engineering proceedings

MDPI

Proceeding Paper

Calamary: A Web Application for the Construction of Multi-Criteria Decision Models †

Ivan-Fredy Jaramillo *, Jose Alvarez-Mora, María Mercedes Alarcón Bermúdez and Valeria Torres-Lindao

Facultad de Ciencias de la Computación, Universidad Técnica Estatal de Quevedo, Quevedo 120501, Ecuador; jalvarezm7@uteq.edu.ec (J.A.-M.); malarconb@uteq.edu.ec (M.M.A.B.); vtorresl@uteq.edu.ec (V.T.-L.)

* Correspondence: ijaramillo@uteq.edu.ec

† Presented at the Sixth International Conference on Information Systems and Software Technologies (ICI2ST 2026), Quito, Ecuador, 23–25 September 2026.

Abstract

Multi-criteria decision-making (MCDM) methods are increasingly applied to complex real-world decisions, yet the available software ecosystem remains dominated by desktop applications that are inaccessible, method-specific, or require programming expertise. This paper presents Calamary, an open-source, web-based decision support system that integrates five MCDM methods (AHP, MAUT, SAW, MOORA, and TOPSIS) within a single modular platform supporting the complete decision analysis cycle: hierarchical model construction, weight assignment (direct, pairwise, and collaborative expert modes), utility function construction, alternative evaluation, and sensitivity analysis. The system was developed following a Component-Based Software Engineering methodology using React and Supabase. Algorithmic validation was conducted through replication of 13 cases from 10 peer-reviewed studies across diverse domains, achieving a global Mean Relative Error (MRE) of 0.2325% and 100% rank concordance. A usability study with 12 academic participants yielded a System Usability Scale (SUS) score of 79.8 (SD = 8.4, classified as Good) and a perceived relevance score of 4.6/5.0. These results confirm that Calamary provides a reliable, accessible, and extensible platform for MCDM research and practice.

Keywords: multi-criteria decision-making (MCDM); analytic hierarchy process (AHP); web application; decision support system; sensitivity analysis

1. Introduction

Multi-criteria decision-making (MCDM) encompasses formal methods for evaluating a finite set of alternatives against multiple, often conflicting criteria [1–3]. Since the formalization of the Analytic Hierarchy Process (AHP) by Saaty in the 1970s [4], MCDM methods have been applied across engineering, supply chain management, healthcare, environmental science, and public policy. A systematic analysis of 393 articles between 2000 and 2014 confirmed that AHP is the most widely used individual method, followed by hybrid approaches combining multiple techniques [5]. A more recent bibliometric study analyzing 3655 peer-reviewed articles between 2004 and 2024 confirmed an accelerating trend toward hybrid MCDM models integrating artificial intelligence and fuzzy logic, with Computer Science leading application domains with 2552 publications [6]. A complementary analysis of MCDM scientific production further corroborates the sustained growth of AHP and the emergence of multi-method approaches [7].

Despite this research growth, the software ecosystem for MCDM practice exhibits a persistent structural gap. A systematic review identified 190 software resources related

Check for updates

Academic Editor: Fredy Gavilanes

Published: 18 September 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

to MCDM, of which only 29 met rigorous inclusion criteria [8]. A separate review of 16 tools documented in 25 scientific studies found that 62.5% are desktop applications, with web-based solutions representing only 12.5% of the ecosystem [9]. Widely adopted platforms such as Expert Choice and SuperDecisions are restricted to AHP, while systems like GMAA (Generic Multi-Attribute Analysis) are limited to additive utility models [10]. No single tool integrates AHP, Multi-Attribute Utility Theory (MAUT), Simple Additive Weighting (SAW), Multi-Objective Optimization on the basis of Ratio Analysis (MOORA), and the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) within a unified, accessible, web-based environment. This fragmentation forces practitioners to learn multiple tools, manually transfer data, or resort to spreadsheets, reducing reproducibility and limiting MCDM adoption in research and practice.

A secondary barrier is accessibility. As documented by Weistroffer and Li [9], the predominance of desktop-only architectures (installation requirements, commercial licensing, and limited mobile compatibility) restricts existing tools to desktop environments, excluding users who work from shared computers, tablets, or constrained institutional networks, a particularly acute constraint in developing-country academic contexts where MCDM education is expanding.

This paper presents Calamary, an open-source web-based MCDM platform designed to address these gaps. Calamary operationalizes the problem-structuring stage of Howard's decision analysis framework [11], adapted to Toskano's multi-criteria decision-making cycle [12]. The decision problem is structured by defining objectives, criteria, and alternatives; weights are elicited under one of several modes; alternatives are evaluated with the selected MCDM method; and sensitivity analysis identifies the criteria with the greatest influence on the final ranking, allowing the user to revisit and refine the model before acting on the result. Section 3 details the six-stage user-facing workflow that implements this cycle.

This paper makes four specific contributions. First, Calamary integrates AHP [4], MAUT [13], SAW [14], MOORA [15], and TOPSIS [16] within a modular, component-based architecture accessible from any device, supporting a shared hierarchical modeling framework, collaborative expert elicitation, and sensitivity analysis. Second, an algorithmic validation protocol based on replication of 13 published cases from 10 peer-reviewed studies demonstrates a global MRE of 0.2325% and 100% rank concordance. Third, a usability evaluation using the System Usability Scale (SUS) with 12 domain-expert participants yields a score of 79.8 (*Good*). Finally, the system is released as open-source software with publicly available validation datasets.

The remainder of this paper is organized as follows: Section 2 reviews existing MCDM software and its limitations. Section 3 describes system development and validation design. Section 4 presents results. Section 5 discusses findings, limitations, and future directions. Section 6 concludes the paper.

2. Related Work

The landscape of MCDM software has been documented in several surveys. Weistroffer and Li [9] provided a comprehensive review identifying a clear predominance of desktop applications. Figueroa-Perez et al. [8] conducted a recent systematic literature review identifying 190 resources, with AHP and TOPSIS as the predominantly implemented methods and 29 tools meeting rigorous quality criteria.

Among the most widely used tools, Expert Choice and SuperDecisions specialize exclusively in AHP, offering hierarchical modeling and pairwise comparison but lacking support for other method families. M-MACBETH implements the MACBETH method with qualitative judgment elicitation [17,18]. Visual PROMETHEE provides advanced visualization for PROMETHEE rankings [19]. The GMAA system supports additive utility

models [10]. DEXi and its successor DEXiSuite implement qualitative decision tables. None of these tools provides a unified environment integrating AHP, MAUT, SAW, MOORA, and TOPSIS within a web-based interface.

Web-based MCDM solutions are scarce. PROMETHEE-Cloud implements the PROMETHEE method with Monte Carlo sensitivity analysis, but is restricted to a single method family and has minimal adoption in the literature (three citations as of 2024). Web-DSS-MAUT [20], the direct successor to GMAA, supports MAUT with interactive utility function definition but lacks pairwise comparison methods; a related multi-expert elicitation extension has since been applied to environmental decisions [21]. Programmatic solutions such as pymcdm (Python v3.14.6) and RMCDA (R 4.6.0 package) offer broader method coverage but require programming expertise, excluding non-technical practitioners.

Usability has received limited attention in MCDM software evaluation. The majority of published tool descriptions rely on expert opinion rather than formal metrics. Where the System Usability Scale (SUS) has been applied, scores for academic tools typically cluster around the 68-point acceptability threshold [22], making structured usability benchmarking uncommon in this domain.

Figueroa-Perez et al. [8] analyzed their 29 selected resources along five dimensions (resource type, implemented methods, operating-system support, license/openness, and design for reuse) that directly informed Calamary's design. Most resources are domain-specific desktop frameworks mainly implementing AHP/TOPSIS, being free and open-source but rarely designed for reuse beyond the originating project. Calamary is instead general-purpose, web-based, and extends coverage to MAUT, SAW, and MOORA; its CBSE architecture (Section 3) directly responds to the reusability gap identified by Figueroa-Perez et al., implementing methods as independent modules extensible without refactoring the core application.

The gap addressed by this paper, a single web-based platform integrating multiple MCDM methods with formal algorithmic validation and structured usability assessment, remains unfilled in the reviewed literature.

3. Materials and Methods

The study was structured in three phases: (1) identification and analysis of requirements, (2) development of the web application, and (3) validation of the system. This design integrates empirical evidence from existing tools with specific user needs, ensuring that the proposed solution is both technically sound and practically useful (see Figure 1).

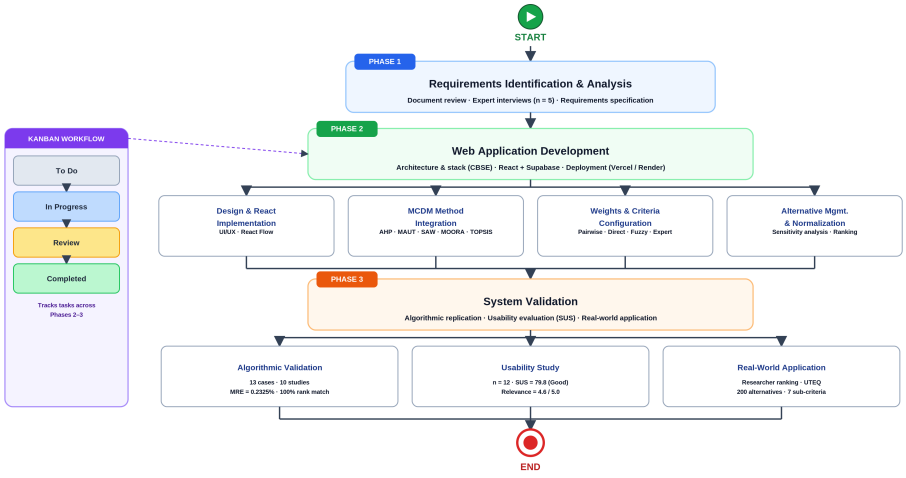


Figure 1. Project methodology.

3.1. Requirement Analysis

3.1.1. Document Review

A systematic document review of existing MCDM software was conducted using Scopus, IEEE Xplore, and ACM Digital Library, with search terms including "MCDM software", "multi-criteria decision support system", "AHP software", and "TOPSIS implementation". Inclusion criteria were as follows: (a) software documented in academic literature, (b) specific focus on MCDM, and (c) sufficient functional documentation. Twenty-five studies describing 16 tools were identified. For each tool, the following were documented: supported methods, main functionalities, access type, and limitations.

3.1.2. Interviews with Potential Users

Semi-structured interviews were conducted with five domain experts: two university students with formal MCDM coursework, two software engineers with decision modeling experience, and one practitioner with professional decision-making responsibilities. Interviews lasted 30–45 min and explored prior tool experience, desired functionalities, visualization preferences, and data entry expectations. Transcripts were analyzed thematically to identify recurring requirement patterns.

3.2. MCDM Methods Implemented

Calamary implements five MCDM methods, selected over other prevalent methods such as PROMETHEE, ELECTRE, VIKOR, or MACBETH because SAW, AHP, TOPSIS, and MOORA were classified as must-have requirements in the document review and user interviews (Section 4.1), while MAUT closes the additive utility gap left by desktop-only tools such as GMAA; further methods (ELECTRE, VIKOR, COPRAS) are left for future work (Section 5). Their core mathematical formulations are described below.

3.2.1. Simple Additive Weighting (SAW)

SAW applies vector normalization followed by a weighted sum. For criteria to maximize, normalization is $r_{ij} = x_{ij}/\max_i(x_{ij})$; for criteria to minimize, $r_{ij} = \min_i(x_{ij})/x_{ij}$. Negative values are shifted to positive using $r_{ij} = (x_{ij} - \min_i)/(\max_i - \min_i)$. The aggregate score is as follows:

$$S_i = \sum_{j=1}^{n} w_j \cdot r_{ij} \tag{1}$$

where w_j is the normalized weight of criterion j.

3.2.2. Multi-Objective Optimization on Ratio Analysis (MOORA)

MOORA applies ratio-based normalization: $r_{ij} = x_{ij}/\sqrt{\sum_i x_{ij}^2}$. The aggregate score separating beneficial and non-beneficial criteria is as follows:

$$y_i = \sum_{j \in \mathcal{B}} w_j \cdot r_{ij} - \sum_{j \in \mathcal{N}} w_j \cdot r_{ij} \tag{2}$$

where $\mathcal{B}$ and $\mathcal{N}$ denote beneficial and non-beneficial criteria sets.

3.2.3. Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)

After vector normalization and weighting, TOPSIS computes Euclidean distances to the positive ideal solution A^+ and negative ideal solution A^-:

$$d_i^+ = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_j^+)^2}, \quad d_i^- = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_j^-)^2} \tag{3}$$

The closeness coefficient $C_i = d_i^- / (d_i^+ + d_i^-)$ ranks alternatives in descending order.

https://doi.org/10.3390/engproc2026153016

3.2.4. Analytic Hierarchy Process (AHP)

AHP derives weights from pairwise comparison matrices using the principal eigenvector method. The Consistency Ratio CR = CI/RI must satisfy CR < 0.10, where CI $= (\lambda_{\max} - n)/(n-1)$ and RI is the random index for matrix order n. For collaborative expert elicitation, individual eigenvectors are aggregated using the geometric mean:

$$\bar{w}_i = \frac{\left(\prod_{j=1}^{m} w_{ij}\right)^{1/m}}{\sum_{k=1}^{n}\left(\prod_{j=1}^{m} w_{kj}\right)^{1/m}} \tag{4}$$

where m is the number of experts, and w_{ij} is the weight assigned by expert j to criterion i.

3.2.5. Multi-Attribute Utility Theory (MAUT)

MAUT elicits a utility function $u_j(x)$ for each criterion, mapping values to $[0,1]$. Calamary supports linear and piecewise-linear functions defined by reference points at minimum and maximum utility values. The aggregate utility is as follows:

$$U_i = \sum_{j=1}^{n} w_j \cdot u_j(x_{ij}) \tag{5}$$

3.3. System Development

3.3.1. Architecture and Technology Stack

Calamary was developed using a Component-Based Software Engineering (CBSE) methodology, selected for its ability to produce modular, reusable, and maintainable systems, characteristics essential for integrating multiple MCDM methods with independent computational logic. CBSE enables new methods to be added as independent modules without refactoring the core application.

The technology stack comprises the following: React (v18) for the frontend single-page application; React Flow for interactive drag-and-drop hierarchical model visualization; Supabase (PostgreSQL) for persistent model storage and real-time expert collaboration; Python (server-side) for MCDM computations; and Vercel/Render for cloud deployment. The system follows a three-tier architecture separating presentation, business logic, and data layers. Development workflow was managed using Kanban with four states: To Do, In Progress, Review, and Completed.

3.3.2. Key System Components

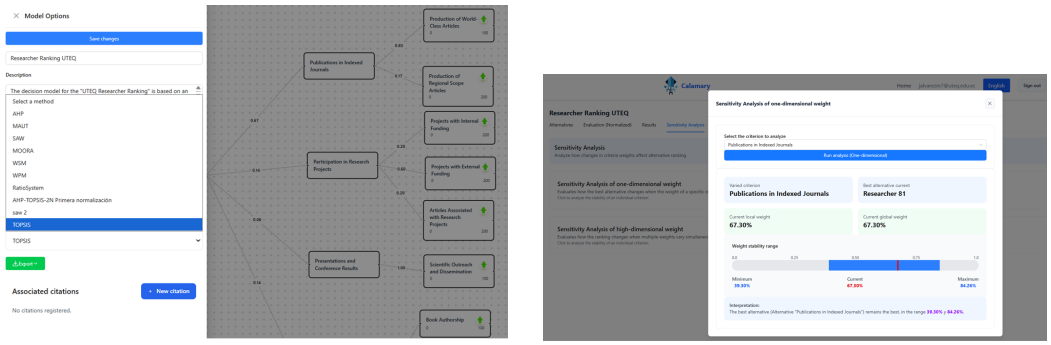
The platform implements six functional components aligned with the decision analysis lifecycle:

- Hierarchy Board: This is a visual drag-and-drop interface for constructing multi-level criteria hierarchies using React Flow. It supports parent–child relationships, criterion type definition (quantitative/qualitative, maximize/minimize), and model templates.
- Direct Weight Assignment: This is a normalized weight input with real-time consistency feedback for each hierarchy node.
- Saaty Matrix Component: This is a pairwise comparison matrix with automatic eigenvector computation and Consistency Ratio validation (CR < 0.10 enforced).
- Expert Collaboration Component:This is an asynchronous expert elicitation via unique token-based invitations, with geometric mean aggregation of individual weight vectors.
- MAUT Utility Editor: This gives an interactive definition of linear and piecewise utility functions per criterion.

https://doi.org/10.3390/engproc2026153016

- Sensitivity Analysis: This is a unidimensional weight perturbation. The system varies the weight of any selected criterion across its feasible range, redistributes remaining weights proportionally, and recomputes rankings to identify stability thresholds.

Figure 2 illustrates two representative views of the Calamary interface. The Hierarchy Board (Figure 2a) provides an interactive drag-and-drop canvas for constructing multi-level criteria trees using React Flow nodes, enabling users to define criterion types, weights, and parent–child relationships without programming. The Sensitivity Analysis panel (Figure 2b) displays how ranking positions of alternatives change as the weight of a selected criterion is varied across its feasible range, allowing decision-makers to identify robust choices and critical criteria.



(**a**) (**b**)

Figure 2. Representative interface views of Calamary. (**a**) Hierarchy Board: visual drag-and-drop canvas for multi-level criteria construction using React Flow; (**b**) sensitivity analysis panel: ranking variation as a function of selected criterion weight, with stability threshold detection.

From the user's perspective, working with Calamary follows six sequential stages—authentication and model creation, hierarchy design, weight assignment (direct, Saaty, expert, or Fuzzy-AHP, plus utility configuration if MAUT is selected), alternative loading, evaluation and results, and sensitivity analysis, with an optional loop from sensitivity analysis back to hierarchy design for iterative refinement (Figure 3).

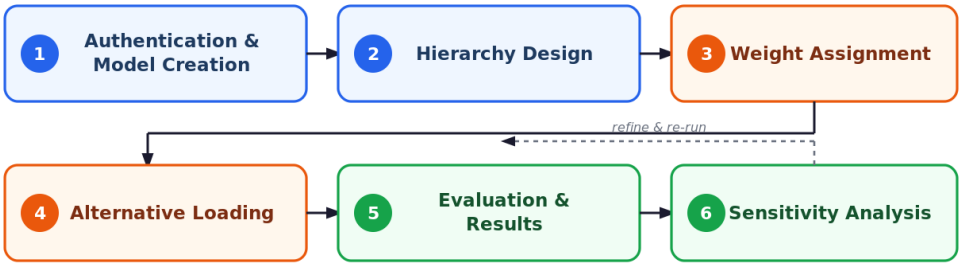


Figure 3. Calamary user workflow.

3.4. System Validation

3.4.1. Algorithmic Validation by Replication of Published Cases

To assess computational accuracy, 13 cases were selected from 10 peer-reviewed publications covering diverse domains: city selection (SAW, MOORA) [23]; supplier selection under sustainability criteria (AHP) [24]; electric demand forecasting model selection (SAW) [25]; weapon system evaluation (Fuzzy-AHP) [26]; supplier selection integrating AHP and SAW [27]; mobile banking app quality assessment (AHP) [28]; investment portfolio selection from the Philippine Stock Exchange (AHP) [29]; gearmotor supplier selection (Fuzzy-AHP) [30]; strategic priority setting with SWOT–Fuzzy-AHP [31]; and

cloud computing load balancing (AHP) [32]. For each case, input data (alternatives, criteria, pairwise matrices) were replicated identically, and results were compared using Mean Absolute Error (MAE) and Mean Relative Error (MRE). An MRE < 1% was established as the precision threshold.

3.4.2. Usability Evaluation

An intentional sample of 12 participants was formed, distributed among 8 Software Engineering students who completed the "Decision and Optimization Models" course, 3 practicing software engineers, and 1 professional decision analyst. Inclusion criteria required either prior MCDM knowledge (students and analyst) or documented experience in software evaluation (engineers). The sample size is justified by established usability testing guidelines suggesting that 10–12 participants identify 85–95% of significant usability problems in qualitative evaluations.

Evaluations were conducted individually under supervised remote modality. Each participant completed a representative task: constructing a complete multi-criteria decision model with five criteria, three sub-criteria, and six alternatives using both direct weight assignment and AHP pairwise comparison within a 30-min session. Participants provided informed consent prior to participation. Usability was assessed with the validated 10-item System Usability Scale (SUS) [22,33], yielding a 0–100 score. A 5-item Likert (1–5) instrument assessed perceived relevance of the system. Open-ended questions explored qualitative perceptions and intention of future use.

4. Results

4.1. Requirement Characterization

The document review identified 16 specialized MCDM tools documented in 25 scientific studies. The analysis revealed a marked predominance of desktop applications (62.5%) versus web solutions (12.5%), frameworks (6.25%), R packages (6.25%), and Python libraries (12.5%) (see Table 1). Desktop tools such as Expert Choice (1015 citations) and M-MACBETH (106 citations) dominate in academic recognition, while PROMETHEE-Cloud (3 citations) and Web-DSS-MAUT (0 citations) remain largely unrecognized. A total of 63 MCDM methods were detected across the surveyed tools, with AHP and PROMETHEE predominating.

Table 1. MCDM tools identified in the document review.

Tool	Type	Methods	Main Functionalities
Expert Choice	Desktop	AHP	Hierarchies; pairwise comparison; sensitivity analysis
Excel	Desktop	Any method	Manual calculations; reusable templates
Super Decisions	Desktop	AHP, ANP	Hierarchical models; influence matrices
MATLAB	Desktop	Scripts	MCDM via scripts; scientific visualization
1000 minds	Desktop	PAPRIKA	Preference surveys; qualitative criteria
M-MACBETH	Desktop	MACBETH	Qual./quant. modeling; sensitivity analysis
Visual PROMETHEE	Desktop	PROMETHEE, GAIA	Multi-criteria ranking; visualization
GMAA System	Desktop	Additive MAUT	Uncertainty handling; incomplete info
DEXi/DEXi Suite	Desktop	DEX	Decision tables; what-if analysis
MCDA-ULAVAL	Desktop	ELECTRE TRI	Classification; sensitivity analysis
Desdeo	Framework	NIMBUS, Nautilus	Interactive multi-objective optimization
PROMETHEE-Cloud	Web	PROMETHEE	Monte Carlo sensitivity; Excel export
Web-DSS-MAUT	Web	MAUT	Interactive hierarchies; utility functions
RMCDA	R Package	AHP, TOPSIS, VIKOR	Web apps; radar/heatmaps; CSV
pymcdm	Python Lib	TOPSIS, MOORA, VIKOR, AHP	Normalization types; no GUI
Scikit-Criteria	Python Lib	ELECTRE, MOORA, TOPSIS	Python ecosystem integration

Based on the document review and user interviews, nine priority functional requirements were defined. Essential (must) requirements include visual construction of criteria hierarchies and implementation of SAW, AHP, TOPSIS, and MOORA. Important (should) requirements include pairwise comparison with consistency checking, sensitivity analysis, template reuse, and support for qualitative and quantitative criteria. Six non-functional requirements were established, covering scalability (5 to >100 alternatives), browser portability, data persistence, architectural extensibility, integrated documentation, and 99% minimum availability.

Table 2 presents the comparative feature matrix between Calamary and representative existing tools.

Table 2. Feature comparison: Calamary versus representative existing tools.

Feature	Calamary	Expert Choice	M-MACBETH	Visual PROMETHEE	Web-DSS-MAUT	RMCDA
Web access (no install)	✓	×	×	×	✓	✓
Multiplatform (mobile/tablet)	✓	×	×	×	~	~
No programming requirement	✓	✓	✓	✓	✓	✓
SAW/MOORA/TOPSIS	✓	×	×	×	×	✓
AHP (pairwise comparisons)	✓	✓	×	×	×	✓
Fuzzy-AHP	✓	×	×	×	×	✓
MAUT (utility functions)	✓	×	×	×	✓	×
Visual hierarchical editor	✓	✓	✓	×	✓	×
Collaborative expert mode	✓	×	×	×	×	×
Sensitivity analysis	✓	✓	✓	✓	×	~
Open source	✓	×	×	×	×	✓

~ = partial support; ✓ = supported; × = not supported.

4.2. Algorithmic Validation

The precision of the system was measured through the replication of 13 cases from 10 published scientific studies covering diverse domains including supplier selection, weapon evaluation, finance, and cloud computing (see Table A1 in Appendix A). The results are as follows:

- Global Mean Absolute Error (MAE): 0.000467;
- Global Mean Relative Error (MRE): 0.2325%;
- Ranking match: 100% in all 13 replicated cases:
- Consistency Ratio (AHP cases): 0.092 and 0.022, identical to reference values.

All values remain well below the 1% MRE threshold. The two Fuzzy-AHP cases achieved zero error (MRE = 0.000%), and the most complex case, 30-alternative portfolio selection from the Philippine Stock Exchange, achieved MRE = 0.015%, confirming scalability of the algorithmic implementation.

4.3. Usability Findings

4.3.1. SUS Scores

The SUS evaluation yielded a mean score of 79.8 (SD = 8.4), classified as *Good* according to the Bangor et al. adjective scale [22]. Individual scores ranged from 65 to 92.5, with 10 of 12 participants (83.3%) scoring above the 68-point acceptability threshold. The score of 79.8 exceeds the general SUS mean of 70.14 reported by Bangor et al. across all evaluated systems and surpasses the 75.4 mean typically reported for educational and research tools. Table 3 summarizes results by participant group.

Table 3. SUS scores and perceived relevance by participant group.

Group	*n*	SUS Mean (SD)	SUS Range	Relevance (1–5)
SE Students (MCDM course)	8	80.3 (9.1)	65–92.5	4.7
Software Engineers	3	78.3 (6.4)	72.5–85.0	4.4
Decision Professional	1	80.0 (—)	80.0	4.5
Overall	12	79.8 (8.4)	65–92.5	4.6

4.3.2. Qualitative Findings

The best-rated functionalities were the graphical hierarchy representation (React Flow board) and the flexibility of weight assignment modes, highlighted by participants across all groups. The collaborative expert elicitation feature was identified as a key differentiator, addressing a practical need not covered by any other tool participants had previously used.

The primary usability concern, reported by 7 of 12 participants (58.3%), was the absence of inline tutorial guidance for first-time users. Participants without formal MCDM training required clarification on sensitivity analysis interpretation. The software engineers (without MCDM background) showed the lowest SUS scores (mean 78.3), suggesting that the interface assumes some domain familiarity. Two participants noted that the absence of PDF/Excel export functionality would limit adoption in professional reporting workflows.

5. Discussion

5.1. Addressing the MCDM Software Gap

The primary motivation for developing Calamary was the absence of a web-based, multi-method MCDM platform accessible to non-programmer practitioners. The results confirm that this gap can be effectively addressed through CBSE: the modular design allowed five computationally distinct methods to share a single hierarchical modeling interface while maintaining independent algorithmic implementations. This contrasts with tools such as Expert Choice, which requires a complete platform change when a problem calls for TOPSIS or MOORA. The dynamic function injection pattern used in Calamary enables users to switch methods or compare results across methods within a single session, a workflow not supported by any reviewed tool. The collaborative expert elicitation feature further differentiates Calamary, since no existing web-based MCDM tool provides asynchronous multi-expert weight aggregation with token-based access control.

5.2. Algorithmic Fidelity and Usability in Context

The global MRE of 0.2325% across 13 replicated cases demonstrates that the browser-based JavaScript/Python implementation matches academic-grade precision. The 100% rank concordance across all cases is the most practically significant metric: it confirms that the ordering of alternatives produced by Calamary is always consistent with the original publication, which is the decision-relevant outcome. The near-zero error in Fuzzy-AHP cases (MRE = 0.000%) reflects triangular number arithmetic being executed server-side in Python, where numerical libraries provide higher-precision operations than browser-side floating-point.

The SUS score of 79.8 (SD = 8.4) situates Calamary above the general SUS benchmark of 70.14 and above typical scores for academic research tools [22]. The *Good* classification (rather than *Excellent*, which requires >90) primarily reflects the identified gap in tutorial support and the learning curve for domain-unfamiliar users. The perceived relevance score of 4.6/5.0 indicates strong fit between the tool and its intended context.

5.3. Limitations and Threats to Validity

Several limitations must be acknowledged. First, the usability sample ($n = 12$) is appropriate for exploratory evaluation but insufficient for statistical inference about the

broader population of MCDM users; future work should target $n \geq 30$ with stratified sampling across professional and academic contexts. Second, all participants were from an academic context in Ecuador; generalizability to industrial or governmental settings remains to be assessed. Third, the validation protocol focused on replication accuracy rather than decision quality whether Calamary helps users make better decisions compared to unassisted analysis was not evaluated. Fourth, the absence of a controlled comparison condition (e.g., Expert Choice versus Calamary on the same task) limits the ability to make comparative usability claims.

Threats to internal validity include the possibility that voluntary participants have above-average comfort with web tools. Threats to external validity include the academic and developing-country context, which may not represent the full diversity of MCDM user populations.

6. Conclusions

This paper presented Calamary, a web-based multi-method MCDM platform addressing the accessibility and fragmentation limitations of existing software. The four contributions are as follows: (i) a unified open-source platform integrating AHP, MAUT, SAW, MOORA, and TOPSIS with hierarchical modeling, collaborative expert elicitation, and sensitivity analysis; (ii) algorithmic validation through replication of 13 published cases with a global MRE of 0.2325% and 100% rank concordance; (iii) a usability evaluation yielding SUS = 79.8 (*Good*); and (iv) an open-source release at https://github.com/srmora-ec/Calamary (accessed on 26 May 2026).

The results confirm that it is technically feasible to build a web platform matching academic-grade algorithmic precision while delivering usability above general benchmarks. The component-based architecture successfully accommodates five computationally distinct methods without architectural coupling.

Future Work

Several extensions are planned, encompassing (1) expanding the method library to include ELECTRE, VIKOR, and COPRAS; (2) implementing PDF and Excel export for professional reporting workflows, the most frequently requested feature in the usability study; (3) conducting a larger-scale usability study with $n \geq 30$ participants from diverse professional backgrounds; (4) evaluating decision quality outcomes in a controlled experiment comparing Calamary against spreadsheet-based approaches; and (5) assessing performance under large-scale models (>100 criteria and alternatives).

Author Contributions: Conceptualization, J.A.-M. and I.-F.J.; methodology, J.A.-M. and I.-F.J.; software, J.A.-M.; validation, J.A.-M. and I.-F.J.; formal analysis, J.A.-M.; investigation, J.A.-M.; resources, M.M.A.B.; data curation, J.A.-M. and M.M.A.B.; writing—original draft preparation, J.A.-M. and V.T.-L.; writing—review and editing, M.M.A.B. and V.T.-L.; visualization, J.A.-M. and V.T.-L.; supervision, I.-F.J.; project administration, I.-F.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study due to the non-sensitive nature of the usability evaluation conducted. No clinical procedures, sensitive personal data, or vulnerable populations were involved. All participants were adult volunteers who provided informed consent prior to participation.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The source code of Calamary is available at https://github.com/srmora-ec/Calamary (accessed on 26 May 2026). The validation dataset supporting the reported results is included in Appendix A.

Acknowledgments: The authors acknowledge the support of the Faculty of Computer Science at Universidad Técnica Estatal de Quevedo. During the preparation of this manuscript, generative AI tools were used for grammar and spelling revision. The authors have reviewed and edited all outputs and take full responsibility for the content of this publication.

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

Abbreviations

The following abbreviations are used in this manuscript:

AHP	Analytic Hierarchy Process
CBSE	Component-Based Software Engineering
MAE	Mean Absolute Error
MAUT	Multi-Attribute Utility Theory
MCDM	Multi-Criteria Decision-Making
MOORA	Multi-Objective Optimization on the basis of Ratio Analysis
MRE	Mean Relative Error
SAW	Simple Additive Weighting
SUS	System Usability Scale
TOPSIS	Technique for Order of Preference by Similarity to Ideal Solution

Appendix A. Validation Results by Replication of Scientific Models

Table A1. Validation results by replication of scientific models.

Model	Method	Alt.	MAE	MRE (%)	Ranking Match
City selection (SAW)	SAW	8	0.000325	0.042	100%
City selection (MOORA)	MOORA	8	0.000275	0.083	100%
Sugar suppliers ($Ca(OH)_2$)	AHP	2	0.003550	0.678	100%
Sugar suppliers (Flocculant)	AHP	3	0.004150	0.881	100%
Sugar suppliers (Alcohol)	AHP	2	0.002050	0.428	100%
Electrical forecasting	SAW	3	0.000433	0.055	100%
Weapon systems	Fuzzy-AHP	5	0.000000	0.000	100%
UD BSA Suppliers	AHP/SAW	3	0.000000	0.000	100%
Mobile banking apps	AHP	6	0.000833	0.523	100%
PSE Portfolios	AHP	30	0.000007	0.015	100%
Gearmotor suppliers	Fuzzy-AHP	3	0.000100	0.030	100%
MSC Priorities	SWOT/Fuzzy-AHP	4	0.000325	0.147	100%
Load balancing	AHP	4	0.000325	0.141	100%
Global			0.000467	0.2325	100%

MAE = Mean Absolute Error; MRE = Mean Relative Error; Alt. = number of alternatives. Sources: [23–32].

References

1. Qin, Y.; Qi, Q.; Shi, P.; Lou, S.; Scott, P.J.; Jiang, X. Multi-Attribute Decision-Making Methods in Additive Manufacturing: The State of the Art. *Processes* **2023**, *11*, 497. [CrossRef]
2. Yu, B.; Xu, Z.; Dai, J.; Yang, T. A novel multi-attribute decision-making method based on neighborhood approximations and its application. *Expert Syst. Appl.* **2022**, *199*, 116946. [CrossRef]
3. Yoon, K.; Hwang, C.L. Multiple Attribute Decision Making. In *Multiple Attribute Decision Making*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 16–41. [CrossRef]
4. Saaty, T.L.; Kearns, K.P. The Analytic Hierarchy Process. *Anal. Plan.* **1985**, *41*, 19–62. [CrossRef]
5. Mardani, A.; Jusoh, A.; Nor, K.M.; Khalifah, Z.; Zakwan, N.; Valipour, A. Multiple criteria decision-making techniques and their applications—A review of the literature from 2000 to 2014. *Econ. Res.-Ekon. Istraz.* **2015**, *28*, 516–571. [CrossRef]
6. Kumar, R.; Pamucar, D. A Comprehensive and Systematic Review of Multi-Criteria Decision-Making (MCDM) Methods to Solve Decision-Making Problems: Two Decades from 2004 to 2024. *Spectr. Decis. Mak. Appl.* **2025**, *2*, 177–196. [CrossRef]

7. Jaramillo, I.F.; Torres-Lindao, V.; Cevallos, J. Bibliometric Analysis of Scientific Production on Multicriteria Decision-Making (MCDM) Methods. In *Proceedings of the International Conference on Applied Technologies*; Botto-Tobar, M., Lema Moreta, L., Zambrano Vizuete, M., Montes León, S., Torres-Carrion, P., Durakovic, B., Eds.; Springer: Cham, Switzerland, 2026; Volume 2949, pp. 222–235. [CrossRef]
8. Figueroa-Perez, J.F.; Leyva-Lopez, J.C.; Ramirez-Noriega, A.D.; Perez-Contreras, E.O. Software Resources for Programming Multicriteria Decision Support Systems: A Systematic Literature Review. In *Proceedings of the Lecture Notes in Networks and Systems*; Springer: Cham, Switzerland, 2025; Volume 1447, pp. 337–353. [CrossRef]
9. Weistroffer, H.R.; Li, Y. Multiple criteria decision analysis software. In *International Series in Operations Research and Management Science*; Springer: New York, NY, USA, 2016; Volume 233, pp. 1301–1341. [CrossRef]
10. Jiménez, A.; Ríos-Insua, S.; Mateos, A. A generic multi-attribute analysis system. *Comput. Oper. Res.* **2006**, *33*, 1081–1101. [CrossRef]
11. Howard, R.A. Decision Analysis: Applied Decision Theory. In *Proceedings of the Fourth International Conference on Operational Research*; Hertz, D.B., Melese, J., Eds.; Wiley-Interscience: New York, NY, USA, 1966; pp. 55–71.
12. Hurtado Toskano, G.B. El Proceso de Análisis Jerárquico (AHP) como Herramienta para la Toma de Decisiones en la Selección de Proveedores. Master's Thesis, Universidad Nacional Mayor de San Marcos, Lima, Peru, 2005.
13. Alinezhad, A.; Khalili, J. MAUT Method. In *New Methods and Applications in Multiple Attribute Decision Making (MADM)*; Springer International Publishing: Cham, Switzerland, 2019; *International Series in Operations Research & Management Science, 277*; pp. 127–131. [CrossRef]
14. Taherdoost, H. Analysis of Simple Additive Weighting Method (SAW) as a MultiAttribute Decision-Making Technique: A Step-by-Step Guide. *J. Manag. Sci. Eng. Res.* **2023**, *6*, 21–24. [CrossRef]
15. Brauers, W.K.; Zavadskas, E.K. Robustness of the multi-objective moora method with a test for the facilities sector; [Daugiatikslio moora method patikimumo tikrinimas paslaugų sektoriuje]. *Technol. Econ. Dev. Econ.* **2009**, *15*, 352–375. [CrossRef]
16. Chakraborty, S. TOPSIS and Modified TOPSIS: A comparative analysis. *Decis. Anal. J.* **2022**, *2*, 100021. [CrossRef]
17. Bana e Costa, C.A.; Vansnick, J.C. MACBETH—An interactive path towards the construction of cardinal value functions. *Int. Trans. Oper. Res.* **1994**, *1*, 489–500. [CrossRef]
18. Bana e Costa, C.A.; Vansnick, J.C. The MACBETH Approach: Basic Ideas, Software, and an Application. In *Advances in Decision Analysis*; Meskens, N., Roubens, M., Eds.; Springer: Dordrecht, The Netherlands, 1999; *Mathematical Modelling: Theory and Applications, 4*; pp. 131–157. [CrossRef]
19. Behzadian, M.; Kazemzadeh, R.B.; Albadvi, A.; Aghdasi, M. PROMETHEE: A comprehensive literature review on methodologies and applications. *Eur. J. Oper. Res.* **2010**, *200*, 198–215. [CrossRef]
20. Gómez Jiménez, A. Desarrollo de un Sistema de Ayuda a la Decisión Web Multicriterio Basado en un Modelo en Utilidad Multi-Atributo Aditivo: El Sistema WEB-DSS-MAUT. Master's Thesis, Universidad Politécnica de Madrid, Madrid, Spain, 2024.
21. Gómez-Jiménez, A.; Jiménez-Martín, A.; Chergui, Z.; Marín, S. A hierarchical elicitation process based on the combination of different weighting methods within multi-attribute utility theory in a group decision-making context. *Ann. Oper. Res.* **2026**. [CrossRef]
22. Bangor, A.; Kortum, P.T.; Miller, J.T. An empirical evaluation of the system usability scale. *Int. J. Hum.-Comput. Interact.* **2008**, *24*, 574–594. [CrossRef]
23. Landeta, J.M.I.; Saldaña, J.A.N.; Camacho, L.A.Z.; Landeta, M.L.I. Selección De La Ciudad Para Vivir Mediante Modelos De Decisión Multicriterio. *Investig. Oper.* **2024**, *45*, 339–353.
24. Ocampo-Murillo, H.F.; Quintero-Garzón, M.L. Selección de proveedores de insumos críticos en términos de sostenibilidad, a través de la metodología multicriterio, en una empresa del sector azucarero. *Entramado* **2020**, *16*, 24–44. [CrossRef]
25. Yajure Ramírez, C. Selección multicriterio del modelo óptimo para el pronóstico de la demanda eléctrica máxima diaria considerando la satisfacción de la demanda real. *Rev. Investig. Tecnol. Inf.* **2025**, *13*, 38–49. [CrossRef]
26. Cheng, C.H.; Mon, D.L. Evaluating weapon system by Analytical Hierarchy Process based on fuzzy scales. *Fuzzy Sets Syst.* **1994**, *63*, 1–10. [CrossRef]
27. Florencia, J.P.T. The Integration of AHP and SAW Methods with Multiple Decision-Makers for Supplier Selection. A Case Study of UD BSA, Surabaya. In *Proceedings of the International Conference on Industrial Engineering and Operations Management*; IEOM Society International: Southfield, MI, USA, 2023; pp. 386–396. [CrossRef]
28. Goceri, M.S. Quality Assessment of Mobile Banking Apps with Expert Users Using AHP Decision Making Tool. In *Proceedings of the 2024 8th International Conference on Information System and Data Mining*; Association for Computing Machinery: New York, NY, USA, 2024; pp. 117–121. [CrossRef]
29. Duran, M.K.L.; Young, M.N.; Mariñas, K.A. Portfolio Selection Utilizing Analytical Hierarchy Process (AHP) with Mean Variance Theory and Safety-first Model: A Study on the Top 30 Companies in the Philippine Stock Exchange. In *2024 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*; IEEE: Piscataway, NJ, USA, 2024; pp. 638–642. [CrossRef]
30. Ayhan, M.B. A Fuzzy Ahp Approach for Supplier Selection Problem: A Case Study in a Gearmotor Company. *Int. J. Manag. Value Supply Chain.* **2013**, *4*, 11–23. [CrossRef]

31. Ohoitimur, J.; Krejci, J.; Raco, J.R.; Raton, Y.; Jamlean, A.; Welerubun, I.; Tanod, R.H. Strategic Management Study Based on Buckely'S Fuzzy-Ahp and Swot: Example of the Congregation of the Missionaries of the Sacred Heart. *Int. J. Anal. Hierarchy Process* **2021**, *13*, 329–354. [CrossRef]
32. Amandeep; Mohammad, F.; Yadav, V. Automatic decision making for multi-criteria load balancing in cloud environment using AHP. In *2015 International Conference on Computing, Communication & Automation*; IEEE: Piscataway, NJ, USA, 2015; pp. 569–576. [CrossRef]
33. Brooke, J. SUS: A retrospective. *J. Usability Stud.* **2013**, *8*, 29–40. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.