

Article

Forecasting the Number of Passengers on Hungarian Railway Routes Using a Similarity and Fuzzy Arithmetic-Based Inference Method

Marcell Fetter [†] and Tamás Jónás ^{*,†} 

Faculty of Economics, Eötvös Loránd University, 1053 Budapest, Hungary; mfetter@student.elte.hu

* Correspondence: jonas@gtk.elte.hu

[†] These authors contributed equally to this work.

Abstract: In this study, we present a similarity and fuzzy arithmetic-based fuzzy inference method and show how effectively it can be used to forecast the number of passengers on a railway route. We introduce a novel fuzzy similarity measure that is derived from the so-called epsilon function, which may be viewed as an alternative to the exponential function. After demonstrating the most important properties of the new similarity measure, we construct a fuzzy inference method that is founded on arithmetic operations over triangular fuzzy numbers. This inference method utilizes the proposed similarity measure to derive weight values for the above-mentioned arithmetic operations. The motivation behind the proposed method is twofold. On the one hand, we aim to construct a method that is simple and easy to implement. On the other hand, we intend to ensure that this method meets the practical requirements for rail passenger forecasts. Using a real-life case study, we demonstrate how well our method can predict the expected number of passengers on a new railway route based on characteristics of this relation. With respect to the studied case, we may conclude that although the similarity and fuzzy arithmetic-based fuzzy inference system has only two adjustable parameters, it may be regarded as a viable alternative to Sugeno-type fuzzy inference systems with a much greater number of adjustable parameters tuned by various optimization techniques.



Academic Editors: Yu-Wang Chen, Mi Zhou and Tao Wen

Received: 8 March 2025

Revised: 3 April 2025

Accepted: 7 April 2025

Published: 8 April 2025

Citation: Fetter, M.; Jónás, T.Forecasting the Number of Passengers on Hungarian Railway Routes Using a Similarity and Fuzzy Arithmetic-Based Inference Method. *Mathematics* **2025**, *13*, 1221. <https://doi.org/10.3390/math13081221>

Copyright: © 2025 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 (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: fuzzy arithmetic operations; fuzzy similarity; similarity and fuzzy arithmetic-based inference; predicting number of passengers; passenger demand forecasting; railway transportation planning

MSC: 68U01

1. Introduction

Nowadays, when mobility plays a major role in the economy, in business and in private life, understanding the magnitude of the demand for a transport service has become an important factor for service providers. This is also the case for railway operators, for whom forecasting the expected number of passengers on a railway route that does not yet exist is of great importance. Accordingly, several methods have been developed to forecast rail passenger traffic. Some of them are mentioned here, although the list is not exhaustive.

There are some examples of fuzzy models used for short-term passenger demand prediction. A fuzzy model has been developed in a study to predict peak-time demand on the Beijing–Shanghai high-speed railway, which predicts peak cross sections during holiday seasons, helping rail dispatchers to minimize costs and the number of passengers

who miss free capacity (see [1]). The application of this method has been validated in practice, showing that fuzzy forecasting has led to significant improvements compared to the previously used methods. There are also some examples of looking for passenger demand trends in a longer-term historical passenger flow database, and learning the characteristics of historical passenger flows can be used to predict future demands using regression techniques (see [2,3]). In other studies, models for air transport demand forecasting were developed by combining neural and fuzzy logic models, with a strong emphasis on the seasonality of transport demands (see [4,5]). A model based on fuzzy logic has also been developed in a study on the prioritization of transport investments, with the aim of prioritizing governmental investments that can have the greatest impact on the shift to public and sustainable transport modes (see [6]). There are also examples of research topics that deal with the human perception of public transport systems, with the aim of demonstrating the potential to increase public transportation ridership (see [7]). Many research articles deal with the implementation of artificial neural networks (ANNs) in transportation sciences and transportation forecasting. The findings and the results comparisons show that these applications have a better prediction ability than the standard methods; however, in most cases it is pointed out that there is more to come in the development and use of artificial neural networks for demand forecasting (see [8–10]). Recent studies on traffic forecasting suggest that combinations of existing travel demand forecasting methods can produce more accurate results than using different methods separately. Forecasting techniques ranging from regression models to artificial neural networks and support vector machines (SVMs) incorporating soft computing methods are discussed and compared in [11]. There are many examples where fuzzy forecasting models are combined with neural networks, combinations that can yield more accurate predictions, but the background to the functioning of neural networks might cause difficulties in some cases (see [12,13]). There are also examples of comparing different types of transport demand forecasting models and identifying factors that influence rail transport demand. Examples of such explanatory variables include average rail passenger travel distance, car ownership index, number of buses working in interurban routes per year, or a competition variable, expressed as the ratio of unit cost by bus to the unit cost by rail (see [14]). In some cases, Grey Relational Analysis is also adopted in transportation demand forecasting as well as other consumer demand models (see [15,16]). In other studies, the characteristics of total passenger demand, rail passenger demand and personal car demand are also discussed, with the possible modeling of the modal split (see [17,18]). Passenger demands can also be examined through capacity allocation and timetable planning models. There are examples for the development of timetable optimizations based on the characteristics of railway infrastructure and vehicles (see [19,20]).

The difficulties of standardizing different demand forecasting models is also highlighted, indicating the difficulties of choosing the right type of modeling methods that could be used in different phases of a transportation planning period. An article from recent years dealt with creating a standardized transport demand forecasting knowledge base from which the optimal forecasting method can be easily selected during each planning phase (see [21]).

In this study, we introduce a similarity and fuzzy arithmetic-based fuzzy inference method and demonstrate its effectiveness in predicting the expected number of future passengers on a railway connection. We present a new fuzzy similarity measure derived from the so-called epsilon function, which serves as an alternative to the exponential function (see [22,23]). After highlighting the key properties of this new similarity measure, we develop a fuzzy inference technique that employs arithmetic operations on triangular fuzzy numbers. This method uses the proposed similarity measure to determine weight

values for the aforementioned operations. Our motivation for this approach is twofold: firstly, we aim to create a method that is straightforward and easy to apply; secondly, we aspire to ensure that this method fulfills the practical requirements for rail passenger forecasting. Through a real-world case study that uses data from the Hungarian State Railways (MÁV), we illustrate how effectively our method can estimate the average number of passengers for a new railway connection based on the specific characteristics of that connection. It is worth noting that while our similarity and fuzzy arithmetic-based inference system has only two adjustable parameters, it can serve as a practical alternative to Sugeno-type fuzzy inference systems, which typically have a much larger number of adjustable parameters optimized through various techniques such as the Adaptive Neuro Fuzzy Inference System (ANFIS) (see, e.g., [24,25]), genetic algorithm (GA) (see, e.g., [26–28]), particle swarm optimization (PSO) (see, e.g., [29]) and pattern-search (PS) method (see, e.g., [30]). Our model and the Sugeno-type fuzzy inference systems with the above-mentioned parameter optimization techniques were implemented in the MATLAB (R2024b) software package [31]. The respective MATLAB source files ‘SFAFIS_example.m’ and ‘FIS_similarity_and_fuzzy_arithmetic_class.m’, which are available at <https://jonast.web.elte.hu/>, include the example of our case study. With these two files, the results of our case study can be reproduced.

The remaining part of this paper is structured as follows. In Section 2, we provide an overview of the basic concepts we employ in our study. Here, we present the arithmetic operations that will be later utilized over triangular fuzzy numbers and describe a function that will be used to construct a similarity measure. In Section 3, we propose a new fuzzy similarity measure and demonstrate its main properties. A novel fuzzy inference method called the similarity and fuzzy arithmetic-based fuzzy inference method is then presented in Section 4. In Section 5, by the means of a real-life case study, we demonstrate how this new method can be used to predict the number of passengers on a railway route. Here, we also compare our method with some much more sophisticated fuzzy and neuro-fuzzy methods. Lastly, we make some concluding remarks and outline possible directions for future research.

2. Preliminaries

Here, we will provide a brief overview of the concepts and notations that we will make use of later on. We will use the common notation $\mathbb{R}$ for the real line.

Definition 1 (cf. [32]). *Let X be a non-empty set. We say that the function $\mu_A: X \rightarrow [0, 1]$ is the membership function of the fuzzy set A on the universe X . For any $x \in X$, $\mu_A(x) \in [0, 1]$ is the membership value of x in the fuzzy set A .*

Definition 2 (cf. [33]). *Let X be a non-empty set and let $\mu_A: X \rightarrow [0, 1]$ be the membership function of the fuzzy set A on the universe X . For any $\alpha \in (0, 1]$, we say that the ordinary set*

$$A_\alpha = \{x \in X: \mu_A(x) \geq \alpha\} \quad (1)$$

is the α -cut of the fuzzy set A .

2.1. Some Arithmetic Operations on Triangular Fuzzy Numbers

We will construct a fuzzy arithmetic-based inference method in which we will utilize triangular fuzzy numbers. We will use the following definition of triangular fuzzy numbers.

Definition 3 (cf. [34]). We say that a fuzzy set A on the real line is a triangular fuzzy number with the parameters $a, b, c \in \mathbb{R}$, $a < b < c$, if its membership function $\mu_A(\cdot; a, b, c): \mathbb{R} \rightarrow [0, 1]$ is given by

$$\mu_A(x; a, b, c) = \begin{cases} 0, & \text{if } x \leq a \\ \frac{x-a}{b-a}, & \text{if } a < x \leq b \\ \frac{c-x}{c-b}, & \text{if } b < x \leq c \\ 0, & \text{if } x > c. \end{cases} \quad (2)$$

From now on, we will use the notation $T(a, b, c)$ for a triangular fuzzy number with the real-valued parameters a, b and c , where $a < b < c$. That is, $A = T(a, b, c)$ means that A is a triangular fuzzy number with the parameters a, b and c , and so the membership function of A is the function $\mu_A(\cdot; a, b, c): \mathbb{R} \rightarrow [0, 1]$ given in Definition 3.

Remark 1. It should be added that fuzzy numbers are a class of fuzzy sets (see, e.g., [34,35]). Namely, a fuzzy set A with the membership function $\mu_A: \mathbb{R} \rightarrow [0, 1]$ is said to be a fuzzy number if

- (a) A is normal, i.e., there exists a $x_0 \in \mathbb{R}$, such that $\mu_A(x_0) = 1$.
- (b) A is fuzzy convex, i.e., for any $t \in [0, 1]$ and $x, y \in \mathbb{R}$,

$$\mu_A(tx + (1-t)y) \geq \min\{\mu_A(x), \mu_A(y)\}.$$

- (c) μ_A is upper semi-continuous on $\mathbb{R}$.
- (d) The support of A , i.e., the set $\{x \in \mathbb{R}: \mu_A(x) > 0\}$, is bounded.

It can be verified that a fuzzy set with the membership function $\mu_A(\cdot; a, b, c): \mathbb{R} \rightarrow [0, 1]$ given in Equation (2) satisfies the criteria for a fuzzy number. This means that triangular fuzzy numbers are a class of fuzzy numbers.

It is well-known (see, e.g., [33,34]) that using the concept of α -cut and interval arithmetic operations, the addition, subtraction, multiplication by scalar, multiplication and division operations over triangular fuzzy numbers can be defined so that any set of triangular fuzzy numbers is closed under these operations. Noting the definition of the α -cut of a fuzzy set given in Definition 2, after direct calculation, we find that for any $\alpha \in (0, 1]$, the α -cut of the triangular fuzzy number $T(a, b, c)$, $T(a, b, c)_\alpha$, is the interval

$$T(a, b, c)_\alpha = [(1-\alpha)a + \alpha b, \alpha b + (1-\alpha)c], \quad (3)$$

where $a, b, c \in \mathbb{R}$ and $a < b < c$. Notice that in Definition 2, $\alpha \in (0, 1]$; however, for $\alpha = 0$, the α -cut of $T(a, b, c)$ is the interval $[a, c]$. Hence, for triangular fuzzy numbers, we interpret the α -cut for any $\alpha \in [0, 1]$ as the interval given by Equation (3).

Now, we will show that the weighted arithmetic mean of triangular fuzzy numbers is a triangular fuzzy number whose parameters are the weighted arithmetic means of the corresponding parameters of the triangular fuzzy numbers in question. We will utilize the following operations over intervals.

Definition 4. For an interval $[x, y]$ and a scalar $t \in \mathbb{R}$, $t > 0$, the operation $t \odot_{\mathcal{I}} [x, y]$ is defined as

$$t \odot_{\mathcal{I}} [x, y] = [tx, ty]. \quad (4)$$

Definition 5. For the intervals $[x_1, y_1], [x_2, y_2], \dots, [x_n, y_n]$, $n \in \mathbb{N}$, $n \geq 2$ the operation $\bigoplus_{i=1}^n [x_i, y_i]$ is defined as

$$\bigoplus_{i=1}^n [x_i, y_i] = \left[\sum_{i=1}^n x_i, \sum_{i=1}^n y_i \right]. \quad (5)$$

Using the scalar multiplication operation given in Definition 4 and the summation operator given in Definition 5, we can state the following theorem.

Theorem 1. Let $T(a_1, b_1, c_1), T(a_2, b_2, c_2), \dots, T(a_n, b_n, c_n)$ be triangular fuzzy numbers, where $n \in \mathbb{N}$, $n \geq 2$, and $a_i < b_i < c_i$ for all $i \in \{1, 2, \dots, n\}$. Furthermore, let $w_1, w_2, \dots, w_n \in [0, 1]$ such that $\sum_{i=1}^n w_i = 1$. Then, for any $\alpha \in [0, 1]$,

$$\bigoplus_{i=1}^n (w_i \odot_{\mathcal{I}} T(a_i, b_i, c_i)_{\alpha}) = T \left(\sum_{i=1}^n w_i a_i, \sum_{i=1}^n w_i b_i, \sum_{i=1}^n w_i c_i \right)_{\alpha}, \quad (6)$$

where $T(p, q, r)_{\alpha}$ denotes the α -cut of a triangular fuzzy number with the parameters $p, q, r \in \mathbb{R}$, $p < q < r$.

Proof. Let $\alpha \in [0, 1]$. Based on Equation (3), the α -cut of $T(a_i, b_i, c_i)$ is the interval

$$T(a_i, b_i, c_i)_{\alpha} = [(1 - \alpha)a_i + \alpha b_i, \alpha b_i + (1 - \alpha)c_i], \quad (7)$$

where $i \in \{1, 2, \dots, n\}$. Now, using Equation (7), and the scalar multiplication and addition operations over intervals given by Equation (4) and Equation (5), respectively, we find that

$$\begin{aligned} \bigoplus_{i=1}^n (w_i \odot_{\mathcal{I}} T(a_i, b_i, c_i)_{\alpha}) &= \bigoplus_{i=1}^n (w_i \odot_{\mathcal{I}} [(1 - \alpha)a_i + \alpha b_i, \alpha b_i + (1 - \alpha)c_i]) \\ &= \bigoplus_{i=1}^n [w_i((1 - \alpha)a_i + \alpha b_i), w_i(\alpha b_i + (1 - \alpha)c_i)] \\ &= \bigoplus_{i=1}^n [(1 - \alpha)w_i a_i + \alpha w_i b_i, \alpha w_i b_i + (1 - \alpha)w_i c_i] \\ &= \left[\sum_{i=1}^n ((1 - \alpha)w_i a_i + \alpha w_i b_i), \sum_{i=1}^n (\alpha w_i b_i + (1 - \alpha)w_i c_i) \right] \\ &= \left[(1 - \alpha) \left(\sum_{i=1}^n w_i a_i \right) + \alpha \left(\sum_{i=1}^n w_i b_i \right), \alpha \left(\sum_{i=1}^n w_i b_i \right) + (1 - \alpha) \left(\sum_{i=1}^n w_i c_i \right) \right]. \end{aligned} \quad (8)$$

Next, noting Equation (7) again, we see that

$$\begin{aligned} &\left[(1 - \alpha) \left(\sum_{i=1}^n w_i a_i \right) + \alpha \left(\sum_{i=1}^n w_i b_i \right), \alpha \left(\sum_{i=1}^n w_i b_i \right) + (1 - \alpha) \left(\sum_{i=1}^n w_i c_i \right) \right] \\ &= T \left(\sum_{i=1}^n w_i a_i, \sum_{i=1}^n w_i b_i, \sum_{i=1}^n w_i c_i \right)_{\alpha}. \end{aligned} \quad (9)$$

Hence, with Equations (8) and (9), we obtain

$$\bigoplus_{i=1}^n (w_i \odot_{\mathcal{I}} T(a_i, b_i, c_i)_{\alpha}) = T \left(\sum_{i=1}^n w_i a_i, \sum_{i=1}^n w_i b_i, \sum_{i=1}^n w_i c_i \right)_{\alpha}.$$

□

We will make use of the following scalar multiplication and summation operations over triangular fuzzy numbers.

Definition 6. For a triangular fuzzy number $T(a, b, c)$ and a $t \in \mathbb{R}$, $t > 0$, the scalar multiplication operation $t \odot T(a, b, c)$ is defined as

$$t \odot T(a, b, c) = T(ta, tb, tc). \quad (10)$$

Definition 7. For the triangular fuzzy numbers $T(a_1, b_1, c_1), T(a_2, b_2, c_2), \dots, T(a_n, b_n, c_n)$, the summation operation $\bigoplus_{i=1}^n T(a_i, b_i, c_i)$ is defined as

$$\bigoplus_{i=1}^n T(a_i, b_i, c_i) = T\left(\sum_{i=1}^n a_i, \sum_{i=1}^n b_i, \sum_{i=1}^n c_i\right). \quad (11)$$

Exploiting Theorem 1 and using Definitions 6 and 7, we can state the following corollary.

Corollary 1. Let $T(a_1, b_1, c_1), T(a_2, b_2, c_2), \dots, T(a_n, b_n, c_n)$ be triangular fuzzy numbers, where $n \in \mathbb{N}$, $n \geq 2$, and $a_i < b_i < c_i$ for all $i \in \{1, 2, \dots, n\}$. Furthermore, let $w_1, w_2, \dots, w_n \in [0, 1]$ such that $\sum_{i=1}^n w_i = 1$. Then, the following hold:

(a) The weighted arithmetic mean of $T(a_1, b_1, c_1), T(a_2, b_2, c_2), \dots, T(a_n, b_n, c_n)$ with respect to the weights $w_1, w_2, \dots, w_n$, i.e., $\bigoplus_{i=1}^n (w_i \odot T(a_i, b_i, c_i))$, is

$$\bigoplus_{i=1}^n (w_i \odot T(a_i, b_i, c_i)) = T\left(\sum_{i=1}^n w_i a_i, \sum_{i=1}^n w_i b_i, \sum_{i=1}^n w_i c_i\right). \quad (12)$$

(b) For any $\alpha \in [0, 1]$, the α -cut of the weighted arithmetic mean $\bigoplus_{i=1}^n (w_i \odot T(a_i, b_i, c_i))$, i.e., $\left(\bigoplus_{i=1}^n (w_i \odot T(a_i, b_i, c_i))\right)_\alpha$, is equal to the weighted arithmetic mean of the α -cuts $T(a_1, b_1, c_1)_\alpha, T(a_2, b_2, c_2)_\alpha, \dots, T(a_n, b_n, c_n)_\alpha$ with respect to the weights $w_1, w_2, \dots, w_n$. That is,

$$\left(\bigoplus_{i=1}^n (w_i \odot T(a_i, b_i, c_i))\right)_\alpha = \bigoplus_{i=1}^n (w_i \odot T(a_i, b_i, c_i))_\alpha. \quad (13)$$

Proof.

(a). Using Equations (10) and (11), we immediately get

$$\bigoplus_{i=1}^n (w_i \odot T(a_i, b_i, c_i)) = \bigoplus_{i=1}^n T(w_i a_i, w_i b_i, w_i c_i) = T\left(\sum_{i=1}^n w_i a_i, \sum_{i=1}^n w_i b_i, \sum_{i=1}^n w_i c_i\right).$$

(b). Noting Equations (6) and (12) and Theorem 1, we can write

$$\begin{aligned} \left(\bigoplus_{i=1}^n (w_i \odot T(a_i, b_i, c_i))\right)_\alpha &= T\left(\sum_{i=1}^n w_i a_i, \sum_{i=1}^n w_i b_i, \sum_{i=1}^n w_i c_i\right)_\alpha \\ &= \bigoplus_{i=1}^n (w_i \odot T(a_i, b_i, c_i))_\alpha. \end{aligned}$$

□

Note that the operators $\odot_{\mathcal{I}}$ and $\oplus_{\mathcal{I}}$ are defined over intervals, while $\odot$ and $\oplus$ are defined over triangular fuzzy numbers. Corollary 1 (b) tells us that any α -cut interval of the weighted arithmetic mean of some triangular fuzzy numbers is none other than the weighted arithmetic mean of the α -cut intervals of the respective triangular fuzzy numbers.

We will utilize this property of triangular fuzzy numbers in our fuzzy arithmetic-based inference method.

2.2. The Epsilon Function

We will use the so-called epsilon function, which was introduced by Dombi et al. in [22], to construct a fuzzy similarity measure. The epsilon function is defined as follows.

Definition 8 (cf. [22]). Let $\lambda \in \mathbb{R}$, $\lambda \neq 0$ and $\Delta \in \mathbb{R}$, $\Delta > 0$. We say that the function $\varepsilon_{\lambda,\Delta}: (-\Delta, +\Delta) \rightarrow \mathbb{R}^+$, which is given by

$$\varepsilon_{\lambda,\Delta}(x) = \left(\frac{\Delta + x}{\Delta - x} \right)^{\lambda \frac{\Delta}{2}}, \quad (14)$$

is an epsilon function with the parameters λ and Δ .

For more details on the epsilon function and its generalization, see [23]. With direct calculation, we find that

$$\varepsilon_{\lambda,\Delta}(x) \Big|_{x=0} = 1 \quad (15)$$

and

$$\frac{d\varepsilon_{\lambda,\Delta}(x)}{dx} \Big|_{x=0} = \lambda \frac{\Delta^2}{\Delta^2 - x^2} \left(\frac{\Delta + x}{\Delta - x} \right)^{\lambda \frac{\Delta}{2}} \Big|_{x=0} = \lambda \frac{\Delta^2}{\Delta^2 - x^2} \varepsilon_{\lambda,\Delta}(x) \Big|_{x=0} = \lambda. \quad (16)$$

Since

$$e^{\lambda x} \Big|_{x=0} = 1 \quad (17)$$

and

$$\frac{de^{\lambda x}}{dx} \Big|_{x=0} = \lambda, \quad (18)$$

from Equations (15)–(18) we see that $\varepsilon_{\lambda,\Delta}(x)$ and $e^{\lambda x}$ are identical to first order at $x = 0$. Moreover, we have the following result.

Theorem 2 (cf. [22]). For any $x \in (-\Delta, +\Delta)$,

$$\lim_{\Delta \rightarrow \infty} \varepsilon_{\lambda,\Delta}(x) = e^{\lambda x}.$$

Based on Theorem 2 and the fact that $\varepsilon_{\lambda,\Delta}(x)$ and $e^{\lambda x}$ are identical to first order at $x = 0$, the epsilon function can be regarded as a good approximation of the exponential function. Exploiting this property of the epsilon function, we will use it to construct a similarity measure.

3. A Fuzzy Similarity Measure Derived from the Epsilon Function

It is well known that if $d: [0, 1]^n \times [0, 1]^n \rightarrow [0, 1]$ is a normalized distance measure in the vector space $[0, 1]^n$ and $\lambda > 0$, then for any vectors $\mathbf{x}, \mathbf{y} \in [0, 1]^n$,

$$s_{\lambda}^*(\mathbf{x}, \mathbf{y}) = e^{-\lambda d(\mathbf{x}, \mathbf{y})} \quad (19)$$

may be viewed as a measure of similarity between the vectors $\mathbf{x}$ and $\mathbf{y}$. Noting Equation (19), we see that the similarity measure $s_{\lambda}^*: [0, 1]^n \times [0, 1]^n \rightarrow [0, 1]$ has the following elementary properties:

- (a) For any $\mathbf{x}, \mathbf{y} \in [0, 1]^n$, if $d(\mathbf{x}, \mathbf{y}) = 0$, then $s_{\lambda}^*(\mathbf{x}, \mathbf{y}) = 1$ (maximal similarity).
- (b) For any $\mathbf{x}, \mathbf{y} \in [0, 1]^n$, if $d(\mathbf{x}, \mathbf{y}) = 1$, then $s_{\lambda}^*(\mathbf{x}, \mathbf{y}) = e^{-\lambda}$ (minimal similarity).

- (c) For any $\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}' \in [0, 1]^n$, $d(\mathbf{x}', \mathbf{y}') > d(\mathbf{x}, \mathbf{y})$ implies $s_\lambda^*(\mathbf{x}', \mathbf{y}') < s_\lambda^*(\mathbf{x}, \mathbf{y})$ (monotonocity).

Notice that if λ is large, i.e., $\lambda \gg 0$, then $e^{-\lambda} \approx 0$. That is, if $\lambda \gg 0$ and $d(\mathbf{x}, \mathbf{y}) = 1$, then $s_\lambda^*(\mathbf{x}, \mathbf{y}) \approx 0$.

It is well-known that for any $\lambda > 0$ and $x \in [0, 1]$, the function $g_\lambda(x) = (1 - x)^\lambda$ is a good approximation of $e^{-\lambda x}$. Furthermore, $g_\lambda(0) = 1$, $g_\lambda(1) = 0$ and $g_\lambda(x)$ and $e^{-\lambda x}$ are identical to first order at $x = 0$.

Now, let $\lambda > 0$ and define $h_\lambda: [0, 1] \rightarrow [0, 1]$ as

$$h_\lambda(x) := \varepsilon_{-\lambda, 1}(x), \quad x \in [0, 1],$$

i.e., $h_\lambda(\cdot)$ is an epsilon function on $[0, 1]$ with the parameters $-\lambda$ and $\Delta = 1$. Hence, noting Definition 8, $h_\lambda(x)$ is given by

$$h_\lambda(x) = \left(\frac{1+x}{1-x} \right)^{-\frac{\lambda}{2}} = \left(\frac{1-x}{1+x} \right)^{\frac{\lambda}{2}}, \quad x \in [0, 1]. \quad (20)$$

Taking into account the properties of the epsilon function, we readily find that for any $\lambda > 0$, $h_\lambda(0) = 1$, $h_\lambda(1) = 0$ and $h_\lambda(x)$ and $e^{-\lambda x}$ are identical to first order at $x = 0$. Hence, we see that $g_\lambda(\cdot)$ and $h_\lambda(\cdot)$ have similar properties on $[0, 1]$. The following theorem states that for any $\lambda > 0$ and $x \in [0, 1]$, $h_\lambda(x)$ is even a better approximation of $e^{-\lambda x}$ than $g_\lambda(x)$.

Theorem 3. For any $\lambda > 0$ and $x \in [0, 1]$, it holds that

$$(1-x)^\lambda \leq \left(\frac{1-x}{1+x} \right)^{\frac{\lambda}{2}} \leq e^{-\lambda x}. \quad (21)$$

Proof. It is sufficient to show that

$$(1-x)^2 \leq \frac{1-x}{1+x} \leq e^{-2x} \quad (22)$$

holds because noting that $\lambda > 0$ and raising both members of Equation (22) to $\frac{\lambda}{2}$, we get Equation (21).

First, we will prove the left hand side inequality in Equation (22). Since $x \in [0, 1]$, we have $1 - x^2 \leq 1$, and as $1 - x^2 = (1 - x)(1 + x)$, we find that

$$(1 - x)(1 + x) \leq 1,$$

from which, taking into account the fact that $1 - x \geq 0$ and $1 + x > 0$,

$$(1 - x)^2 \leq \frac{1 - x}{1 + x}$$

follows.

Now, we will prove the right hand side inequality in Equation (22). With the Maclaurin series of e^x , we have

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!},$$

and since $x \in [0, 1]$, we see that

$$e^x \geq 1 + x + \frac{x^2}{2},$$

from which

$$e^{-2x} \leq \frac{1}{\left(1 + x + \frac{x^2}{2}\right)^2} \quad (23)$$

follows. As $x \in [0, 1]$, we also have

$$\frac{1}{\left(1 + x + \frac{x^2}{2}\right)^2} \leq \frac{1}{(1+x)^2}, \quad (24)$$

and so based on Equations (23) and (24), we find that

$$\frac{1}{(1+x)^2} \geq e^{-2x} \quad (25)$$

holds for any $x \in [0, 1]$, and the equality in Equation (25) holds only if $x = 0$. Now, define $f(x)$ as

$$f(x) = e^{-2x} - \frac{1-x}{1+x}, \quad x \in [0, 1].$$

The first derivative of f is

$$\frac{df(x)}{dx} = 2 \left(\frac{1}{(x+1)^2} - e^{-2x} \right),$$

and taking into account the inequality given in Equation (25), we see that for any $x \in [0, 1]$,

$$\frac{df(x)}{dx} \geq 0.$$

Since the equality in Equation (25) holds only if $x = 0$, we find that for any $x \in (0, 1]$,

$$\frac{df(x)}{dx} > 0,$$

which means that $f(\cdot)$ is a strictly increasing function on $(0, 1]$. As $f(0) = 0$ and $f(\cdot)$ is a strictly increasing function on $(0, 1]$, we see that for any $x \in [0, 1]$, $f(x) \geq 0$. Therefore, for any $x \in [0, 1]$, we have

$$\left(\frac{1-x}{1+x} \right)^{\frac{\lambda}{2}} \leq e^{-\lambda x}.$$

□

Remark 2. We should add that an immediate practical consequence of Theorem 3 is that if $\lambda \gg 0$, then $h_\lambda(x) \approx e^{-\lambda x}$ on $[0, 1]$.

Figure 1 shows some example plots of $e^{-\lambda x}$ and its approximation by $g_\lambda(x)$ and the epsilon function $h_\lambda(x) = \varepsilon_{-\lambda,1}(x)$ on $[0, 1]$.

Based on the above considerations, we define the epsilon-function-based normalized fuzzy similarity measure as follows.

Definition 9. Let $d: [0, 1]^n \times [0, 1]^n \rightarrow [0, 1]$ be a normalized distance measure in the vector space $[0, 1]^n$. We say that the function $s_\lambda: [0, 1]^n \times [0, 1]^n \rightarrow [0, 1]$ is an epsilon-function-based

normalized fuzzy similarity measure with a parameter $\lambda > 0$ in the vector space $[0, 1]^n$, if for any $\mathbf{x}, \mathbf{y} \in [0, 1]^n$, $s_\lambda(\mathbf{x}, \mathbf{y})$ is given by

$$s_\lambda(\mathbf{x}, \mathbf{y}) = \left(\frac{1 - d(\mathbf{x}, \mathbf{y})}{1 + d(\mathbf{x}, \mathbf{y})} \right)^{\frac{\lambda}{2}}. \quad (26)$$

Notice that $s_\lambda(\mathbf{x}, \mathbf{y}) = h_\lambda(d(\mathbf{x}, \mathbf{y}))$, where $h_\lambda(\cdot)$ is the epsilon function on $[0, 1]$ given by Equation (20). The following lemma summarizes the main properties of the epsilon-function-based normalized fuzzy similarity measure given in Definition 9.

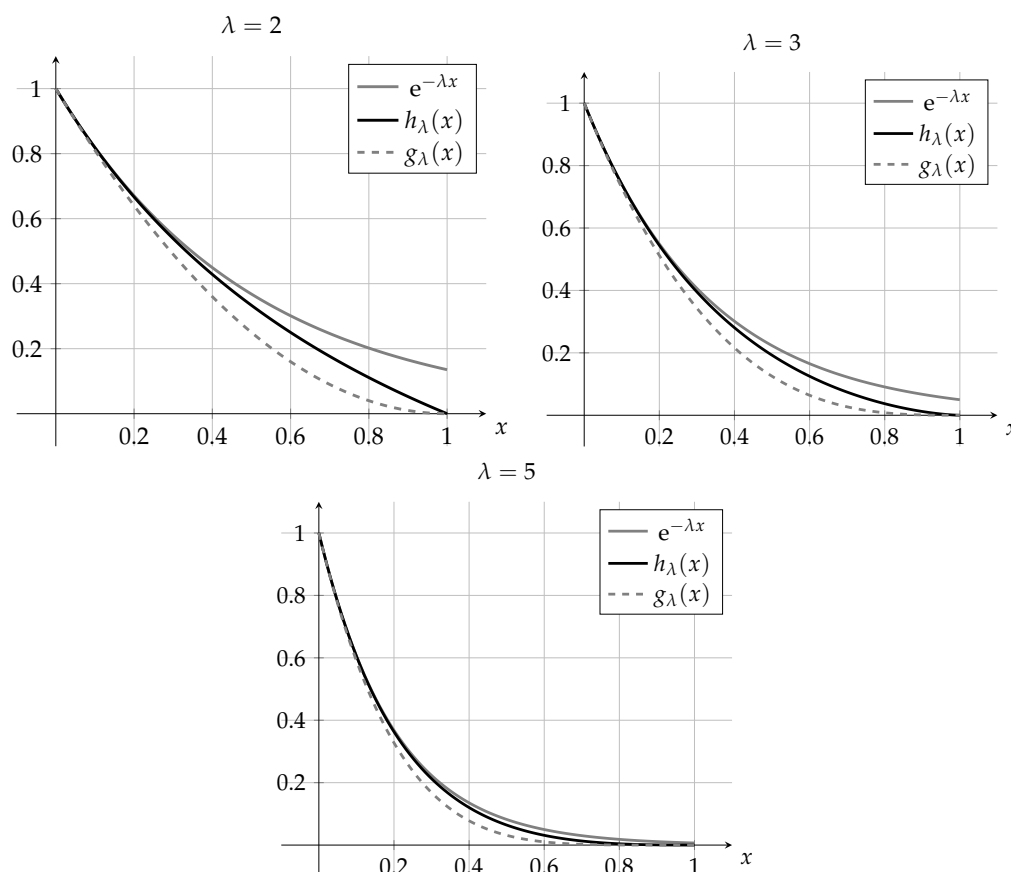


Figure 1. Example plots of $e^{-\lambda x}$ and its approximation by $g_\lambda(x) = (1-x)^\lambda$ and the epsilon function $h_\lambda(x) = \left(\frac{1-x}{1+x} \right)^{\frac{\lambda}{2}}$ on $[0, 1]$.

Lemma 1. Let $d: [0, 1]^n \times [0, 1]^n \rightarrow [0, 1]$ be a normalized distance measure in the vector space $[0, 1]^n$ and let $s_\lambda: [0, 1]^n \times [0, 1]^n \rightarrow [0, 1]$ be the corresponding epsilon-function-based normalized fuzzy similarity measure with a parameter $\lambda > 0$ given by Equation (26). $s_\lambda(\cdot)$ has the following properties:

- For any $\mathbf{x}, \mathbf{y} \in [0, 1]^n$, if $d(\mathbf{x}, \mathbf{y}) = 0$, then $s_\lambda(\mathbf{x}, \mathbf{y}) = 1$ (maximal similarity).
- For any $\mathbf{x}, \mathbf{y} \in [0, 1]^n$, if $d(\mathbf{x}, \mathbf{y}) = 1$, then $s_\lambda(\mathbf{x}, \mathbf{y}) = 0$ (minimal similarity).
- For any $\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}' \in [0, 1]^n$, $d(\mathbf{x}', \mathbf{y}') > d(\mathbf{x}, \mathbf{y})$ implies $s_\lambda(\mathbf{x}', \mathbf{y}') < s_\lambda(\mathbf{x}, \mathbf{y})$ (monotonocity).
- If $\lambda \gg 0$, then for any $\mathbf{x}, \mathbf{y} \in [0, 1]^n$, $s_\lambda(\mathbf{x}, \mathbf{y}) \approx s_\lambda^*(\mathbf{x}, \mathbf{y})$, where $s_\lambda^*(\mathbf{x}, \mathbf{y})$ is given by Equation (19).
- If $d(\mathbf{x}, \mathbf{y}) \in (0, 1)$, then

$$\lim_{\lambda \rightarrow \infty} s_\lambda(\mathbf{x}, \mathbf{y}) = 0. \quad (27)$$

Proof. Properties (a), (b) and (c) readily follow from the definition of $s_\lambda(\cdot)$ given in Definition 9, while property (b) is a consequence of Theorem 3. $\square$

4. A Similarity and Fuzzy Arithmetic-Based Fuzzy Inference Method

In our model, we assume that the value of a dependent variable $y \in \mathbb{R}$ is a function of an n -dimensional normalized feature vector $\mathbf{x} \in [0, 1]^n$ as follows:

$$y = f(\mathbf{x}) + \varepsilon, \quad (28)$$

where $n \in \mathbb{N}$, $n \geq 1$, f is an unknown mapping from $[0, 1]^n$ to $\mathbb{R}$ and ε is a normally distributed random variable with mean of 0. Our goal is to build a fuzzy inference method to approximate function f .

4.1. Forming a Feature Vector–Linguistic Value of the Number of Passenger Pairs

We will construct a Similarity and Fuzzy Arithmetic-based Fuzzy Inference System (SFAFIS) to approximate the mapping f given in Equation (28). For this purpose, we will use a set of training data given by the ordered pairs

$$(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_m, y_m), \quad (29)$$

where $m \in \mathbb{N}$, $m \geq 2$,

$$\mathbf{x}_i = (x_{i,1}, x_{i,2}, \dots, x_{i,n}) \in [0, 1]^n,$$

$i = 1, 2, \dots, m$, $n \in \mathbb{N}$, $n \geq 1$ and $y_i \in \mathbb{R}$, $y_i > 0$. Here, the vector $\mathbf{x}_i$ and the scalar y_i represent the normalized feature vector (explanatory or input vector) and the corresponding output in the i th observation, respectively, in the sample given by Equation (29).

Let

$$q_0 = \min_{i=1,2,\dots,m} (y_i) \quad \text{and} \quad q_1 = \max_{i=1,2,\dots,m} (y_i),$$

and suppose that $y \in [q_0, q_1]$. Now, let $L_1, L_2, \dots, L_\ell$ be the linguistic values like ‘very small’, ‘small’, ‘large’, ‘very large’, etc., of y where $\ell \in \mathbb{N}$, $\ell \geq 3$. This means that $L_1, L_2, \dots, L_\ell$ are fuzzy sets on the interval $[q_0, q_1]$. Furthermore, let $q_{\frac{k}{\ell-1}}$ denote the $\frac{k}{\ell-1}$ -quantile of the $y_1, y_2, \dots, y_m$ values, where $k = 1, 2, \dots, \ell - 2$.

In our approach, the k th linguistic level L_k of y is a triangular fuzzy number with the parameters $a_k, b_k, c_k \in \mathbb{R}$, $a_k < b_k < c_k$. That is,

$$L_k = T(a_k, b_k, c_k),$$

and the membership function $\mu_{L_k}(\cdot; a_k, b_k, c_k): \mathbb{R} \rightarrow [0, 1]$ of L_k is given by

$$\mu_{L_k}(y; a_k, b_k, c_k) = \begin{cases} 0, & \text{if } y \leq a_k \\ \frac{y-a_k}{b_k-a_k}, & \text{if } a_k < y \leq b_k \\ \frac{c_k-y}{c_k-b_k}, & \text{if } b_k < y \leq c_k \\ 0, & \text{if } y > c_k \end{cases}$$

such that

if $k = 1$, then

$$a_k = q_0 - \left(q_{\frac{1}{\ell-1}} - q_0 \right) = 2q_0 - q_{\frac{1}{\ell-1}},$$

$$b_k = q_0,$$

$$c_k = q_{\frac{1}{\ell-1}},$$

if $k \in \{2, \dots, \ell - 1\}$, then

$$a_k = q_{\frac{k-2}{\ell-1}},$$

$$b_k = q_{\frac{k-1}{\ell-1}},$$

$$c_k = q_{\frac{k}{\ell-1}},$$

if $k = \ell$, then

$$a_k = q_{\frac{\ell-2}{\ell-1}},$$

$$b_k = q_1,$$

$$c_k = q_1 + (q_1 - q_{\frac{\ell-2}{\ell-1}}) = 2q_1 - q_{\frac{\ell-2}{\ell-1}}.$$

Figure 2 shows an example of the plots of membership functions of the triangular fuzzy numbers (linguistic values) $L_1, L_2, \dots, L_\ell$.

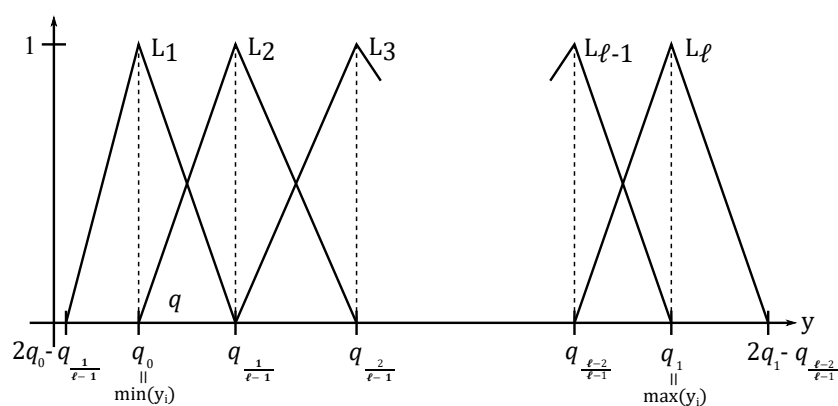


Figure 2. Example plots of the triangular fuzzy numbers (linguistic values) $L_1, L_2, \dots, L_\ell$.

Next, for all $i = 1, 2, \dots, m$, we transform the training data pair $(\mathbf{x}_i, y_i)$ into the pair

$$(\mathbf{x}_i, L_{k(i)}) = (\mathbf{x}_i, T(a_{k(i)}, b_{k(i)}, c_{k(i)}))$$

such that $k(i) \in \{1, 2, \dots, \ell\}$ is the smallest index for which the membership value $\mu_{L_{k(i)}}(y_i; a_{k(i)}, b_{k(i)}, c_{k(i)})$ equals the maximum of the membership values

$$\mu_{L_1}(y_i; a_1, b_1, c_1), \mu_{L_2}(y_i; a_2, b_2, c_2), \dots, \mu_{L_\ell}(y_i; a_\ell, b_\ell, c_\ell).$$

That is, $k(i)$ is given as

$$k(i) = \min \left\{ r: r \in \{1, 2, \dots, \ell\}, \mu_{L_r}(y_i; a_r, b_r, c_r) = \max_{t=1, \dots, \ell} (\mu_{L_t}(y_i; a_t, b_t, c_t)) \right\}.$$

In other words, $k(i) \in \{1, 2, \dots, \ell\}$ is the smallest index for which, among the linguistic values $L_1, L_2, \dots, L_\ell$, the linguistic value $L_{k(i)}$ best represents the y_i crisp value.

Remark 3. Alternatively, instead of the $\frac{k}{\ell-1}$ -quantile, $q_{\frac{k}{\ell-1}}$ can be defined as

$$q_{\frac{k}{\ell-1}} = \frac{k}{\ell-1},$$

where $k = 1, 2, \dots, \ell - 2$. In this case, we apply a grid partitioning, i.e., with $\ell - 1$ equidistant points, we divide the range $[q_0, q_1]$ into ℓ consecutive intervals with the same Lebesgue measure (i.e., same length).

4.2. The Inference Mechanism

Suppose that $\mathbf{x} \in [0, 1]^n$ is an input feature vector for which we aim to predict the value of y . Furthermore, let $s_\lambda: [0, 1]^n \times [0, 1]^n \rightarrow [0, 1]$ be an epsilon-function-based normalized fuzzy similarity measure with a parameter $\lambda > 0$ given in Definition 9. That is, for any $\mathbf{x}, \mathbf{x}' \in [0, 1]^n$, $s_\lambda(\mathbf{x}, \mathbf{x}')$ characterizes how much vector $\mathbf{x}$ is similar to the vector $\mathbf{x}'$. Now, for all $i = 1, 2, \dots, m$, let w_i be given by

$$w_i = \frac{s_\lambda(\mathbf{x}, \mathbf{x}_i)}{\sum_{j=1}^m s_\lambda(\mathbf{x}, \mathbf{x}_j)}.$$

Notice that $w_i \in [0, 1]$ and $\sum_{i=1}^m w_i = 1$, and the weight w_i may be regarded as the relative measure of how well the input feature vector $\mathbf{x}$ fits to the feature vector $\mathbf{x}_i$ of the i th training data pair $(\mathbf{x}_i, y_i)$, where $i \in \{1, 2, \dots, m\}$.

Since

$$L_{k(i)} = T(a_{k(i)}, b_{k(i)}, c_{k(i)}),$$

for all $i = 1, 2, \dots, m$, based on Corollary 1 (a), the weighted arithmetic mean of the triangular fuzzy numbers $L_{k(1)}, L_{k(2)}, \dots, L_{k(m)}$ is the triangular fuzzy number $T(a, b, c)$, which is given by

$$\begin{aligned} T(a, b, c) &= \bigoplus_{i=1}^m (w_i \odot L_{k(i)}) = \bigoplus_{i=1}^m (w_i \odot T(a_{k(i)}, b_{k(i)}, c_{k(i)})) \\ &= T\left(\sum_{i=1}^m w_i a_{k(i)}, \sum_{i=1}^m w_i b_{k(i)}, \sum_{i=1}^m w_i c_{k(i)}\right), \end{aligned} \quad (30)$$

where $\odot$ and $\oplus$ are the scalar multiplication and summation operators over triangular fuzzy numbers given in Definition 6 and Definition 7, respectively. Here, we treat the triangular fuzzy number $Y = T(a, b, c)$ as the fuzzy output of our similarity and fuzzy arithmetic-based fuzzy inference method, where

$$a = \sum_{i=1}^m w_i a_{k(i)}, \quad b = \sum_{i=1}^m w_i b_{k(i)}, \quad c = \sum_{i=1}^m w_i c_{k(i)}.$$

With this approach, our method takes the linguistic value of y_i (i.e., the triangular fuzzy number $L_{k(i)}$) into account in the aggregate fuzzy output Y as much as the input feature vector $\mathbf{x}$ fits to the i th feature vector $\mathbf{x}_i$. Hence, the defuzzified value of Y may be treated as a $\hat{y}$ prediction of y for the feature vector $\mathbf{x} \in [0, 1]^n$. Using the center of gravity defuzzification method, the $\hat{y}$ prediction of y is

$$\hat{y} = \frac{a + b + c}{3}.$$

Figure 3 shows the schematic diagram of the similarity and fuzzy arithmetic-based fuzzy inference system (SFAFIS) that we have described so far.

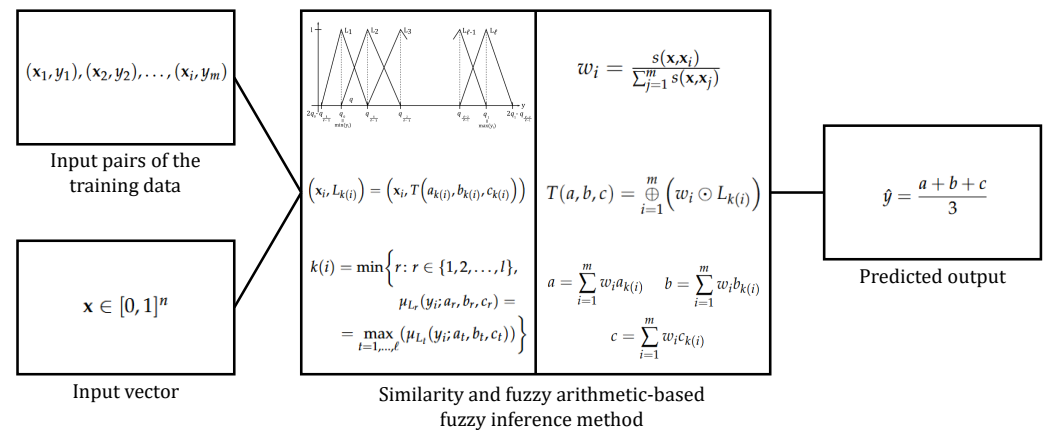


Figure 3. Schematic diagram of the similarity and fuzzy arithmetic-based fuzzy inference system (SFAFIS).

4.3. Tuning the Hyperparameters

The similarity and fuzzy arithmetic-based fuzzy inference system presented above may be treated as a two-parameter function $f_{\lambda, \ell}: [0, 1]^n \rightarrow \mathbb{R}$ that models the relationship between $\mathbf{x}_i$ and y_i , i.e.,

$$y_i \approx f_{\lambda, \ell}(\mathbf{x}_i), \quad (31)$$

where $(\mathbf{x}_i, y_i)$ is the i th training data pair, $i = 1, 2, \dots, m$. Function $f_{\lambda, \ell}$ has two parameters, a $\lambda > 0$ and an $\ell \in \mathbb{N}$, $\ell \geq 3$, which may be viewed as the hyperparameters of our SFAFIS. Recall that λ is the parameter of the epsilon-function-based normalized fuzzy similarity measure and ℓ is the number of linguistic levels utilized in the SFAFIS. The optimal values of λ and ℓ , denoted by λ_{opt} and ℓ_{opt} , respectively, can be determined by solving the minimization problem

$$\sum_{i=1}^m (y_i - f_{\lambda, \ell}(\mathbf{x}_i))^2 \rightarrow \min.$$

That is, λ_{opt} and ℓ_{opt} are parameter values for which

$$\sum_{i=1}^m (y_i - f_{\lambda_{\text{opt}}, \ell_{\text{opt}}}(\mathbf{x}_i))^2 = \min_{\substack{\lambda \in (0, \infty) \\ \ell \in \mathbb{N}, \ell \geq 3}} \left(\sum_{i=1}^m (y_i - f_{\lambda, \ell}(\mathbf{x}_i))^2 \right) \quad (32)$$

holds.

The quasi-optimal values of λ and ℓ can be found using various numerical methods such as particle swarm optimization or genetic algorithms. In a case study, which we will present in the next section, we are going to demonstrate that, since there are just two tunable parameters in our SFAFIS, even a simple grid search method is an effective way for finding the quasi-optimal values of these parameters.

We should emphasize that although our inference system has only two tunable parameters, in practice, this may serve as a viable alternative to much more sophisticated multi-parametric modeling methods.

5. Case Study

In this case study, we demonstrate how our similarity and fuzzy arithmetic-based model can be applied in practice. The aim of this study was to predict the number of passengers on railway relations using real-life data. For this purpose, historical passenger traffic data of the Hungarian State Railways (MÁV) were collected.

In order to utilize the SFAFIS method, we had to identify those characteristics of the examined relations that mostly affect the number of passengers on these relations. That is, we had to identify the components of an input vector $\mathbf{x}$. Based on the findings of previous research articles (see, e.g., [14,36–43]) and preliminary studies of Hungarian State Railways, the following characteristics of railway routes were taken in to account as explanatory variables:

- p_1 : Population of the the first city.
- p_2 : Population of the the second city.
- l : Distance between the cities (km).
- t_r : Fastest possible travel time between the cities by railway (min).
- t_b : Fastest possible travel time between the cities by bus (min).
- t_c : Fastest possible travel time between the cities by car (min).
- f_{IC} : Number of average InterCity train departures between the cities.
- f_b : Number of average bus departures between the cities.
- f_r : Number of average train departures between the cities.
- c_r : Unit cost of transport by rail (per passenger-km) (HUF).
- c_b : Unit cost of transport by bus (per passenger-km) (HUF).
- c_c : Unit cost of transport by car (per passenger-km) (HUF).
- Ico : Car ownership index of the region.
- d_r : Distance of the railway station from the city centers (km).
- d_b : Distance of the bus station from the city centers (km).

Since the above fifteen input characteristics were taken into consideration, an input vector $\mathbf{x} \in [0, 1]^{15}$ contained the normalized values of the above characteristics. The normalization of the inputs was performed using the well-known min–max normalization method on the data given in Appendix A.

For the sake of simplicity, we stipulated that the railway connections are examined independently, i.e., the number of passengers on different sections of the same route are not added up. Using the company’s sales database for the year 2022, we calculated the rounded value of the average monthly number of passengers for each examined connection. These are shown in column y in Appendix A. Here, we utilized data from 60 railway connections in Hungary. However, it turned out, that even with these real data, the data range is too large; therefore, more data would be required to accurately test the proposed model. Unfortunately, because of the size and usual passenger flows in Hungary, more real-world data could not be gathered; therefore, we selected the densest part of our database (i.e., connections with an average monthly passenger number between 800 and 3500) and generated more records based on the original data. First, we generated new records from the original ones such that each new record was derived as a weighted average of two original records with a randomly selected average ratio. Next, using the same approach, we generated records using the previously generated ones. In Appendix A, we summarize our database, which includes both original and generated data and has 81 records in total. The original records are marked by letters from A to Z, while generated data are marked by a combination of letters corresponding to the records that they were generated from. For example, record AB was generated from records A and B and record ABBC was generated from the records AB and BC.

We evaluated the performance of the SFAFIS method and compared it with some well-known methods using randomly selected training and test data sets from the sample S given in Appendix A. We performed the method evaluations repeatedly so that, in every iteration, the training data set S_{tr} and the test data set S_{te} were randomly selected from the sample S as

$$S_{tr} = \left\{ (\mathbf{x}_{i_1}^{(tr)}, y_{i_1}^{(tr)}), (\mathbf{x}_{i_2}^{(tr)}, y_{i_2}^{(tr)}), \dots, (\mathbf{x}_{i_{n_{tr}}}^{(tr)}, y_{i_{n_{tr}}}^{(tr)}) \right\},$$

$$S_{te} = \left\{ (x_{j_1}^{(te)}, y_{j_1}^{(te)}), (x_{j_2}^{(te)}, y_{j_2}^{(te)}), \dots, (x_{j_{n_{te}}}^{(te)}, y_{j_{n_{te}}}^{(te)}) \right\},$$

such that $S_{tr} \cap S_{te} = \emptyset$ and $S_{tr} \cup S_{te} = S$, where

- n_{tr} and n_{te} are the number of training and test data pairs, respectively;
- $x_{i_u}^{(tr)}$ and $x_{j_v}^{(te)}$ are the n -dimensional normalized feature vectors of the i_u th and j_v th railway connections in the training and test data sets, respectively (i.e., $x_{i_u}^{(tr)}, x_{j_v}^{(te)} \in [0, 1]^n$);
- $y_{i_u}^{(tr)}$ and $y_{j_v}^{(te)}$ are the average number of passengers of the i_u th and j_v th railway relations in the training and test data sets, respectively;
- $\hat{y}_{i_u}^{(tr)}$ and $\hat{y}_{j_v}^{(te)}$ are the estimated (computed) values of $y_{i_u}^{(tr)}$ and $y_{j_v}^{(te)}$, respectively;
- $u = 1, 2, \dots, n_{tr}$ and $v = 1, 2, \dots, n_{te}$.

To characterize the goodness of the investigated methods, the following Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) metrics were taken into account:

$$MAPE_{tr} = \frac{100}{n_{tr}} \sum_{u=1}^{n_{tr}} \left| \frac{y_{i_u}^{(tr)} - \hat{y}_{i_u}^{(tr)}}{y_{i_u}^{(tr)}} \right|, \quad (33)$$

$$MAPE_{te} = \frac{100}{n_{te}} \sum_{v=1}^{n_{te}} \left| \frac{y_{j_v}^{(te)} - \hat{y}_{j_v}^{(te)}}{y_{j_v}^{(te)}} \right|, \quad (34)$$

$$MAPE_{tr,te} = \frac{100}{n_{tr} + n_{te}} \left(\sum_{u=1}^{n_{tr}} \left| \frac{y_{i_u}^{(tr)} - \hat{y}_{i_u}^{(tr)}}{y_{i_u}^{(tr)}} \right| + \sum_{v=1}^{n_{te}} \left| \frac{y_{j_v}^{(te)} - \hat{y}_{j_v}^{(te)}}{y_{j_v}^{(te)}} \right| \right), \quad (35)$$

$$RMSE_{tr} = \sqrt{\frac{1}{n_{tr}} \sum_{u=1}^{n_{tr}} \left(y_{i_u}^{(tr)} - \hat{y}_{i_u}^{(tr)} \right)^2}, \quad (36)$$

$$RMSE_{te} = \sqrt{\frac{1}{n_{te}} \sum_{v=1}^{n_{te}} \left(y_{j_v}^{(te)} - \hat{y}_{j_v}^{(te)} \right)^2}, \quad (37)$$

$$RMSE_{tr,te} = \sqrt{\frac{1}{n_{tr} + n_{te}} \left(\sum_{u=1}^{n_{tr}} \left(y_{i_u}^{(tr)} - \hat{y}_{i_u}^{(tr)} \right)^2 + \sum_{v=1}^{n_{te}} \left(y_{j_v}^{(te)} - \hat{y}_{j_v}^{(te)} \right)^2 \right)}. \quad (38)$$

Here, $MAPE_{tr}$ and $MAPE_{te}$ are the MAPE metrics for the training and test data sets, respectively, while $MAPE_{tr,te}$ is the MAPE value that takes into account both the training and test data sets. Similarly, $RMSE_{tr}$ and $RMSE_{te}$ are the RMSE metrics for the training and test data sets, respectively, while $RMSE_{tr,te}$ is the RMSE value that takes into consideration both the training and test data sets.

Notice that using the Equations (33), (34), (36) and (37), the $MAPE_{tr,te}$ and $RMSE_{tr,te}$ given in Equation (35) and Equation (38), respectively, can also be calculated as

$$MAPE_{tr,te} = \frac{n_{tr} MAPE_{tr} + n_{te} MAPE_{te}}{n_{tr} + n_{te}} \quad (39)$$

$$RMSE_{tr,te} = \sqrt{\frac{n_{tr} RMSE_{tr}^2 + n_{te} RMSE_{te}^2}{n_{tr} + n_{te}}}. \quad (40)$$

In every iteration, the number of randomly selected training data points was $n_{tr} = 66$, while the remaining $n_{te} = 15$ data points were utilized as test data.

For every randomly selected training data set, the proposed SFAFIS method was applied with $\lambda = 30$, $\lambda = 60$ and $\lambda = 90$ parameter values. For each value of the λ parameter, the quasi-optimal value of the number of linguistic variables, i.e., ℓ_{opt} , was

determined by minimizing the $\text{RMSE}_{\text{tr}}^2$ quantity as described in Section 4.3, such that ℓ_{opt} was searched for in the set $\{5, 6, \dots, 200\}$. That is, for a given value of λ , ℓ_{opt} is the value of parameter ℓ for which

$$\sum_{u=1}^{n_{\text{tr}}} \left(y_{i_u}^{(\text{tr})} - f_{\lambda, \ell_{\text{opt}}}(\mathbf{x}_{i_u}^{(\text{tr})}) \right)^2 = \min_{\ell \in \{5, 6, \dots, 200\}} \left(\sum_{u=1}^{n_{\text{tr}}} \left(y_{i_u}^{(\text{tr})} - f_{\lambda, \ell}(\mathbf{x}_{i_u}^{(\text{tr})}) \right)^2 \right),$$

where $(\mathbf{x}_{i_u}^{(\text{tr})}, y_{i_u}^{(\text{tr})})$ is the i_u th data pair in the training data set and $f_{\lambda, \ell}(\mathbf{x}_{i_u}^{(\text{tr})})$ is the output of the SFAFIS method for input vector $\mathbf{x}_{i_u}^{(\text{tr})}$.

With the *genfis* method of the MATLAB (R2024b) software package [31], in every iteration, utilizing the training data, a Sugeno-type fuzzy inference system (FIS) was generated using the subtractive clustering (SC) algorithm. This Sugeno-type FIS had 15 inputs ($\mathbf{x} \in [0, 1]^{15}$) with 8 Gaussian membership functions for each input, i.e., the number of parameters on the input side was $15 \times 8 \times 2 = 240$. The number of rules in the rule base, i.e., the number of clusters formed, was eight, and so the output side of the FIS consisted of eight linear functions of the inputs. That is, each of these linear functions had 16 parameters (number of dimensions in the input vectors plus one for the constant term in the linear function). Hence, the total number of tunable parameters in such a Sugeno-type FIS was $15 \times 8 \times 2 + 8 \times 16 = 368$. These parameters were tuned (optimized) using the following techniques, along with a $k = 5$ -fold cross validation:

- Adaptive Neuro Fuzzy Inference System (ANFIS) (see, e.g., [24,25]);
- Genetic Algorithm (GA) (see, e.g., [26–28]);
- Particle Swarm Optimization (PSO) (see, e.g., [29]);
- Pattern-search (PS) method (see, e.g., [30]).

The parameter tuning (optimization) of the Sugeno-type FIS was performed using the *tunefis* method of the MATLAB (R2024b) software package. The modeling goodness of the SFAFIS method and that of the Sugeno-type FIS tuned by the above-listed methods were expressed in terms of MAPE and RMSE metrics. These goodness results are shown in Tables 1–4. The MATLAB source codes that were used to generate the results in these tables are available at <https://jonast.web.elte.hu/>. The ‘SFAFIS_example.m’ file contains the example of our case study, and the ‘FIS_similarity_and_fuzzy_arithmetic_class.m’ file is the class file in which the method itself is implemented. Using these two files, the results of our case study can be reproduced. We should note that in the MATLAB implementation of SFAFIS, instead of the similarity measure $s_{\lambda}(\mathbf{x}, \mathbf{y})$ given in Equation (26), the following modified variant was employed:

$$s'_{\lambda}(\mathbf{x}, \mathbf{y}) = \left(\frac{1 - d(\mathbf{x}, \mathbf{y})}{1 + d(\mathbf{x}, \mathbf{y})} \right)^{\lambda},$$

i.e., $s'_{\lambda}(\mathbf{x}, \mathbf{y}) = s_{2\lambda}(\mathbf{x}, \mathbf{y})$.

Using the goodness results shown in Tables 1–4, Mann–Whitney median tests were conducted to compare the modeling performance of the SFAFIS method with that of the Sugeno-type FIS tuned using the ANFIS, GA, PSO and PS methods. From the three SFAFIS metric values (for $\lambda = 30, 60$ and 90 parameter values) corresponding to each iteration in Tables 1–4, the best one always was taken into account in the hypothesis tests. The results of the Mann–Whitney median tests are shown in Table 5. In this table, the notation m_M stands for the median of the corresponding metric for method M.

Table 1. Modeling results of SFAFIS and the Sugeno-type FIS tuned using ANFIS.

Iteration	Method	MAPE _{tr}	MAPE _{te}	MAPE _{tr,te}	RMSE _{tr}	RMSE _{te}	RMSE _{tr,te}	ℓ_{opt}	λ
1	SFAFIS	0.2202	4.4183	0.99761	5.0363	135.89	58.654	123	30
		0.15377	4.7255	1.0004	4.213	139.69	60.233	128	60
		0.15324	4.8394	1.021	4.2055	141.68	61.087	128	90
2	ANFIS	0.0014637	12.094	2.2407	0.04118	302.94	130.37	-	-
2	SFAFIS	0.22453	5.8265	1.2619	5.0409	126.06	54.438	119	30
		0.18751	5.186	1.1132	4.4291	130.86	56.454	119	60
		0.18602	4.9695	1.0719	4.4058	133.72	57.682	119	90
3	ANFIS	0.0015201	9.7151	1.8003	0.08549	197.57	85.02	-	-
3	SFAFIS	0.32908	3.6451	0.94316	9.2391	72.864	32.446	101	30
		0.1394	2.8135	0.6346	3.8243	65.986	28.605	116	60
		0.13703	2.5159	0.57755	3.806	63.059	27.353	116	90
4	ANFIS	0.0037684	6.3676	1.1823	0.097153	144.38	62.132	-	-
4	SFAFIS	0.22989	2.0713	0.57088	5.4742	69.623	30.366	129	30
		0.16468	1.4503	0.40276	3.833	52.779	22.974	129	60
		0.16229	1.3263	0.37785	3.8257	45.738	19.983	129	90
5	ANFIS	0.0019279	4.5712	0.84808	0.11433	120.42	51.821	-	-
5	SFAFIS	0.32783	0.72639	0.40164	9.5702	15.558	10.929	125	30
		0.1617	0.72237	0.26553	4.1248	19.621	9.2282	125	60
		0.16172	0.76868	0.27412	4.1371	20.693	9.6562	125	90
6	ANFIS	0.0016693	2.6455	0.49127	0.045098	68.991	29.689	-	-
6	SFAFIS	0.25164	5.5006	1.2237	5.8787	98.651	42.783	118	30
		0.1556	4.2795	0.91928	3.9032	79.832	34.535	118	60
		0.15255	3.8394	0.8353	3.8559	71.451	30.944	118	90
7	ANFIS	0.0013355	5.9663	1.106	0.032256	90.34	38.876	-	-
7	SFAFIS	0.29038	3.0833	0.80759	8.4814	116.93	50.897	112	30
		0.17055	2.9978	0.69411	4.4023	114.11	49.267	132	60
		0.16441	3.1599	0.71913	4.2877	115.73	49.951	132	90
8	ANFIS	0.0049698	4.4572	0.82945	0.51938	96.282	41.436	-	-
8	SFAFIS	0.3094	2.9724	0.80254	8.984	106.33	46.47	117	30
		0.1567	1.4212	0.39087	3.6034	47.162	20.554	129	60
		0.15741	0.98954	0.31151	3.6139	41.72	18.248	129	90
9	ANFIS	0.0021185	8.9088	1.6515	0.066603	236.31	101.69	-	-
9	SFAFIS	0.30447	3.7847	0.94895	8.9396	65.489	29.315	113	30
		0.16889	1.4384	0.40398	4.0767	41.3	18.15	117	60
		0.15815	1.4636	0.39989	3.9599	43.411	19.02	117	90
10	ANFIS	0.0013597	8.3017	1.5385	0.054377	146.21	62.918	-	-
10	SFAFIS	0.19831	4.9907	1.0858	5.0846	129.59	55.955	125	30
		0.16201	4.5909	0.98218	4.291	138.01	59.518	125	60
		0.16043	4.6778	0.99697	4.3038	139.36	60.095	125	90
11	ANFIS	0.0014798	6.9232	1.2833	0.047309	140.56	60.487	-	-

the ANFIS, GA, PSO and PS methods both in terms of MAPE and RMSE (see the Mann–Whitney test results for $\text{MAPE}_{\text{te,tr}}$ and $\text{RMSE}_{\text{te,tr}}$ in Table 5).

Noting the modeling results shown in Tables 1–4, in most of the cases, the MAPE_{te} metric for the SFAFIS method is less than 5. This means that using the above-considered characteristics of a new railway connection, the SFAFIS method can predict the average number of passengers on this connection with an absolute relative error of less than 5%. This accuracy exceeds the practical expectations.

Table 2. Modeling results of SFAFIS and the Sugeno-type FIS tuned using genetic algorithm (GA).

Iteration	Method	MAPE_{tr}	MAPE_{te}	$\text{MAPE}_{\text{tr,te}}$	RMSE_{tr}	RMSE_{te}	$\text{RMSE}_{\text{tr,te}}$	ℓ_{opt}	λ
1	SFAFIS	0.2202	4.4183	0.99761	5.0363	135.89	58.654	123	30
		0.15377	4.7255	1.0004	4.213	139.69	60.233	128	60
		0.15324	4.8394	1.021	4.2055	141.68	61.087	128	90
	GA	5.939×10^{-13}	54.011	10.002	1.2304×10^{-11}	1612.8	694.02	-	-
2	SFAFIS	0.31285	4.9046	1.1632	8.9279	124.97	54.379	118	30
		0.16985	3.5869	0.80263	4.5693	82.416	35.705	126	60
		0.15862	3.5916	0.79436	4.4574	87.145	37.717	126	90
	GA	8.2149×10^{-13}	30.979	5.7368	1.7463×10^{-11}	584.05	251.34	-	-
3	SFAFIS	0.31248	2.1682	0.65613	8.9226	60.615	27.3	119	30
		0.17735	1.7383	0.46642	4.4278	32.421	14.513	119	60
		0.17931	1.8002	0.47946	4.4923	34.058	15.207	119	90
	GA	3.9052×10^{-13}	37.424	6.9304	7.9338×10^{-12}	933.91	401.89	-	-
4	SFAFIS	0.266	3.1635	0.80258	8.4559	86.529	38.01	119	30
		0.14044	3.2805	0.72193	3.9928	108.97	47.032	128	60
		0.13474	3.5954	0.77561	3.9833	119.37	51.495	128	90
	GA	1.0777×10^{-12}	107.74	19.953	2.7237×10^{-11}	2169.8	933.72	-	-
5	SFAFIS	0.21337	3.796	0.87681	5.0801	177	76.305	118	30
		0.15085	4.1158	0.8851	4.0728	178.61	76.949	118	60
		0.14864	4.2597	0.90996	4.0587	188.58	81.236	118	90
	GA	2.3839×10^{-13}	23.942	4.4337	4.6876×10^{-12}	507.75	218.5	-	-
6	SFAFIS	0.20716	4.8887	1.0741	5.0419	120.99	52.264	129	30
		0.15365	4.9371	1.0395	4.305	131.88	56.884	128	60
		0.15182	4.9114	1.0332	4.2387	135.37	58.378	128	90
	GA	4.1628×10^{-13}	33.713	6.2431	7.9228×10^{-12}	748.46	322.09	-	-
7	SFAFIS	0.28009	2.9498	0.77448	8.277	55.694	25.104	117	30
		0.15527	1.8633	0.47157	4.0176	37.824	16.676	110	60
		0.15237	1.7295	0.44444	3.9751	32.959	14.63	132	90
	GA	1.0958×10^{-13}	48.42	8.9667	2.4363×10^{-12}	1099.8	473.27	-	-
8	SFAFIS	0.20807	6.7026	1.4108	4.7065	427.77	184.13	123	30
		0.16001	7.7727	1.5698	3.8988	473.47	203.78	123	60
		0.15815	7.88	1.5881	3.8891	477.16	205.37	123	90
	GA	1.3733×10^{-13}	20.04	3.7112	3.2496×10^{-12}	797.99	343.4	-	-
9	SFAFIS	0.29691	6.4588	1.438	8.9759	129.42	56.282	109	30
		0.1628	3.7829	0.83319	4.2356	73.225	31.742	132	60
		0.15557	3.1536	0.71077	4.1189	65.904	28.603	132	90
	GA	1.9898×10^{-13}	16.164	2.9932	5.7774×10^{-12}	476.88	205.22	-	-
10	SFAFIS	0.28708	1.2686	0.46884	8.1722	39.502	18.53	118	30
		0.15661	0.99323	0.31154	3.6767	40.177	17.605	129	60
		0.16178	1.0292	0.32242	3.7043	41.457	18.151	129	90
	GA	1.8214×10^{-13}	14.027	2.5976	4.2808×10^{-12}	413.89	178.11	-	-

Table 6 shows the average training and inference times of the investigated methods examined on the same computer under the same conditions. In this table, the average inference time represents the mean value of the time needed to predict both the training

and test output data for the iterations shown in Tables 1–4. In every iteration, from the three SFAFIS methods (for $\lambda = 30, 60$ and 90 parameter values), the one with the best $RMSE_{tr,te}$ value was taken into account in computing the average inference time of the SFAFIS method.

Table 3. Modeling results of SFAFIS and the Sugeno-type FIS tuned using particle swarm optimization (PSO).

Iteration	Method	MAPE _{tr}	MAPE _{te}	MAPE _{tr,te}	RMSE _{tr}	RMSE _{te}	RMSE _{tr,te}	ℓ_{opt}	λ
1	SFAFIS	0.2202	4.4183	0.99761	5.0363	135.89	58.654	123	30
		0.15377	4.7255	1.0004	4.213	139.69	60.233	128	60
		0.15324	4.8394	1.021	4.2055	141.68	61.087	128	90
	PSO	5.939×10^{-13}	54.011	10.002	1.2304×10^{-11}	1612.8	694.02	-	-
2	SFAFIS	0.30661	2.6871	0.74744	9.3238	104.95	45.943	115	30
		0.16299	1.4458	0.40054	4.3656	49.901	21.832	132	60
		0.1574	1.173	0.34548	4.268	46.291	20.289	132	90
	PSO	4.1799×10^{-13}	32.744	6.0637	1.0934×10^{-11}	967.97	416.55	-	-
3	SFAFIS	0.18747	4.5946	1.0036	4.7889	110.26	47.645	121	30
		0.18007	4.511	0.9821	4.3273	122.34	52.792	105	60
		0.17921	4.2945	0.94131	4.3083	129.23	55.75	105	90
	PSO	1.0973×10^{-13}	8.6157	1.5955	2.0855×10^{-12}	183.57	78.995	-	-
4	SFAFIS	0.21829	9.7979	1.9923	4.9939	440.71	189.71	110	30
		0.16126	10.554	2.0858	4.1652	488.16	210.1	132	60
		0.15819	10.75	2.1196	4.1102	492.26	211.87	132	90
	PSO	6.0409×10^{-13}	38.055	7.0473	1.2908×10^{-11}	842.16	362.41	-	-
5	SFAFIS	0.19639	11.809	2.3468	4.5202	453.9	195.37	116	30
		0.1456	11.998	2.3404	3.5729	495.23	213.14	116	60
		0.1434	12.038	2.3461	3.5365	498.18	214.41	116	90
	PSO	2.7137×10^{-13}	48.339	8.9517	5.7643×10^{-12}	1072.2	461.42	-	-
6	SFAFIS	0.33054	3.5667	0.92982	9.8091	70.693	31.684	108	30
		0.16586	3.2494	0.73688	4.1939	74.449	32.261	126	60
		0.15436	3.6957	0.81016	4.0768	84	36.334	126	90
	PSO	1.7145×10^{-13}	18.289	3.3869	3.4361×10^{-12}	293.82	126.44	-	-
7	SFAFIS	0.28899	9.9587	2.0797	9.5004	431.5	185.89	126	30
		0.1397	9.593	1.8903	3.8897	475.58	204.69	126	60
		0.13334	9.4787	1.864	3.8658	479.45	206.35	126	90
	PSO	1.5571×10^{-13}	23.123	4.282	3.8223×10^{-12}	486.1	209.18	-	-
8	SFAFIS	0.19582	8.9888	1.8241	4.5347	441.01	189.82	109	30
		0.15987	8.5128	1.7067	4.129	484.7	208.62	132	60
		0.15786	8.5872	1.7188	4.1076	489.37	210.62	132	90
	PSO	3.6489×10^{-13}	31.322	5.8003	6.6528×10^{-12}	1599.4	688.29	-	-
9	SFAFIS	0.25163	6.8997	1.4827	7.7123	388.25	167.22	124	30
		0.14193	5.8837	1.2052	3.3971	461.81	198.75	124	60
		0.14254	6.0215	1.2312	3.4167	475.42	204.61	124	90
	PSO	4.8984×10^{-13}	58.248	10.787	8.6798×10^{-12}	1805.7	777.05	-	-
10	SFAFIS	0.19379	7.9628	1.6325	4.5137	394.07	169.63	132	30
		0.14739	8.1986	1.6384	4.0815	470.95	202.7	116	60
		0.14556	8.5111	1.6947	4.0623	486.14	209.23	116	90
	PSO	3.1618×10^{-13}	31.842	5.8967	8.6715×10^{-12}	1804	776.3	-	-

Based on the average training and inference times shown in Table 6, we see that, for the data utilized in our case study, the average training time for the SFAFIS method was similar to that of the ANFIS method, and the average training times of these two methods (2.288 and 3.695 s) were considerably shorter than the average training times of the other three methods. At the same time, the average inference time for the SFAFIS method turned out to be longer than that of the other investigated methods. Here, we should emphasize that

since we aim to predict the number of passengers on railway connections and do not need to do this real-time, neither the training time nor the inference time of the method used is critical.

Table 4. Modeling results of SFAFIS and the Sugeno-type FIS tuned using pattern-search (PS) optimization.

Iteration	Method	MAPE _{tr}	MAPE _{te}	MAPE _{tr,te}	RMSE _{tr}	RMSE _{te}	RMSE _{tr,te}	ℓ_{opt}	λ
1	SFAFIS	0.2202	4.4183	0.99761	5.0363	135.89	58.654	123	30
		0.15377	4.7255	1.0004	4.213	139.69	60.233	128	60
		0.15324	4.8394	1.021	4.2055	141.68	61.087	128	90
	PS	5.939×10^{-13}	54.011	10.002	1.2304×10^{-11}	1612.8	694.02	-	-
2	SFAFIS	0.23605	11.846	2.3861	5.8128	159.09	68.663	108	30
		0.15636	6.8386	1.3938	3.8409	121.46	52.384	119	60
		0.18602	4.9695	1.0719	4.4058	133.72	57.682	119	90
	PS	1.5614×10^{-13}	26.582	4.9227	3.8216×10^{-12}	482.6	207.68	-	-
3	SFAFIS	0.28364	8.512	1.8074	8.396	434.57	187.16	112	30
		0.1476	7.8513	1.5742	3.7915	474.26	204.12	126	60
		0.13703	2.5159	0.57755	3.806	63.059	27.353	116	90
	PS	1.5834×10^{-13}	41.438	7.6738	3.239×10^{-12}	782.9	336.91	-	-
4	SFAFIS	0.28301	8.1045	1.7314	8.5418	406.78	175.22	109	30
		0.15165	6.4616	1.3202	3.6838	464.64	199.97	132	60
		0.16229	1.3263	0.37785	3.8257	45.738	19.983	129	90
	PS	3.797×10^{-13}	73.306	13.575	7.2282×10^{-12}	1413.5	608.27	-	-
5	SFAFIS	0.29353	2.2379	0.65361	8.6	98.406	43.053	119	30
		0.15202	1.4259	0.38793	4.658	61.988	27.005	124	60
		0.16172	0.76868	0.27412	4.1371	20.693	9.6562	125	90
	PS	2.494×10^{-13}	52.274	9.6804	5.0776×10^{-12}	1196.8	515.02	-	-
6	SFAFIS	0.20946	4.5473	1.0128	4.9107	127.29	54.954	132	30
		0.16973	4.8916	1.0441	4.3344	140.26	60.484	132	60
		0.15255	3.8394	0.8353	3.8559	71.451	30.944	118	90
	PS	3.2409×10^{-13}	24.475	4.5324	5.7138×10^{-12}	926.2	398.57	-	-
7	SFAFIS	0.27993	2.9145	0.76781	8.0152	93.395	40.837	129	30
		0.13906	3.2011	0.7061	3.6172	111.41	48.054	128	60
		0.16441	3.1599	0.71913	4.2877	115.73	49.951	132	90
	PS	3.7447×10^{-13}	21.939	4.0628	7.1633×10^{-12}	542.97	233.66	-	-
8	SFAFIS	0.30154	4.6014	1.0978	8.8431	420.07	180.95	124	30
		0.17052	5.0451	1.0732	3.8546	459.5	197.77	123	60
		0.15741	0.98954	0.31151	3.6139	41.72	18.248	129	90
	PS	1.8569×10^{-13}	11.627	2.1531	3.8067×10^{-12}	265.96	114.45	-	-
9	SFAFIS	0.28664	4.8999	1.1409	8.684	117.96	51.363	117	30
		0.13164	2.2561	0.52507	3.2993	54.869	23.799	128	60
		0.15815	1.4636	0.39989	3.9599	43.411	19.02	117	90
	PS	2.6935×10^{-13}	31.185	5.7749	6.8908×10^{-12}	1237.4	532.5	-	-
10	SFAFIS	0.21633	5.5482	1.2037	5.1291	132.86	57.36	132	30
		0.16746	4.3179	0.93606	4.2094	133.92	57.754	132	60
		0.16043	4.6778	0.99697	4.3038	139.36	60.095	125	90
	PS	5.4371×10^{-13}	19.465	3.6046	1.1614×10^{-11}	402.97	173.41	-	-

Table 5. Results of Mann–Whitney median tests for the modeling results of SFAFIS and the Sugeno-type FIS tuned using the ANFIS, GA, PSO and PS methods and the samples in Tables 1–4.

Metric	Null and Alternative Hypotheses and p -Values				
MAPE _{tr}	H ₀ :	$m_{\text{SFAFIS}} = m_{\text{ANFIS}}$	$m_{\text{SFAFIS}} = m_{\text{GA}}$	$m_{\text{SFAFIS}} = m_{\text{PSO}}$	$m_{\text{SFAFIS}} = m_{\text{PS}}$
	H ₁ :	$m_{\text{SFAFIS}} > m_{\text{ANFIS}}$	$m_{\text{SFAFIS}} > m_{\text{GA}}$	$m_{\text{SFAFIS}} > m_{\text{PSO}}$	$m_{\text{SFAFIS}} > m_{\text{PS}}$
	p -value:	0.000	0.000	0.000	0.000
MAPE _{te}	H ₀ :	$m_{\text{SFAFIS}} = m_{\text{ANFIS}}$	$m_{\text{SFAFIS}} = m_{\text{GA}}$	$m_{\text{SFAFIS}} = m_{\text{PSO}}$	$m_{\text{SFAFIS}} = m_{\text{PS}}$
	H ₁ :	$m_{\text{SFAFIS}} < m_{\text{ANFIS}}$	$m_{\text{SFAFIS}} < m_{\text{GA}}$	$m_{\text{SFAFIS}} < m_{\text{PSO}}$	$m_{\text{SFAFIS}} < m_{\text{PS}}$
	p -value:	0.001	0.000	0.000	0.000
MAPE _{tr,te}	H ₀ :	$m_{\text{SFAFIS}} = m_{\text{ANFIS}}$	$m_{\text{SFAFIS}} = m_{\text{GA}}$	$m_{\text{SFAFIS}} = m_{\text{PSO}}$	$m_{\text{SFAFIS}} = m_{\text{PS}}$
	H ₁ :	$m_{\text{SFAFIS}} < m_{\text{ANFIS}}$	$m_{\text{SFAFIS}} < m_{\text{GA}}$	$m_{\text{SFAFIS}} < m_{\text{PSO}}$	$m_{\text{SFAFIS}} < m_{\text{PS}}$
	p -value:	0.003	0.000	0.000	0.000
RMSE _{tr}	H ₀ :	$m_{\text{SFAFIS}} = m_{\text{ANFIS}}$	$m_{\text{SFAFIS}} = m_{\text{GA}}$	$m_{\text{SFAFIS}} = m_{\text{PSO}}$	$m_{\text{SFAFIS}} = m_{\text{PS}}$
	H ₁ :	$m_{\text{SFAFIS}} > m_{\text{ANFIS}}$	$m_{\text{SFAFIS}} > m_{\text{GA}}$	$m_{\text{SFAFIS}} > m_{\text{PSO}}$	$m_{\text{SFAFIS}} > m_{\text{PS}}$
	p -value:	0.000	0.000	0.000	0.000
RMSE _{te}	H ₀ :	$m_{\text{SFAFIS}} = m_{\text{ANFIS}}$	$m_{\text{SFAFIS}} = m_{\text{GA}}$	$m_{\text{SFAFIS}} = m_{\text{PSO}}$	$m_{\text{SFAFIS}} = m_{\text{PS}}$
	H ₁ :	$m_{\text{SFAFIS}} < m_{\text{ANFIS}}$	$m_{\text{SFAFIS}} < m_{\text{GA}}$	$m_{\text{SFAFIS}} < m_{\text{PSO}}$	$m_{\text{SFAFIS}} < m_{\text{PS}}$
	p -value:	0.006	0.000	0.002	0.000
RMSE _{tr,te}	H ₀ :	$m_{\text{SFAFIS}} = m_{\text{ANFIS}}$	$m_{\text{SFAFIS}} = m_{\text{GA}}$	$m_{\text{SFAFIS}} = m_{\text{PSO}}$	$m_{\text{SFAFIS}} = m_{\text{PS}}$
	H ₁ :	$m_{\text{SFAFIS}} < m_{\text{ANFIS}}$	$m_{\text{SFAFIS}} < m_{\text{GA}}$	$m_{\text{SFAFIS}} < m_{\text{PSO}}$	$m_{\text{SFAFIS}} < m_{\text{PS}}$
	p -value:	0.006	0.000	0.003	0.000

Table 6. Training and inference times vs. method.

Method	Average Training Time (s)	Average Inference Time (s)
SFAFIS	2.288	0.018157
Sugeno-type FIS tuned by ANFIS	3.695	0.003318
Sugeno-type FIS tuned by GA	545.985	0.005004
Sugeno-type FIS tuned by PSO	310.319	0.003858
Sugeno-type FIS tuned by PS	1185.936	0.005838

Possibilities for Applying the SFAFIS Method to Other Railway Systems

As mentioned before, in our case study, fifteen characteristics of railway connections were taken into consideration, and so the input vector $\mathbf{x} \in [0, 1]^{15}$ contained the normalized values of these characteristics. Here, we should note that these characteristics were identified based on previous research articles (see, e.g., [14,36–43]) and preliminary studies of Hungarian State Railways (MÁV). This means that we identified those characteristics of the examined railway connections that mostly affected the number of passengers on these connections. Hence, these characteristics are specific to the Hungarian railway system in the sense that these inputs are the most explanatory in the given system. At the same time, it is worth noting that the SFAFIS method presented in Section 4 can be applied in the same way to any another railway system, even if the input characteristics in that system are different. That is, once the input features of railway routes, and hence the explanatory variables, in a railway system are identified, and the historical number of passengers associated with each input vector is given, the SFAFIS method can be applied in the same way as described. The number of passengers can be determined from ticket sales data.

6. Conclusions and Plans for Future Research

In this study, we developed a new fuzzy inference method that utilizes a novel similarity measure and fuzzy arithmetic operations to model the relation between inputs and outputs. Using known input–output pairs, for a given input vector, we determine its similarity to all the known input vectors, and from these similarities, we derive weights. Using these weights, the output of our system is computed as the defuzzified value of the weighted average of triangular fuzzy numbers that represent the linguistic values of the

outputs. Based on the results of a real-life case study, which we presented in Section 5, we may conclude that although the proposed method has only two adjustable parameters, its forecasting capability is comparable with that of Sugeno-type fuzzy inference systems, which have a much larger number of adjustable parameters.

In Section 3, we introduced the epsilon-function-based normalized fuzzy similarity measure and demonstrated its main properties. Later, we showed how this similarity measure can be used in the presented SFAFIS method. However, we did not study how the SFAFIS method would perform using other similarity measures. Since comparing the proposed similarity measure with other similarity measures is an interesting topic, we plan to study this in our future research.

We proposed the use of a simple grid search method to find the quasi-optimal values of the λ and ℓ parameters of an SFAFIS system. However, as part of our future research, we plan to investigate how more effective optimization methods can be utilized for this purpose. In the proposed inference system, we perform arithmetic operations over triangular fuzzy numbers. We would also like to examine how our method can be adapted to other types of fuzzy numbers.

In the presented case study, we utilized passenger traffic data from the Hungarian State Railways (MÁV). This naturally raises the question of how the SFAFIS method can be applied to other transportation systems, such as high-speed rail networks, urban metro systems or air routes. We would therefore like to further explore this question in our future research.

In the wider field of transportation planning, in our future research, we would also like to investigate how the choice of transport mode can be modeled using fuzzy methods. Moreover, we are also motivated to answer the question of how and by what evaluation methods limited transport capacities can be allocated on a socioeconomic basis.

Author Contributions: Conceptualization, M.F. and T.J.; methodology, M.F. and T.J.; software, M.F. and T.J.; validation, M.F. and T.J.; formal analysis, M.F. and T.J.; data curation, M.F. and T.J.; writing—original draft preparation, M.F. and T.J.; writing—review and editing, M.F. and T.J.; visualization, M.F. and T.J.; supervision, T.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

Appendix A. The Data Set Used in the Case Study

City 1	City 2	p_1	p_2	l	t_r	t_b	t_c	f_{IC}	f_b	f_r	c_r	c_b	c_c	I_{co}	d_r	d_b	y
A	B	55,910	54,428	86	99	180	100	8	9	24	2520	2520	4356	0.4602	4.5	1.3	824
A	C	55,910	78,190	98	103	230	100	8	17	10	2700	2790	4140	0.4496	4.6	1.6	1759
D	E	55,164	66,061	87	68	185	90	16	19	6	1860	2610	3996	0.4241	2.9	2	1772
F	G	199,858	3530	67	52	116	80	8	10	2	1980	1680	3060	0.3888	2.5	1.5	2459
H	I	8389	116,282	45	39	110	60	9	18	1	1230	1395	1944	0.4055	2.9	1.7	2040
J	K	18,039	23,742	64	73	117	65	8	10	13	1490	1490	3096	0.5117	1.7	1.8	885
L	M	23,116	158,797	91	76	153	80	17	19	10	1860	2180	4104	0.4075	3.5	1.1	2758
N	O	7190	147,533	48	37	130	60	9	23	18	1770	1800	2772	0.3987	2.5	1.1	1390
P	O	15,504	147,533	36	28	65	50	9	23	32	1490	930	1548	0.3987	3.9	1.2	3149
Q	R	60,334	65,830	137	95	250	110	7	8	12	3630	3905	5796	0.4422	1.6	1.3	1521
S	T	1,658,342	24,371	145	165	200	143	1	11	8	3249	3410	7596	0.4044	2.7	0.5	2634
S	U	1,658,342	29,889	130	122	115	90	1	10	25	3480	2830	5616	0.4458	1	1	3325
F	V	199,858	127,599	299	269	680	218	19	0	3	6240	9190	13,608	0.4314	1.9	2	1208
F	M	199,858	158,797	181	192	308	190	0	18	8	4398	3950	8100	0.3894	2.7	1.8	1725
F	W	199,858	49,182	102	160	150	85	0	8	9	2200	2520	4752	0.411	2.7	1.7	839
V	A	127,599	55,910	68	117	105	80	0	8	10	1490	1490	2916	0.45	4.1	1.8	1970
V	E	127,599	66,061	200	185	403	140	20	0	2	4380	5820	8244	0.4326	2.3	1.5	1194
X	E	108,120	66,061	47	49	80	65	17	14	8	1120	1300	2124	0.4406	2.4	1.4	3327
O	W	147,533	49,182	38	75	95	65	8	11	20	1490	1300	2988	0.3995	3.7	1.4	1466
O	V	147,533	127,599	239	242	665	190	17	0	1	4698	7690	11,880	0.4199	2.9	1.7	838
O	T	147,533	24,371	21	27	35	30	0	13	87	550	550	1008	0.3783	4.8	1.2	1904
O	Y	147,533	31,022	39	65	90	60	0	13	23	1120	1120	2160	0.3503	3.3	1.6	1042
I	E	116,282	66,061	145	109	290	150	18	18	4	2399	4720	6552	0.4182	3.2	1.8	1731
Z	V	139,330	127,599	186	226	270	190	8	2	5	5160	4200	8784	0.4328	2	1.8	1160
Z	C	139,330	78,190	179	261	300	210	0	9	5	4620	4200	8640	0.4325	2.5	1.6	1308
A1	C	95,045	78,190	136	145	225	120	8	8	6	3330	3880	5832	0.4627	2.8	1.2	959

City 1	City 2	p_1	p_2	l	t_r	t_b	t_c	f_{IC}	f_b	f_r	c_r	c_b	c_c	I_{co}	d_r	d_b	y
A1	A	95,045	55,910	40	42	55	40	16	21	33	840	840	1692	0.4506	4.8	1.2	3092
C	B	78,190	54,428	47	55	75	55	0	21	18	1330	1120	2016	0.4723	2.5	1.3	2013
AA	BC	55,910	70,567	94	102	214	100	8	14	14	2642	2703	4209	0.453	4.6	1.5	1459
AD	CE	55,551	72,345	93	86	208	95	12	18	8	2295	2703	4071	0.4373	3.8	1.8	1765
DF	EG	102,007	45,817	81	63	163	87	13	16	5	1899	2309	3693	0.4127	2.8	1.8	1994
FH	GI	140,296	38,605	60	48	114	74	8	12	2	1747	1591	2713	0.394	2.6	1.6	2329
HJ	IK	11,953	82,106	52	52	113	62	9	15	5	1326	1430	2369	0.4447	2.5	1.7	1613
JL	KM	20,636	92,825	78	75	135	73	13	15	11	1679	1843	3612	0.4584	2.6	1.4	1843
LN	MO	17,430	154,775	76	62	145	73	14	20	13	1828	2044	3628	0.4044	3.1	1.1	2269
NP	OO	9038	147,533	45	35	116	58	9	23	21	1708	1607	2500	0.3987	2.8	1.1	1781
PQ	OR	48,257	87,840	110	77	200	94	8	12	17	3054	3104	4652	0.4305	2.2	1.3	1960
QS	RT	1220139	35,740	143	146	214	134	3	10	9	3353	3546	7102	0.4148	2.4	0.7	2329
SS	TU	1,658,342	28,825	133	130	131	100	1	10	22	3435	2942	5998	0.4378	1.3	0.9	3192
SF	UV	1,223,448	59,024	180	166	283	128	6	7	18	4303	4726	7999	0.4415	1.3	1.3	2694
FF	VM	199,858	156,814	188	197	332	192	1	17	8	4515	4283	8450	0.392	2.6	1.8	1692
FF	MW	199,858	143,995	170	188	287	176	0	17	8	4101	3757	7648	0.3923	2.7	1.8	1605
FV	WA	160,745	52,824	84	137	126	82	0	8	10	1816	1962	3758	0.4321	3.5	1.8	1451
VV	AE	127,599	61,929	146	157	282	116	12	3	5	3204	4058	6075	0.4397	3	1.6	1510
VX	EE	109,263	66,061	56	57	99	69	17	13	8	1311	1565	2483	0.4401	2.4	1.4	3202
XO	EW	124,506	59,044	43	60	86	65	13	13	13	1274	1300	2483	0.4235	2.9	1.4	2553
OO	WV	147,533	122,287	225	231	626	182	16	1	2	4481	7257	11,278	0.4185	3	1.7	881
OO	VT	147,533	115,007	212	216	588	170	15	2	11	4192	6819	10,554	0.4148	3.1	1.6	968
OO	TY	147,533	30,240	37	61	84	56	0	13	31	1053	1053	2024	0.3536	3.5	1.6	1143
OI	YE	143,063	36,034	54	71	119	73	3	14	20	1303	1635	2788	0.36	3.3	1.6	1141
IZ	EV	117,752	69,987	148	116	289	153	17	17	4	2575	4687	6694	0.4192	3.1	1.8	1695
ZZ	VC	139,330	84,583	180	256	296	207	1	8	5	4690	4200	8659	0.4325	2.4	1.6	1289
ZA1	CC	101,635	78,190	142	162	236	133	7	8	6	3522	3928	6250	0.4582	2.8	1.3	1011

City 1	City 2	p_1	p_2	l	t_r	t_b	t_c	f_{1c}	f_b	f_r	c_r	c_b	c_c	I_{co}	d_r	d_b	y
A1A1	CA	95,045	76,443	128	137	212	114	9	9	8	3135	3642	5507	0.4617	3	1.2	1126
A1C	AB	80,528	54,634	46	53	72	53	2	21	20	1262	1081	1971	0.4693	2.8	1.3	2163
AAAD	BCCE	55,731	71,453	93	94	211	98	10	16	11	2469	2703	4140	0.4452	4.2	1.6	1611
ADDF	CEEG	91,601	51,759	83	68	173	89	13	17	5	1988	2397	3778	0.4182	3	1.8	1943
DFFH	EGGI	121,914	42,068	70	55	137	80	11	14	3	1820	1936	3183	0.403	2.7	1.7	2168
FHHJ	GIJK	49,984	69,215	54	50	113	65	9	14	4	1451	1478	2471	0.4297	2.5	1.7	1825
HJJL	IKKM	17,763	89,278	69	67	128	69	11	15	9	1562	1706	3201	0.4539	2.6	1.5	1767
JLLN	KMMO	17,959	144,544	76	64	143	73	14	19	13	1803	2011	3626	0.4133	3.1	1.2	2199
LNNP	MOOO	14,177	151,968	64	52	133	67	12	21	16	1781	1875	3191	0.4022	3	1.1	2080
NPPQ	OOOR	45,314	92,319	105	74	194	91	8	13	18	2953	2991	4490	0.4281	2.3	1.3	1946
PQQS	ORRT	499,570	67,775	123	103	205	109	6	11	14	3169	3274	5595	0.4244	2.3	1.1	2102
QSSS	RTTU	1,240,160	35,424	142	145	210	132	3	10	10	3357	3518	7052	0.4158	2.3	0.7	2368
SSSF	TUUV	1,402,059	46,622	161	151	221	117	4	8	20	3947	3993	7177	0.44	1.3	1.1	2898
SFFF	UVVM	301,763	147,078	188	194	327	185	2	16	9	4494	4327	8405	0.397	2.5	1.8	1792
FFFF	VMMW	199,858	150,020	179	192	308	183	1	17	8	4296	4004	8025	0.3922	2.7	1.8	1646
FFV	MWWA	165,584	64,105	94	143	146	94	0	9	9	2098	2185	4239	0.4272	3.4	1.8	1470
FVVV	WAAE	142,111	57,943	119	148	213	101	7	5	7	2596	3140	5061	0.4364	3.2	1.7	1484
VVVX	AEEE	111,291	65,604	66	68	119	75	17	12	7	1521	1841	2880	0.44	2.5	1.4	3015
VXXO	EEEW	123,063	59,708	44	60	87	65	14	13	12	1277	1325	2483	0.4251	2.9	1.4	2615
XOOO	EWVW	138,378	97,142	153	163	412	135	15	6	7	3206	4889	7781	0.4205	2.9	1.6	1546
OOOO	WVVT	147,533	118,693	219	223	608	176	16	1	7	4338	7041	10,920	0.4167	3	1.7	924
OOOO	VTTY	147,533	58,030	94	111	249	94	5	9	24	2082	2943	4821	0.3737	3.4	1.6	1086
OOOI	TYYE	143,360	35,648	53	71	116	72	2	14	21	1286	1596	2737	0.3596	3.3	1.6	1141
OIIZ	YEEV	132,995	49,539	91	89	186	105	8	15	14	1809	2849	4342	0.3836	3.2	1.7	1361
IZZZ	EVVC	138,306	83,891	178	250	296	205	2	9	5	4590	4223	8565	0.4319	2.5	1.6	1308
ZZZA1	VCCC	131,028	83,175	172	236	283	191	2	8	5	4433	4140	8128	0.4382	2.5	1.5	1228
ZA1A1A1	CCCA	100,432	77,871	140	158	232	130	7	8	6	3451	3875	6114	0.4588	2.8	1.2	1032
A1A1A1C	CAAB	89,200	67,661	95	103	156	89	6	14	13	2381	2611	4083	0.4648	2.9	1.2	1544

References

1. Dou, F.; Wang, L.; Jia, L. A train dispatching model based on fuzzy passenger demand forecasting during holidays. *J. Ind. Eng. Manag.* **2013**, *6*, 320–335. [\[CrossRef\]](#)
2. Fowkes, A.S.; Nash, C.A.; Whiteing, A.E. Understanding trends in inter-city rail traffic in Great Britain. *Transp. Plan. Technol.* **1985**, *10*, 65–80. [\[CrossRef\]](#)
3. Butkevičius, J.; Mazūra, M.; Ivankovas, V.; Mazūra, S. Analysis and forecast of the dynamics of passenger transportation by public land transport. *Transport* **2004**, *19*, 3–8. [\[CrossRef\]](#)
4. Xiao, Y.; Liu, J.J.; Hu, Y.; Wang, Y.; Lai, K.K.; Wang, S. A neuro-fuzzy combination model based on singular spectrum analysis for air transport demand forecasting. *J. Air Transp. Manag.* **2014**, *39*, 1–11.
5. Srisaeng, P.; Baxter, G.; Wild, G. An adaptive neuro-fuzzy inference system for forecasting Australia's domestic low cost carrier passenger demand. *Aviation* **2015**, *19*, 150–163. [\[CrossRef\]](#)
6. Pamucar, D.; Deveci, M.; Canitez, F.; Božanić, D. A fuzzy Full Consistency Method-Dombi-Bonferroni model for prioritizing transportation demand management measures. *Appl. Soft Comput.* **2020**, *87*, 105952. [\[CrossRef\]](#)
7. Dorsey, B. Mass transit trends and the role of unlimited access in transportation demand management. *J. Transp. Geogr.* **2005**, *13*, 235–246. [\[CrossRef\]](#)
8. Sharif Azadeh, S.; Marcotte, P.; Savard, G. Railway demand forecasting in revenue management using neural networks. *Int. J. Revenue Manag.* **2013**, *7*, 18–36. [\[CrossRef\]](#)
9. Milosavljević, N.; Milenkovic, M.; Bojovic, N.; Svadlenka, L.; Avramovic, Z. A hybrid model for forecasting the volume of passenger flows on Serbian railways. *Oper. Res.* **2015**, *16*, 271–285. [\[CrossRef\]](#)
10. Zhang, J.; Guo, W. Research on Railway Passenger Flow Prediction Method Based on GA Improved BP Neural Network. *J. Comput. Commun.* **2019**, *7*, 283–292. [\[CrossRef\]](#)
11. Ghalekhondabi, I.; Ardjmand, E.; Young, W.A.; Weckman, G.R. A review of demand forecasting models and methodological developments within tourism and passenger transportation industry. *J. Tour. Futur.* **2019**, *5*, 75–93. [\[CrossRef\]](#)
12. Noersasongko, E.; Julfia, F.; Syukur, A.; Purwanto, P.; Pramunendar, R.; Supriyanto, C. A Tourism Arrival Forecasting using Genetic Algorithm based Neural Network. *Indian J. Sci. Technol.* **2016**, *9*, 1–5. [\[CrossRef\]](#)
13. Huarng, K.H.; Yu, T.; Moutinho, L.; Wang, Y.C. Forecasting tourism demand by fuzzy time series models. *Int. J. Cult.* **2012**, *6*, 377–388. [\[CrossRef\]](#)
14. Profillidis, V.; Botzoris, G. A Comparative Analysis of Performances of Econometric, Fuzzy and Time-series Models for the Forecast of Transport Demand. In Proceedings of the 2007 IEEE International Fuzzy Systems Conference, London, UK, 23–26 July 2007; pp. 1–6. [\[CrossRef\]](#)
15. Rong, W.; Liu, D.; He, X. Prediction of High Speed Railway Passenger Demand Volume Based on Grey Relational Analysis. In Proceedings of the ICTE 2015, Hong Kong, China, 2–4 July 2015; pp. 173–179. [\[CrossRef\]](#)
16. Tsekeris, T.; Tsekeris, C. Demand Forecasting in Transport: Overview and Modeling Advances. *Econ. Res.-Ekonom. Istraživanja* **2011**, *24*, 82–94. [\[CrossRef\]](#)
17. Profillidis, V.A.; Botzoris, G.N. Econometric models for the forecast of passenger demand in Greece. *J. Stat. Manag. Syst.* **2006**, *9*, 37–54. [\[CrossRef\]](#)
18. Nijkamp, P.; Reggiani, A.; Tritapepe, T. Modelling Inter-Urban Transport Flows in Italy: A Comparison between Neural Network Analysis and Logit Analysis. *Transp. Res. Part C* **1996**, *4*, 323–338.
19. Yang, L.; Li, K.; Gao, Z. Train Timetable Problem on a Single-Line Railway With Fuzzy Passenger Demand. *IEEE Trans. Fuzzy Syst.* **2009**, *17*, 617–629. [\[CrossRef\]](#)
20. Ghoseiri, K.; Szidarovszky, F.; Asgharpour, M.J. A multi-objective train scheduling model and solution. *Transp. Res. Part B Methodol.* **2004**, *38*, 927–952. [\[CrossRef\]](#)
21. Banerjee, N.; Morton, A.; Akartunalı, K. Passenger demand forecasting in scheduled transportation. *Eur. J. Oper. Res.* **2020**, *286*, 797–810.
22. Dombi, J.; Jónás, T.; Tóth, Z.E. The epsilon probability distribution and its application in reliability theory. *Acta Polytech. Hung.* **2018**, *15*, 197–216. [\[CrossRef\]](#)
23. Jónás, T. The Generalized Epsilon Function: An Alternative to the Exponential Function. *Acta Cybern.* **2022**, *25*, 703–716. [\[CrossRef\]](#)
24. Jang, J.S. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [\[CrossRef\]](#)
25. Abraham, A. Adaptation of Fuzzy Inference System Using Neural Learning. In *Fuzzy Systems Engineering: Theory and Practice*; Nedjah, N., Macedo Mourelle, L.d., Eds.; Springer: Berlin/Heidelberg, Germany, 2005; pp. 53–83. [\[CrossRef\]](#)
26. Banzhaf, W.; Nordin, P.; Keller, R.E.; Francone, F.D. *Genetic Programming: An Introduction: On the Automatic Evolution of Computer Programs and Its Applications*; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1998.
27. Vose, M.D. *The Simple Genetic Algorithm: Foundations and Theory*; MIT Press: Cambridge, CA, USA, 1999.

28. Alhijawi, B.; Awajan, A. Genetic algorithms: Theory, genetic operators, solutions, and applications. *Evol. Intell.* **2024**, *17*, 1245–1256. [CrossRef]
29. Bonyadi, M.R.; Michalewicz, Z. Particle Swarm Optimization for Single Objective Continuous Space Problems: A Review. *Evol. Comput.* **2017**, *25*, 1–54. [CrossRef]
30. Dolan, E.D.; Lewis, R.M.; Torczon, V. On the Local Convergence of Pattern Search. *SIAM J. Optim.* **2003**, *14*, 567–583. [CrossRef]
31. The MathWorks Inc. Fuzzy Logic Toolbox Version: 24.2 (R2024a). Available online: <https://www.mathworks.com> (accessed on 1 November 2024).
32. Zadeh, L. Fuzzy sets. *Inf. Control* **1965**, *8*, 338–353. [CrossRef]
33. Dubois, D.; Prade, H. *Fuzzy Sets and Systems: Theory and Applications*; Academic Press: Cambridge, MA, USA, 1980.
34. Cheng, C.B. Group opinion aggregation based on a grading process: A method for constructing triangular fuzzy numbers. *Comput. Math. Appl.* **2004**, *48*, 1619–1632. [CrossRef]
35. Bede, B. Fuzzy Numbers. In *Mathematics of Fuzzy Sets and Fuzzy Logic*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 51–64. [CrossRef]
36. Nag, A.; Sarkar, S. Integrating choice freedom, economic health, and transportation infrastructure to forecast tourism demand: A case study of Bishnupur and its alignment with sustainable development goals. *Transp. Policy* **2024**, *147*, 198–214. [CrossRef]
37. Kuby, M.; Barranda, A.; Upchurch, C. Factors influencing light-rail station boardings in the United States. *Transp. Res. Part A Policy Pract.* **2004**, *38*, 223–247. [CrossRef]
38. Chiou, Y.C.; Jou, R.C.; Yang, C.H. Factors affecting public transportation usage rate: Geographically weighted regression. *Transp. Res. Part A Policy Pract.* **2015**, *78*, 161–177. [CrossRef]
39. Iseki, H.; Liu, C.; Knaap, G. The determinants of travel demand between rail stations: A direct transit demand model using multilevel analysis for the Washington D.C. Metrorail system. *Transp. Res. Part A Policy Pract.* **2018**, *116*, 635–649. [CrossRef]
40. Liu, X.; Wu, J.; Huang, J.; Zhang, J.; Chen, B.Y.; Chen, A. Spatial-interaction network analysis of built environmental influence on daily public transport demand. *J. Transp. Geogr.* **2021**, *92*, 102991. [CrossRef]
41. Ma, J.; Liu, H.; Chen, L. Demand-Responsive Transportation Vehicle Routing Optimization Based on Two-Stage Method. *Comput. Mater. Contin.* **2024**, *81*, 443–469. [CrossRef]
42. Toro-González, D.; Cantillo, V.; Cantillo-García, V. Factors influencing demand for public transport in Colombia. *Res. Transp. Bus. Manag.* **2020**, *36*, 100514. [CrossRef]
43. Konečný, V.; Břídíková, M.; Marienka, P. Research of bus transport demand and its factors using multicriteria regression analysis. *Transp. Res. Procedia* **2021**, *55*, 180–187. [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.