Received: 24 December 2024 | Revised: 16 June 2025 | Accepted: 30 June 2025 | Published online: 28 August 2025

RESEARCH ARTICLE

Journal of Data Science and Intelligent Systems
2025, Vol. 00(00) 1-10
DOI: 10.47852/bonviewJDSIS52025076

BON VIEW PUBLISHING

Handling Uncertain Information with Fuzzy Logic in Locating New Infrastructure

Manoj K. Jha[1,*]

[1] *Department of Information Technology, University of Maryland Global Campus, USA*

Abstract: Considerable research has been conducted in location optimization for new infrastructures such as highways and rail lines, as well as fixed facilities such as sports arenas, warehouses, and airports. Uncertainty plays a crucial role in determining the final location of such facilities. For example, there may be uncertainties about land-use and site characteristics as well as about demand for the new facilities. There may be disagreements among various stakeholders that complicate reaching a consensus due to budgetary constraints and differing political views. While some uncertainties can be quantitatively represented, others can only be represented qualitatively (e.g., low, medium, or high). Not all of these uncertainties can be precisely and mathematically modeled. While deep learning and other probabilistic techniques have been developed to deal with uncertainties that can be represented numerically, fuzzy logic has been recognized as a preferred choice for handling qualitative uncertainties. This paper identifies situations with uncertainties in locating new infrastructure and offers solutions for handling the qualitative uncertainties with fuzzy logic. An example is presented to illustrate the approach to handling uncertainties that can only be represented qualitatively. The results are promising for future research dealing with uncertainties represented as linguistic variables, thereby improving the decision-making process. The method can be applied in other domains involving uncertainty.

Keywords: infrastructure optimization, uncertainty theory, fuzzy logic, linguistic hedges

1. Introduction

Considerable research has been conducted on locating infrastructure such as fixed and continuous facilities. In the transportation context, such infrastructure includes highways, rail lines, airports, and canals [1]. Typically, mathematical formulations and solution algorithms are developed to optimize the location or alignment of such infrastructure [2, 3]. These formulations often consist of an objective function that must be minimized or maximized subject to several user-specified constraints [2, 4]. The objective function is generally expressed as a function of relevant decision variables. For example, Yin et al. [2] developed an objective function representing the difference between transportation revenue and costs to optimize transportation services using the case of the China railway system. Some of the decision variables in their model included the delivery volume factor, revenue factor, and the cost of transporting goods. Jong and Schonfeld [4] developed an objective function consisting of location and user costs for optimizing highway alignments.

A major drawback of existing approaches is that the objective functions typically do not account for uncertainties associated with decision variables that arise during different stages of the planning and implementation process. Uncertainties may involve various factors such as land-use data, topographical features, future demand, construction and maintenance costs, budget availability, implementation timelines, technological advances, and the potential for consensus among stakeholders. If these uncertainties are not integrated into the optimization process, the outcomes may be overly optimistic, rigid, or infeasible in practice.

***Corresponding author:** Manoj K. Jha, Department of Information Technology, University of Maryland Global Campus, USA. Email: manoj.jha@faculty.umgc.edu

Among the different types of uncertainty that can arise in infrastructure planning, one important category is information uncertainty. Information uncertainty refers to vagueness, subjectivity, or incompleteness in available input data or expert judgment. It is often present during early planning stages when exact numerical data are unavailable or when assessments are qualitative in nature—for example, when sites are described as having "low suitability" or "moderate cost." This type of uncertainty differs from aleatory uncertainty, which is associated with inherent randomness (e.g., weather variability), and epistemic uncertainty, which stems from a lack of knowledge and could potentially be reduced with more data. Information uncertainty, by contrast, is often inherently qualitative and cannot be adequately modeled using traditional probabilistic or statistical techniques.

To address numeric uncertainty, researchers have applied probabilistic approaches and Machine Learning (ML) methods [5]. These models can process large, structured datasets and represent uncertainties in quantitative terms. However, they are limited when dealing with uncertainties that are not easily quantifiable or when linguistic input (such as expert or stakeholder evaluations) is a critical component of the decision-making process. ML methods also require large, well-labeled datasets, which are often unavailable in the early stages of infrastructure development.

This paper presents a modeling framework that explicitly addresses qualitative information uncertainty by integrating fuzzy logic with linear programming. Fuzzy logic has been recognized as a preferred approach for solving problems that involve qualitative uncertainties [1, 6–12]. It enables decision-makers to represent vague or subjective data—such as environmental impact, constructability, or stakeholder preferences—using fuzzy sets and membership functions. These are then incorporated into the optimization process through fuzzy inference and rule-based reasoning. The fuzzy-augmented model allows infrastructure alignment decisions to reflect the ambiguous, often linguistic nature of available information, without requiring overly simplified numerical assumptions.

The novelty of the proposed approach lies in its ability to embed qualitative, linguistically defined uncertainties directly into a formal optimization process. Most conventional methods, including probabilistic models and ML approaches, assume that uncertainties can be numerically specified. In contrast, this method introduces a flexible and robust way to capture real-world decision ambiguity, offering a more realistic treatment of uncertainty in infrastructure location optimization tasks. This makes the approach particularly relevant during the conceptual and preliminary planning stages of infrastructure projects, where qualitative assessments dominate and data scarcity is common.

The goals of this research are to develop and demonstrate a fuzzy logic-based optimization framework that improves the reliability and practicality of infrastructure location optimization decisions under uncertain conditions. The scope includes defining appropriate fuzzy sets and membership functions for key qualitative decision variables, integrating them into a linear programming model, and evaluating the method through both artificial test scenarios and a real-world case study. The methodology is designed to be transparent, interpretable, and easily adaptable to varying project contexts. The paper contributes to the literature by:

- defining and formalizing information uncertainty in infrastructure planning,
- identifying the limitations of conventional models in handling such uncertainty,
- proposing a novel fuzzy logic-based optimization method tailored to linguistic and qualitative uncertainty, and
- demonstrating its practical application and advantages through comparative case studies.

The approach can be extended in future research by embedding fuzzy logic within ML architectures to further enhance adaptability in dynamic and data-scarce environments.

2. Literature Review

Optimizing infrastructure locations is quite complex since many conflicting factors and their relative weights need to be considered simultaneously, which makes it difficult to obtain an optimal result [12–14]. For example, Yin et al. [2] developed a time–space network-based model for transportation service optimization of the China Railway Express. The authors developed an optimization function as a difference between transportation revenue and transportation costs. Many decision variables, such as quantity of goods supplied by a specific source point and a specific source category, and agreed delivery volume, revenue factor, and cost of transporting goods were considered. Optimization models have been developed for locating transportation infrastructure, such as highways [4, 15], rail-lines and rail transit lines [16] as well as for locating fixed facilities, such as airports and warehouses [17]. While formulations for various cost functions were provided in these studies [18], uncertainties with respect to the decision-making process were not considered. Jha [19] developed a criteria-based decision support system for locating highways. In that work, various conflicting preferences of stakeholders were considered within an optimization framework. However, the uncertainties about information were not considered in the decision-making process.

Pang et al. [20] developed a stochastic route optimization method under dynamic ground risk uncertainties. The authors proposed a two-stage stochastic optimization method for unmanned autonomous system risk management. However, the likelihood of risk in selecting a route was handled using a probabilistic method which was not capable of handling uncertainties consisting of linguistic hedges, such as the inability to mathematically represent the instances of “low”, “medium”, or “high”. Other stochastic optimization methods based on a probability function also suffered the same weaknesses in handling uncertainty.

ML-based methods, including deep learning, have been used in recent works to handle uncertainty [5, 21]. However, those methods handle uncertainty using crisp and numerical data. They are unable to handle and map uncertainty information that is represented qualitatively.

Uncertainties can be attributed to, among other things, poor knowledge about the terrain and topographic features, climate change issues (e.g., hydraulic instabilities due to excessive rain, wildfires, and hurricanes), as well as decision-making and project delays due to scope changes. Fang and Zhu [22] use an active learning framework to address uncertainties in labeling of unstructured datasets. The framework is based on the concept of diverse density which is a machine learning approach. The diverse density defines the density of instances or situations, in terms of how many positive bags are within a region and how far the region is from the negative bags, to help predict whether an instance is positive or not. A bag is a cluster of densely populated points representing the instances or situations. Maron and Lozano-Perez [23] provide insights into diversity density functions in the context of ML and the concept of bagging applied to unstructured datasets. The active learning framework can refine the intermediate task durations before optimizing a critical path for the end-to-end execution of transportation and construction projects.

Fuzzy logic has been widely used to model uncertainty in various domains, such as water security assessment [24], gas lift design [25], and analysis of heterogenous data [26]. It has also been used in modeling and controlling nonlinear systems [27, 28].

Uncertainties about climate change have been recognized in a recent study in the hydrologic context [29]. Uncertainties due to measurement errors have primarily been handled using statistical and probability theories [30]. Tennøe et al. [31] developed a Python toolbox for uncertainty quantification and sensitivity analysis in computational neuroscience. They based the uncertainty analysis on polynomial chaos expansions which are more efficient than the Monte-Carlo based approaches. Uncertainties regarding railroad alignments have been analyzed with robust optimization methods [32, 33].

Fuzzy logic has been proposed in previous works to handle linguistic situations which are difficult to model mathematically, for optimization purposes. Some examples in handling uncertainty include “somewhat”, “may be”, and “about right”. Some earlier works on the foundation and applications of fuzzy logic can be found in Zadeh [34] as well as in Yager and Filev [35]. In other works [36, 37], methods using fuzzy logic to handle uncertain information were proposed. Zimmermann [38] developed the concept of fuzzy set theory and offered some applications. Kikuchi and Milkovic [39] applied fuzzy logic to preprocess observed traffic data for consistency. Kikuchi et al. [40] examined some methods, including fuzzy logic to adjust observed traffic volumes on a network. Kikuchi and Jha [41] developed a method for reconciling the values of parameters using fuzzy logic. They developed a fuzzy-logic approach to address uncertainty associated with decision-making considering diverse viewpoints of the decision makers. A fuzzy set was used to represent the notion of desire, and a fuzzy optimization approach based on linear programming was used to obtain the optimal solution. The approach is extended in this paper to deal with qualitative uncertainties in locating infrastructures.

3. Research Methodology

Route optimization for highways usually involves finding the 3-dimensional points of intersections (PIs) and fitting appropriate curves to connect the tangents between PIs [42]. A similar concept is followed for optimizing railway routes, except that different transition curves are used.

For large-scale analysis, some fixed facilities can be modelled as discrete points [17]. The objective is to optimally locate a facility (of desired shape) subject to some constraints. For example, a sports arena can be considered a fixed facility whose location will be largely affected by its accessibility to sports fans. An airport can also be considered as a fixed facility as a point for a macroscopic analysis, whose location will be driven not only by the accessibility to the travelers but also by keeping a buffer distance from the residential neighborhoods to minimize the noise impacts on the neighborhoods due to flights. Many other factors affect airport location including topography, land use, land availability and costs, wind patterns, obstructions to flights, relations to other airports (especially in market shares and flight paths), and various environmental factors.

Figure 1 shows the concept of locating highways, rail lines, and fixed facilities (e.g., sports arenas and airports).

3.1. Uncertainties with respect to terrain, topographic features, and property values

While terrain, topographic features, and property values are available from GIS maps, the extent of their relative impacts cannot be ascertained, and a trade-off analysis may be needed. For example, while it is preferable to minimize the total cost of a highway, it may be desirable or necessary to also minimize the wetland or sensitive watershed impacts at the expense of increasing the total cost to some desirable extent [19]. Figure 2 shows a conceptual GIS map of soil characteristics. It can be seen that if a highway is to be constructed between the bottom left corner to top right corner, it is impossible to avoid average or poor soil characteristics.

Uncertainties may arise due to the inherent variability in the computational parameters used for solving a mathematical optimization problem. This may include certain algorithm-specific tuning parameters in genetic algorithms or swarm intelligence that have been used for infrastructure location optimization problems.

There could be many situations with uncertainties regarding terrain and topographic features. For example, there could be missing information on soil characteristics or property values. Previous models assumed that these values are available. Construction cost is affected by soil characteristics and right-of-way cost is affected by property values. In this paper, we present a generalized methodology for handling qualitative uncertainties in the route optimization process using fuzzy logic.

Figure 1
Conceptual illustration of highways, rail lines, and fixed facilities

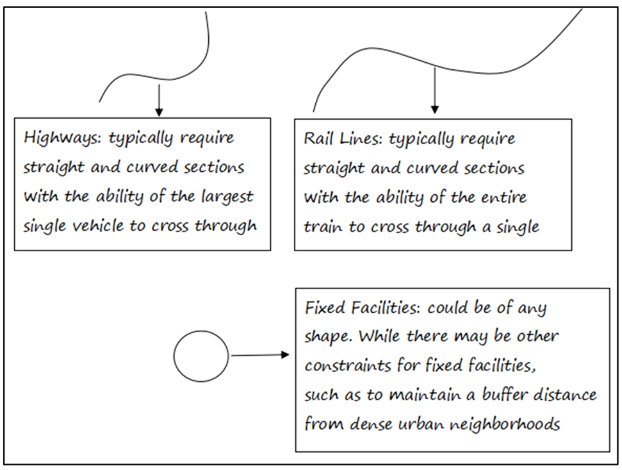
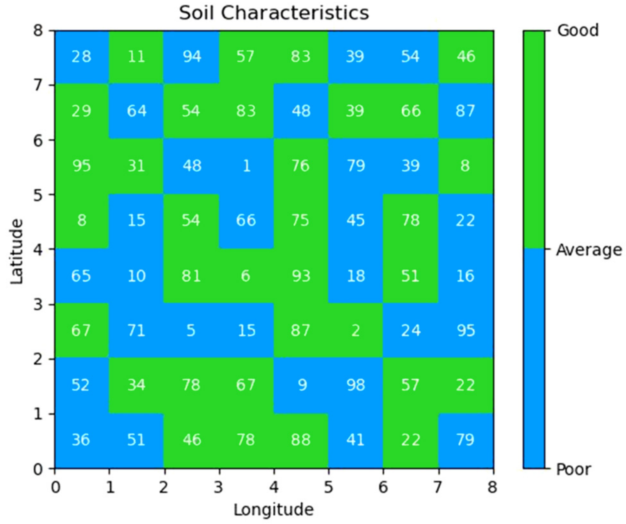


Figure 2
A conceptual GIS map of soil characteristics (soil characteristics are represented on a scale of 1–100)



3.2. Modeling uncertainty with fuzzy logic

As explained in Section 2, fuzzy logic is especially suited to handle uncertainty when linguistic variables, such as “low”, “medium”, and “high” cannot be mathematically represented. Such linguistic uncertainties can be encountered at many stages of infrastructure optimization problems. For example, the future values of properties impacted by infrastructure construction can be rated as low, medium, or high. The terrain can be susceptible to future climate changes, such as flood, earthquake, or wetland formation. The likelihood of such changes could only be rated in different geographic domains as low, medium, or high. In Figure 2, while indices of soil characteristics in various cells are available, their initial values are fixed with the option of readjustment within an allowable range. In the same overlapping search space, there will be values for terrain elevations, hydrological characteristics, land (or property values), and other topographical features. These values could be related by a certain set of relations. For example, high elevation areas should be avoided if less expensive alternate terrain is available for road or rail construction. The problem is to find the optimal values that satisfy the relations. A trade-off among the values of these attributes will be needed in order to find an optimized solution.

The values given initially and to be optimized may be observed (or measured) values, estimated values, or desired values, and uncertainty is associated with each value with respect to its acceptability and tolerance. The relations that the values must satisfy may include a “hard” equality or inequality, or a “soft” equality or inequality. The relations may (1) satisfy a set of equality (exact or approximate) relations; or (2) satisfy a set of approximate inequality relations (e.g., “much greater than”).

The following situations may require analysts to adjust or optimize the values:

- When the values are used as inputs to another set of models, or inputs to a computer package.
- When analysts wish to change one or more of the already agreed upon design values to reflect their or the client’s desire for design modification.
- When the data points are known, the parameters of an equation that best fit the data points must be found.

The method must satisfy the following general requirements:

- The final adjusted value should be close to the initially presented value as possible.
- The method should allow analysts the opportunity to interact with those who are directly involved in deciding the acceptability of the values and allow dialogues and readjustment.
- The method should be able to measure the level of satisfaction (or acceptability) of the adjusted values.

3.2.1. Traditional approaches and their limits

Traditionally, this type of problem has been handled in two ways: (1) by a trial-and- error method, in which an approximate band of tolerable deviation is assumed for each value; within the band, each value is adjusted by trial-and-error (sometimes involving negotiation); and (2) by the least-squares method, in which the squared sum of the differences between the initial values and the adjusted values is minimized. While ML methods, such as deep learning, have been used to model uncertainty (generally referred to as "noise"), those methods are not capable of handling linguistic hedges.

In summary, existing approaches cannot accommodate in a systematic manner, the soft notions of "desire" and "tolerance" in the mind of an analyst or concerned party. Such a method is particularly useful for the early stage of planning or design in which many parameters of decision and design are uncertain to the analyst or to the concerned parties.

3.2.2. The proposed method

While the study focuses on fuzzy logic-based optimization for handling uncertainty, it acknowledges the importance of alternative uncertainty-handling techniques such as probabilistic models and ML approaches. A comparative analysis is shown in Table 1 to justify the methodological choice and clarify its domain of relevance. Bayesian networks assume the availability of conditional probabilities between variables, which are often unavailable or subjective during early-stage infrastructure planning.

Deep learning methods require extensive labeled data and perform poorly when input features are ambiguous or described using natural language.

Fuzzy logic, in contrast, allows direct modeling of vagueness and stakeholder perceptions without the need for empirical frequency data, making it a more realistic choice for the early, data-sparse stages of infrastructure decision- making.

In the current research, fuzzy logic was chosen because it uniquely addresses information uncertainty expressed in linguistic terms, a frequent occurrence in early infrastructure planning stages. Unlike probabilistic methods that rely on numerical distributions, fuzzy systems can accommodate subjective human judgment (e.g., "moderate cost," "low consensus") and integrate it directly into an optimization framework. Furthermore, the resulting model can be solved using standard linear programming techniques, making it both interpretable and computationally efficient.

We propose a fuzzy logic-based optimization process as shown in the flow-diagram shown in Figure 3. As illustrated earlier, fuzzy logic was chosen over probabilistic methods, like Bayesian networks or deep learning because the problem domain primarily involves qualitative uncertainty, rather than randomness or statistical variation. While probabilistic approaches require large, well-structured datasets to model distributions, fuzzy logic is uniquely suited to handle imprecise, linguistically expressed knowledge, such as "low soil stability" or "moderate stakeholder resistance," where numeric probabilities are difficult or impossible to obtain.

Linguistic descriptors such as "low," "medium," and "high" were transformed into numerical values using triangular membership functions, a common approach in fuzzy systems. Each linguistic label was assigned a fuzzy number with a peak (representing maximum membership) and a base that defines the acceptable range:

For instance, "medium" property cost might have a peak at 200 and extend from 120 to 280. The degree of satisfaction or compatibility

Figure 3
Flow-diagram showing the fuzzy logic-based optimization process

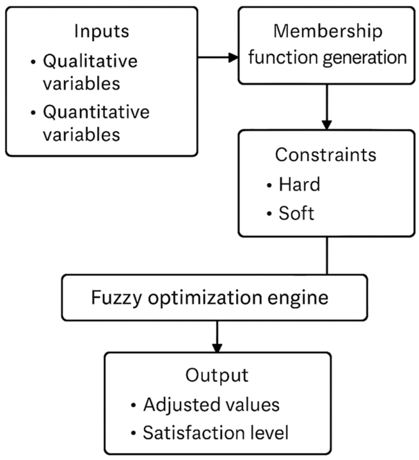


Table 1
Comparison of fuzzy logic with other methods for handling uncertainty

Approach	Key feature	Limitations
Fuzzy logic (this study)	Models qualitative uncertainty using linguistic variables; enables integration of vague, imprecise knowledge via membership functions	Limited flexibility in learning from data; requires expert-defined functions
Bayesian networks	Captures conditional probabilities and belief updates; suitable for probabilistic causal inference	Requires known prior probabilities; difficult to apply when data is sparse or qualitative
Deep learning/ML	Learns from large datasets; can model nonlinear relationships and uncertainty quantitatively	Needs large, labeled datasets; poorly handles ambiguity or subjective inputs
Robust optimization	Focuses on worst-case scenarios; ensures solution feasibility under bounded uncertainty	Often pessimistic; lacks flexibility to model nuanced linguistic hedges
Stochastic programming	Integrates randomness through probability distributions; useful for cost–benefit trade-offs	Requires well-defined distributions; unsuitable for purely qualitative inputs

is then computed using piecewise-linear functions depending on how far an actual value deviates from this range.

These membership functions enable smooth, interpretable mapping from qualitative expressions to a numerical framework suitable for optimization. Only triangular membership functions were used in this study. This choice was made for the following reasons:

- Interpretability: Triangular functions are intuitive and easy to communicate to stakeholders.
- Computational Simplicity: They yield linear constraints in the fuzzy optimization model, allowing the use of efficient linear programming solvers.

Previous studies [41, 43] have demonstrated the adequacy of triangular functions in similar infrastructure problems.

Although alternative functions (e.g., trapezoidal, Gaussian) can offer smoother transitions, the simplicity and transparency of the triangular form were considered appropriate for this exploratory application. Future studies could compare different shapes to explore sensitivity and robustness.

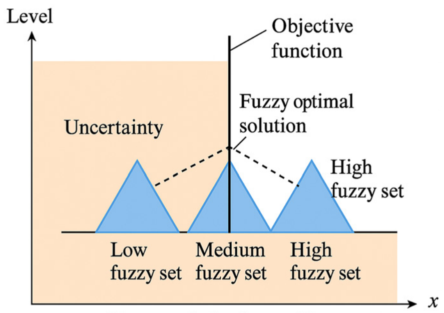
The initial value harbors the notion of desire, approximation, or tolerance. A fuzzy number is introduced to represent these notions [43]. The membership function of a fuzzy set *A, A(x)*, characterizes the degree of "compatibility," or "satisfaction" as a function of x such that the value of $A(x)$ is between 0 and 1. $A(x) = 0$ representees the least degree of compatibility or satisfaction whereas $A(x) = 1$ represents the highest degree of compatibility or satisfaction.

The relations that the values must satisfy may be a combination of equations and inequalities, both in the exact and the inexact sense. The "hard" equality or inequality relations are usually associated with the physical principles, such as the conservation of flow in which the total input must be equal to the total output. The "soft" equality and inequality relations arise when analysts are unsure about the exact relation. An approximate relation, such as "x is 'much' greater than y," and "x and y are 'approximately' equal," are examples of fuzzy relations.

A flow-diagram depicting the fuzzy logic-based optimization process used in this study is shown in Figure 3. A fuzzy optimization model is developed (see Figure 4 for a diagrammatic representation of the fuzzy optimization model) with the following assumptions and calculations:

1. Input Uncertainty as Fuzzy Sets: Each input variable (e.g., soil condition, stakeholder support) is assumed to have uncertainty best described by a triangular membership function, capturing its possible deviation from an initially preferred value.
2. Tolerances Defined by Experts: The bounds (±) for each variable are defined by experts or decision-makers and reflect acceptable deviation ranges from nominal values.

Figure 4
Diagrammatic representation of the fuzzy optimization problem



3. Hard vs. Soft Constraints:
   a. Hard constraints (e.g., total length or cost) must be satisfied exactly (e.g., x1 + x2 + x3 + x4 = z1).
   b. Soft constraints (e.g., "x2 is approximately 2x3") are expressed using fuzzy logic, allowing approximate satisfaction with decreasing degrees of confidence as deviations grow.
   c. Satisfaction Level Maximization: The goal is to find variable values that maximize the minimum satisfaction level (h) across all fuzzy constraints and initial preferences (Bellman-Zadeh principle).
   d. Each variable's deviation from its initial value is modeled using a triangular membership function.
   e. These functions are piecewise-linear and yield values from 0 (fully unsatisfactory) to 1 (fully satisfactory).
   f. The fuzzy optimization model is transformed into a linear programming (LP) problem by expressing these membership functions and fuzzy inequalities as linear constraints.

In the optimization model, the individual values of the variables x should be close to their initial values, but they are "pulled" from them in order to satisfy the required relations. Thus, the solution is found at the point of compromise or trade-off. Such an optimization process follows the Bellman-Zadeh principle; that is, the solutions, or the decision set, lies in the confluence of the goal and the constraints: **D** = **C** ∩ **G**, where **D** is the solution set, **C** is the constraint set, and **G** being the goal set (Bellman and Zadeh 1970). The solution set, **D**, is a set of values that satisfies both **C** and **G**. The optimization objective is to select a value that "best" satisfies both **C** and **G**.

The process can be formalized as:

$$\mathbf{D}(x^*)\mathrm{Max}\{\mathrm{Min}[\mathbf{C}(x), \mathbf{G}(x)]\},\ \text{for } x \in \mathbf{X}, \tag{1}$$

where x^* is the best solution. $\mathbf{D}(x^*)$ indicates the degree of overall satisfaction with the best solution, and $\mathbf{C}(x)$ and $\mathbf{G}(x)$ indicate the degree to which x satisfies the constraint and the goal, respectively. Finding the value of x^* in Equation (1) is equivalent to solving the following optimization problem, where the unknowns are x and h. The quantity, Z is the objective function to be maximized.

$$\mathrm{Max}\ Z \tag{2}$$

subject to:

$$\mathbf{C}(x) \geq h \tag{2a}$$

$$\mathbf{G}(x) \geq h \tag{2b}$$

$$x \geq 0 \tag{2c}$$

$$h \geq 0 \tag{2d}$$

Assume the following:

- $w_1, w_2, \ldots$ are the initial values, which the analyst (or stakeholder) wishes to adjust.
- $W_1(x_1), W_2(x_2), \ldots$ are the membership functions of the fuzzy numbers derived from the initial value.
- $f_{R1}(x_1, x_2, x_3, \ldots), f_{R2}(x_1, x_2, x_3, \ldots)$ are the "hard" relations and requirements that the values must satisfy.
- $F_{R1}(x_1, x_2, x_3 \ldots), F_{R2}(x_1, x_2, x_3 \ldots)$ are the membership functions of the "soft (fuzzy)" relations that the adjusted values must satisfy.
- Z is the objective function (to be maximized).

The general formulation for the optimization process is:

$$\text{Max } Z \tag{3}$$

subject to:
For initial values:

$$W_1(x_1) \geq h,\ W_2(x_2) \geq h,\ \ldots \tag{4}$$

For hard relations:

$$f_{R1}(x_1, x_2, \ldots) = 0,\ f_{R2}(x_1, x_2, x_3, \ldots) =, >, \text{or } < 0, \ldots \tag{5}$$

For soft relations:

$$F_{R1}(x_1, x_2, \ldots) \ldots h,\ F_{R2}(x_1, x_2, \ldots) \geq \mathrm{h}, \ldots . \tag{6}$$

$$x_1, x_2, \ldots \geq 0, h \geq 0 \tag{7}$$

As an example, consider the following five values of the decision parameters in an infrastructure location optimization problem that need to be adjusted: X_1, X_2, X_3, X_4, and Z_1. This example, while artificial in nature, mimics real-world situations of uncertainty. We consider these five variables in the optimization process. The example can easily be applied to a real-world scenario, for example, in assessing the soil characteristics, availability of skilled labor, availability of materials, variation of future climate change parameters, and susceptibility of decision-makers in reaching a consensus in deciding on the final location for the infrastructure construction. Such an infrastructure could be a metro line (e.g., the red line under consideration in the City of Baltimore, Maryland, USA). The method will yield an optimal solution with any real-world variables.

In an optimization sense, these decision parameters are similar to being the decision variables whose values must be adjusted in order to reach the optimal solution. Their initial values are 450 ($=x_1^0$), 200 ($=x_2^0$), 120 ($=x_3^0$), 160 ($=x_4^0$), and 1,100 ($=z_1^0$), respectively. The tolerance for deviation from the value is for X_1, 450 ± 100; for X_2, 200 ± 80; for X_3, 120 ± 50; for X_4, 160 ± 90, and for Z_1, 1100 ± 100. The corresponding membership functions are shown in the upper layer of Figure 5.

For the above problem, the following four relations are considered:

$$\text{Relation 1}:\ x_1 + x_2 + x_3 + x_4 = z_1, \tag{8}$$

$$\begin{aligned}&\text{Relation 2}:\ x_1 \text{ is somewhat greater than } (x_2 + x_3)\\ &\text{with the acceptable difference } \leq 50\end{aligned} \tag{9}$$

$$\begin{aligned}&\text{Relation 3}:\ x_3 \approx x_4,\ (x_3 \text{ is approximately equal to } x_4)\\ &\text{with the acceptable difference } \leq 100\end{aligned} \tag{10}$$

$$\begin{aligned}&\text{Relation 4}:\ x_2 \approx 2x_3,\ (x_2 \text{ is approximately twice of } x_3)\\ &\text{with the acceptable difference } \leq 50\end{aligned} \tag{11}$$

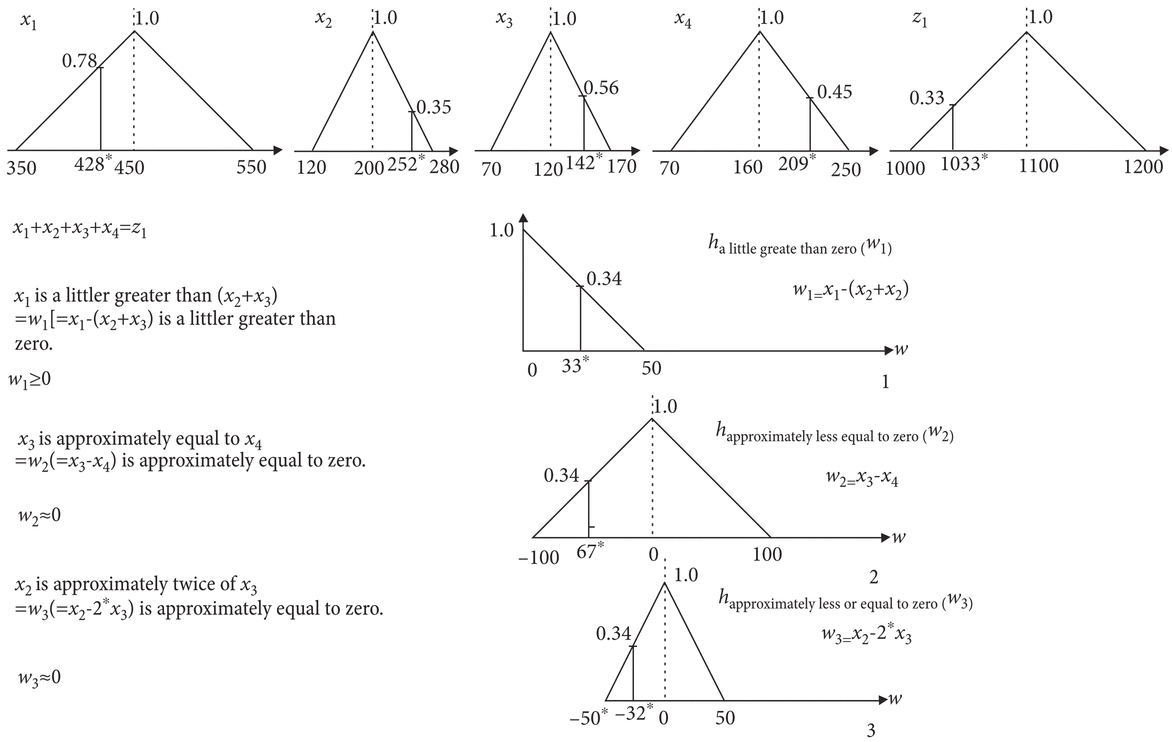
The objective is to maximize the value of Z according to the formulation in Equation (3). The membership functions for parameter values, x_1, x_2, x_3, x_4, and Z_1 are defined as a triangular function. For the right- and left-hand side of the triangle, the corresponding equations are $h_{xi}+(x_i)$ and $h_{xi}-(x_i)$, respectively.

$$\mathrm{h_{xi}} + (\mathrm{x_i}) = -\frac{1}{RXi(or\ RZ_1)}\left(\mathrm{x_i} - \mathrm{x_i^0}\right) + 1 \geq \mathrm{h} \tag{12}$$

$$\mathrm{h_{xi}} - (\mathrm{x_i}) = \frac{1}{L{:}Xi(or\ LZ_1)}\left(\mathrm{x_i} - \mathrm{x_i^0}\right) + 1 \geq \mathrm{h},\ \text{for } \mathrm{i} = 1, ..4 \tag{13}$$

where $x_1^0 = 450$, $x_2^0 = 200$, $x_3^0 = 120$, $x_4^0 = 160$, $z_1^0 = 1100$ and $RX_1 = LX_1 = 50$, $RX_2 = LX_2 = 40$, $RX_3 = LX_3 = 25$, $RX_4 = LX_4 = 40$, $RZ_1 = LZ_1 = 50$.

Figure 5
Illustrations for solutions of the example problem



For the constraints based on the required relations:

$$\text{Relation 1}: \ x_1 + x_2 + x_3 + x_4 = z_1 \tag{14}$$

$$\text{Relation 2}: \ \text{h}_{\text{somewhat greater than zero}}(\text{w}_1) = -\frac{1}{RW_1}w_1 + 1 \geq h \text{ where } w_1 = x_1 - (x_2 + x_3), \text{ and } RW_1 = 50 \tag{15}$$

$$\text{Relation 3}: \ \text{h}_{\text{aproximately zero+}}(\text{w}_3) = \frac{-1}{RW_3}w_3 + 1 \geq \text{h}, \ \text{h}_{\text{aproximately zero}-}(\text{w}_3) = \frac{1}{LW_3}w_3 + 1 \geq h \tag{16}$$

$$\text{where } w_3 = x_3 - x_4, \text{ and } RW_3 = LW_3 = 100 \tag{17}$$

$$\text{Relation 4}: \ \text{h}_{\text{near equal to zero+}}(\text{w}_4) = \frac{-1}{RW_4}w_4 + 1 \geq \text{h}, \ \text{h}_{\text{near equal to zero}-}(\text{w}_4) = \frac{1}{LW_4}w_4 + 1 \geq h \tag{18}$$

$$\text{where } w_4 = x_2 - 2x_3, \text{ and } RW_4 = LW_4 = 50 \tag{19}$$

The above formulation can be reformulated in the linear programming format as:

$$\begin{array}{lll} \text{Max } Z & & \\ \text{subject to} & & \\ \text{parameter} & \text{Right hand side} & \text{Left hand side} \\ x_1 & -100h + x_1 \geq 350 & 100h + x_1 \leq 550 \end{array} \tag{20}$$

$$x_2 \quad -80h + x_2 \geq 120 \quad 80h + x_2 \leq 280 \tag{21}$$

$$x_3 \quad -50h + x_3 \geq 70 \quad 50h + x_3 \leq 170 \tag{22}$$

$$x_4 \quad -90h + x_4 \geq 70 \quad 90h + x_4 \leq 250 \tag{23}$$

$$Z_1 \quad -100h + z_1 \geq 1000 \quad 100h + z \leq 1200 \tag{24}$$

The constraints related to the relations can be expressed as follows:

$$\text{Relation 1}: x_1 + x_2 + x_3 + x_4 - z_1 = 0 \tag{25}$$

Relation 2 :

$$\text{for the left-hand side, } x_1 - x_2 - x_3 \geq 0 \tag{26}$$

$$\text{for the right-hand side, } x_1 - x_2 - x_3 + 50h \leq 50 \tag{27}$$

Relation 3 :

$$\text{for the left-hand side, } x_3 - x_4 + 100h \leq 100 \tag{28}$$

$$\text{For the right-hand side, } x_3 - x_4 + 100h \geq -100 \tag{29}$$

Relation 4 :

$$\text{for the left-hand side, } x_2 - 2x_3 + 50h \leq 50 \tag{30}$$

$$\text{for the right-hand side, } x_2 - 2x_3 - 50h \geq -50 \tag{31}$$

$$x_1, \ x_2, \ x_3, \ x_4, \ z_1, \text{ and } h \geq 0 \tag{32}$$

4. Solution Procedure

Applying linear programming, the optimized values, the satisfaction level relative to the initial value, and also the satisfaction level relative to the required relations, are obtained as shown in Tables 2 and 3.

5. Results and Discussion

The above optimization problem may or may not have a feasible solution. When a feasible solution is found, the value of Z indicates the degree to which all the constraints are satisfied. The satisfaction of individual constraint relations can be found by inserting the derived xi's into the respective membership functions, Equations (5) and (6). If the analysts require that the solution must yield at least a given level of satisfaction, α, then an additional constraint, h ≥ α, can be added. The optimized result is shown in Figure 5.

If the solution does not exist, then a set of values that satisfies the initial desire for the parameters and the required relations do not exist. A possible action is to revise some (or all) of the initial values of the fuzzy variables, or to change the corresponding membership functions of the fuzzy variables, and/or fuzzy relations by widening or shifting their bases. The readjustment of the membership functions may involve further investigating desired tolerance intervals of the decision variables.

The example shown here can be used for handling uncertainty in soil characteristics, topographical features, future demand patterns, or conflicting viewpoints of various stakeholders, including politicians involved in the decision process of building infrastructures. For example, before building a metro line, locating stations, or expanding a major highway, townhall meetings are generally conducted to accommodate the viewpoints of as many stakeholders as possible. These viewpoints are often conflicting in nature. The mathematical method presented here can serve as a useful tool for handling such uncertainty. The inherent benefit of fuzzy logic is its ability to convert linguistic hedges (such as, good, not so good, ok, or kind of) into numerical terms.

The proposed fuzzy optimization method has strong real-world relevance, especially during the early stages of infrastructure planning, where data scarcity and stakeholder ambiguity are common. Here is how policymakers and planners can benefit:

1. Decision Support Under Uncertainty

Policymakers often face incomplete or conflicting information (e.g., environmental risk, stakeholder resistance). The fuzzy logic framework allows them to incorporate qualitative input (e.g., "high risk," "moderate support") directly into the planning model, yielding solutions that reflect practical realities.

Table 2
Optimal parameters

Parameter	Initial value (range)	Adjusted value	Satisfaction (*h*-value)
X_1	450 (±100)	428	0.78
X_2	200 (±80)	252	0.35
X_3	120 (±50)	142	0.56
X_4	160 (±90)	209	0.45
Z_1	1100 (±100)	1031	0.33* (min satisfaction)

*: Max–Min solution of *h*.

Table 3
Optimal relations

Relations	With the adjusted values	Satisfaction *h*–value
1. $x_1 + x_2 + x_3 + x_4 = z_1$	428 + 252 + 142 + 209 = 1031	1.0
2 x_1 is close somewhat greater or equal to $(x_2 + x_3)$	428 is close but somewhat greater or equal to [252 + 142] (=396)	0.34
3 $x_3 \approx x_4$	142 ≈ 209	0.34
4. $x_2 \approx 2x_3$	252 ≈ 2 x 142 (=284)	0.34

2. Stakeholder Engagement
   By modeling stakeholder inputs as linguistic variables, the method helps planners reconcile diverse opinions (e.g., environmentalists, engineers, community members) into a unified, mathematically tractable optimization framework. This promotes transparent and inclusive decision-making.

3. Policy Trade-off Analysis
   The model supports trade-offs among competing objectives (e.g., cost vs. environmental impact), allowing planners to evaluate multiple scenarios while ensuring solutions remain within acceptable tolerance bands.

4. Adaptability to Real Data
   Once qualitative assessments are converted to fuzzy numbers, they can be updated as more precise data becomes available, enabling the model to evolve alongside project development.

5. Integration into Existing Tools
   Fuzzy optimization can be integrated into GIS-based tools, transportation planning platforms, or used in tandem with rule-based or simulation models already in use by agencies.

6. Conclusions and Future Extensions

We studied the infrastructure location optimization problem considering uncertainties which are qualitative in nature. We reviewed Machine Learning (ML) literature in the context of handling uncertainty. Two recent papers are especially relevant: Russel and Reale [5] and Loquercio et al. [44]. We found that ML-based methods can only handle quantitative uncertainties referred to as epistemic uncertainty and aleatoric uncertainty. Epistemic uncertainty reflects uncertainty in the model parameters used to perform prediction in machine learning problems. Aleatoric uncertainty reflects the noise inherent to the data.

We advocated the use of fuzzy logic to handle uncertainties consisting of linguistic hedges (e.g., impact of soil characteristics being labeled as low, medium, or high) which cannot be formulated by existing ML techniques due to inherent limitations of those techniques in handlining linguistic hedges. The integration of the proposed fuzzy logic approach with ML-based deep learning technique should be pursued in future works.

The proposition of fuzzy logic presented here is an effective tool for handling uncertainty because it takes a range of possible values into account in reaching an optimized solution. In many actual situations the tolerance of decision variables resides within specified bounds, and a best traded-off optimized solution is desired. A traditional deterministic model where each input parameter has a chosen fix value, gives a single output of the model. An uncertainty quantification of the model accounts for the distributions of the input parameters, and the output of the model becomes a range of possible values.

Future work may include exploring ML models for handling uncertainty. ML concepts have recently become very popular for performing prescriptive, descriptive, and predictive analytics. Predictive analytics is also known as a classification problem in which data-driven decision-making is performed to predict a single independent variable (or multiple independent variables) dependent on a set of dependent variables. The process mimics the structure of a typical mathematical optimization problem, except that ML algorithms are purely data driven. If the data are incomplete or information (data) on stronger predictor variables are missing (or incomplete), predictions are less accurate. Nevertheless, ML concepts can be exploited to handle uncertainties associated with the route optimization process when integrated with a fuzzy logic approach presented here.

Some of the ML methods applied for handling uncertainty include neural networks, deep ensembles, and deep learning. While each of these methods has its advantages and disadvantages, they can be integrated with the proposed fuzzy logic approach in future works. We intend to explore these methods in future studies.

The study has several limitations. First, the numerical example used is synthetic and illustrative, not drawn from a real-world case. Second, the model relies on expert-defined membership functions and tolerance thresholds, which may introduce subjectivity. Third, the current framework does not include automated learning of fuzzy parameters from data, which may limit scalability. Lastly, field validation and testing in large-scale infrastructure projects remain to be conducted.

These limitations offer opportunities for future work, including the integration of fuzzy logic with data-driven machine learning models, refinement of membership function estimation methods, and application to real-world transportation or infrastructure datasets.

Acknowledgement

Some of the work presented here was jointly carried out by the first author and late Professor Shinya Kikuchi of Virgina Tech. That work has been significantly expanded here by the author for the route optimization problem and dealing with uncertainty. No funding was provided for this work.

The author confirms contribution to the paper as follows: study conception and design: M. Jha; developing the methodology and analyzing the data: M. Jha; draft manuscript preparation: M. Jha; editing and preparing the final manuscript: M. Jha. The author reviewed the results and approved the final version of the manuscript.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

Manoj K. Jha is an Editorial Board Member for *Journal of Data Science and Intelligent Systems* and was not involved in the editorial review or the decision to publish this article. The author declares that he has no conflicts of interest to this work.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Author Contribution Statement

Manoj K. Jha: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration.

References

[1] Saldanha-da-Gama, F. (2022). Facility location in logistics and transportation: An enduring relationship. Transportation Research Part E: Logistics and Transportation Review, 166, 102903. https://doi.org/10.1016/j.tre.2022.102903

[2] Yin, W., Hu, W., Yan, X., Peng, B., & Yang, X. (2024). A time-space network-based model for transportation service optimization of China Railway Express. High-speed Railway, 2(3), 153–163.

[3] Torres, M. N., Fontecha, J. E., Walteros, J. L., Zhu, Z., Ahmed, Z., Rodríguez, J. P., & Rabideau, A. J. (2021). City-scale optimal

location planning of Green Infrastructure using piece-wise linear interpolation and exact optimization methods. Journal of Hydrology, 601, 126540.

[4] Jong, J. C., & Schonfeld, P. (2003). An evolutionary model for simultaneously optimizing three-dimensional highway alignments. Transportation Research Part B: Methodological, 37(2), 107–128.

[5] Russell, R. L., & Reale, C. (2021). Multivariate uncertainty in deep learning. IEEE Transactions on Neural Networks and Learning Systems, 33(12), 7937–7943.

[6] Miah, M. M., AlArjani, A., Rashid, A., Khan, A. R., Uddin, M. S., & Attia, E. A. (2023). Multi-objective optimization to the transportation problem considering non-linear fuzzy membership functions. AIMS MATHEMATICS, 8(5), 10397–10419. https://doi.org/10.3934/math.2023527

[7] Rishabh, R., & Das, K. N. (2025). A fusion of decomposed fuzzy based decision-making and metaheuristic optimization system for sustainable planning of urban transport. Knowledge-Based Systems, 113823. https://doi.org/10.1016/j.knosys.2025.113823

[8] Eshuis, R., Firat, M., & Kaymak, U. (2021). Modeling uncertainty in declarative artifact-centric process models using fuzzy logic. Information Sciences, *579*, 845–862. https://doi.org/10.1016/j.ins.2021.07.075

[9] Eti, S., Yüksel, S., Dinçer, H., Pamucar, D., Deveci, M., & Olaru, G. O. (2024). A machine learning and fuzzy logic model for optimizing digital transformation in renewable energy: Insights into industrial information integration. *Journal of Industrial Information Integration*, *42*, 100734. https://doi.org/10.1016/j.jii.2024.100734

[10] Zhang, M. Y., Yang, S. C., Feng, X. J., Chen, Y. Y., Lu, J. Y., & Cao, Y. G. (2022). Route planning for autonomous driving based on traffic information via multi-objective optimization. *Applied Sciences*, *12*(22), 11817. https://doi.org/10.3390/app122211817

[11] Manna, T., & Anitha, A. (2023). Precipitation prediction by integrating rough set on fuzzy approximation space with deep learning techniques. *Applied Soft Computing*, *139*, 110253. https://doi.org/10.1016/j.asoc.2023.110253

[12] Chen, X., Dong, C., Tao, S., Peng, Q., & Liu, J. (2024). Integrated Transit-Oriented Development (TOD) with suburban rail network design problem for maximizing profits. *Transportation Letters*, *16*(10), 1318–1337. https://doi.org/10.1080/19427867.2024.2304999

[13] Gao, Y., Zhang, T., Zhu, C., Yang, S., Schonfeld, P., Zou, K., ... & He, Q. (2024). Biobjective optimization for railway alignment fine-grained designs with parallel existing railways. *Computer-Aided Civil and Infrastructure Engineering*, *39*(3), 438–457.

[14] Yao, Y., Li, P., Mo, P., D'Ariano, A., & Appolloni, A. (2024). Bi-objective optimization of timetable and rolling stock schedule for an urban rail passenger and freight line. *Computers & Industrial Engineering*, *194*, 110394. https://doi.org/10.1016/j.cie.2024.110394

[15] Kang, M. W., Jha, M. K., & Schonfeld, P. (2012). Applicability of highway alignment optimization models. *Transportation Research Part C: Emerging Technologies*, *21*(1), 257–286. https://doi.org/10.1016/j.trc.2011.09.006

[16] Lai, X., & Schonfeld, P. (2016). Concurrent optimization of rail transit alignments and station locations. *Urban Rail Transit*, *2*(1), 1–15.

[17] Jha, M. K. (2002). Using a GIS and genetic algorithms to select optimal facility locations. In *Proceedings of the 6th World MultiConference of Systemics, Cybernetics, and Informatics*, *XVII*, 130–135.

[18] Jong, J. C., & Schonfeld, P. (1999). Cost functions for optimizing highway alignments. *Transportation Research Record*, *1659*(1), 58–67.

[19] Jha, M. K. (2003). Criteria-based decision support system for selecting highway alignments. *Journal of Transportation Engineering*, *129*(1), 33–41.

[20] Pang, B., Hu, X., Dai, W., & Low, K. H. (2024). Stochastic route optimization under dynamic ground risk uncertainties for safe drone delivery operations. *Transportation Research Part E: Logistics and Transportation Review*, *192*, 103717.

[21] Loquercio, A., Segu, M., & Scaramuzza, D. (2020). A general framework for uncertainty estimation in deep learning. *IEEE Robotics and Automation Letters*, *5*(2), 3153–3160. https://doi.org/10.1109/LRA.2020.2974682

[22] Fang, M., & Zhu, X. (2014). Active learning with uncertain labeling knowledge. *Pattern Recognition Letters*, *43*, 98–108.

[23] Maron, O., & Lozano-Pérez, T. (1997). A framework for multiple-instance learning. *Advances in Neural Information Processing Systems*, *10*.

[24] Yuan, L., Zhou, Z., He, W., Wu, X., Degefu, D. M., Cheng, J., ... & Ramsey, T. S. (2025). A fuzzy logic approach within the DPSIR framework to address the inherent uncertainty and complexity of water security assessments. *Ecological Indicators*, *170*, 112984. https://doi.org/10.1016/j.ecolind.2024.112984

[25] Ohia, N. P., Paul, C., Asolo, E., Adewa, T. A., Chukwu, C. F., Ubi, P. A., ... & Nnaji, D. U. (2025). Artificial intelligence-driven fuzzy logic approach for optimal well selection in gas lift design: A brown field case study. *Results in Engineering*, *25*, 103927. https://doi.org/10.1016/j.rineng.2025.103927

[26] Yu, K., Wang, D., Yu, Z., & Li, Z. (2025). Design of granular fuzzy relation models for heterogeneous data. *Expert Systems with Applications*, *269*, 126474. https://doi.org/10.1016/j.eswa.2025.126474

[27] Hoshino, Y., Rathnayake, N., Dang, T.L., & Rathnayake, U. (2024). Cascaded-ANFIS and its successful real-world applications, In P. Balasubramaniam & N. Ramesh Babu (Eds.), Fuzzy logic controllers and applications. *IntechOpen*. https://doi.org/10.5772/intechopen.1006491

[28] Rathnayake, N., Dang, T. L., & Hoshino, Y. (2021). A novel optimization algorithm: Cascaded adaptive neuro-fuzzy inference system. *International Journal of Fuzzy Systems*, *23*(7), 1955–1971. https://doi.org/10.1007/s40815-021-01076-z

[29] Zeng, Q., Zhao, Q., Luo, Y. T., Ma, S. G., Kang, Y., Li, Y. Q., ... & Xu, C. Y. (2024). Uncertainty estimation of hydrological modelling using gridded precipitation as model inputs in the Gandaki River Basin. *Journal of Hydrology: Regional Studies*, *53*, 101825.

[30] Yu, Q., Simões, J. T., & Muttoni, A. (2023). Model uncertainties and partial safety factors of strain-based approaches for structural concrete: Example of punching shear. *Engineering Structures*, *292*, 116509.

[31] Tennøe, S., Halnes, G., & Einevoll, G. T. (2018). Uncertainty: A python toolbox for uncertainty quantification and sensitivity analysis in computational neuroscience. *Frontiers in Neuroinformatics*, *12*, 49. https://doi.org/10.3389/fninf.2018.00049

[32] Song, T., Pu, H., Schonfeld, P., & Hu, J. (2022). Railway alignment optimization under uncertainty with a minimax robust method. *IEEE Intelligent Transportation Systems Magazine*, *15*(1), 333–346.

[33] Song, T., Pu, H., Schonfeld, P., Hu, J., & Liu, J. (2022). Robust optimization method for mountain railway alignments considering preference uncertainty for costs and seismic risks. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, *8*(1), 04021077.

[34] Zadeh, L. A. (1994). Fuzzy logic, neural networks, and soft computing. In G. J. Klir & B. Yuan (Eds.), *Fuzzy sets, fuzzy logic, and fuzzy systems* (pp. 775–782). World Scientific. https://doi.org/10.1145/175247.175255

[35] Yager, R. R., & Filev, D. P. (1994). Essentials of fuzzy modeling and control. *SIGART Bulletin*, *6*(4), 22–23.

[36] Klir, G. J., & T. Folger (1988). *Fuzzy sets, uncertainty and information*. USA: Prentice-Hall Inc.

[37] Klir, G.J. , U. H. St. Clair, & B. Yuah (1997). *Fuzzy set theory: Foundation and applications*. USA: Prentice-Hall Inc.

[38] Zimmermann, H. J. (1996). *Fuzzy set theory and its applications*. Germany: Springer Dordrecht.

[39] Kikuchi, S., & Miljkovic, D. (1999). Method to preprocess observed traffic data for consistency: Application of fuzzy optimization concept. *Transportation Research Record*, *1679*(1), 73–80.

[40] Kikuchi, S., Miljkovic, D., & van Zuylen, H. J. (2000). Examination of methods that adjust observed traffic volumes on a network. *Transportation Research Record*, *1717*(1), 109–119.

[41] Kikuchi, S. and Jha, M.K. (2006). A method for reconciling values of parameters. In *the INFORMS AI/DM Workshop*.

[42] Jha, M. K., & Schonfeld, P. (2004). A highway alignment optimization model using geographic information systems. *Transportation Research Part A: Policy and Practice*, *38*(6), 455–481. https://doi.org/10.1016/j.tra.2004.04.001

[43] Bellman, R. E., & Zadeh, L. A. (1970). Decision-making in a fuzzy environment. *Management Science*, *17*(4), B-141.

[44] Loquercio, A., Segu, M., & Scaramuzza, D. (2020). A general framework for uncertainty estimation in deep learning. *IEEE Robotics and Automation Letters*, *5*(2), 3153–3160.

How to Cite: Jha, M. K. (2025). Handling Uncertain Information with Fuzzy Logic in Locating New Infrastructure. *Journal of Data Science and Intelligent Systems*. https://doi.org/10.47852/bonviewJDSIS52025076