GRD Journal for Engineering | Volume 7 | Issue 2 | January 2022
ISSN: 2455-5703

A Fuzzy Time Inference Prototype for Rice Crop Watering

Daniela López De Luise
Department of Computer Science and Engineering
CI2S Labs

Juan Ignacio Marandino
Department of Computer Science and Engineering
Autonomous University of Entre Ríos (Entre Ríos, Argentina)

Federico David Roldán
Department of Computer Science and Engineering
Autonomous University of Entre Ríos (Entre Ríos, Argentina)

Pedro Juan Labonia
Department of Information Technology
CAETI - UAI

Diego Rafael Kracoff
Department of Information Technology
Inter-American Open University

Abstract

Approaches for time mining are frequently used to build a model of how the dynamic of certain (productive or non-productive) activities, statuses, and events occurring in a system under study. Sometimes it is also used to test hypotheses concerning that behavior. But these approaches are mostly applied to model complex activities, with a large number and variety of parameters and multiple activity combinations that are relevant. As a consequence, the algorithms applied are slow and cannot be used for real-time processes, where time allocation for making the analysis and choices is really hard. This paper presents a first prototype for applying an advanced heuristic for time mining called Harmonic Systems, able to work in real-time, for solving the problem of rice yield watering using Fuzzy Patterns. The project involves three main stages for gradually adjusting the hardware to a robust and lightweight outdoor architecture, and the variables to be relevant according to the mining previous findings and newly introduced data sources. The testing and statistics performed confirm that the main characteristics of the ground stations, with a reduced device set, are adequate to perform a 24 x 7 sampling with the appropriate accuracy.
Keywords- Rice Crops, Precision Farm, Arduino, Time Mining, Harmonic Systems

I. INTRODUCTION

The task of sequential pattern mining is useful for various applications, like marketing, Decision Support Systems, and inferences depending on time in like radar image processing [1], smartphone data processing [2], econometrics [3], risk predictions [4], and precision crops [5]. An important issue is to discover typical patterns that follow a precise sequence according to corresponding intervals. Chang et al. propose [6] a Time Mining with fuzzy time intervals (SPFTI). The main idea of SPFTI algorithm is to use an Apriori-like heuristic. According to the authors it outperforms with the fuzzy sequential patterns mining with fixed time interval. Of course the problem is here the variable determination, in contrast to the proposal of this paper. Other interesting approach is centered in the degree of distance between time-series patterned attributes [7] named Fuzzy clustering and Dynamic Time Warping (DTW), where the perspective differs from the current paper proposal in the hypothesis of having static time distances during the model building, so they have their own pattern that doesn't evolve.

There are many proposal for performing inferencing by fuzzy variables [8] - [17] but in all cases the platforms and heuristics require lots of restrictions, pre-encoded information, or complex steps to derive meta-information.

This paper introduces the design and first prototyping of KRONOS, a time mining model for rice crops, from on-site devices to a novel platform that collects and monitors data, applies HS, and develops many reports. Some of the activities involve self-testing devices, cross-validation of data, smart data security, local and remote access and control, data filtering, and data learning. It combines Fuzzy Logic and Harmonic Systems (HS). HS is a good heuristic for Real-Time prediction in the context of precision crops with scarce information. Specifically, there are some contributions in the field of rice [18].

Almost half of the world population considers rice as vital nourishment. Rice is a staple food that provides between 35 and 60% of the required calories consumed. More than 75% of the world's rice production is obtained from irrigated farms. The water flood where rice plants are raised is the key difference from the remaining relevant crops. But a severe problem emerges since hydric sources are dramatically being reduced [19].

Recently, a shortage of water resources is evident around the world. There is an increasing pressure to reduce water usage in irrigation agriculture. Rice production is an obvious target for water preservation since more than 30% of irrigated lands are with this plant, representing about 50% of the total amount of used water for irrigation. A 10% reduction of the hydric usage for rice crops could result in 150.000 million cubic meters. This is approximately 25% of the freshwater used in the entire planet for other uses [20].

The rice productive structure in Entre Ríos province, in northeastern Argentina, is mainly by border irrigation with a constant level during a period from 90 to 100 days. Most of the water comes underground, extracted from the Salto Chico formation aquifer. Extraction is carried out by Wells 60 to 80 meters deep with energy-intensive centrifugal pumps. Most of the Entre Ríos extractions are with deep wells, so the energy impact is much greater than in the rest of the rice-producing provinces of Argentina. Furthermore, the average water consumption in an entire production cycle is around 1,300 mm [21].

As a consequence of cited characteristics, rice crops in this context require careful monitoring to avoid economic losses. This project aims to technically present a prototype based on data collected with simple and cheap on-site stations specifically designed to collect a minimum set of data in order to evaluate soil moisture. It is intended to cross data and metadata from radar and satellite in order to model with HS the humidity condition at every moment.

A proper set of alarms and reports allows to conveniently handle water usage according to land requirements, for water-saving and better rice plant yielding. The goal of KRONOS is to study the spatial and temporal humidity/ climate balance and to develop tuned behavior patterns. Also, specific sampling rates are suitable for different soils, seasons, and weather. Traditional approaches have standardized certain variables which have been considered as well for this proposal. The inference analyses a reduced set of sampling data in Real-Time, in order to gain rapid control with a few critical data. Other data is used to provide redundancy and robustness, as fault tolerance and precision are the main concerns as well. The prototype manages extra data that allows to mine crop behavior with other perspectives. A set of patterns define the key relationships detected during Data Mining (DM) processing, and are the first input to the HS. Upon pattern detection on data, a set of proper alarms allow for early detection and acknowledge different critical situations for optimized water usage.

Time Mining Systems are being used in diverse fields. Some approaches are mainly statistical evaluations of the yielding variability [22]. Others aim to design long term strategies [23], to estimate the impact of climate change on water requirements [24], predict moisture conditions in crops [25], even to model the phenology of rice [26] and crop productivity in general [27]. In some cases, satellite information is used as well to model crop data [28] and to adjust the precision where the water condition is critical for the success of the crop [29]. The trend for the future is to sustain this type of technology to support sustainable production systems in an optimized way.

In the current world context, there is freshwater demand. The Food and Agriculture Organisation of the United Nations, also known as FAO [30] expresses that agriculture competes in the consumption of this resource globally. In general, 69% of the water consumed corresponds to agriculture, and the rest to domestic (10%) and industrial (20%) usage. Therefore, the observation of its consumption is a priority.

The proposal of this paper not only aims to be a tool for a better saving of water in the irrigation process but also involves a prototype suitable for systematic replication since the design is on a low-cost equipment base. Extensions and adaptations could be made relatively easily since the developments are generic and the specific data is processed with generic software and hardware technologies.

The rest of this paper is organized as follows: Harmonics Systems introduction (section II), the technical description of the proposal (section III), a detailed consideration on powering constraints, low consumption electronics and hardware architecture (section IV), some information about communications interfaces (section V), testing (section VI) and conclusions (section VII).

II. HARMONICS SYSTEMS

Information management usually involves mining techniques. Among others, special interests for this paper are those related to process temporality. It allows being able to make useful predictions for business, WEB, science and industry [31]. From the traditional approaches, the focus is on the features of certain events under analysis [32] [33]. Harmonics Systems (HS) focuses on the temporality of the process. It could be considered as the opposite stance to the traditional, taking variables describing event's characteristics as a known stage and evaluating rhythm, accelerations, static periods, and other time-dependent behavior.

A. Harmonic Systems Purpose

The main goal of HS is to determine the evolution during a certain period of well-known events typically described as sets of patterns. Each of them consists of a reduced set of variables with values ranging in predetermined ranges.

B. Harmonic Systems Applicability

As an on-the-fly pattern matching process it can be used to detect hardware failures, software in-progress hacking, risky events on the fly, and any other software or hardware-specific processes under interest. The main difference with other approaches is the plastic parameter adaptation process performed in the background. These characteristics allow HS to perform a kind of just-in-time modelling, self-adapting to eventual peculiarities as could be: changes in weather due to different seasons of special crop circumstances.

C. Harmonic Systems Components

The patterns are predefined aggregations of properties identified with an ID, with a specific order in time and duration.

When a pattern is matched against data produced in a system, it is a pattern resonance and the model adapts internal parameters to the current circumstances.
Main HS components are [34]:

- Pattern threshold U: used to consider minor timing differences between current and theoretical pattern
- Pattern variables: set of features under interest
- Variable time t: lapse of the variable holding a value
- Variable lambda: coefficient that indicates how much the declared value for a variable could change
- Sample size n: a cut-off number to indicate if the inferences in the model follow Binomial or Poisson distributions
- Filters: optionally, a set of high-pass, low-pass and band-pass filters could be enabled to process just part of the real-time incoming data.

III. THE PROPOSAL

The project is performed in the IDTI Lab at UADER University (Entre Ríos province in Argentina) in collaboration with the CAETI Lab at UAI University, National Commission for Aerospatial Activities (CONAE by its name in Spanish), and National Institute of Agronomy Technology (INTA). While IDTI Lab and CAETI provide the design and implementation of the collecting and inference systems, CONAE provides, collects, analyzes, and models satellite products covering the region under study. The information covers biophysical parameters and their relationship with the studied crops. These later data are gathered by using images from Synthetic and optic Aperture Radar. Mostly, data come from images and information from satellites SAOCOM 1A and 1B.

INTA is responsible for the coordination of the activities with the owner of the farm, the monitoring of water usage and the proper variable collection. Part of the calibration of the sensors and data sampling validation are also part of the tasks. Also provides expert knowledge during the mining and knowledge drifting, specifically regarding the type of crops and soils.

UADER is in charge of the technological project from Data Mining and software production. The activities are in collaboration with CAETI labs. The infrastructure, prototyping, and technical decisions are also part of this UADER and CAETI. The remainder of this section technically describes the prototype as a tool for data collecting and HS predictions in the context of rice crops.

A. Global Architecture

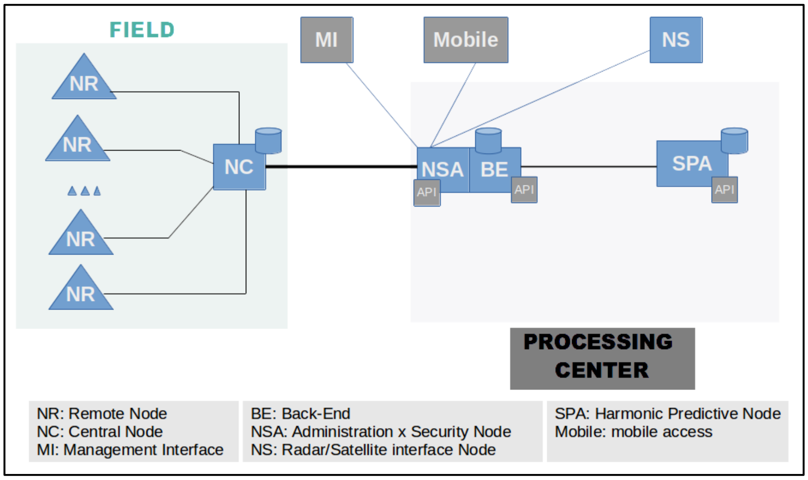


Fig. 1. KRONOS architecture

There is a back-end node (BE) implemented in Python with WEB and wireless connectivity, complementing direct interactions with entities and other external subsystems. Among other considerations, some of the main global parameters of the prototype are:

1) Type and name of data source
2) Type of data being collected
3) Sampling timing
4) Data relevance for the main DM goal

The entire prototype works in a star architecture as a first implementation stage. Nodes perform different activities for BE, which is attached to Administration Security Node (NSA). Remote Nodes (NR) are in-land nodes with specific sensors, each feeding a central node (NC)in a round-robin fashion. The Central Node at its time also has certain extra sensors, and sends its data plus NR collected data to BE. NS provides radar and satellite data as well but independently. Several APIS are for mobile (Mobile), WEB (MI), PC, and host accesses (denoted with grey tags API in Figure 1).
CONAE measurements are under a strict protocol combining biophysical variables with optical data from satellite and radar and are working to add information from an extended electromagnetic spectrum.

INTA determines the best approach to monitor the process of watering, devices and sensor administration in the field. Also complements these tasks with on-site validations of sensor samplings and inferences precision over the entire growing process of the crop.

B. Connectivity Architecture

The communications between nodes are as follows:

- Radio and/or WLAN between NR and NC
- Mobile connection between MI and NC
- Mobile connection between BE and BC to provide SMS alarms
- Internet Wifi between BE and APIs

There are many logs: for NR-NC interchanges, NR faults, bad connectivity, etc.

C. Back-End Node

The Back-End manages all data including remote-diagnose from NR and NC (information about battery, memory, and device status). Status can be accessed through NSA from mobile and MI interfaces. The inference module (SPA) has a dedicated interface. Users can also access special stats according to its privileges:

- Instantaneous flow rate (flow rate/ surface): useful to detect high-risk scenarios.
- Recording of key dates (sowing, fertilization, irrigation starting, etc.)
- Recording of 2-weekly phenology and growth rate per irrigation sector (to assess if it is worth being included in the predictive model)
- Recording of soil samplings to determine gravimetric humidity (per irrigation sector)
- Recording of irrigation sector productivity.
- Derivation and management of Kc coefficient being:

$$Kc = Kcb + K\epsilon \quad \text{(eq. 10)}$$

Where Kcb stands for basal crop coefficient, and Ke is coefficient of soil water evaporation. It is used to infer the evolution of the humidity map.

D. SPA Node

This node performs inference tasks in order to advise on watering timing. It has several modules for: data loading, patterns administration, reporting, parameters management, filtering administration and alarm set-up.

1) Data Loading

Loads data sent by NC. Source, security information, variables, and metadata. Validates data and processes timestamps. Registers its activity in a log.

2) Pattern Management

Handles the lifecycle of patterns, from variable to pattern parameters and related information: Lambda, U threshold, name, ID, sample size (n) and variables. The patterns are managed following fuzzy logic in order to extend the model to problem language and uncertainty. Any activity is recorded in a log.

3) Alarms

Interacts with the resonance administration routines. It handles sensibility parameters, and a set of possible alarms: relevance of the alarm, information channel, receivers, etc. Alarms are included in a special log.

4) Reports

It is an interface (API) to request logs, data, and any system's parameters information, statistics, and reports. Can handle dates and coverage of the data in the query. Its activity is held in report.log

5) Filters

Manages any information to define, enable or disable data filtering prior to being sent to the SPA node. Only high-access users can access it. Its activity is recorded in a special log.

6) Parameters

API for global inferences administration. The parameters that are handled are: nU (threshold for U sensibility), U (resonance threshold), nL (variables timing stability), NC (type of statistical consideration for inferences). Its activity is recorded in a log.

7) Resonance

It has a detector and a coach. The first detects patterns following the pattern states in Table 1. The second checks whether the sample size is big or not, to be able to accordingly update U.

8) Patterns

Monitors resonance and patterns according to different modes (“signal” for sending alarms with timestamps and ID of patterns, “control” to hold down the system activity until a second pattern is activated, “recording” for recording resonances, “monitor” for handling a subset of patterns with higher priority)

Table 1: Pattern states

	Target state			
Source State	*OFF (value 0)*	*ON (value 1)*	*WIP (value 2)*	*RES (value 3)*
OFF		*Activate (pattern(i))*	*NA*	*NA*
ON	*DeActivate (pattern(i))*		*checkPatternVariable(pattern(i))=1 & verifyVariable(pattern(i))= YES*	*NA*
WIP	*DeActivate (pattern(i))*	*checkVariable (pattern(i))= NO*	*checkPatternVariable(pattern(i))<= checkSIZEpattern & checkVariable(pattern(i))= YES*	*checkPatternVariable(pattern(i))> checkSIZEpattern & checkVariable(pattern(i))= YES*
RES	*NA*	*patternActivity(pattern(i))*	*NA*	

E. Administrative Security Node

The security node (NSA) is in charge of user administration, and security violations at any node.

F. Remote Node

The Remote Node (NR) is an on-site sampling of field humidity conditions, sensing water surface depth, soil humidity, air temperature, water temperature, and any other variable determined in every case. It also performs self-calibration and self-diagnostics. There is one NR every 200 m. It is mainly a simple ESP32 and LORA with a battery and solar cells.

G. Central Node

The Central Node (NC) is more complex than NR. It adds devices to sample wind, temperature, radiation, and pluviometer. As in NR, it is an ESP32 connected with NR, forming a WLAN. It coordinates every calibration process in its own devices, diagnostics and data sampled reporting to BE node.

IV. THE POWER CONSTRAINTS

Both NR and NC nodes are off-grid systems. It is important to note that the architecture is able to work in any outdoor sunlight condition. That is a power isolated system in a rural region. It is typically composed of photo-voltaic panels, charge regulators, at least a battery, and a power regulator. Panels and batteries are key parts and must have a proper dimension in order to enable 24 x 7 activity. This calculation is determined by the radiation on a determined surface, which depends on the incidence slope (β) and orientation (O). β is the angle between panels and horizontal (in the range [0:90]). O corresponds to the angle between the panel surface and meridian (in the range [-90:+90], being value 0° in the case of placement in the southern hemisphere. Positioning for the field under study in this paper is performed with a NASA site [35]. The location in Concepción Del Uruguay is Latitude -32.4056, longitude -58.3708. It is used to get specific weather information (cloud and lighting conditions, β and O values, etc.). This data is relevant for improving solar radiation usage and very specific to determine the temperature and the shadows due to obstacles in the neighborhood. The design of modules for battery administration, input regulation, and power input switching are based on this and other information.

A solar tracker is implemented to maximize the number of lighting hours on panels. The boards for NC and NR are PVC and their dimensions are smaller than previous versions in order to let them be portable and practical for handling during the operations in the field (see Fig. 2 and 3 respectively).

Fig. 2: Board for module

Fig. 3: Board for module NR

As the in-field processing is battery dependent, it is very important to save power. For that reason, controllers are taken to deep suspension whenever it is possible, and components are low consumption as well. The highest power load is during NC control and communication activities. The prototype is designed to evolve in three main stages: beta (with a few devices working at the in-door lab), alpha1 (with a few devices working at a small yard), and alpha2 (with an extended set of devices working in a rice crop).

This paper is in the beta stage with soil humidity, air humidity, canopy temperature, and water level in NR. Regarding NC there are canopy temperature, anemometer and global radiation sensors. The battery level is dummy data since the test runs in the lab.

V. TEST CASE

The testing performed for this paper corresponds to the beta stage described earlier. Just 2 nodes were installed (NC and NR) in the backyard of the IDTI Lab. Each one is sensing a sample plant in pots 5m away.

The setting of the devices are:

1) LaRa Wan for connecting both nodes
2) Sampling every 30 minutes
3) An instant connection between NC and the Server
4) NC with a local storage for recording both NC and NR information
5) NC working as a client and a PC as server

After sampling, the client (NC) sends a PUBLISH message with topic “values”. The server checks the information and processes it: that is to show the collected data (as it is a preliminary stage).

The activity is registered in the LaRaWan log as in Figure 4.

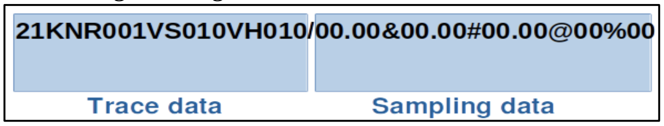


Fig. 4: LaRa Wan Log structure

Any sampling information is located after LaRa Wan trace information. All sample data are positive numbers in float format.

The dynamics is as follows:

1) Every 30 minutes NC and NR performs samplings
2) NR sends data to NC
3) NC sorts NR data and formats it as CSV
4) NC records CSV information in a local memory
5) Afterwards, NC uploads combined information to the cloud using MQTT protocol.
6) Central server (NSA) collects samplings and stores them locally.

Figure 5 shows the structure of the streaming in step 4.

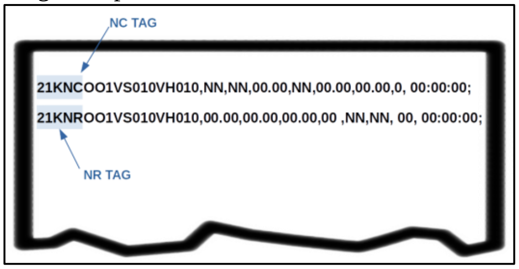


Fig. 5: Streaming in MQTT

As can be seen in the figure, the first entry is for NC data and the second for NR. After interchanging data both nodes set to Sleep mode.

Table 2: presents a summary of the testing log

Status	Stream
RTC timestamp < Sampling timestamp *NC and NR ready for sampling*	*LoRa Receiver* *LoRa Initializing OK* *Resetting WDT* *Starting the function: lora received packet();* *Lora packet received: 21KNR001VS010VH010/19.30&41.00# 22.60@0%90 with RSSI -72* *Starting the function: lora send confirmation data(); Sending confirmation received packet* *Sleep LoRa zZZzzzZZ* *Starting the function: RTC get Date Time();* *08-30-2021 19:11:51* */08-30-2021.csv* *19:11:51*
Sampling performed *NR reporting to NC*	*Starting the function: Environment Sensing();* *Starting the function: SD Card init.();* *SD Card Type: SDHC* *SD Card Size: 15279MB* *[E, [vfs api.cpp: 64, open() : / sd/08-30-202 1. csv does not exist* *File doesn't exist Creating file... Writing file: /08-30-2021.csv* *File written* *Ordering Data...* *Starting the function: SD Append File();* *Appending to file: /08-30-2021.csv* *Message appended*
NC ready for persistence in local SD	*Saving data of central node: 21KNCOO1VS010VH010,NN,NN,22.60,NN,3.00,0,80, 19:11:51* *Starting the function: SD Append File(); Appending to file: /08-30-2021.csv* *Message appended* *Saving data of remote node: 21KNROO1VS010VH010,19.30,22.60,34.20,22 ,NN,NN, 90, 19:11:51* *Going to Sleep Now*

The log starts denoting that RTC is newer than sampling timestamp, which is expected since it is a first formal testing. Then it describes the steps performed by NC, its coordination with NR and the formatting prior to uploading everything with MQTT. Table 3 is the data validation process, which shows that obtained streams are correct.

Table 3: Server validations for data structure

Data	*Expected Structure*	*Received Data Structure*
LoRa NR	*21KNR001VS010VH010/00.00&00.00#00.00@00%00*	*21KNR001VS010VH010/23.40&61.90#34.20@22%90*
NC Csv	*21KNROO1VS010VH010,00.00,00.00,00.00,00 ,NN,NN, 00, 00:00:00;*	*21KNCOO1VS010VH010,NN,NN10.00,NN,40.20,20.20,80,19:11:51*
NR Csv	*21KNCOO1VS010VH010,NN,NN,00.00,NN,00.00,00.00,0,00:00:00;*	*21KNROO1VS010VH010,23.40,61.90,34.20,22 ,NN,NN, 90, 19:11:51*

VI. FUZZY PATTERNS DERIVED

The next step for the prototype is to derive fuzzy patterns from the csv received in the server when the subsystem from MR to SIA module is fully integrated.

To do that every variable must be fuzzyfied. The process starts with the analysis of each one and determining the labels and membership functions. Table 4 shows a summary of the variables considered for current patterns. The variables in previous tests [3] were Soil Humidity (ST), Ambient Humidity (AH), Soil Temperature (ST), Canopy Temperature (CT), Water Sheet height WSL), Anemometer (AN), UV radiation (UVR), Global Radiation (GR), Pluviometer (PM). Variables for stage 1 are the same except for ST and PM. It is important to note that in previous tests patterns work with ground stations in a preliminary architecture, that is, without many functionalities like autotest, cross validation of data and a more extensive set of sensors. The results (see Figure 6) derive from Expectation Maximization and J48 Induction Tree.

Table 4: Pattern obtained with prototype stage-1

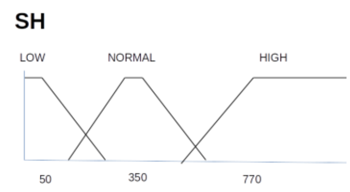
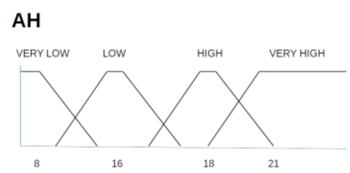
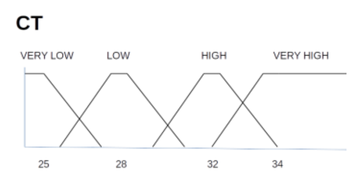
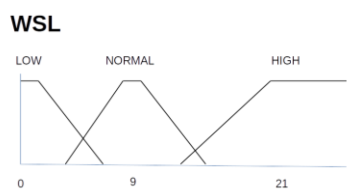
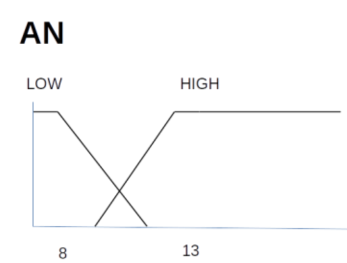
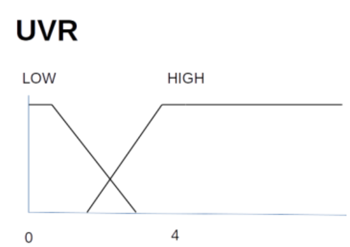
Variable	*Membership*
Soil Humidity (SH)	
Air Humidity (A)	
Canopy Temperature (CT)	
Water Sheet Height (WSL)	
Anemometer (AN)	
UV radiation (UVR)	

Table 5 shows a total of 9 patterns derived with a first arrangement of sensors. Some of them are no longer available for stage 1 (the ones with dark background).

Table 5: Data in NSA sampled with prototype stage-1

IF	*THEN*
AH is VERY LOW and CT is not (VERY LOW) and UVR is LOW and AN is HIGH and WSL is (NORMAL or HIGH)	*LOW_WATER_LEVEL*
AH is (VERY HIGH or HIGH) and CT is (VERY HIGH or HIGH) and WSL is HIGH	*LOW_WATER_LEVEL*

IF	THEN
AH is not (VERY LOW) and CT is VERY HIGH and AN is not(LOW) and WSL is HIGH	MEDIUM_WATER_LEVEL
AH is not (VERY LOW) and CT is not(VERY LOW) and WSL is HIGH and UVR is LOW	LOW_WATER_LEVEL
AH is not (VERY LOW) and CT is not (VERY LOW) and WSL is not(LOW) and ST is not (VERY LOW) and UVR is HIGH and AN is HIGH	LOW_WATER_LEVEL
AH is VERY HIGH and CT is not (VERY LOW) and WSL is not(LOW) and ST is not (VERY LOW) and AN is not (HIGH)	LOW_WATER_LEVEL
AH is VERY LOW and CT is VERY HIGH and WSL is not(LOW) and ST is not (VERY LOW)	LOW_WATER_LEVEL
AH is not(VERY LOW) and CT is not(VERY LOW) and ST is not (VERY LOW) and AN is HIGH	LOW_WATER_LEVEL
AH is VERY HIGH and ST is not (VERY LOW)	MEDIUM_WATER_LEVEL
AH is not(VERY LOW) and CT is not(VERY LOW) and ST is not (VERY LOW) and AN is HIGH	LOW_WATER_LEVEL
AH is VERY HIGH and ST is not (VERY LOW)	MEDIUM_WATER_LEVEL

Several considerations deserve attention. The first is the fact that AH is now denoted as **λ**. Behind these patterns runs a model with a 98.2% of correctly Classified Instances, Kappa statistic: 0.97 and Mean absolute error of 0.0165.

Those metrics confirm that these patterns have the critical variables for a good time prediction. The new prototype in stage 1, is replacing UVR by GR since it is the most common type of sensor in the field, and left aside ST because the model is intended to be complemented with radar and satellite maps with much better precision. Table 6 shows part of the edited data from the prototype.

Table 6: Data in NSA sampled with prototype stage-1

TR	SH	AH	CT	WSL	AN	GR	TST
21KNCOO1VS010VH010	NN	NN	23.4	NN	3.0	680.0	18.41.10
21KNCOO1VS010VH010	852.0	40.0	23.5	30.0	NN	NN	18.41.10
21KNCOO1VS010VH010	NN	NN	22.6	NN	3.0	660.0	19.11.51
21KNCOO1VS010VH010	774.0	41.0	22.6	30.0	NN	NN	19.11.51
21KNCOO1VS010VH010	NN	NN	22.4	NN	3.0	640.0	19.41.47
21KNCOO1VS010VH010	746.0	39.5	22.5	27.0	NN	NN	19.41.47
21KNCOO1VS010VH010	NN	NN	20.3	NN	3.0	600.0	20.11.30
21KNCOO1VS010VH010	736.0	39.30	20.4	27.0	NN	NN	20.11.30

The improved dataset shows 2 rows for every sampling set: one for NC and the second for NR. Several cells are NN since they correspond to sensors not implemented in the node. The testing for stage 2 will extend and add redundant data. The structure and rate are compatible with patterns in Table 5.

Regarding the validation of results, it is performed as any other approach in the field: the yield decrease when the crop suffers water stress. Also, the comparison of the information with radar and satellite is relevant for the precision.

VII. Conclusion

This paper presents the main characteristics of a prototype for time mining of rice crops watering requirements. The hardware is organized in a lightweight and portable set of nodes capable of sensing some of the initial parameters. This work describes the first stage testing, with a low sampling rate, and several changes from previous versions: reduced set of sensors, migration from global radiation to UV radiation sensors, a new balanced sensing with two types of nodes (NR and NC), and the perspective of adding radar/satellite information replacing certain soil information from direct sampling. The patterns derived in previous work are compatible in format with new structure, from the data tested. The hardware design is suitable for coordinating nodes with NSA

in the server, preserving connectivity and low powering consumption, resulting in data like the presented in Table 6. Also the patterns obtained are coherent with previous findings, as expected.
The next step will be to implement and test the prototype at stage 2, adding radar and satellite data that complete and validate weather and soil conditions. This is expected to improve the quality and complexity of patterns derived with time mining.

REFERENCES

[1] H. C. Strifors, A. Gustafsson, S. Abrahamson and G. C. Gaunaurd, "Fuzzy-cluster representation of time-frequency signatures as a means for automatic classification of buried mine-like targets," Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis (Cat. No.98TH8380), 1998, pp. 597-600, doi: 10.1109/TFSA.1998.721495.
[2] Imam Mukhlash, Desna Yuanda, Iqbal Mohammad. Mining Fuzzy Time Interval Periodic Patterns in Smart Home Data. International Journal of Electrical and Computer Engineering 8(5):3374-3385. 2018.
[3] Xiaoyue Zhou, Kaiqi Zou and Yanfang Wang, "Fuzzy variable time series based on fuzzy membership function and econometrics," 2010 Third International Symposium on Knowledge Acquisition and Modeling, 2010, pp. 225-228, doi: 10.1109/KAM.2010.5646254.
[4] I. Acuña Barrios, E. García, D. López De Luise, C. Paredes, A. Celayeta, M. Sandillú, W. Bel. Traffic & Pedestrian risk inference using Harmonic Systems. Soft computing applications. Springer. 2014.
[5] D. López De Luise, E. Ledesma, W. Bel, E.Velazquez. Analysis and Identification of Relevant Variables for Precision Farming using Harmonic Systems. FTNCT-2019. 2019.
[6] C. Chang, H. Chueh and N. P. Lin, "Sequential Patterns Mining with Fuzzy Time-Intervals," 2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery, 2009, pp. 165-169, doi: 10.1109/FSKD.2009.733.
[7] Guoqing Chen, Qiang Wei and Hong Zhang, "Discovering similar time-series patterns with fuzzy clustering and DTW methods," Proceedings Joint 9th IFSA World Congress and 20th NAFIPS International Conference (Cat. No. 01TH8569), 2001, pp. 2160-2164 vol.4, doi: 10.1109/NAFIPS.2001.944404.
[8] Y. Bai and T. Li, "Robust time series forecasting using fuzzy inference systems," 2012 24th Chinese Control and Decision Conference (CCDC), 2012, pp. 2703-2706, doi: 10.1109/CCDC.2012.6244430.
[9] C. Chang, H. Chueh and Y. Luo, "An integrated sequential patterns mining with fuzzy time-intervals," 2012 International Conference on Systems and Informatics (ICSAI2012), 2012, pp. 2294-2298, doi: 10.1109/ICSAI.2012.6223511.
[10] V. Novák and I. Perfilieva, "Time Series Mining by Fuzzy Natural Logic and F-Transform," 2015 48th Hawaii International Conference on System Sciences, 2015, pp. 1493-1502, doi: 10.1109/HICSS.2015.181.
[11] Wai-Ho Au and K. C. C. Chan, "Mining fuzzy rules for time series classification," 2004 IEEE International Conference on Fuzzy Systems (IEEE Cat. No.04CH37542), 2004, pp. 239-244 vol.1, doi: 10.1109/FUZZY.2004.1375726.
[12] P. Ravikumar and V. Susheela Devi, "Fuzzy classification of time series data," 2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2013, pp. 1-6, doi: 10.1109/FUZZ-IEEE.2013.6622571.
[13] Z. Li, F. Bu and F. Yu, "Temporal fuzzy association rules mining based on fuzzy information granulation," 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2017, pp. 1168-1174, doi: 10.1109/FSKD.2017.8392930.
[14] S. S. Liao, T. H. Tang and Wei-Yi Liu, "Finding relevant sequences in time series containing crisp, interval, and fuzzy interval data," in IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 34, no. 5, pp. 2071-2079, Oct. 2004, doi: 10.1109/TSMCB.2004.833597.
[15] D. Owuor, A. Laurent and J. Orero, "Mining Fuzzy-Temporal Gradual Patterns," 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2019, pp. 1-6, doi: 10.1109/FUZZ-IEEE.2019.8858883.
[16] Chun-Hao Chen, Tzung-Pei Hong, Vincent S. Tseng. Fuzzy data mining for time-series data. Applied Soft Computing. Volume 12, Issue 1, January 2012, Pages 536-542. 2012.
[17] Marlis Fulgueira Camilo,William Reyes Burunate. Fuzzy Miner algorithm parallelization with OpenCL. Rev cuba cienc informat vol.10 no.4 La Habana oct.-dic. 2016
[18] D. López De Luise, E. Ledesma, W. Bel, E. Velazquez, J. Pirchi. Minería Para Cultivos de Precisión de Arroz: Modelo y Parametrización del Riesgo. ARGENCON 2020. 2021.
[19] W. van der Hoek, R. Sakthivadivel, M. Renshaw, J. Silver, M. Birley, F. Konradsen. Alternate wet/dry irrigation in rice cultivation: A practical way to save water and control malaria and Japanese encephalitis? Research Report 47. Colombo, Sri Lanka: International Water Management Institute. 2001.
[20] T. Tuong, B. Bouman. Rice production in water-scarce environments. In: Kijne, J.W., Barker, R. and Molden, D., Ed., Water Productivity in Agriculture: Limits and Opportunities for Improvement, CABI Publishing, Wallingford, 53-67. 2003.
[21] O. Duarte, E. Diaz, L. Lenzi, R. Valenti, R. Benavidez. Water lost evaluation in rice crops irrigation systems (Evaluación de las pérdidas de agua en un sistema de riego de arroz). In: Rice: growing and Entre Ríos sustainability (El arroz: su cultivo y sustentabilidad en Entre Ríos). National University of Entre Ríos. 2006.
[22] F. Esfandiary. Wheat Yield Prediction through Agro Meteorological Indices for Ardebil District. World Academy of Science, Engineering and Technology. 2009.
[23] M. McCandless, H. Venema. Applying the Water Soft Path Concept for Agro-ecosystem Adaptation in Prairie Canada. International Conference on Adaptive and Integrated Water Management. Basel, Switzerland. 2007.
[24] W. SchlenkeR, W. Michael Hanemann, C. FISHER. Will U.S. Agriculture Really Benefit from Global Warming? Accounting for Irrigation in the Hedonic Approach. The American Economic Review, pp. 395 -406. 2005. [8] ref13=13 (Esfandiary, 2009)
[25] H. Jaafar, F. Ahmad. Relationships Between Primary Production and Crop Yields in Semi-Arid and Arid Irrigated Agro-Ecosystems. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-7/W3, 36th International Symposium on Remote Sensing of Environment, pp. 11–15, Berlin, Germany. 2015.
[26] E. Aghdam, H. Babazadeh, M. Vazifedoust, F. Kaveh. Regional Modeling of Wheat yield production Using The Distributed Agro-hydrological SWAP. Advances in Environmental Biology, 7(1): 86-93, ISSN 1995-0756. pp 86-93. 2013.
[27] S. Nagini, T. V. R. Kanth and B. V. Kiranmayee, "Agriculture yield prediction using predictive analytic techniques," 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), pp. 783-788, DOI: 10.1109/IC3I.2016.7918789. 2016.
[28] R. Carrasco, I. Soto, F. Seguel, L. Osorio-Valenzuela, C. Lagos, C. Flores. Water balance analysis in plantations of strawberries, in the commune of San Pedro. 2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON), pp. 1-5, DOI: 10.1109/CHILECON. 2017.
[29] M. Jihua, L. Zhongyuan, W. Bingfang, X. Jin. Design, development and application of a satellite-based field monitoring system to support precision farming. The Third International Conference on Agro-Geoinformatics, pp. 1-9, DOI: 10.1109/Agro-Geoinformatics. 2014.
[30] FAO. Making the best use of water for agriculture. FAO AQUASTAT Report. 2002.
[31] A. Fazel, M. Saraee, P. Shamsinejad. Mining time series data: Case of predicting consumption patterns in steel industry. The 2nd International Conference on Software Engineering and Data Mining. 2010.
[32] J. CryerKung, S. Chan. Time Series Analysis. Springer. 2008.

[33] Y. FangSang. A review on the applications of wavelet transform in hydrology time series analysis. Atmospheric Research. Elsevier. Volume 122, Pages 8-15. 2013.
[34] D. López De Luise. PageInternational Journal of Modern Engineering Research (IJMER). Vol. 3, Issue. 6. pp-2719-2727, ISSN: 2249-6645. 2013.
[35] NASA. Prediction Of Worldwide Energy Resources: https://power.larc.nasa.gov/. Taken on Nov. 16th 2021.