Empirical study of fuzzy quantification models for linguistic descriptions of meteorological data

Carlos Heble-Lahera*, Andrea Cascallar-Fuentes†, Alejandro Ramos-Soto† and Alberto Bugarín Diz†
* Máster Universitario en Investigación en Inteligencia Artificial, Universidad Internacional Menéndez Pelayo
†Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS), Universidade de Santiago de Compostela
Emails: *carlos.heble@posgrado.uimp.es, † {andrea.cascallar.fuentes, alejandro.ramos, alberto.bugarin.diz}@usc.es

***Abstract*—In this work we present an experimental comparison of six widely used quantification methods (Zadeh's scalar and fuzzy cardinality, Yager's OWA, Delgado's GD, Sugeno integral and Vila's VQ) when evaluating Type-1 and Type-2 linguistic descriptions of data generated from meteorological data provided by the Galician Meteorological Agency MeteoGalicia. The objective of this study is to evaluate if there are significant differences among these models for the data considered. We ranked the generated descriptions based on their degree of truth for each quantification model and we analyzed those results calculating the Pearson correlation coefficient. Results show that there are not significant differences in the models when evaluating Type-1 descriptions. However, in Type-2 evaluation the methods can be grouped in three clusters with a significantly different behavior among them: *i)* Zadeh's scalar cardinality, Delgado's GD and Zadeh's fuzzy cardinality, *ii)* Yager's method and *iii)* Vila's VQ.**

***Index Terms*—fuzzy quantification, linguistic descriptions of data, natural language generation**

I. Introduction

Nowadays, the analysis and interpretation of data is becoming an increasingly difficult task for humans due to its exponential growth. Therefore, computational methods that can perform these tasks are in high demand.

In the natural language generation field (NLG), many systems have been developed in order to generate comprehensible texts with useful information from several data sources [1].

Besides, in the fuzzy logic field, several approaches were proposed to generate data descriptions using linguistic terms. Following Zadeh's computing with words and perceptions paradigms [2], [3], linguistic descriptions of data (LDD) [4] summarize in a linguistic form one or more numerical variables and their values, using the general notion of protoform [5]. These protoforms can follow several structure types (e.g temporal or comparative [6], [7]), being Type-1 and Type-2 fuzzy quantified statements [8]–[10] with absolute or relative quantifiers the most common in the literature (e.g. *"In some places the temperature is low"*).

Type-1 descriptions have the following structure: *"Q X are A"*, where *Q* is a linguistic quantifier, *X* is a linguistic variable defined on a given referential and *A* is a fuzzy linguistic value (property) of *X*. For instance, in *"Most temperatures are normal," "Most"* is the quantifier, *"temperatures"* is the linguistic variable and *"normal"* is a linguistic value of *temperatures*. Type-2 descriptions follow the structure *"Q DX are A"* where an additional fuzzy property *"D"* is defined on the same referential of X. For example, in *Most temperatures in the North are normal "North"* is the additional fuzzy property (D).

Relative quantifiers express the proportion of elements over the total which fulfill a condition, e.g *"**half** locations have low temperature"*. Absolute quantifiers express quantities over the total of elements which fulfill a condition, e.g *"**15** locations have low temperature"*.

Evaluating a quantified sentence involves obtaining its truth value (a value in the range [0, 1]) on a given data set. It is obtained by calculating the compatibility between the number of elements in the referential which fulfill the sentence (its cardinality) and the quantifier in the sentence. Therefore, this compatibility measure depends on the data, the quantifier definition and the linguistic terms defined from the properties in the referential. Besides, it also depends on the quantification model used for evaluating the sentence. There exist several quantification methods [11]–[16] in the literature that differ from each other in the way they calculate the truth value.

The aim of this work is to experimentally compare six widely used quantification methods for the evaluation of Type-1 and Type-2 quantified descriptions, in order to assess their empirical behavior when applied for the evaluation of fuzzy quantified sentences. To generate these descriptions we use meteorological data provided by the Galician (NW Spain) Meteorological Agency (MeteoGalicia) [17].

Our objective in this work is to perform a comparison among the selected fuzzy quantification models described in Section II with the aim of analyze the correlation between them when evaluating type-1 and type-2 quantified statements. If differences between the methods behavior are detected, the selection of a fuzzy quantification model for a specific case should take this difference into consideration. Conversely, if the results show a similar behavior between the models, the selection of a fuzzy quantification model for an specific case should follow another criteria (e.g., its theoretical properties or computational cost).

This paper is structured as follows: in Section II we describe the methods included in this experimentation. In Section III we describe the used data set and the definition of the linguistic variables for the quantified sentences generation. In Section IV we describe the experiments we performed for comparing the quantification methods behavior. Finally, Section V includes some final remarks.

978-1-7281-6932-3/20/$31.00 ©2020 IEEE

II. Fuzzy quantification methods

In this section, we present the set of quantification methods empirically compared in this study. The selection of these methods rely on their presence in the literature and their different cardinalities, since this feature can have such a great impact on the quantification methods performance.

The first approaches for fuzzy quantification were proposed by Zadeh [11], who identified the need of extending the concept of the quantifiers *exist* and *all* to other imprecise ones, with more expressiveness.

Evaluating a quantified sentence involves calculating its truth value. In this evaluation, two aspects have to be considered: *i)* the (fuzzy or scalar) cardinality, i.e., how many elements in the referential fulfill belong to the fuzzy set defined from a variable and *ii)* the compatibility between the cardinality measure and the quantifier. Quantification models differ in the type of cardinality and in how the compatibility is evaluated.

Several studies [15], [16], [18]–[21] theoretically analyzed quantification methods by checking the properties they fulfill. These analysis show that, in general, quantifiers should behave (very) differently, although, for Type-1 descriptions and coherent[1] quantifiers, it was theoretically proved [14] that GD [14] is a generalization of Yager's method [12] and ZS [15] is a generalization of the Sugeno integral base method [13].

Aiming to go beyond of these theoretical results, to the best of our knowledge, there are no experimental analysis that assess the behavior of quantification models from an empirical point of view.

A. Type-1 models

As indicated in the previous sections, Type-1 descriptions follow the "Q X are A" protoform.

1) Zadeh's method: [11] is based on the scalar cardinality "power" defined by Zadeh as $P(A)=\sum_{i=1}^{n} A(x_i)$.

The evaluation of Type-1 sentences for relative quantifiers is defined (for the t-norm minimum, the most popular operator used in the literature) as:

$$Z_Q(A)=Q\left(\frac{P(A)}{|X|}\right) \quad (1)$$

2) Yager's method based on OWA operators: [12] can only be used with coherent and relative quantifiers. The evaluation is:

$$Y_Q(A)=\sum_{i=1}^{n} A(w_i b_i) \quad (2)$$

where b_i is the i-th higher value of the degree of truth to the fuzzy set A and w_i a coefficient obtained between the quantifier and $|X|$ to ensure that the evaluation is coherent.

[1] A quantifier Q is called coherent (or monotonically nondecreasing quantifiers [12]) if $Q(x_i)\leq Q(x_{i+1}) \forall x_i < x_{i+1}$, $Q(0)=0$, $Q(1)=1$ [14]

3) Sugeno integral based method: [13] is another method to evaluate quantified sentences which also requires coherent quantifiers. In the relative quantifier case, the evaluation is:

$$S_Q(A)=\max_{1\leq i\leq n} \min\left(Q\left(\frac{P(A)}{|X|}\right)\right) \quad (3)$$

4) Delgado's GD method: [14] uses a fuzzy cardinality E.

The evaluation of a Type-1 description with relative quantifiers is as follows:

$$GD_Q(A)=\sum_{i=0}^{n} ED(A,i)\times Q\left(\frac{i}{n}\right) \quad (4)$$

where any t-norm and t-conorm can be used. In our evaluation, we selected the product t-norm and Łukasiewicz's t-conorm, which are the ones defined in [14].

On the other hand, $ED(A,k)=b_k-b_{k+1}$ with $b_0=1$ and $b_{n+1}=0$ is a particular case of the E cardinality, using the minimum t-norm, Łukasiewicz's t-norm, the maximum t-conorm and the standard negation. Considering a set of $\alpha-cuts$ of A plus $\emptyset$, the possibility that a A_α is a subset of A is α and an integer k with $1\leq k\leq n$ such that $\alpha=b_k$.

5) ZS method: This method [15] is based on Zadeh's fuzzy cardinality:

$$Z(A,k)=\begin{cases} 0 & \text{if } \nexists\alpha \mid |A_\alpha|=k, \\ sup\{\alpha \mid |A_\alpha|=k\} & otherwise \end{cases} \quad (5)$$

Its evaluation for relative quantifiers is:

$$ZS_Q=\max_{k\in\{0,\ldots,n\}} \min\left(Z(A,k),Q\left(\frac{k}{n}\right)\right) \quad (6)$$

B. Type-2 models

As indicated in the previous sections, Type-2 descriptions follow the "Q DX are A" protoform.

1) Zadeh's method: [11] defined (as in Type-1, for the minimum t-norm) as:

$$Z_Q(A/D)=Q\left(\frac{P(A\cap D)}{P(D)}\right) \quad (7)$$

where $P(A\cap D)=\sum_{i=1}^{n} A(x_i)\wedge D(x_i)$.

2) Yager's method based on OWA operators: [12] which can only be generalized to Type-2 sentences using coherent and relative quantifiers.

The evaluation is:

$$Y_Q(A/D)=\sum_{i=1}^{n}=w_i c_i \quad (8)$$

where w_i is a coefficient obtained between the quantifier and $|X|$ to ensure that the evaluation is coherent calculated as follows:

$$w_i=Q(S_i)-Q(S_{i-1})\ i\in\{1,\ldots,n\} \quad (9)$$

and $S_0=0$ and c_i is the i-th high value of the $\neg D\vee A$ set's truth value.

3) Vila, Cubero, Medina and Pons' method: [16] uses the "or" or "orness" degree defined for coherent quantifiers. $orness(\exists) = 1$ and $orness(\forall) = 0$.

The evaluation for a Type-2 sentence is:

$$\begin{aligned} V_Q(A/D) = o_Q \max_{x \in X}(D(x) \wedge A(x)) \\ +(1 - o_Q) \min_{x \in X}(A(x) \vee (1 - D(x))) \end{aligned} \tag{10}$$

where o_Q is the orness degree. The definition on this method uses the minimum as t-norm and the maximum as t-conorm so we also selected this criteria in our evaluation.

4) Delgado's GD method: The generalization of this method [14] is defined as the compatibility between the ER cardinality and the quantifier by means of te product and the Łukasiewicz's t-conorm, as follows:

$$GD_Q(A/D) = \sum_{c \in CR(A/D)} ER(A/D, c) \times Q(c) \tag{11}$$

where

$$CR(A/D) = \{\frac{|(A \cap D)_\alpha|}{|D_\alpha|}\ with\ \alpha\ \in\ M(A/D)\} \tag{12}$$

and

$$\begin{aligned} M(A/D) = M(A \cap D) \cup M(D),\ and \\ M(A) = \{\alpha \in (0, 1]\ |\exists x_i \in X\ with\ A(x_i) = \alpha\} \end{aligned} \tag{13}$$

5) ZS method: [15], [19] uses the fuzzy cardinality ES, which consists in a max-min composition between such cardinality and the quantifier, and can be defined as:

$$ZS_Q(A/D) = \max_{\alpha \in M(A/D)} \min(\alpha, Q\left(\frac{|(A \cap D)_\alpha|}{|D_\alpha|}\right) \tag{14}$$

III. Materials and methods

A. Data set

As mentioned above, in this study we evaluate Type-1 and Type-2 quantified sentences from meteorological data. A meteorological situation is calculated by complex numerical models including a high number of variables and is usually represented with maps, which are often not intuitive due to the high amount of icons used for representing the different weather situations.

In our case, we used data from a real-time observation service for each Galician council provided by the Galician Meteorology Agency (MeteoGalicia). This service provides information for the following meteorological variables of interest: sky state, wind and temperature. Figure 1 shows a meteorological real example map for these three variables. Both the sky and wind icons are the standard ones used in MeteoGalicia whereas in the temperature maps the possible values are: VL (very low), L (low), N (normal), H (high), VH (very high). These are printed in different colors from red, associated to high temperatures, to dark blue, associated to low temperature values.

1) Sky state: This variable describes the state of the sky based on two variables: cloud coverage and rainfall. Meteorologists labeled the values of this variable with 42 integer codes used to describe the day (21 integer numbers in the range [101, 121]) and the night situations (21 integer numbers in the range [201, 221]). For example, 101 means *"clear sky"* whereas 211 means *"night with clear sky"*.

2) Wind: This variable that comprises the wind direction and speed, and is labeled with integer 34 codes in the range [299, 332]. Meteorologists consider eight wind directions (N, S, E, W, NW, NE, SE, SW) combined with four wind speed values (weak, moderate, strong, very strong). Also the calm and variable direction situations are considered. For instance, 305 code means "South direction and weak speed".

3) Temperature: Represents the temperature in degrees Celsius.

B. Linguistic descriptions

In this section, we present the definition of the linguistic variables we used in this study, based on the three meteorological variables previously described.

1) Sky: It is treated as a crisp variable, therefore the values of the meteorological variable are the values of the resulting linguistic variable. Each value in the meteorological variable is defined as a singleton having an integer code as label and a degree of truth in the set $\{0, 1\}$.

2) Wind: It is also a crisp variable, so the resulting linguistic variable has values in the range [299, 332] defined as singletons. Likewise, each value has a fulfillment degree in the set $\{0, 1\}$.

3) Temperature: This numerical variable represents the temperature in degrees Celsius. We modeled the linguistic variable as a fuzzy variable with the following five labels: {"very low", "low", "normal", "high", "very high"}, which are defined as fuzzy sets.

In order to provide meaning and contextualization for these labels, meteorologists define a reference temperature value. This reference is taken from the average temperature, $\bar{x}$, and its standard deviation, σ, for the last twenty years, for each location and each month of the year. These two values were used to model the labels presented in Figure 2. For instance, the label "normal" is defined with the trapezoid with support $[\bar{x} - \sigma,\ \bar{x} + \sigma]$ and core $[\bar{x} - 0.5\sigma,\ \bar{x} + 0.5\sigma]$. Thus, for a specific location with, for instance, $\bar{x} = 14.2$ and $\sigma = 4.8$, its "normal" label has support [9.4, 19] and core [11.8, 16.6].

4) Quantifiers: As we mentioned above, we generate quantified sentences (e.g. *"In most locations the wind has North direction and moderate speed"*). Therefore, quantifiers are necessary to count the number of elements in the referential that fulfills the condition.

We defined the five quantifiers ("at least 25%", "at least 50%", "at least 75%", "most", "all"), represented in Figure 3. All of them are modeled as coherent (as defined in Section II-A), since some quantification methods (Yager's method, Sugeno integral based method, and Vila et al. method) only support quantifiers that fulfill this property.

(a) Sky state (b) Wind (c) Temperature

Fig. 1: Real Map representing the meteorological state for the three variables of interest

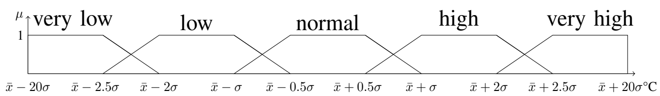

Fig. 2: Linguistic variable temperature definition.

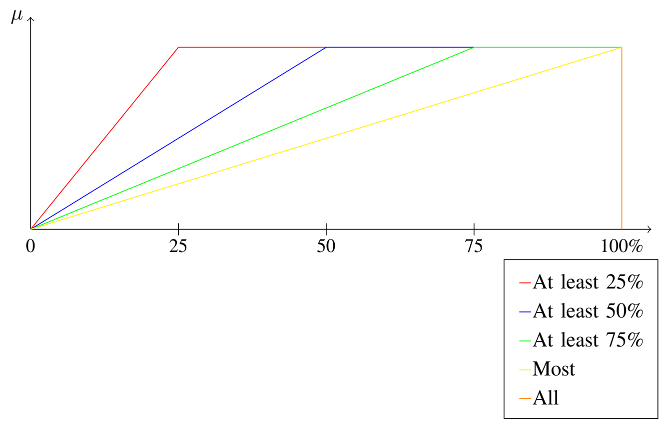

Fig. 3: Quantifiers definition for the percentage of locations.

5) Geographical descriptors: for Type-2 sentences we added a geographical qualifier. This allows us to describe smaller regions instead of the whole territory, which happens only in Type-1 descriptions. For instance, "In some locations in the North the temperature is low."

We defined nine linguistic geographical descriptors (N, S, E, W, Center, NE, NW, SE and SW), using longitude and latitude as reference as shown in Figure 4.

In the defined descriptors set, we can classify them into two subsets: simple or composite. A geographical descriptor is simple if only uses one dimension (latitude or longitude) in its definition, for instance "North" uses the latitude. On the contrary, a composite descriptor uses both dimensions, for example, "SW" uses both longitude and latitude.

With these components, we generated the quantified statements. In Type-1 descriptions (*"Q X are A"*) Q is the set of defined quantifiers, X is the set of described points and A is a combination of one or more of the linguistic variables created from the meteorological variables. For instance, "In some places the sky is clear and the temperature is high". Besides, in Type-2 K is the set of geographical descriptors previously defined. For example, "In most places in the North the sky is covered".

C. Experiments

The experimentation we performed in this study consisted mainly in two steps: *i)* generating the linguistic descriptions and *ii)* analyzing the quantification models behavior to evaluate their similarity. Figure 5 describes the performed stages in this experimentation and their corresponding inputs and outputs.

1) Linguistic descriptions generation: in the first stage, we generated all Type-1 and Type-2 descriptions from several real meteorological situations.

Firstly, we collected data from 15 different days and times of the day from 30th July 2019 to 30th August 2019. Thus, we ensured having a wide variety of situations, all independent from each other.

We generated all possible descriptions from this resulting data set, obtaining for each meteorological situation 45,145 Type-1 descriptions and 406,305 Type-2 descriptions.

For each generated sentence, we assessed how descriptive it was in terms of the meteorological situation described by the data. Thus, at this stage, we evaluated the descriptions with the selected quantification methods.

2) Quantification methods correlation comparison: We performed correlation coefficient tests in order to determine whether the quantification models have similar behaviors when evaluating quantified descriptions.

Among the different correlation tests, we selected the Pearson test, which is applied over the degree of truth, giving information about the similarity of quantification methods results, instead of ranking the descriptions based on the associated truth value. We performed this test in two stages: *i)* the test was performed over the data set from 30th July 2019 and *ii)* test was performed for each description over the average of evaluation result of the 15 data sets. Besides, in this second stage we performed two separate tests: *i)* applying the correlation coefficient test only to those quantified sentences which described the temperature and *ii)* applying test to quantified sentences which describe the three meteorological variables (temperature, sky state and wind). Sky state and wind values were filtered in order to only generate descriptions that

(a) Latitude definition.

(b) Longitude definition.

Fig. 4: Geographical descriptors definition.

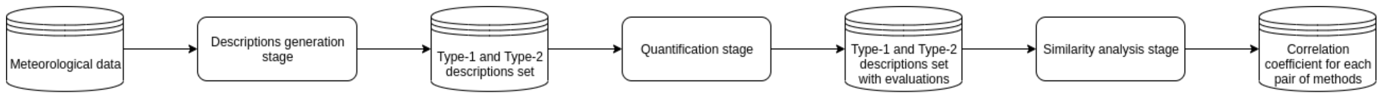

Fig. 5: Description of the experimentation stages with their inputs and outputs.

described actual situations in the data sets. So descriptions were generated with 5 different sky states (101, 103, 104, 105 and 111) and with 18 different wind states (from 300 to 316 and 318).

IV. Experimental results

In this section, we present the obtained results applying the Pearson correlation test to the different sets of descriptions for Type-1 and Type-2 descriptions.

A. Type-1 descriptions

1) 30 July 2019 results: results of the Pearson correlation coefficient with the 30 July 2019 data set (Table I) confirm there is a very high correlation between the quantification methods, with correlations higher than 0.9. A total correlation is obtained between the pairs Y_Q - GD_Q and S_Q - ZS_Q, which is consistent with [14], where it is stated that these methods coincide for coherent quantifiers.

	Z_Q	Y_Q	S_Q	GD_Q	ZS_Q
Z_Q	1	0.9995	0.9925	0.9995	0.9925
Y_Q		1	0.9945	1	0.9945
S_Q			1	0.9945	1
GD_Q				1	0.9945
ZS_Q					1

TABLE I: Pearson correlation coefficient for quantification methods evaluating Type-1 descriptions from 30 July 2019.

2) Average results describing only temperature: in Table II we present the results of the Pearson test with the resulting average data set describing only the temperature situation.

Also in this case, the correlation coefficient between all quantification methods pairs is very high and the total correlation between the pairs Y_Q - GD_Q and S_Q - ZS_Q was also obtained.

3) Average results including all meteorological variables: table III shows the correlations in this experiment. This experiment also confirms the conclusions of the previous tests. Therefore, there are not significant differences between the quantification methods when evaluating Type-1 quantified descriptions.

Analyzing these three different tests results, we can conclude when evaluating Type-1 descriptions the compared methods do not show differences between their behaviors since

	Z_Q	Y_Q	S_Q	GD_Q	ZS_Q
Z_Q	1	0.9983	0.9987	0.9998	0.9987
Y_Q		1	0.9992	1	0.9992
S_Q			1	0.9992	1
GD_Q				1	0.9992
ZS_Q					1

TABLE II: Pearson correlation coefficient for quantification methods evaluating Type-1 descriptions from the average data set describing temperature.

	Z_Q	Y_Q	V_Q	GD_Q	ZS_Q
Z_Q	1	0.9999	0.9976	0.9999	0.9976
Y_Q		1	0.9979	1	0.9979
S_Q			1	0.9979	1
GD_Q				1	0.9979
ZS_Q					1

TABLE III: Pearson correlation coefficient for quantification methods evaluating Type-1 descriptions from the average data set including all meteorological variables.

their correlation coefficient is higher than 0.9 between all models pairs. This is consistent with the analysis of properties described in [19] for Type-1 evaluation. We compile in Table IV a summary of the most relevant properties, where it can be seen that all methods fulfill the six properties considered, with the only exceptions of Z_Q (which does not fulfill properties 4 and 6) and Y_Q, which does not fulfill property 7.

Furthermore, the method pairs GD_Q - Y_Q and ZS_Q - S_Q had a correlation coefficient of 1, which are consistent with the theoretical previous work [14] when it is proved these pair of methods are equivalent with coherent quantifiers.

B. Type-2 descriptions

1) 30 July 2019 results: Results of the Pearson correlation coefficient (Table V) of the compared quantification models when evaluating Type-2 quantified descriptions show a high correlation between the methods Z_Q, GD_Q and ZS_Q. However, the correlation between V_Q with the others is lower than 0.85. Also Y_Q has low correlation coefficient when comparing its behavior with the other four methods.

2) Average results describing only temperature: results presented in Table VI, show high correlation between Z_Q, GD_Q, V_Q and ZS_Q. Also in this case, Y_Q have lower correlation with the others methods than when evaluating Type-1 descriptions.

Property	Description	Encoding as in [19]	Z_Q	Y_Q	GD_Q	S_Q	V_Q	ZS_Q
1	Absolute-relative transformation and correct generalization	I.1	✓	✓	✓	✓	NA	✓
2	Quantifiers monotonicity	I.3	✓	✓	✓	✓	NA	✓
3	Correct generalization	II.1	✓	✓	✓	NA	✓	✓
4	Coherence with fuzzy logic	I.2	X	✓	✓	✓	NA	✓
		II.6 (∃)	NI	X	✓	NA	✓	✓
		II.6 (∀)	NI	✓	X	NA	✓	✓
5	Computational complexity	I.4	O(n)	O(n log n)	O(n log n)	O(n log n)	NA	O(n log n)
		II.3	O(n)	O(n log n)	O(n log n)	NA	O(n)	O(n log n)
6	Not too strict evaluation	I.5	X	✓	✓	✓	NA	✓
		II.8	NI	✓	✓	NA	NI	✓
7	Valid for any quantifiers	I.6	✓	X	✓	✓	NA	✓
		II.7	✓	NI	✓	NA	✓	✓
8	Type-2 to Type-1 transformation	II.2	NI	NI	✓	NA	NI	✓
9	If D is a normal fuzzy set and $D \subseteq A$ the evaluation is 1	II.4	✓	NI	✓	NA	NI	✓
10	If $D \cap A = \emptyset$, the evaluation is 0	II.5	✓	X	✓	NA	NI	✓

TABLE IV: Recompilation of the properties analyzed in [19] with the corresponding encoding for Type-1 (I.x) and Type-2 (II.x) evaluations (as in [19], NI means no information is available and NA means property does not apply).

	Z_Q	Y_Q	V_Q	GD_Q	ZS_Q
Z_Q	1	0.3899	0.7610	0.9862	0.9396
Y_Q		1	0.3330	0.3811	0.3825
V_Q			1	0.7727	0.8338
GD_Q				1	0.9595
ZS_Q					1

TABLE V: Pearson correlation coefficient for quantification methods evaluating Type-1 descriptions from the 30 July 2019 data set describing temperature.

	Z_Q	Y_Q	V_Q	GD_Q	ZS_Q
Z_Q	1	0.7716	0.9694	0.9870	0.9806
Y_Q		1	0.7711	0.7508	0.7514
V_Q			1	0.9834	0.9850
GD_Q				1	0.9935
ZS_Q					1

TABLE VI: Pearson correlation coefficient for quantification methods evaluating Type-2 descriptions from the average data set describing temperature.

3) Average results including all meteorological variables: these experiments, as the performed with the 30 July 2019 data set, show a high correlation between Z_Q, GD_Q and ZS_Q whereas V_Q and Y_Q have lower correlation with the others. Results are presented in Table VII.

	Z_Q	Y_Q	V_Q	GD_Q	ZS_Q
Z_Q	1	0.2686	0.8168	0.9926	0.9759
Y_Q		1	0.2875	0.2547	0.2754
V_Q			1	0.8256	0.8703
GD_Q				1	0.9831
ZS_Q					1

TABLE VII: Pearson correlation coefficient for quantification methods evaluating Type-2 descriptions from the average data set including all meteorological variables.

The results of this tests set show lower correlation between the analyzed models in comparison with the Type-1 case. Besides, results show V_Q and Y_Q have low correlations with the compared methods whereas Z_Q, GD_Q and ZS_Q have a similar behavior. One possible explanation for the correlation between Z_Q and GD_Q is their type, since both models are sum-based methods. On the other hand, the correlated behavior between GD_Q and ZS_Q can be explained by their cardinality type, since both of them are based on fuzzy cardinalities.

Regarding the Type-2 properties described in Table IV, the three fuzzy quantification models share properties 7, 9 and 10. Besides, GD_Q and ZS_Q also fulfill properties 4 (∃) (ZS_Q also fulfills property 4 (∀)), 6 and 8 whereas Z_Q and ZS_Q share property 3.

Regarding Y_Q, this model is the only one which does not fulfill properties 4 (∃) and 10. This non-compliance could explain the low correlation between this method and the other four quantification methods.

On the other hand, V_Q fulfills properties 3, 4 (∃), 4 (∀) and 7. Therefore, the correlation between this method and Z_Q, GD_Q and ZS_Q could be explained by these shared properties.

V. Conclusions

In this paper, we presented an experimental study to compare the behavior of six well-known quantification methods when evaluating Type-1 and Type-2 quantified descriptions.

We performed our experiments using meteorological data provided by the Galician Meteorology Agency. Three different tests were performed both for Type-1 and Type-2 descriptions.

Test results show there are no significant differences between the behavior of the quantification models when evaluating Type-1 descriptions. Besides, our results are consistent with the theoretical results that state that GD_Q [14] and [12] methods, on one hand, and [15] and [13] methods, on the other, coincide for coherent quantifiers, since for these two pairs we obtained a correlation coefficient of 1.The results also prove a very similar behavior between all pairs of methods, with a correlation coefficient higher than 0.9 in all cases.

In the Type-2 scenario, correlation between the quantification models is lower than for Type-1, but methods can be clustered into three categories according to their correlation: *i)* Z_Q, GD_Q and ZS_Q, *ii)* V_Q and *iii)* Y_Q.

As future work, we are extending our experimentation in the Type-2 scenario in different ways: *i)* increasing the number of data sets in order to confirm whether the correlation between the quantification methods is independent of the data; *ii)* considering other definitions or partitions of quantifiers and other criteria if possible; and *iii)* extending the current experimentation including new quantification methods and extending the study of their properties and results to verify if their similarities are by chance or not.

ACKNOWLEDGMENTS

This research was funded by the Spanish Ministry for Science, Innovation and Universities (grants TIN2017-84796-C2-1-R and TIN2017-90773-REDT) and the Galician Ministry of Education, University and Professional Training (grants ED431C 2018/29, "accreditation 2016-2019, ED431G/08" and ED431G2019/04). All grants were co-funded by the European Regional Development Fund (ERDF/FEDER program).

REFERENCES

[1] A. Gatt and E. Krahmer, "Survey of the State of the Art in Natural Language Generation: Core tasks, applications and evaluation," *Journal of Artificial Intelligence Research*, vol. 61, pp. 65–170, 2018. [Online]. Available: https://doi.org/10.1613%2Fjair.5477

[2] L. A. Zadeh, "Fuzzy logic = computing with words," *IEEE Transactions on fuzzy systems*, vol. 4, no. 2, pp. 103–111, 1996. [Online]. Available: https://doi.org/10.1109%2F91.493904

[3] ——, "A new direction in AI: Toward a computational theory of perceptions," *AI magazine*, vol. 22, no. 1, p. 73, 2001. [Online]. Available: https://doi.org/10.1609/aimag.v22i1.1545

[4] J. Kacprzyk and R. R. Yager, "Linguistic summaries of data using fuzzy logic," *International Journal of General System*, vol. 30, no. 2, pp. 133–154, 2001.

[5] L. A. Zadeh, "A prototype-centered approach to adding deduction capability to search engines-the concept of protoform," in *Intelligent Systems, 2002. Proceedings. 2002 First International IEEE Symposium*, vol. 1. IEEE, 2002, pp. 2–3. [Online]. Available: https://doi.org/10.1109%2Fis.2002.1044219

[6] R. Castillo-Ortega, N. Marín, and D. Sánchez, "Time series comparison using linguistic fuzzy techniques," in *IPMU*, ser. Lecture Notes in Computer Science, vol. 6178. Springer, 2010, pp. 330–339.

[7] A. Wilbik and U. Kaymak, "Gradual linguistic summaries," in *IPMU (2)*, ser. Communications in Computer and Information Science, vol. 443. Springer, 2014, pp. 405–413.

[8] J. Kacprzyk and S. Zadrozny, "Computing with words is an implementable paradigm: fuzzy queries, linguistic data summaries, and natural-language generation," *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 3, pp. 461–472, 2010. [Online]. Available: https://doi.org/10.1109%2Ftfuzz.2010.2040480

[9] N. Marín and D. Sánchez, "On generating linguistic descriptions of time series," *Fuzzy Sets and Systems*, vol. 285, pp. 6–30, 2016. [Online]. Available: https://doi.org/10.1016%2Fj.fss.2015.04.014

[10] A. Alvarez-Alvarez and G. Trivino, "Linguistic description of the human gait quality," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 1, pp. 13–23, 2013. [Online]. Available: https://doi.org/10.1016%2Fj.engappai.2012.01.022

[11] L. A. Zadeh, "A computational approach to fuzzy quantifiers in natural languages," in *Computational linguistics*. Elsevier, 1983, pp. 149–184.

[12] R. R. Yager, "On ordered weighted averaging aggregation operators in multicriteria decisionmaking," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 18, no. 1, pp. 183–190, 1988. [Online]. Available: https://doi.org/10.1109/21.87068

[13] P. Bosc and L. Lietard, "Monotonic quantified statements and fuzzy integrals," in *NAFIPS/IFIS/NASA'94. Proceedings of the First International Joint Conference of The North American Fuzzy Information Processing Society Biannual Conference. The Industrial Fuzzy Control and Intellige*. IEEE, 1994, pp. 8–12. [Online]. Available: https://doi.org/10.1109/IJCF.1994.375159

[14] M. Delgado, D. Sánchez, and M. A. V. Miranda, "Fuzzy cardinality based evaluation of quantified sentences," *Int. J. Approx. Reasoning*, vol. 23, no. 1, pp. 23–66, 2000. [Online]. Available: https://doi.org/10.1016/S0888-613X(99)00031-6

[15] M. D. Calvo-Flores, D. Sánchez, and M. A. Vila, "Un método para la evaluación de sentencias con cuantificadores lingüísticos," in *Actas del VIII Congreso Español sobre Tecnologías y Lógica Fuzzy: Pamplona, 8-10 de septiembre de 1998*. Departamento de Automática y Computación, 1998, pp. 193–198.

[16] M.-A. Vila, J.-C. Cubero, J.-M. Medina, and O. Pons, "Using OWA operator in flexible query processing," in *The ordered weighted averaging operators*. Springer, 1997, pp. 258–274. [Online]. Available: https://doi.org/10.1007/9781461561231_20

[17] "MeteoGalicia," 2019, [Accessed December 2019]. [Online]. Available: www.meteogalicia.gal

[18] S. Barro, A. Bugarín, P. Cariñena, and F. Díaz-Hermida, "A framework for fuzzy quantification models analysis," *IEEE Trans. Fuzzy Systems*, vol. 11, no. 1, pp. 89–99, 2003. [Online]. Available: https://doi.org/10.1109/TFUZZ.2002.806319

[19] M. Delgado, M. D. Ruiz, D. Sánchez, and M. A. Vila, "Fuzzy quantification: a state of the art," *Fuzzy Sets and Systems*, vol. 242, pp. 1–30, 2014. [Online]. Available: https://doi.org/10.1016/j.fss.2013.10.012

[20] F. Díaz-Hermida, D. E. Losada, A. Bugarín, and S. Barro, "A probabilistic quantifier fuzzification mechanism: The model and its evaluation for information retrieval," *IEEE Trans. Fuzzy Systems*, vol. 13, no. 5, pp. 688–700, 2005. [Online]. Available: https://doi.org/10.1109/TFUZZ.2005.856557

[21] F. Díaz-Hermida, M. Pereira-Fariña, J. C. Vidal, and A. Ramos-Soto, "Characterizing Quantifier Fuzzification Mechanisms: A behavioral guide for applications," *Fuzzy Sets and Systems*, vol. 345, pp. 1–23, 2018. [Online]. Available: https://doi.org/10.1016/j.fss.2017.07.017