Quantification of Perception Clusters Using R-Fuzzy Sets and Grey Analysis

Arjab Singh Khuman[*], Yingjie Yang[†]
[*†§]Centre for Computational Intelligence (CCI)
School of Computer Science and Informatics
De Montfort University
Leicester, United Kingdom

Robert John[‡]
[‡]Lab for Uncertainty in Data and Decision Making (LUCID)
School of Computer Science
University of Nottingham
Nottingham, United Kingdom

Sifeng Liu[§]
[§]College of Economics and Management
Nanjing University of Aeronautics and Astronautics
Nanjing, China

Abstract—This paper investigates the use of the R-fuzzy significance measure hybrid approach introduced by the authors in a previous work; used in conjunction with grey analysis to allow for further inferencing, providing a higher dimension of accuracy and understanding. As a single observation can have a multitude of different perspectives, choosing a single fuzzy value as a representative becomes problematic. The fundamental concept of an R-fuzzy set is that it allows for the collective perception of a populous, and also individualised perspectives to be encapsulated within its membership set. The introduction of the significance measure allowed for the quantification of any membership value contained within any generated R-fuzzy set. Such is the pairing of the significance measure and the R-fuzzy concept, it replicates in part, the higher order of complex uncertainty which can be garnered using a type-2 fuzzy approach, with the computational ease and objectiveness of a typical type-1 fuzzy set. This paper utilises the use of grey analysis, in particular, the use of the absolute degree of grey incidence for the inspection of the sequence generated when using the significance measure, when quantifying the degree of significance fore each contained fuzzy membership value. Using the absolute degree of grey incidence provides a means to measure the metric spaces between sequences. As the worked example will show, if the data contains perceptions from clusters of cohorts, these clusters can be compared and contrasted to allow for a more detailed understanding of the abstract concepts being modelled.

I. INTRODUCTION

The research presented in this paper is yet another enhancement the authors have implemented with the purpose of making R-fuzzy sets (RFS) more applicable, versatile and robust. Since the introduction of RFS by Yang and Hinde in [1], the authors have proposed the significance measure [2], [3], a bridge which connects RFS to that of type-2 fuzzy sets. This facilitates a feasible means to express and infer from complex uncertainty without the inherent difficulties often associated with type-2 fuzzy sets. The authors have also created a heuristic derived approach for establishing a convex hull for the returned degrees of significance, based on the grey whitenisation weight function [4]. The notion of R-fuzzy has still yet to gather serious momentum in terms of usage, which given its capabilities is rather surprising, ergo, the premise of this paper. As the membership set of an RFS, is itself a set, more specifically a rough set, a greater amount of detail can be encapsulated. As it has already been highlighted in previous works, the major draw backs of existing approaches is that a membership value can be lost to an interval or shadow region. In doing so, one is no longer able to ascertain the object's relevance relative to its interval. In certain instances this may not be too much of a concern, but for domains where perception is being modelled, it should always be preferred that each and every membership value be accounted for and have its relevance quantified. For the likes of; Atanassov intuitionistic fuzzy sets [5], where a membership degree and non-membership degree are given. Shadowed sets [6], where the membership value can either belong to the set (1), not to the set (0), or belong to the shadow region $[0, 1]$, to an unknown capacity. Interval-valued fuzzy sets [7], where the use of an interval is used to characterise the object itself. Type-2 fuzzy sets [8], where the secondary grade of membership is a type-1 fuzzy set. These *new* approaches will not be able to recognise the difference between the values which are contained within their intervals or shadow regions. As an RFS makes use of a rough set for its membership set, the lower approximation will contain all fuzzy membership values that have absolutely been agreed upon by all in the consensus. Whereas, the upper approximation will contain all fuzzy membership values that have at least one vote from the populous. The implementation of the significance measure then allows for each and every fuzzy membership value to be quantified [2], [3]. The results of which can be further investigated using techniques from grey system theory.

Grey theory is yet another approach for handling uncertainty, first proposed by Deng in [9]. The paradigm places particular emphasis on domains associated with small samples and poor information, where the information may be partially known and partially unknown, a common trait of uncertain systems. The purpose of which is to garner an informed and accurate conclusion based on what little, uncertain information is available. This is generally achieved through the processes of generating, excavating and extracting meaningful content. In doing so, the system's operational behaviours and its laws governing its evolution can be accurately described and acutely

monitored [10]. The use of sequences in grey modelling is heavily favoured, it is this component and the absolute degree of grey incidence that will provide the additional level of inspection of the results returned by the RFS and the significance measure.

Section II will present the preliminaries for R-fuzzy sets and the significance measure, also introduced is the absolute degree of grey incidence. Section III presents the observations, using a worked example to demonstrate the added benefit of using grey techniques for the inspection and further analysis of the results. Section IV will conclude the paper, providing an overall summary.

II. PRELIMINARIES

We first present the definitions for the *approximations*, the bounding component of all RFS.

A. Approximation Preliminaries

Definition 1 (*Approximations [11]*): Assume that $\Lambda = (\mathbb{U}, A)$ is an information system and that $B \subseteq A$ and $X \subseteq \mathbb{U}$. Set X can be approximated based on the information contained in B, via the use of a lower and upper approximation set.

The lower approximation contains all observed objects that wholeheartedly belong to the set X with regards to the information contained in B. It is the union of all equivalence classes in $[x]_B$ which are absolutely contained within set X, and is given by:

$$\underline{B}X = \{x \mid [x]_B \subseteq X\} \tag{1}$$

$$\underline{B}(x) = \bigcup_{x \in \mathbb{U}} \{B(x) : B(x) \subseteq X\}$$

The upper approximation contains all observed objects that have a *possible* affinity to the set X with regards to the information contained in B. It is the union of all equivalence classes that have a non-empty intersection with set X, and is given by:

$$\overline{B}X = \{x \mid [x]_B \cap X \neq \emptyset\} \tag{2}$$

$$\overline{B}(x) = \bigcup_{x \in \mathbb{U}} \{B(x) : B(x) \cap X \neq \emptyset\}$$

B. R-Fuzzy Set Preliminaries

We now present the concept of R-fuzzy sets, which makes use of the approximations as given in *Definition 1*.

Definition 2 (*R-fuzzy sets [1]*): Let the pair $apr = (J_x, B)$ be an approximation space on a set of values $J_x = \{v_1, v_2, \ldots, v_n\} \subseteq [0, 1]$, and let J_x/B denote the set of all equivalence classes of B. Let $\left(\underline{M}_A(x), \overline{M}_A(x)\right)$ be a rough set in apr. The membership set of an R-fuzzy set A is a rough set $\left(\underline{M}_A(x), \overline{M}_A(x)\right)$, where $x \in \mathbb{U}$, given by:

$$A = \{\langle x, \left(\underline{M}_A(x), \overline{M}_A(x)\right)\rangle \mid$$
$$\forall x \in \mathbb{U}, \underline{M}_A(x) \subseteq \overline{M}_A(x) \subseteq J_x\} \tag{3}$$

$$A = \sum_{x \in \mathbb{U}} \left(\underline{M}_A(x), \overline{M}_A(x)\right)/x$$

Where $\sum$ is the union of all admissible x elements over the universe of discourse. Each $x_i \in \mathbb{U}$ will have an associated description of membership $d(x_i)$, which describes the belongingness of each x_i with regards to the set $A \subseteq \mathbb{U}$. The set C is the available evaluation criteria from which the consensus of the populous is contained. For each pair $((x_i), c_j)$ where $x_i \in \mathbb{U}$ and $c_j \in C$, a subset $M_{cj}(x_i) \subseteq J_x$ is created, given by:

$$M_{cj}(x_i) = \{v \mid v \in J_x, v \xrightarrow{(d(x_i),c_j)} \text{YES}\} \tag{4}$$

The lower approximation for the rough set $M(x_i)$ is given by:

$$\underline{M}(x_i) = \bigcap_j M_{cj}(x_i) \tag{5}$$

The upper approximation for the rough set $M(x_i)$ is given by:

$$\overline{M}(x_i) = \bigcup_j M_{cj}(x_i) \tag{6}$$

Therefore the rough set approximating the membership $d(x_i)$ for x_i is given as:

$$M(x_i) = \left(\bigcap_j M_{cj}(x_i), \bigcup_j M_{cj}(x_i)\right) \tag{7}$$

C. Significance Measure

We now present the significance measure, originally proposed by Khuman et al. in [2], [3].

Definition 3 (*Degree of significance*): Assume that an RFS has already been created using the same notation given in *Definition 2*. This also implies that we have a criteria set C, and in turn, have an established fuzzy membership value set J_x. The total number of all generated subsets for a given R-fuzzy set is denoted by $|N|$. The number of subsets that contain the specific membership value one is inspecting is given by S_v. Each value $v \in J_x$ is evaluated by $c_j \in C$, the frequency of which is the number of times v occurred over $|N|$, this results in the degree of significance given by:

$$\gamma_{\bar{A}}\{v\} = \frac{S_v}{|N|} \tag{8}$$

If the returned degree of significance for any given fuzzy membership value is $\gamma_{\bar{A}}\{v\} = 1$, this implies that the value was absolutely agreed upon by all in the criteria set C, meaning that it belongs to the lower approximation:

$$\underline{M}_A = \{\gamma_{\bar{A}}\{v\} = 1 \mid v \in J_x \subseteq [0, 1]\} \tag{9}$$

For any membership value to be given a $\gamma_{\bar{A}}\{v\} = 1$, one will know that it will also be included in the upper approximation. This is due to that fact that Eq. (3) states that the lower approximation is a subset of the upper approximation $\underline{M}_A(x) \subseteq \overline{M}_A(x)$. Any returned degree of significance greater than 0 will also be included in the upper approximation:

$$\overline{M}_A = \{\gamma_{\bar{A}}\{v\} > 0 \mid v \in J_x \subseteq [0, 1]\} \tag{10}$$

TABLE I. HUMAN PERCEPTION BASED ON THE VARIATIONS FOR THE COLOUR RED

	Age	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9
p_1	20	R	DR	LR	DR	LR	DR	LR	R	LR
p_2	30	R	DR	LR	DR	LR	DR	LR	R	R
p_3	20	R	DR	LR	DR	LR	DR	LR	R	LR
p_4	25	R	DR	LR	DR	LR	DR	LR	R	LR
p_5	25	R	DR	LR	DR	LR	DR	LR	R	LR
p_6	20	R	DR	LR	DR	LR	DR	LR	R	LR
p_7	20	R	DR	LR	DR	LR	DR	LR	R	LR
p_8	25	R	DR	LR	DR	LR	DR	LR	R	R
p_9	25	R	DR	LR	DR	LR	DR	LR	R	R
p_{10}	30	R	DR	LR	DR	LR	DR	LR	R	R
p_{11}	20	R	DR	LR	DR	LR	DR	LR	R	LR
p_{12}	25	R	DR	LR	DR	LR	DR	LR	R	LR
p_{13}	30	R	DR	LR	DR	LR	DR	LR	R	R
p_{14}	30	R	DR	LR	DR	LR	DR	LR	R	R
p_{15}	30	R	DR	LR	DR	LR	DR	LR	R	R

D. Grey Theory

The concept of the absolute degree of grey incidence (ADGI) from grey theory, comes under the remit of grey analysis. It takes into consideration the characteristic sequences of a system $Y_1, Y_2, \ldots, Y_n$, against its behavioural factor sequences $X_1, X_2, \ldots, X_m$, all of which must be of the same magnitude. This is to ascertain how much the sequences are alike, or how much the behaviour factors impact upon the characteristic sequence itself. This information can then be used in terms of identifying if more emphasis should be applied to a particular behaviour or not. This can order the importance of factors which affect the overall performance, and also identify factors that could be seen as detrimental. However, the authors have adopted the use of ADGI but implemented it in a not so conventional way. Much like the work the authors did with regards to natural language processing using grey analysis in [12], the ADGI was utilised to measure the metric spaces of the sequence curves for an optimal string, against the input strings. The returned degree of incidence scored the overall similarity, the higher the value was to 1, the greater the similarity of the two strings.

The traditional degree of grey incidence provides the basis for all variances of the degree of incidence; $\Gamma = [\gamma_{ij}]$, where each entry in the ith row of the matrix is the degree of grey incidence for the corresponding characteristic sequence Y_i, and relevant behavioural factors $X_1, X_2, \ldots, X_m$. Each entry for the jth column is reference to the degrees of grey incidence for the characteristic sequences $Y_1, Y_2, \ldots, Y_n$ and behavioural factors X_m. The ADGI $A = [\epsilon_{ij}]_{n \times m}$, is defined as follows:

Definition 4 (*Absolute degree of grey incidence [13][10]*): Assume that X_i and $X_j \in \mathbb{U}$ are two sequences of data with the same magnitude, that are defined as the sum of the distances between two consecutive time points, whose zero starting points have already been computed:

$$s_i = \int_1^n (X_i - x_i(1)) dt \tag{11}$$

$$s_i - s_j = \int_1^n (X_i^0 - X_j^0) dt \tag{12}$$

Therefore the absolute degree of incidence is given as:

$$\epsilon_{ij} = \frac{1 + |s_i| + |s_j|}{1 + |s_i| + |s_j| + |s_i - s_j|} \tag{13}$$

This will provide the coefficient values between the contained clusters belonging to the criteria set C.

III. OBSERVATIONS

This section will bring together the use of RFS, the significance measure and ADGI. An example is put forward to further explain the advantages of such a framework.

Example 1: Given that $F = \{f_1, f_2, \ldots, f_9\}$ is a set containing 9 different colour swatches, all of which are a variations on the colour red:

$f_1 \rightarrow \qquad [204, 0, 0] \rightarrow$

$f_2 \rightarrow \qquad [153, 0, 0] \rightarrow$

$f_3 \rightarrow \qquad [255, 102, 102] \rightarrow$

$f_4 \rightarrow \qquad [51, 0, 0] \rightarrow$

$f_5 \rightarrow \qquad [255, 153, 153] \rightarrow$

$f_6 \rightarrow \qquad [102, 0, 0] \rightarrow$

$f_7 \rightarrow \qquad [255, 204, 204] \rightarrow$

$f_8 \rightarrow \qquad [255, 0, 0] \rightarrow$

$f_9 \rightarrow \qquad [255, 51, 51] \rightarrow$

The colours themselves are given by their [RGB] values, from which the average is worked out and stored in N. The values contained are given as $N = \{68, 51, 153, 17, 187, 34, 221, 85, 119\}$. Each average N_i value will correspond to a specific colour swatch F_i. For example, the swatch associated with f_3 has a value of 153, f_5 will be related to 187, and so on. Assume that the criteria set $C = \{p_1, p_2, \ldots, p_{15}\}$ contains the perceptions of 15 individuals, all of whom gave their own opinions based on the available descriptors and the swatches themselves. These values have been collected, along with their ages and are presented in Table I.

The terms contained within the table can be understood as meaning:

$$LR \rightarrow \text{Light Red} \qquad R \rightarrow \text{Red} \qquad DR \rightarrow \text{Dark Red}$$

The fuzzy membership set J_x is created using a simple linear function:

$$\mu(f_i) = \frac{l_i - l_{min}}{l_{max} - l_{min}} \tag{14}$$

The resulting fuzzy membership set is given as follows:

$$J_x = \{0.25, 0.17, 0.67, 0.00, 0.83, 0.08, 1.00, 0.33, 0.50\}$$

Using *Definition 2*, the final generated RFS based on the collected subsets for LR, R and DR, respectively, are given as:

$$LR = (\{0.67, 0.83, 1.00\}, \{0.50, 0.67, 0.83, 1.00\})$$

$$R = (\{0.25, 0.33\}, \{0.25, 0.33, 0.50\})$$

$$DR = (\{0.00, 0.08, 0.17\}, \{0.00, 0.08, 0.17\})$$

TABLE II. THE DEGREES OF SIGNIFICANCE FOR EACH OF THE GENERATED R-FUZZY SETS BASED ON TABLE I

LR		R		DR	
J_x	γ	J_x	γ	J_x	γ
$\gamma_{\overline{LR}}\{0.00\}=$	0.00	$\gamma_{\overline{R}}\{0.00\}=$	0.00	$\gamma_{\overline{DR}}\{0.00\}=$	1.00
$\gamma_{\overline{LR}}\{0.08\}=$	0.00	$\gamma_{\overline{R}}\{0.08\}=$	0.00	$\gamma_{\overline{DR}}\{0.08\}=$	1.00
$\gamma_{\overline{LR}}\{0.17\}=$	0.00	$\gamma_{\overline{R}}\{0.17\}=$	0.00	$\gamma_{\overline{DR}}\{0.17\}=$	1.00
$\gamma_{\overline{LR}}\{0.25\}=$	0.00	$\gamma_{\overline{R}}\{0.25\}=$	1.00	$\gamma_{\overline{DR}}\{0.25\}=$	0.00
$\gamma_{\overline{LR}}\{0.33\}=$	0.00	$\gamma_{\overline{R}}\{0.33\}=$	1.00	$\gamma_{\overline{DR}}\{0.33\}=$	0.00
$\gamma_{\overline{LR}}\{0.50\}=$	0.53	$\gamma_{\overline{R}}\{0.50\}=$	0.47	$\gamma_{\overline{DR}}\{0.50\}=$	0.00
$\gamma_{\overline{LR}}\{0.67\}=$	1.00	$\gamma_{\overline{R}}\{0.67\}=$	0.00	$\gamma_{\overline{DR}}\{0.67\}=$	0.00
$\gamma_{\overline{LR}}\{0.83\}=$	1.00	$\gamma_{\overline{R}}\{0.83\}=$	0.00	$\gamma_{\overline{DR}}\{0.83\}=$	0.00
$\gamma_{\overline{LR}}\{1.00\}=$	1.00	$\gamma_{\overline{R}}\{1.00\}=$	0.00	$\gamma_{\overline{DR}}\{1.00\}=$	0.00

TABLE III. THE DEGREES OF SIGNIFICANCE FOR EACH OF THE GENERATED R-FUZZY SETS FOR THE CLUSTER -20 YEAR OLDS

LR		R		DR	
J_x	γ	J_x	γ	J_x	γ
$\gamma_{\overline{LR}}\{0.00\}=$	0.00	$\gamma_{\overline{R}}\{0.00\}=$	0.00	$\gamma_{\overline{DR}}\{0.00\}=$	1.00
$\gamma_{\overline{LR}}\{0.08\}=$	0.00	$\gamma_{\overline{R}}\{0.08\}=$	0.00	$\gamma_{\overline{DR}}\{0.08\}=$	1.00
$\gamma_{\overline{LR}}\{0.17\}=$	0.00	$\gamma_{\overline{R}}\{0.17\}=$	0.00	$\gamma_{\overline{DR}}\{0.17\}=$	1.00
$\gamma_{\overline{LR}}\{0.25\}=$	0.00	$\gamma_{\overline{R}}\{0.25\}=$	1.00	$\gamma_{\overline{DR}}\{0.25\}=$	0.00
$\gamma_{\overline{LR}}\{0.33\}=$	0.00	$\gamma_{\overline{R}}\{0.33\}=$	1.00	$\gamma_{\overline{DR}}\{0.33\}=$	0.00
$\gamma_{\overline{LR}}\{0.50\}=$	1.00	$\gamma_{\overline{R}}\{0.50\}=$	0.00	$\gamma_{\overline{DR}}\{0.50\}=$	0.00
$\gamma_{\overline{LR}}\{0.67\}=$	1.00	$\gamma_{\overline{R}}\{0.67\}=$	0.00	$\gamma_{\overline{DR}}\{0.67\}=$	0.00
$\gamma_{\overline{LR}}\{0.83\}=$	1.00	$\gamma_{\overline{R}}\{0.83\}=$	0.00	$\gamma_{\overline{DR}}\{0.83\}=$	0.00
$\gamma_{\overline{LR}}\{1.00\}=$	1.00	$\gamma_{\overline{R}}\{1.00\}=$	0.00	$\gamma_{\overline{DR}}\{1.00\}=$	0.00

By using Eq. (8), one is able to calculate the degree of significance for each and every encapsulated fuzzy membership value, from J_x that has an affinity to its RFS. The returned degree of significance for all generated RFS are presented in Table II.

Fig. 1 collectively displays all the generated RFS for *Example 1*, along with the degree of significance for each fuzzy membership value, in accordance to its relative association to each RFS. Referring back to Table I, one will see that the age of each individual was also collected; there are three ages of note: 20, 25 and 30 (5 from each criterion). As these are all from the same original criteria set C, the membership set J_x will stay the same, we do not need to recalculate. We can now reconstruct our RFS so that they now represent the three individual age groups collected in Table I. From this we can then compare the perspective of each age cluster to that of another, gaining a more detailed understanding of the concept being modelled.

Furthermore, as the membership set J_x does indeed remain the same, we can use this as the sequence needed for the ADGI component. The membership values themselves act as the discretised points along the x axis, whereas the *varying* significance degrees give the associated amplitude. For example, if one refers to Table IV which contains the data for the age cluster 25 year olds, the membership value 0.50 for the RFS LR, has a returned degree of significance of 0.60. The same membership value and RFS for the age cluster 20 year olds in Table III, returns a degree of significance of 1.00. The same associated degree of significance for the age cluster 30 year olds in Table V, returns a 0.00.

Regardless of how small or large the difference between the returned degrees of significance for comparable fuzzy membership values, the fact that there can be a difference should provide one the motivation to explore further. It is precisely this aspect of wanting to investigate that warrants the use of the ADGI. Since we now have established sequences indicative of the fuzzy membership set J_x, we can now measure the difference between the metric spaces of comparable sequences based on the returned degrees of significance.

TABLE IV. THE DEGREES OF SIGNIFICANCE FOR EACH OF THE GENERATED R-FUZZY SETS FOR THE CLUSTER -25 YEAR OLDS

LR		R		DR	
J_x	γ	J_x	γ	J_x	γ
$\gamma_{\overline{LR}}\{0.00\}=$	0.00	$\gamma_{\overline{R}}\{0.00\}=$	0.00	$\gamma_{\overline{DR}}\{0.00\}=$	1.00
$\gamma_{\overline{LR}}\{0.08\}=$	0.00	$\gamma_{\overline{R}}\{0.08\}=$	0.00	$\gamma_{\overline{DR}}\{0.08\}=$	1.00
$\gamma_{\overline{LR}}\{0.17\}=$	0.00	$\gamma_{\overline{R}}\{0.17\}=$	0.00	$\gamma_{\overline{DR}}\{0.17\}=$	1.00
$\gamma_{\overline{LR}}\{0.25\}=$	0.00	$\gamma_{\overline{R}}\{0.25\}=$	1.00	$\gamma_{\overline{DR}}\{0.25\}=$	0.00
$\gamma_{\overline{LR}}\{0.33\}=$	0.00	$\gamma_{\overline{R}}\{0.33\}=$	1.00	$\gamma_{\overline{DR}}\{0.33\}=$	0.00
$\gamma_{\overline{LR}}\{0.50\}=$	0.60	$\gamma_{\overline{R}}\{0.50\}=$	0.40	$\gamma_{\overline{DR}}\{0.50\}=$	0.00
$\gamma_{\overline{LR}}\{0.67\}=$	1.00	$\gamma_{\overline{R}}\{0.67\}=$	0.00	$\gamma_{\overline{DR}}\{0.67\}=$	0.00
$\gamma_{\overline{LR}}\{0.83\}=$	1.00	$\gamma_{\overline{R}}\{0.83\}=$	0.00	$\gamma_{\overline{DR}}\{0.83\}=$	0.00
$\gamma_{\overline{LR}}\{1.00\}=$	1.00	$\gamma_{\overline{R}}\{1.00\}=$	0.00	$\gamma_{\overline{DR}}\{1.00\}=$	0.00

TABLE V. THE DEGREES OF SIGNIFICANCE FOR EACH OF THE GENERATED R-FUZZY SETS FOR THE CLUSTER -30 YEAR OLDS

LR		R		DR	
J_x	γ	J_x	γ	J_x	γ
$\gamma_{\overline{LR}}\{0.00\}=$	0.00	$\gamma_{\overline{R}}\{0.00\}=$	0.00	$\gamma_{\overline{DR}}\{0.00\}=$	1.00
$\gamma_{\overline{LR}}\{0.08\}=$	0.00	$\gamma_{\overline{R}}\{0.08\}=$	0.00	$\gamma_{\overline{DR}}\{0.08\}=$	1.00
$\gamma_{\overline{LR}}\{0.17\}=$	0.00	$\gamma_{\overline{R}}\{0.17\}=$	0.00	$\gamma_{\overline{DR}}\{0.17\}=$	1.00
$\gamma_{\overline{LR}}\{0.25\}=$	0.00	$\gamma_{\overline{R}}\{0.25\}=$	1.00	$\gamma_{\overline{DR}}\{0.25\}=$	0.00
$\gamma_{\overline{LR}}\{0.33\}=$	0.00	$\gamma_{\overline{R}}\{0.33\}=$	1.00	$\gamma_{\overline{DR}}\{0.33\}=$	0.00
$\gamma_{\overline{LR}}\{0.50\}=$	0.00	$\gamma_{\overline{R}}\{0.50\}=$	1.00	$\gamma_{\overline{DR}}\{0.50\}=$	0.00
$\gamma_{\overline{LR}}\{0.67\}=$	1.00	$\gamma_{\overline{R}}\{0.67\}=$	0.00	$\gamma_{\overline{DR}}\{0.67\}=$	0.00
$\gamma_{\overline{LR}}\{0.83\}=$	1.00	$\gamma_{\overline{R}}\{0.83\}=$	0.00	$\gamma_{\overline{DR}}\{0.83\}=$	0.00
$\gamma_{\overline{LR}}\{1.00\}=$	1.00	$\gamma_{\overline{R}}\{1.00\}=$	0.00	$\gamma_{\overline{DR}}\{1.00\}=$	0.00

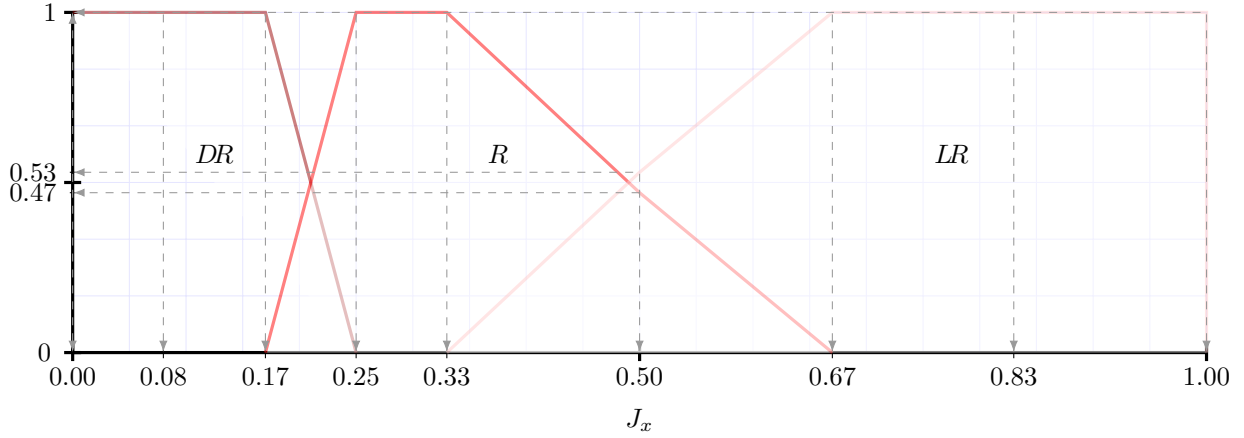Fig. 2 provides a visualisation of the comparison between the RFS and significance measure sequence generated for

Fig. 1. A continuous visualisation for *Example 1*, based on all the generated significance measures for the R-fuzzy sets of *Dark Red*, *Red* & *Light Red*
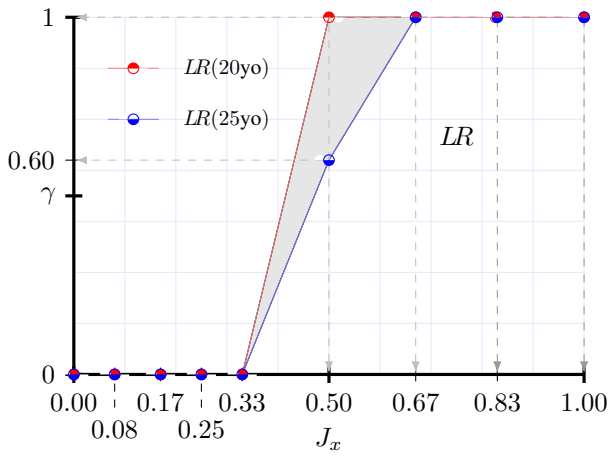

Fig. 2. The comparability between two *LR* R-fuzzy sets, one generated for the age cluster 20 year olds, the other, 25 year olds
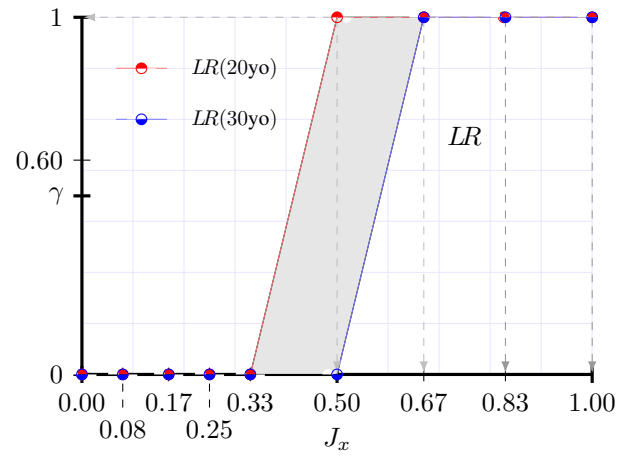

Fig. 3. The comparability between two *LR* R-fuzzy sets, one generated for the age cluster 20 year olds, the other, 30 year olds

LR, with relation to 20 year olds from Table III, against the RFS and significance measure sequence generated for *LR*, with relation to 25 year olds from Table IV. The shaded area of the plot signifies the difference in the metric spaces between the two sequences. The greater the difference the smaller the ADGI value, equally, the more similar the sequences, the greater the value. Using *Definition 4* one can now apply the ADGI to quantify the overall difference between the two sequences based on their metric spaces. As the ADGI uses *absolute* values, the order in which the sequences are passed through will not alter the final results. The sequence associated with *LR* for 20 year olds is given as: $s_i = \{0.00, 0.00, 0.00, 0.00, 0.00, 1.00, 1.00, 1.00, 1.00\}$. The sequence associated with *LR* for 25 year olds is given as: $s_j = \{0.00, 0.00, 0.00, 0.00, 0.00, 0.60, 1.00, 1.00, 1.00\}$. Using Eq. (13), the returned result is $\epsilon(0.950)$. A very high scoring value, indicating that the two sequences are indeed very similar. Referring back to Fig. 2, it can quite easily be inferred that the similarities are in indeed there, the sequences score almost exactly the same for their returned degrees of significance, except for one point. The fuzzy membership value of $J_x\{0.50\}$ differs between the two RFS for the two different age clusters. Comparing the RFS and associated significance

measure sequences with that of differing age clusters, one can quantify the rate of change in perception as we propagate through each cluster.

Fig. 3 provides a visualisation of the comparison between the RFS and significance measure sequence generated for *LR*, with relation to 20 year olds, against the RFS and significance measure sequence generated for *LR*, with relation to 30 year olds from Table V. In exactly the same way as before, the sequences are based on the returned degrees of significance and are compared using Eq. (13). In this case, the sequence associated with *LR* for 30 year olds is given as: $s_j = \{0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 1.00, 1.00, 1.00\}$, and the returned ADGI is $\epsilon(0.875)$. Another high scoring value but not as high as the comparison between the RFS and significance measures for 20 and 25 year olds. This comparing and contrasting is repeated for all clusters against all generated RFS and significance measure sequences. Table VI provides a summary of the collected ADGI values for all generated RFS and significance measure sequences, against each of the three age clusters. One can see that the age clusters 20 and 25 share the closest correlated perception for the concept *LR*, with an ADGI value of $\epsilon(0.950)$. The age clusters 20 and 30 share the

TABLE VI. A COMPARABLE SUMMARY OF THE ABSOLUTE DEGREE OF GREY INCIDENCES FOR THE R-FUZZY SETS: LR, R AND DR, AGAINST THE AGE CLUSTERS: 20, 25 AND 30

LR	20yo	25yo	30yo	R	20yo	25yo	30yo	DR	20yo	25yo	30yo
20yo	$\epsilon(1.00)$	$\epsilon(0.950)$	$\epsilon(0.875)$	20yo	$\epsilon(1.00)$	$\epsilon(0.931)$	$\epsilon(0.857)$	20yo	$\epsilon(1.00)$	$\epsilon(1.00)$	$\epsilon(1.00)$
25yo	-	$\epsilon(1.00)$	$\epsilon(0.916)$	25yo	-	$\epsilon(1.00)$	$\epsilon(0.914)$	25yo	-	$\epsilon(1.00)$	$\epsilon(1.00)$
30yo	-	-	$\epsilon(1.00)$	30yo	-	-	$\epsilon(1.00)$	30yo	-	-	$\epsilon(1.00)$

least correlated perception with an ADGI value of $\epsilon(0.875)$. The age clusters 25 and 30 share an intermediary correlation with an ADGI value of $\epsilon(0.916)$.

Applying the same level of inspection, one can see that the correlation for each R significance measure sequence, between 20 and 25 years, is again stronger than between 20 and 30 year olds. With regards to the DR, it is logical to infer that all clusters correlated exactly with one another. Simply inspecting Table III, Table IV and Table V, one can see that each significance measure sequence is absolutely the same, meaning for this instance, all age clusters agreed upon the RFS that was created for DR.

IV. Conclusion

The RFS and significance measure hybrid approach is further improved upon by the work presented in this paper. Given that an RFS allows for the encapsulation of a general consensus and also individual perspectives, the wealth of information an RFS can contain is a great deal. The introduction of the significance measure by the authors in [2], [3] has allowed for the R-fuzzy concept to model more complex uncertainty, returning a higher dimension of results for better inferencing. With the introduction of grey analysis, specifically the use of the ADGI, it has been shown that even more information and inference can be garnered from the same initial data set.

In the example presented in this paper, the data contained three separate age groups. The use of the ADGI allowed for one to further inspect the change in perception as one propagated through each cluster. If changes did occur, the corresponding significance measure sequence when compared to that of another, would quantify the amount of difference between the metric spaces of the sequences. As the sequences are all based on the degrees of significance for the same fuzzy membership values in J_x, the magnitude for each sequence is guaranteed to be the same with the one to be compared against. It's noteworthy to extend a mention to what was being compared. For example, this paper compared and contrasted RFS LR from the age cluster 20, with that of the RFS LR from the age cluster 25. Although it is completely acceptable to compare LR from the age cluster 20, with that of the RFS DR from the age cluster 25, it would not make logical sense, as they are two different abstract concepts being inspected. RFS of the same likeness should be the only concern for when conducting comparisons. If the data also contained the sex of the individuals, this too could also be compared, contrasted and inspected. This would identify, if any, the change in perceptions for a given observation. Ethnicity, geographical location, occupation and so on, all varying combinations of facets could be factored into the data set. The robustness of RFS means that varying numbers participants can be factored into the criteria set C, not just equal quantities. In much the same way, the RFS and significance measure approach, along

with the ADGI would allow for even greater levels of detail to be obtained. The more understood the problem, the better informed the solution.

References

[1] Y. Yang and C. Hinde, "A new extension of fuzzy sets using rough sets: R-fuzzy sets," *Information Sciences*, vol. 180, no. 3, pp. 354 – 365, 2010.

[2] A. S. Khuman, Y. Yang, and R. John, "A significance measure for r-fuzzy sets," in *Fuzzy Systems (FUZZ-IEEE), 2015 IEEE International Conference on*, Aug 2015, pp. 1–6.

[3] A. Khuman, Y. Yang, and R. John, "Quantification of r-fuzzy sets," *Expert Systems with Applications*, vol. 55, pp. 374 – 387, 2016.

[4] A. S. Khuman, Y. Yang, and R. John, "R-fuzzy sets and grey system theory," in *2016 IEEE International Conference on Systems, Man, and Cybernetics*, Oct 2016, p. in press.

[5] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 20, no. 1, pp. 87 – 96, 1986.

[6] W. Pedrycz, "Shadowed sets: representing and processing fuzzy sets," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 28, no. 1, pp. 103–109, 1998.

[7] R. Sambuc, *Fonctions Φ foues*. Application a l'aide au Diagnostic en Pathololologie Thyroidienne. PhD thesis, Univ. Marseille, France, 1975.

[8] J. Mendel and R. John, "Type-2 fuzzy sets made simple," *Fuzzy Systems, IEEE Transactions on*, vol. 10, no. 2, pp. 117–127, Apr 2002.

[9] D. Ju-Long, "Control problems of grey systems," *Systems & Control Letters*, vol. 1, no. 5, pp. 288 – 294, 1982.

[10] S. Liu and Y. Lin, *Grey Information: Theory and Applications*, 1st ed., ser. Advanced Information and Knowledge Processing. Springer-Verlag London, 2006.

[11] Z. Pawlak, "Rough set theory and its applications to data analysis," *Cybernetics & Systems*, vol. 29, no. 7, pp. 661–688, 1998.

[12] A. S. Khuman, Y. Yang, and S. Liu, "Grey relational analysis and natural language processing," in *2015 IEEE International Conference on Grey Systems and Intelligent Services (GSIS)*, Aug 2015, pp. 107–112.

[13] L. Sifeng, J. Forrest, and Y. Yingjie, "A brief introduction to grey systems theory," in *Grey Systems and Intelligent Services (GSIS), 2011 IEEE International Conference on*, Sept 2011, pp. 1–9.