Iranian Journal of Fuzzy Systems

Volume 19, Number 3, (2022), pp. 69-87

Original paper

University of Sistan and Baluchestan

A novel Kumaraswamy interval type-2 TSK fuzzy logic system for subway passenger demand prediction

Z. Saghian[1], A. Esfahanipour[2] and B. Karimi[3]

[1,2,3] *Department of Industrial Engineering and Management Systems, Amirkabir University of Technology, Tehran, Iran*

Saghian@aut.ac.ir, esfahaa@aut.ac.ir, b.karimi@aut.ac.ir

Abstract

Fuzzy logic systems (FLSs) are proper tools for learning and predicting of real-world problems. Type-2 fuzzy sets are developments of the conventional type-1 fuzzy sets which are applied for prediction problems with uncertainty. Interval type-2 fuzzy logic system (IT2 FLSs) is the most wildly used type-2 FLS due to its efficiency and simplicity. Passenger demand prediction has a crucial role in the public transportation sector. Because of the nonlinearity and instability of the passenger arrivals prediction, IT2 FLS can be an appropriate method for solving this problem. In this paper, we develop a fuzzy logic system named KIT2 TSK for passenger arrivals prediction in subway stations. In our proposed model, we utilize the Kumaraswamy distribution in the construction of an IT2 TSK FLS. Furthermore, we develop a new input selection measure that applies the SchweizerSklar t-conorm operator in the variable selection process. The flexibility of the Kumaraswamy distribution leads to the ability to approximate several distributions using the same equation by different values for its shape parameters. Utilizing this property, we adopt our proposed model for passenger arrivals prediction of one line of the Tehran subway system as a case study. Moreover, to see the results on unusual days, passenger demand on public holidays, weekends, and special events are also taken into account. The results demonstrate that our proposed methodology has better performance in the hourly prediction of passenger arrivals compared to the benchmarks. The results for the chaotic Mackey-Glass problem also show the superiority of our proposed model.

Keywords: Fuzzy logic system, interval type-2 TSK FLS, Kumaraswamy distribution, passenger flow prediction, public transportation, subway systems.

1 Introduction

A fuzzy rule-based system (FRBS) as a learning and reasoning method is applied to represent the complexity of a systems behavior [60] and to tackle various real-world problems, mainly in the presence of ambiguity or vagueness [11]. It consists of membership functions (MFs), rule-bases, and inference. The inference in fuzzy logic systems maps the inputs and outputs of a model in reliance on fuzzy logic [33]. Mamdani [43], and TSK [61] inferences are two frequently used inference engines, both with fuzzy inputs [9]. The outputs are fuzzy for the Mamdani inference [11], and a weighted combination of the inputs such as the linear regression for the TSK inference [53, 56, 63]. In a data-driven TSK FRBS, the problem domain is divided into rule clusters, which are represented by TSK rules [40]. Because of the uncertainty and randomness in data, in many types of real world problems, the type-1 (T1) FLS is insufficient to obtain the required performance, which results in applying the type-2 (T2) FLS with the capability of modeling higher levels of uncertainty [48]. Instead of crisp values for membership function in the conventional T1 FLS, the membership functions in the T2 FLS are T1 fuzzy sets [35, 36]. The footprint of uncertainty (FOU) derives the T2 fuzzy sets the capability to handle more complex uncertainties; it is determined by lower and upper membership functions (LMF and UMF) [34]. IT2 FLS is the extension of T2 FLS to tackle rule uncertainties more flexibly and effectively. In this type of FLS, membership functions have crisp intervals bounded by [0, 1] [39]. Because of easy implementation, IT2 FLS is the most wildly used

Corresponding Author: A. Esfahanipour

Received: April 2021; Revised: September 2021; Accepted: January 2021.

https://doi.org/10.22111/IJFS.2022.6944

T2 FLS [13, 24]. It has been applied in many areas. Aziz Khater et al. [5], introduce an AIT2-TSK-FLC-RL[1] model to control nonlinear systems. In [49] a methodology for designing MAM-IT2 FLS[2] is proposed. Ashrafi et al. [4], present an IT2-GSETSK [3] model to handle rule uncertainties in a fuzzy neural system. Eyoh et al. [17] present an IT2IFLS[4] -TSK approach for time series prediction to better model the vagueness of knowledge, compared to the classical fuzzy methods.

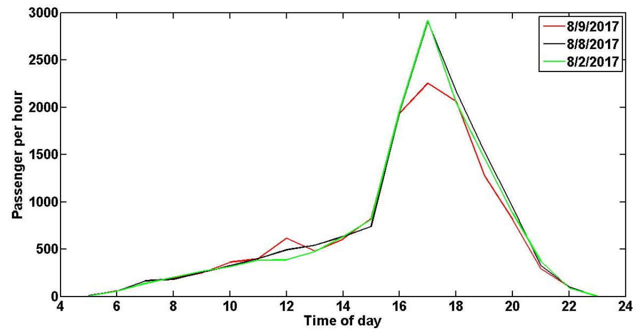

Figure 1: Passenger arrivals at a subway station for three different days

Figure1 depicts passenger arrivals at a subway station on August 9, August 8 (a day before), and August 2 (a week before), 2017. Because of the uncertainties in passenger demand, utilizing a fuzzy rule-based system may be useful for modeling demand with imprecision. In this study, we propose a new KIT2 TSK FLS (Kumaraswamy interval type-2 TSK fuzzy logic system) for passenger arrivals prediction in subway stations. We apply the Kumaraswamy distribution in the construction of the IT2 TSK FLS. The Kumaraswamy distribution, which is introduced by Kumaraswamy [32], is a double-bounded probability density function (DBPDF). It is very close to the Beta distribution, but the primary benefit of the Kumaraswamy is to have a closed-form cumulative distribution function. It has more simple implementation and efficiency [47]. It is a very flexible model and can approximate several distributions using the same equation by different values for its shape parameters. Utilizing this property in our proposed KIT2 TSK model, we design the consequent part of our proposed model via Kumaraswamy distribution.

In recent years, the expansion of subway systems has motivated a severe growth in ridership [74], and therefore it causes several problems, including train congestion and the inadequate capacity of the facilities [75]. During peak hours, because of overcrowdedness, passengers may face problems for boarding trains successfully. Providing proactive passenger flow prediction information is a possible solution for passengers to adjust their departure times and their travel paths. Reliable estimation of passenger demand is applicable for transit managers in optimizing services schedules as well [37]. The weekday, the hour, and special events are some criteria that make variations in demand for subway systems.

There are generally two types of techniques regarding the functional dependency between the dependent and independent variables for passenger flow prediction: parametric and non-parametric methods. In comparison to parametric methods, non-parametric approaches are very flexible to uncover the features of the observed time series [22]. Historical average, ARIMA[5], and regression models are popular parametric techniques, and Neural networks, Kalman filtering models, support vector machines, Bayesian Networks, and fuzzy logic systems are popular non-parametric techniques for passenger flow prediction. In some researches, a combination of parametric and non-parametric approaches as hybrid models are used [68].

The prediction of Passenger flows is mainly derived from historical data. Wei and Chen [68] use a combination of EMD[6] and BPNN[7] to 15-min ahead forecasts. Dou et al., [15] present a short-term passenger flow prediction model based on fuzzy logic (FTLPFFM[8]) to predict the demand for high-speed railways. In a study by Anvari et al. [3], the authors predict short-term passenger arrivals using ARMA models for Istanbul metro stations. Milenkovic et al.[46] apply the SARIMA[9] method for monthly passenger flows prediction of the Serbian railway network. Toque et al.

[1]Adaptive Interval Type-2 Takagi-Sugeno-Kang Fuzzy Logic Controller Based on Reinforcement Learning
[2]Mamdani based Interval type-2
[3] Generic Self-Evolving Takagi-Sugeno-Kang
[4]Interval Type-2 Intuitionistic Fuzzy Logic System
[5]Autoregressive Integrated Moving Average
[6]Empirical Mode Decomposition
[7]Back Propagation Neural Networks
[8]Fuzzy Temporal Logic-Based Passenger Flow Forecast Model
[9]Seasonal AutoRegressive Integrated Moving Average

[62] propose a model to predict dynamic public transportation origin-destination demand for a line in France, using LSTM[10] recurrent neural networks. Liu and Chen [42] propose a new SAE-DNN[11] model for the hourly prediction of passenger flows for some bus rapid transit stations. Li et al. [37] propose a model for passenger demand prediction of a subway system in the presence of special events, using the MSRBF[12] network. Noursalehi et al. [54] use univariate and multivariate state-space models to forecast passenger arrivals of 32 stations of a line in the London subway system. Identification of the stations with similar demand patterns helps to model the effect of events. Liang et al., [41] use a combination of FNN[13] with deep learning to introduce a STEF-Net[14] model for short-term passenger demand prediction, which is the combination of deep learning with a fuzzy neural network. Li et al. [38] propose a dynamic RBFNN[15] approach to forecast outbound volumes of three stations in the Beijing metro. He et al. [21] introduce an Ada-GWL[16] methodology for modeling subway ridership using Shenzhen Metro data; the model adopts a network-based distance metric. Saghian et al., [57] apply the TSK fuzzy logic system modified with the GKPCM [17] Clustering method for forecasting passenger flows of seven important stations of a line in the Tehran Metro system.

In most existing passenger flow prediction models, since the models are developed for only a few stations and a few days, the scalability issues are not discussed [54]. Furthermore, in many of the proposed models, the flows on public holidays or weekends are not taken into account [42].

A fuzzy rule-based system is proposed by Zadeh [73], which may be useful for modeling demand with imprecision. Using fuzzy sets, we can model qualitative knowledge and handle the uncertainties [64]. Because of the randomness, instability, and nonlinearity of passenger arrivals predictions [42], interval type-2 FLS can be applied to solve these problems. Fuzzy rule-based systems have been efficiently utilized for forecasting in the transportation sector [1, 10, 57, 70, 72], but to the best of our knowledge, it is the first time to apply the Kumaraswamy distribution in designing fuzzy systems. Because of its flexibility and goodness-of-fit property, and its ability to approximate several types of distributions depending on different values of the models parameters, the Kumaraswamy distribution is applied in the designing of our proposed KIT2 TSK FLS. The contributions of this study are as follows:

- We develop a novel KIT2 TSK FLS, in which the Kumaraswamy distribution is applied in modeling the interval Type-2 TSK FLS to predict hourly passenger arrivals in subway stations. For the consideration of scalability issues, we also consider passenger demands for public holidays, weekends, and days with special events.
- We develop a new feature selection method for choosing proper inputs from the candidates for passenger demand prediction. In this method, the SchweizerSklar t-conorm operator is utilized in the feature selection process.

The rest of this paper is structured in the following. Section 2 explains the interval type-2 TSK FLS, Kumaraswamy distribution, and the Eigen-decomposition method. In Section 3, we describe the overall problem of subway passenger arrivals prediction, the solution framework, and our proposed KIT2 TSK fuzzy logic system. In Section 5, passenger arrivals prediction of the stations of line 3 of the Tehran subway system as a case study is presented, and the prediction performances are analyzed. The conclusions are represented in Section 6.

2 Preliminary concepts

We explain the interval type-2 TSK FLS, and the Kumaraswamy distribution as basic concepts of designing our proposed model, and the Eigenvalue decomposition technique as a way of removing linear dependencies in variables.

2.1 Interval type-2 TSK FLS

Suppose there are n rules, each one has m antecedents (i.e., input variables). For a rule base with multiple inputs and a single output (MISO), let $X = \{x_1, ..., x_m\}$, and $Y = fn(x_1, ..., x_m)$ be the sets of input and output vectors, respectively. An IT2 TSK IF-THEN rule would be in the following form:

$$\begin{aligned} R_r : IF\, x_1\, is \tilde{A_{r1}}\, and \ldots and\, x_m\, is\, \tilde{A_{mn}} \quad r = 1, ..., n,\ s = 1, ..., m. \\ THEN \quad \tilde{y}_r = fn(x_1, ..., x_m) = \tilde{B_{r0}} + \tilde{B_{r1}} x_1 + ... + \tilde{B_{nm}} x_m, \end{aligned} \tag{1}$$

where $\tilde{A_{r1}}$,..., $\tilde{A_{mn}}$ are antecedent IT2 fuzzy sets, and $\tilde{B_{r0}}$,$\tilde{B_{r1}} x_1$,...,$\tilde{B_{nm}} x_m$ are the parameters of IT2 rule consequences. Since various fine-tuned membership functions do not make considerable differences in inference results [8], and because

[10] Long Short-Term Memory
[11] Stacked Autoencoders Deep Neural Network
[12] Multiscale Radial Basis Function
[13] Fuzzy Neural Network
[14] Spatiotemporal Fuzzy Neural Network
[15] Radial Basis Function Neural Network
[16] Adapted Geographically Weighted Absolute Shrinkage and Selection Operator Least Absolute Shrinkage and Selection Operator
[17] Gustafson-Kessel Possibilistic c-Means

of simplicity, continuity, easy representation, and optimization [69], we utilize the Gaussian IT2 membership function. The uncertain mean or standard deviations are two types of Gaussian IT2 fuzzy sets [24]. Figure2 is an example of an IT2 fuzzy membership function with an uncertain mean.

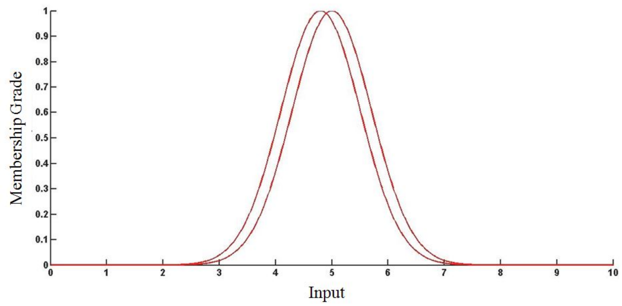

Figure 2: An example of an interval type-2 Gaussian membership function

Eq.2 represents the membership function with an uncertain mean, which takes on values $[m^1{}_{rs}, m^2{}_{rs}]$. The membership functions are defined as follows.

$$\mu_{\tilde{A}_{rs}}(x_s) = exp(-\frac{x_s - m_{rs}}{2\sigma_{rs}})^2 \equiv N(x_s; m_{rs}; \sigma_{rs}), \quad m_{rs} \in [m^1{}_{rs}, m^2{}_{rs}], \tag{2}$$

where $N(x_s; m_{rs}; \sigma_{rs})$ is a normal distribution which is also known as the Gaussian distribution with a mean $m_{rs} \in [m^1{}_{rs}, m^2{}_{rs}]$, and a standard deviation σ_{rs} of input variables. $m^1{}_{rs}$ and $m^2{}_{rs}$ are the mean values of lower and upper MFs of input variables. The value of $\mu_{\tilde{A}_{rs}}(x_s)$ is bounded between the lower$\underline{\mu}_{A_{rs}}(x_s)$ and upper $\overline{\mu}_{A_{rs}}$ membership functions [13]:

$$\underline{\mu}_{A_{rs}}(x_s) = \begin{cases} N(x_s; m^2{}_{rs}; \sigma_{rs}), & x_s \le \frac{m^2{}_{rs} + m^1{}_{rs}}{2} \\ N(x_s; m^1{}_{rs}; \sigma_{rs}), & x_s > \frac{m^2{}_{rs} + m^1{}_{rs}}{2} \end{cases} \tag{3}$$

$$\overline{\mu}_{A_{rs}}(x_s) = \begin{cases} N(x_s; m^1{}_{rs}; \sigma_{rs}), & x_s < m^1{}_{rs} \\ 1 & m^1{}_{rs} \le x_s \le m^2{}_{rs} \\ N(x_s; m^2{}_{rs}; \sigma_{rs}), & x_s > m^2{}_{rs} \end{cases} \tag{4}$$

For each rule, we can obtain the firing strength interval using the integration of the similarity degrees between the antecedent fuzzy sets and the inputs. In other words, for each rule utilizing t-norm operators, weighting parameters are obtained corresponding to the aggregated condition membership functions [50].

For each rule, the firing strength interval would be as [24]:

$$[\underline{f_r}, \overline{f_r}] = [\prod_{s=1}^{m} \underline{\mu}_{A_{rs}}(x_s), \prod_{s=1}^{m} \overline{\mu}_{A_{rs}}(x_s)], \quad r = 1, ..., n. \tag{5}$$

The lower output $\underline{f_r}$ and upper output $\overline{f_r}$ are normalized in the following way:

$$\underline{f} = \frac{\sum_{r=1}^{n} \underline{f_r} * \underline{y_r}}{\sum_{r=1}^{n} \underline{f_r}}, \tag{6}$$

$$\overline{f} = \frac{\sum_{r=1}^{n} \overline{f_r} * \overline{y_r}}{\sum_{r=1}^{n} \overline{f_r}}, \tag{7}$$

Finally, the output y_t is obtained using the BegianMelekMendel (BMM) method [7]:

$$y_t = \theta\overline{f} + (1 - \theta)\underline{f}, \quad 0 \le \theta \le 1, \tag{8}$$

2.2 Kumaraswamy distribution

The Kumaraswamy distribution is a continuous probability distribution and is utilized for modeling double bounded random processes and uncertainties [47]. It is a very flexible model and can approximate several types of distributions depending on different values of its parameters [6, 66]. The application of the Kumaraswamy distribution to broad domains of problems, including many natural phenomena, like hydrological data, shows well goodness-of-fit [16, 29]. The probability density function of the Kumaraswamy distribution Kum(p,q,b,c) would be as follows [47]:

$$f_z(z) = \frac{1}{c-b} pq(\frac{z-b}{c-b})^{p-1}[1-(\frac{z-b}{c-b})^p]^{q-1}, \quad b < z < c,\, p > 0,\, q > 0, \tag{9}$$

in which p and q are shape parameters that affect the general shape of the distribution, and b and c are boundary parameters. The standard form of the Kumaraswamy density function $Kum(p, q) \equiv Kum(p, q, 0, 1)$ is acquired by transforming $x = \frac{z-b}{c-b}$ [47]. Examples of Kum (p, q) for different values of p and q are represented in Figure 3.

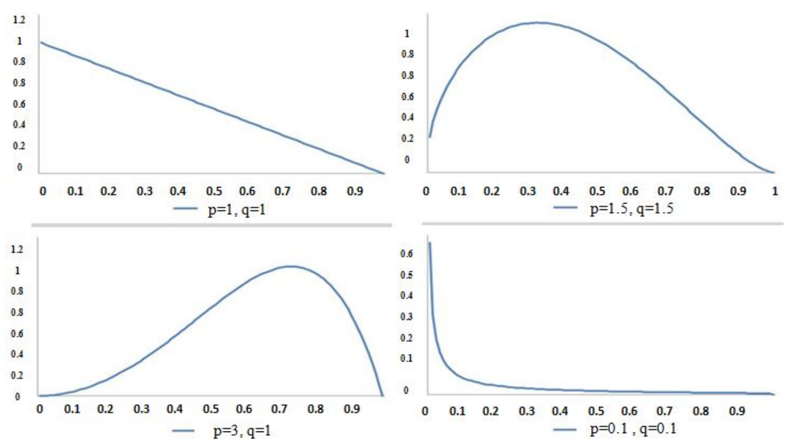

Figure 3: Examples of Kumaraswamy density function for different values of parameters p and q

The cumulative distribution function of the Kumaraswamy density function would be as:

$$F(x) = 1 - (1 - x^p)^{q-1}, 0 < x < 1, \tag{10}$$

The quantile function$F^{-1}(u)$ is also available as closed form:

$$x = [1 - (1 - u)^{\frac{1}{q}}]^{\frac{1}{p}}, \quad 0 < u < 1, \tag{11}$$

where u is a uniformly distributed random variable. Using Eq.11, the random variable z would be:

$$z = z_{min} + (z_{max} - z_{min})x, \tag{12}$$

where z_{min} and z_{max} are lower and upper bounds of random variables [32].

The Kumaraswamy probability distribution produces a broad range of distributional shapes over a bounded interval, and it is an appropriate model in data analysis. The standard form of the probability density function of the Kumaraswamy distribution Kum(p,q,0,1) would be as follows [19]:

$$f(x) = pqx^{p-1}(1 - x^p)^q, \quad 0 < x < 1, p > 0,\, q > 0. \tag{13}$$

According to [45], one of the generalizations of standard Beta distribution is [51]:

$$f(x) = \frac{px^{pa-1}(1 - (x/b)^p)^{q-1}}{b^{ap}B(a, q)}, \quad 0 < x < b,\, p > 0, q > 0. \tag{14}$$

If b=1 and a=1 , Eq.14 reduces to Eq.15.

$$f(x) = \frac{px^{p-1}(1 - (x)^p)^{q-1}}{B(1, q)} = p\frac{(1 + q - 1)!}{(1 - 1)!(q - 1)!}(x)^{p-1}[1 - x^p]^{q-1} = pq(x)^{p-1}[1 - x^p]^{q-1}. \tag{15}$$

In which Eq.15 is the same as Eq.13.

In the literature, the superiority of the results of the Kumaraswamy-based distribution compared to the other distributions such as Gumble [14], Beta [27], and Gamma distributions [20] are demonstrated. Beta distribution has a higher performance in comparison to the other types of distributions, because of its global approximation and allowance to model many finite support distributions [51]. The Kumaraswamy distribution has many similarities and has some advantages over the Beta distribution. Both distributions have identical shape properties such as the following [29]:

- $p > 1, q > 1$ unimodal;
- $p < 1, q < 1$ unianimodal;
- $p > 1, q \leq 1$ increasing;
- $p \leq 1, q > 1$ decreasing;
- $p = q = 1$ constant;

Some advantages of the Kumaraswamy distribution are easy and clear formulas for its distribution and quantile function without the involvement of any special functions; a simple formula for the random variates generation [29]. The advantages of the Kumaraswamy distribution over the Beta and consequently other types of distributions encourage us to apply it in the construction of an IT2 TSK FLS; instead of linear functions of the conventional IT2 rule consequences.

2.3 Eigenvalue-decomposition method

In this work, we adopt the Eigenvalue-decomposition technique known as PCA [18] as a way of removing linear dependencies in the given variables to acquire independent sample variables. It is a variance-covariance matrix decomposition technique and works for matrices in which they are not even positive definite [25]. The Eigen- decomposition technique works as follows.

$$Av = \lambda v, \tag{16}$$

where A is the variance-covariance matrix,v represents the eigenvectors of the linear transformation, in which it keeps the same direction when multiplied by matrix A. λ represents the eigenvalues corresponding to that eigenvectors. To find the eigenvalues and eigenvectors of matrix A, we have to solve the following equations [25].

$$|A - \lambda \mathbb{I}| = 0, \tag{17}$$

$$Var(X)V = V\Lambda \rightarrow Var(X) = V\Lambda V', \tag{18}$$

where $\mathbb{I}$ is the identity matrix, V is an orthogonal matrix, and Λ is a matrix with only non-zero elements on the diagonal. These elements are eigenvalues of matrix A, that is [25]:

$$\mathbb{I} = VV', \tag{19}$$

For a given dataset, axes of the original variable lie along the lines of variation in data. Each axis is independent of the other [59]. The new variable W is the value for each sample on new axes, with expected value E(W) and variance Var(W).

$$W = V'X, \tag{20}$$

$$E(W) = V'E(X), \tag{21}$$

$$Var(W) = V'Var(X)V \rightarrow V'(V\Lambda V')V, \tag{22}$$

Here the covariance matrix is diagonal, which results in a generation of the desired number of independent samples. Then the new independent samples are transformed into their initial coordinates.

$$X = VW. \tag{23}$$

[18] Principal Components Analysis

3 Proposed methodology for subway passenger arrival prediction

3.1 Problem definition and solution framework

We aim to introduce a data-based methodology for the hourly passenger arrivals prediction of subway stations. Therefore, passengers will be able to decide about making proper trip plans. The main data is historical data for passengers arriving at subway stations.

In the following, we describe the overall framework of passenger arrivals prediction. Firstly, input features are determined. They are the calendar date, the month, the weekday, the hour, holiday, and special events. In the feature selection procedure, for each station, the historical data for passenger arrivals are preprocessed, and proper variables are applied as the inputs of our proposed model. In the next step, we implement our proposed KITS TSK model. The outputs are the predicted number of passengers arriving at different hours of the day for the selected station. Figure4. illustrates the overall framework of our proposed methodology for the passenger arrivals prediction. We describe more details about the steps of our proposed approach, next.

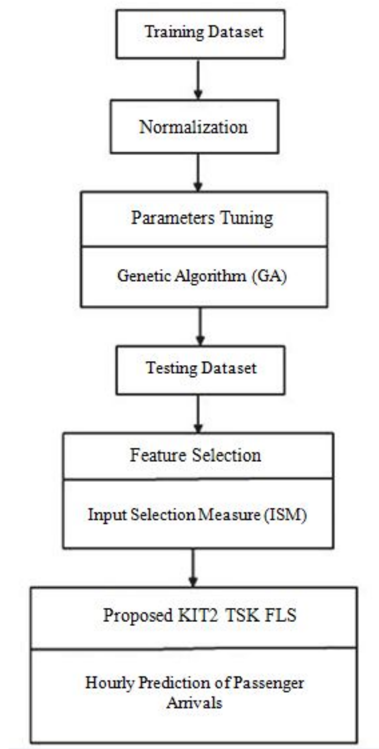

Figure 4: The framework of our proposed Methodology

3.2 Preprocessing and parameters tuning

The first step is to reprocess the historical data. Firstly, we apply the min-max normalization, which results in the mapping of the data into the [0, 1] intervals.

In a TSK FLS as a data-driven approach, utilizing a training dataset, the parameters of the rules and membership functions are optimized for training error reduction [30]. Since different types of fine-tuned membership functions do not make considerable differences in inference results [8], in this study, we utilize Gaussian membership functions. In fuzzy system modeling, Gaussian membership functions are widely used. Because the mean and the standard deviation are parameters of these Gaussian membership functions, the uncertainties of a system are associated with them [24].

Here, we use a genetic algorithm (GA) as a general optimization searching algorithm for tuning the models parameters. GA is based on evolutionary ideas and has been implemented with success in rule base optimization. In a genetic algorithm, firstly, the population is initialized randomly. Then, genetic operators are applied to determine individuals for reproduction. The algorithm repeats until the termination criterion is reached. The optimal solution is the most suitable individual in the current population. The termination criterion is a satisfactory solution or reaching the maximum number of generations [40]. The population size is an adjustable parameter, which is commonly being used as 20- 30 [52]. In this work, the roulette wheel selection method is used for the selection of individuals for reproduction. Using mutation, and crossover operators, offsprings are generated. For two individuals, the crossover operator swaps adjoining parts of the genes. By changing one gene value, the mutation operator is used to escape from local optima and to produce diversity through generations. Crossover and mutation rates must be predefined to control the percentage of their occurrence [40]. In this study, we use GA for tuning the parameters of our proposed model.

We use the mean absolute percentage error (MAPE) as the objective function for determining the quality of individuals for the GA model.

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}|\frac{y_i - \hat{y}_i}{y_i}| \times 100, \tag{24}$$

where y_i and $\hat{y}_i$ are the output observations, and the model prediction, respectively. The number of observational samples equals N.

3.3 Feature Selection

Our proposed input selection process is as follows. Firstly, two input candidates are selected, and one of them is chosen as the input of the model. Then, deviations ($D^{candidate}$) of these variables from the average are computed.

$$D_i^{candidate} = |d_i^{candidate} - avh_i|, \tag{25}$$

where $d_i^{candidate}$ and avh_i are the input candidate and the average of historical data during the ith hour, respectively. The input selection measure (ISM) will be the difference between the $D_i^{current-input}$ of current input and the $D_i^{candidate}$ of input candidate. If $ISM > 0$, the new input $d_i^{new-input}$ will be computed utilizing the t-conorm (i.e. s-norm) operator.

$$ISM_i = D_i^{current-input} - D_i^{candidate}, \tag{26}$$

$$d_i^{new-input} = S(d_i^{candidate}, d_i^{current-input}, p), \tag{27}$$

where $S(d_i^{candidate}, d_i^{current-input}, p)$ is the SchweizerSklar s-norm operator[31], $d_i^{candidate}$, and $d_i^{current-input}$ are the historical data for the candidate, and the current input during the ith hour, and p is a parameter that must be tuned. The SchweizerSklar s-norm operator would be as follows.

$$S(d_i^{candidate}, d_i^{current-input}, p) = [(d_i^{candidate})^p + (d_i^{current-input})^p - (d_i^{candidate})^p(d_i^{current-input})^p]^{\frac{1}{p}}. \tag{28}$$

Triangular norms and conorms (i.e. t-norm and t-conorm) produce a great class of aggregation operators [2, 31, 58] to implement the conjunction and disjunction, respectively. Reinforcement is a feature that differentiates these operators from mean type operators which are always bounded by the maximum and the minimum of their arguments. Utilizing the t-conorm, the aggregated value goes above the maximum of its arguments (i.e. upward reinforcement) [71]. A fuzzy union (t-conorm) u satisfies the following axioms for all $a, b, d \in [0, 1]$ [31]:

- $u(a, 0) = a$ (boundary condition).
- $b \leq d$ Implies $u(a, b) \leq (b, d)$ (monotonicity).
- $u(a, b) = u(b, a)$ (commutativity).
- $u(a, u(, a, b)) = u(u(b, a), d)$ (associativity).
- u is a continuous function (continuity).
- $u(a, a) > a$(superidempotency).
- $a_1 < a_2$ and $b_1 < b_2$ implies $u(a_1, b_1) < u(a_2, b_2)$ (strict monotonicity).

Schweizer-Sklar t-norm and t-conorm are significant classes of the t-norm and t-conorm with high flexibility in information fusion processes. Their adjustable parameter leads to having more flexibility in coping with fuzzy information [67]. The SchweizerSklar t-conorm operator defined as follows [31]:

$$u(a,b) = [a^p + b^p - a^p b^p]^{\frac{1}{p}}. \tag{29}$$

For all $a, b \in [0,1]$ in which p is a parameter.

$$max(a,b) < u(a,b) < u_{max}(a,b), \tag{30}$$

where u_{max} is the drastic union.Eq.30 denotes the upward reinforcement characteristic of the t-conorm operator, and Eq.29 denotes the flexibility of the SchweizerSklar t-conorm operator, in which setting up different values for p results in different values of $u(a,b)$. Some examples of SchweizerSklar t-conorm with different values of the parameter p are represented in Table 1.

Table 1: Some examples of SchweizerSklar fuzzy t-conorm

p	Converges to 1	Converges to 0	converges to ∞
u(a,b)	a+b-ab	$u_{max}(a,b)$	$max(a,b)$

For the problem of subway passenger flow prediction, we utilize both characteristics of Eq.29 and Eq.30 and compute the deviations of the current input and the input candidate from the historical average. If the absolute deviation of the current input from the historical average is greater than the absolute deviation of the input candidate from the historical average, SchweizerSklar t-conorm will be applied to compute the new value of the input. This input selection process is declared in equations Eq.25- Eq.28.

In the next step, utilizing input variables, the proposed KIT2 TSK fuzzy model is implemented. The outputs of the model are the predicted numbers of passengers arrive at each station.

3.4 Proposed KIT2 TSK FLS

In this section, our proposed KIT2 TSK FLS is introduced, in which we adopt the Kumaraswamy distribution for designing the interval type-2 TSK FLS. Suppose there are n rules, each with m antecedents (i.e., input variables). For a MISO rule base, let $X = x_1, ..., x_m$ and $\tilde{Y} = fn(x_1, ..., x_m)$ be the sets of input and output vectors, respectively. As stated in Eq.1 in section 2.1, the general form of the consequent part of IT2 FLS is: $\tilde{y}_r = \tilde{B_{r0}} + \tilde{B_{r1}} x_1 + ... + \tilde{B_{nm}} x_m$.

In our proposed model, instead of linear functions of the conventional form of IT2 rule consequences $\tilde{y}_r$, the functions of IT2 rule consequences are derived from the Kumaraswamy distribution $\tilde{y}_r^{kum}$.

$$\begin{aligned} &R_r : IF\, x_1\, is\, \tilde{A_{r1}}\, and ...\, and\, x_m\, is\, \tilde{A_{mn}} \quad r = 1, ..., n,\ s = 1, ..., m. \\ &THEN \quad \tilde{y}_r^{kum} \sim Kum(p,q), \end{aligned} \tag{31}$$

The flexibility of the Kumaraswamy distribution under different values of the parameters helps to have a better approximation of the rule consequences. Here, we utilize the output values obtained from the linear functions of conventional IT2 TSK FLS to construct the quantile function of the Kumaraswamy distribution. The new values are generated based on the probability density functions of the Kumaraswamy distribution.

$\tilde{y}_r^{kum}$ is computed in the following steps:

- Let us introduce $\tilde{X_{sr}}$,

$$\tilde{X_{sr}} = \begin{cases} \tilde{B_{sr}} x_s & s > 0, r = 1, ..., n, s = 0, ..., m, \\ \tilde{B_{sr}} & s = 0 \end{cases} \tag{32}$$

  where s and r are variables and rules, respectively.
- In our proposed KIT2 TSK FLS, the outputs of the conventional IT2 TSK are transformed into the outputs based on the probability density functions of the Kumaraswamy distribution. To estimate the output variable that follows the Kumaraswamy distribution more precisely, we need to generate independent random samples with normal distributions $N(\mu, \sigma^2)$; μ and σ^2 are derived from datasets in which each set ($\tilde{X_s}$) contains values of a variable in different rule consequences, $\tilde{X_s} = \tilde{X_{s1}}, \tilde{X_{s2}}, ..., \tilde{X_{sn}}$, and $s = 0, ..., m$. Therefore, the Eigenvalue-decomposition technique is applied to remove linear dependencies in given variables and acquire the desired numbers of independent sample variables. Through this approach, the original variable axes lie along lines of variation in

the data, and for each sample, the values for new axes are computed. Then, the variables are projected onto the new axes with the diagonal covariance matrix, which leads to the generation of the required independent samples. Then, these new independent samples are converted into the initial coordinates [25, 59]. $\tilde{X_s^{sim}}$ s are independent random samples with normal distributions which are generated utilizing Eq.16- Eq.23, where s and sim are the variables and the simulations, respectively.

- Then, the simulated outputs for each rule ($\tilde{y}_r^{sim}$) are computed as follows.

$$\tilde{y}_r^{sim} = \sum_{s=0}^{m} \tilde{X}_{sr}^{sim}, \quad r = 1, ..., n, \ sim = 1, ..., sv, \ and\, s = 0, ..., m. \tag{33}$$

where $\tilde{X}_{sr}^{sim}$ is an independent random sample, $\tilde{y}_r^{sim}$ is also normally distributed with mean $\sum_{s=0}^{m} \mu_{sr}^{sim}$, and variance $\sum_{s=0}^{m} {\sigma^2}_{sr}^{sim}$ and m, n, and sv are the total numbers of input variables, rules, and simulation, respectively.

- Utilizing Eq.13 in section 2.2, the new output ($\tilde{y}_r^{kum}$) would be as:

$$\tilde{y}_r^{kum} = \tilde{mn}_r + (\tilde{mx}_r - \tilde{mn}_r)[1 - (1 - u)^{\frac{1}{q}}]^{\frac{1}{p}}, \tag{34}$$

where $\tilde{y}_r^{kum}$ is a continuous random variable that follows the Kumaraswamy distribution. For each rule, $\tilde{mn}_r$ and $\tilde{mx}_r$ are lower and upper bounds of random samples $\tilde{y}_r^{sim}$, u is a uniformly distributed random variable; p and q are shape parameters. Different values of p and q produce various shapes of density functions. The tuned values of parameters lead to the estimation of the proper density function and improve the prediction accuracy of the outputs. In our proposed KIT2 TSK FLS, $\tilde{y}_r^{kum}$ is replaced by $\tilde{y}_r$ in the consequent part of the IT2 TSK FLS. Then, the normalized lower output $\underline{y}^{kum}$ and the upper output $\overline{y}^{kum}$ would be:

$$\underline{y}^{kum} = \frac{\sum_{r=1}^{n} \underline{f_r} * \underline{y_r}^{kum}}{\sum_{r=1}^{n} \underline{f_r}}. \tag{35}$$

$$\overline{y}^{kum} = \frac{\sum_{r=1}^{n} \overline{f_r} * \overline{y_r}^{kum}}{\sum_{r=1}^{n} \overline{f_r}}, \tag{36}$$

where $\underline{f_r}$ is a lower bound and $\overline{f_r}$ is an upper bound of firing interval of rule r, which are obtained from Eq.5, in section 2.1. Then the final output y^{KIT2_TSK} can be presented as follows:

$$y^{KIT2-TSK} = \theta(\overline{y}^{kum}) + (1 - \theta)(\underline{y}^{kum}), 0 \leq \theta \leq 1, \tag{37}$$

where θ is a weight parameter that must be tuned.

3.5 Performance evaluation of our proposed model

The dataset is divided into training and test sets. The parameters of the model are tuned using the training dataset. To estimate the performance of our proposed model, we adopt the root mean square error (RMSE), along with the MAPE. Then we compare the results of our proposed KIT2 TSK FLS model with the results of the IT2 TSK, the ANFIS[19], and the conventional TSK, using both MAPE and RMSE measures.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2}. \tag{38}$$

In this study, we apply the GA as a tool for tuning the parameters of the antecedent and the consequent parts of the TSK FLSs. We adopt the same optimized parameters for the TSK, IT2 TSK, and also KIT2 TSK inference systems. The adaptive neuro-fuzzy inference system (ANFIS) is a robust modeling method that combines ANN[20] and fuzzy logic theory. It is a kind of artificial neural network that is based on a TSK fuzzy inference system, its inference system consists of a set of fuzzy IF-THEN rules [26]. We apply the fuzzy toolbox of MATLAB (ANFIS) as a robust modeling method for time series prediction. Fuzzy Logic Toolbox software applies a command-line function and an interactive app (Neuro-Fuzzy Designer) for training the ANFIS [44]. In this paper, the same training and test data sets are used to compare the prediction results of our proposed KIT2 TSK model with the results of ANFIS, TSK, and IT2 TSK models containing two input variables with 5 Gaussian membership functions, and a single output variable. The Conventional TSK FLS, IT2 TSK FLS, and the proposed KIT2 FLS codes are implemented in the MATLAB software environment.

[19] Adaptive Neuro-Fuzzy Inference System
[20] Artificial Neural Network

4 Experimentation

In this section, first of all, the proposed KIT2 TSK model is utilized to predict values of the chaotic Mackey-Glass time series as one of the well-known benchmark problems. The time-series data is generated from the following equation.

$$x(t) = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t). \tag{39}$$

For 1000 input-output samples from 118 to 1117, the 500 first samples are used as the training data and the remaining are the testing data for performance evaluation. The results are evaluated in terms of the RMSE, by considering $x(t-18)$, $x(t-12)$, $x(t-6)$, $x(t)$ as inputs to predict $x(t+6)$ [65]. For training the dataset, the GA is run for the population size 10 and the maximum number of iteration of 1000. The crossover and mutation rates are 0.8 and 0.1 respectively. We set the number of rules equal to 4. Then, the performance of our proposed KIT2 TSK FLS is evaluated in comparison to some typical methods for time series prediction. The RMSE results are shown in Table 2. The results show the better performance of our proposed model in comparison to studies of [12, 18]and the ANFIS.

Table 2: the RMSE results for the Mackey-Glass time series

Model	RMSE
LSTM[21] [18]	0.7173
BPNN[22] [12]	0.02
ANFIS	0.0025
Proposed model	0.0023

Then, the approach is demonstrated utilizing the historical hourly entrance transaction data for the Tehran metro line 3 from 6:00 AM to 10:00 PM. The Tehran Metro (Tehran subway system) comprises six operational, and one additional commuter rail line, with 229 kilometers of subway network, and carries more than 3 million passengers a day. Figure5 shows the layout of the Tehran metro network. The blue color represents line 3; It connects northeast to southwest of Tehran [23].

In this section, our proposed KIT2 TSK FLS model is validated and evaluated in 2 cases. The dataset consists of 28 days for training, which are 4-17 April, 23-29 July, and 25-31 May 2017. The test data are for March, April, May, June, July, and August 2017. Data from 14 days in April 2017 from 4-17 April are selected to compute the passenger arrival historical average for each station. In Case1 the performance of the proposed ISM method is evaluated. In Case2 the hourly prediction is done, which helps passengers and policymakers to decide about their plans. In this case, we consider two days with unusual passenger arrivals patterns for hourly prediction of passenger demand. Data selection is based on the Iranian calendar.

To achieve an optimal trade-off between model complexity and prediction accuracy, it is vital to select appropriate parameters. In this study, the model parameters are trained using a genetic algorithm (GA). The GA is used for tuning the parameters of the antecedent fuzzy sets (Gaussian membership functions) and the parameters of the rule consequences $\tilde{B}_{r0}, \tilde{B}_{r1}, ..., \tilde{B}_{nm}$ of the TSK FLSs. We set up the population size, crossover, and mutation rates of 20, 0.8, and 0.05; these values are commonly used in the literature. The termination threshold is 0.01, and the maximum number of generations for termination is 200; these values have been obtained by a trial and error process. Firstly, the population is initialized randomly, and offsprings are generated utilizing roulette-wheel selection operator, as well as mutation, and one-point crossover operators. The MAPE is the objective function to determine the quality of individuals for the GA model. The most suitable results in terms of the MAPE are selected as the parameters of the KIT2 and IT2, and the conventional TSK FLSs. The parameters of the Kumaraswamy distribution are tuned utilizing the Lsqcurvefit function in the MATLAB software, in which it is applied to solve nonlinear data-fitting problems in the least-squares concept [44]. The remaining dataset is used as a test dataset to assess model performance using the RMSE and the MAPE measures. In this study, all experiments are implemented on a 2.53 GHz Intel(R) core i5 PC with 4 GB RAM on MATLAB R2014a environment.

4.1 Results and discussion

We evaluate our proposed KIT2 TSK FLS model in 2 cases. In Case1, the performance of our proposed model with a new feature selection method (ISM) is evaluated. Then, in Case2 we evaluate the model in comparison to the conventional first-order TSK, the IT2 TSK, and the ANFIS models. Due to the curse of dimensionality issues, for each

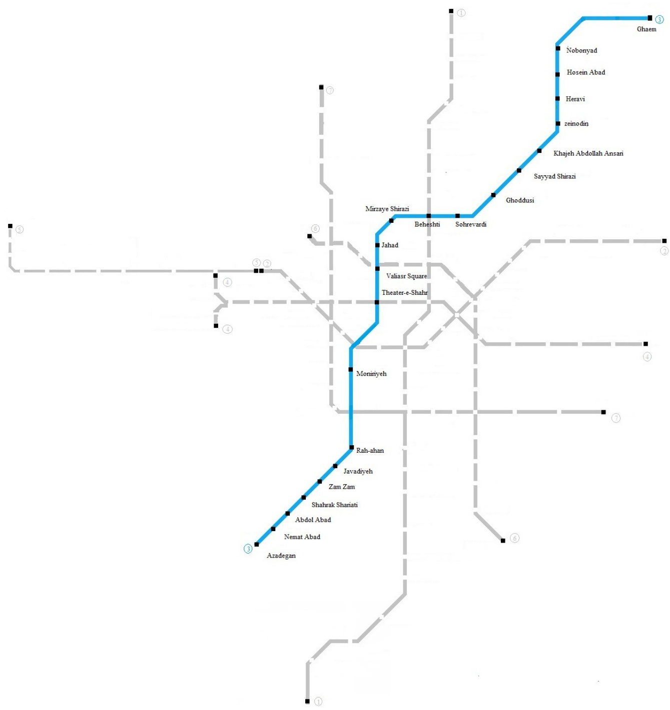

Figure 5: The layout of the Tehran Metro network

station, two variables based on historical data for passenger arrivals, along with general features such as the calendar date, the month, the weekday, the hour, being a holiday or not, and special events are selected as the inputs of our proposed model. The Inputs of our proposed methodology are represented in Table3.

Table 3: the inputs of our proposed KIT2 TSK model

	General features					Historical data for the selected station			
Case	Date	Month	Weekday	Hour	Holiday/ Ordinary day	Special event	Two weeks before	Previous seven days	Previos hour
1	•	•	•	•	•		•	•	•
2	•	•	•	•	•	•	•	•	•

4.1.1 Case 1:

To evaluate the effectiveness of our proposed SchweizerSklar s-norm input selection measure (SS ISM), we test our proposed model for 2-3 August 2017, and then the results are compared to the results of the model with the simple ISM, and the historical average method in terms of both MAPE and RMSE. The historical average is the average of the historical data. The simple ISM means that if $ISM > 0$, instead of using SchweizerSklar s-norm operator, the input candidate will be replaced with the current input. The average results are represented in Table 4.

To detect significant differences between the results of our proposed model with both SS ISM, and simple ISM, and also the historical average method, we perform t-student tests with 95% confidence levels in terms of MAPE, and define the Null Hypotheses as no differences between the prediction results of the models. The p-values associated with these tests equal 0.0001, and 0, and the Null Hypotheses are rejected. The statistical t-student tests and results of Table 4 show the significance of our proposed model with the SS ISM in comparison to the simple ISM and the historical average method.

Table 4: The average MAPE and RMSE results for each station

Stations	KIT2 TSK with SS ISM		KIT2 TSK with simple ISM		Historical Average	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
Ghaem	**0.225**	**1.435**	0.275	1.52	15.07	881.13
Nobonyad	**0.22**	**1.64**	0.24	1.745	15.67	846.82
Hosein Abad	**0.36**	**0.63**	0.45	0.83	17.08	771.56
Heravi	**0.435**	**0.755**	0.515	0.97	13.12	69.96
Zeinodin	**0.36**	**1.75**	0.36	2.12	14.2	434.55
Khajeh Abdollah Ansari	**0.315**	**2.01**	0.36	2.25	11.4	303.65
Sayyad Shirazi	**0.295**	**0.78**	0.365	1.05	12.19	303.65
Ghoddusi	**0.255**	**1.03**	0.305	1.145	17.34	452.95
Sohrevardi	**0.355**	**3.405**	0.435	4.665	19.43	401.1
Beheshti	**0.485**	**2.09**	0.425	1.88	18.5	280.31
Mirzaye Shirazi	**0.31**	**3.72**	0.35	4.31	22.72	1322.27
Jahad	**0.185**	**3.64**	0.225	3.77	14.43	1714.1
Valiasr Square	**0.175**	**2.345**	0.22	3.125	12.83	815.18
Theater-e-Shahr	**0.195**	**1.44**	0.205	1.655	11.26	957.51
Moniriyeh	**1.72**	**11.125**	1.785	11.605	10.59	706.25
Rah-ahan	**0.28**	**1.96**	0.31	2.265	10.29	794.26
Javadiyeh	**0.47**	**2.12**	0.615	3.15	11.37	875.9
Zam Zam	**0.595**	**4.91**	0.685	5.64	13.13	596.12
Shahrak shariati	**0.19**	**0.87**	0.205	1.08	10.38	210.84
Abdol Abad	**0.295**	**1.1**	0.3	1.255	10.68	319.76
Nemat Abad	**0.35**	**1.48**	0.365	1.83	10.28	477.73
Azadegan	**1.51**	**24.03**	1.55	25.73	7.17	2239.69

4.1.2 Case 2:

In this case, KIT2 TSK FLS is evaluated for 16 test days. Events may have different influences on transit demand patterns. Prediction of these patterns helps transit operators to implement effective strategies. In this case, the days with special events are 14 and 18 June 2017, in which the offices started two hours later. As shown in Figure6. these events change the passenger arrivals' behavior.

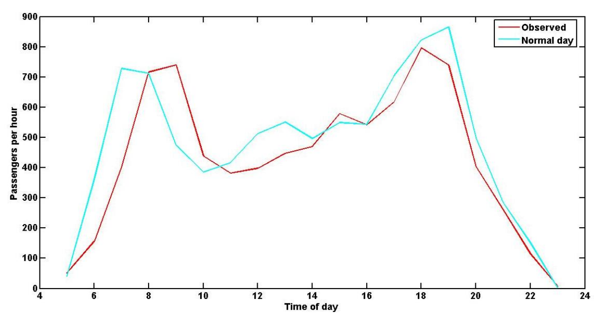

Figure 6: The comparison between observed passenger arrivals on Jun.14, 2017 (i.e., a day with a special event), and passenger arrivals on an ordinary day for the Moniriyeh station.

To measure the performance of our proposed KIT2 TSK model, we employ the conventional first-order TSK, the IT2 TSK, and the ANFIS models for 16 testing days.

Tables 5-6 depict the performance results. We perform t-student tests with 95% confidence levels in terms of MAPE, to detect significant differences between the proposed KIT2 TSK FLS and the ANFIS, the conventional TSK, and the IT2 TSK FLSs. We define the Null Hypothesis as no differences between the prediction results of the models. The p-values associated with these tests are equal to 1.84e-11, 3.78e-16, and 2.34e-12, respectively, the Null Hypotheses are rejected. Figurs 7-8 are examples of hourly arrival predictions. As they are depicted in Tables 5- 6 , the proposed model outperforms overall in comparison to the ANFIS, the conventional TSK, and the IT2 TSK method for all stations of the Tehran metro line 3, for busy and less busy stations, ordinary days, weekends, holidays, and days with special events.

Table 5: The average MAPE and RMSE results for 16 testing days

	KIT2 TSK		IT2 TSK		TSK		ANFIS	
Stations	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
Ghaem	**5.93**	**17.2**	6.75	25.07	7.16	23.64	19.04	56.59
Nobonyad	**4.24**	**29.64**	5.25	46.14	5.47	41.71	14.76	133.44
Hosein Abad	**7.39**	**12.29**	8.5	14.95	9.13	14.45	25.73	48.28
Heravi	**6.93**	**9.9**	7.78	11.97	8.47	12.59	36.77	43
Zeinodin	**5.14**	**16.61**	6.39	24.78	6.94	25.19	19.4	59.06
Khajeh Abdollah Ansari	**4.62**	**19.03**	5.32	25.88	5.58	24.95	17.31	68.68
Sayyad Shirazi	**6.54**	**11.7**	7.59	15.82	7.72	15.29	25.46	58.45
Ghoddusi	**6.21**	**22**	6.97	27.22	7.38	28.6	21.6	71.03
Sohrevardi	**6.98**	**17.72**	8.72	25.72	8.78	24.98	26.51	71.11
Beheshti	**5.62**	**18.27**	7.27	25.58	7.21	24.34	20.07	60.93
Mirzaye Shirazi	**5.8**	**41.04**	7.03	45.5	7.34	45.52	19.22	112.93
Jahad	**4.33**	**60.39**	5.24	79.5	5.35	75.23	16.31	210.82
Valiasr Square	**4.38**	**43.89**	5.47	65.26	5.47	56.5	14.43	155.14
Theater-e-Shahr	**3.96**	**15.63**	4.66	20.56	4.96	21.01	16.18	60.55
Moniriyeh	**5.62**	**31**	6.41	38.48	7.12	39.86	17.05	125.81
Rah-ahan	**6.13**	**40.26**	7.78	63.57	7.33	56.71	13.39	96.21
Javadiyeh	**7.52**	**12.02**	8.1	13.7	8.68	13.58	16.96	26.79
Zam Zam	**5.79**	**18.45**	6.75	25.25	7.26	27.49	16.78	48.33
Shahrak shariati	**6.4**	**25.63**	7.52	32.6	7.88	33.58	18.23	99.03
Abdol Abad	**5.82**	**17.72**	7.01	21.71	6.97	21.6	21.72	82.41
Nemat Abad	**5.03**	**14.72**	6.26	22.02	6.46	21.53	15.07	46.44
Azadegan	**5.15**	**61.12**	5.62	69.18	6.01	71.5	12.49	202.16

The results of our experiments represent the superiority of our proposed KIT2 TSK FLS in comparison to the IT2 TSK with 15.4%, the conventional TSK with 18.8%, and the ANFIS with 70.4% of MAPE. The proposed KIT2 TSK FLS outperforms the IT2 TSK with 24.88%, the conventional TSK with 22.7%, and the ANFIS with 71.28% of RMSE on average.

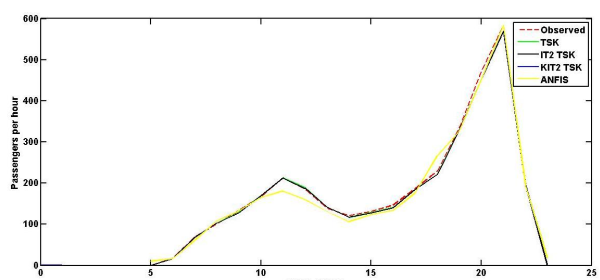

Figure 7: The arrivals predictions of Sayyad Shirazi station on Jul. 25, 2017

5 Conclusion

In this paper, we propose a new fuzzy logic system named KIT2 TSK, in which we apply the Kumaraswamy distribution in the construction of the interval type-2 TSK FLS. Besides, we present a new input selection measure that uses SchweizerSklar t-conorm operator in the input selection procedure. Then, utilizing the proposed KIT2 TSK FLS, we present a general methodology for the hourly prediction of passenger arrivals at subway stations. Earlier understanding of incoming passenger flows helps governmental agencies in policymaking, such as the implementation of proper fare strategies or providing appropriate service levels, and to prevent the decline in the service quality. Utilizing the prediction results, we inform passengers, and transit operators, an hour earlier.

Table 6: The average MAPE and RMSE results for each day

Date		KIT2 TSK		IT2 TSK		TSK		ANFIS	
		MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
21-Mar-17	Holiday	**0.3**	**0.35**	1.4	1.94	1.3	1.89	37	21.16
22-Mar-17	Holiday	**0.3**	**0.41**	1.4	2	1.3	1.95	31.5	32.07
8-APR-17	Working day	**11.1**	**48.3**	11.99	54.98	11.69	73.73	21.79	129.8
13-Apr-17	Weekend	**2.8**	**13.64**	4.13	18.07	2.96	14.7	34.75	130.6
24-May-17	Working day	**10.9**	**51.83**	11.2	54.98	11.2	55.95	16.1	143.52
14-Jun-17	Special event	**10.83**	**45.94**	12.31	54.87	11.67	48.25	16.09	70.12
19-Jun-17	Special event	**10.85**	**45**	11.83	52.42	11.49	45.25	16.79	89.99
25-Jun-17	Working day	**0.3**	**2.17**	1.4	10.83	1.4	11.87	9.1	49.38
29-Jun-17	Working day	**0.6**	**4.26**	1.4	10.99	1.5	11.69	10.3	53.23
2-Aug-17	Weekend	**0.44**	**3.34**	1.36	10.83	1.45	10.7	9.7	45.78
3-Aug-17	Weekend	**0.43**	**3.49**	1.39	10.99	1.33	9.71	16.7	99.32
8-Aug-17	Working day	**7.1**	**33.96**	7.8	38.11	8.4	41.39	14	83.83
9-Aug-17	Working day	**0.6**	**2.64**	1.4	10.87	1.4	11.59	10.6	55.3
12-Aug-17	Working day	**8.5**	**42.73**	10	47.13	10.1	47.02	20.3	112.58
13-Aug-17	Working day	**8**	**39.25**	9	43.79	9.1	44.12	16	97.63
14-Aug-17	Working day	**7.5**	**38.13**	8.8	47.8	9.5	45.72	15.9	121.98

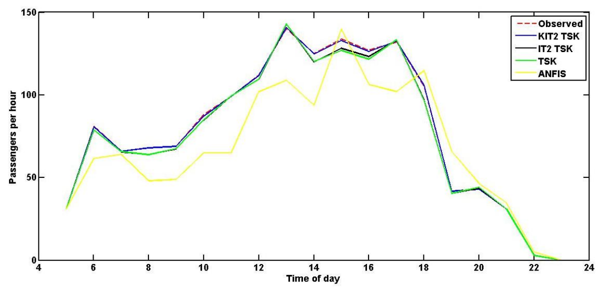

Figure 8: The arrivals predictions of Nemat Abad station on March 21, 2017

The prediction results of 22 stations of the Tehran Metro line 3 indicate that KIT2 TSK FLS works well to forecast hourly passenger arrivals. Furthermore, the application of our proposed model for the chaotic Mackey-Glass benchmark problem shows the better performance of the KIT2 TSK model in comparison to the conventional prediction models in the literature. For further work, we can apply the proposed KIT2 TSK model in other prediction areas and other forms of transportation modes. The weather data and the whole subway network can be included in passenger flow prediction methodology as well.

References

[1] T. Adetiloye, A. Awasthi, *Predicting short-term congested traffic flow on urban motorway networks*, Handbook of Neural Computation, Chapter 8, (2017), 145-165.

[2] C. Alsina, E. Trillas, L. Valverde, *On some logical connectives for fuzzy set theory*, Journal of Mathematical Analysis and Applications, **93** (1983), 15-26.

[3] S. Anvari, S. Tuna, M. Canci, M. Turkay, *Automated Box-Jenkins forecasting tool with an application for passenger demand in urban rail systems*, Journal of Advanced Transportation, **50**(1) (2015), 25-49.

[4] M. Ashrafi, D. K. Prasad, C. Quek, *IT2-GSETSK: An evolving interval Type-II TSK fuzzy neural system for online modeling of noisy data*, Neurocomputing, **407** (2020), 1-11.

[5] A. Aziz Khater, A. M. El-Nagar, M. El-Bardini, N. M. El-Rabaie, *Online learning of an interval type-2 TSK fuzzy logic controller for nonlinear systems*, Journal of the Franklin Institute, **356**(16) (2019), 9254-9285.

[6] F. M. Bayer, D. M. Bayer, G. Pumi, *Kumaraswamy autoregressive moving average models for double bounded environmental data*, Journal of Hydrology, **555** (2017), 385-396.

[7] M. B. Begian, W. W. Melek, J. M. Mendel, *Stability analysis of type-2 fuzzy systems*, IEEE International Conference on Fuzzy Systems, IEEE, (2008), 947-953.

[8] P. C. Chang, C. Y. Fan, *A hybrid system integrating a wavelet and TSK fuzzy rules for stock price forecasting*, IEEE Transactions on Systems, Man, and Cybern Part C Appl Rev., **38**(6) (2008), 802-815.

[9] M. Y. Chen, D. Linkens, *Rule-base self-generation and simplification for data-driven fuzzy models*, Fuzzy Sets and Systems, **142**(2) (2004), 243-265.

[10] Y. P. Chi, Y. Han, L. Rui, Y. P. Wei, *Study of bus incident prediction based on dynamic fuzzy-neural network*, IEEE 2010 International Conference on E-Product E-Service and E-Entertainment Henan China, 7-9 Nov, (2010), 1-6.

[11] O. Cosgun, Y. Ekinci, S. Yank, *Fuzzy rule-based demand forecasting for dynamic pricing of a maritime company*, Knowledge-Based Systems, **70** (2014), 88-96.

[12] S. P. Day, M. R. Davenport, *Continuous-time temporal backpropagation with adaptable time delays*, IEEE Transactions on Neural Networks, **4**(2) (1993), 348-354.

[13] Z. Deng, C. Kup-Sze, C. Longbing, W. Shitong, *T2FELA: Type-2 fuzzy extreme learning algorithm for fast training of interval type-2 TSK fuzzy logic system*, IEEE Transactions on Neural Networks and Learning Systems, **25**(4) (2014), 664-676.

[14] S. Dey, J. Mazucheli, S. Nadarajah, *Kumaraswamy distribution: Different methods of estimation*, Journal of Computational and Applied Mathematics, **37**(2) (2017), DOI: 10.1007/s40314-017-0441-1.

[15] F. Dou, L. Jia, L. Wang, J. Xu, Y. Huang, *Fuzzy temporal logic based railway passenger flow forecast model*, Computational Intelligence and Neuroscience, **2014** (2014), 9 pages, 950371, DOI:10.1155/2014/950371.

[16] M. S. El-Deen, G. Al-Dayian, A. El-Helbawy, *Statistical inference for Kumaraswamy distribution based on generalized order statistics with applications*, British Journal of Mathematics and Computer Science, **4**(12) (2014), 1710-1743.

[17] I. Eyoh, R. John, G. De Maere, *Time series forecasting with interval type-2 intuitionistic fuzzy logic systems*, IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Naples, (2017), 1-6.

[18] F. A. Gers, D. Eck, J. Schmidhuber, *Applying LSTM to time series predictable through time-window approaches*, Neural Nets WIRN Vietri-01, R. Tagliaferri and M. Marinaro, Eds. London, U.K.: Springer, (2002), 193-200.

[19] Y. Ghozzi, N. Baklouti, H. Hagras, M. Ben Ayed, A. Alimi, *Interval type-2 beta fuzzy near set based approach to content based image retrieval*, IEEE Transactions on Fuzzy Systems, (2021), DOI: 10.1109/TFUZZ.2021.3049900.

[20] I. Gosh, G. Hamedani, *The Gamma-Kumaraswamy distribution: An A iternative to gamma distribution*, Communication in Statistics-Theory and Methods, **47**(9) (2015), DOI:10.1080/03610926.2015.1122055.

[21] Y. He, Y. Zhao, K. L. Tsui, *An adapted geographically weighted LASSO (Ada-GWL) model for predicting subway ridership*, Transportation, **48**(3) (2021), 1185-1216.

[22] G. Hesamian, F. Torkian, M. Yarmohammadi, *A fuzzy non-parametric time series model based on fuzzy data*, Iranian Journal of Fuzzy Systems, **19**(1) (2022), 61-72.

[23] Http://Metro.tehran.ir/

[24] S. Huang, M. Chen, *Constructing optimized interval type-2 TSK neuro-fuzzy systems with noise reduction property by quantum inspired BFA*, Neurocomputing, **173**(3) (2016), 1839-1850.

[25] H. T. Huynh, V. S. Lai, I. Soumare, *Stochastic simulation and applications in finance with MATLAB programs*, Wiley Finance, (2011), 60-61.

[26] J. S. R. Jang, *ANFIS adaptive-network-based fuzzy inference system*, IEEE Transactions on Systems Man and Cybernetics, **23**(3) (1993), 665-685.

[27] Z. Javanshiri, A. Habibi Rad, N. R. Arghami, *Exp-Kumaraswamy distributions: Some properties and applications*, Journal of Sciences, Islamic Republic of Iran, **26**(1) (2015), 57-69.

[28] W. Jiang, Z. Ma, H. N. Koutsopoulos, *Deep learning for short-term origin-destination passenger flow prediction under partial observability in urban railway systems*, Neural Computing and Applications, (2022), DOI: 10.1007/s00521-021-06669-1.

[29] M. C. Jones, *Kumaraswamys distribution: A beta-type distribution with some tractability advantages*, Statistical Methodology, **6**(1) (2009), 70-81.

[30] D. R. Keshwani, D. D. Jones, G. E. Meyer, M. B. Rhonda, *Rule-based Mamdani-type fuzzy modeling of skin permeability*, Applied Soft Computing, **8**(1) (2008), 285-294.

[31] G. J. Klir, B. Yuan, *Fuzzy sets and fuzzy logic: Theory and applications*, Prentice-Hall, Inc, (1995), 78-82.

[32] P. Kumaraswamy, *A generalized probability density function for double-bounded random processes*, Journal of Hydrology, **46**(1-2) (1980), 79-88.

[33] W. H. Lai, C. Tsai, *Fuzzy rule-based analysis of firms technology transfer in Taiwans machinery industry*, Expert Systems with Application, **36**(10) (2009), 12012-12022.

[34] J. Leski, *TSK-fuzzy modeling based on ϵ-insensitive learning*, IEEE Transactions on Fuzzy Systems, **13**(2) (2005), 181-193.

[35] R. Li, C. Jiang, F. Zhu, X. Chen, *Traffic flow data forecasting based on interval type-2 fuzzy sets theory*, IEEE/CAA Journal of Automatica Sinica, **3**(2) (2016), 141-148.

[36] L. Li, W. Lin, H. Liu, *Type-2 fuzzy logic approach for short-term traffic forecasting*, IEEE Proceedings-Intelligent Transport Systems, **153**(1) (2006), 33-40.

[37] Y. Li, X. Wang, S. Sun, X. Ma, G. Lu, *Forecasting short-term subway passenger flow under special events scenarios using multiscale radial basis function networks*, Transportation Research Part C, **77** (2017), 306-328.

[38] H. Li, Y. Wang, X. Xu, L. Qin, H. Zhang, *Short-term passenger flow prediction under passenger flow control using a dynamic radial basis function network*, Applied Soft Computing Journal, **83** (2019), 105620, DOI: 10.1016/j.asoc.2019.105620.

[39] J. Li, L. Yang, X. Fu, F. Chao, Y. Qu, *Interval Type-2 TSK+ fuzzy inference system*, IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Rio de Janeiro, (2018), 1-8.

[40] J. Li, L. Yang, Y. Qu, G. Sexton, *An extended Takagi-Sugeno-Kang inference system (TSK+) with fuzzy interpolation and its rule base generation*, Soft Computing, **22** (2018), 3155-3170.

[41] X. Liang, G. Wang, M. Min, Y. Qi, Z. Han, *A deep Spatio-Temporal fuzzy neural network for passenger demand prediction*, Proceedings of the 2019 SIAM International Conference on Data Mining, (2019), 9 pages, DOI: 10.1137/1.9781611975673.12.

[42] L. Liu, R. C. Chen, *A novel passenger flow prediction model using deep learning methods*, Transportation Research Part C: Emerging Technologies, **84** (2017), 74-91, DOI:10.1016/j.trc.2017.08.001.

[43] E. H. Mamdani, *Application of fuzzy algorithms for control of simple dynamic plant*, Proceedings of the Institution of Electrical Engineers, **121**(12) (2009), 1585-1588.

[44] MATLAB and Statistics Toolbox Release 2014a, The MathWorks, Inc., Natick, Massachusetts, United States.

[45] J. B. McDonald, *Some generalized functions for the size distribution of income*, Econometrica, **52**(3) (1984), 647-664.

[46] M. Milenkovic, L. Švadlenka, V. Melichar, N. Bojovic, Z. Avramovic, *SARIMA modeling approach for railway passenger flow forecasting*, Transport, **33**(5) (2018), 1113-1120.

[47] P. A. Mitnik, *New properties of the Kumaraswamy distribution*, Communications in Statistics-Theory and Methods, **42**(5) (2013), 741-755.

[48] K. Mittal, A. Jain, K. S. Vaisla, O. Castillo, J. Kacprzyk, *A comprehensive review on type 2 fuzzy logic applications: Past, present and future*, Engineering Applications of Artificial Intelligence, **95** (2020), 103916.

[49] J. E. Moreno, M. A. Sanchez, O. Mendoza, A. Rodríguez-Díaz, O. Castillo, P. Melin, J. R. Castro, *Design of an interval type-2 fuzzy model with justifiable uncertainty*, Information Sciences, **513** (2020), 206-221.

[50] S. Mousavi, A. Esfahanipour, M. H. Zarandi, *MGP-INTACTSKY: Multitree genetic programming-based learning of INTerpretable and ACcurate TSK sYstems for dynamic portfolio trading*, Applied Soft Computing, **34** (2015), 449-462.

[51] S. Nadarajah, *Discussion on the distribution of Kumaraswamy*, Journal of Hydrology, **348** (2008), 568-569.

[52] N. Naik, R. Diao, Q. Shen, *Genetic algorithm-aided dynamic fuzzy rule interpolation*, IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Beijing, China, 6-11 July 2014, 2198-2205.

[53] A. K. Nandi, F. Klawonn, *Detecting ambiguities in regression problems using TSK models*, Soft Computing, **11**(5) (2007), 467-478.

[54] P. Noursalehi, H. Koutsopoulos, J. N. Zhao, *Real time transit demand prediction capturing station interactions and impact of special events*, Transportation Research Part C., **97** (2018), 277-300.

[55] A. Piegat, M. Landowski, *Multidimensional interval type 2 epistemic fuzzy arithmetic*, Iranian Journal of Fuzzy Systems, **18**(5) (2021), 19-36.

[56] B. Rezaee, M. H. Zarandi, *Data-driven fuzzy modeling for Takagi- Sugeno-Kang fuzzy system*, Information Sciences, **180**(2) (2010), 241-255.

[57] Z. Saghian, A. Esfahanipour, B. Karimi, *Passenger flow prediction of subway systems utilizing TSK fuzzy modeling based on Gustafson-Kessel Possibilistic c-Means Clustering approach*, 17th Iranian International Industrial Engineering Conference held in Mashhad, (2021), 7 pages.

[58] B. Schweizer, A. Sklar, *Statistical metric spaces*, Pacific Journal of Mathematics, **10**(1) (1960), 313-334.

[59] C. Syms, *Principal components analysis*, In: Jorgensen, Sven Erik, and Fath, Brian D., (eds.) Encyclopedia of Ecology, Elsevier, Oxford, (2008), 2940-2949.

[60] K. Tahera, R. N. Ibrahim, P. B. Lochert, *A fuzzy logic approach for dealing with qualitative quality characteristics of a process*, Expert Systems with Applications, **34**(4) (2008), 2630-2638.

[61] T. Takagi, M. Sugeno, *Fuzzy identification of systems and its applications to modeling and control*, IEEE Transactions on Systems, Man, and Cybernetics, **15**(1) (1985), 116-132.

[62] F. Toque, E. Cŏme, M. K. E. Mahrsi, L. Oukhellou, *Forecasting dynamic public transport origin-destination matrices with long-short term memory recurrent neural networks*, In IEEE Conference on Intelligent Transportation Systems, Proceedings, 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), Rio de Janeiro, Brazil, (2016), 1071-1076.

[63] N. L. Tsakiridis, J. B. Theocharis, G. C. Zalidis, *DECO 3 RUM: A differential evolution learning approach for generating compact Mamdani fuzzy rule-based models*, Expert Systems with Applications, **83** (2017), 257-272.

[64] A. Ustundag, M. S. Kilinc, E. Cevikcan, *Fuzzy rule-based system for the economic analysis of RFID investments*, Expert Systems with Applications, **37**(7) (2010), 5300-5306.

[65] B. B. Ustundag, A. Kulaglic, *High-performance time series prediction with predictive error compensated wavelet neural networks*, IEEE Access, **8** (2020), 210532-210541.

[66] L. Wang, S. Dey, Y. M. Tripathi, S. J. Wu, *Reliability inference for a multicomponent stress-strength model based on Kumaraswamy distribution*, Journal of Computational and Applied Mathematics, **376**(1) (2020), 112823.

[67] P. Wang, P. Liu, *Some Maclaurin symmetric mean aggregation operators based on Schweizer-Sklar operations for intuitionistic fuzzy numbers and their application to decision making*, Journal of Intelligent and Fuzzy Systems, **36** (2019), 3801-3824.

[68] Y. Wei, M. C. Chen, *Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks*, Transportation Research Part C: Emerging Technologies, **21**(1) (2012), 148-162.

[69] D. Wu, J. M. Mendel, *Recommendations on designing practical interval type-2 fuzzy systems*, Engineering Applications of Artificial Intelligence, **85** (2019), 182-193.

[70] Y. Xiao, J. J. Liu, Y. Hu, Y. F. Wang, K. K. Lai, S. Wang, *A neuro-fuzzy combination model based on singular spectrum analysis for air transport demand forecasting*, Journal of Air Transport Management, **39** (2014), 1-11.

[71] R. R. Yager, *Generalized triangular norm and conorm aggregation operators on ordinal spaces*, International Journal of General Systems, **32**(5) (2003), 475-490.

[72] H. T. Yu, C. J. Jiang, R. D. Xiao, H. O. Liu, W. Lv, *Passenger flow prediction for new line using region dividing and fuzzy boundary processing*, IEEE Transactions on Fuzzy Systems, **27**(5) (2019), 994-1007.

[73] L. A. Zadeh, *Fuzzy sets*, Information and Control, **8**(3) (1965), 338-353.

[74] D. P. Zhang, X. K. Wang, *Transit ridership estimation with network Kriging: A case study of second avenue subway, NYC*, Journal of Transport Geography, **41** (2014), 107-115.

[75] C. Zhong, M. Batty, E. Manley, J. Wang, Z. Wang, F. Chen, G. Schmitt, *Variability in regularity: Mining temporal mobility patterns in London, Singapore and Beijing using smart-card data*, PLoS One, **11**(2) (2016), DOI: 10.1371/journal.pone.0149222.