Method for Measurement of Uncertainty Applied to the Formation of Interval Type-2 Fuzzy Sets

Mauricio A. Sanchez, Oscar Castillo and Juan R. Castro

Abstract This paper proposes a new method for directly discovering the uncertainty from a sample of discrete data, which is then used in the formation of an Interval Type-2 Fuzzy Inference System. A Coefficient of Variation is used to measure the uncertainty on a finite sample of discrete data. Based on the maximum possible coverage area of the Footprint of Uncertainty of Gaussian membership functions, with uncertainty on the standard deviation, which then are modified according to the found index values, obtaining all antecedents in the process. Afterwards, the Cuckoo Search algorithm is used to optimize the Interval Sugeno consequents of the Fuzzy Inference System. Some sample datasets are used to measure the output interval coverage.

1 Introduction

Uncertainty, as it is currently perceived, is still something of a mistified topic. Being defined as something that is doubtful or unknown, in which by nature cannot be directly measured, therefore showing a first problem in making use of it. Although by nature, uncertainty is an unknown, it has not stopped engineers, scientists, mathematicians, etc. from using it. That is, although directly not known, an approximate of it can be modeled and used, improving the models in which it is used. By using uncertainty in a model, that model will improve its resilience, thus obtaining a better model in the end.

M.A. Sanchez · J.R. Castro
Autonomous University of Baja California, Tijuana, Mexico
e-mail: mauricio.sanchez@uabc.edu.mx

J.R. Castro
e-mail: jrcastror@uabc.edu.mx

O. Castillo (✉)
Tijuana Institute of Technology, Tijuana, Mexico
e-mail: ocastillo@hafsamx.org

P. Melin et al. (eds.), *Design of Intelligent Systems Based on Fuzzy Logic, Neural Networks and Nature-Inspired Optimization*,
Studies in Computational Intelligence 601, DOI 10.1007/978-3-319-17747-2_2

Most current literature on uncertainty [1–6] is mainly based on having previous knowledge of the confidence interval around certain measurements, which translates into what is the probable uncertainty which exists within certain measurements, usually expressed with the plus-minus symbol ± (e.g. 10.4 ± 0.02, with in interval representation of [10.38, 10.42]).

As for models with uncertainty, there exists a logic which directly manages uncertainty, this being Interval Type-2 Fuzzy Logic (IT2 FL) [7], which infers Interval Type-2 Fuzzy Sets (IT2 FS) and ultimately obtains an interval or a crisp value [8]. IT2 FS manage uncertainty directly into its logic by means of confidence intervals [9], the best solution could be anywhere within such interval, and as such is an excellent tool for directly applying and inference when dealing with uncertainty. And as stated, the output interval can be used as the end result and a deffuzification process can be computed upon such interval in the case that a crisp value is required, and not an interval.

In this paper, a link is proposed between a measure of dispersion and uncertainty, which is ultimately used in the formation of IT2 FS. The platform for the model is created by a Fuzzy C-Means algorithm [10], afterwards using the Coefficient of Variation is used to calculate the Fingerprint Of Uncertainty (FOU) of each individual IT2 FS in the antecedents of the Interval Type-2 Fuzzy Inference System (IT2 FIS), and finally, a Cuckoo Search algorithm [11] is used to optimize Interval Type-2 Sugeno linear consequents [12]. The proposed method can be categorized as a hybrid algorithm because it requires multiple steps/algorithm to work in sequence for the final result to be obtained.

This paper is divided into three sections, the first is a brief introduction to the definition of Interval Type-2 Fuzzy Sets; the following section describes in detail both the premises and the proposed method; afterwards, some experimental results are shown and discussed which asses the viability of the proposed method; finally, concluding remarks are given as well as a couple of open questions as future work.

2 Interval Type-2 Fuzzy Sets

With the introduction of Fuzzy Sets in 1965 [13], it improved upon formal hard logic, where instead of only having two choices of truth values {0, 1}, any value between [0, 1] was now possible. This set an unprecedented involvement in research that up to today is still very strong, first came Type-1 Fuzzy Sets [14], which can only represent vagueness or imprecision, later came Interval Type-2 Fuzzy Sets, which could now, apart from vagueness, also represent a degree of uncertainty (which is the focus of the proposed method in this paper), although recently General Type-2 Fuzzy Sets [15] are starting to gain traction in research, is still far from maturity when compared to Type-1 or Interval Type-2 Fuzzy Sets.

By nature, IT2 FS directly integrate uncertainty into its reasoning. This behavior is best applied in the case of when it is expected to deal with uncertainty in the

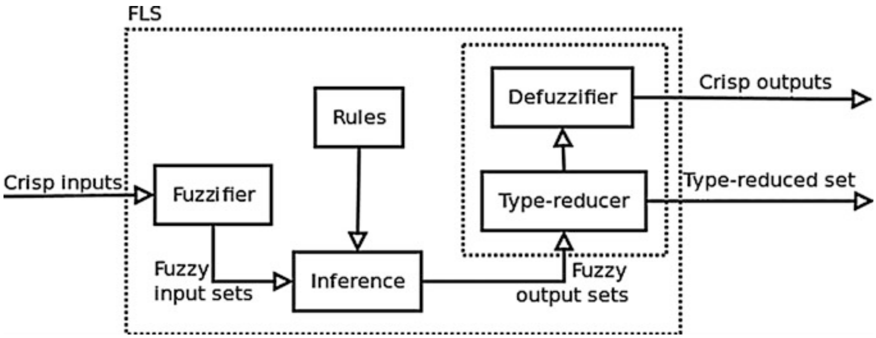


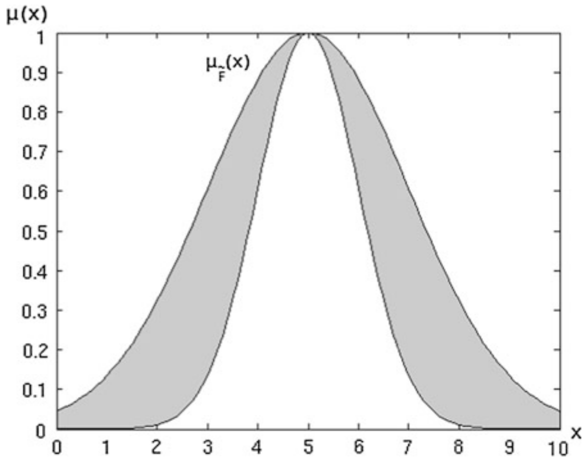
Fig. 1 Block diagram describing an IT2 FLS. With a cripst input, two outputs are possible, a confidence interval in which any possible point within such interval is a correct answer, or a crisp value, in the case a single real number is required as output

system that it is modeling, or when certain confidence intervals (uncertainty) are known a priori to designing the IT2 FIS.

The most general descriptive form of an IT2 FIS is through a block diagram, as shown in Fig. 1, which describes the basic inner functions of the complete inference. The Fuzzifier block may or may not transform the crisp input into a FS, this is chosen depending on the intended behavior of the system; the Inference block takes from the Rules block and reasons upon each input´s compatibility; the Type-reducer block processes the outputs into an interval; finally, the Defuzzifier block reduces the interval from the previous block and obtains a single real number.

An IT2 FS $\tilde{A}$ is represented by $\underline{\mu}_{\tilde{A}}(x)$ and $\overline{\mu}_{\tilde{A}}(x)$ which are the lower and upper membership functions respectively of $\mu_{\tilde{A}}(x)$, and is expressed as $\tilde{A} = \int_{w^l \in [\underline{\mu}_{\tilde{F}^l_k}(x_k), \overline{\mu}_{\tilde{F}^l_k}(x_k)]} 1/w^l$. Where, $x \in X$, k is the kth antecedent, and l the lth rule. A sample IT2 FS is shown in Fig. 2, here a Gaussian membership function with uncertainty in the standard deviation.

Fig. 2 Sample IT2 FS membership function. A Gaussian membership function is shown which has uncertainty through the standard deviation



The representation for rules in an IT2 FIS is formatted as shown in Eq. (1), where, $l = 1, \ldots, M$ rules, $p = 1, \ldots, q$ inputs, $\tilde{F}$ is an antecedent IT2 FS, and $\tilde{G}$ a consequent IT2 FS.

$$R^l : \text{IF}\, x_1 \text{ is } \tilde{F}_1^l \text{ and } \ldots x_p \text{ is } \tilde{F}_p^l, \text{ THEN } y \text{ is } \tilde{G}^l \tag{1}$$

3 Proposed Method for Measuring Uncertainty

Before giving a detailed description of the proposed method, the input data must first be defined. As a starting point, a dataset is required, these data pairs, defined as Eq. (2), where φ is a set of ordered input values, and γ is a set of ordered output values, such that Γ forms a tuple of ordered sets of inputs with their respective outputs.

$$\Gamma = \langle \varphi, \gamma \rangle \tag{2}$$

Having a dataset Γ, first some pre-processing must be done in order to obtain the required inputs to the proposed method, this process is executed in order to acquire a description of the IT2 FIS, that is, to obtain the rule description ω as well as each membership function's base description, and the set of data pairs which affected the formation of each membership function $\gamma \in \Gamma$. As this is are the required inputs $\{\omega, \gamma\}$ to the proposed method, a Fuzzy C-Means (FCM) algorithm was chosen to process the raw dataset Γ into the listed required inputs $\{\omega, \gamma\}$. The FCM provides a description of rules by means of a center for each membership function for each rule. Although the FCM can define consequents for the rules in a Fuzzy System, only the antecedents are used. As the other required input is a set of data pair sets which affected the definition of each center, this can be obtained from the partition matrix that is given by the FCM; for each data pair there exists a membership value $[0,\ 1]$ which defines how much a certain data pair belongs to a cluster, or rule of the found FIS, to simplify building the sets of data pairs, a simple competitive rule is used: *the cluster with the highest value decides that said data pair belongs to its formation set*. With both required inputs obtained, the proposed method can now begin.

3.1 Dispersion in Data

Data dispersion in a sample of data pairs can be interpreted as a case of uncertainty. An example of varying degrees of dispersion is shown in Fig. 3, where low, medium, and high data dispersion, in relation to its center can, be perceived.

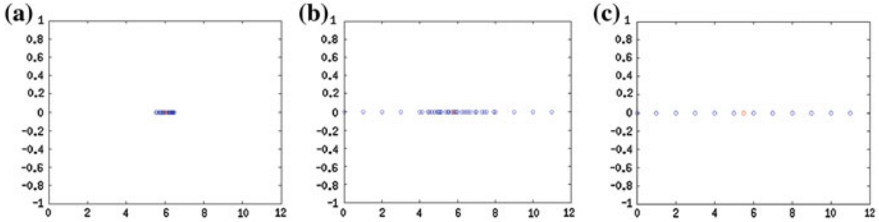


Fig. 3 Example of data dispersion. **a** Low data dispersion, **b** medium data dispersion, and **c** high data dispersion

When there is low dispersion of data samples near it representative center point, most data points are bound by only a small distances as the standard deviation is very small. This being interpreted directly into uncertainty in data, low dispersion is low uncertainty because its numerical evidence concludes that there is near zero possibility that further singular samples will fall far from the central point, unto which all previous numerical evidence is very close to. In the case of medium data dispersion, although there is a concentration of numerical evidence near its central point, there are still data points farther from its center, this leads to knowing that although future reading might obtain evidence which is far from the center, the probabilities of this occurring is low when compared to having future readings fall near the center, although not as near in the case of lower dispersion, this behavior points to having a medium amount of uncertainty. On the extreme case of high dispersion, where every sampled data point is evenly distributed throughout the range, the available numerical evidence gives way to conclude that any future sample may equally land on any section, therefore a high amount of uncertainty exists.

3.2 Relation Coefficient of Variation with Uncertainty

For the purpose of converting dispersion into uncertainty, a measure is first required which can identify a degree of dispersion in a given set, preferably a normalized value, and as such requirement, the Coefficient of Variation c_v, shown in Eq. (3), where σ is the standard deviation, and μ is the mean of the set.

$$c_v = \frac{\sigma}{\mu} \quad (3)$$

This coefficient has some limitations which can be avoided by applying some modifications. First, c_v should only be computed on non-negative values, for the case of existing negative values, the solution is to remap all values unto the positive side of the axis. Second, if μ has a value of 0 (zero), this would case an error in computation, the solution is to add ε, which is a very small value, assuring a non-division by zero. Another note on the behavior of c_v, is that in normal distributions,

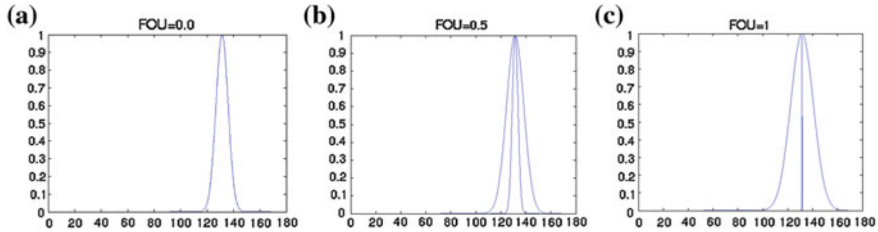


Fig. 4 Examples of varying degrees of FOU. Where **a** FOU = 0, **b** FOU = 0.5, and **c** FOU = 1

values of [0, 1] are most likely to be obtained, but non-normal distributions can obtain values above 1. Fortunately, with the FCM, all calculated sets are normal distributions, so this is a non-issue with the current implementation.

With the known limitations of c_v, an equation which modifies a set D is proposed which addresses the issue negative values, shown in Eq. (4), where if a value exists that is negative then the absolute value of the minimum is added to the set, thus remapping all values into the domain of positive values.

$$\text{IF}\,(\exists x \in D), \{x|x < 0\}\,\text{THEN}\; D = D + |\min(D)| \tag{4}$$

In addition, a modification of Eq. (3) to suppress a possible division by zero, as shown in Eq. (5), where ε is a very small value.

$$c_v = \frac{\sigma}{\mu + \varepsilon} \tag{5}$$

To express a relation dispersion-uncertainty, when dealing with IT2 FS, the Footprint of Uncertainty (FOU) is used. This relation is a direct proportion $FOU \propto c_v$. When there is low dispersion, there is a small FOU, when there is a medium amount of dispersion, there is a medium amount of FOU, and when there exists a high amount of dispersion, there is a high amount of FOU. This is better expressed in Fig. 4, where varying degrees of a measure of dispersion has been converted into a FOU which directly forms an IT2 FS, explained in the following sub-section.

3.3 Proposed Method

To form IT2 FS for the antecedents of a FIS, the first step is to obtain rule configuration, and data pair sets for each inputs on each rule, via a FCM algorithm. Afterwards each set of data pairs is worked on independently of each other. First, a standard deviation σ is found for the set in relation to its μ, which was found by the FCM, then the c_v is calculated. This value is now used to search for the optimal FOU area in an IT2 FS. Considering Fig. 4c, this would be the highest possible area. The initial search is done by first considering the highest possible area and the σ which was already calculated,

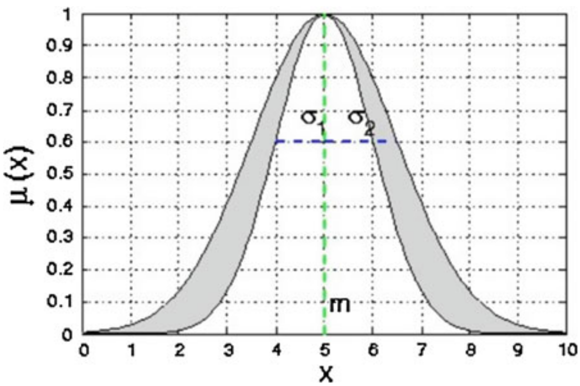


Fig. 5 IT2 FS represented by a Gaussian membership function with uncertainty in the standard deviation

with discrete small steps a search is performed for the FOU value which equals c_v. The smallest value is set as $\sigma_1 = \sigma_2$, shown in Fig. 4a. Each increment step λ affects σ as shown in Eq. (6), this is done iteratively while $\sigma_i \leq \|\mu, \sigma_0\|$.

$$\sigma \pm \lambda \tag{6}$$

Once the search has found the values of σ_1 and σ_2 which represent the desired FOU, the IT2 FS can be formed. Which has the form of Fig. 5, this can be formed with the values which have been calculated, by the FCM, μ, and by the proposed method, σ_1 and σ_2. This concludes the proposed method for building the antecedents of an IT2 FIS.

3.4 IT2 Sugeno Fuzzy Consequents

The proposed method only obtains the IT2 FS for the antecedents of a FIS, the next required step is to obtain the consequents of the FIS. This is done by optimizing the IT2 Sugeno linear parameters via a Cuckoo Search algorithm. Although any other optimization algorithm can be used.

4 Experimental Results and Discussion

To test the proposed method, various datasets were used. The validation method was to verify that the interval output of the IT2 FIS had good coverage of the reference targets and at the same time not overreaching too far with the output interval.

Among the used datasets, three were used. A synthetic dataset of a 5th Order curve [16], with 1 input (x) and 1 output (y), and 94 total samples. And two real datasets; engine behavior [16], with 2 inputs (fuel rate, speed) and 2 outputs (torque, nitrous oxide emissions), and 1199 total samples; and Hahn1 [16], with 1 input (temperature) and 1 output (thermex), with 236 total samples.

4.1 Experimental Results

The obtained IT2 FIS for each dataset is shown in Figs. 6, 7 and 8, making emphasis on the FOU of the individual membership functions in the antecedents, where varying degrees of uncertainty can be seen.

As for the output coverage for each dataset, using 40 % training and 60 % training, Table 1 show the summary of the obtained coverage results.

The last set of results show graphical representations of the respective outputs for each dataset, these are shown in Figs. 9, 10 and 11. Where the blue points represent the output targets, and the lower and green lines represent the coverage of the FOU (Fig. 12).

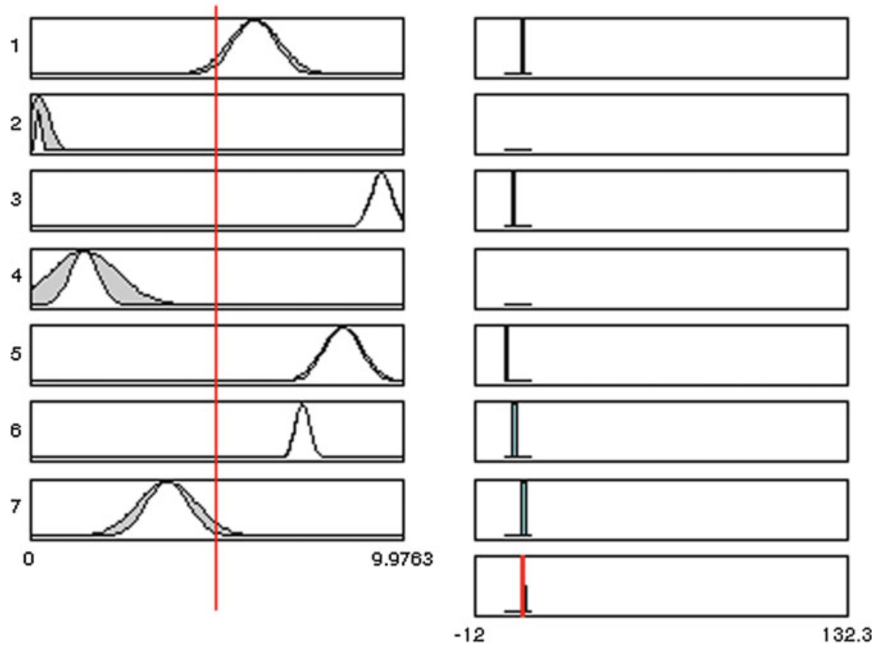


Fig. 6 IT2 FIS for solving the 5th order curve dataset

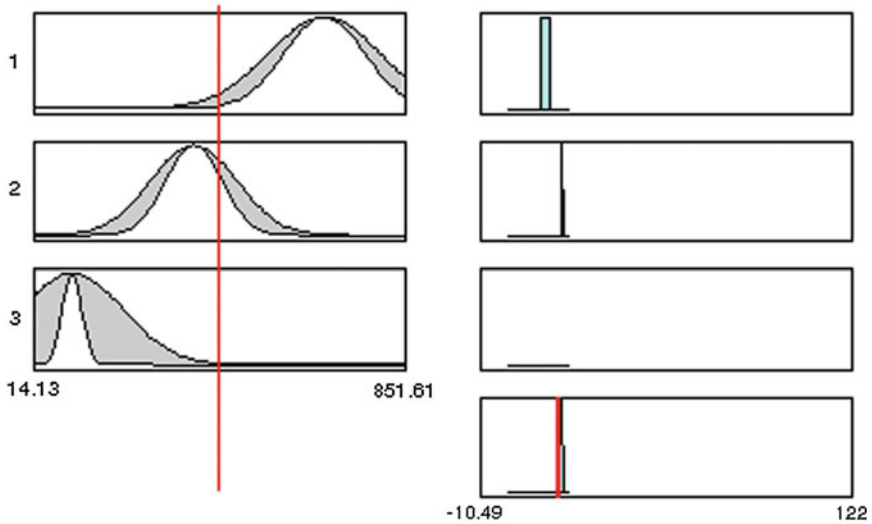


Fig. 7 IT2 FIS for solving the Hahn1 dataset

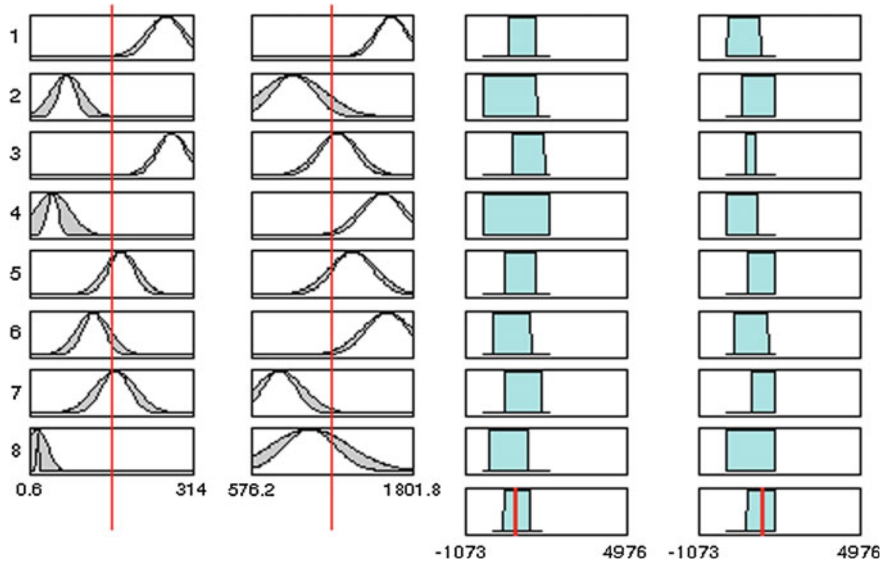


Fig. 8 IT2 FIS for solving the engine behavior dataset

Table 1 Obtained output coverage results for the chosen datasets

Dataset name	Coverage (%)
5th Order curve	100
Hahn1	100
Engine behavior	99.88/99.66

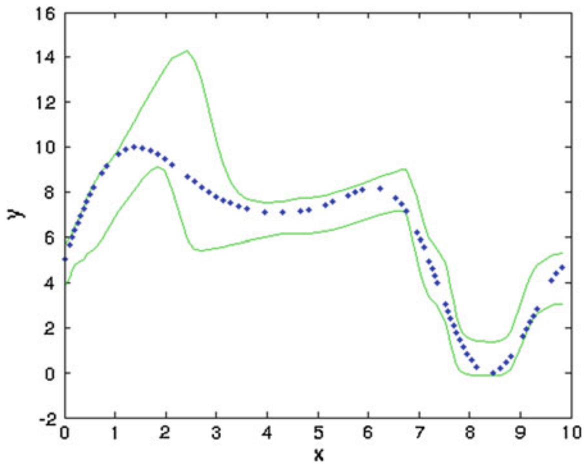


Fig. 9 Output coverage for the 5th order dataset

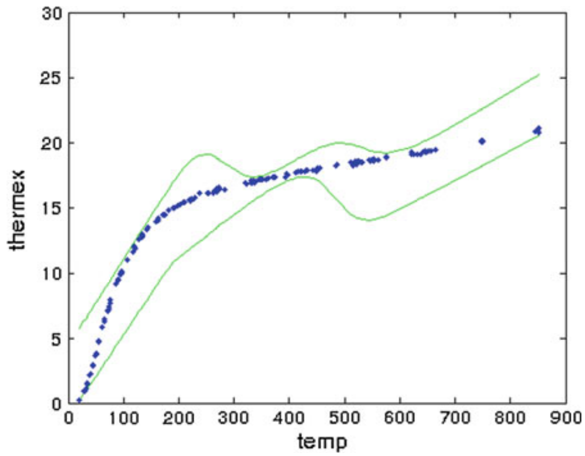


Fig. 10 Output coverage for the Hahn1 dataset

4.2 Results Discussion

With the obtained results, two facets of discussion arise, on the individual level and on the general level. On the individual level; for the 5th Order curve, a full coverage of the target is achieved although there are spikes where the curve changes slope, this is caused by the linear consequents which cannot follow abrupt changes in the

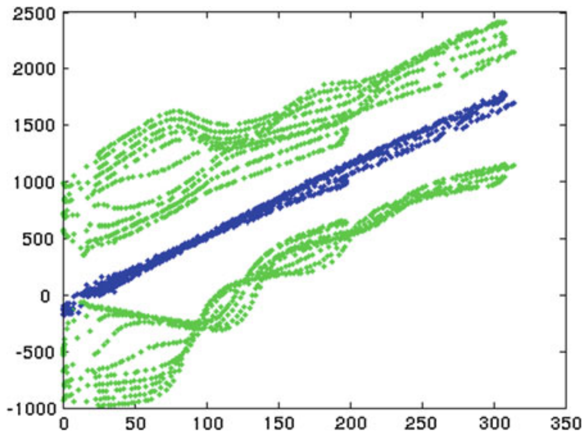


Fig. 11 Output coverage for the Hahn1 dataset. For the first output of the FIS

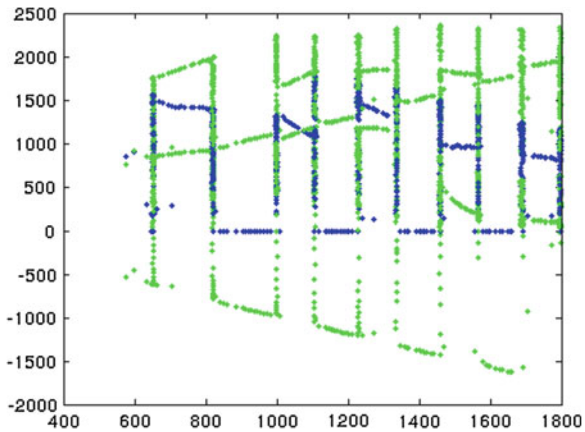


Fig. 12 Output coverage for the Hahn1 dataset. For the second output of the FIS

curve grade due to the small amount of rules used for this FIS. A solution would be to use more rules to compensate, but this would also cause an additional, and unnecessary, complexity in the system. Yet the overall behavior is acceptable as there is sufficient coverage as well as a controlled width of the output uncertainty. For the Hahn1 solution, it has the same curve behavior of the 5th Order curve, where with only three rules there is a pronounced visible behavior in the linear output of the consequents. Yet there is good coverage, of 100 %, of the target outputs via the controlled output uncertainty. Finally, for the engine, having two outputs, each ones behavior was slightly different. The first output has a more predictable behavior by better following the output targets with its coverage of 99.88 % of reference targets, whereas the second output´s behavior is not as linear, such that it holds a tendency to expand as the *x* axis increases, although it has a coverage of 99.66 % of its reference targets. It must be noted that this specific behavior is more in line with how the Cuckoo Search algorithm optimized the consequents, because the spreads on each individual consequent control the output interval behavior. The solution would be to adjust the Cuckoo Search for better performance or use another optimization algorithm that obtains a better solution.

On the general level of the obtained results, the formed antecedents give a good representation of uncertainty based on the dispersion of the individual sets which affected the creation of the rule configurations found by the FCM. It also depicts a behavior that IT2 FS are not always necessary, with low to no dispersion, and a T1 FS would be more than enough.

Being dependent on other algorithms can limit the general performance of the proposed method. Yet it also adds more possibilities, such as interchanging clustering algorithms to one that can obtain better rule configurations and belonging sets to be used by the proposed method. As for the optimization of the IT2 Sugeno linear consequents, there is a vast amount of optimization algorithms which could also be used for acquiring better results and thus improving the output interval behavior.

5 Conclusion and Future Work

5.1 Conclusions

With the suggested relation dispersion-uncertainty, direct uncertainty extraction is possible from existing data. This relation is found through the Coefficient of Variation, an existing equation used to measure the amount of dispersion in a set, this measure is a normalized value between 0 and 1, that although higher values than 1 are possible, this is only for non-normal distributions, which, for the purposed application, are non existent considering that the sets are created by a clustering algorithm which only groups in normal distributions of data.

The application shown in this paper, of forming IT2 FS through the suggested equation, which relates dispersion-uncertainty, finds this relation based on the maximum possible achievable FOU, valued at 1, and relates to the maximum possible Coefficient of Variation, in a normal distribution, valued also at 1. This relation of dispersion-uncertainty-FOU is the main contribution of this paper.

With a deeper examination of the experimental results, there is much dependence on the FCM algorithm, where if such algorithm fails to provide a good model, the proposed method would fail also, since the proposed method depends on the performance of the clustering algorithm. Fortunately, if the FCM fails, other clustering algorithms could be used.

5.2 Future Work

Considering the limitation, as well as dependence, of the clustering algorithm, which other clustering or non-clustering techniques could be used to create a better pairing with the proposed method?

With the other high dependence on optimization algorithms for the consequent section of the IT2 FIS, what other optimization algorithm could be used to best pair with the proposed method?

In this paper an IT2 FS was formed, represented by a Gaussian membership function with uncertainty in the standard deviation. How would other IT2 FS membership function be adapted to use the proposed method?

How the area was directly correlated to the FOU by means of its maximum possible area was proposed. Is this the best approach?

Acknowledgments We thank the MyDCI program of the Division of Graduate Studies and Research, UABC, and Tijuana Institute of Technology the financial support provided by our sponsor CONACYT contract grant number: 314,258.

References

1. Yu, X., Mehrotra, S.: Capturing uncertainty in spatial queries over imprecise data. In: Database and Expert Systems Applications, pp. 192–201. Springer, Berlin (2003)
2. Klir, G.J.: Uncertainty and Information: Foundations of Generalized Information Theory. Wiley, New York (2005)
3. Weise, K., Woger, W.: A Bayesian theory of measurement uncertainty. Meas. Sci. Technol. **3** (1992)
4. Jurado, K., Ludvigson, S.C., Ng, S.: Measuring uncertainty. Am. Econ. Rev. (AEA) **105**(3), 1177–1216 (2013)
5. Chen, G., Ying, M., Liu, Y.: Dealing with uncertainty and fuzziness in intelligent systems. Int. J. Intell. Syst. **24**, 223–225 (2009)
6. Klir, G.J., Wierman, M.J.: Uncertainty-Based Information. Springer, Heidelberg (1999)
7. Mendel, J.M., John, R.I., Liu, F.: Interval type-2 fuzzy logic systems made simple. IEEE Trans. Fuzzy Syst. **14**, 808–821 (2006)
8. Castillo, O., Melin, P.: Recent Advances in Interval Type-2 Fuzzy Systems. Springer, Berlin (2012)
9. Mo, H., Wang, F.-Y., Zhou, M., Li, R., Xiao, Z.: Footprint of uncertainty for type-2 fuzzy sets. Inf. Sci. (Ny) **272**, 96–110 (2014)
10. Bezdek, J.C., Ehrlich, R., Full, W.: FCM: the fuzzy c-means clustering algorithm. Comput. Geosci. **10**, 191–203 (1984)
11. Yang, X.-S.: Cuckoo search via lévy flights. In: World Congress on Nature and Biologically Inspired Computing (NaBIC) 2009, pp. 210–214. IEEE (2009)
12. Ying, H.: Interval type-2 Takagi-Sugeno fuzzy systems with linear rule consequent are universal approximators. In: NAFIPS 2009—2009 Annual Meeting of the North American Fuzzy Information Processing Society, pp. 1–5. IEEE (2009)
13. Zadeh, L.A.: Fuzzy Sets. Inf. Control **8**, 338–353 (1965)
14. Melin, P., Castillo, O.: Fuzzy modeling fundamentals. Wiley Encycl. Comput. Sci. Eng. (2007)
15. Mendel, J.: General type-2 fuzzy logic systems made simple: a tutorial. IEEE Trans. Fuzzy Syst. **22**, 1 (2013)
16. The MathWorks, Inc.: Natick. MATLAB Release, Massachusetts, U.S. 2013b, (2013)

http://www.springer.com/978-3-319-17746-5

Design of Intelligent Systems Based on Fuzzy Logic, Neural Networks and Nature-Inspired Optimization
Melin, P.; Castillo, O.; Kacprzyk, J. (Eds.)
2015, XII, 637 p. 307 illus., 226 illus. in color., Hardcover
ISBN: 978-3-319-17746-5