learning; survey analytics; Projects

ABSTRACT

IoT technologies have the potential to provide real-time visibility of labour, materials, and plant and machinery on construction and infrastructure projects, but the knowledge gained from practical experience is rarely turned into reusable decision-making knowledge. Most existing studies offer only descriptive observations or technical pilot projects and rarely convert survey findings into clear, rule-based models that can support asset management decisions. In this study, knowledge is drawn from a survey of practitioners and a fuzzy expert-system model is developed which links the perceived benefits of IoT for labour, materials, and plant and machinery to the perceived project performance. The questionnaire was completed by 75 construction professionals (96% of them working in India) and included eleven five-point Likert items, which were analysed using the relative importance index, Cronbach's alpha, exploratory factor analysis, non-parametric tests, and clustering. A Wang–Mendel Mamdani rule base and a Takagi–Sugeno–Kang model were trained and compared with linear regression, random forest, and a mean-only baseline using leave-one-out cross-validation. Perceived benefits were high (mean 3.96 of 5; relative importance index 0.79), and 97.3% of respondents would recommend IoT-based solutions. Materials management (mean 4.19) and plant and machinery management (mean 4.17) ranked highest, whereas labour productivity ranked lowest (mean 3.57). Three factors explained 63.1% of variance, and internal consistency across all items was high ($\alpha = 0.89$). Plant and machinery benefits were the dominant predictor of perceived project performance (standardised beta = 0.56, $p < 0.001$). Cross-validated mean absolute error was 0.49 for regression, 0.53 for the Takagi–Sugeno–Kang model, 0.59 for the Mamdani model, and 0.64 for the baseline, although differences between models were not statistically significant. The fuzzy approach yielded 15 interpretable, data-derived rules. Practitioners are strongly positive about IoT-enabled asset management, with plant and machinery the most credible entry point. The rule base provides transparent decision knowledge, but the small, India-centred sample and perception-based outcomes limit generalisation; validation with sensor data from live projects is required.

1. Introduction

* Corresponding author.

E-mail address: author.mail@gmail.com

<https://doi.org/10.59543/s4h40a28>

1.1 Background and Motivation

Construction and infrastructure projects depend on the coordinated use of three classes of physical assets: labour, materials, and plant and machinery. Delays, rework, idle equipment, material losses, and safety incidents typically arise when information about these assets reaches decision-makers late, incompletely, or not at all. The Internet of Things (IoT), in which uniquely identifiable devices sense, communicate, and act through networks, was proposed as a way to close this information gap [1, 2].

In the construction context, IoT is realized by integrating RFID and Bluetooth tags in materials, satellite and telematics equipment in machinery, wearables in workers, and embedded sensors in structures; it is being increasingly integrated with Building Information Modeling (BIM) and Digital Twins [3]. Technology availability does not guarantee its adoption. Research shows perceived usefulness is a primary driver of intention to adopt new information systems [4, 5]. Understanding what practitioners expect IoT to deliver for each asset class is a prerequisite to prioritize digital investments, particularly in emerging construction markets with heterogeneous digital maturity. The knowledge also has a secondary value. When captured in a structured form, it can be embedded in expert and decision-support systems that recommend where IoT deployment is likely to pay off first.

1.2 IoT-Enabled Asset Management in Construction

The reviews of real-time site monitoring describe a wide and developing range of sensors and analytical tools for use in positioning, identifying hazards, and monitoring progress [6]. According to studies on Industry 4.0, the Internet of Things (IoT) is seen as a key enabler of digital and automated construction [7], and research into the integration of BIM with IoT demonstrates how sensor data can be connected to semantic building models via open standards [3, 8]. Approaches based on digital twins take this concept further by keeping a continuously updated virtual model of the project which aids in monitoring and decision-making [9, 10].

For Labour, wearable and positioning systems have been developed to warn workers of proximity to hazards and to monitor their location in real time, as demonstrated by Kanan *et al.*, [11]. For plant and machinery, data-driven predictive-maintenance frameworks that combine BIM, IoT data, and machine learning illustrate how equipment condition can be anticipated rather than merely observed, as shown by Cheng *et al.*, [12]. More broadly, deep-learning methods are being applied to many construction data streams [13]. Most of this literature, however, consists of technical designs and case studies. Fewer studies capture how practitioners themselves perceive the benefits of IoT across labour, materials, and plant and machinery within a single instrument, and fewer still convert those perceptions into an explicit, inspectable model.

1.3 Soft Computing and Expert Systems for Construction Decisions

Practitioner judgements are expressed in linguistic terms such as “agree” or “strongly agree”, which are naturally represented by fuzzy sets [14]. Fuzzy inference systems turn such linguistic variables into IF–THEN rules; the Mamdani controller [15] and the Takagi–Sugeno–Kang (TSK) model [16] are the two most widely used forms, and neuro-fuzzy extensions such as ANFIS tune them from data [17]. Wang and Mendel [18] showed that fuzzy rules can be generated directly from numerical examples, which makes the approach a practical tool for knowledge extraction from survey records. A general overview of fuzzy set theory and its applications is given by Zimmermann [19].

In contrast to black-box ensembles such as random forests [20], rule-based fuzzy models give up a certain amount of flexibility but do produce knowledge that a project manager can read, question, and modify, which is a fundamental requirement of expert systems.

1.4 Research Gap, Objectives and Contributions

Three gaps motivate this study. First, perception evidence for labour, materials, and plant and machinery is rarely analyzed alongside project-level cost and quality outcomes. Second, survey results are usually reported as descriptive statistics and not converted into reusable rule-based knowledge. Third, where predictive models are built from small survey samples, they are rarely validated against simple benchmarks, which makes it difficult to judge whether added model complexity is warranted.

The study therefore pursues five objectives: (i) to quantify practitioners' perceptions of IoT benefits for labour, materials, and plant and machinery, and for project cost and quality; (ii) to assess the reliability and dimensional structure of the instrument; (iii) to test whether perceptions differ by respondent characteristics; (iv) to develop a fuzzy expert-system model and compare it with regression and random-forest benchmarks under leave-one-out cross-validation; and (v) to extract an interpretable rule base and respondent profiles. The contributions are an integrated measurement and modeling workflow for IoT-enabled asset management, an openly specified fuzzy rule base derived from practitioner data, and a candid appraisal of how much predictive value such a model has when the sample is small.

2. Methodology

2.1 Research Design and Instrument

A cross-sectional questionnaire survey was used. The instrument contained nine demographic and professional items (age, country of residence, country of work, gender, employment status, occupation, organization size, education, and years of experience), eleven attitudinal statements, and one binary item asking whether the respondent would recommend IoT-based solutions in construction projects. The eleven statements were rated on a five-point Likert scale (1 = strongly disagree, 5 = strongly agree) and covered three asset classes with three statements each: labour (L1 management, L2 productivity, L3 safety), materials (M1 management, M2 efficiency, M3 waste reduction), and plant and machinery (P1 management, P2 efficiency, P3 downtime reduction). Two further statements addressed project-level outcomes: reduction of overall project cost (C1) and improvement of project quality (Q1). The full wording of each statement is given in Table 2.

2.2 Sample and Data Preparation

Responses were obtained from 75 construction professionals and contained no missing values. Because no sampling frame was available, the sample is treated as a non-probability sample and the results as exploratory. Free-text variants were harmonized before analysis (for example, "INDIA" and "Indi" were coded as India, and "Strongly Agree" and "Strongly agree" were merged). Four composite scores were computed as arithmetic means: labour (LAB), materials (MAT), plant and machinery (PLT), and perceived project performance (PERF, the mean of C1 and Q1). An overall IoT benefit index (IBI) was computed as the mean of all eleven items. Participation was voluntary and anonymous, and no personally identifying information was analyzed.

2.3 Statistical Analysis

The relative importance index (RII) of each item was calculated as in Eq. (1), where i is the scale value, n_i is the number of respondents choosing it, and N is the sample size.

$$RII = \frac{\sum_{i=1}^5 i \cdot n_i}{5 \cdot N} \quad (1)$$

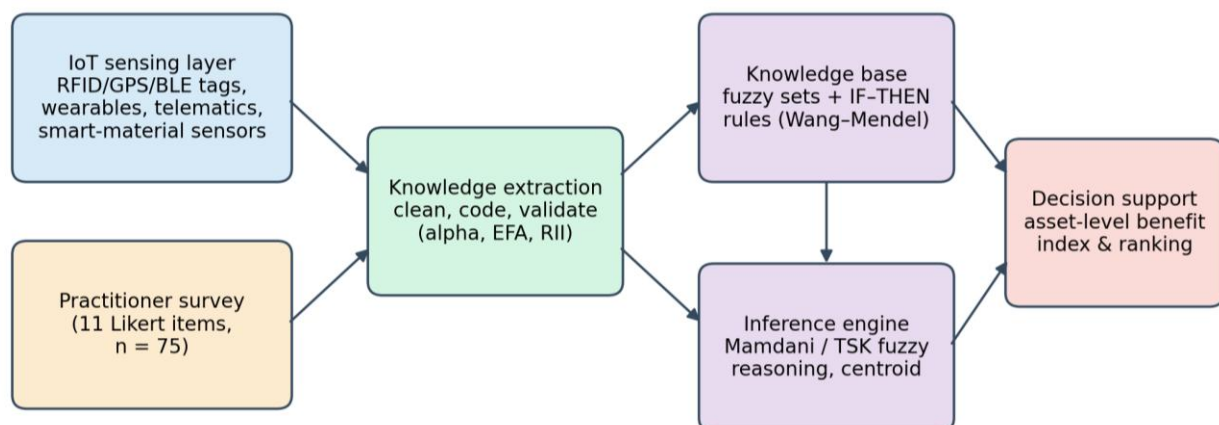
Each item was tested for deviations from the neutral midpoint (3) using a one-sample Wilcoxon signed-rank test [21]. In addition, internal consistency was estimated using Cronbach's alpha coefficient [22], which is expressed in Eq. (2), where k is the number of items, σ_j^2 is the variance of item j , and σ_T^2 is the variance of the total score.

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum_{j=1}^k \sigma_j^2}{\sigma_T^2} \right) \quad (2)$$

The initial factorability of the data was estimated using the Kaiser–Meyer–Olkin index [23] and Bartlett's test of sphericity. The dimensionality of the data was examined using a principal component analysis of the Spearman correlation matrix with varimax rotation. Components with an eigenvalue of greater than 1 were extracted. Differences among the four composite dimensions were tested with the Friedman test followed by Bonferroni-adjusted pairwise Wilcoxon tests. Differences across respondent groups were tested with the Kruskal–Wallis test [24] for three groups and the Mann–Whitney U test for two groups. Respondent profiles were identified with k-means clustering on the eleven items, with the number of clusters guided by the silhouette coefficient. The significance level was 0.05. All computations were performed in Python using NumPy, SciPy, and scikit-learn.

2.4 Fuzzy Expert-System Model

The proposed expert-system workflow is shown in Fig. 1. Survey records pass through a knowledge-extraction stage (cleaning, coding, and validation), after which fuzzy sets and IF–THEN rules are stored in a knowledge base; an inference engine then converts new asset-level scores into a project-performance estimate for decision support. The inputs are LAB, MAT, and PLT, and the output is PERF, all on the 1–5 scale.



Labour (LAB) · Materials (MAT) · Plant & machinery (PLT) → Project performance (PERF)

Fig. 1. Knowledge-extraction and fuzzy expert-system framework for IoT-enabled asset management

For the Mamdani model, each variable was described by three linguistic terms (Low, Medium, High) with the membership functions in Eq. (3a)–(3c) and Fig. 2a. The terms form a partition of unity, so memberships of any score sum to one.

$$\mu_{\text{Low}}(x) = \max\{0, \min\{1, 3.5 - x\}\} \quad (3a)$$

$$\mu_{\text{Med}}(x) = \max\{0, \min\{1, x - 2.5, 4.5 - x\}\} \quad (3b)$$

$$\mu_{\text{High}}(x) = \max\{0, \min\{1, x - 3.5\}\} \quad (3c)$$

Rules were learned with the Wang–Mendel procedure [18]. Each respondent defines one candidate rule by assigning every variable to the term with the highest membership; the rule degree is the product of these maximum memberships. When several respondents generate rules with the same antecedent but different consequents, the rule with the highest degree is kept, and the number of respondents sharing the antecedent is recorded as its support. Of the 27 possible antecedent combinations, those not observed in the data were filled with a monotonic expert default in which the consequent term equals the rounded mean of the antecedent terms (better asset outcomes imply better project performance). Finally, the Mamdani algorithm was used with min–max rules. The firing strength of rule r is defined in Eq. (4). The aggregated output set is defined in Eq. (5). The crisp output is the centroid of the aggregated output set, as defined in Eq. (6).

$$w_r = \min\{\mu_{Ar1}(LAB), \mu_{Ar2}(MAT), \mu_{Ar3}(PLT)\} \quad (4)$$

$$\mu_{\text{agg}}(y) = \max_r \min\{w_r, \mu_{Br}(y)\} \quad (5)$$

$$y^* = \frac{\int_1^5 y \cdot \mu_{\text{agg}}(y) dy}{\int_1^5 \mu_{\text{agg}}(y) dy} \quad (6)$$

Due to the coarse nature of the three-term Mamdani partition, a first-order Takagi–Sugeno–Kang (TSK) model was also implemented. This model employs two terms (Low, High; Fig. 2b), resulting in eight rules with linear consequents estimated by ridge-penalized least squares (penalty $\lambda = 1$, predetermined). The output is calculated as the weighted average in Eq. (7), where the normalized firing strengths are $\bar{w}_r = w_r / \sum w_r$.

$$\widehat{PERF} = \sum_{r=1}^8 \bar{w}_r (a_{r0} + a_{r1}LAB + a_{r2}MAT + a_{r3}PLT) \quad (7)$$

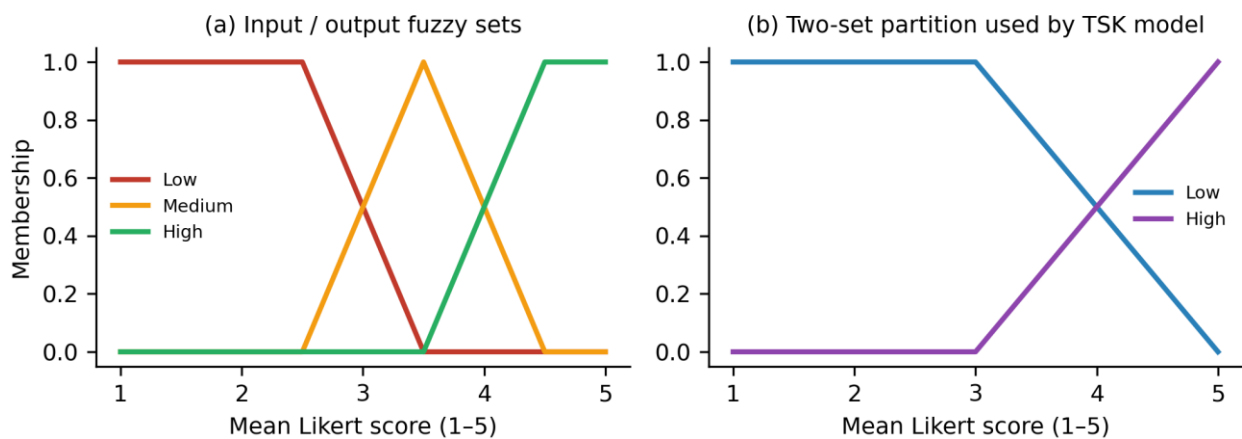


Fig. 2. Membership functions: (a) three-term partition used by the Mamdani model; (b) two-term partition used by the TSK model

The positioning of the Mamdani terms corresponds to the concentration of scores between 3 and 5. An initial equally spaced partition, with term peaks at 1, 3, and 5, resulted in predictions converging

toward the neutral value during a pilot analysis on the same dataset. Consequently, the partition was adjusted. Since this adjustment was based on the same sample, the cross-validated accuracy of the fuzzy models reported below may be somewhat optimistic.

2.5 Benchmark Models and Validation

Three benchmarks were used: a mean-only baseline, ordinary least squares regression (Eq. 8), and a random forest with 200 trees and a minimum leaf size of three [20]. All models predicted PERF from LAB, MAT, and PLT.

$$PERF = \beta_0 + \beta_1 LAB + \beta_2 MAT + \beta_3 PLT + \varepsilon \quad (8)$$

Given the sample size of 75 observations, a single hold-out split would be unreliable. Therefore, leave-one-out cross-validation (LOOCV) was employed, with each step of rule learning or model fitting repeated within each fold. The performance of the model was summarized using the mean absolute error (MAE), root mean squared error (RMSE), Spearman's ρ between observed and predicted values, and the cross-validated coefficient of determination in Eq. (9) where $\hat{y}_i$ indicates a prediction for respondent i from a model estimated without that respondent.

$$R_{cv}^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (9)$$

The paired differences in absolute error between models were analyzed using Wilcoxon signed-rank tests, and a 5,000-sample bootstrap was used to estimate a confidence interval for the reduction in error relative to a baseline. A sensitivity analysis was performed after excluding two respondents who exhibited contradictory response patterns: in particular, both strongly agreed with item L1 but strongly disagreed with the remaining ten statements while indicating that they would recommend IoT.

3. Results

3.1 Respondent Profile

Table 1 provides an overview of the 75 respondents. Most respondents were young, male, highly-educated individuals based in India: specifically, 70.7% were aged 25–34 years, 86.7% were male, 96.0% were working in India (with the rest three in United Arab Emirates and Qatar), and 86.7% had a master's degree. Most respondents were employed full-time (92.0%) and had more than 500 employees as their workplace (68.0%). Their work experience was limited (69.3% had less than five years), and most respondents belonged to "Other" category (56.0%) in terms of their occupation, with remainder distributed between project managers (21.3%) and construction managers (17.3%).

Table 1

Profile of survey respondents (N = 75)

Characteristic	n	%
Age (years)		
18–24	11	14.7
25–34	53	70.7
35–44	9	12.0
45–54	2	2.7
Gender		
Male	65	86.7
Female	10	13.3
Country of work		
India	72	96.0
United Arab Emirates / Qatar	3	4.0
Occupation		
Other	42	56.0
Project manager	16	21.3
Construction manager	13	17.3
Contractor	3	4.0
Site engineer	1	1.3
Organization size		
Large (>500 employees)	51	68.0
Medium (50–500 employees)	19	25.3
Small (<50 employees)	5	6.7
Highest education		
Master's degree	65	86.7
Bachelor's degree	6	8.0
High school or equivalent	2	2.7
Other	2	2.7
Experience		
Less than 1 year	8	10.7
1–5 years	44	58.7
6–10 years	9	12.0
11–15 years	9	12.0
16–20 years	3	4.0
More than 20 years	2	2.7
Employment status		
Employed full-time	69	92.0
Self-employed	4	5.3
Employed part-time	1	1.3
Student	1	1.3

3.2 Perceived Benefits at Item Level

Table 2 and Fig. 3 show the item level results. All eleven statements were scored significantly above the neutral midpoint (Wilcoxon signed-rank tests, all $p < 0.001$), and the distributions were negatively skewed (skewness -0.61 to -1.57), suggesting a ceiling effect. Improved materials management (M1) had the highest mean score (4.19, RII = 0.837; 88.0% agreement) followed by improved plant and machinery management (P1: 4.17, RII = 0.835), and plant and machinery efficiency (P2: 4.11, RII = 0.821). Labour management (L1) and material-management efficiency (M2) tied for fourth place (4.09, RII = 0.819). Labour safety (L3) ranked sixth (4.00), ahead of plant downtime reduction (P3: 3.92), project quality (Q1: 3.85), and material waste reduction (M3: 3.77). The two lowest-ranked items were overall cost reduction (C1: 3.73) and labour productivity (L2: 3.57),

the latter attracting the largest dispersion ($SD = 1.10$), the lowest agreement (56.0%), and the highest neutral share (30.7%).

Table 2

Item-level descriptive statistics and ranking ($N = 75$)

Code	Statement	Mean	SD	RII	Rank	Agree (%)	Neutral (%)	Disagree (%)
L1	IoT-based solutions can improve the management of labour	4.09	0.81	0.819	4	86.7	8.0	5.3
L2	... can increase the productivity of labour	3.57	1.10	0.715	11	56.0	30.7	13.3
L3	... can improve the safety of labour	4.00	0.87	0.800	6	78.7	17.3	4.0
M1	... can improve the management of materials	4.19	0.88	0.837	1	88.0	6.7	5.3
M2	... can increase the efficiency of material management	4.09	0.82	0.819	4	86.7	9.3	4.0
M3	... can reduce the waste of materials	3.77	0.89	0.755	9	72.0	21.3	6.7
P1	... can improve the management of plant and machinery	4.17	0.83	0.835	2	86.7	10.7	2.7
P2	... can increase the efficiency of plant and machinery management	4.11	0.89	0.821	3	84.0	10.7	5.3
P3	... can reduce the downtime of plant and machinery	3.92	1.00	0.784	7	77.3	13.3	9.3
C1	... can reduce the overall cost of construction projects	3.73	0.99	0.747	10	68.0	21.3	10.7
Q1	... can improve the quality of construction projects	3.85	0.90	0.771	8	70.7	24.0	5.3

Note: Agree = "Agree" plus "Strongly agree"; Disagree = "Disagree" plus "Strongly disagree". All items differ from the neutral value 3 at $p < 0.001$ (Wilcoxon signed-rank test).

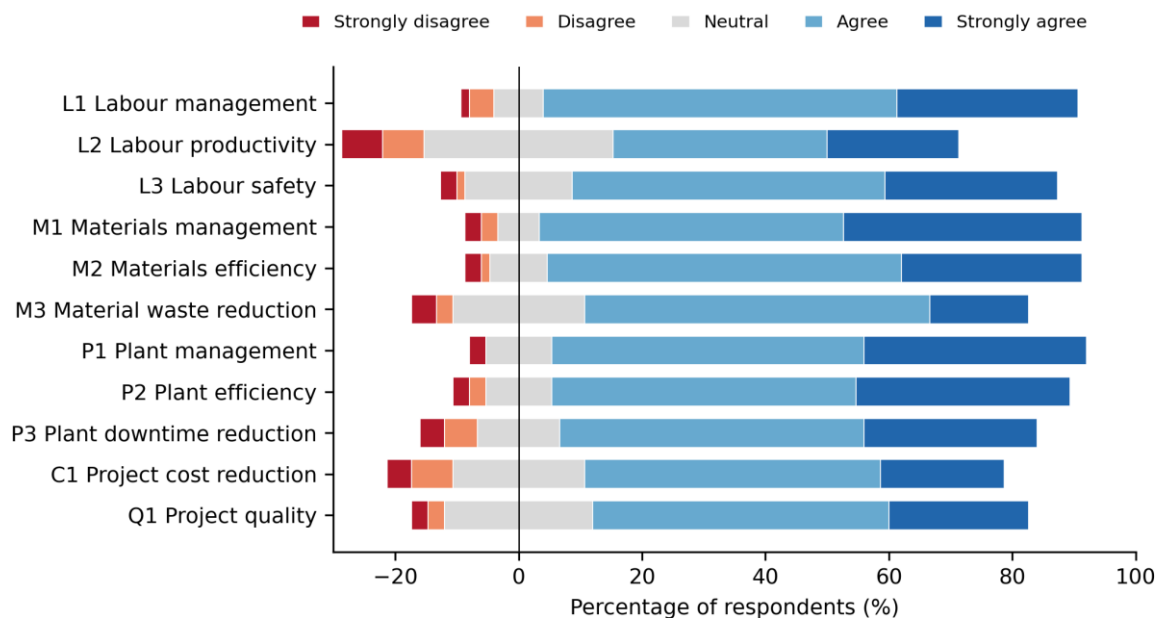


Fig. 3. Distribution of responses to the eleven Likert statements (diverging stacked bars; percentages of respondents)

3.3 Reliability, Dimensions and Factor Structure

Table 3 reports the composite dimensions. Internal consistency was acceptable to good for materials ($\alpha = 0.78$) and plant and machinery ($\alpha = 0.86$), lower for labour ($\alpha = 0.66$, below the conventional 0.70 threshold), and high for all eleven items together ($\alpha = 0.89$). The two outcome items (C1 and Q1) correlated moderately (Spearman $\rho = 0.53$), so no alpha was computed for PERF. Mean scores differed across the four dimensions (Friedman $\chi^2(3) = 12.05$, $p = 0.007$), but the effect was small (Kendall's $W = 0.054$); after Bonferroni correction, only plant and machinery exceeded perceived project performance ($p = 0.002$), while materials versus performance ($p = 0.073$) and labour versus plant and machinery ($p = 0.123$) were not significant.

Table 3

Composite dimensions: reliability and descriptive statistics (N = 75)

Dimension	Items	α	Mean	SD	RII	95% CI of mean
Labour (LAB)	L1–L3	0.66	3.89	0.72	0.778	3.72–4.05
Materials (MAT)	M1–M3	0.78	4.02	0.72	0.804	3.85–4.18
Plant & machinery (PLT)	P1–P3	0.86	4.07	0.80	0.813	3.88–4.25
Project performance (PERF)	C1, Q1	$\rho = 0.53$	3.79	0.84	0.759	3.60–3.99
Overall index (IBI)	All 11	0.89	3.96	0.62	0.791	3.81–4.10

Note: α = Cronbach's alpha (Spearman ρ for the two-item PERF). CI = t-based confidence interval.

Spearman correlations among the dimensions were all positive: LAB–MAT = 0.34, LAB–PLT = 0.47, LAB–PERF = 0.32, MAT–PLT = 0.36, MAT–PERF = 0.37, and PLT–PERF = 0.55. At item level, correlations ranged from -0.02 to 0.75 (mean 0.32 ; Fig. 4), with the strongest block among the plant and machinery items. The data were suitable for factor analysis (KMO = 0.750 ; Bartlett's $\chi^2(55) = 313.9$, $p < 0.001$). Three components had eigenvalues above one (4.27 , 1.45 , and 1.22), jointly explaining 63.1% of the variance. After varimax rotation (Table 4), the components separate into a plant-and-outcome factor (25.8% of variance), a materials factor (20.2%), and a labour factor (17.0%). Cost reduction (C1) and quality (Q1) load with the plant and machinery items, and Q1 shows a secondary loading (0.43) on materials. The a-priori labour, materials, and plant structure is therefore largely recovered, with the project-outcome items grouping with plant and machinery. Given the sample of 75 and a ratio of under seven respondents per item, this structure is tentative.

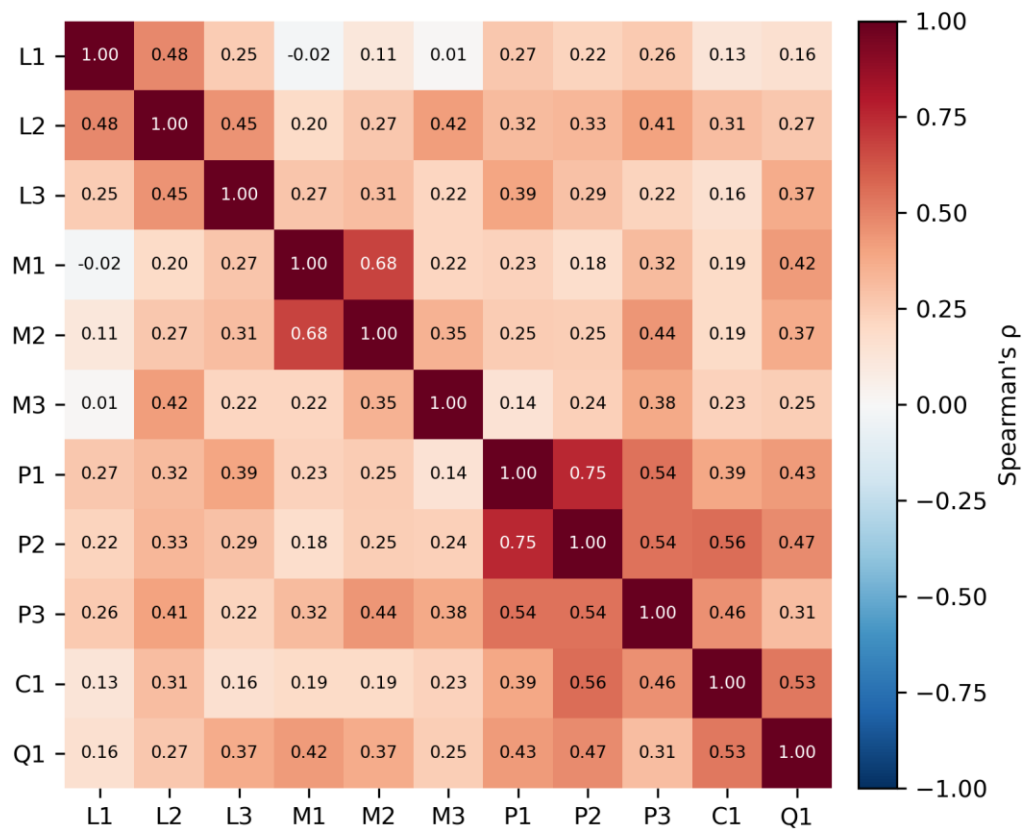


Fig. 4. Spearman correlation matrix of the eleven items

Table 4

Varimax-rotated component loadings (loadings ≥ 0.50 in bold)

Item	F1: Plant & outcomes	F2: Materials	F3: Labour
P1	0.79	0.09	0.25
P2	0.88	0.07	0.17
P3	0.60	0.34	0.28
C1	0.77	0.13	0.03
Q1	0.59	0.43	0.08
M1	0.13	0.86	-0.04
M2	0.14	0.85	0.14
M3	0.13	0.49	0.30
L1	0.15	-0.15	0.79
L2	0.21	0.23	0.80
L3	0.18	0.34	0.56
Variance explained (%)	25.8	20.2	17.0

Note: Principal components of the Spearman correlation matrix; KMO = 0.750; cumulative variance = 63.1%.

3.4 Differences by Respondent Characteristics

Table 5 summarizes 30 comparisons (six respondent characteristics by five scores). None reached the 0.05 level. The smallest p-values concerned the plant and machinery score by experience group ($p = 0.071$), education ($p = 0.091$), and country of work ($p = 0.130$). Mean IBI was 3.94 for respondents with up to five years' experience ($n = 52$), 4.06 for those with 6–15 years ($n = 18$), and 3.71 for those with more than 15 years ($n = 5$; $SD = 1.38$, inflated by one contradictory respondent). By organization size, mean IBI was 4.09 (small, $n = 5$), 3.82 (medium, $n = 19$), and 3.99 (large, $n = 51$). Perceptions

were thus broadly homogeneous, although the small sizes of several groups mean that the absence of significant differences is not evidence of equivalence.

Table 5

Group comparisons of the overall index (IBI) and smallest p-value across the four dimensions

Characteristic	Groups (n)	Test	Statistic (IBI)	p (IBI)	Min. p (LAB, MAT, PLT, PERF)
Experience	≤5 y (52); 6–15 y (18); >15 y (5)	Kruskal–Wallis H	1.12	0.572	0.071
Role	Construction mgr (13); Project mgr (16); Other (46)	Kruskal–Wallis H	0.53	0.767	0.537
Organization size	Small (5); Medium (19); Large (51)	Kruskal–Wallis H	1.28	0.527	0.140
Education	Master's (65); Other (10)	Mann–Whitney U	386.0	0.345	0.091
Gender	Male (65); Female (10)	Mann–Whitney U	288.5	0.574	0.351
Country	India (72); UAE/Qatar (3)	Mann–Whitney U	57.0	0.171	0.130

Note: Experience was grouped as up to 5, 6–15, and over 15 years; “Other” occupation includes contractors and a site engineer.

3.5 Predictive Modeling of Perceived Project Performance

Regression results are in Table 6. The model explained 53.9% of the variance in PERF (adjusted $R^2 = 0.519$; $F(3, 71) = 27.6$, $p < 0.001$). Plant and machinery was the dominant predictor ($b = 0.589$, standardized $\beta = 0.561$, $p < 0.001$), materials had a smaller but significant effect ($b = 0.261$, $\beta = 0.224$, $p = 0.034$), and labour contributed nothing distinguishable from zero ($b = 0.030$, $\beta = 0.025$, $p = 0.792$). Variance inflation factors were below 1.9, so multicollinearity is not a concern. The random-forest importance scores gave the same ordering (PLT 0.63, MAT 0.21, LAB 0.16).

Table 6

Ordinary least squares regression of perceived project performance (PERF)

Predictor	b	SE	t	p	Std. β	VIF
Intercept	0.234	0.448	0.52	0.604	—	—
Labour (LAB)	0.030	0.112	0.26	0.792	0.025	1.42
Materials (MAT)	0.261	0.120	2.17	0.034	0.224	1.65
Plant & machinery (PLT)	0.589	0.114	5.15	< 0.001	0.561	1.82

Note: $R^2 = 0.539$; adjusted $R^2 = 0.519$; $F(3, 71) = 27.6$, $p < 0.001$; $N = 75$.

Table 7 and Fig. 5 compare the models under LOOCV. Regression achieved the lowest errors (MAE = 0.49, RMSE = 0.60, $R^2_{cv} = 0.48$), followed by the TSK model (MAE = 0.53, $R^2_{cv} = 0.38$), the random forest (MAE = 0.58), and the Mamdani model (MAE = 0.59, $R^2_{cv} = 0.25$). All four were below the baseline error (MAE = 0.64), whose R^2_{cv} is slightly negative, as expected for a mean-only predictor. In-sample, the TSK model reached $R^2 = 0.58$, which falls to 0.38 under cross-validation and indicates mild overfitting. Paired tests on absolute errors found no statistically significant differences: regression versus baseline $p = 0.064$, TSK versus baseline $p = 0.256$, random forest versus baseline $p = 0.326$, Mamdani versus baseline $p = 0.583$, Mamdani versus regression $p = 0.063$, and TSK versus regression $p = 0.083$. The bootstrap interval for the TSK error reduction relative to the baseline was -0.01 to 0.25 (mean 0.11). The evidence therefore favours the parsimonious linear model, while showing that the fuzzy models are competitive and, unlike the linear model, expose their reasoning as rules.

Table 7

Leave-one-out cross-validated performance for predicting PERF (N = 75)

Model	MAE	RMSE	R^2_{cv}	Spearman ρ
Mean-only baseline	0.638	0.848	-0.027	—
OLS regression	0.492	0.601	0.484	0.503
Random forest	0.578	0.706	0.288	0.412
TSK fuzzy model (8 rules)	0.527	0.658	0.383	0.411
Mamdani fuzzy model (Wang–Mendel)	0.591	0.727	0.245	0.414

Note: Lower MAE and RMSE are better. Spearman ρ is undefined for the baseline because its predictions are nearly constant.

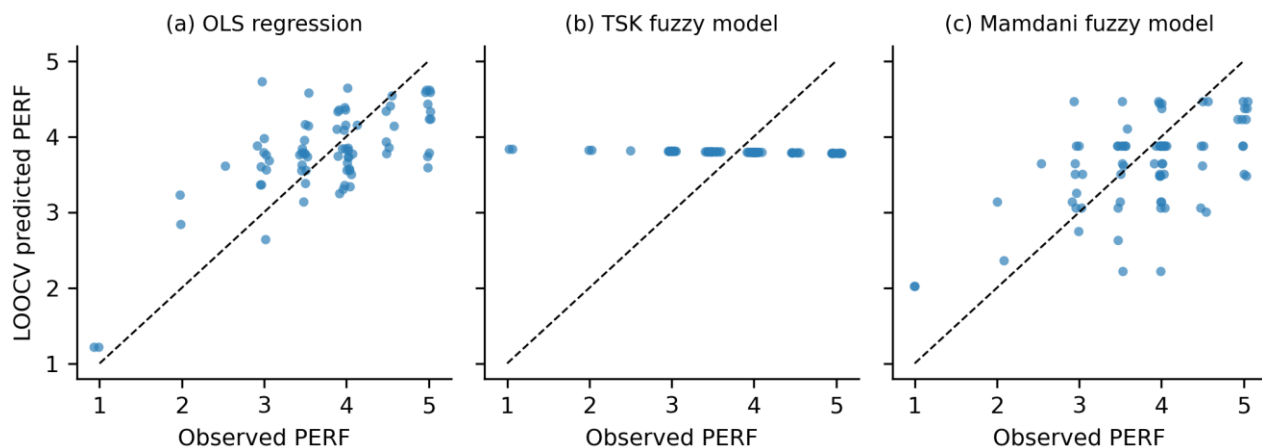


Fig. 5. Leave-one-out predicted versus observed PERF: (a) OLS regression; (b) TSK fuzzy model; (c) Mamdani fuzzy model (small jitter added to observed values; dashed line = perfect prediction)

3.6 Extracted Rule Base and Decision Scenarios

Table 8 lists the 15 rules obtained from the data by the Wang–Mendel procedure; the other 12 of the 27 antecedent combinations were never observed and were completed with the monotonic default. The dominant rule (High, High, High $\rightarrow$ High) is supported by 29 respondents (38.7%). Taken together, the rules show that plant and machinery acts as the gating variable: when PLT is High and MAT is High, PERF is High (rules 1 and 2, 42 respondents), whereas a High PLT combined with only medium materials yield medium performance (rules 4 and 5). If PLT is Medium then output remains Medium despite High labour or materials (rules 3 and 6). Rules involving Low plant or Low overall scores also lead to Low performance. Several rules were supported by only one respondent and so may be regarded as tentative.

Table 8

Data-derived Mamdani rule base (Wang–Mendel): IF LAB is ... AND MAT is ... AND PLT is ... THEN PERF is ...

Rule	LAB	MAT	PLT	PERF	Support (n)	Support (%)
1	High	High	High	High	29	38.7
2	Medium	High	High	High	13	17.3
3	Medium	High	Medium	Medium	6	8.0
4	High	Medium	High	Medium	6	8.0
5	Medium	Medium	High	Medium	5	6.7
6	High	High	Medium	Medium	4	5.3
7	Low	High	Medium	Low	3	4.0
8	Low	Low	Low	Low	2	2.7
9	Low	Low	Medium	Low	1	1.3
10	Low	Medium	Medium	Medium	1	1.3
11	Low	High	High	High	1	1.3
12	Medium	Medium	Low	Low	1	1.3
13	Medium	Medium	Medium	Medium	1	1.3
14	High	Medium	Medium	Medium	1	1.3
15	High	High	Low	Low	1	1.3

Note: Rules sorted by support. The remaining 12 of 27 antecedent combinations were filled by the monotonic expert default and are not shown.

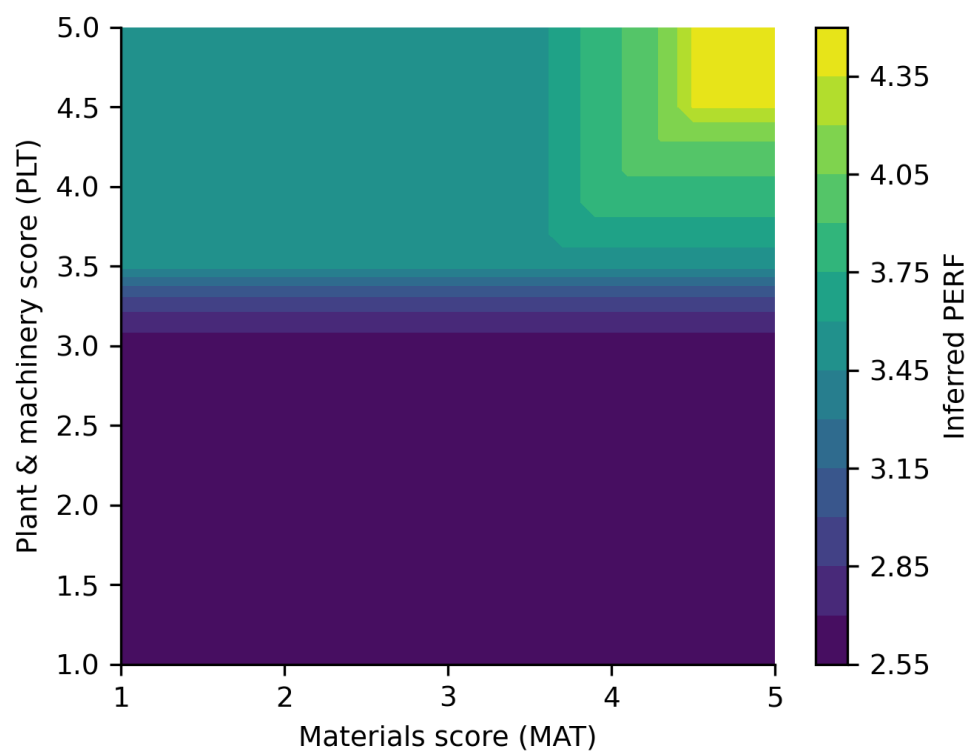


Fig. 6. Mamdani response surface of inferred PERF as a function of MAT and PLT (LAB fixed at 4)

Figure 6 shows the inference surface with LAB fixed at 4, and it can be seen that inferred performance increases mainly along the plant and machinery axis. The steps are typical of the course three term rule base. Table 9 compares the responses of the two fuzzy models to hypothetical asset level scores. If all three scores are 4 then both models infer a performance of approximately 3.8; if all are 5 the inferred value is about 4.4. Leaving PLT alone at 4 and reducing it to 2 decreases the Mamdani output from 3.87 to 2.63 but only 3.01 for the TSK model, which has less stepwise response

to changes in labour and materials. Scenarios featuring low scores lie in a sparsely populated region of the data since few respondents gave assets below 3, and so these outputs are extrapolations and should be treated with caution.

Table 9

Illustrative inference results for hypothetical asset-level scores

LAB	MAT	PLT	Mamdani PERF	TSK PERF
3	3	3	2.63	2.26
4	4	4	3.87	3.80
5	5	5	4.46	4.39
4	4	2	2.63	3.01
4	4	5	3.87	4.26

Note: Scores on the 1–5 scale. Outputs are model estimates, not observed outcomes.

3.7 Respondent Profiles

The silhouette coefficient was highest (0.59) at two clusters, predominantly differentiating three sceptical respondents from the remaining 72. A three-cluster solution (silhouette = 0.19) was also considered (Fig. 7). This revealed a large group of cautious optimists ($n = 43$; mean IBI = 3.78), a group of enthusiasts ($n = 29$; mean IBI = 4.45) unanimously willing to recommend IoT, and a very small group of sceptics ($n = 3$; mean IBI = 1.79). The low silhouette for three clusters implies that the difference between enthusiasts and cautious optimists is more of degree than kind. The cautious group is mainly distinguished by less agreement on labour productivity ($L2 = 3.14$) and cost reduction ($C1 = 3.53$). Remarkably, two out of three sceptics were willing to recommend IoT, suggesting the recommendation item captures a general attitude rather than agreement with the statements. Only two respondents (2.7%) answered “No” to the recommendation item, which is too small a sample for meaningful classification.

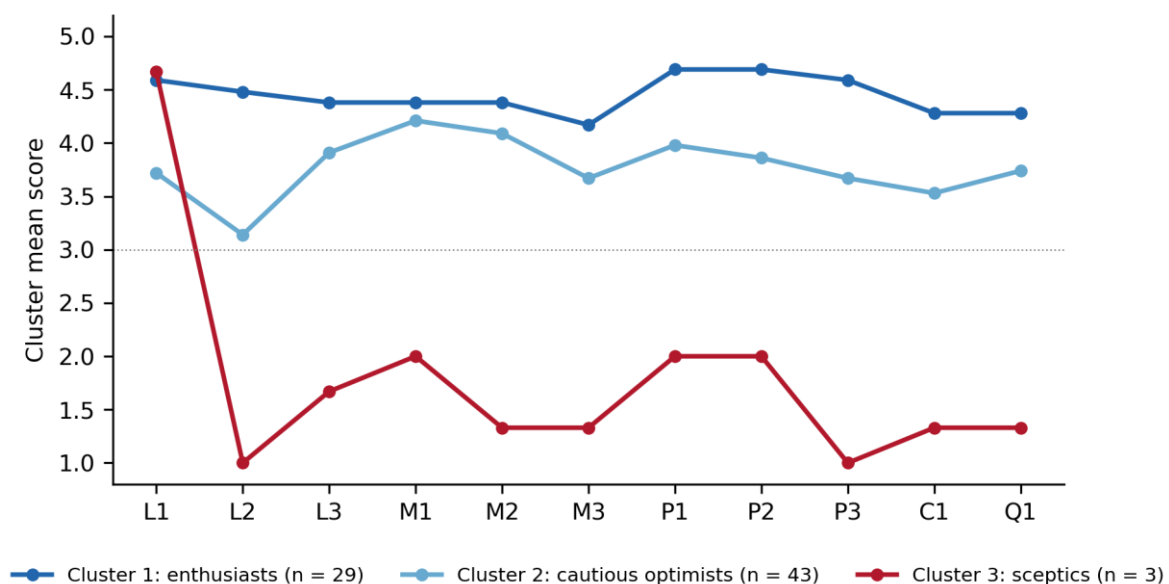


Fig. 7. Cluster mean scores on the eleven items for the three respondent profiles

3.8 Sensitivity Analysis

Two respondents (numbers 6 and 61 in the data file) strongly agreed with L1 but strongly disagreed with all other ten statements while recommending IoT, an inconsistency that could reflect inattentive responding. They were retained in the main analyses as there was no independent reason to exclude them. Excluding them ($n = 73$) raised the mean IBI to 4.03 ($SD = 0.46$), lowered α for all items from 0.89 to 0.80, and reduced the regression R^2 from 0.54 to 0.34; plant and machinery nevertheless remained the dominant predictor ($b = 0.564$, compared with 0.221 for materials and 0.034 for labour). LOOCV MAE was 0.56 for the baseline, 0.51 for regression, and 0.55 for the TSK model, so the advantage of the models over the baseline is modest. The two respondents thus inflate the apparent internal consistency and explanatory power, but they do not alter the main conclusion about the role of plant and machinery.

4. Discussion

4.1 Interpretation of Findings

Practitioners in this sample expect IoT to benefit all three asset classes, with mean scores close to 4 on a five-point scale and 97.3% willing to recommend adoption. This is consistent with the technology-acceptance view that perceived usefulness is the main driver of favorable attitudes [4, 5]. The ranking is informative: materials management and plant and machinery management received the strongest endorsement, which matches the maturity of tracking and telematics solutions and the growing use of IoT data for predictive maintenance, as discussed in [6, 12]. Labour management and safety were also well rated, in line with the evidence on wearable safety monitoring [11], whereas labour productivity was rated lowest and showed the largest disagreement. A plausible explanation, which this study did not test, is that productivity monitoring of individual workers raises concerns about privacy, trust, and workforce acceptance. This links asset-management analytics with the information-security and trust agenda that expert systems must address.

The factor analysis suggests respondents associated cost and quality improvements most closely with equipment-related visibility since C1 and Q1 loaded on the plant and machinery factor. Plant and machinery were the main predictor of perceived performance according to both regression analysis and random forest. This finding should be read as a statement about perceptions: respondents who see strong equipment benefits also expect strong project-level benefits, which may partly reflect a shared tendency to respond favorably (common-method variance) rather than causal effects on outcomes.

On modeling, a simple linear model was hard to beat. With 75 observations, a ceiling effect, and three predictors, additional flexibility offered no accuracy gain, which agrees with the general expectation that complex learners need more data. The fuzzy models were nevertheless competitive and provide a different benefit: they convert the data into explicit IF–THEN knowledge that a project manager can inspect, such as the gating role of plant and machinery in Table 8. For an expert system, this transparency is part of the value, but it should not be mistaken for superior accuracy.

4.2 Implications for Practice and Expert-System Deployment

For practitioners planning IoT investments, the results suggest starting with plant and machinery and materials tracking, where perceived benefits and agreement are highest, and introducing worker-level analytics later and with explicit governance on data use. The extracted rule base and

the TSK model can be embedded in a decision-support tool that converts site-level readiness or benefit scores for labour, materials, and plant into a performance outlook and highlights which asset class limits that outlook. In a deployed system, the survey-derived rules would act as an initial knowledge base to be updated with sensor-derived indicators, such as equipment utilization, downtime, and material-loss rates, as projects generate data.

4.3 Limitations

Several limitations apply. The sample is small, non-probabilistic, and strongly concentrated in India (96%), among young (71% aged 25–34), highly qualified (87% with a master's degree), and relatively inexperienced (69% with up to five years) respondents, and more than half chose the occupation “Other”. The results may therefore not generalize to other regions, senior managers, or site-based trades. Outcomes (cost and quality) are perceptions measured with single items, not observed project performance. Likert data were averaged into composites and analyzed with parametric regression, although non-parametric tests were used for inference on single items. The factor analysis used fewer than seven respondents per item. Subgroup sizes were small, limiting the power of the group comparisons, and the fuzzy-set partition was tuned on the same data, so the cross-validated estimates for the fuzzy models may be somewhat optimistic. Finally, two respondents gave contradictory answers, and while sensitivity analysis shows that the main conclusion is robust, the reliability and explanatory power depend on them to a noticeable degree.

5. Conclusions

This study analyzed a practitioner survey on IoT-enabled asset management and converted it into a fuzzy expert-system model. In response to the five objectives, the following conclusions are drawn. First, perceived benefits are high across labour, materials, and plant and machinery (mean IBI = 3.96; all items above neutral at $p < 0.001$), led by materials management (4.19) and plant and machinery management (4.17), with labour productivity (3.57) the weakest. Second, the instrument is reliable overall ($\alpha = 0.89$) and has a three-factor structure (63.1% of variance) that largely mirrors the labour, materials, and plant and machinery design, although the labour scale is only moderately consistent ($\alpha = 0.66$). Third, perceptions did not differ significantly by experience, role, organization size, education, gender, or country, within the power limits of this sample. Fourth, plant and machinery benefits are the dominant predictor of perceived project performance ($\beta = 0.56$); under leave-one-out validation, linear regression (MAE = 0.49) was most accurate, the TSK (0.53) and Mamdani (0.59) models were competitive but not significantly different, and all models improved only modestly on the mean-only baseline (0.64). The Wang–Mendel procedure yielded 15 interpretable rules, with plant and machinery acting as the gating variable. Clustering analysis revealed a large majority of optimistic respondents alongside a very small group of sceptics.

Further work should include replicating the survey with larger, multi-country samples incorporating senior managers and site roles, linking perceptions to objective performance indicators from live IoT implementations, and evaluating type-2 fuzzy and ANFIS-style learning with optimized membership functions on larger datasets. Open-ended responses could be mined using language models to extend the knowledge base while addressing privacy and trust concerns around worker-level monitoring warrants further attention.

Author Contributions

Conceptualization, S.J.; methodology, S.J.; software, S.J.; validation, S.J.; formal analysis, S.J.; investigation, S.J.; resources, S.J.; data curation, S.J.; writing—original draft preparation, S.J.; writing—review and editing, S.J.; visualization, S.J.; supervision, S.J.; project administration, S.J. The author read and approved the final version of this manuscript.

Funding

This research received no external funding.

Data Availability Statement

The anonymized questionnaire data used for this research (75 responses) are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced this study.

Acknowledgement

The authors thank construction professionals who participated in the study by filling out a questionnaire. This research was unfunded.

References

- [1] Atzori, L., Iera, A., & Morabito, G. (2010). The internet of things: A survey. *Computer Networks*, 54(15), 2787–2805. <https://doi.org/10.1016/j.comnet.2010.05.010>
- [2] Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generations Computer Systems*, 29(7), 1645–1660. <https://doi.org/10.1016/j.future.2013.01.010>
- [3] Tang, S., Shelden, D. R., Eastman, C. M., Pishdad-Bozorgi, P., & Gao, X. (2019). A review of building information modeling (BIM) and the internet of things (IoT) devices integration: Present status and future trends. *Automation in Construction*, 101, 127–139. <https://doi.org/10.1016/j.autcon.2019.01.020>
- [4] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- [5] Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward A unified View1. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>
- [6] Rao, A. S., Radanovic, M., Liu, Y., Hu, S., Fang, Y., Khoshelham, K., Palaniswami, M., & Ngo, T. (2022). Real-time monitoring of construction sites: Sensors, methods, and applications. *Automation in Construction*, 136(104099), 104099. <https://doi.org/10.1016/j.autcon.2021.104099>
- [7] Oesterreich, T. D., & Teuteberg, F. (2016). Understanding the implications of digitisation and automation in the context of Industry 4.0: A triangulation approach and elements of a research agenda for the construction industry. *Computers in Industry*, 83, 121–139. <https://doi.org/10.1016/j.compind.2016.09.006>
- [8] Dave, B., Buda, A., Nurminen, A., & Främling, K. (2018). A framework for integrating BIM and IoT through open standards. *Automation in Construction*, 95, 35–45. <https://doi.org/10.1016/j.autcon.2018.07.022>
- [9] Sacks, R., Brilakis, I., Pikas, E., Xie, H. S., & Girolami, M. (2020). Construction with digital twin information systems. *Data-Centric Engineering*, 1(e14). <https://doi.org/10.1017/dce.2020.16>
- [10] Boje, C., Guerriero, A., Kubicki, S., & Rezgui, Y. (2020). Towards a semantic Construction Digital Twin: Directions for future research. *Automation in Construction*, 114(103179), 103179. <https://doi.org/10.1016/j.autcon.2020.103179>
- [11] Kanan, R., Elhassan, O., & Bensalem, R. (2018). An IoT-based autonomous system for workers' safety in construction sites with real-time alarming, monitoring, and positioning strategies. *Automation in Construction*, 88, 73–86. <https://doi.org/10.1016/j.autcon.2017.12.033>

- [12] Cheng, J. C. P., Chen, W., Chen, K., & Wang, Q. (2020). Data-driven predictive maintenance planning framework for MEP components based on BIM and IoT using machine learning algorithms. *Automation in Construction*, 112(103087), 103087. <https://doi.org/10.1016/j.autcon.2020.103087>
- [13] Akinosho, T. D., Oyedele, L. O., Bilal, M., Ajayi, A. O., Delgado, M. D., Akinade, O. O., & Ahmed, A. A. (2020). Deep learning in the construction industry: A review of present status and future innovations. *Journal of Building Engineering*, 32(101827), 101827. <https://doi.org/10.1016/j.jobbe.2020.101827>
- [14] Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353. [https://doi.org/10.1016/s0019-9958\(65\)90241-x](https://doi.org/10.1016/s0019-9958(65)90241-x)
- [15] Mamdani, E. H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies*, 7(1), 1–13. [https://doi.org/10.1016/s0020-7373\(75\)80002-2](https://doi.org/10.1016/s0020-7373(75)80002-2)
- [16] Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-15(1), 116–132. <https://doi.org/10.1109/tsmc.1985.6313399>
- [17] Jang, J.-S. R. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3), 665–685. <https://doi.org/10.1109/21.256541>
- [18] Wang, L.-X., & Mendel, J. M. (1992). Generating fuzzy rules by learning from examples. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(6), 1414–1427. <https://doi.org/10.1109/21.199466>
- [19] Zimmermann, H.-J. (2010). Fuzzy set theory: Fuzzy set theory. *Wiley Interdisciplinary Reviews. Computational Statistics*, 2(3), 317–332. <https://doi.org/10.1002/wics.82>
- [20] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/a:1010933404324>
- [21] Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6), 80. <https://doi.org/10.2307/3001968>
- [22] Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://doi.org/10.1007/bf02310555>
- [23] Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36. <https://doi.org/10.1007/bf02291575>
- [24] Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260), 583–621. <https://doi.org/10.1080/01621459.1952.10483441>