CEUR-WS.org/Vol-3834/paper62.pdf

Addressing Uncertainty according to the Annotator's Expertise in Archaeological Data Collections: an Approach from Fuzzy Logic

Patricia Martin-Rodilla[1,*], Leticia Tobalina-Pulido[2]

[1] Padre Sarmiento Institute of Galician Studies, Spanish National Research Council (CSIC), (Spain)
[2] Departamento de Ciencias Históricas, Universidad de Cantabria (Spain)

Abstract

Archaeological data allow us to synthetically represent the past of individuals and communities over time. This complex representation task requires an amalgamation of variables and makes the intrinsic data vagueness. The study of vagueness as an archaeological data dimension has become a dynamic focus of archaeologists' work in recent years, presenting theoretical and practical approaches for the representation, mainly with fuzzy logic, of archaeological variables. Vagueness in archaeological data can occur due to different reasons: non-existence of evidence, imprecision, errors, subjectivity, etc. Furthermore, the data is usually managed in groups, shared or recovered for subsequent investigations, so the vagueness traceability that is injected due to these management phases is lost.

In this paper we present the ongoing work carried out in modeling under fuzzy formal theory the explicit representation of the expertise of the annotator (understood as the professional who introduces archaeological data into a certain system, giving value to the defined variables) in a decoupled way from the value attributed to each variable. The first experiments with chronological and use variables of the sites show how making the annotator's expertise explicit in the fuzzy model allows maintaining the traceability of the uncertainty injected into the archaeological data due to the definition and management of the datasets by different people, as well as establishes a base for implementing archaeological fuzzy decision-based systems.

Keywords

vagueness, annotator expertise, archaeological data, fuzzy logic, Mamdani

1. Introduction

The large and heterogeneous volume of current archaeological data is often collected over long periods (weeks, months, and even years) and by different people and at various times (inheriting legacy data) and places. These aspects make difficult to perform compilation studies with a coherent data analysis. Since already known vagueness aspects of the archaeological data such as imprecision (for example in dating systems, where the chronologies assigned to a stratum, a site phase, or a material are sometimes imprecise [12] or present some errors), or the vagueness in the categorical variables (i.e. the case of the interpretation of the uses attributed to a site or a phase, since we do not always have data sufficient to assign a precise use value), there is also

CHR 2024: Computational Humanities Research Conference, December 4–6, 2024, Aarhus, Denmark
*Corresponding author.
p.m.rodilla@iegps.csic.es (P. Martin-Rodilla); leticia.tobalina@unican.es (L. Tobalina-Pulido)
0000-0002-1540-883X (P. Martin-Rodilla); 0000-0002-3315-5506 (L. Tobalina-Pulido)
© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

the vagueness injected into data depending on the person who gives it value, who may have different levels of training and expertise. This ”archaeological annotator” can present several degrees of knowledge and security in its categorization depending on its previous background, even assigning values for different variables representing the same archaeological site or object. The high number of sites and the archaeological intervention’s data available, mainly since the 1980s, only increased the problem. All these data appear today as ”legacy data”, not only in terms of the technological format and the different definitions inherited but also because we do not have any information about the knowledge of the annotators that produce them and the degree of security in the values assigned that they presented.

Based on previous works to address the problem of vagueness in archaeological data through fuzzy logic ([29], [7], [31], [25]), this paper presents the work in progress and its first validation and results to address the specific problem of including the expertise of the annotator as a fuzzy variable. The model presented allows expressing in a decoupled manner the value that the annotator assigns within the fuzzy membership function regarding a certain use or chronology to an archaeological element and its knowledge and security about the assigned value, facilitating a more complete analysis of the vagueness situation of a specific archaeological site or object and to implement measurement and inference/decision-making systems based on it. The paper also implements an illustrative inference fuzzy system with a definable threshold to decide whether certain archaeological sites or objects meet certain characteristics of vagueness in terms of the expertise of the annotator that may be proposed for an expert review.

The paper is structured as follows: section 2 reviews the existing works of fuzzy logic applications in archaeology. As far as we know, the expertise of the archaeological annotator has not been formally discussed in existing models. Section 3 briefly details the Mamdani formal theory of fuzzy logic and justifies its choice. Furthermore, the theoretical bases of our model are presented. In section 4, the proposed model is applied to a list of real archaeological sites (through scripts in the Math Works software publicly available at [17]), showing the results of the fuzzy inference system: a list of archaeological sites proposed for expert review due to their vagueness. Section 5 discusses the limitations of the model and the conclusions and future work of the proposal.

2. Previous works

In the Humanities, there are limited reflections on the vagueness issue [30]. It is worth mentioning the project ”Faire Science avec l’Incertitude” started in 2013 at the Université Nice Sophia Antipolis, which aimed to carry out an analysis and reflection on uncertainty from the epistemological and methodological view [6]. This project was a starting point proposal for a more mature reflection on the different types of vagueness and how to integrate them into scientific production [6].

Regarding historical studies, J. Owens [20] indicates that historians often face the problem of fitting their data into specific “boxes.” Historical periods, for example, are ambiguous and diffuse, and we cannot always frame an event in a precise chronology. Thus, fuzzy logic has been the most used approach since the 1980s for vagueness data treatment. In this sense, we must mention the DICTOMAGRED project [16], which uses ConML [11] as a language to model information, including markers managed through fuzzy logic to express uncertainty [18]. For the modern era, the works of J.-B. Owens [21] use fuzzy logic to measure some inaccuracy of historical data and adapt to a computer-wise method. In Geography, most of the work is focused on surfaces and diffuse borders modeling, using "geographical fuzzy models" for calculating the degree of membership of each point to the geometry (lines, polygons...). [28]. However, in all these works there is no annotator's expertise in vagueness treatment.

In archaeology, over the last decades, the management of data vagueness has become increasingly important in research; Some authors have tried to apply it practically in their projects ([25], [26], [7], [32]), and others have reflected on a more theoretical level [33]. In general, when addressing this problem, there are three approaches used: Bayesian statistical procedures, the aoristic method [23], and fuzzy logic. The first is the method that is used recurrently for the calibration of 14C dating [23]. The second allows temporal uncertainties to be quantified and incorporated into subsequent analysis [12]; It has been applied in archaeology mainly as a means of statistically representing the temporal in-definition of archaeological artifacts ([4], [34], [3]). Regarding fuzzy logic, it allows working with non-Boolean labels, something necessary in humanistic disciplines when working with qualitative data with a strong linguistic character; It has been the most used both theoretically and practically in research [10] or applied to materials, both lithic and ceramic ([2], [31]) and bone (for the determination of sex in human bones, for example: [29]). Also the analysis of road networks, where one of the best examples of application is the work of C. de Runz and E. Desjardin ([27]), who focus on the characterization of the imperfection of data on Roman roads in their SIGRem project; or to the study of the settlement [7].

Although all these attempts allow the archaeologist to express the vagueness contained in archaeological data for explicit reasons (such as known errors or inaccuracies in instrumental measurements, especially in spatial data) however as far as we know, the vagueness due to the degree of annotator's expertise has not been formally addressed. As previously stated, the work of data production in archaeology over long periods of time, with multidisciplinary teams and with a large volume of inherited data motivates us to propose a model that includes this data dimension and illustrates their use in a fuzzy inference system as a validation.

3. A fuzzy logic model for addressing annotator's expertise in archaeological data collections

3.1. Fuzzy logic strengths

Fuzzy logic is intended to model logical reasoning with vague or imprecise statements [22], emerging inside the theory of fuzzy sets introduced by Lotfi Zadeh [35]. A fuzzy set assigns a degree of membership, typically a real number from the interval [0,1], to elements of a universe [35]. A fuzzy set also allows the extension of this approach, allowing the possibility of partially belonging to the same element to several sets considering its degree of membership in a group and allowing human knowledge to be represented quantitatively through the theory of fuzzy sets. For the case of fuzzy sets defined on continuous numerical domains, there are different membership function shapes, being more appropriate depending on the fuzzy logic problem that we face.

From a practical point of view, especially due to their balance between simplicity and human expressive capacity, we decide to use Mamdani fuzzy theory in this work. Mamdani's fuzzy theory is considered the most intuitive, rule-based, and well-suited to human input fuzzy model [13]. Since Mamdani systems have more intuitive and easier-to-understand rule bases, they are well-suited to expert system applications where the rules are created from human expert knowledge, such as medical diagnostics, decision support systems, or, in our case, archaeological data management. It was proposed by Ebrahim Mamdani in 1975 [15, 14]. Table 1 presents Mandani's five steps:

Table 1
Mandani's fuzzy system steps

Step	Description
1. Fuzzification	Convert the crisp inputs (precise numerical values) into fuzzy sets. This is achieved by membership functions that map the input values to degrees of membership between 0 and 1.
2. Fuzzy Rule Base	If-then rules formulated linguistically, as 1) an antecedent (premise) i.e. the "if" conditions, and 2) a consequent (conclusion) i.e. the actions "then".
3. Fuzzy Inference	The system evaluates the fuzzy rules. For each rule, the degree of compliance is determined based on the membership functions of the inputs. Then, the logical operator (usually "AND" or "OR") is applied to combine the degrees of compliance of the antecedents.
4. Aggregation	The consequents of the activated rules are combined to form an output fuzzy set. This fuzzy set represents the combined action of all relevant rules.
5. Defuzzification	Convert the output fuzzy set into a crisp (precise) value for the system output. There are several defuzzification methods, the most common being the "centroid" (center of mass), which calculates the center of gravity of the fuzzy set.

3.2. Capturing archaeological annotator's expertise

As we previously detailed, we can model several inputs in Mamdani systems. Thus, the membership value functions and the annotator expertise functions of the logic implementation in this work are both represented by classical Mamdani functions. The experimentation in terms of performance and diversity of final results in terms of other fuzzy models are part of our future work. Following Mandani's model, each of our archaeological elements (archaeological sites or objects) presents four input membership functions:

- **Use membership function**: we defined 6 possible uses (1.Basilica, 2.Roman Villa, 3.Fortification, 4.Habitat, 5.Farm, 6.Funeral Space).
- **Chronology membership function**: we define 6 possible chronological phases (1.Middle Ages, 2.Late Middle Ages, 3.Early Middle Ages, 4.Roman Imperial, 5.Roman High Imperial, 6.Late Antiquity).

- **Annotator expertise membership function regarding use**: the annotator defines their expertise for each value given in the use membership function. For example, if a site has a value of 0.2 for the Basilica and 0.7 for the Farm use, the annotator specifies their knowledge about these periods separately.
- **Annotator expertise membership function regarding chronology**: the annotator defines their expertise for each value given in the chronology membership function.

The definition of the membership functions' possible values is determined for the archaeological data of the next illustrative example, being able to expand and/or adapt according to the case. These inputs are fuzzificated. In our case, the resultant fuzzy inference system should produce one output function (the value for the revision proposal for each archaeological site or object). Note that there are 81 possible combinations in our system (4 vector inputs and linguistic labels (LOW, MEDIUM, and HIGH) for determining the membership function's values). To manage these 81 possible combinations, the table 2 illustrated 10 system's rules for our application:

Table 2
Rule-based system for our next illustrative archaeological example

Rule	Condition 1: use expertise	Condition 2: chronology expertise	Decision Value
Rule1 [0 0 1 1 2 1 1]	IF is Low	AND is Low,	THEN Review.
Rule2 [0 0 1 2 2 1 1]	IF is Low	AND is Medium,	THEN Review.
Rule3 [0 0 2 1 2 1 1]	IF is Medium	AND is Low,	THEN Review.
Rule4 [0 0 3 3 1 1 1]	IF is High	AND is High,	THEN No Review.
Rule5 [0 0 3 2 1 1 1]	IF is High	AND is Medium,	THEN No Review.
Rule6 [0 0 3 3 1 1 1]	IF is High	AND is High,	THEN No Review.
Rule7 [0 0 3 1 2 1 1]	IF is High	AND is Low,	THEN No Review.
Rule8 [0 0 1 3 2 1 1]	IF is Low	AND is High,	THEN No Review.
Rule9 [0 0 2 3 2 1 1]	IF is Medium	AND is High,	THEN No Review.
Rule0 [0 0 2 2 1 1 1]	IF is Medium	AND is Medium	THEN No Review.

4. Illustrative example: a Mamdani fuzzy inference system for expert revision proposal of the archaeological sites

4.1. List of archaeological sites

In this section, we detailed the intrinsic characteristics of the archaeological information of each of the 4 sites chosen to illustrate the fuzzy system behavior. Also, we present here the motivations for the different values in the membership function: use and chronology evidence and the annotator's expertise for both. We select 4 real cases with different characteristics to better illustrate the various possibilities.

1. **Fortunatus** archaeological site [24]: Villa Fortunatus site is located near Fraga in "Pilaret de Santa Quiteria", Aragón (Spain). It is a Roman villa whose construction dates back to the 2nd century, although was expanded and decorated with rich mosaic pavements at

the end of the 3rd century or the beginning of the 4th century BF. Archaeological research has shown that after the house was abandoned, a part was reused and transformed into a basilica for Christian worship at the end of the 4th century or the beginning of the 5th century. Signs of Visigoth occupation have also been found at the site. All this information about uses and chronology is extracted from the official archaeological catalog of the Government of Aragon (Spain). The data was entered by workers at the archaeology service, and the archaeologist attributed a medium level of expertise due to their archaeological generalist training.

2. **Guzmán** archaeological site: A medieval fortification site currently being excavated by the Universidad Rey Juan Carlos de Madrid (URJC) Archaeology and Digital Humanities Laboratory. There is not high evidence of the site's occupation as a habitat. The annotator defined 4 chronological phases. For two of them (2 and 3) they have absolute radiocarbon dating, for 1 and 4 they are based on the archaeological materials found. For the last three phases there is sufficient evidence to categorise it as a fortification. The annotator presents a good knowledge of the materiality of the area.
3. **El Mandalor** archaeological site: Site discovered and excavated by an archaeology private company during the works on the Camino de Santiago highway in Navarra, Spain [1]. We consider the data from the publication by an expert in the period, adding new possibilities of use and changes in the certainty of uses. The certainty data on chronologies provided by the excavation team are maintained. The site is possibly a villa with several phases of renovation, although there are indications that it could also be a farm, as there are no elements that could be linked to the aristocratic classes.
4. **Las Penas** archaeological site: Cave-like site, it is located near the urban center of Mortera, Cantabria (Spain), discovered by members of the G.E.I.S./C-R, in 2003 and excavated by Angeles Valle Gómez's team [9]. At this time, human remains appeared in the cavity jointed with objects of personal adornment, specifically the set of 5 Visigoth clasps related to Visigoth clothing (belt garnishes). Subsequently, José Ángel Hierro Gárate compiles information about the site, proposing a chronology of the late 7th century and 8th century for the belt clasps that appeared. Additionally, 4 samples were taken for radiocarbon dating with chronologies that coincide with the chrono-typological analyses of the metal pieces [8].

4.2. Building a fuzzy inference system for archaeological revision

The information about sites presented in the previous section allows us to define, for each of the 4 archaeological sites, their 4 membership functions by site. Due to space limitations, we detail the membership functions defined for one of the sites studied (El Mandalor). The rest of the membership functions are described in the Matlab script, available for query, execution, and edition [17].

- Use membership function: [0, 0.75, 0, 0.8, 0.5, 0]: 6 membership values assigned for the 6 use categories. In this site's case, the use cases are, with different degrees in certainty, habitat, Roman villa, and farm.

- Chronology membership function:[0, 0, 0, 0.75, 0.75, 0.75]: 6 membership values assigned for the 6 chronology categories. In this site's case, the chronological cases are, with different membership degrees in High Imperial Roman, Imperial Roman, and Late Antiquity.
- Annotator expertise membership function regarding use: [0, 1, 0, 1, 0.5, 0]: 6 annotators expertise values assigned for the 6 use categories. In this site's case, the annotator declares with different degrees, their expertise in terms of all use assigned values.
- Annotator expertise membership function regarding chronology: [0, 0, 0, 0,4, 0.4, 0.25]: 6 annotators expertise values assigned for the 6 chronological categories. In this site's case, the annotator declares with different degrees, their expertise in terms of all chronology-assigned values.

Once we have defined the four membership functions, the next step is to implement the fuzzy inference system that takes these membership functions into account to recommend whether or not to review each archaeological site. Thus, a script has been implemented in MathWorks Matlab software[19]. This public script [17], along with the results produced, contains all the logic natural language-based rules defined in the 3.2 section but codified as a rule-list matrix. This matrix implements the AND and OR operators as part of Mandani's step 3. Then, the consequences of the activated rules are combined to form an output fuzzy set. This fuzzy output set is defuzzified using the centroid Mamdani's method for offering a revision value for each archaeological site.

4.3. Results

The table 3 shows the numeric results obtained from the fuzzy inference system implemented. For each site, the system generates a centroid value to defuzzify the 4 fuzzy inputs following Mamdani's method. Once these centroid values are calculated, the system generates a decision value. This decision value is evaluated to recommend the archaeological site as Review or No Review.

Table 3
Output results form the fuzzy inference system. Note that columns 1-4 show centroid values for each input. Final columns show decision value and final decision obtained for each archaeological site

Site	Use	Chronology	Use Expertise	Chronology Expertise	Decision Value
1.Fortunatus	0.1000	0.8000	0.1000	0.2714	0.8400 REVIEW
2.Guzmán	0.4400	0.2700	0.4347	0.4000	0.5000 NO REVIEW
3.El Mandalor	0.5024	0.8000	0.4800	0.7714	0.1700 NO REVIEW
4. Las Penas	0.9200	0.4000	0.8000	0.5214	0.1700 NO REVIEW

More detailed results and the archaeological discussion are provided in Appendix A. System limitations, possibilities and future work are addressed in the next section.

5. Discussion and conclusions

As we detailed before, this is the first approach taken for addressing the annotator's expertise in fuzzy inference systems as a useful approach for archaeological data management. As the results show, the rule-based Mamdani's system allows to customize the rules in function not only based on the annotator's expertise but also on combinations of the 4 membership functions defined. This allows us to create fuzzy inference systems with our Matlab script, easily reusable. In addition, we often overlook our shortcomings when using archaeological data from other authors. Modeling "expertise" in annotations can indicate the certainty level of interpretations and researchers' confidence. Implementing this in archaeology departments or research centers would provide a useful filter for site information. While administrations often include a "location security" field in inventories, similar annotations for typology and dating are rare. Expert annotations would enable handling larger datasets with indicated certainty, enhancing data evaluation in GIS analyses and allowing researchers to assess results and data corpus.

Regarding complementary studies and future work, Mamdani's system is not the only fuzzy modeling alternative. In addition, the generalization when defuzzifying, using centroids, is useful for practical application but simplifies the capabilities of the representation system inputs. We have recently explored the integration in the fuzzy inference system of more than one annotator information, using vector distance theories between annotators. Similar approaches are explored for capturing the natural evolving of the expertise in one annotator. Also, the study should complete it with the evaluation of other fuzzy logic models, even considering type-2 fuzzy logic [5] to gain expressiveness and improve the performance of the fuzzy inference system.

Finally, some future work is needed regarding non-archaeological application of the approach, addressing the question if it is possible to generalize the model to near scientific areas and similar practices to the archaeological ones.

Acknowledgments

The work presented in this paper has been partially funded by the Spanish National Research Council (CSIC) through grant "Ayudas Extraordinarias a la incorporación de Científicos Titulares OPE 20-21: Estudio, generación, evaluación y análisis de modelos y tecnologías lingüísticas para la comprensión de la conexión pasado-presente desde fuentes textuales".

References

[1] M. R. Aguirre. "Arqueología en la Autovía del Camino". In: *Trabajos de Arqueología Navarra* 21 (2009), pp. 5–119.

[2] J. A. Barceló. "Heuristic classification and fuzzy sets. New tools for archaeological typologies". In: *Analecta Praehistorica Leidensia* 28 (1996), pp. 155–164.

[3] A. Bevan, J. Conolly, C. Henning, A. Johnston, A. Quercia, L. Spencer, and J. Vroom. "Measuring chronological uncertainty in intensive survey finds: a casse study from Antikythera, Greece". In: *Archaeometry* 55 (2 2013), pp. 312–328. DOI: 10.1111/j.1475-4754.2012.00674.x.

[4] C. Cáceres-Puerto and J. G. Sánchez. "El uso funerario en Augusta Emerita, siglos I a.C. a III d.C. Una propuesta de empleo de modelos probabilísticos y su repercusión espacial". In: *Anas* 31-32 (2019), pp. 1130–1929. URL: http://hdl.handle.net/10261/230192.

[5] O. Castillo, P. Melin, J. Kacprzyk, and W. Pedrycz. "Type-2 Fuzzy Logic: Theory and Applications". In: *2007 IEEE International Conference on Granular Computing (GRC 2007)*. 2007, pp. 145–145. DOI: 10.1109/GrC.2007.118.

[6] G. Fusco, F. Bertoncello, J. Candau, K. Emsellem, T. Huet, C. Longhi, S. Poinat, J.-L. Primon, and C. Rinaudo. "Faire science avec l'incertitude : réflexions sur la production des connaissances en Sciences Humaines et Sociales". In: *Incertitude et connaissances en SHS : production, diffusion, transfert*. 2014. URL: https://halshs.archives-ouvertes.fr/halshs-01166287.

[7] J. Fusco. "Analyse des dynamiques spatio-temporelles des systèmes de peuplement dans un contexte d'incertitude : Application à l'archéologie spatiale". 2016.

[8] J. Á. H. Gárate. "La utilización sepulcral de las cuevas en época visigoda : los casos de Las Penas, La Garma y el Portillo del Arenal (Cantabria)". In: *Munibe. Antropología-Arqueología* 62 (1 2011), pp. 351–402.

[9] Á. V. Gómez, J. Á. H. Gárate, and M. L. S. Gancedo. "Broches de cinturón hispanovisigodos y otros materiales tardoantiguos de la Cueva de Las Penas (Mortera, Piélagos)". In: *Sautuola: Revista del Instituto de Prehistoria y Arqueología Sautuola* 11 (2005), pp. 247–277.

[10] S. Hermon and F. Niccolucci. "A Fuzzy Logic Approach to Typology in Archaeological Research". In: *CAA 2002. The Digital Heritage in Archaeology. Computer Applications and Quantitative Methods in Archaeology*. 2002.

[11] C. S. de Investigaciones Científicas (CSIC) Instituto de Ciencias del Patrimonio (Incipit). *Conml*. 2024. URL: http://www.conml.org/.

[12] V. Kaše, A. Sobotková, and P. Heřmánková. "Modeling temporal uncertainty in historical datasets". In: *Workshop on Computational Humanities Research*. 2023. URL: https://api.semanticscholar.org/CorpusID:265429284.

[13] E. Mamdani. "Advances in the linguistic synthesis of fuzzy controllers". In: *International Journal of Man-Machine Studies* 8.6 (1976), pp. 669–678. DOI: 10.1016/s0020-7373(76)80028-4.

[14] E. Mamdani. "Application of fuzzy algorithms for control of simple dynamic plant". In: *Proceedings of the Institution of Electrical Engineers* 121 (12 1974), 1585–1588(3). URL: https://digital-library.theiet.org/content/journals/10.1049/piee.1974.0328.

[15] E. Mamdani and S. Assilian. “An experiment in linguistic synthesis with a fuzzy logic controller”. In: *International Journal of Man-Machine Studies* 7.1 (1975), pp. 1–13. DOI: https://doi.org/10.1016/S0020-7373(75)80002-2.

[16] M. Á. Manzano. *Diccionario de Toponimia Magrebí en la Red (DICTOMAGRED).* 2014. URL: https://dictomagred.usal.es/.

[17] P. Martin-Rodilla. *ArchaeoFuzzyExpertise GitHub repository.* ©2024. URL: https:/github.com/patriciaMRodilla/ArchaeoFuzzyExpertise/.

[18] P. Martin-Rodilla and C. González-Pérez. “Conceptualization and non-relational implementation of ontological and epistemic vagueness of information in digital humanities”. In: *Informatics* 6 (2 2019).

[19] MathWorks. *MatLab Fuzzy inference Systems The MathWorks Inc. ©1994-2024.* 2024. URL: https://es.mathworks.com/help/fuzzy/mamfis.html.

[20] J. B. (Owens and R. Jones. *The Cooperation of Humans and Killer Whales (Orcinus orca): The Application of a Simple Fuzzy Rule-Based Model to a Historical System.* 2012.

[21] J. Owens and E. Coppola. *Fuzzy set theory (or fuzzy logic) to represent the messy data of complex human (and other) systems.* 2012.

[22] C. Petr, C. G. Fermüller, and C. Noguera. *Fuzzy Logic.* Ed. by E. Zalta and U. U. Nodelman. 2023. URL: https://plato.stanford.edu/archives/sum2023/entries/logic-fuzzy/.

[23] L. Prada. “Métodos estadísticos aplicados en Arqueología, Trabajo Fin de Grado”. Universidad de Sevilla, 2015. URL: https://idus.us.es/handle/11441/40814.

[24] J. S. Ràfols. “La Villa Fortunatus, de Fraga”. In: *Empúries: revista de món clàssic i antiguitat tardana* 5 (1939), pp. 5–35. URL: https://www.raco.cat/index.php/Empuries/article/view/97101.

[25] C. D. Runz. “Imperfection, temps et espace : modélisation, analyse et visualisation dans un SIG archéologique”. Université de Reims Champagne-Ardenne, 2008. URL: https://tel.archives-ouvertes.fr/tel-00560668.

[26] C. D. Runz and É. Desjardin. “Exploration de données temporelles floues: application en archéologie”. In: *Revue des Nouvelles Technologies de l’Information (RNTI), numéro spécial fouille de données complexes* 21 (2011), pp. 249–266.

[27] C. D. Runz and E. Desjardin. *Imperfect Spatiotemporal Information Analysis in a GIS Application to Archæological Information Completion Hypothesis.* Ed. by R. Jeansoulin, O. Papini, H. Prade, and S. Schockaert. 2010. DOI: 10.1007/978-3-642-14755-5_13.

[28] M. Schneider. *Modelling Spatial Objects with Undetermined Boundaries using the Realm/ROSE Approach.* 2020. DOI: 10.1201/9781003062660-13.

[29] S. M. Taheri, F. I. Ghadim, and M. Kabirian. “Application of Fuzzy Inference Systems in Archaeology”. In: *7th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS).* 2019, pp. 1–4. DOI: 10.1109/cfis.2019.8692167.

[30] S. Termini. “On some “family resemblances” of fuzzy set theory and human sciences”. In: *Soft Computing in Humanities and Social Sciences.* Springer, 2011, pp. 39–54.

[31] A. Tirpáková, J. Vojteková, M. Vojtek, and I. Vikolinská. “Using Fuzzy Logic to Analyze the Spatial Distribution of Pottery in Unstratified Archaeological Sites: The Case of the Pobedim Hillfort (Slovakia)”. In: *Land MDPI* 10 (2 2021), pp. 1–16. DOI: 10.3390/land10020103.

[32] L. Tobalina-Pulido and C. González-Pérez. “Valoración de la calidad de los datos arqueológicos a través de la gestión de su vaguedad. Aplicación al estudio del poblamiento tardorromano”. In: *Complutum* 31-2 (2020), pp. 343–360. DOI: 10.5209/cmpl.72488.

[33] F. S. Trigueros. “Reflexiones sobre la calidad y la incertidumbre de los datos arqueológicos”. In: *Ligustinus* 1 (2013), pp. 9–19.

[34] R. Willet. “Experiments with Diachronic Data Distribution Methods Applied to Eastern Sigillata a in the Eastern Mediterranean”. In: *Herom* 3 (1 2014), pp. 39–69. DOI: 10.11116/herom.3.3.

[35] L. A. Zadeh. “Fuzzy Sets”. In: *Information and control* 8 (1965), pp. 338–353.

A. Extended results

This appendix condensed the complete results obtained from the fuzzy inference system implementation illustrated in section 4 of this paper. Note that the final goal of the illustrative example is to show how the proposed fuzzy model based on Mamdani’s system allows us to apply rules about the annotator’s expertise in a decoupled way from the membership function assigned value, both in use and chronology, for six categories of each one of the variables.

After the complete implementation of the system, we proceeded to validate the system with a list of 4 archaeological sites, previously detailed. The final results indicate that the system can apply fuzzy rules on each site and make a defuzzified review decision with Mamdani’s method (based on the centroid) that allows the archaeologist to have a review fuzzy assistance system.

Focusing case by case, we also obtain for each archaeological site studied: 1) a bar graph with each value assigned to its 4 input membership functions for all the defined categories and 2) the fuzzy surface graph for that specific archaeological site, that is, the situation of the archaeological site on the decision surface of the fuzzy inference system. These data are offered below for each of the sites studied.

Note in Figure 1 that the use is focused on the two main categories (basilica and Roman villa) but with no expertise in basilica by the annotator and a medium expertise in Roman villa.

Regarding chronology, the annotator indicates more expertise in the first three temporal phases than in the last three temporal phases. However, the most probable temporal phase (finally assigned with a 1 in the membership function) is the sixth one. These divergences between the final assignment and the expertise of the annotator cause the inference system, with the defined rules, to decide to review the site for the presence of high vagueness.

In addition, Figure 2 shows the overall decision surface of the system. Note that green and yellow areas of the graphic represent values near to revision, that is, archaeological sites with more vagueness in terms of the annotator’s expertise, whereas the darkest colors represent areas with less annotator’s vagueness. The location of the Fortunatus site in function of their

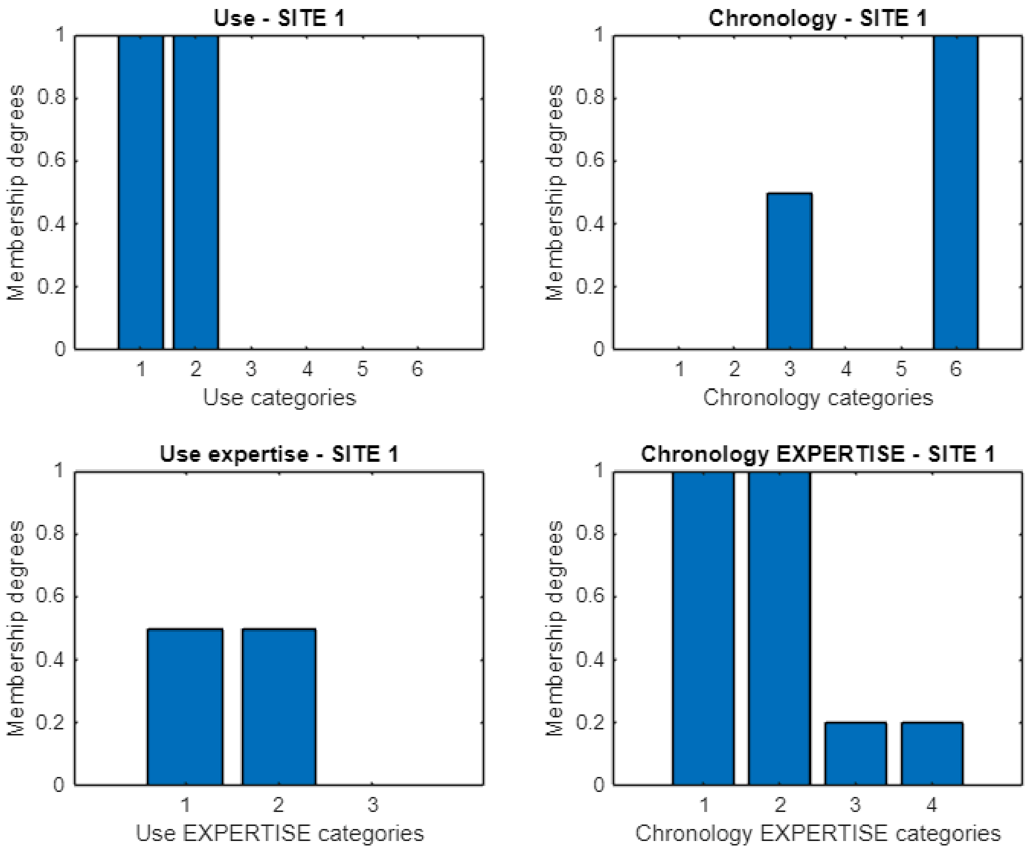

Figure 1: Site 1: Fortunatus's membership input values.

values for expertise is plotted as a red point in the shape, in a yellow area that explains the final decision of the system: the revision recommendation.

Note in Figure 3 (Guzmán site) that the use is focused on the two main categories (fortification and habitat). Regarding chronology, the annotator indicates a good expertise for all the temporal phases. However, the most probable temporal phases (assigned with a 1 in the membership function) are phases one and three. The alignment between the final assignment and the expertise of the annotator makes the inference system, with the defined rules, decide not to review the site for the presence of low vagueness. In addition, in Figure 4 (Guzmán, site 2) the location of the red point in a green-blue area explains the final decision of the system: not the revision recommendation. The site annotator is a chronology specialist, and the data's certainty is good, so it corresponds with the decision.

The El Mandalor site 3 case is similar to the previous site (Figure 5). The use of the site of Mandalor is focused on the two main categories (villa and farm) with a high expertise in these use categories by the annotator. However, the certainty of use of the phases is medium. Regarding chronology, the annotator indicates expertise in all the temporal phases assigned. This means that the high expertise in use and the medium expertise focus on the specific phases

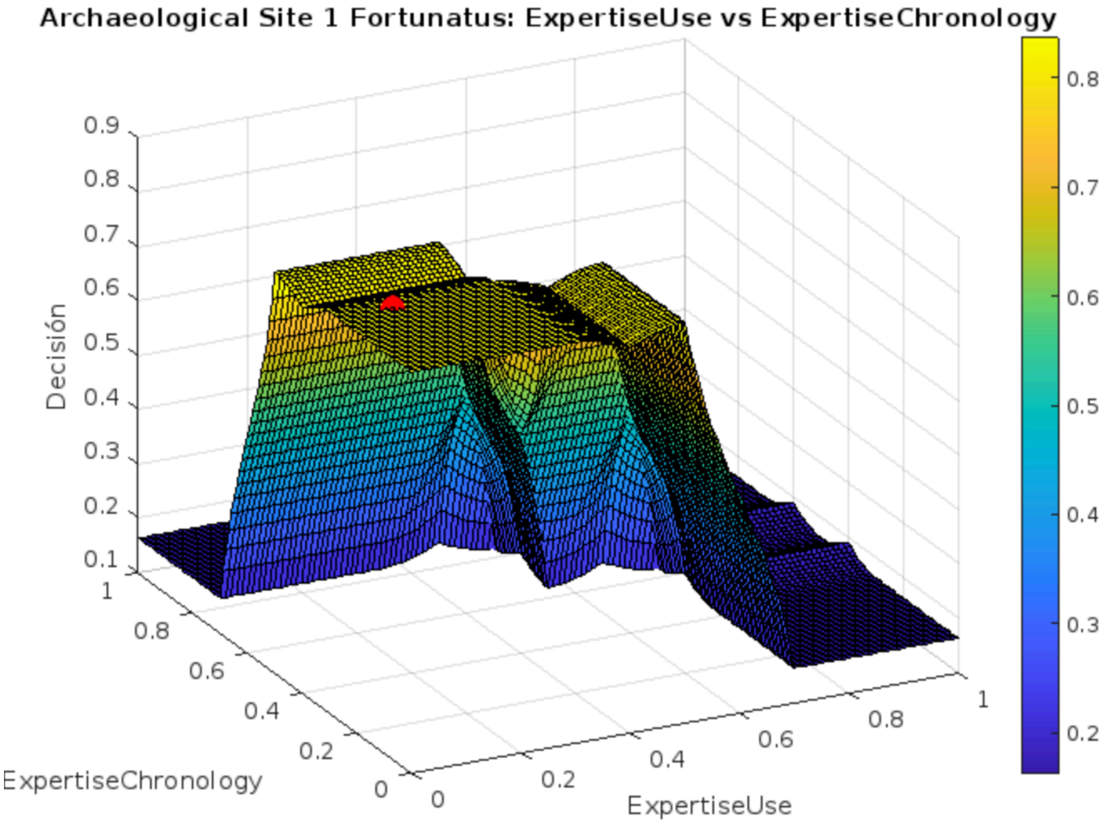

Figure 2: Site 1 location in the decision surface of the system implemented.

assigned allowing the system to output a non-revision recommendation. Figure 6 shows the location of the site in a dark blue area.

Finally, for case 4 Las Penas, we can see in Figure 7 that the use of the Las Penas site is focused on only one category (funeral space). Although there is no vagueness in the use membership function assigned, there is heterogeneous information in the annotator's expertise. Thus, note that the level of expertise of the annotator both in chronology and use is from medium to high in most of the categories, and high in the categories assigned. Applying the fuzzy rules defined, Figure 8 shows that the location of the Las Penas site in a dark blue area, with a no-revision recommendation final system decision.

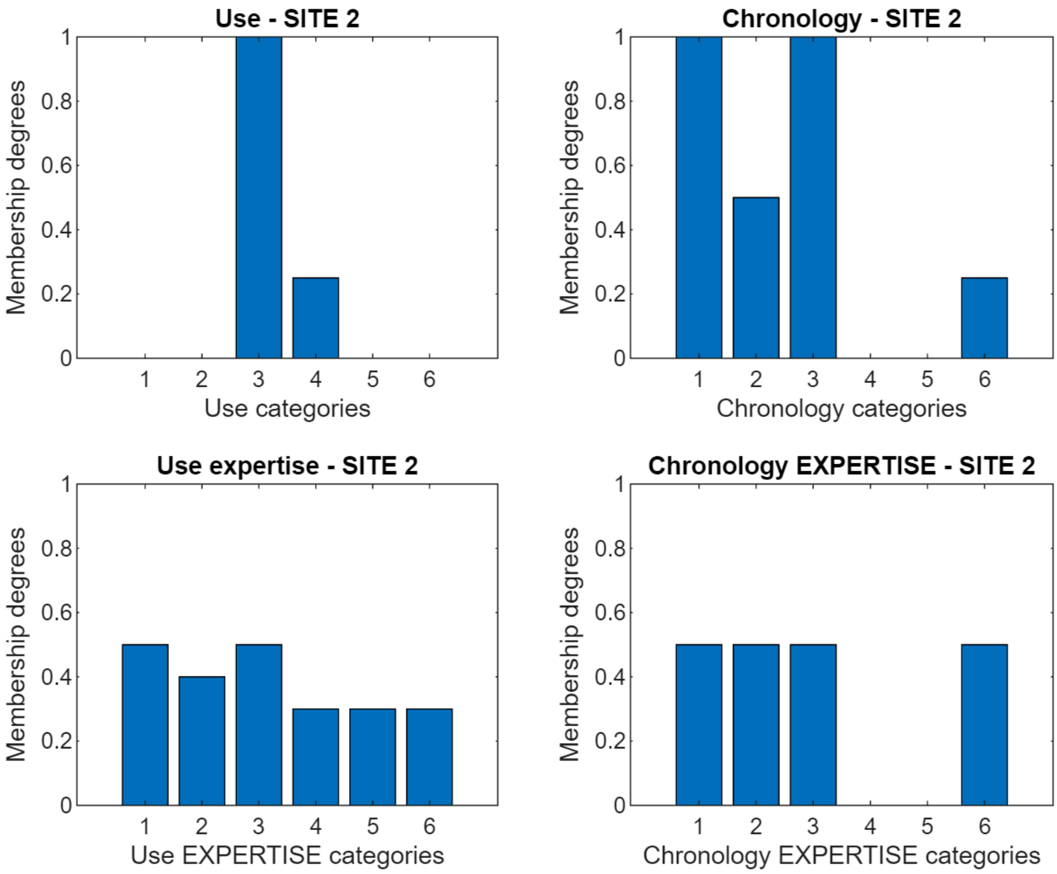

Figure 3: Site 2: Guzmán's membership input values.

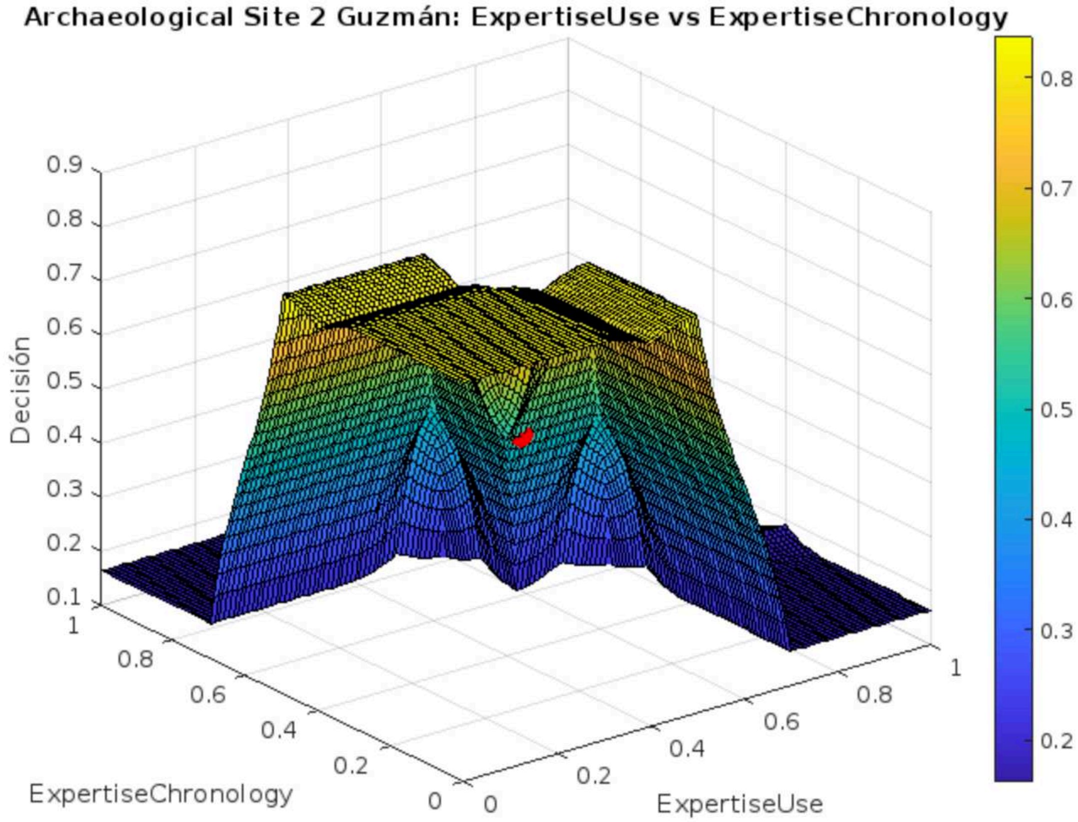

Figure 4: Site 2 location in the decision surface of the system implemented.

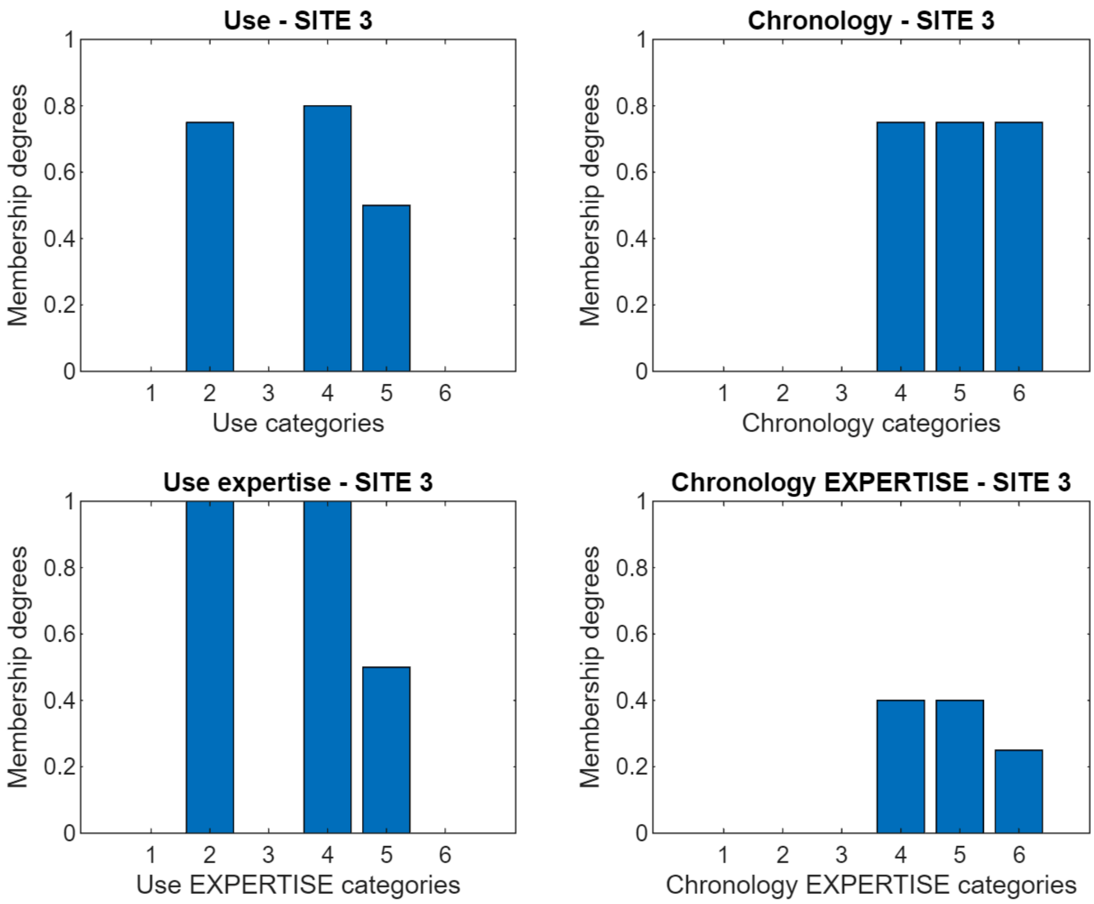

Figure 5: Site 3: El Mandalor's membership input values.

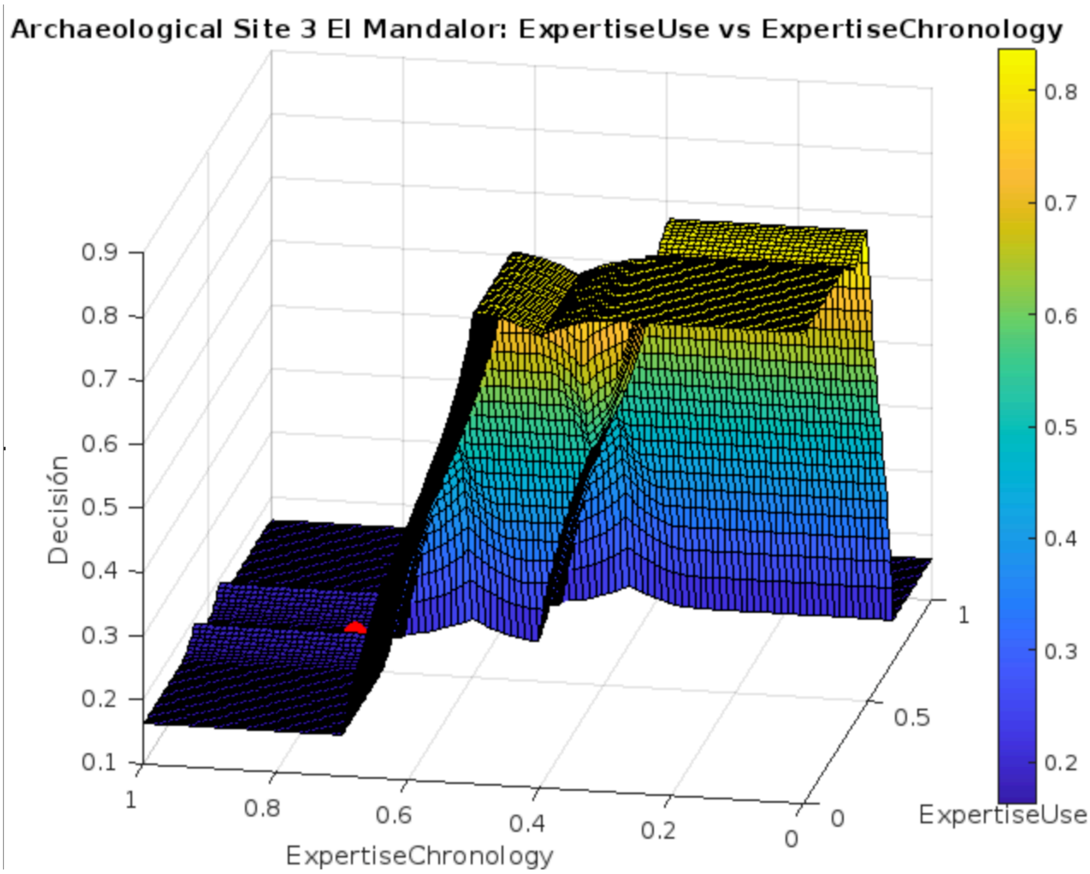

Figure 6: Site 3 location in the decision surface of the system implemented.

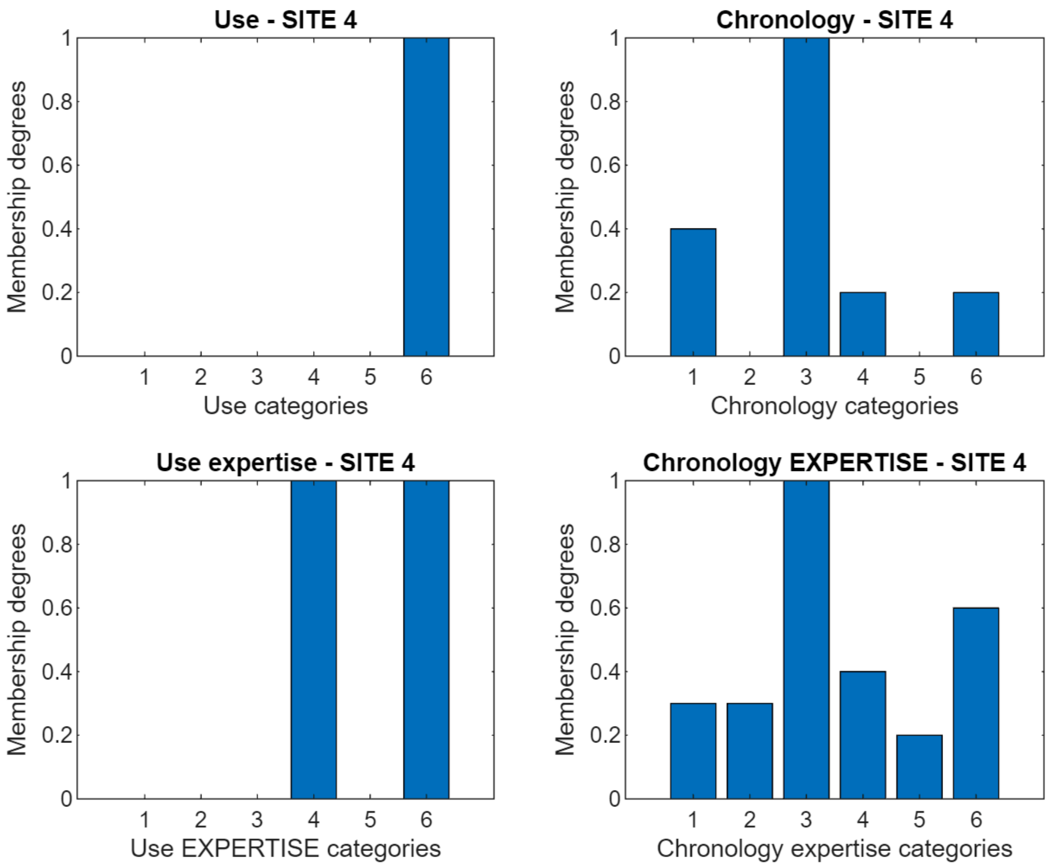

Figure 7: Site 4: Las Penas's membership input values.

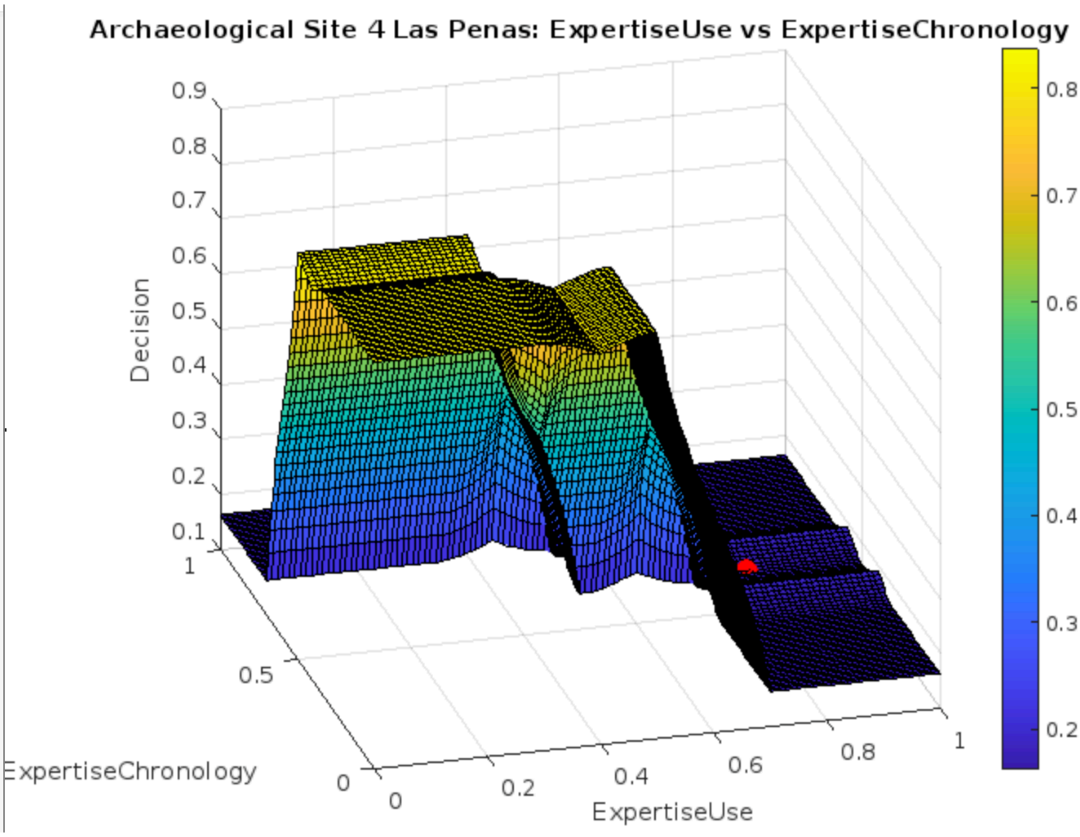

Figure 8: Site 4 location in the decision surface of the system implemented.