

**Open Innovation Platform for IoT-Big data**

**in Sub-Sahara Africa**

# Introduction

University of Gaston Berger has performed an experiment about sensor-based irrigation, lasting from February to end of May 2019. The objective of the experiment is to have concluding elements in the monitoring of soil water content using soil moisture sensors. The experiment is held by experts and farmers in the UGB farm. A lot of data is being collected from soil moisture sensors and from a weather station. This data is published as open data on our Cloud platform. We took the opportunity to launch a “data competition” over the open data, with our partner Zindi.

Zindi is the first data science competition platform in Africa. Zindi hosts an entire data science ecosystem of scientists, engineers, academics, companies, NGOs, governments and institutions focused on solving Africa’s most pressing problems. Zindi works with companies, non-profit organizations, and government institutions to develop, curate, and prepare data-driven challenges. Solutions are ranked automatically by the accuracy achieved. Zindi is a place to access African datasets and solve African problems. Data scientists can find all the tools they need on Zindi to compete and share their ideas. Zindi is a team of data scientists and creators based in Cape Town, Johannesburg, and Accra committed to a better Africa.

The challenge set up with Zindi is named "Soil moisture prediction with low-cost DYI Internet of Things in Africa". In the face of climate change, it is very important for the agricultural sector in Africa to adapt. Sensor-based irrigation provides a solution to manage water more efficiently. However, there is a lack of algorithms that can predict efficiently the soil humidity and provide recommendations for irrigation schedules, using data from unreliable hardware. Current algorithms need a lot of data for proper training, which is difficult to obtain in rural Africa: many problems arise such as accessibility, battery problem, Internet problem, humidity/heat problem…

A part of the solution is to design algorithms that are resilient and can be trained with incomplete data (e.g. missing data points) and unclean data (e.g. lot of outliers). This is the objective of this competition. In the following sections, we summarize the context of the experiment, present the data collected, provide the elements of the competition.

## Experiment context

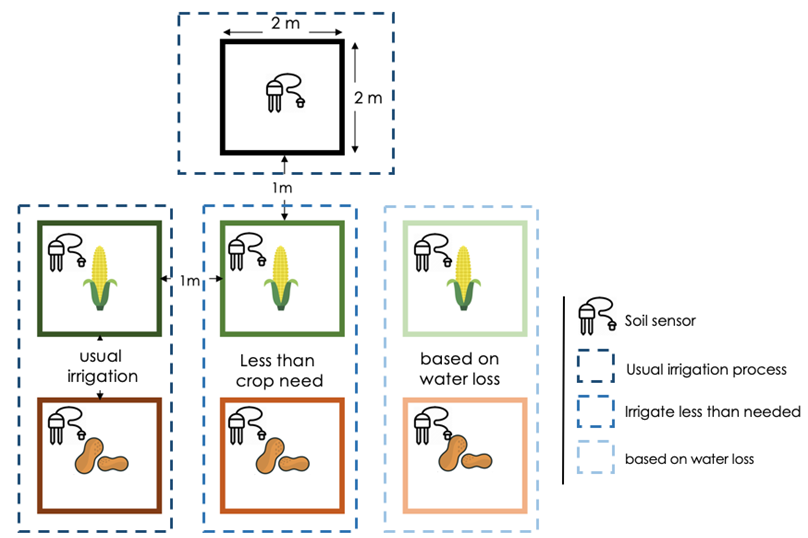


Figure 25 Map of the experiment

As shown in Figure 11, the experiment is divided into 7 plots of 2X2 square meters each. Nothing will be sown in the first plot. This one will serve for comparing the level of water usage with the other plots, so as to be able to estimate moisture loss without crops water usage. The next six plots will be spread over 2 kinds of crops: maize and peanuts. In the second range of plots, the same amount of maize is sown in each of the 3 plots. The same for the last range of plots where peanuts will be sown. Plots are between each other separated by 1 meter.

The irrigation process is done every 2 days, according to some earlier results provided by the Faculty of Agronomy in UGB. Indeed, it has been found that bringing water every 2 days was more efficient in terms of yields. However, the amount of water brought in each one of the plots will not be the same for the purpose of the experiment. Considering the 3 plots for maize as an example, the idea is to irrigate the first one every 2 days, with the normal irrigation way, i.e. bringing water every 2 days according to usual farmer practices. In the second plot, a reduced amount of water is provided compared to the previous plot. In the third plot, an amount based on the different parameters is provided. Parameters such as evapotranspiration, soil moisture level, are combined to estimate water loss, before bringing the required amount to the plot.

The data collected from the sensors is sent to the WAZIUP gateways and then to the WAZIUP Cloud platform. Two gateways are used in the experiment to ensure data reliability. The different parameters are listed below:

* SH: Soil Humidity,
* PA: Atmospheric Pressure,
* TP: Temperature,
* WD: Wind Direction,
* WS: Wind Speed,
* WG: Wind Gust,
* HD: Ambient Humidity,
* RA: Rain Amount.

The first parameter is obtained from the soil moisture sensors, while the next is provided by the weather station daily. This latter is deployed in the farm near the pilot field, and data is collected by the gateway every 5 minutes. The use of all these sensors help determine the requirements in water, that will be added to maize and peanut plots.

## Data collected

Two kinds of data are collected from the experiments: sensor data and context data. Context data is collected by hand in an Excel file. It consists of:

* Min temperature min (°C) jour j-1
* Max temperature (°C) jour j-1
* Relative humidity (%) jour j-1
* Wind speed (m/s) jour j-1
* Sun
* Solar Radiation (Mj/jour) jour j-1
* Coefficient cultural (Kc) jour j-1
* Evapotranspiration reference (ETo) jour j-1
* Evapotranspiration measured (Etc) jour j-1
* Rain amount per day j-1
* Efficient Rain (Pe) jour j-1
* Water need 100% BE / 1j
* Water need 100% BE / 2j
* Water need 100% BE / 3j
* T0

For example, on the day 02/27/2019, the following values were collected:

|  |  |
| --- | --- |
| Min temperature (°C) jour j-1 | 17.3 |
| Max temperature (°C) jour j-1 | 33.7 |
| Relative humidity (%) jour j-1 | 13 |
| Wind speed (m/s) jour j-1 | 6.0 |
| Sun | 9.9 |
| Solar radiation (Mj/jour) jour j-1 | 21.9 |
| Coefficient cultural (Kc) jour j-1 | 0.80 |
| Evapotranspiration reference (ETo) jour j-1 | 11.01 |
| Evapotranspiration measured (Etc) jour j-1 | 8.81 |
| Rain amount per day j-1 | 0 |
| Efficient rain (Pe) jour j-1 | 0 |
| Water need 100% BE / 1j | 35.232 |
| Water need 100% BE / 2j | 82.144 |
| Water need 100% BE / 3j | 0.000 |
| T0 | 82.144 |

Examples of data collected by the sensors are presented below.

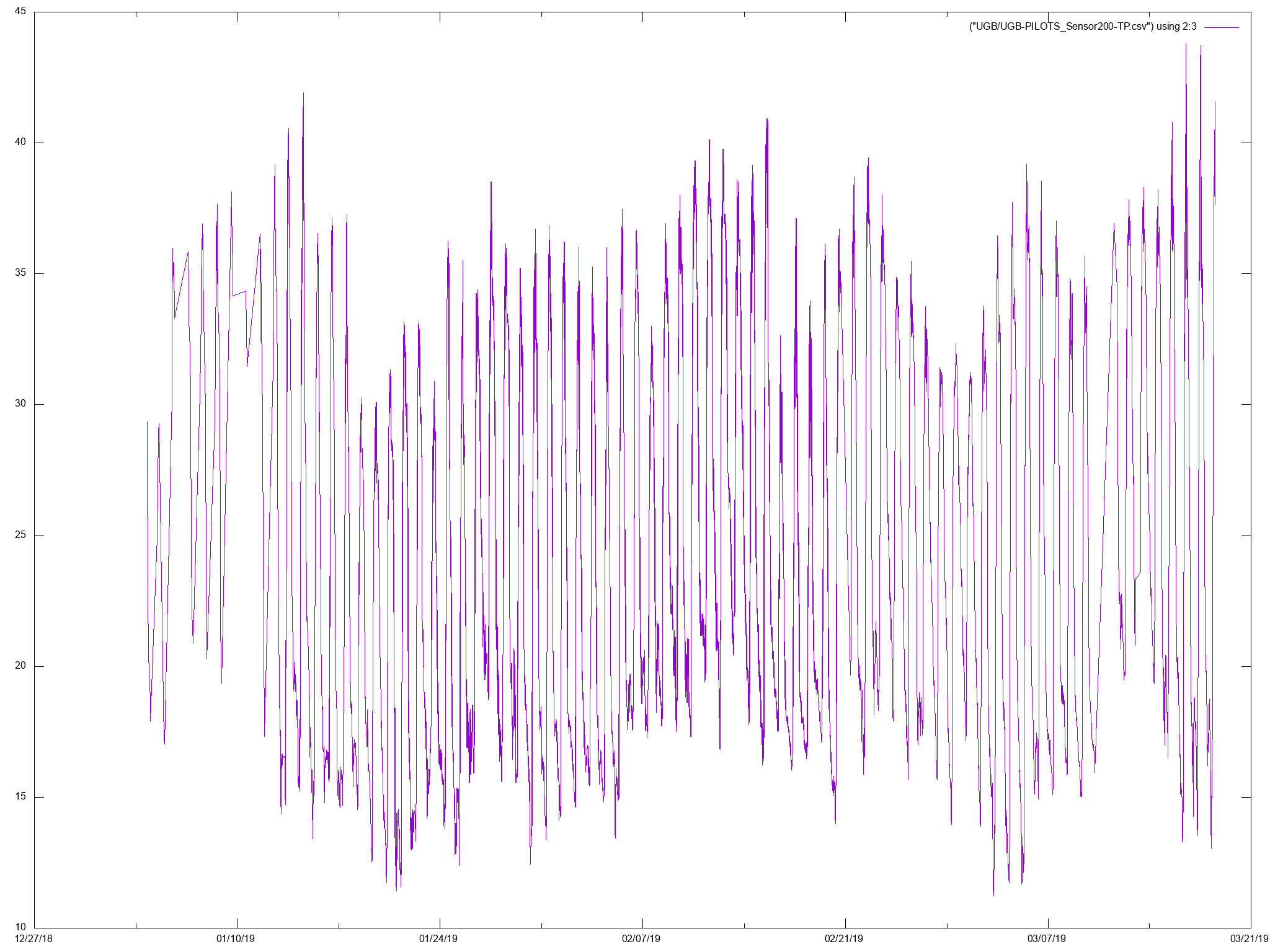


Figure 26 Temperature as measured by the weather station during the experiment

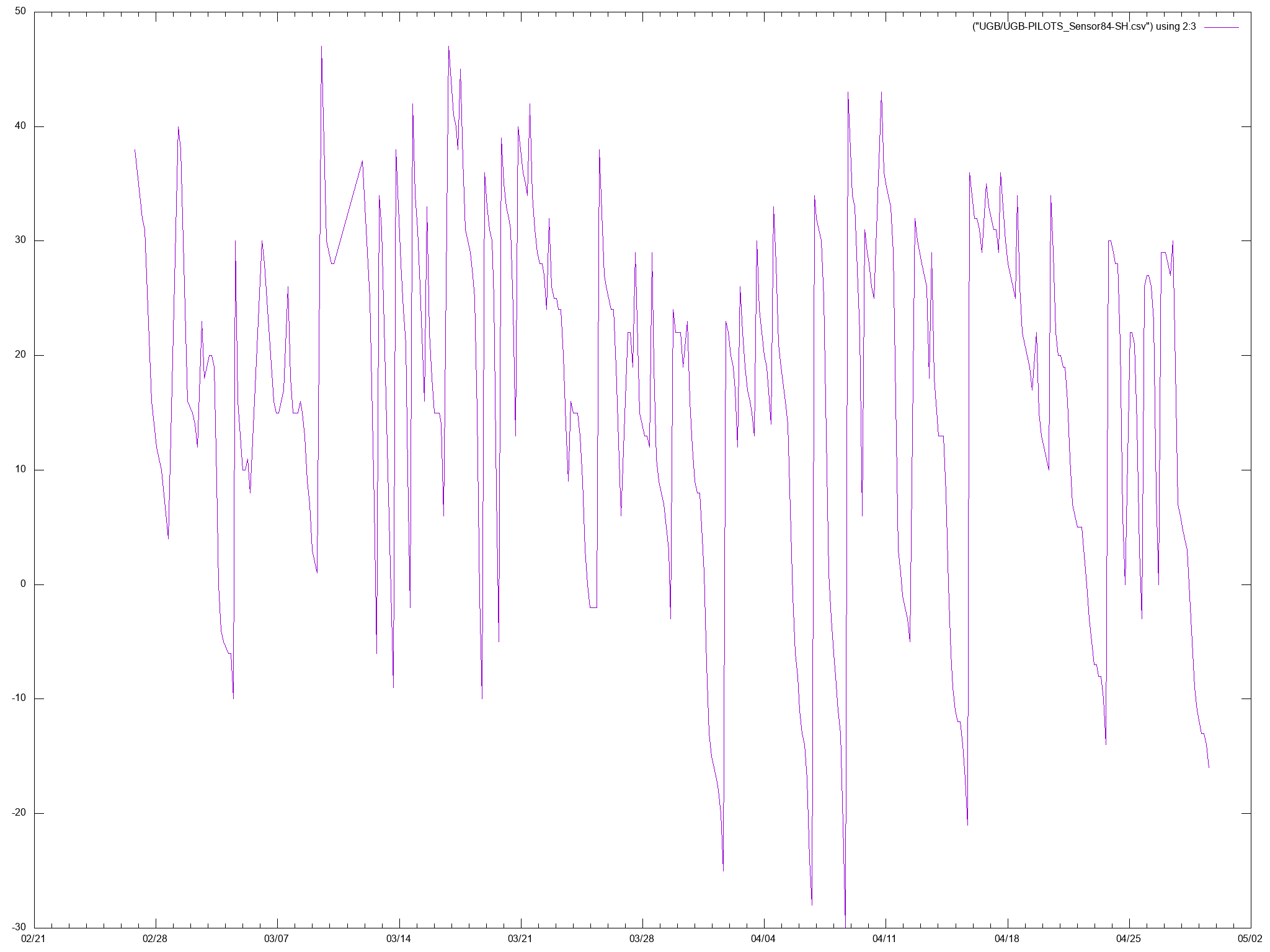
****

Figure 27 Soil humidity as measured by Sensor 84 (uncalibrated)

The Figure 26 shows the temperature collected by the weather station, while Figure 27 gives the humidity measured by one of the height soil sensors. For this experiment, around 100.000 sensor data points has been recorded on our platform (weather station sensors + soil moisture sensors).

## Big data competition

Using the data mentioned above, we are launching the Zindi competition. The objective of this competitionis to create a machine learning model capable of predicting the humidity for a particular plot in the next few days, using data from the past. This kind of algorithm is very useful for farmers in order to prepare their irrigation schedules.

The data will be separated into two sets: the training set and the test set. The participants will be provided only with the training set. The participants may, if they desire, train their model with additional data coming from other datasets, such as the Iowa State University dataset[[1]](#footnote-1). However, the algorithms should compete using the Waziup data.

Once the participants have created and trained their algorithms, they will have to run them to predict the humidity for the next 5 days. The result will be uploaded on the Zindi website. Based on this data, Zindi will rank the participants. A second round can then start: Zindi will provide additional data to the participants covering more time, and the participants will have the opportunity to re-train and fine tune their algorithms. A leader board will be created, in order to show the ranking of the participants from round to round.

The results and algorithms will be released as public good, i.e. the winners will be encouraged to share their code on GitHub. Cash prizes for this competition will be provided by Microsoft.

1. [↑](#footnote-ref-1)