Intelligent Automation And Soft Computing, 2019
Copyright © 2019 , TSI® Press
Vol.25,no.1,1–14

Short-term Forecasting of Air Passengers based on the Hybrid Rough Set and the Double Exponential Smoothing Model

Haresh Kumar Sharma, Kriti Kumari, Samarjit Kar

[1]Department of Mathematics, National Institute of Technology Durgapur, West Bengal 713 209, India
[2]Department of Mathematics and Statistics, Banasthali Vidyapith, Jaipur, Rajasthan, 304022, India
[2]kriti.kri89@gmail.com, [1]kar_s_k@yahoo.com

ABSTRACT

This article focuses on the use of the rough set theory in modeling of time series forecasting. In this paper, we have used the double exponential smoothing (DES) model for forecasting. The classical DES model has been improved by using the rough set technique. The improved double exponential smoothing (IDES) method can be used for the time series data without any statistical assumptions. The proposed method is applied on tourism demand of the air transportation passenger data set in Australia and the results are compared with the classical DES model. It has been observed that the forecasting accuracy of the proposed model is better than that of the classical DES model.

KEY WORDS: Air passengers, DES model, Hybrid model, Rough set theory, Time series forecasting.

1 INTRODUCTION

A perfect and accurate tourism prediction plays a vital role in the commercial and business airline industry because of its various real-life applications. In the airline sector, short-term, medium-term and long-term predictions are useful for market planning, daily business operation, and fleet arrangement. Highly precise forecasts help in making effective and successful policies for the government/private agencies. This feature of planning relates to the decision-making process regarding future demand. Therefore, more highly reliable and accurate forecasting tools are required to plan effectively.

Time series forecasting has an effective impact on the planning of all activities of the airline industry and it also defines the relationship with other factors. Time series methods include univariate using previous data points as well as multivariate using the relationship between dependent and independent variables. Different univariate and multivariate econometric techniques such as; autoregressive integrated moving average (ARIMA), seasonal autoregressive integrated moving average (SARIMA), seasonal naïve and simple exponential smoothing have been used to develop some models for forecasting non-stationary airline passenger's data along with trend components (Hsu & Wen, 1998; Ippolito, 1981; Kaemmerle, 1991; Russon & Riley, 1993). Nancy et al. (2016) presented an efficient clinical decision-making system related to the temporal rough set induced neuro-fuzzy (TRiNF) framework for mining clinical time series data obtained from electronic health records. The study of Nancy et al. (2017) also proposed a bio-statistical mining framework to tackle the complexities related to mining in the clinical time series data, which is an effective classification model.

Most of the commonly used time series and econometric models are based on some statistical assumptions, which have certain limitations such as normal distribution and independent identical distribution (iid) requirements. Correspondingly, such models (Regression model, ARIMA, Error correction model, etc.) are unable to give any kind of prediction about the data or its analysis specifically when interdependent variables are showing unknown probability distribution, they can only provide a few conclusions. Some soft computing methods including neural networks models have been commonly used in airline passenger traffic data (Alekseev & Seixas, 2009; Nam & Schaefer, 1995). A neural network has its own advantages and limitations. The advantage of this model is that it does not need many statistical assumptions compared to earlier statistical time series models (Li et al. 2006; Martin & Witt, 1989). In addition, these models still require a large sample for modeling and forecasting time series data (Khashei et al., 2012). However, it turns out to be difficult to

CONTACT Haresh Kumar Sharma harishanuu60@yahoo.com

obtain such large samples. Therefore, this study employs the double exponential smoothing forecasting model for air transport passengers.

The double exponential smoothing (DES) model is proposed by Holt (1957) for level and the trend estimation process (Holt, 2004; Gardner, 1981). The main advantage of the smoothing model is future prediction, which has been obtained by using previously estimated levels and trends in time series data with some appropriate smoothing parameters as mentioned in Chatfield (1995). In addition, statistical distributional assumptions are not necessary for time series when the double exponential smoothing model is used. For this reason, the double exponential smoothing model has frequently been adopted in empirical studies related to forecasting of tourist arrivals. For, instance Yu & Schwartz (2006) compared the forecasting performance of two simple models i.e. double exponential smoothing, and double moving average with the artificial intelligence (AI) model under the mean absolute percentage error (MAPE) indicator. In the study of Law & Au (1999) and Law (2000), the authors have applied exponential smoothing to forecast tourist arrivals in Hong Kong, Japan and Taiwan. In addition, Wu et al. (2016) applied the grey double exponential smoothing model to forecast pig price.

The problem for evaluating and selecting appropriate criteria using different forecasting models has been studied in various domains including, stock exchange (Shen & Loh, 2004; Yao & Herbert, 2009; Pal & Kar, 2017), tourism (Xiaoya & Zhiben, 2011), hospitality (Xu et al., 2016), etc. Therefore, evaluation and selection of criteria can be considered as a multi-criteria decision making (MCDM) problem, which concerns many factors ranging from customer needs to the resource constraints of the enterprise. These real-world problems become more complicated due to imprecise data, decision-makers' subjective judgments using linguistic terms, and multiple sources of information. Hence, a lot of research on MCDM has been done on fuzzy sets (Cosgun et al., 2014, Korol, 2014), rough sets (Zhai et al. 2010), hesitant fuzzy sets (Ye, 2014), etc. Recently Alcantud (2016) conducted a study, which reveals that the hesitant fuzzy set can be considered as interval type-2 fuzzy sets and interval-valued fuzzy sets can be considered as soft sets over the universe [0, 1]. In addition, Alcantud et al. (2016) proposed a ranking methodology for the hesitant fuzzy set and applied it in an MCDM problem.

This article attempts to employ the double exponential smoothing model to forecast air transport passengers. Here we propose a novel approach that combines the double exponential smoothing and rough set to improve the forecasting accuracy of the classical forecasting model. The rough set theory (Lingras, 1996; Pawlak, 1982; Pawlak, 1991; Pawlak, 1996) pays attention to incomplete knowledge to classify the objects in an information system. This methodology is different from other previously used methods, which do not require any distributional assumption (Chen et al. 2011; Pawlak & Slowinski, 1994; Salamo & Lopez-Sanchez, 2011; Yao & Lin, 1996). Nowadays, the rough set theory is applied in various fields, such as acoustical analysis (Kostek, 1996), machine learning (Kaya & Uyar, 2013; Tsumoto & Tanaka, 1996), hospitality (Sharma & Kar, 2018), supplier evaluation (Omurca, 2013), disease diagnosis (Xu & Liu, 2012), feature selection (Aijun et al. 2004), data analysis (Mahapatra & Sreekumar, 2010; Nassiri & Rezaei, 2012; Tay et al. 2003) and so on. Recently, there are some articles where the rough set theory is used for the generation of decision rules for forecasting tourism demand (Goh & Law, 2003; Law & Au, 2000; Li et al. 2011; Liou et al. 2016). In the literature, forecasting accuracy of the rough set theory has been compared with other time series models. For example, in Faustino et al. (2011), a comparison has been done between the accuracy of the rule-based model and the Holt-Winters exponential smoothing model for the electrical charge demand in the United States and the level of Sapucal River in Brazil. However, in the literature, a study on forecasting air transport passenger's data, using the combination of the double exponential smoothing time series model and the rough set theory is yet to be conducted.

This research investigates on modeling and forecasting of air transport passengers and presents a new integrated approach for forecasting by considering both the double exponential smoothing and rough set theory. The empirical results of the combined improved double exponential smoothing model have been compared to the classical double exponential smoothing model using the percentage of corrected classified accuracy (PCCA) criterion. The rest of this paper is as follows: The preliminary concepts for the study are discussed in Section 2. In Section 3, the IDES model has been proposed by employing the Rough set theory. The proposed work related to the study is discussed in Section 4. Two case studies from the Department of Infrastructure and the Regional Development in Australia and their empirical findings are provided in Section 5 and conclusions in Section 6.

2 PRELIMINARIES

In this section, we introduce some basic concepts of the double exponential smoothing (DES), the rough set theory (RST) and related properties.

2.1 Double Exponential Smoothing Model

The double exponential smoothing model is the augmentation of the simple exponential smoothing to forecasting time series data with trend components. This model is used when the trend element is present in the time series data. It consists of two smoothing equations to estimate level and trend. The estimated level and trend of the double exponential smoothing model for time series A_t is given by:

$$L_t = \mu A_t + (1-\mu)(L_{t-1} + T_{t-1})$$
$$T_t = \rho(L_t - L_{t-1}) + (1-\rho)T_{t-1}$$

In these equations, L_t and T_t show the estimated level and trend of the time series. Also, μ and ρ are smoothing parameters for level and trend, respectively, which lie in the range 0 and 1. The forecasted value of time series the series A_t for period h is obtained as

$$\hat{A}_{t+h} = L_t + hT_t$$

2.2 Rough set theory

The concept of the rough set is based on the assumption that every element of the universal set X is related to information and the associated attributes for each object, and describes its relevant information. Information will have indiscernibility when the object has the same description. For the evaluation of a vague description of the member rough set theory is an excellent mathematical tool. The adjective vague expresses the information quality that is uncertainty or ambiguity, that chase from information granulation. The indiscernibility relation developed in this manner is a mathematical foundation of the rough set theory; it induces a separation of the universe into pieces of indiscernible (similar) objects, named the elementary set. The rough set theory can be expressed in the form of two approximation sets called lower and upper approximation of a set.

Let X be the non-empty finite set of objects referred to as universe and A be a non-empty finite set of attributes, then $S = (X, A)$ is called an information system where C, D are two subsets of A, C and D are condition and decision attributes, respectively. For any $\mathrm{P} \subseteq \mathrm{A}$ there exist an indiscernibility relation $\mathrm{K(P)}$ defined as $\mathrm{K(P)} = \{(\mathrm{p},\mathrm{q}) \in \mathrm{X} \times \mathrm{X} \mid \forall \mathrm{b} \in \mathrm{P}, \mathrm{b(p)} = \mathrm{b(q)}\}$, where (p, q) is a couple of instances, $b(p)$ represents the value of attribute b for instance p and $K(P)$ indicate the indiscernibility relation. For $S = (X, A)$ and $P \subseteq A$, $R \subseteq X$ can be approximated based on the knowledge having in P by assembling the P-lower and P-upper approximations of R, represented by $\underline{\mathrm{P}}(\mathrm{R})$ and $\overline{\mathrm{P}}(\mathrm{R})$ respectively; where

$$\underline{\mathrm{P}}(\mathrm{R}) = \{x \mid [x]_P \subseteq \mathrm{R}\} \quad (1)$$

$$\overline{\mathrm{P}}(\mathrm{R}) = \{x \mid [x]_P \cap \mathrm{R} \neq \emptyset\} \quad (2)$$

The objects in $\underline{\mathrm{P}}(\mathrm{R})$ is known as the set of all members of X, which can be surely classified as the member of R in knowledge P, whereas objects in $\overline{\mathrm{P}}(\mathrm{R})$ is the set of all elements of X that can be classified as member of R involving knowledge P. The boundary region of R is expressed as; $\mathrm{BN_P(R)} = \overline{\mathrm{P}}(\mathrm{R}) - \underline{\mathrm{P}}(\mathrm{R})$ is the set of member, which cannot decisively classify into R consisting of knowledge P. If the lower approximation and upper approximation sets are the same then, R becomes an exact set with the boundary region representing an empty set. In the adverse case, if the boundary region contains some objects then set R is referred to as the rough set with respect to P.

2.3 Reduction of Attributes

Attributes reduction procedure removes the surplus attributes from information set and generates 'minimal good enough' subset of attributes for an information system. Such 'minimum good enough' subset of attributes is called a reduct. Reduct is an essential segment of the information system, which can understand all objects discernible by way of the data set and cannot be minimized anymore. Condition attributes (C) of any information system may consist of one or more reducts. The set of condition attributes similar to all reducts of C is known as the 'core' of condition attributes.

2.4 Positive Region and Reduct

The positive region is an essential concept of the classical rough set theory (Pawlak, 1991). The C-Positive region of D consists of all cases of universe U that are definitely classified into partition of U/D by involving attributes from C. We suppose that C and D be the condition and decision attribute of U, which represents the C-Positive region of D, as

$$POS_C(\mathrm{D}) = \cup_{R \in U/D} \underline{C}R$$

If E is a subset of attribute set C, then E becomes a reduct of C with respect to decision attribute D, if and only if the following two conditions are satisfied

$(i)\ POS_C(D) = POS_E(D),$

$(ii)\ POS_C(D) \neq POS_{E-\{e\}}(D), for\ any\ e \in E.$

The core is known as the set of all common reducts of C. The core consists of the set of all indispensable attributes.

2.5 Accuracy of Approximation and Dependency of Attributes (Pawlak, 1991)

The accuracy of approximation of any subset can be denoted in the following manner

$$\alpha_P(R) = \frac{|\underline{\mathrm{P}}(\mathrm{R})|}{|\overline{\mathrm{P}}(\mathrm{R})|}, \quad (3)$$

Where $0 \leq \alpha_P(R) \leq 1$. If the accuracy of approximation of set is equal to 1 then set is called exact otherwise the set is rough.

The dependency of attributes (DA) is based on the total member in the lower approximation to the total member in universe and it is described as follows:

$$\gamma_C(D) = \frac{|POS_C(D)|}{|U|} \quad (4)$$

If $\gamma_C(D) = 1$ we say that D depends completely on C, and if $0 \leq \gamma_C(D) \leq 1$, we say that D depends partially on C. Furthermore, if $\gamma_C(D) = 0$ then D is entirely independent from C.

2.6 Decision Rules

Decision rules are used to preserve the core semantics of the feature set, from the provided information of particular problem, which is an additional significant aspect of the rough set theory. The reduction of needless situations from the decision rules is called as attribute reduction. Hence it can also be called as the generation of decision rules from the data (Omurca, 2013).

The following steps are used for the information table exploration:

1. Collection of data.
2. Computation of lower and upper approximation of the universal set.
3. Obtain C-positive region of D.
4. Obtain reduct and core of attribute sets.

The decision rules can be obtained from the information table. Rules can be considered as "if $p_j = r$ then $d = q$". Where $p_j \subseteq C$ having attribute value *r* and d is decision attribute accept attribute value *q*.

3 IMPROVED DES MODEL BASED ON RST

In this section, the rough set theory is used to improve the forecasting accuracy of the double exponential smoothing model. First, an information system of the rough set is constructed using three attributes, A_t, L_t and T_t of the double exponential smoothing model as described in the previous section. Then the decision rules are created from the information table. Creation of decision rules for a given information system is the most important stage of the rough set theory for the prediction of new objects (samples) by employing the 'IF and THEN' logical statement(s). Further, the information system of the rough set is discussed below.

Let, the data table $\mathcal{D} = \{A, B\}$, where A is the universe of samples, i.e. $A = \{A_t^{(1)}, A_t^{(2)}, \dots, A_t^{(i)}\}$. A_t, L_t and T_t parameters are used as the conditional attributes. $\hat{A}_t$ is the decision attribute to construct the n number of 'IF…THEN' rules (R_n). In this way, we can have many decisions rules to classify the objects. The decision rules are obtained by using the IDES model based on the rough set theory. The information system for the rough set with their conditional and decision attributes are reported in Table1

Table 1. The Information System for IDES Model based on the Rough Set Approach.

Condition attribute			Decision attribute
A_t	L_t	T_t	$\hat{A}_t$
$A_t^{(1)}$	$L_t^{(1)}$	$T_t^{(1)}$	$\hat{A}_t^{(1)}$
$A_t^{(2)}$	$L_t^{(2)}$	$T_t^{(2)}$	$\hat{A}_t^{(2)}$
.	.	.	.
.	.	.	.
.	.	.	.
$A_t^{(i)}$	$L_t^{(i)}$	$T_t^{(i)}$	$\hat{A}_t^{(i)}$

From Table 1, the rules are represented as

If $A_t = A_t^{(i)}$, $L_t = L_t^{(i)}$ and $T_t = T_t^{(i)}$, then we have $\hat{A}_t = \hat{A}_t^{(i)}$ defined by R_n.

And if $A_t = A_t^{(i)}$, $L_t = L_t^{(i)}$ and $T_t = T_t^{(i)}$, then we have $\hat{A}_t = \hat{A}_t^{(i)}$ defined by n numbers of rules R_n for the i_{th} samples.

R_1: If $A_t = A_t^{(1)}$, $L_t = L_t^{(1)}$ and $T_t = T_t^{(1)}$, then $\hat{A}_t = \hat{A}_t^{(1)}$

R_2: If $A_t = A_t^{(2)}$, $L_t = L_t^{(2)}$ and $T_t = T_t^{(2)}$, then $\hat{A}_t = \hat{A}_t^{(2)}$

.
.
.

R_n: If $A_t = A_t^{(i)}$, $L_t = L_t^{(i)}$ and $T_t = T_t^{(i)}$, then $\hat{A}_t = \hat{A}_t^{(i)}$.

4 PROPOSED APPROACH FOR MODELING AND FORECASTING AIR PASSENGERS

Here we propose an improved double exponential smoothing model to forecast time series of the air passenger's data using the rough set theory. The designed structure of this research consists of several stages; time series data collection, double exponential smoothing modeling, generation of improved DES model based on rough set theory, analysis though quality of approximations and dependency of attributes, correlation analysis, creation of decision rules, comparison of the double exponential smoothing time series and improved double exponential smoothing (IDES) models via the rough set. The different stages of the proposed method are described in Figure 1. Accordingly, we discuss the proposed approach in the following subsections.

4.1 Data Collection and Problem Classification

Data related to air passengers are collected from the Australian Department of Infrastructure and Regional Development (https://www.otexts.org/fpp/7/2). The time series data for the air passengers are carefully studied for analysis purpose.

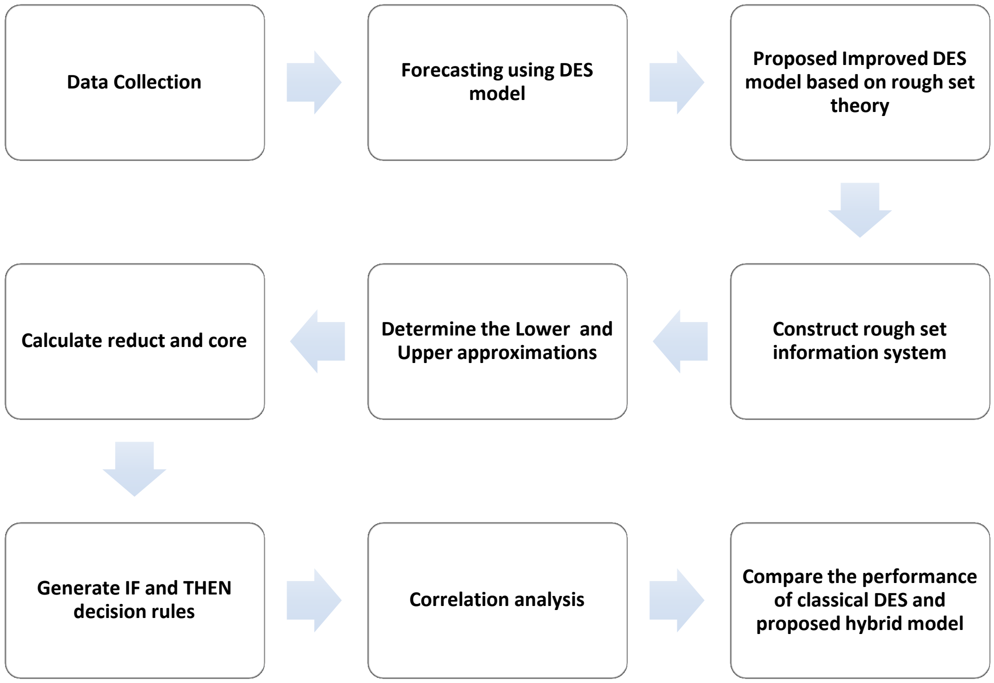

Figure 1. **Stages for the Air Transport Passengers Forecasting based on DES and IDES Models.**

4.2 Rough Set Theory Modeling

After data collection, the DES model is applied to the forecast the time series data related to the air passengers. Then an information system can be built to further analyse the data using the rough set theory. Actual time series (A_t), level (L_t) and trend (T_t) are used as the condition attributes and forecasted value $(\hat{A}_t)$ is used as the decision attribute. The proposed IDES, which is based on the rough set theory, is applied on the data set for forecasting purposes.

4.3 Indispensable Attributes

To identify the important attribute of the information table related to the data set, the reduct and core are determined by employing the rough set theory. Moreover, the correlation analysis has been performed to determine the essential attributes of information table.

4.4 Generation of the IDES Model via the Decision Rules

The decision rules are derived from the information table using the rough set theory. Subsequently, the percentage of corrected classified accuracy criterion is used to evaluate the accuracy of the double exponential smoothing and improved double exponential smoothing models for forecasting. The different stages of the proposed work are depicted in Figure 1.

4.5 Empirical Results of the DES and Improved DES Analysis

Two case studies are selected based on the Australian airline passenger's data. The data related to the air transport passengers (ATP) and airport traffic movement passengers (ATMP) in Australia are considered within the periods of 1992 to 2004, and 1987 to 2014.

The empirical study explains the efficiency of the proposed method. The main goal of this research is to induce a set of decision rules from the rough set information system, to improve the accuracy of the double exponential smoothing model. The experimental results of two different passenger series show, ATP and ATMP are evaluated using R-3.0.3 software for the double exponential smoothing model. The correlation analysis has been performed by employing Minitab 16 software. The rough set modeling for forecasting has been explored with Rough Set Data Explorer (ROSE2) software to produce decision rules (Predki et al. 1998).

4.6 Indispensable Attributes

The double exponential smoothing model is used to estimate the level (mean) and trend from the time series data using the smoothing parameters μ and ρ, respectively. The previously estimated level and trend are utilized in predicting the future of the air transport passengers. The results of the estimated parameters for the double exponential smoothing model are given in Table 2. Figures 2 and 3 explain the relationship between the actual, level, trend and forecasted values for ATP and ATMP based on the double exponential smoothing model. In the next section, the rough set analysis has been performed to improve the forecasting accuracy of the double exponential smoothing model.

4.7 Rough Set Analysis

To generate information system of the rough set, estimated A_t, L_t, T_t and $\hat{A}_t$ of the double exponential smoothing forecasting is used as conditional and decision variables. In the next step the estimated results are normalized for the rough set analysis. The information system of the normalized values (NV) is classified into three qualitative classes; low (IF $0 < NV \leq 0.4$), average (IF $0.4 < NV \leq 0.8$), and high (IF $NV > 0.8$). The normalized decision table for ATP and ATMP are given in Tables 3 and 4 respectively. Therefore, the information system is used for the rough set analysis is to study the forecasting accuracy of the improved double exponential smoothing model.

Table 2. Estimated Parameters of the DES Model for the ATP and ATMP Series.

Data	μ	ρ	Mean	Trend
ATP	0.8	0.2	41.93	1.87
ATMP	0.8	0.2	147664021	4578307

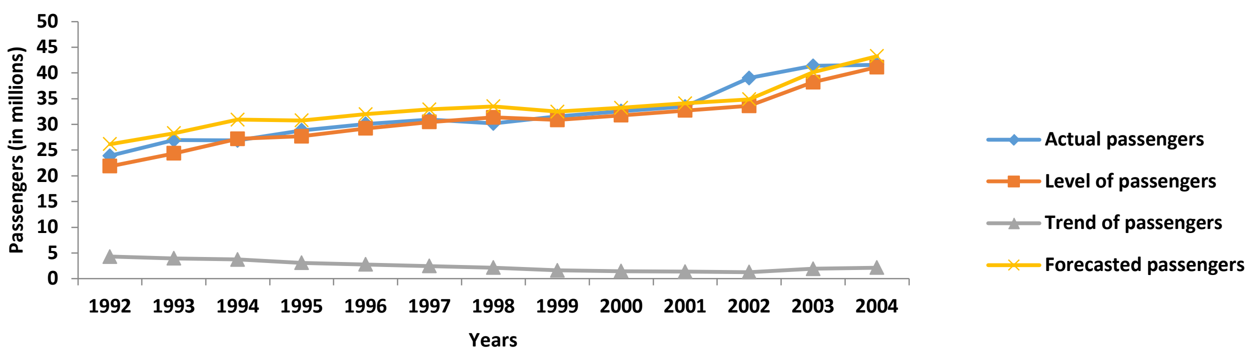

Figure 2. The Relationship between the Actual, Level, Trend and Forecasted Passengers of the ATP.

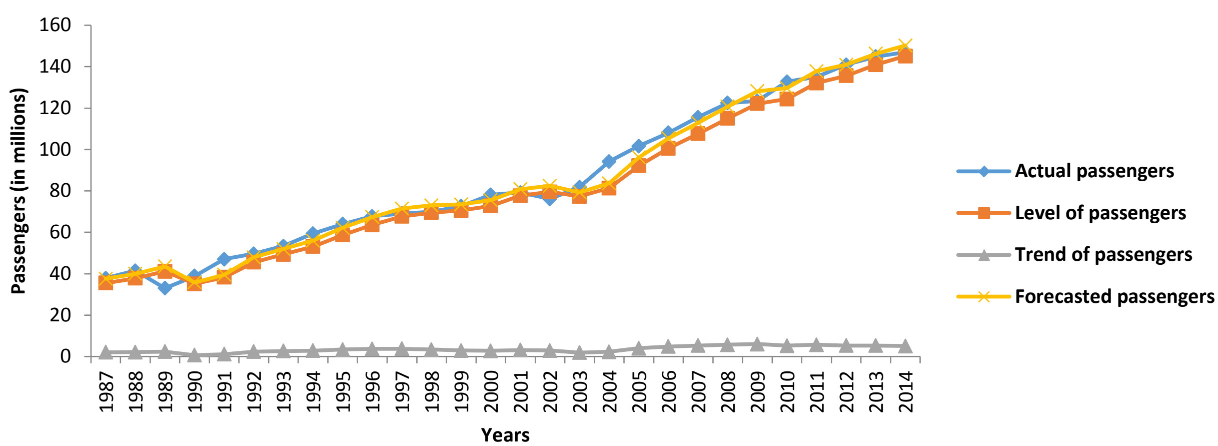

Figure 3. The Relationship between Actual, Level, Trend and Forecasted Passengers of the ATMP.

Table 3. **Information System for the ATP series.**

Universe	Year	A_t	L_t	T_t	$\hat{A}_t$
1	1992	Low	Low	High	Low
2	1993	Low	Low	High	Low
3	1994	Low	Low	High	Low
4	1995	Low	Low	Average	Low
5	1996	Low	Low	Average	Low
6	1997	Low	Average	Low	Low
7	1998	Low	Average	Low	Average
8	1999	Average	Average	Low	Low
9	2000	Average	Average	Low	Average
10	2001	Average	Average	Low	Average
11	2002	High	Average	Low	Average
12	2003	High	High	Low	High
13	2004	High	High	Low	High

Table 4. **Information System for the ATMP Series.**

Universe	Year	A_t	L_t	T_t	$\hat{A}_t$
1	1987	Low	Low	Low	Low
2	1988	Low	Low	Low	Low
3	1989	Low	Low	Low	Low
4	1990	Low	Low	Low	Low
5	1991	Low	Low	Low	Low
6	1992	Low	Low	Low	Low
7	1993	Low	Low	Low	Low
8	1994	Low	Low	Average	Low
9	1995	Low	Low	Average	Low
10	1996	Low	Low	Average	Low
11	1997	Low	Low	Average	Low
12	1998	Low	Low	Average	Low
13	1999	Low	Low	Average	Low
14	2000	Low	Low	Low	Low
15	2001	Average	Low	Average	Low
16	2002	Low	Average	Average	Average
17	2003	Average	Low	Low	Low
18	2004	Average	Average	Low	Average
19	2005	Average	Average	Average	Average
20	2006	Average	Average	Average	Average
21	2007	Average	Average	High	Average
22	2008	Average	Average	High	Average
23	2009	Average	Average	High	High
24	2010	High	High	High	High
25	2011	High	High	High	High
26	2012	High	High	High	High
27	2013	High	High	High	High
28	2014	High	High	High	High

4.7.1 Indiscernibility Relations for the Information System of the ATP Data

The indiscernibility relations and equivalence classes of the rough set information system are described as follows:

$A/A_t = \{(1,2,3,4,5,6,7),(8,9,10),(11,12,13)\}$, since the first seven 1-7 are contained as the 'low' of attribute A_t, cases 8, 9 and 10 having the same value 'average' and cases 11, 12 and 13 consist of the value 'high' for attribute A_t.

Similarly, we can calculate other indiscernibility relations.

$A/L_t = \{(1,2,3,4,5),(6,7,8,9,10,11),(12,13)\}$,

$A/T_t = \{(6,7,8,9,10,11,12,13),(4,5),(1,2,3)\}$.

Next, an indiscernibility relation of the set of all condition attributes is

$A/(A_t, L_t, T_t)$
$= \{(1,2,3),(4,5),(6,7),(8,9,10),(11),(12,13)\}$.

4.7.2 The Concept of Reduct and Core

For calculating the reduct and core, initially, we determine the indiscernibility relation for the sequence of attribute sets.

$A/(L_t, T_t)$
$= \{(1,2,3),(4,5),(6,7,8,9,10,11),(12,13)\}$,

$A/(A_t, T_t)$
$= \{(1,2,3),(4,5),(6,7),(8,9,10),(11,12,13)\}$,

$A/(A_t, L_t)$
$= \{(1,2,3,4,5),(6,7),(8,9,10),(11),(12,13)\}$,

Thus the indiscernibility relation of the decision attribute $\hat{y}_t$ is as follows:

$A/\hat{A}_t = \{(1,2,3,4,5,6,8),(7,9,10,11),(12,13)\}$.

Now we calculate the positive region of information in Table 4.

$A/(A_t, L_t, T_t) =$
$\{(1,2,3),(4,5),(6,7),(8,9,10),(11),(12,13)\}$, and

$A/\hat{A}_t = \{(1,2,3,4,5,6,8),(7,9,10,11),(12,13)\}$,

We have,

$POS_{(A_t, L_t, T_t)}(\hat{A}_t) = \underline{C}(R_{\{\hat{A}_t=\text{low}\}}) \cup \underline{C}(R_{\{\hat{A}_t=\text{average}\}}) \cup \underline{C}(R_{\{\hat{A}_t=\text{high}\}})$, Where

(i) $\underline{C}(R_{\{\hat{A}_t=\text{low}\}}) = (1,2,3,4,5)$.

(ii) $\underline{C}(R_{\{\hat{A}_t=\text{average}\}}) = (11)$.

(iii) $\underline{C}(R_{\{\hat{A}_t=\text{high}\}}) = (12,13)$.

Hence, $POS_{(A_t, L_t, T_t)}(\hat{A}_t) = (1,2,3,4,5,11,12,13)$.

The positive region theory reveals that the set of all indispensable attribute is called core. Therefore, we have

$(i)\ If\ POS_{(L_t, T_t)}(\hat{A}_t) \neq POS_{(A_t, L_t, T_t)}(\hat{A}_t)$, then the attribute A_t is called indispensable.

$(ii)\ If\ POS_{(A_t, T_t)}(\hat{A}_t) \neq POS_{(A_t, L_t, T_t)}(\hat{A}_t)$, then the attribute L_t is called indispensable.

$(iii)\ If\ POS_{(A_t, L_t)}(\hat{A}_t) = POS_{(A_t, L_t, T_t)}(\hat{A}_t)$, then the attribute T_t is called dispensable.

Similarly, the decision of Table 4 has been analysed by employing the rough set theory. A_t and L_t are selected as the indispensable attributes for information in Table 4. Thus, the quality and accuracy of approximation has been discussed in next section.

4.7.3 Accuracy of Approximation and Dependency of Attributes

The attribute selection is one of most important steps in this study, which can reveal the efficiency of indispensable attributes to forecasting the air passenger's data. We can find a smaller attribute set, which describes their important role in the decision table for forecasting. From decision Table 3, the core is the set of (A_t, L_t) as discussed in the previous section. Therefore, these two attributes are essential for forecasting ATP data. Similarly, the core element of decision Table 4 is (L_t, T_t). Hence, L_t and T_t are indispensable attributes to forecast the ATMP series for the double exponential smoothing model. The accuracy of approximation for the three decision classes (Low, Average and High) is shown in Table 5. For the ATP series, the accuracy of approximation for the decision class 'high' is more consistent than the decision classes 'low' and 'average'. Whereas, for the ATMP series, the decision class 'low' has more consistent accuracy of approximation compared to the decision classes 'average' and 'high'. The overall dependency between conditional and decision attributes for both the series are 61% and 89% respectively. In our analysis, it is assumed that all the attributes are of equal importance for the air transport passengers forecasting based on the double exponential smoothing model. Some of the attributes are more essential than the others during the data analysis.

Table 5. Accuracy of Approximation.

Class	No. of objects	Lower approximation	Upper approximation	Accuracy
ATP				
1	7	5	10	0.5
2	4	1	6	0.16
3	2	2	1	1
ATMP				
1	16	16	16	1
2	6	4	7	0.57
3	6	5	8	0.62

4.7.4 Correlation Analysis

In this section, the correlation analysis has been performed to find the correlation coefficient between the conditional and decision variables. We considered the most indispensable attributes based on the largest correlation coefficient with the decision attribute. The results of the correlation coefficient for the different parameters are shown in Table 6. There is a very high degree of association in terms of the positive correlation between the various parameters. The correlation coefficient between $\hat{A}_t$ on $\mathrm{A_t}$ and $\mathrm{L_t}$ is positive i.e., 0.93 and 0.98, it indicates that $\hat{A}_t$ has highly positively correlation with $\mathrm{A_t}$ and $\mathrm{L_t}$. It is also found that T_t has negative low correlation with $\mathrm{A_t}$, $\mathrm{L_t}$ and $\hat{A}_t$ attributes. Table 7 shows correlation coefficients between $\hat{A}_t$ on $\mathrm{L_t}$ and $\mathrm{T_t}$ positive i.e. 0.99 and 0.89, which indicates that $\hat{A}_t$ has highly positive correlation with $\mathrm{L_t}$ and T_t. The correlation between $\hat{A}_t$ on $\mathrm{L_t}$ is 99%, which shows that strong relation of $\hat{A}_t$ with $\mathrm{L_t}$. Hence, the results of attribute reduction statistical inference are the same and that the attribute sets (A_t, L_t) and (L_t, T_t) play an important role in forecasting ATP and ATMP. So we choose (A_t, L_t) and (L_t, T_t) attribute sets as the indispensable attributes in an information system.

Table 6. Correlation between Conditional and Decision Attributes of the ATP Series.

	$\mathrm{A_t}$	$\mathrm{L_t}$	T_t	$\hat{A}_t$
$\mathrm{A_t}$	1	0.95	-0.72	0.93
$\mathrm{L_t}$	0.95	1	-0.74	0.98
T_t	-0.72	-0.74	1	-0.62
$\hat{A}_t$	0.93	0.98	-0.62	1

Table 7. Correlation between Conditional and Decision Attributes of the ATMP series.

	$\mathrm{A_t}$	$\mathrm{L_t}$	T_t	$\hat{A}_t$
$\mathrm{A_t}$	1	0.99	0.87	0.99
$\mathrm{L_t}$	0.99	1	0.89	1
T_t	0.87	0.89	1	0.89
$\hat{A}_t$	0.98	0.99	0.89	1

4.7.5 Improved DES Forecasting

This section formulated the improved double exponential smoothing model to forecasting both the air transport passengers' data using the decision rules. To create the decision rules, information Tables 3 and 4 have been analysed by applying the rough set theory technique. The minimal decision rules describe the reduced information table, which contains strong information about the passengers' data. Now are the decision rules of the improved double exponential smoothing model to calculate the forecasting accuracy of air passenger's data. Five decision rules are produced from Table 3 and Table 4, which are given below for both data series.

Rules for the ATP series:

$R_1. IF\ [0 \leq L_t \leq 0.4]\ \ THEN\ \ \hat{A}_t = low.$

$R_2. IF\ [A_t > 0.8]\ AND\ [0.4 < L_t \leq 0.8]\ \ THEN\ \ \hat{A}_t = average.$

$R_3. IF\ [L_t > 0.8]\ \ THEN\ \ \hat{A}_t = high.$

$R_4. IF\ [0.4 < A_t \leq 0.8]\ \ THEN\ \ \hat{A}_t = low\ or\ \hat{A}_t = average.$

$R_5. IF\ [0 \leq A_t \leq 0.4]\ AND\ [0.4 < L_t \leq 0.8]\ THEN\ \ \hat{A}_t = low\ or\ \hat{A}_t = average.$

Rules for the ATMP series:

$R_1. IF\ [\ 0 \leq L_t \leq 0.4]\ \ THEN\ \ \hat{A}_t = low.$

$R_2. IF\ [0.4 < L_t \leq 0.8]\ AND\ [0.4 < T_t \leq 0.8]\ \ THEN\ \ \hat{A}_t = average.$

$R_3. IF\ [0.4 < L_t \leq 0.8]\ AND\ [0 \leq T_t \leq 0.4]\ \ THEN\ \ \hat{A}_t = average.$

$R_4. IF\ [A_t > 0.8]\ \ THEN\ \ \hat{A}_t = high.$

$R_5. IF\ [0.4 \leq A_t \leq 0.8]\ AND\ [T_t > 0.8]\ \ THEN\ \ \hat{A}_t = average\ or\ \hat{A}_t = high.$

These rules are used to predict the air transport passengers to Australia. From Figure 4, R_1 has the highest support for both series. It indicates that R_1 is the strongest rule for the prediction. The corrected classified prediction for the three classes (Low, Average, and High) is shown in Table 8 using IDES. Thus the forecasting results of the improved double exponential smoothing model are highly accurate for the prediction of the passenger's data. In the next section the accuracy of the improved double exponential smoothing model is compared with the classical DES model under PCCA criterion.

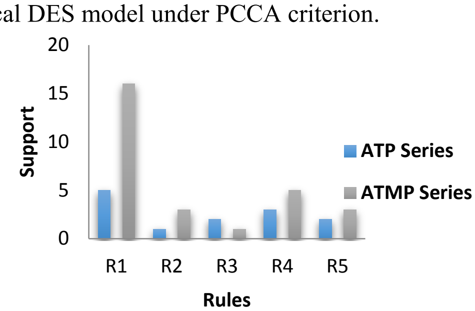

Figure 4. Support for the Decision Rules of the ATP and ATMP series.

Table 8: Validation of the IDES Model.

ATP	Class1	Class2	Class 3
Class1	5	2	0
Class2	1	3	0
Class3	0	0	2
ATMP	Class1	Class2	Class 3
Class1	16	0	0
Class2	0	5	0
Class3	0	1	5
None	NA	1	NA

4.8 *Evaluation and Comparison of the Proposed Approach with the DES Model*

To compare the forecasting performance of the double exponential smoothing and improved double exponential smoothing models, percentage of corrected classified accuracy (PCCA) criterion is used (Goh, & Law, 2003). In general, the accuracy represents the percentage of the corrected classified (CC) instances in an information system. Percentage of the corrected classified accuracy is the ratio of the corrected classified instances and the total number of classified instances. The empirical results show that the average accuracy of the double exponential smoothing and the rough set based improved the double exponential smoothing models being 76.92%, 84.61%, 85.71% and 88.88%, respectively, for both series. Table 10 reveals that the forecasting accuracy of the rough set based improved the double exponential smoothing model being superior to that of the classical double exponential smoothing model for both the passengers' series. Therefore, the improved double exponential smoothing model is more accurate for forecasting the air transport passenger as compared to the double exponential smoothing model. The comparison of actual and forecasted values for the ATP and ATMP series is shown in Figures 5 and 6 based on the double exponential smoothing and improved double exponential smoothing models.

For validation purposes, the classical double exponential smoothing and improved double exponential smoothing models are applied on the data set related to the tourism demand of Luxembourg (Yu & Schwartz, 2006). Subsequently, the performance of both models has been compared with respect to PCCA criterion. The accuracy of the double exponential smoothing and improved double exponential smoothing are respectively determined as 62.5 and 75 and therefore based on the values of accuracy, we can conclude that the forecasting of the improved double exponential smoothing is better than the double exponential smoothing. The results of the actual and forecasted values for Luxembourg data is shown in Figure 7 based on the double exponential smoothing and improved double exponential smoothing model. The corrected classified prediction on the data related to the tourism demand of Luxembourg is shown in Table 9 for the improved double exponential smoothing.

Table 9. Validation for the IDES Model.

Luxembourg	Class 1	Class 2	Class 3
Class 1	5	1	0
Class 2	0	5	0
Class 3	1	2	2

Table 10. PCCA Criterion of the Classical DES and IDES Models (in %).

Series	DES	IDES
ATP	76.92	84.61
ATMP	85.71	88.88
Luxembourg	62.50	75.00

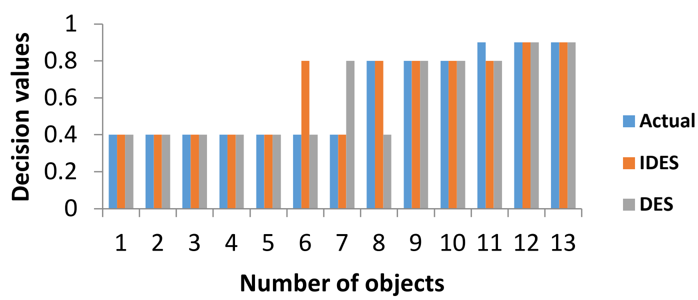

Figure 5. Actual and Forecasted Values of the ATM Series for the DES and IDES Models.

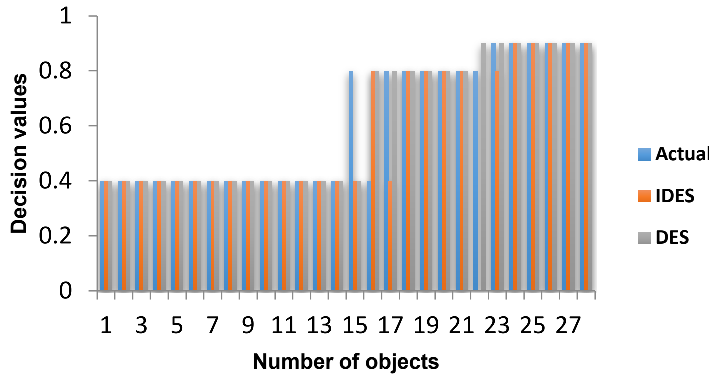

Figure 6. **Actual and Forecasted Values of ATMP Series for the DES and IDES Models.**

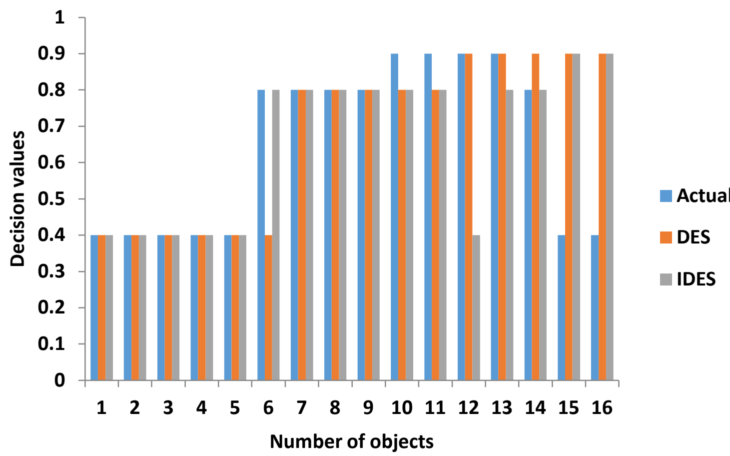

Figure 7. **Actual and Forecasted Values of Luxembourg Data Set for the DES and IDES Models.**

5 CONCLUSIONS

This study explains the usefulness of the rough set approach for improving the accuracy of the DES model to forecast air transport passengers in Australia. We have applied the double exponential smoothing time series model for predicting air transport passengers in Australia. The DES is a highly efficient model to forecast time series data without any particular statistical distributional assumptions. In this research, we have proposed the IDES model by integrating the classical DES model with the rough set theory. The working principle of the proposed model depends on the decision rules generated from an information system. Moreover, we have evaluated the performance of the DES and improved double exponential smoothing models using a percentage of corrected classified accuracy (PCCA) criterion.

In this study, we have considered the indispensable factors for the prediction of air transport passengers by using attribute reduction and correlation analysis. According to the attribute reduction and correlation analysis (A_t, L_t) and (L_t, T_t) these are the most essential factors to forecast the future of air transport passengers. Correlation is positively high between conditional and decision attributes. The forecasting conducted by the improved double exponential smoothing indicates that the prediction accuracy of the proposed approach is superior to that of the classical DES time series model. Therefore, our empirical analysis reveals that the improved double exponential smoothing has a better prediction capacity compared to that of the classical DES model. Moreover, the derived decision rules are easier to understand with respect to the statistical time series methods without any distributional assumptions.

The limitation of this study is that we have used only air transport passengers' time series data for the future prediction. The time series data from other resources are necessary to be considered in order to analyse more robustly the performance of the proposed IDES model in the future.

6 DISCLOSURE STATEMENT

The authors report no conflict of interest.

7 REFERENCES

J. C. R. Alcantud. (2016). Some formal relationships among soft sets, fuzzy sets, and their extensions. *International Journal of Approximate Reasoning*, 68, 45–53.

J. C. R. Alcantud, R. D. A. Calle, & M. J. M. Torrecillas. (2016). Hesitant Fuzzy Worth: An innovative ranking methodology for hesitant fuzzy subsets. *Applied Soft Computing*, 38, (2016) 232–243.

K. P. G. Alekseev & J. M. Seixas. (2009). A multivariate neural forecasting modeling for air transport – Pre-processed by decomposition: A Brazilian application. *Journal of Air Transport Management*, 15, 212-216.

A. An, Y. Huang, X. Huang, & N. Cercone. (2004). Feature Selection with Rough Sets for Web Page Classification. *James F. Peters, A. Skowron. Transactions on Rough Sets II (Rough Sets and Fuzzy Sets). Springer*, 1-13.

Z. Changqing & L. Chengzhi. (2009). Research on Railway Passenger Traffic Volume Forecasting Method. *International Forum on Information Technology and applications*, 279-282.

C. Chatfield. (1995). The Analysis of Time Series an Introduction. Chapman and Hall.

R. R. Chen, Yen-I Chiang, P. P. Chong, Y. H., Lin, & H. K. Chang. (2011). Rough set analysis on call center metrics. *Applied Soft Computing,* 11, 3804-3811.

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

P. C. Faustino, A. C. Pinheiro, A. O. Carpinteiro, & I. Lima. (2011). Time series forecasting through rule-based models obtained via rough sets. *Artif Intell Rev,* 36, 299-310.

D. E. Gardner. (1981). Weight Factor Selection in Double Exponential Smoothing Enrolment forecasts. *Research in Higher Education,* 14(1), 49-56.

C. Goh & R. Law. (2003). Incorporating the rough sets theory into travel demand analysis. *Tourism Management,* 24, 511-517.

J. W. Grzymała-Busse. (1992). LERS–A system for learning from examples based on rough sets. In: R. Slowinski (Ed.). Intelligent Decision Support: Handbook of Application and Advances of Rough set theory. Kluwer Academic Publisher. Dordrecht, 3–18.

S. Hansun. (2016). A new approach of brown's double exponential smoothing method in time series analysis. *Balkan journal of Electrical & Computer Engineering,* 4(2), 75-78.

C. C. Holt. (2004). Forecasting seasonal and trends by exponentially weighted moving averages. *International Journal of Forecasting,* 20, 5– 10.

C. I. Hsu & Y. H. Wen. (1998). Improved Grey prediction models for the trans-pacific air passenger market. *Transportation Planning and Technology,* 22, 87-107.

R. A. Ippolito. (1981). Estimating airline demand with quality of service variables. *Journal of transport Economics and Policy,* 7-15.

K. C. Kaemmerle. (1991). Estimating the demand for small community air service. *Transportation Research 25A (2/3)*, 101-112.

Y. Kaya & M. Uyar. (2013). A hybrid decision support system based on rough set and extreme learning machine for diagnosis of hepatitis disease. *Applied soft computing*, 13(8), 3429-3438.

M. Khashei, M. Bijari, & S. R. Hejazi. (2012). Combining seasonal ARIMA models with Computational intelligence techniques for time series forecasting. *Soft Computing*, 16, 1091-1105.

T. Korol. (2014). A fuzzy logic model for forecasting exchange rates. *Knowledge-Based Systems,* 67, 49-60.

B. Kostek. (1996). Rough Set and Fuzzy Set Methods Applied to Acoustical Analyses. *Intelligent Automation & Soft Computing*, 2(2), 147-160.

R. Law & N. Au. (1999). A neural network model to forecast Japanese demand for travel to Hong Kong. *Tourism Management*, 20(1), 89-97.

R. Law & N. Au. (2000). Relationship modeling in tourism shopping: a decision rules induction approach. *Tourism Management*, 21, 241-249.

R. Law. (2000). Back-propagation learning in improving the accuracy of neural network based tourism demand forecasting. *Tourism Management,* 21(4), 331-340.

J. Li, F. Li, & G. Zhou. (2011). Travel Demand Prediction in Tangshan City of China Based on Rough Set. *Springer-Verlag Berlin Heidelberg,* 440-446.

G. Li, K. K. F. Wong, H. Song, & S. F. Witt. (2006). Tourism demand forecasting: A time-varying parameter error correction model. *Journal of Travel Research,* 45, 175-185.

J. Li, C. Huang, C. Mei, & Y. Yin. (2017). An intensive study on rule acquisition in formal decision contexts based on minimal closed label concept lattices. *Intelligent Automation & Soft Computing, 23(3)*, 519-533.

P. Lingras. (1996). Evidential Comparisons Using Belief Functions, Rough Sets and No monotonic Preferences. *Intelligent Automation & Soft Computing,* 2(2), 203-210.

J. J. H. Liou, Y. C. Chuang, & C. C. Hsu. (2016). Improving airline service quality based on rough set theory and flow graphs. *Journal of Industrial and Production Engineering*, 33(2), 123-133.

S. Mahapatra & S. S. Sreekumar. (2010). Attribute selection in marketing: A rough set approach. *IIMB Management Review*, 22(1-2), 16-24.

A. Martin & F. Witt. (1989). Forecasting tourism demand: A comparison of the accuracy of several quantitative methods. *International Journal of Forecasting*, 5, 1-13.

K. Nam & T. Schaefer. (1995). Forecasting international airline passenger traffic using neural networks. *Logistics and Transportation Review*, 31(3), 239-252.

J. Y. Nancy, N. H. Khanna, & A. Kannan. (2017). A bio-statistical mining approach for classifying Multivariate clinical time series data observed at irregular intervals. *Expert Systems with Applications*, 78, 283–300.

J. Y. Nancy, N. H. Khanna, & A. Kannan. (2016). A Temporal Mining Framework for Classifying Un-Evenly Spaced Clinical Data. *Applied Clinical Informatics*, 7, 1-21.

H. Nassiri & A. Rezaei. (2012). Air itinerary choice in a low-frequency market: A decision rule approach. *Journal of Air Transport Management*, 18, 34-37.

S. I. Omurca. (2013).An intelligent supplier evaluation, selection and development system. *Applied Soft Computing,* 13, 690–697.

S. S. Pal & S. Kar. (2017).Time series forecasting using fuzzy transformation and neural network with back propagation learning, *Journal of Intelligent & Fuzzy Systems*, 33 (1), 467 - 477.

Z. Pawlak. (1982). Rough sets. International Journal of Computer and Information Science, 11, 341-356.

Z. Pawlak. (1991). Rough sets. A Theoretical Aspect of Reasoning about data. *Kluwer Academic Publisher, Boston.*

Z. Pawlak. (1996). Rough Sets Present State and Further Prospects. *Intelligent Automation & Soft Computing*, 2(2), 95-102.

Z. Pawlak & R. Slowinski. (1994). Rough set approach to multi-attribute decision analysis. *European Journal of Operational Research*, 72, 443-459.

Z. Pawlak & A. Skowron. (2007). Rudiments of rough sets. Information Sciences. *An International Journal, 177(1),* 3-27.

B. Predki, S. K. M. Wong, J. Stefanowski, R. Susmaga, & Sz. Wilk, ROSE-software implementation of the rough set theory. In L. Pollkowski, A. Skowron (Eds.). (1998). Rough Sets and Current Trends in Computing. Lecture Notes in Artificial Intelligence. Berlin: Springer, 605-608.

M. G. Russon & N. F. Riley. (1993). Airport substitution in short haul model of air transportation. *International Journal of Transportation Economics*, 20 (2), 157-174.

M. Salamo & L. M. Sanchez. (2011). Rough set based approaches to feature selection for Case-Based Reasoning classifiers. *Pattern Recognition Letters*, 32, 280-292.

H. K. Sharma & S. Kar. (2018). Decision Making for Hotel Selection using Rough Set Theory: A case study of Indian Hotels. *International Journal of Applied Engineering Research,* 13(6), 3988-3998.

L. Shen & H. T. Loh. (2004). Applying rough sets to market timing decisions. *Decision Support Systems*, 37 (4), 583–597.

B. Siregar, I. A. B. Butar, R. F. Rahmat, U. Andayani, & F. Fahmi. (2016). Comparison of exponential smoothing methods in forecasting palm oil real production. *International Conference on Computing and Applied Informatics,* Conf. Series 801.

F. E. H. Tay & L. Shen. (2003). Fault diagnosis based on Rough Set Theory. *Engineering Applications of Artificial Intelligence*, 16, 39-43.

The OTEXTS (https://www.otexts.org/fpp/7/2).

S. Tsumoto & H. Tanaka. (1996). Machine Discovery of Functional Components of Proteins from Amino Acid Sequences Based on Rough Sets and Change of Representation. *Intelligent Automation & Soft Computing*, 2(2), 169-180.

L. Wu, S. Liu, & Y. Yang. (2016). Grey double exponential smoothing model and its application on pig price forecasting in China. *Applied Soft Computing,* 39, 117-123.

H. Xiaoya & J. Zhiben. (2011). Research on Econometric Model for Domestic Tourism Income Based on Rough Set. In: Dai M. (eds) Innovative Computing and Information. ICCIC 2011. Communications in Computer and Information Science, vol 232. Springer, Berlin, Heidelberg, 259-266.

L. Xu & S. Liu. (2012). Intelligent Pearl Disease Diagnosis Based on Rough Set - Neural Network *Intelligent Automation & Soft Computing*, 18(5), 469-476.

X. Xu, R. Law, W. Chen, & L. Tang. (2016). Forecasting tourism demand by extracting fuzzy Takagie Sugeno rules from trained SVMs. *CAAI Transactions on Intelligence Technology,* 1, 30-42.

J. Yao & J. P. Herbert. (2009). Financial time-series analysis with rough sets. *Applied Soft Computing*, 9, 1000–1007.

J. Ye. (2014). Correlation coefficient of dual hesitant fuzzy sets and its application to multiple attribute decision making. *Applied Mathematical Modelling*, 38(2), 659-666.

G. Yu & Z. Schwartz. (2006). Forecasting short time-series tourism demand with Artificial Intelligence models. *Journal of Travel Research*, 45, 194-203.

L. Zadeh. (1965). Fuzzy sets. *Information and Control*, 8, 338-353.

L. Y. Zhai, L. P. Khoo, & Z. W. Zhong. (2010). Towards a QFD-based expert system: A novel extension to fuzzy QFD methodology using rough set theory. *Expert Systems with Applications*, 37 (12), 8888-8896.

8 NOTES ON CONTRIBUTORS

Haresh Kumar Sharma is a Ph.D. Researcher. He received his M.Tech. in Operations Research. His current research interests include rough set, time series forecasting, and multi-attribute decision making.

Kriti Kumari has Masters in Statistics. Her current research interests include time series modeling and rough sets.

Samarjit Kar Ph.D., associate professor. Dr. Kar has published over 80 referred articles in international journals. His research interest includes operations research and optimization, soft computing, uncertainty theory and financial modeling.