Decifarm: A Fuzzy Decision-support Environment for Smart Farming

Jérôme Dantan[1][a], Hajer Baazaoui Zghal[2,3][b] and Yann Pollet[4][c]
[1]*INTERACT Research Unit, UniLaSalle, 3 rue du Tronquet, 76130 Mont-Saint-Aignan, France*
[2]*ETIS UMR 8051, CY University, ENSEA, CNRS, F-95000, Cergy, France*
[3]*Riadi Laboratory, ENSI, Manouba University, Manouba, Tunisia*
[4]*CEDRIC Laboratory, Cnam, Paris, France*

Keywords: Decision Support, Fuzzy Logic, Sensor Network, Agriculture.

Abstract: Farmers need smart tools to optimize their crops and production. They need other agricultural experts such as advisors, accounting companies and also systems and software tools for decision support. The proposed solution is a fuzzy decision support Environment for smart farming (Decifarm) intended to ensure better data structuration extracted from farms, automated calculations, reducing the risk of missing operations, while ensuring data security. We designed a modular architecture to carry out these problems: we first provide crops phenological stages from both historical and forecast weather open data as well as historical data from sensors previously implemented, located at the parcels, with large amounts of data stored into a No-SQL document database; second, we provide control of an automatic water system based on fuzzy logic; finally, a prototype of hardware and software environments was designed from open hardware components, open source languages and open data, promoting both interoperability and extensibility.

1 INTRODUCTION

Agriculture is a strategic area in many countries, one main objective being to improve production in both quality and quantity, both globally and through individual efforts. In addition, agriculture is currently facing major challenges such as climate change, increasing in the world population, reduction of agricultural areas for the benefit of artificial land, more and more drastic regulations, etc. (American Farm Bureau Federation, 2016). The work of farmers benefited from mechanization during the 20th century but it is still difficult, with parameters that may be predicted, such as wage costs or inputs quantity, but also unpredictable ones, such as the weather or diseases. Major issues remain the significant use of fertilizers without intra-parcel modulation, the high risk of error, especially with many types of crops at the same time, and then losses in terms of resources (energy, water), when their estimated quantities are higher than necessary, generating additional costs.

In this work, we propose a fuzzy decision support Environment for smart farming (Decifarm), aiming at providing better information on environment, reducing risk of missing operations with automated calculation, and, at last, ensuring data security. We designed a modular architecture to carry out these problems. First, we provide estimations of crops phenological stages based on both historical and forecast weather open data and data issued from sensors located on the parcels, large amounts of data being stored into a No-SQL document database; second, we provide a an automatic water control process based on fuzzy logic; finally, a hardware and software prototype was designed based on open hardware components, open source languages and open data, promoting both interoperability and extensibility.

The remainder of this paper is organized as follows. Section 2 presents the background and existing solutions. Then, in section 3, we detail our proposal and its main components. Section 4 is

[a] https://orcid.org/0000-0001-7007-8725
[b] https://orcid.org/0000-0002-1224-1202
[c] https://orcid.org/0000-0002-5819-355X

Dantan, J., Zghal, H. and Pollet, Y.
Decifarm: A Fuzzy Decision-support Environment for Smart Farming.
DOI: 10.5220/0009828901360143
In *Proceedings of the 15th International Conference on Software Technologies (ICSOFT 2020)*, pages 136-143
ISBN: 978-989-758-443-5

dedicated to a presentation of the implementation with an evaluation of our proposal. Finally, we conclude and present our future work in section 5.

2 BACKGROUND

In the following, we present sensor networks, the Precision Farming (PF) existing solution, and mathematical background.

2.1 Wireless Sensor Networks

A Wireless Sensor Network (WSN) is an *ad hoc* network, most nodes of which being micro-sensors collecting and transmitting independent environmental data. The locations of nodes are not necessarily predetermined, and sensors may be randomly scattered on a geographical area of interest.

Data are routed to a node recognized as a "collection point", called *sink node*, that can be connected to the user network via the Internet, a satellite, etc. Users can send requests to other nodes of the network, specifying the type of data required and then collecting the captured environmental data through the sink node (Cambra C. et al., 2017). Joint advances in microelectronics, wireless networks and software have made possible to produce *micro-sensors* with only a few cubic millimetres volume, at a reasonable cost. Such an embedded device can integrate sensors providing digital measures of physical quantities (temperature, humidity, vibrations, radiation, ...), a processor with memory, input/output peripheral devices as well as a wireless transmission module.

In agriculture, micro-sensors may be integrated in the soil, and can answer to queries about the state of the field (e.g., detect driest areas in order to water them first). Digital agriculture is an essential step towards precision farming, which is one of the basic elements of resource-saving technologies (Skobelev P.O. et al., 2019). It requires both data geolocation and characterization of intra-plot heterogeneity (spatial distribution, number of sensors per plot, …). Many environments have been designed, especially in the form of Web and/or mobile applications, to meet agricultural needs and answer to problems faced by farmers in modern agriculture.

2.2 Existing Solutions

Terre-net [4] offers to follow live news related to agriculture while choosing a specific theme: news, machinery, culture, market flows. It is also possible to get the latest agricultural weather forecast (precipitation radar sensor).

Agriaffaires [5] offers various types of agro-equipment, forestry, handling or maintenance of green spaces and includes a real estate for sale and job offers.

There is a plethora of weather forecast applications, but Agricultural Weather [6] stands out because it is strictly designed for farmers. Information is refreshed every hour, and includes forecasts about precipitation, humidity, temperature, wind speed and direction, atmospheric pressure and cloudy movements, etc.

Field Navigator [7] application is an environment for drawing parcels, taking into account obstacles. The editor guarantees precision, even in poor visibility conditions.

(Cambra et al., 2019) present a powerful tool enabling real-time decisions based on data such as variable rate irrigation, or selected parameters from field and weather conditions.

These current solutions, where data is retrieved by means of services provided by external companies are unsatisfactory to farmers, for several reasons.

First, such solutions are exclusively based on technological advances, yet the farmers' participation in the innovation process and the technology customization on their needs appear to be quite limited. Next, farmers have generally to adapt to standard solutions suited for the greatest market share. Therefore, the proposed solutions do not fully suit the local heterogeneous agricultural needs. Nevertheless, customized solutions realized by businesses would be too expensive.

The role of farmers in the innovation process is not clearly defined: proposed solutions (software, innovations, data involved, and decisions via a "black box") are often proprietary ones, the farmer being considered as an end-user more than as an innovation actor, which would promote their autonomy.

Then, farmers' collaboration / participatory control on hardware / data / knowledge sharing / decision support is low. Indeed, providers centralize both data and "black box" decision tools, without a real collaboration between farmers in a region. In

[4] https://www.terre-net.fr/
[5] https://www.agriaffaires.com/
[6] https://www.lameteoagricole.net/
[7] https://play.google.com/store/apps/details?id=lt.noframe.farmisfieldnavigator.free&hl=en_US

addition, according to the American Farm Bureau Federation, farmers are growing awareness and concerns about the access to and the use of their farm data and the related major shift in role and power relationships (American Farm Bureau Federation, 2016). Finally, (Kritikos, 2017) noticed that "*As a result of these asymmetries, farmers' own particular needs and rights may be ignored, and inequalities are at risk of growing due to data-driven insights, rather than be reduced*"

2.3 Weather Data Sources

Weather data can be extracted from surface synoptic observations (SYNOP) circulating on the Global Telecommunication System (GTS) of the World Meteorological Organization (WMO). The available atmospheric parameters are either measured (temperature, humidity, wind direction and force, atmospheric pressure, precipitation depth) or observed (sensitive weather, description of clouds, visibility) from the Earth's surface. Data are extracted from both Opendatasoft [8] (historical data) and Openweathermap [9] (weather forecast) open data platforms, via a dedicated Application Programing Interface (API).

2.4 Mathematical Foundations

2.4.1 Fuzzy Logic

The fuzzy set theory has been initially introduced by L. Zadeh (Zadeh, 1965). It may be seen as an extension of the classic set theory to imprecisely defined sets and provides the basis for fuzzy logic. A fuzzy set F of a set E is defined by a membership function μ_F from E to $[0; 1]$, which associates each element x of E its membership degree $\mu_F(x)$, to the subset F, i.e.: x belongs "more or less" to F. When this membership function is normalized (i.e. a x value from E such as $\mu_F(x) = 1$ exists), $\mu_F(x)$ is then interpreted as the chance that F takes the value x ($\mu_F(x)$ is then a possibility distribution). We call it a "type 1 fuzzy set". In practice, we associate symbolic variables (e.g. words in everyday language such as "small", "large", etc.) to fuzzy subsets, then called "linguistic variables", thus automating reasoning during the implementation of fuzzy systems. *Linguistic predicates* (Zadeh, 1975) enable to reason on fuzzy (i.e. imprecise) concepts issued from experts' advices. To automate the aggregation process of these linguistic predicates, classical fuzzy operators, namely *T-norms* and *T-conorms* (e.g. min and max operators) are usually used.

2.4.2 Choquet Integral

Choquet's integral (Choquet, 1953) is widely used in multi-criteria decision support. With respect to a commonly made hypothesis, we consider here that the decision maker can formulate the possible dependencies between factors taken two by two. Under this hypothesis, the associated capacity is *2-additive* and the Choquet's integral may be expressed using formula (1):

$$Choquet_{\mu}(x_1, x_2, \dots x_n) = \sum_{i=1}^{n} (\gamma_i . x_i) - \frac{1}{2} \sum_{i,j \in N} I_{ij} |x_i - x_j| \quad (1)$$

With:
x_i: input value for criterion i
γ_i: average contribution of criterion i to the result.
I_{ij}: coefficient corresponding to a positive or negative interaction between criteria i and j.

More general expressions of Choquet's integral exist (Grabisch, Labreuche, 2008), but the interest of this writing comes from its simplicity of interpretation, clearly appearing as a generalization of the classical weighted sum which may lead to certain contradictions (Capet, Delavallade, 2013).

3 A FUZZY DECISION SUPPORT ENVIRONMENT FOR SMART FARMING

3.1 Challenges and Approach

Farmers need not only third parties to help them in making decisions and optimizing their crop production, but also systems and applications able to provide answers to several questions, as illustrated by the figure 1 below.

Figure 1: A synthesis of problems faced by farmers.

[8] https://data.opendatasoft.com/explore/dataset/donnees-synop-essentielles-omm%40public/table/?sort=date

[9] http://openweathermap.org/api

The proposed solution consists in designing a decision support environment for smart farming, based on open source / open data / open hardware, and specifically dedicated to farmers.

This environment is composed of three modules, which are 1) the sensor network, 2) the module in charge of detecting phenological stages, and 3) the module in charge of estimating required watering duration. Unlike others low-cost systems like (Cambra et al., 2017), our irrigation control system provides automating reasoning thanks to *linguistic predicates* managed by agricultural experts.

Our environment relies on the AgriLab® platform, a "new generation" agronomic laboratory which is, among others, a rapid prototyping platform in digital technologies (robotics, collaboration platforms, big data processing and decision support tools) and in agroequipments. This platform, which promotes the culture of knowledge sharing, is part of the worldwide movement of free knowledge exchange for a more sustainable agriculture.

3.2 Sensor Networks

Data are very expensive in the agriculture domain. Expert companies often refuse to make their data available, even for research. In addition, in order to test our fuzzy decision support system, we need a certain degree of accuracy, which open source data sets cannot provide in agriculture. That's why we decided to collect our own data for more availability and flexibility, according to our needs.

The main idea is to create a network of sensors attached to agricultural plots to measure physical variables, enabling to periodically store them remotely via a long range and low consumption communication protocol in open hardware. In the following parts, data are extracted from "do-it-yourself" humidity and temperature sensors network, based on open hardware and open data.

3.3 Detecting Phenological Stages

Phenological stage (φ) is a significant indicator providing information about the growth rate of a plant. It may be captured in different ways: in addition to direct visual inspection of plants, there are examples of calculations based on growing degree days (GDD), on satellite images or images captured by drones (e.g. normalized difference vegetation index (NDVI)). After their analysis, we can then compute the level of chlorophyll and detect the phenological stages.

3.3.1 Proposed Environment

We opted for a detection of phenological stages via GDD in order to be as close as possible to the farmer's concerns. In fact, plants grow cumulatively in stages strongly influenced by the ambient temperature. This method is widely used in the world of agriculture to identify the growth stages of crops, to predict the occurrence of pests dates, the date of flowering, the date of maturity of crops, so that the farmer to respectively process, provide inputs (fertilizer, water) and harvest at the right time.

The originality of our approach is both to periodically extract data from open source historical data APIs, and sensors located in their fields. GDD are also calculated in advance from weather forecasts.

The calculation method consists in applying a formula to minimum and maximum temperature for a given day and adding them as long as the crop year progresses, from sowing to harvest. The correspondence between GDD values and phenological stages (example: germination, emergence, 1 leaf, 3 leaves, early tillering, late tillering) differ according to the types of culture (maize, corn, etc.). The formula is (2):

$$GDD_i = ((Tmin_i + Tmax_i)/2) - T_{base} \quad (2)$$

With:
$Tmin_i$: minimum temperature on day i ;
$Tmax_i$: maximum temperature on day i ;
T_{base}: zero vegetation temperature.

Figure 2 illustrates the process of phenological stages detection based on two input data sources: on the one hand, data extracted from the database which stores the historical values of temperature sensors (Sensor), and, on the other hand, the temperature data get through APIs invocation (SYNOP model). Phenological stages φ_i are computed from temperature values.

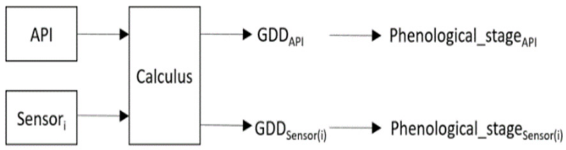

Figure 2: Process of phenological stages calculation.

3.3.2 Analysis

Farmers can look at the phenological stages of their crops and modify TBase after authentication on the

Web platform. Figure 3 shows the corresponding use case diagram:

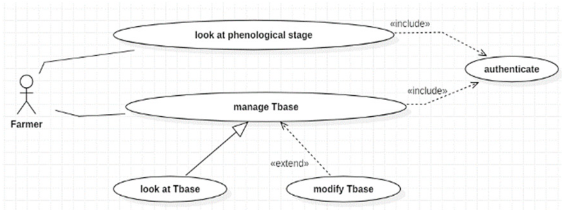

Figure 3: Phenological stages use case diagram.

When a user requests the phenological stage interface, the GDD is computed from the two data sources (API and Sensor) with the same formula.

3.4 Calculating Watering Duration

In this part, we propose a fuzzy system to calculate watering duration, expressed in minutes.

3.4.1 Proposed Approach

A fuzzy system takes real variables as input. It processes them through a fuzzification interface to obtain fuzzy input variables. The inference mechanism applies fuzzy rules to fuzzy input variables and finally sends them to a defuzzification interface to get real output variables.

We apply this principle on the watering system, which may take several input factors such as nature of soil, soil humidity, weather forecast, temperature, the slope of the soil, time, etc. In the case of only 2 factors: soil humidity and air temperature, inference rules apply according to the following rules (3):

IF temperature is {t_value} AND soil_humidity is {h_value} THEN watering_duration is {w_value} (3)

Inference rules associated to values are summarised into table 1:

Table 1: Inference rules for watering duration.

Rule nr	Premise 1: temperature is…	Premise 2: soil humidity is…	Consequent: watering duration is…
1	Boiling hot	Dry	Long
2	Hot	Moist	Medium
3	Hot	Dry	Long
4	Boiling hot	Moist	Medium
5	Freezing cold	Moist	Short
6	Freezing cold	Wet	Short
7	Cold	Dry	Long
8	Mild	Moist	Medium
9	Mild	Dry	Long
10	Cold	Wet	Short

Temperature values are in the range [0 ... 50] (degrees Celsius). For soil humidity, the values belong to [0… 35], and finally watering time, values are between 0 and 100 minutes. The following figures 4, 5, and 6 illustrate the used membership functions based on fuzzy logic applied to irrigation:

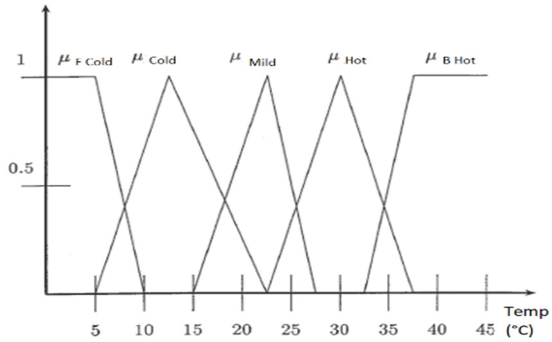

Figure 4: Temperature fuzzy input.

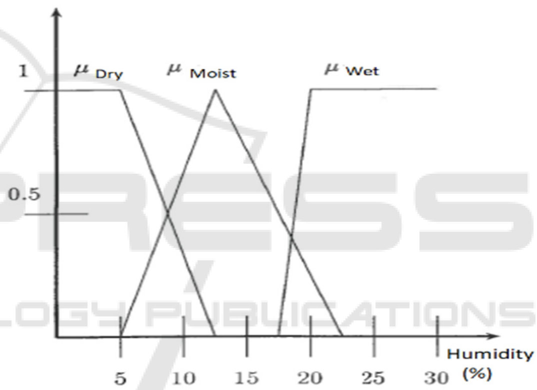

Figure 5: Humidity fuzzy input.

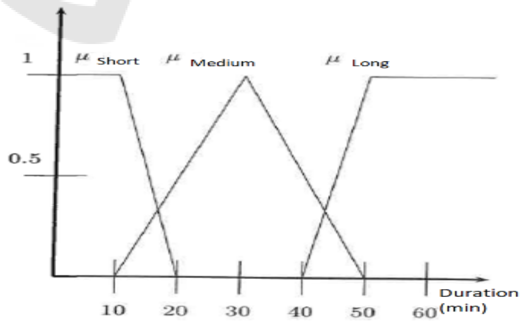

Figure 6: Watering duration fuzzy output.

3.4.2 Analysis

The application enables the farmer to consult the watering time, to receive alerts, to look at weather forecasts, API and sensor data.

Figure 7 illustrates the corresponding relative use case diagram:

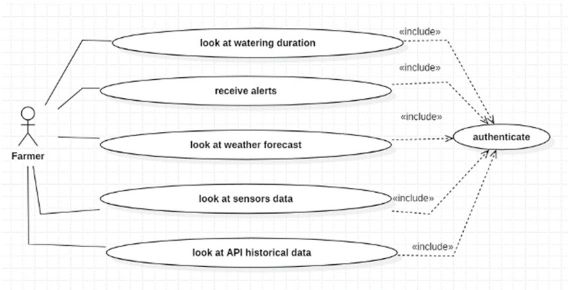

Figure 7: Watering duration use case diagram.

The fuzzy logic algorithm takes the last recorded soil temperature and soil humidity values as input to calculate the required watering duration (in minutes).

4 PROOF OF CONCEPT

This chapter deals with the design and production of a prototype.

4.1 Technical Choices

The Arduino Uno WiFi Rev 2 is designed around the new ATmega4809, u-blox Nina W102 WiFi module and an integrated inertial measurement unit.

The DHT11 sensor[10] measures both humidity and temperature. It is able to measure temperatures from 0 to + 50 ° C with an accuracy of +/- 2 ° C and relative humidity rates from 20 to 80% with an accuracy of +/- 5%. A measurement can be made every second.

The communication protocol depends on which characteristics you base yourself when making your choice, namely: memory requirement, battery life, number of nodes, transfer speed, range, etc.

Many communication protocols exist. For the prototype, we chose Wifi despite its limited range, UniLaSalle's experimental plots being located within a radius of a few hundred meters only. In the future, we will probably use Zigbee which consumes less energy and supports a larger number of nodes, since we don't manage very large datasets in real-time.

A far as software is concerned, we choose Flask, an open-source Python web development framework, and MongoDB, a database created in 2009, commonly defined as a NoSQL document database, and storing data as documents in a JSON format.

Python is an open source interpreted programming language created in 1991 by the programmer Guido van Rossum. Scikit-Fuzzy is a library of fuzzy logic algorithms written in Python language. We chose Spyder as Integrated Development Environment (IDE).

We designed C program for Arduino with the Arduino IDE, with both WiFiNINA.h and DHT.h libraries.

4.2 Sensors Network and Detection of Phenological Stages

This part is devoted to the "Detection of phenological stages" part of the prototype.

Figure 8 shows the class diagram. It is based on two entities which are "measure" (data retrieved from sensors) and "calculatedValue" (GDD).

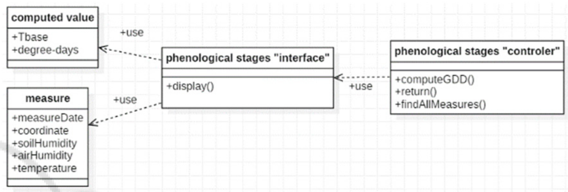

Figure 8: Phenological stages class diagram.

The base temperature (Tbase) depends on the chosen culture by users. After choosing their culture, users can consult GDD and the corresponding phenological stages. Every 24 hours at midnight, a job recovers the temperature values from the previous day, calculates the GDD and determines the corresponding phenological stages, from the information given in the literature. 8 temperature values per sensor and per day are extracted from the database (i.e. one measurement every 3 hours). The maximum and minimum temperatures are determined for each day and the degree-days are calculated from the date of sowing.

After configuration of the sensor network, the next step is to display historical data, as illustrated in figure 9.

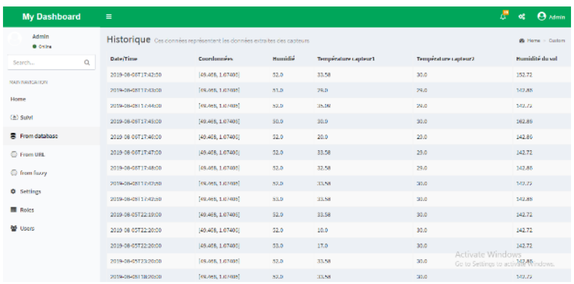

Figure 9: Historical data from sensors.

[10] https://learn.adafruit.com/dht

To compare with the values of the sensors, we extract data from APIs (temperature, humidity and precipitation) to provide several data sources to users (figure 10):

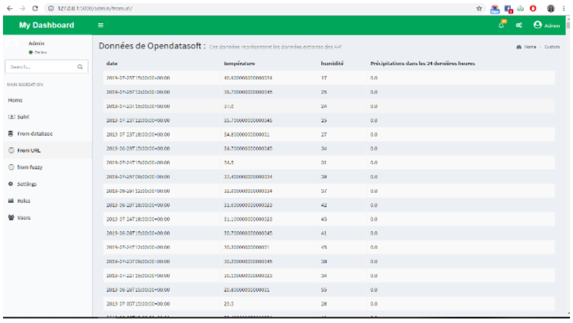

Figure 10: Data get from the opendatasoft API (SYNOP).

The interface below (figure 11) displays graphs with GDDs and phenological stages:

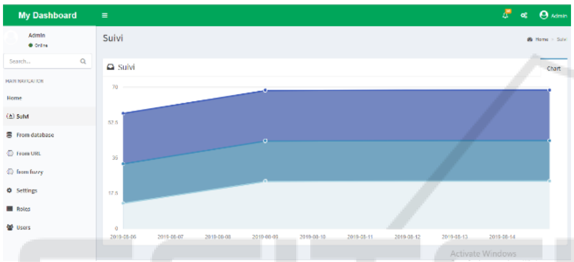

Figure 11: Graphs interface (GDDs and phenological stages).

Phenological stages depend on temperature values coming from different sources (sensors and API) which do not always deliver the same values. We shall therefore let the user choose the reliability of each sources himself, so that he/she can finally choose the most suitable phenological stage. For this, users must first assign values from 1 to 10, according to their preferences, to evaluation criteria, as shown in the table below:

Table 2: Characteristics of data sources according to an expert.

Criteria	1	2	3	4	5	6	7	8	9
C_1									
C_2				●					
C_3		●							
C_1&C_2								●	
C_1&C_3					●				
C_2&C_3					●				

With C_i: evaluation criteria such as proximity, error, etc. of data sources.

Experts can assess data sources by the means of these criteria. Table 3 shows an example of data sources evaluation:

Table 3: Data sources evaluation.

Data source	Properties	Evaluation
Sensor	C1	0.8
	C2	1
	C3	0.5
API	C1	0.2
	C2	0
	C3	0.8

We have applied Choquet's integral to order data sources according to the user's preferences. Phenological stages detection is an essential criterion to know not only the growth rate of the plant but also to give an idea on the actions to do.

4.3 Calculation of the Watering Duration

This section is devoted to the “calculation of the watering duration” part of the prototype, based on fuzzy logic.

Figure 12 shows the class diagram. It is based on two entities which are "measure" (data retrieved from sensors) and "computed value" (fuzzy system).

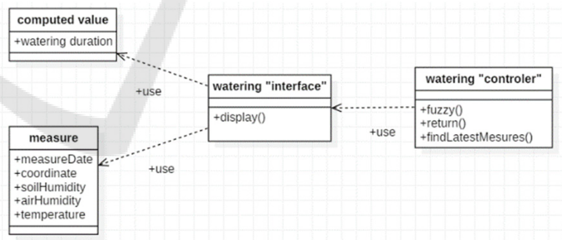

Figure 12: Watering duration class diagram.

4.3.1 Implementation

To display the recommended watering duration, we designed an algorithm based on fuzzy logic, which takes the soil temperature and humidity values as input. The interface presenting the required watering duration is illustrated in figure 13.

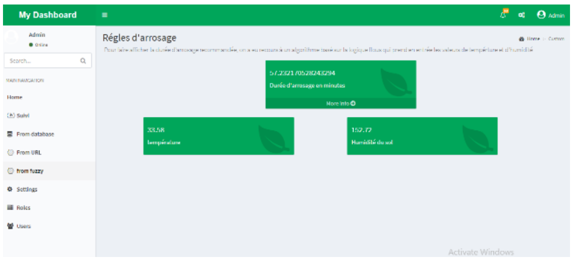

Figure 13: Watering rules.

We defined our fuzzy rules, fuzzification and defuzzification thanks to the "Sikkit fuzzy" python library. Here is an extract of the code, from the watering class:

Watering is strongly linked to humidity and temperature, but it could be interesting to add input variables which can also affect the watering duration such as the inclination of the soil, its nature (clay, sandy, humus, silty, etc.).

5 CONCLUSION

Nowadays, computerization of a domain is no longer considered as an evolution because we are used to finding IT everywhere. The need for computerization is justified by the need to have systems that replace humans to perform complex calculations, instant operations, store, archive, ... Finally, these systems operated to help us make our decisions, like the fuzzy decision support environment that we designed: indeed, its purpose is to guide the users to follow their crop stages and take the necessary precautions to guarantee better production in both quality and quantity. It is structured around three main components: a sensors network, phenological stages detection and watering duration calculation, based on fuzzy logic.

The hardware environment, based on open hardware, was chosen to respect the constraints of maintenance, cost and scalability. We chose Web-based technologies and Python as software environment for reasons of availability and extensibility of the environment. Indeed, to add a module that analyses satellite images or drones images to follow crops, we can simply integrate a python code into a Flask application which limits any interoperability problems.

In the future, we will propose the following evolutions: estimate the final production from the first phenological stages, add a module for analysing satellite images to reinforce the detection of the phenological stages and add other input variables to calculate the watering duration.

ACKNOWLEDGEMENTS

We warmly thank Farah Dronga for the development of the prototype.

All UML diagrams were produced with StarUML software.

REFERENCES

American Farm Bureau Federation, 2016. *Highlights of the Farm Bureau Big Data Survey*.

Cambra C. et al., 2017. *An IoT service-oriented system for agriculture monitoring*. 1-6. 10.1109/ICC.2017.7996640.

Cambra C. et al., 2019. *A Smart Decision System for Digital Farming*. Agronomy, 2019, vol 9(5), p.216. https://doi.org/10.3390/agronomy9050216.

Capet Ph, Delavallade Th, 2013. *L'évaluation de l'information – confiance et defiance*. 301p. Chapter 5, pp. 147-172. Lavoisier, 05/2013.

Choquet G., 1953. *Theory of capacities*. Annales de l'institut Fourier 5 (1954): 131-295.

Grabisch M., Labreuche C., 2008. *A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid*. 4OR 6, 1–44 (2008). https://doi.org/10.1007/s10288-007-0064-2

Kritikos M., 2017. *Precision agriculture in Europe. Legal, social and ethical considerations*. PE 603.207. European Parliamentary Research Service.

Skobelev P.O. et al., 2019. *Development of a Knowledge Base in the "Smart Farming" System for Agricultural Enterprise Management*, Procedia Computer Science, Vol. 150, 2019, pp.154-161, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2019.02.029.

Zadeh L.A., 1965. *Fuzzy Sets*. Information and Control, Vol 8, Academic Press.

Zadeh L.A., 1975. *The Concept of a Linguistic Variable and Its Application to Approximate Reasoning–1*. Information Sciences, vol. 8, pp. 199–249.