

Artificial Intelligence GHG Monitoring for a Voluntary Carbon Certification

Doimi Mauro¹ and Minetto Giorgio²

1. Department of Electronic Biology, and Manufacturing Engineering, D&D Consulting, Venice, Mestre 30171, EU-Italy

2. Department of Biology, D&D Consulting, Venice, Mestre 30171, EU-Italy

Abstract: Generating carbon credits in rural and wetland lagoon environments is important for the economic and social survival of the same. There are many methodologies to study and certificate the Carbon Sink such as the ISO 14064, VCS VERRA, UNI-BNEUTRAL, GOLD STANDARD and others. Many methods done before 2018 are obsolete since research has developed greatly in recent years. The methods are all different, but they share a continuous and real monitoring of the environment to ensure a true CCS (Carbon Capture and Storage) action. In the case of absence of monitoring, the method uses a system of provision of carbon credits called “buffer”. This system allows maintaining a credit-generating activity even in the presence of important anomalies due to adverse weather events. This research shows the complex analytic web of the different sensors in a continuous environmental monitoring system via GSM (Global System for Mobile) Communication and IoT (Internet of Things). By 2011, a monitoring network was installed in the wetland environments of Northern Italy Venetian Lagoon (UNESCO heritage) and used to understand and validate, the CCS action. Thingspeak cloud platform is used to collect data and is used to send alert to the user if the biological sink is reversed to emission. The obtained large dataset was used to prepare a AI (Artificial Intelligence) model “CCS wetland forecast” by Google COLAB. This model can fit the trend to avoid the direct and spot chemical field analysis and demonstrate the real efficacy of the model chosen. This network is now implemented by the Italian national method UNI PdR 99:2021 BNeutral generation of carbon credits.

Key words: AI model, data logger, IoT, CCS, CO₂, UNI BNeutral, VERRA VCS, wetland.

1. Introduction

The significant degree of air-contamination in metropolitan regions, caused by the toxic gas emissions GHG (Greenhouse Gas) into atmosphere, gives, as consequence, a problematic climate change. Many solutions have been studied and one of these is the CCS (Carbon Capture and Storage) in a special rural wetland area. This action can be certified for a carbon credit generation. Recently Doimi [1] has described how the HCWs (Human Controlled Wetlands) inside the Venetian Lagoon, is a very promising and important area with an important CCS activity against the climate change.

Currently, the certification authority has established the importance of a monitoring air and water plan inside the CCS area to obtain information that allows the removal

(or neutralization) of the GHG with no harmful impacts on the environment and people. The UNI-BNeutral method [2] gives us a new approach with new wetland carbon pools to be included in a carbon sequestration and credit generation as described in Tab 1. These wetland carbon pools need an analytic approach to verify the activity. In Tab 2 the main parameters to be checked are summarized. These parameters allow us to prepare a Carbon Sink Index that is the first step towards a correct management of the wetland. We can integrate the quality of the chemical analytical dataset by monitoring the environment using specific electronic sensors. The obtained data are then transmitted to the cloud using technologies like IoT and Wireless Sensor Network. Actually, many IoT data loggers are in activity inside 16 different HCWs (“Valle da Pesca”,

Corresponding author: Doimi M., Ph.D., research fields: marine biology and electronics. E-mail: maurodoimi@ded-consulting.eu.

Table 1 Wetland carbon pool used by the PdR UNI 99:2021.

Algal biomass	Obligatory
Aquatic Plant biomass	Obligatory
Lagoon seabed	Obligatory
Biomass of bivalve molluscs	Obligatory

Table 2 Wetland parameter and range used by the data loggers in the Venetian Lagoon (North Italy).

Sector	Parameter	Lower	High
Air	CO ₂	350 ppm	700 ppm
	CH ₄	1,900 ppb	1,950 ppb
	Temp	0 °C	40 °C
	PAR (Photosynthesis Active Radiation)	10 lux	80 lux
Water	CO ₂	100 ppm	4,000 ppm
	TDS (Total Dissolved Solid)	500 ppm	800 ppm
	TC	100 ppm	800 ppm
	CHL	0 ppb	10,000 ppb

Italian language) and this allows us to create the biggest IoT network in Europe to measure and reveal the Carbon Sink activity in controlled wetlands and lagoons.

Different methods and sensors have been tested since 2010. Several are the main control guidelines like the CO₂, methane (CH₄) and VOC (Volatile Organic Compounds) in air and the analysis of the PAR (Photosynthetic Activity) of CHL (Chlorophyll), TDS (Conductivity/Dissolved Solids), TC (Total Carbon) and CO₂ in water.

The use of data loggers increases the attention of the CCS process inside the wetland. Our selected prototype consists of an embedded microcontroller and sensors to analyze environment, communicate to IoT cloud through GSM, and inform the certification authority of the correct process CCS status [3].

2. Material and Methods

D&D Consulting SAS, has installed by 2013, 20 data loggers in 16 HCWs called “Valli Da Pesca”. Each data logger consists of an Arduino MKR GSM 1400 (Fig. 1). It easily connects devices, visualizes data, controls and shares the data from anywhere in the world using the GSM/3G network that covers the highest percentage of the Italian wetland. This choice is due to the lack of Wi-

Fi communication within brackish water pools. The board's main processor is a low power Arm® Cortex®-M0 32-bit SAMD21. The GSM/3G connectivity is performed with a module from u-blox, the SARA-U201, a low power chipset operating in the different bands of the cellular range (GSM 850 MHz, E-GSM 1,900 MHz, DCS 1,800 MHz, PCS 1,900 MHz).

The daily data are transmitted via the GPRS (General Packet Radio Service) to a ThingSpeak cloud system (www.thingspeak.com). Each board receives the environmental information from sensors selected for affordability, speed of analysis and resistance to salt environments. All these devices, before use, have been tropicalized with a special spray to make them less sensitive to marine weather.

- Embedded wetland system:
 - A. Arduino MKR GSM
- Sensors:
 - B. 2X Sensor CO₂ MHZ-19
 - C. TDS/Conducibility home made
 - D. Sensor triad spectroscopy
- Components required: software
 - E. Arduino IDE
- Cloud system & machine learning model:
 - F. Thingspeak

The sensors communicate with the motherboard via 3.3 V I2C serial except TDS and CO₂ which have a direct analog signal on the appropriate Arduino input PIN.

The gaseous CO₂ sensor is used double, one for the analyses of the atmospheric CO₂ and the second to check the CO₂ released by the water in a closed environment PVC tube (Fig. 2).

**Fig. 1 Arduino board MKR GSM 1400.**

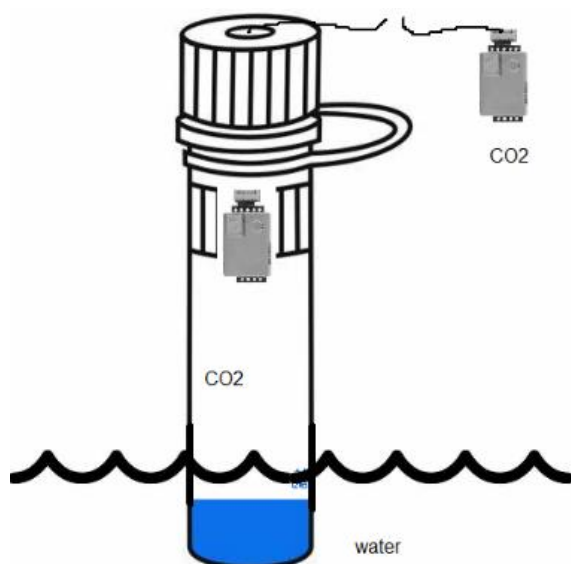


Fig. 2 Layout use of the two CO₂ sensors.

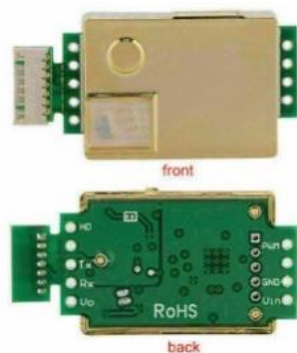


Fig. 3 The CO₂ sensor MHZ19.

MH-Z19CNDIR infrared gas module (<http://www.winsen-sensor.com/>) (Fig. 3) is a common type, small size sensor, using NDIR (Non-Dispersive Infrared) principle to detect the existence of CO₂ in the air; it has a good selectivity, it is non-oxygen dependent

and has a long life, built-in temperature compensation; and it has UART (Universal Asynchronous Receiver-Transmitter) output and PWM (Pulse With Modulation) output. It is developed by the tight integration of mature infrared absorbing gas detection technology, precision optical circuit design and superior circuit design.

The data difference (CO₂ in air - CO₂ in water), gives us the potential of the HCWs brackish water to sink the CO₂.

The conductivity/TDS is analyzed in water using a sensor (Fig. 4). The probe is a maker solution where:
GND -----C B-----A---- [R1] ----- +5V

A = analog in of Arduino

B, C = connectors immersed in water

R1 = 10 K resistance

Water between B & C will have a certain conductivity or resistance. Together with R1 it forms a voltage divider.

A SparkFun (<https://www.sparkfun.com>) triad spectroscopy sensor (Fig. 5) enables reflectance analysis to get the PAR, CHL, TDC and temperature.



Fig. 4 The maker TDS (conductivity) sensor.

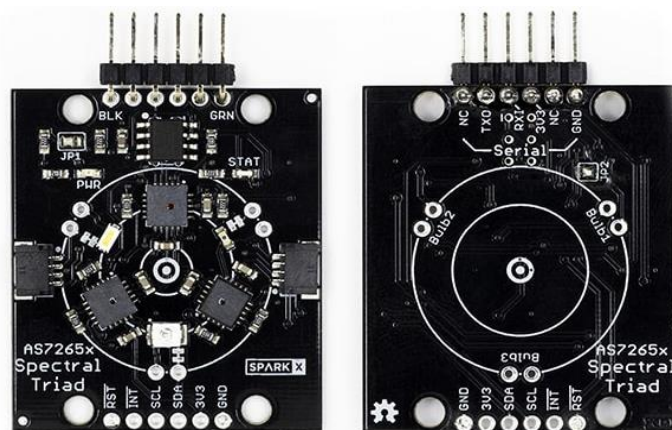


Fig. 5 The triad spectroscopy sensor.

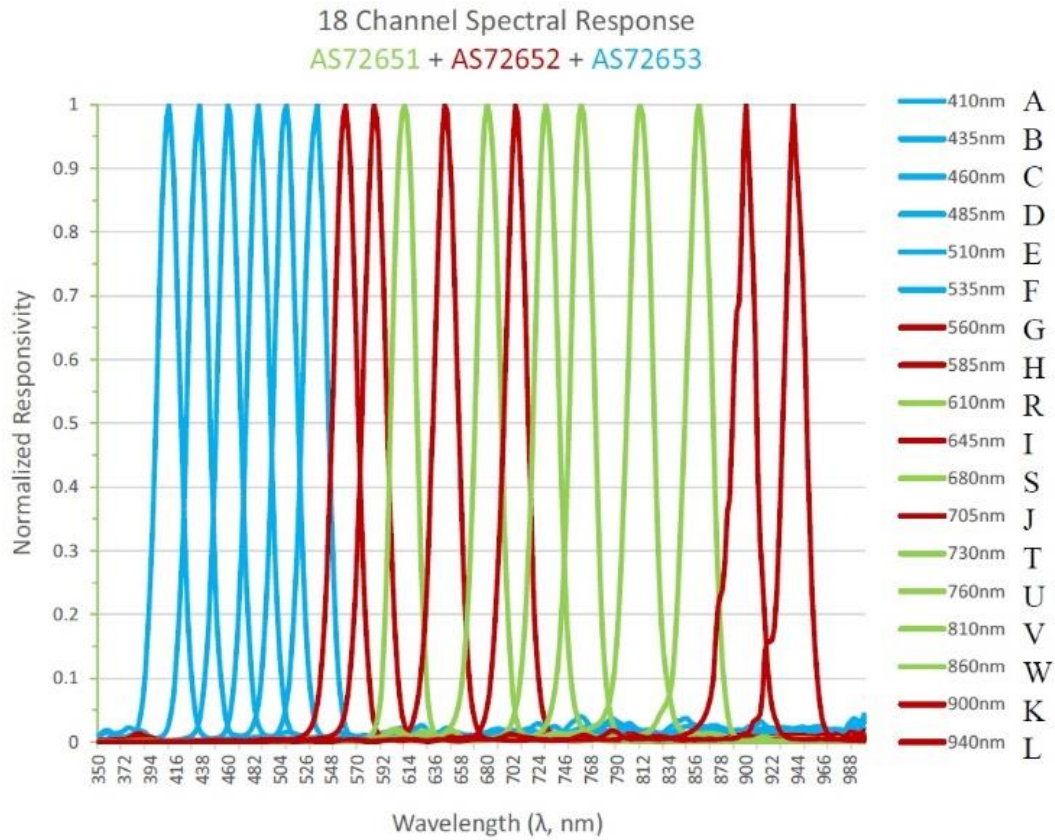


Fig. 6 The triad spectroscopy spectral response.

Three AS7265x spectral sensors are combined alongside visible, UV (Ultraviolet), and IR (Infrared) LEDs (Light-emitting Diode) to illuminate and test various surfaces for light spectroscopy. The Triad is made up of three sensors: the AS72651, the AS72652, and the AS72653 and can detect the light from 410 nm (UV) to 940 nm (IR). In addition, 18 individual light frequencies can be measured with precision down to 28.6 nW/cm^2 and accuracy of $\pm 12\%$ (Fig. 6).

The microprocessor is placed inside a waterproof case and powered at 5 V by a solar panel and battery (Fig. 7).

As a result, a complete CCS wetland monitoring system (Figs. 8-9) based on on-board IoT was successfully built to collect data from many distinct wetland areas (HCWs).

The CO_2 sensor output signal is analogic and linear to the gas presence after a preheat time of $3 \text{ min} \pm (50 \text{ ppm} + 5\% \text{ reading})$.



Fig. 7 The complete assembled system.



Fig. 8 Data logger located in a HCW ("Valle da Pesca").

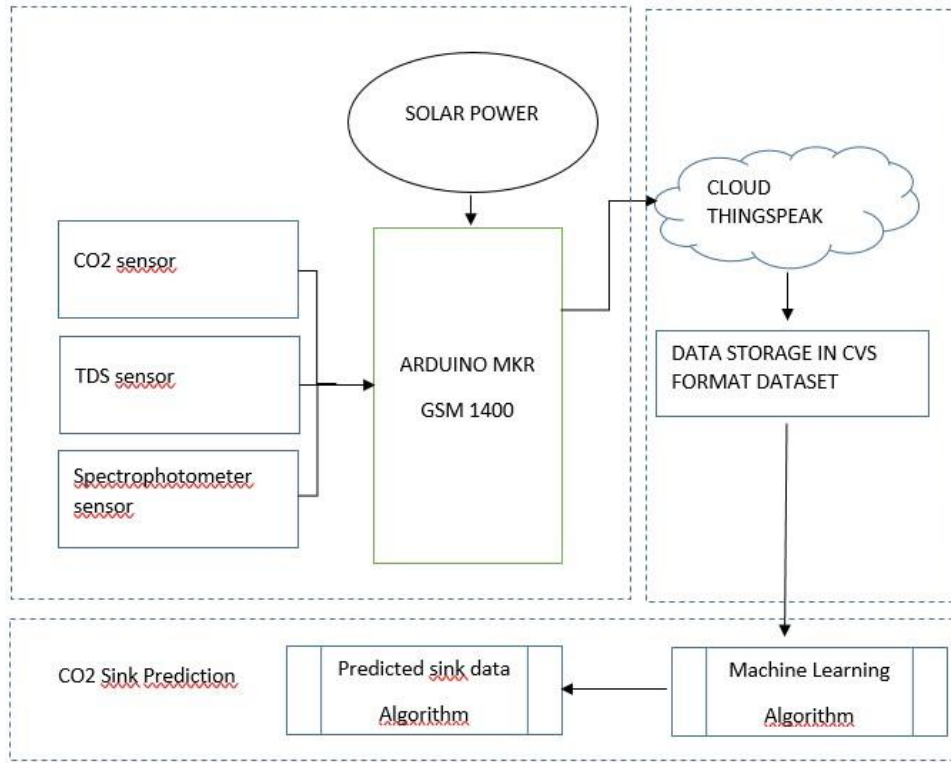


Fig. 9 Project layout.

The conductivity is converted in TDS by the conversion factor of 0.67 and indicates how many milligrams of soluble solids are dissolved in 1 L of water (PPM). In general, the higher the TDS value, the more soluble solids are dissolved in water, and the less clean the water is. Therefore, the TDS value can be used as one of the references to show the cleanliness of water [4]. The more transparent the water and rich in salts, the greater the photosynthetic activity of the algae and therefore the CO₂ sink activity is enhanced [5]. Output voltage is linear to the TDS: 0~2.3 V and the measurement range: 0~1,000 ppm with accuracy: $\pm 10\%$ (25 °C).

PAR is related to photosynthesis.

Green leaves absorb a great deal of the light at red and blue wavelengths [6]. We use the AS7265x spectral sensor as Nathan Seidle, SparkFun Electronics (2018) License MIT to measure the sum of the red light (645 nm) and the blue (460 nm).

$$\text{PAR} = ((\text{sensor.getCalibratedI}()) + (\text{sensor.getCalibratedC}())) / 2 \quad (1)$$

Eq. (1) is the use of the spectro sensor to determine the PAR value.

TC (Total Carbon) is the sum of the POC (Particulate Organic Carbon) and the PIC (Particulate Inorganic Carbon).

We use the NASA OCTS, MODIS-Aqua and Terra, MERIS, SeaWiFS, VIIRS. With satellite algorithm that returns the concentration of POC in mg/m^3 , it uses an empirical measurement of blue-to-green band ratios of sensing reflectance (R_{rs}).

The support for this algorithm is contingent on the availability of bands centered at 443 nm in the blue region and between 547 and 565 nm in the green region [7].

$$\text{POC} = 203.2 \times (443 \text{ nm} / 555 \text{ nm})^{-1.304}$$

Since the $R_{rs}(555)$ and $R_{rs}(443)$ is not available in the sensor an equivalence is estimated from the closest as follow:

$$\text{POC} = 203.2 \times (435 \text{ nm} / 560 \text{ nm})^{-1.304}$$

$$\text{POC} = 203.2 \times ((\text{sensor.getCalibratedB}()) / (\text{sensor.getCalibratedG}()))^{-1.304} \quad (2)$$

Eq. (2) is the use of the spectra sensor to determine the POC value.

In order to estimate the PIC (Particulate Inorganic Carbon) concentrations from the brackish water we use an algorithm presented by Mitchell et al. [8].

$$PIC = 555 \text{ nm} - [443 \text{ nm} + ((555 \text{ nm} - 443 \text{ nm} / 670 \text{ nm} - 443 \text{ nm})) \times (670 \text{ nm} - 443 \text{ nm})]$$

Since some nanometers (nm) are not available by the sensor an equivalence is estimated from the closest one, as follows:

$$PIC = 560 \text{ nm} - [435 \text{ nm} + ((560 \text{ nm} - 435 \text{ nm} / 680 \text{ nm} - 435 \text{ nm})) \times (680 \text{ nm} - 435 \text{ nm})]$$

$$PIC = ((\text{sensor.getCalibratedG()}) - ((\text{sensor.getCalibratedB()}) + ((\text{sensor.getCalibratedG()}) - (\text{sensor.getCalibratedB()}) / ((\text{sensor.getCalibratedS()}) - (\text{sensor.getCalibratedB()})) \times ((\text{sensor.getCalibratedS()}) - (\text{sensor.getCalibratedB()})))$$
 (3)

Eq. (3) is the use of the spectra sensor to determine the PIC value.

The TC is calculated by the sum POC+PIC.

As regards the CHL water concentration, we use the results obtained by the MODIS/Aqua NIR-SWIR ocean color products developed by the NOAA/NESDIS Center for STAR (Satellite Applications and Research). The products give a relationship between the CHL concentration and the remote sensing reflectance at 667 nm [9].

$$CHL = 667 \text{ nm};$$

Because the Rrs (667) is not available in our sensor,

an equivalence is estimated from the closest as follow:

$$CHL = (680 \text{ nm} + 645 \text{ nm})/2;$$

$$CHL = [(\text{sensor.getCalibratedS()}) + (\text{sensor.getCalibratedI()})]/2$$
 (4)

Eq. (4) is the use of the spectra sensor to determine the CHL value.

CHL concentration provides an estimate of the live phytoplankton biomass in the surface layer of the water lagoon.

The data obtained by the sensors, transferred to the cloud via the Arduino MKR 1400 GSM, are displayed in Thingspeak graphs (Fig. 10).

The UNI PdR 99:2021 indicates that the GPP (Gross Primary Production) is a parameter for a correct study of carbon storage in wetlands.

GPP data were obtained online by the ORNL DAAC MODIS/VIIRS system, which uses the MODIS AQUA satellite using the MODIS/VIIRS Land Products Global Subsetting Tool [10]. The big data obtained was used by AI (Artificial Intelligence) to do a model.

As first step, we balance the dataset.

We have selected the 2021-year data and have added all the information in a single dataset. After we filled in any missing data with data augmentation using an algorithm SMOTE (Synthetic Minority Over-Sampling Technique) [11]. SMOTE is an oversampling technique that creates new samples instead of duplicating under-represented instances standardization and brings the input features into the same numerical range. Following the data augmentation, we have normalized



Fig. 10 The cloud Thingspeak chart.

3. Results

Generally, malfunctions can be summarized in salt corrosion of electronic components in contact with water, battery buffer exhausted, solar panel not working, attacks due to wild animals (Fig. 13).

As just described by Doimi [1], there is a correlation between seasonality and the selected parameters (Figs. 14-20).



Fig. 11 North Italy HCW localization.

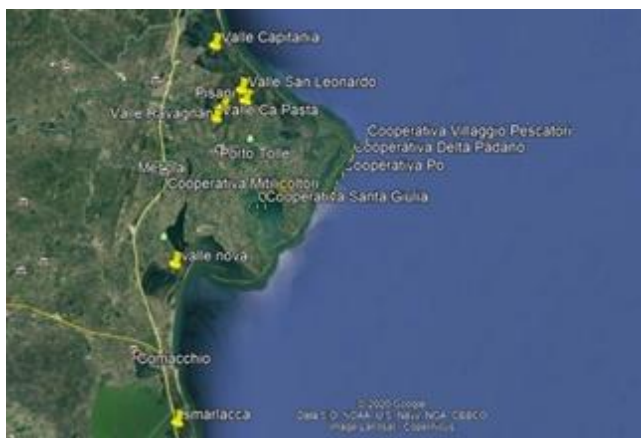
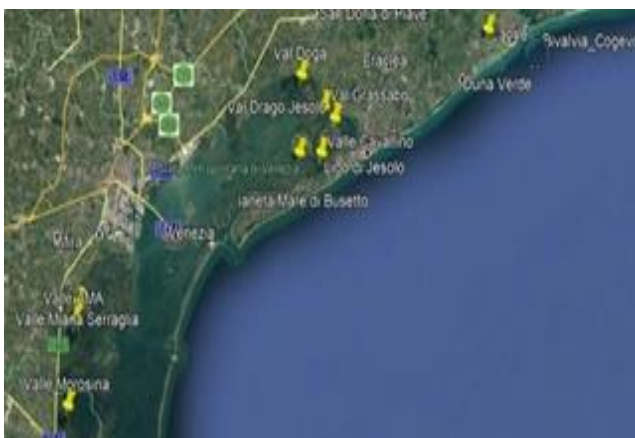


Fig. 12 The data logger localization: yellow points are where the units are active.

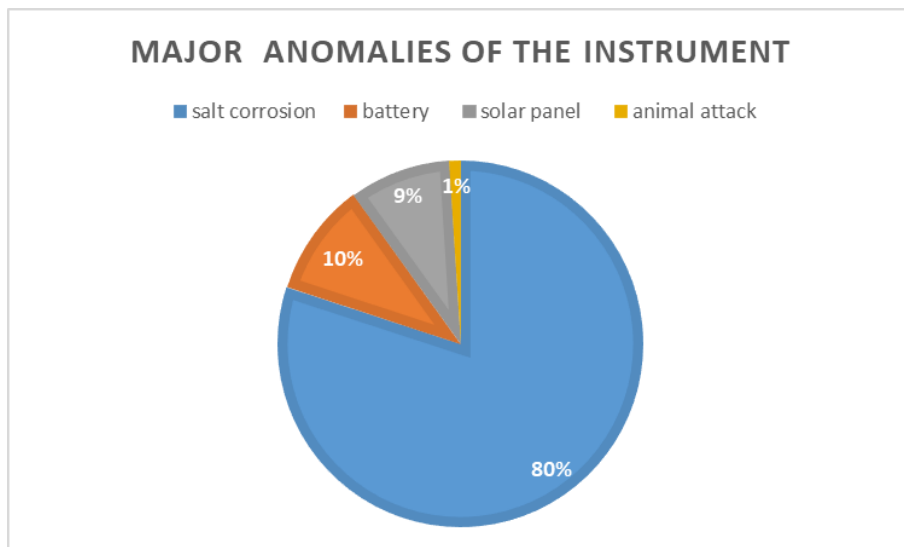


Fig. 13 Main anomalies of the data loggers working in wetland.

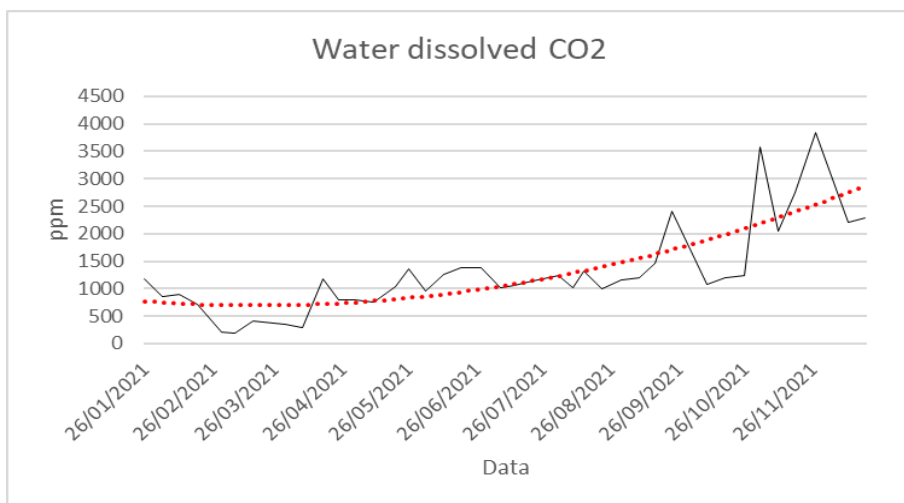


Fig. 14 Yearly dissolved CO₂, it increase during the autumn.

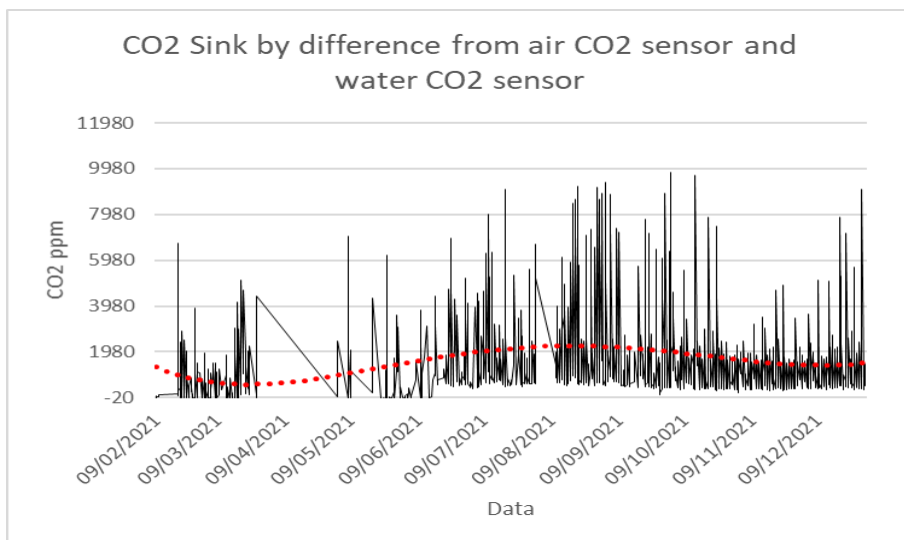


Fig. 15 CO₂ sink data obtained by subtraction of the atmospheric CO₂ by the Water CO₂; it increase during the summer.

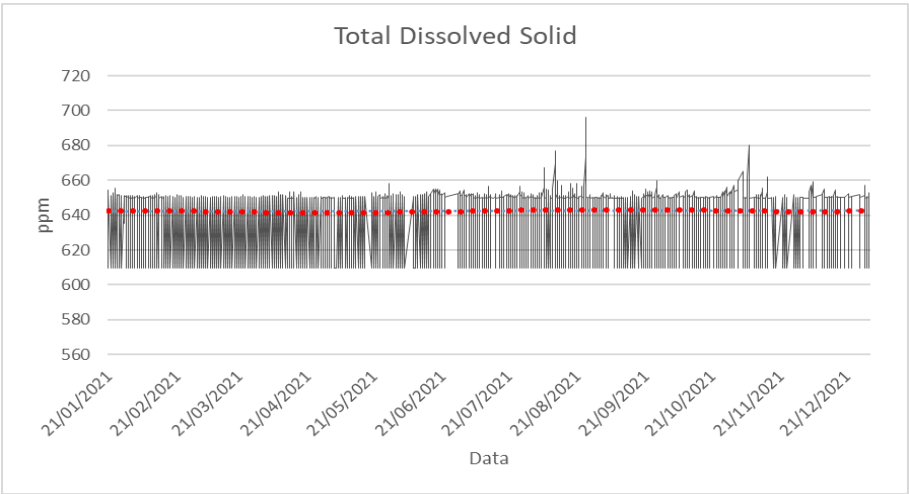


Fig. 16 TDS result analysis; it remains mainly stable but can be different in each wetland.

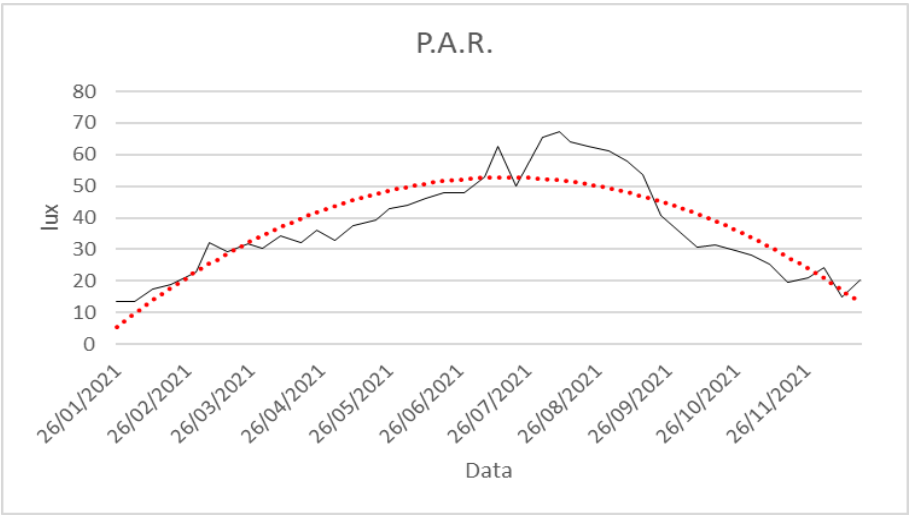


Fig. 17 PAR result analysis; summer higher activity.

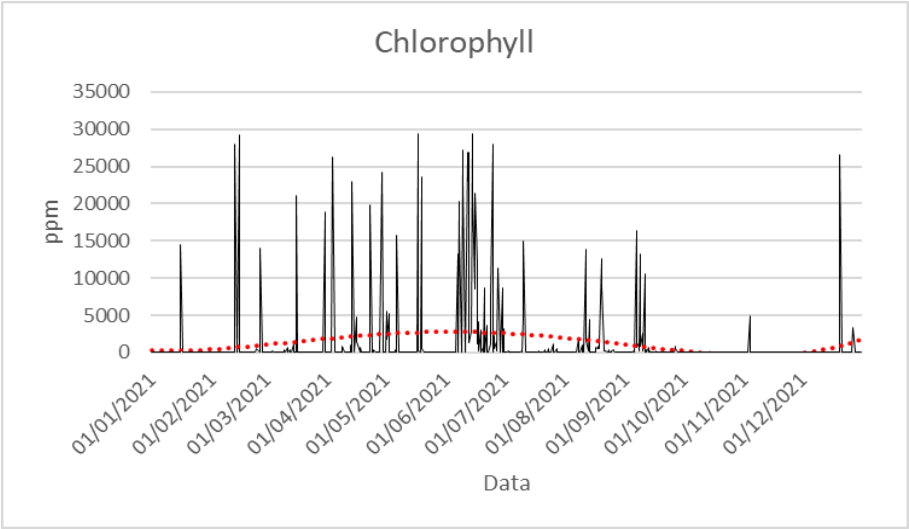


Fig. 18 Water dissolved CHL; summer higher activity.

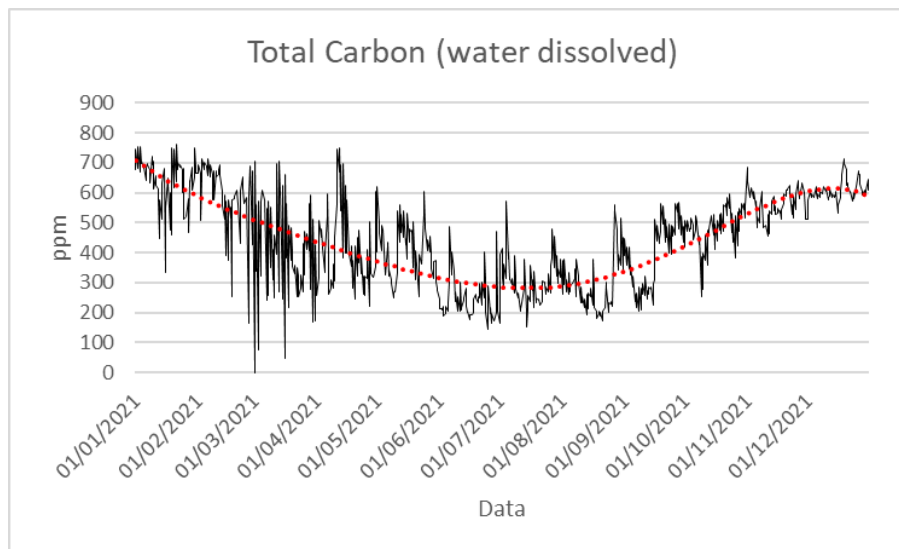


Fig. 19 The TDC in the wetland water; it decreases during the summer.

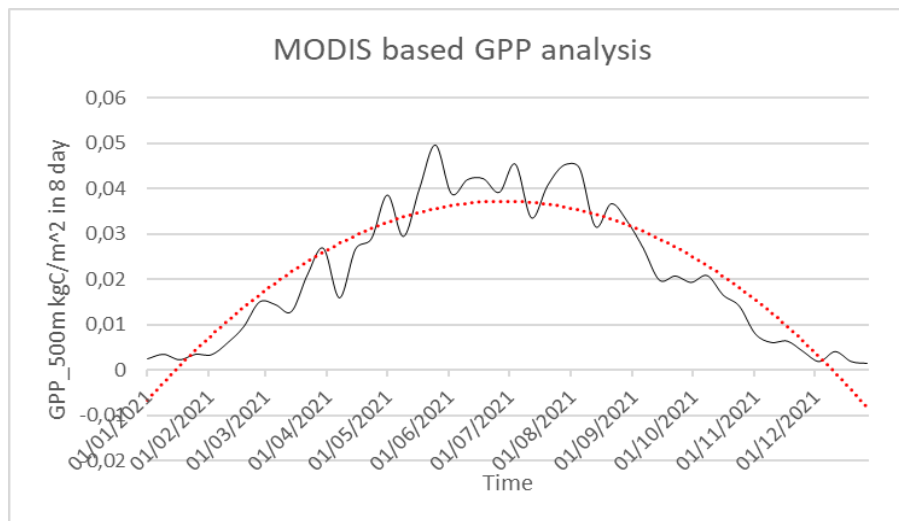


Fig. 20 Satellite GPP analysis; the maximum activity is during the summer.

Many authors supposed that a spot chemical or physical analysis as described in the reference practice UNI 99:2021, is not viable to check a very complex environment like an extensive wetland area (1,000-2,000 ha each site). The necessity to use a deep learning modeling suite for the study of the environment big data and the CO₂ storage is recommend by many authors [14-16]. EDGE IMPULSE is a very useful informatics platform to be used to identify and study a big amount of dataset [17]. It takes raw data, uses signal processing to extract features, and then uses a learning block to classify new data (Fig. 21). Using the spectral features

mode, the system allows us to classify the wetland CCS activity into three groups.

The first group is classified with high CCS activity (-0.018 to -0.03114) kg CO₂/m²/year (dark green).

The second demonstrates an intermediate CCS activity (-0.01069) kg CO₂/m²/year (light green).

The last is with a very low CCS activity (-0.00758 to 0.00288) kg CO₂/m²/year (lighter green).

The accuracy is 100%, the confusion matrix gives good value and the F1 score is 1 (Fig. 22). This demonstrates that the A.I. gives a good model to classify the HCW.

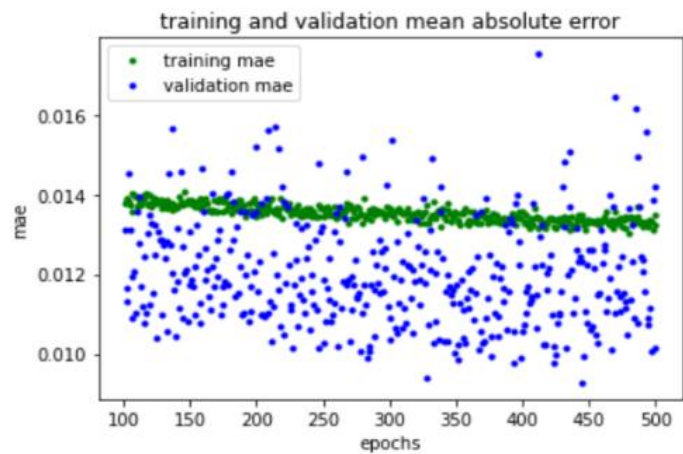


Fig. 21 Comparison of the M.A.E. (Mean Absolute Error) during the Training and Validation test.



Fig. 22 The North Italy HCW “Valli da Pesca” classification elaborate by EDGE IMPULSE platform.

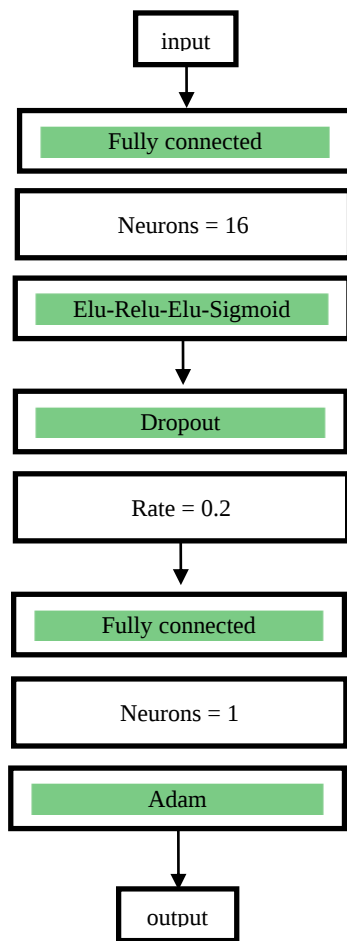


Fig. 23 Binary classifier model designed for forecasting the CCS with different activator and one optimizer.

Meanwhile, several NN (Neural Network) patterns are used to verify the efficiency of the prediction model.

We compare different activator like *SELU* (https://www.tensorflow.org/api_docs/python/tf/keras/activations/selu): Scaled Exponential Linear Unit, *RELU* (https://www.tensorflow.org/api_docs/python/tf/keras/activations/relu): REctified Linear Unit, *ELU* (https://www.tensorflow.org/api_docs/python/tf/keras/activations/elu): Exponential Linear Unit., and *SIGMOID* (https://www.tensorflow.org/api_docs/python/tf/keras/activations/hard_sigmoid): Hard sigmoid activation function. As optimizer we have selected *ADAM*. Adam (Adaptive Moment Estimation) is among the top-most optimization techniques used today (Fig. 23).

The best model was the one with RELU/ADAM.

To verify the quality, the model was studied by comparing the real data to those predicted (Figs. 24-25).

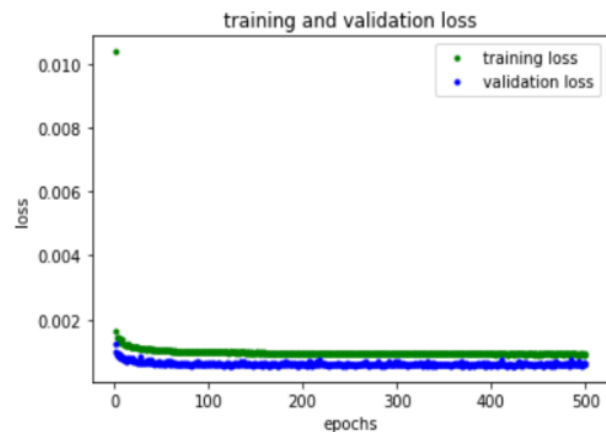


Fig. 24 Comparison of the training and validation loss of the selected model.

20/20 [=====] - 0s 4ms/step - loss: 0.0016 - mae: 0.0149
 20/20 [=====] - 0s 2ms/step

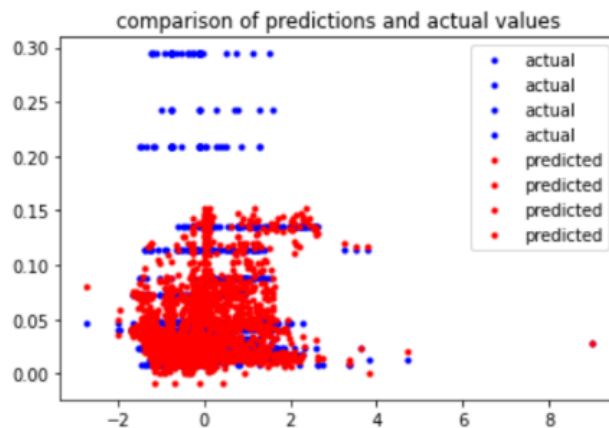


Fig. 25 Real data vs. the predicted data (red).

Table 3 Obtained M.A.E. value during the different mode of training.

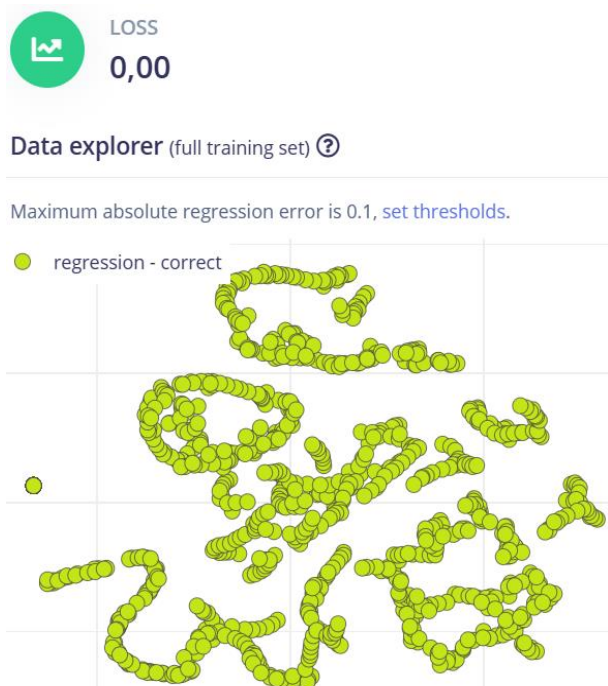
Mode	MAE value
sigmoid/adam	0.0197
selu/adam	0.0230
elu/adam	0.0183
relu/adam	0.0149

The best MAE (Mean Absolute Error) of the training data was obtained using the combination of Relu as activator and Adam as optimizer (Table 3).

The EDGE IMPULSE data explored reveal that the regression check is correct with a low loss and maximum error of 0.1 (Fig. 26).

Another important result is the evaluation of the importance of the type of analysis. The AI studying the huge amount of data acquired by data loggers has highlighted a scale of importance of the parameter selected by the data logger cloud and used for the CCS activity.

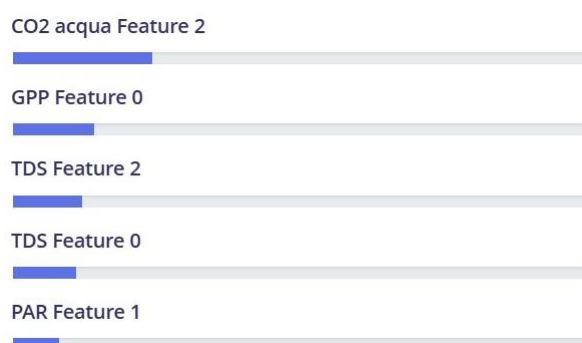
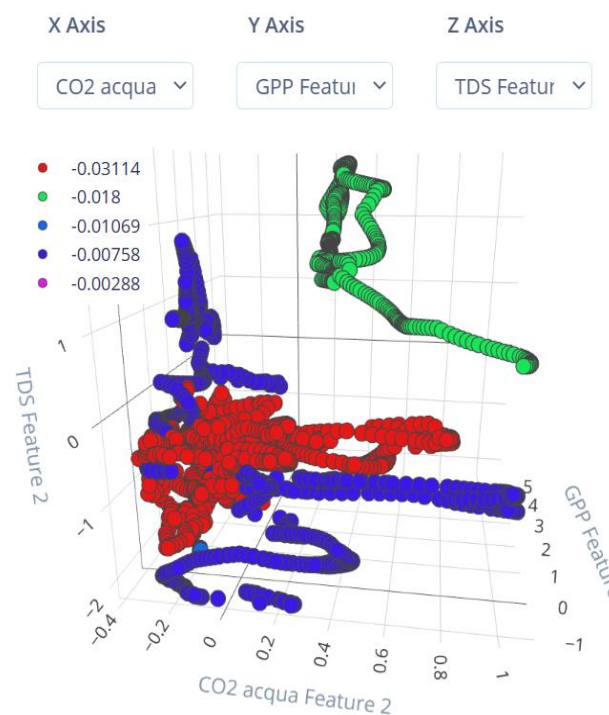
The result of this discovery is that we can decrease the environmental parameter (and expensive probe) to three: (1) the water dissolved CO₂ (CO₂ acqua), (2) the GPP, and (3) the TDS/conductivity (Fig. 27).

**Fig. 26** Regression result of the selected CCS wetland model.

Their relationship on the different CCS class is indicated in Fig. 28.

The models produced by the COLAB and EDGE IMPULSE computer platform have been finally downloaded and used for the programming of the target device as shown in the following diagram (Fig. 29).

The final model was tested again with new lived information in a new Arduino board GSM. The following table shows the accuracy of AI for predicting the wetland CCS (Table 4 and Fig. 30).

**Fig. 27** How important features are for each analysis compared to all other.**Fig. 28** This picture shows a subset of the wetland dataset (2,000 samples on 5,000) using only the best parameters.

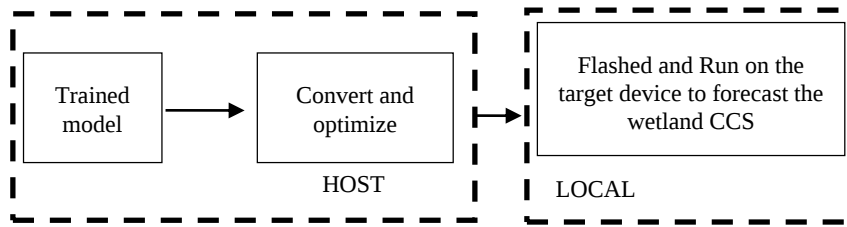


Fig. 29 Flow of the model from the host to local.

Table 4 Prediction test of the CCS model.

Class	Expected	Accuracy	Result
Low CCS	-0.00288	91%	0.01170
Medium CCS	-0.03144	100%	-0.031
High CCS	-0.018	100%	-0.018

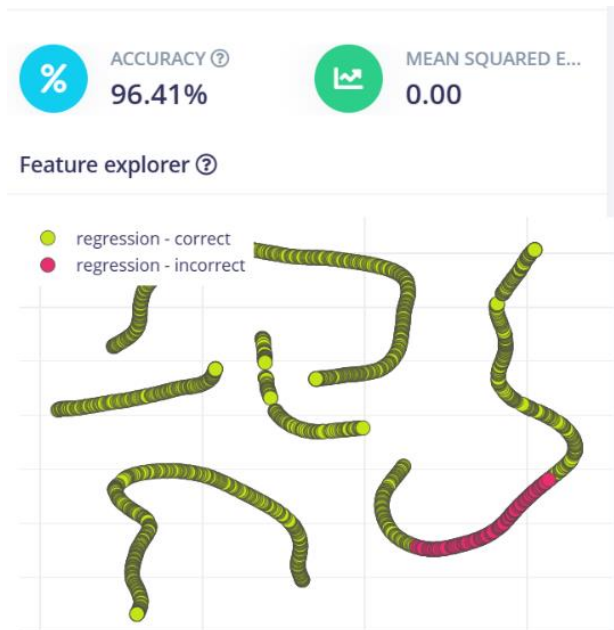


Fig. 30 Model test results in Arduino board using live dataset.

4. Conclusion

It is now consolidated by the scientific community that the problem of climate change due to the uncontrolled increase in GHG in the atmosphere, is a global problem. The contrast actions can be divided into two blocks of activity: the first exploits the ability of certain ecosystems to absorb carbon [18] and the second, the technological one, exploits engineering systems for its transformation and use CO₂ as a resource [19]. In the first block, we find all photosynthetic ecosystems (forests etc.) but also

environments until now no considered such as the seas and lagoons. In recent years, many researchers have considered the importance of wetlands for their strong CO₂ capture and storage capacity [20]. The Venetian lagoon is an important example for the presence of parts with strict human control that take the local and historical name of “Valli da Pesca” or better said “HCW” (Fig. 31) .

Here CO₂ is absorbed from the subsoil and transformed in depth into methane pockets. The mechanism uses both pressure and anaerobic bacterial action. As a result, GHG gas cannot be re-exported into the atmosphere by human action for example drilling that is forbidden for the presence of the historic city of Venice, UNESCO heritage. It acts as the guardian of the irreversibility of the whole natural process. There are multiple standards and or methods for studying carbon storage but mainly are obsolete. For example, the ISO 14064 indicates that the GHG removal can be done by the land use management (table G.4 ISO 14064-1:2018) but does not take into account other environments such as the wetlands and lagoons. Gold Standard protocol (<https://www.goldstandard.org/>) provides methods to the removal of C only from forests or other photosynthetic ecosystems in land. Recently other two methods, UNI PdR 99:2021 and the VERRA VCS (<https://verra.org>) consider other C pools such as wetlands (Verra_VM0033-Tidal-Wetland-and-Seagrass-Restoration-v1.0 and P.d.R. UNI 99:2021 appendix A/A.2 pp. 14-16). It is clear that since 2018, year of publication of ISO 14064, the environmental scientific research has made considerable progress and the new methods published since 2021 include other environments to be used for the activity of CCS. In any case, all actions regulated according to official

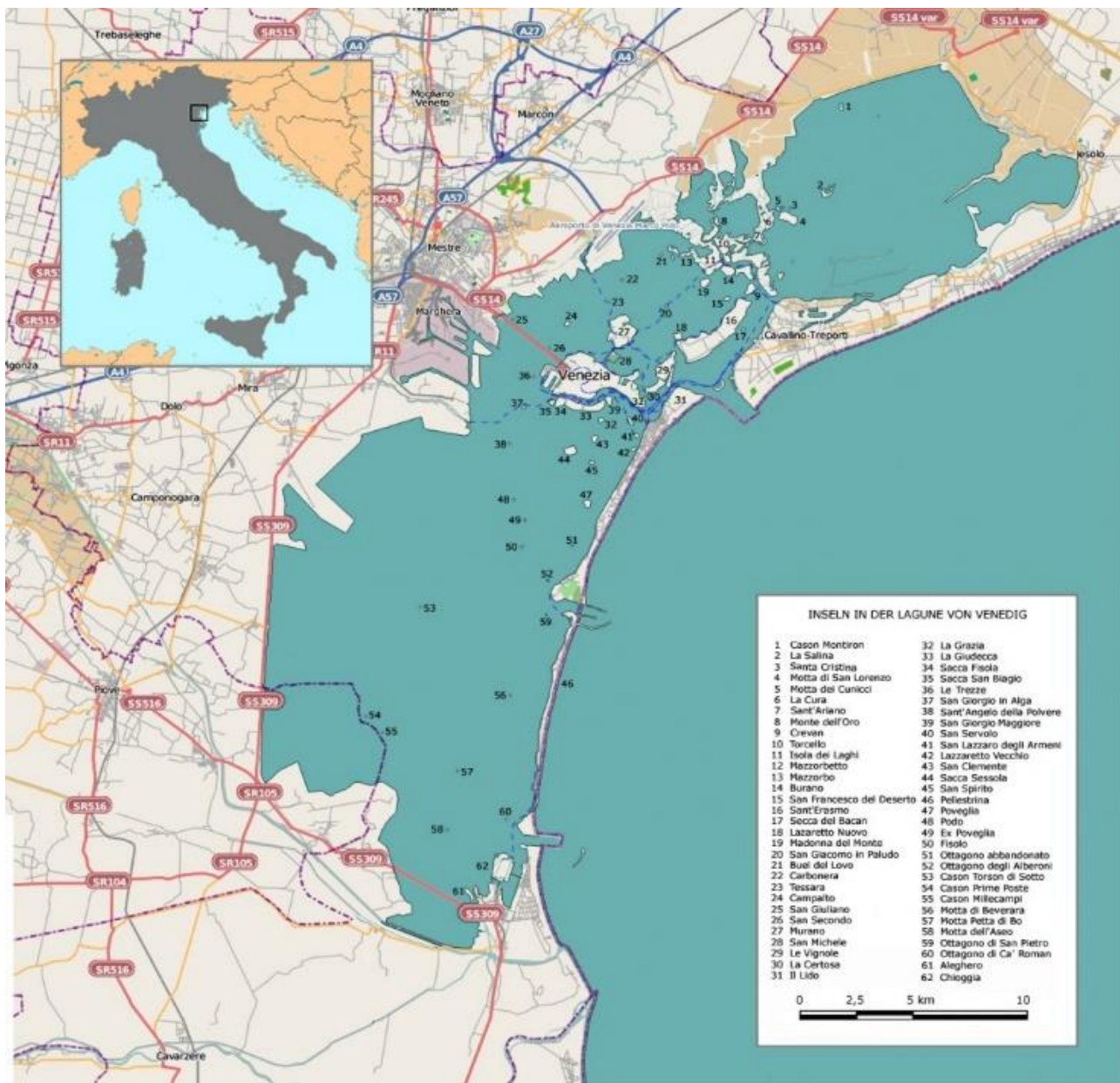


Fig. 31 Venetian Lagoon UNESCO heritage and “valli da pesca” wetlands.

methods, require chemical or physical analysis; but for big surface areas of thousands of hectares, single points for well localized analyses give us a little statistical performance. We must decide “where” and “when”. This problem can be solved by an AI model with few inputs and a CCS value as output. To get this model to be used on the HCW in the Italian Venetian lagoon, we have placed data loggers with data transmission to the cloud. By combining this localized data and satellite data, we get a huge dataset of information that can be

analysed through AI for different types of predictions of carbon absorption and storage in the lagoons. The AI allows us to categorize the HCWs into three groups with a specific and different importance in CO₂ sink. We identify the best parameters to be used in data logger’s probes that can be reduced from the 7 initial to 3 at the end with considerable savings in hardware and relative cost. The combined use of data logger/satellite [21] and the AI allows us to produce a new model called CCS wetland forecast.

