Engenharia Agrícola

ISSN: 1809-4430 (on-line)

www.engenhariaagricola.org.br

Special Issue: Artificial Intelligence

Scientific Paper
Doi: http://dx.doi.org/10.1590/1809-4430-Eng.Agric.v42nepe20210128/2022

FUZZY MODELING IN EVALUATING THE CONSISTENCY AND EFFICIENCY OF DATA REMOTELY MONITORED BY A MULTIPARAMETRIC PROBE

Mádilo L. V. Passos[1*], Alan B. O. de Sousa[1], Adunias dos S. Teixeira[1]

[1*]Corresponding author. Universidade Federal do Ceará/ Fortaleza - CE, Brasil.
E-mail: madilolages@hotmail.com | ORCID ID: https://orcid.org/0000-0003-2495-1612

KEYWORDS

Arduino, Bluetooth, internet of things, artificial intelligence, Wi-Fi.

ABSTRACT

The analysis of water quality for irrigation assists in solving problems with irrigation equipment, such as obstruction in localized systems, being fundamental in precision irrigation. This study aimed to develop and evaluate an affordable multiparametric probe, as well as the performance in the remote data transmission by Bluetooth classic and Wi-Fi. The probe was based on the Arduino Nano platform. The sensors consisted of a pH (potential of hydrogen) sensor, a turbidity sensor (TSW30), and a total dissolved solids sensor. Bluetooth classic (HC-06 module) and ESP8266 module (ESP-01) were implemented for wireless transmission. A fuzzy inference system was used to evaluate the performance of sending data, using the variables bit error rate (BER) and percentage efficiency (Ef). The low-cost multiparametric probe allowed the measurement of pH, turbidity, and total dissolved solids. The Wi-Fi standard (IEEE 802.11 g/b/n), via ESP8266 version 01, presented the best results of consistency and efficiency of information transmission, according to fuzzy modeling.

INTRODUCTION

The analysis of water quality for irrigation can assist in minimizing problems with irrigation equipment, such as obstruction in localized systems (Allende et al., 2015). The sensitivity to clogging depends on the emitter characteristics and the physical, chemical, and biological water quality (Baeza & Contreras, 2020; Zhoua et al., 2019).

Water quality can be obtained through laboratory analysis of samples or the use of portable equipment such as multiparametric probes. Mendonça et al. (2017) recommended the use of probes, as they allow the automatic collection of data through contact with water bodies, not requiring sample collection or transport for analysis in laboratories, in addition to displaying real-time results. High costs and the need for installation in a safe place can be cited as limitations. In this sense, Cunha et al. (2020) emphasized the need to develop multi-parameter meters based on more accessible platforms and sensors, with support for communication technologies that enable data collection and remote access in real-time. The use of embedded micro-controlled platforms and accessible sensors, with connectivity based on basic assumptions of the well-known Internet of Things (IoT), is present in irrigated agriculture (Ogidan et al., 2019; Pandit et al., 2019; García et al., 2020). However, few studies have addressed the performance of the various existing communication architectures.

The application of transmission quality metrics is indispensable for a concise inference. For this purpose, there are indicators such as the bit error rate (BER), received signal strength indication (RSSI), header error check (HEC), and packet error rate (PER) (Conti et al., 2003; Wel & Yan, 2007). In addition, Chiasserini & Rao (2003) emphasized that the joint and simultaneous appreciation of these metrics as an intelligent combination is preponderant for satisfactory inference.

Fuzzy logic is widely applied for applications in multivalued analysis and decision support systems (Sá & Wen, 2019; Veronez et al., 2019). A fuzzy inference system (approximate estimation method) can be used to model the behavior of a process in terms of reliability even

[1] Universidade Federal do Ceará/ Fortaleza - Ceará, Brasil.

Area Editor: Rouverson Pereira da Silva
Received in: 8-7-2021
Accepted in: 2-10-2022

with ambiguous and uncertain information (Mousa et al., 2014; Rizvi et al., 2020). Therefore, it may help in the evaluation of information transmitted by wireless communication protocols.

In this context, this study aimed to develop and evaluate an affordable multiparametric probe, as well as the performance in the remote data transmission by Bluetooth classic and Wi-Fi.

MATERIAL AND METHODS

The development of the multiparametric probe turned to the automatic analysis of water quality physicochemical variables from the pH (potential of hydrogen), turbidity, and total dissolved solids (TDS) sensors. The experiment was carried out at the Laboratory of Hydraulics and Irrigation of the Federal University of Ceará, Fortaleza, CE, Brazil.

Prototyping platform and sensors

The selected hardware model was the Arduino Nano and the IDE (Integrated Development Environment) software, exclusively for this type of development. The Arduino Nano is a small board (45 x 18 mm), complete and compatible with the breadboard based on ATmega328 (Arduino Nano 3.x), functioning as a microcontroller.

The pH (potential of hydrogen) was measured using a Ph4502c sensor module and a Bnc Diymore probe electrode. According to the manufacturer, it has the following characteristics: working voltage of 5 ± 0.2 volts, operating current of 5 to 10 mA, operating temperature range from 0 to 60 °C, analog output (0 to 5 volts), and pH measurements from 0.0 and 14.0, with an alkaline error of 0.2 pH units.

A TSW30 TZT teng Official Store turbidity sensor with interface board and signal conditioning was used for turbidity quantification. According to the manufacturer, it has the following specifications: operating voltage of 5 volts, a maximum current of 30 mA, temperature from 30 to 80 °C, analog output (0 to 4.5 volts), and measurement range from 0.0 to 1000 ± 30 NTU.

A KS0429 Keyestudio total dissolved solids (TDS) sensor was also used. It has a working voltage in the range of 3.3 to 5.5 volts, a working current from 3 to 6 mA, a maximum temperature of 55.0 °C, and analog output (0 to 2.3 volts). The measurement range is from 0.0 to 1000.0 ppm or 0.0 to 1000.0 mg L^{-1}, with a measurement error of ±10% f. s. (25 °C).

Sensor calibration

The direct calibration methodology was used for the Ph4502c sensor module, from pH standard solutions, that is, previously known values (pH values of 1.79, 4.5, 6.88, 12.13, and 13.99) and quantified by a benchtop pH meter.

The indirect method was used for the turbidity sensor calibration. For this purpose, a Digimed DM-TU portable digital turbidimeter previously calibrated in Nephelometric Turbidity Unit (NTU) and solutions obtained from water and soil at different concentrations were used. In the process, the sensor under calibration and the portable meter acted on the same samples. A total of eight solutions with turbidity values of 1.3, 158.0, 255.4, 315.2, 427.4, 571.5, 632.8, and 713.0 NTU were used.

The direct method was applied to obtain the calibration regression of the total dissolved solids (TDS) sensor, using solutions with known electrical conductivity (0.50, 1.0, and 2.0 dS m^{-1}).

The simple linear regression model was chosen for the calibration equations of all sensors. The slopes and intercepts of the models were estimated by two approaches. The ordinary least squares (OLS) method was used to estimate the parameters in the first approach. The parameters in the second approach were estimated using the generalized least squares (GLS) method, which is an efficient method to estimate coefficients of a linear model in the presence of heteroscedasticity and/or correlation between observations.

The general form of the simple linear regression model in matrix notation is given by [eq. (1)].

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

Where:

$\mathbf{Y}$ is the vector of observations (response variable);

$\mathbf{X}$ is a matrix containing the explanatory variable;

$\boldsymbol{\beta}$ is the vector of coefficients, and

$\boldsymbol{\varepsilon}$ is the error vector of the regression model.

The $\boldsymbol{\beta}$ estimator by the ordinary least squares (OLS) method is vector $\mathbf{b}$, obtained according to [eq. (2)].

$$\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} \tag{2}$$

Different assumptions of the OLS method need to be satisfied so that the regression analysis can be properly used, producing the best linear unbiased estimator (BLUE). The independence assessment was performed using the Durbin-Watson test for first-order autocorrelation, homoscedasticity using the Breusch-Pagan test, and normality given by Kolmogorov-Smirnov.

The $\boldsymbol{\beta}$ estimator ($\mathbf{b}_{GLS}$ vector) in the generalized least squares method (GLS) can be expressed by [eq. (3)] when there is a certain residual correlation.

$$\mathbf{b}_{GLS} = (\mathbf{X}^T\mathbf{W}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}^{-1}\mathbf{Y} \tag{3}$$

The matrix $\mathbf{W}$ can be constructed according to [eq. (4)] to correct the correlation between the cases.

$$\mathbf{W} = \begin{bmatrix} 1 & \rho & \cdots & \rho^{N-1} \\ \rho & 1 & \cdots & \rho^{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{N-1} & \rho^{N-2} & \rho^{N-3} & 1 \end{bmatrix} \tag{4}$$

Where:

ρ is a correlation coefficient, in absolute values.

The ρ estimation considered that the error covariance structure was first-order autoregressive, AR(1). Thus, observations $\boldsymbol{\varepsilon}$ at time $\mathbf{t}$, denoted by $\boldsymbol{\varepsilon}_t$, are obtained by [eq. (5)].

$$\boldsymbol{\varepsilon}_t = \rho\boldsymbol{\varepsilon}_{t-1} + \mathbf{a}_t \tag{5}$$

Where:

$\mathbf{a}_t$ is white noise with zero mean and constant variance.

In the case of uncorrelated observations, but with unequal variance, the variance-covariance matrix **W** is diagonal but has non-equal diagonal elements (w_n). In this situation, the **β** estimator (**b_{WLS}** vector) is commonly known as weighted least squares (WLS). Thus, the matrix **W** is given by [eq. (6)].

$$\mathbf{W} = \begin{bmatrix} \frac{1}{w_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{w_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \frac{1}{w_n} \end{bmatrix} \tag{6}$$

Therefore, if W is diagonal, but with unequal diagonal elements, the observations have unequal variance although not correlated. On the other hand, the observations will be correlated if W has elements outside the non-zero diagonals. Furthermore, W can represent heteroscedasticity and the correlation between residues, simultaneously. Thus, if necessary, the models were adjusted considering one of the following ways: GLS method for autocorrelation correction, WLS for non-homogeneous variance, and the generalized least squares method with correction for both dependence and heteroscedasticity.

The adjustments were performed using the R software, more specifically the gls (generalized least squares) function of the nlme library, specifying the argument *corAR1* to correct autocorrelation by the first-order autoregressive structure and/or weights for heteroscedasticity. Details on methodology and implementation in R can be found in Pinheiro & Bates (2000).

The indicators correlation coefficient (r), coefficient of determination (R^2), root mean square error (RMSE), index of agreement (d) by Willmott et al. (1985), and the confidence or performance index (c) were used to measure the statistical performance of the models. The RMSE metrics and the index of agreement were calculated only for OLS adjustments since these parameters are associated with a simple linear regression of ordinary least squares (OLS), according to Willmott et al. (1985), being estimated from the decomposition of the mean error by the referred method.

According to Luiz (2013), the coefficient of determination (R^2) for the GLS and WLS methods is not a good estimator, as the sum of squared residuals and the total quadratic sum consider both sampling errors and model error. Thus, R^2 was not estimated for generalized methodologies.

The significance of the regressions was given using Student's t-test for the slope (b) and linear (a) coefficients. All tests were performed at a 5% significance. The regressions were compared using the Akaike information criterion (AIC) and Bayesian information criterion (BIC) to select the most appropriate estimation methodology.

Bluetooth and Wi-Fi connectivity

The multiparametric probe was built to support wireless data transmission collected by sensors from Bluetooth and Wi-Fi connectivity platforms, technologies that are compatible with Arduino.

Bluetooth classic was implemented by the HC-06 module. The data collected by sensors connected to the Arduino were transferred to HC-06 from the Arduino/HC-06 interface, which uses serial communication, and then to a mobile application on a smartphone. A mobile application, acting as a master, was created for monitoring, while the module functioned as a slave. The basis of application routines was programmed in Delphi language, IDE Rad Studio Rio 10.3.3 Community Edition by Embarcadero®. The compilation was intended for the Android platform. The mobile software was named DTblue and is available at https://play.google.com/store/apps/details?id=br.com.madilopassos.DTblue, whereas the user manual is available at https://dt04.com.br/dtblue/manual.pdf.

Regarding connectivity via wireless fidelity (Wi-Fi), The 802.11 g/b/n standard for wireless fidelity (Wi-Fi) connectivity was used from the ESP8266 ESP-01 module with Wi-Fi V3.0. The module was used as a Wi-Fi Serial Bridge to interface with Arduino Nano V3. The ESP-01 functioned as the master, requesting data from the Arduino (slave) in this connection.

A WebService was built with real-time visualization by web browsers to monitor the traffic data. The Hypertext Markup Language (HTML) and styles (e.g., colors, fonts, and spacing) by CSS (Cascading Style Sheets) were used for front-end development. The PHP Hypertext Preprocessor was used for the back-end. The Hypertext Transfer Protocol (HTTP) was used for information transfer. Paid services offered by the Umbler® platform were used for hosting and domain registration. It supports PHP technologies and MySQL database, which were used in the elaboration of the WebService. The registered domain has the electronic address http://dt04.com.br/.

Probe evaluation and fuzzy modeling

Five collections (tests), denoted by 01, 02, 03, 04, and 05, were performed with the sensors collecting and transmitting data at the same time after setting up the probe and building the tools to verify the consistency and efficiency of the water quality data. Solutions obtained by adding vinegar and sodium bicarbonate aiming at leading to reading variations were used to obtain data from the pH sensor. Turbidity data were obtained from "clean" water and water plus soil solution at different concentrations. Solutions of water and sodium chloride at different proportions were used for measurements with the total dissolved solids sensor.

The Arduino Nano was programmed in each test (01, 02, 03, 04, and 05) to send the collected data every 1.0 minute for 8.0 hours at two monitoring distances (1.0 and 10.0 m) by the Bluetooth classic and Wi-Fi transmission protocols using the HC-06 and ESP-01 modules, respectively. Thus, each database should consist of a total of 480 values per sensor.

The variables bit error rate (BER) and percentage efficiency (Ef) were adopted for the creation of the fuzzy sets and performance classification. BER was calculated according to [eq. (7)].

$$\text{BER} = \frac{\text{Ne}}{\text{Nbits}} \tag{7}$$

Where:

Ne is the number of bits received in error, and

Nbits is the total number of bits received.

Ef was calculated to analyze the data transmission efficiency, that is, recorded for each request during the sending time. It was estimated as the ratio between the number of values received by the number that would be sent/received if there were no errors, according to [eq. (8)].

$$Ef = \frac{Nr}{N} * 100 \tag{8}$$

Where:

Nr is the number of values received, and

N is the total number of values sent (480.0 for 8.0 hours every 1.0 minute).

Once the input variables were defined, the performance classification was performed based on a fuzzy inference system. Its architecture consisted of three modules: fuzzification, inference, and defuzzification. In the first stage, the input variables BER and Ef of three membership functions were modeled with triangular and trapezoidal shapes and linguistic terms low, medium, and high.

The classification proposed by Anderson (2011) was adopted to write BER in terms of linguistic variables, as well as the construction of its universe of discourse, while the clustering approach was used for the percentage efficiency in an attempt to discriminate clusters and guide writing in linguistic terms and because no classification was found in the literature. For this purpose, the Ef data for the Bluetooth and Wi-Fi platforms were applied to the k-means algorithm. Its objective function is defined by [eq. (9)].

$$\text{Minimize } J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^j - c_j \right\|^2 \tag{9}$$

Where:

||xi (j) − cj || is the Euclidean distance between a data vector xi (j) and the center of the cluster cj, indicating the distance from n data points to the centers of k clusters.

The number of clusters was identified by testing k values (k = 1, …, 5) to apply the visual method called the elbow. The number of clusters was determined by the maximum curvature point of the graph (Yuan & Yang, 2019). The percentage efficiencies for the execution of the methodology were calculated for the data transmitted by Bluetooth and Wi-Fi in groups of 8.0 hours (total) and subgroups of 4, 2, and 1 hour.

The number of clusters was fixed at 3 using the elbow method and after some experimentation. Thus, the linguistic terms for Ef were defined from the observed centroids.

Table 1 shows the triangular and trapezoidal numbers for BER (a), Ef (b), and the output. The latter was associated with five membership functions.

TABLE 1. Triangular and trapezoidal numbers for BER (a), Ef (b), and output (performance).

Linguistic terms	Input		Linguistic terms	Output
	BER	Ef		Performance
Low	[0 0 10^{-5} 10^{-5}]	[0 0 50 60]	Bad	[0 0 20 20]
Medium	[10^{-6} 10^{-4} 10^{-3}]	[50 70 80]	Poor	[20 30 40]
High	[10^{-4} 10^{-3} 1 1]	[70 80 100 100]	Fair	[30 45 60]
Universe	0.00 to 1.00	0.00 to 100,00%	Good	[45 70 80]
–	–	–	Excellent	[75 90 100 100]
–	–	–	Universe	0.00 to 100.00%

Regarding the inference mechanism, nine control rules were established relating the three membership functions of BER with the three of Ef. The set of elaborated rules was interpreted by the Mamdani inference method, also known as the Mamdani controller (Mamdani, 1973). The maximum composition of the minima (max-min) was used for the composition of each control rule. The centroid or center of gravity method was chosen for defuzzification.

The Python programming language was used to develop the model, implement the fuzzy logic, and cluster analysis, particularly the scikit-fuzzy and scikit-learn libraries.

RESULTS AND DISCUSSION

Multiparametric probe and sensor calibration

The cost of manufacturing the affordable multiparametric water quality probe with all sensors and accessory components was around US$ 42.82, calculated as of 6/22/2021.

The prototyping of low-cost multiparametric for different purposes probes is recurrent, as observed by Ungureanu et al. (2010), Faustine et al. (2014), and Win et al. (2019). However, Cunha et al. (2020) noted that the authors infer about stability in remote transmission, noise, reduced cost, and portability, without validating the collected information or models.

Regarding the methods for estimating the regression parameters, the evaluation of the assumptions of the ordinary least squares (OLS) method by the Durbin-Watson, Breusch-Pagan, and Kolmogorov-Smirnov tests to verify the premises of independence, homogeneity of variance, and normality for the pH, turbidity, and total dissolved solids (TDS) sensors are shown in Table 2. There was no premise regarding the residuals that must be verified for the generalized least squares (GLS) method (Luiz, 2013).

TABLE 2. Durbin-Watson, Breusch-Pagan, and Kolmogorov-Smirnov test statistics and sample size (n) of simple linear calibration equations.

Sensor	Durbin-Watson	Breusch-Pagan	Kolmogorov-Smirnov	n
pH	0.3755^{NS}	1.1921*	0.1290*	50
Turbidity	0.3932^{NS}	12.3608^{NS}	0.09123*	80
TDS	0.2936^{NS}	11.1543^{NS}	0.3111^{NS}	30

*, ^{NS} – Significant and not significant at 5%, respectively; and TDS – Total dissolved solids.

The assumption of the absence of autocorrelation between cases for calibration equations of all sensors was not met. Thus, a residual first-order correlation, AR(1), was found according to the Durbin-Watson test. Some causes of residual dependence may be due to measurement errors or omissions of important variables in the model. Nevertheless, unbiased coefficients can be obtained, but they will not be efficient and significance tests and confidence intervals are unreliable (Chen, 2016).

The presence of correlation between the errors generated by the models is possibly related to how the calibration data were collected, as ten readings with the sensors were taken in the same experimental units (solutions) at different times, configuring a test with repeated measures. Thus, there may be a correlation between the data obtained at these times, and, overall, these correlations are higher at closer times (Yokoo, 2014).

Heterogeneity of variance of error terms was observed for the adjustments of turbidity and total dissolved solids, with non-normality for the latter. Violation of homoscedasticity can be associated with the presence of measurement errors or the absence of independent variables in the model, promoting errors in the significance tests and confidence intervals and, even though the least-squares coefficients continue to be unbiased, they are not efficient (no minimum variance) (Figueiredo-Filho et al., 2011). According to the Gauss-Markov theorem, the coefficients obtained with the distance from normality for residuals are biased and inefficient.

Thus, the ordinary least squares method did not produce the best linear unbiased estimator for the calibration equations of the pH, turbidity, and TDS sensors.

Importantly, only the Ph4502c sensor module, which has automatic compensation for room temperature, did not provide relaxation of the assumptions of heterogeneity of variance or absence of normality of residuals.

Therefore, the solution temperature can be a key factor to be measured, as it influences several other parameters. Pinto et al. (2011) stated that the potential difference generated between the indication and reference electrodes is not only a function of the solution pH but also sensitive to temperature changes. Thus, the solution temperature interferes with the pH reading, and the higher the temperature, the higher the pH (Pratami et al., 2020). Both measures are necessary and must be integrated into a monitoring system (Patil et al., 2015). The pH meter of the present study has a temperature sensor for automatic compensation of the air temperature, but the liquid temperature is not measured.

Regarding the turbidity measurements, part of the temporal dispersion of the readings may be due to the water + soil solutions used for calibration since a reduction in turbidity was observed with the deposition of material in suspension at the base of the container. In addition, the manufacturer reports temperature correlation with the readings provided by the sensor. Gillett & Marchiori (2019) also highlighted the need for periodic calibration after a certain period of use. In terms of total dissolved solids (TDS), the conductivity of ions in water depends on water temperature (Cloete et al., 2016). Ions move faster when water is hot and, therefore, the apparent conductivity is increased when water reaches higher temperatures (Rietmanm et al., 1985).

Although the temperature of the solutions has not been measured, the developed multiparametric probe has support for a temperature sensor that can perform the proper compensation, and the DS18B20 can be used for this purpose.

Table 3 shows the linear equations fitted by the OLS and GLS methods for the pH, turbidity, and TDS sensors. Only the intercept of the adjustment of total dissolved solids of the GLS method was not significant.

TABLE 3. Models (ŷ) of the relationship between pH, turbidity, and total dissolved solids (TDS) values with voltage (x, volts).

Method	Sensor	Equation
OLS	pH	$\hat{y} = -22.7850^{*}x + 74.3480^{*}$
	Turbidity	$\hat{y} = -436.3956^{*}x + 1811.4481^{*}$
	TDS	$\hat{y} = 604.8903^{*}x - 339.7890^{*}$
GLS	pH	$\hat{y} = -14.0615^{*}x + 48.7629^{*}$
	Turbidity	$\hat{y} = -396.2280^{*}x + 1684.3480^{*}$
	TDS	$\hat{y} = 490.0772^{*}x - 99.6983^{NS}$

*, ^{NS} – Significant and not significant by the t-test at 5% for slopes and intercepts.

As the violation of non-correlation and heteroscedasticity was verified, the GLS method with both corrections was applied to estimate the coefficients of turbidity and total dissolved solids sensor equations. However, only the residual autocorrelation effect was considered for the pH adjustment.

Table 4 shows the calibration performance measures, root means square error (RMSE), coefficient of determination (R^2), correlation coefficient (r), index of agreement (d), and confidence index (c) for sensor adjustments of pH, turbidity, and TDS.

TABLE 4. Statistics of model performance, root mean square error (RMSE), coefficient of determination (R^2), correlation coefficient (r), index of agreement (d), and confidence index (c).

Method	Sensor	RMSE	R^2	r	d	c
OLS	pH	1.1750	0.9342	0.9665	0.9827	0.95
	Turbidity	56.9850	0.9401	0.9696	0.9843	0.95
	TDS	165.1435	0.8335	0.9130	0.9538	0.87

The turbidity and pH calibration equations for the OLS methodology showed the highest precision, according to the coefficients of determination (R^2) of 0.9401 and 0.9342, respectively, while TDS estimated the lowest R^2 (0.8335). The performance statistics express good precision of the adjustments. However, the use of linear models provided by OLS must be careful due to the relaxation of at least one assumption regarding the residuals.

Thus, the evaluation and selection of the method with higher statistical consistency were based on the Akaike (AIC) and Bayesian (BIC) information criteria. The results are shown in Table 5.

TABLE 5. Akaike information criterion (AIC) and Bayesian information criterion (BIC).

Method	Sensor	AIC	BIC
OLS	pH	162.0207	165.8447
	Turbidity	877.8761	882.6402
	TDS	395.5452	398.3476
GLS	pH	80.5616	84.3857
	Turbidity	803.3822	808.1462
	TDS	348.8896	351.6920

The AIC and BIC values for all adjustments were lower for the coefficients estimated by the generalized least squares method. The Akaike criterion weighs between adequacy to data and adjustment complexity (Sobral & Barreto, 2016). As a rule of choice, the lower the AIC and BIC values, the better the model (Pho et al., 2019). In this case, the low AIC and BIC values denote the superiority of linear regressions by the generalized least squares estimator considering heteroscedasticity and/or residual autocorrelation.

Some authors have evaluated only the statistical performance of the models and reported good precision in the sensor calibration equation using the OLS methodology. Cunha et al. (2020) implemented an integrated system (hardware and software) for surface water analysis and found R^2 values between 0.996 and 0.998 for the pH sensor when comparing the readings with equipment available on the market. Gillett & Marchiori (2019) developed four low-cost continuous turbidity monitor sensors, calibrated in the range from 0 to 100 NTU and attached to the outside of a transparent tube at angles of 90 and 180 degrees, and obtained R^2 values ranging from 0.9934 to 1.0000 for a multiple linear fit. Jorge Junior (2019) observed $R^2 = 0.995$ for the calibration of the SKU SEN0189 turbidity sensor.

Probe evaluation and fuzzy modeling

Table 6 shows the results for the parameters of transmission quality (BER), efficiency (Ef), and performance (fuzzy logic) for monitoring data at distances of 1.0 and 10.0 meters with the Bluetooth classic communication technology (BT). Among others, BT comprises short-range transfer technologies (Lia et al., 2018) at initial speeds of 1 to 2 Mbps, implementing the concept of frequency-division multiplexing, i.e., it adopts a frequency-hopping spread spectrum (FHSS) time-slot scheme based on the forward error correction (FEC) coding technique (Conti et al., 2003; Nikitha et al., 2018).

TABLE 6. Transmission quality (bit error rate, BER), efficiency (Ef), sample size (N), and performance for Bluetooth classic during 8 hours of monitoring and 1-min transmission interval.

Test	Bluetooth classic (1 meter)				
	BER	Ef (%)	N	Performance	Membership
01	0.00	97.50	468	Excellent	1.00
02	0.00	97.92	470	Excellent	1.00
03	0.036	98.12	471	Poor	1.00
04	0.00	97.29	467	Excellent	1.00
05	0.00411	97.71	469	Good	0.96
Test	**Bluetooth classic (10 meters)**				
	BER	Ef (%)	N	Performance	Membership
01	1.00×10^{-3}	47.71*	229	Poor	0.67
02	0.00	97.50	468	Excellent	1.00
03	3.26×10^{-3}	97.50	468	Good	0.75
04	0.00	96.87	465	Excellent	1.00
05	0.00	18.96*	91	Good	1.00

*Device disconnected.

Tests 01, 02, and 04 at a distance of 1.0 meter showed no inconsistencies in the packages sent (BER = 0.00), but the sending efficiency (Ef) presented inconsistencies, with excellent performance, with a membership of 1.00, when associated. In the concept of membership, an element belongs to the set with a degree of membership, that is, an element can partially belong to more than one set at the same time (Pessoa et al., 2020).

The IEEE 802.15.1 standard of tests 03 and 05 provided noisy bits BER of 3.60×10^{-2} and 4.11×10^{-3}, respectively. Despite the high transmission efficiency, test 03 presented a classification of poor (membership of 1.00) and good for test 05. Tests 02, 04, and 05 at 10.0 meters showed no information loss (BER = 0.00), only a delay in sending, with excellent performance for tests 02 and 04 and poor for tests 01 and 05.

Shin (2002) reported that unrestricted access to the ISM (Industrial, Scientific, and Medical) spectrum exposes Bluetooth devices to a high level of interference, classified into two categories. The first category is linked to non-Bluetooth devices, such as IEEE 802.15 and HomeRF wireless PANs, and operates at the 2.4 GHz band, with IEEE 802.11 also operating at this band for wireless LAN. The second category is due to other Bluetooth connections (self-interference), as these networks adopt an ad-hoc topology called piconet and scatternet, which allows many Bluetooth devices to coexist in the vicinity. Thus, an open Bluetooth connection nearby may result in interference.

Regarding the problems associated with non-Bluetooth devices, failures observed during the tests (Table 6) can be justified by the presence of the HC-06 (IEEE 802.15.1) and ESP-01 (802.11 g/b/n) modules operating from simultaneously on the same board. In fact, Mathew et al. (2010) highlighted that, as they operate in the same 2.4 GHz ISM band, there is mutual interference, a problem already addressed by Conti et al. (2003), Golmie et al. (2003), and Wel & Yang (2007). Chiasserini & Rao (2003) reported that influences between 802.11 and BT occur whenever the interference energy is sufficient to cause a decrease in the signal-to-noise ratio at the receiver and the two transmissions of the system overlap in both frequency and time.

Conti et al. (2003) emphasized that a careful evaluation should consider the presence of thermal noise, propagation impediments, interference, the relative distance between the interfering systems, modulation formats, coding techniques, frequency hopping, packet structures, and traffic loads.

Table 7 shows the results for readings at distances of 1.0 and 10.0 meters with the standard 802.11 g/b/n Wi-Fi platform (ESP-01). No failures were observed in the transferred packets (BER = 0.00) for all tests at both distances. Thus, distance did not influence BER.

TABLE 7. Transmission quality (bit error rate, BER), efficiency (Ef), sample size (N), and performance for Wi-Fi during 8 hours of monitoring and 1-min transmission interval.

Test	Wi-Fi (1 meter)				
	BER	Ef (%)	N	Performance	Membership
01	0.00	95.42	458	Excellent	1.00
02	0.00	97.71	469	Excellent	1.00
03	0.00	98.12	471	Excellent	1.00
04	0.00	96.87	465	Excellent	1.00
05	0.00	97.29	467	Excellent	1.00
Test	**Wi-Fi (10 meters)**				
	BER	Ef (%)	N	Performance	Membership
01	0.00	93.33	448	Excellent	1.00
02	0.00	97.50	468	Excellent	1.00
03	0.00	96.66	464	Excellent	1.00
04	0.00	96.87	465	Excellent	1.00
05	0.00	96.66	464	Excellent	1.00

Similarly, Ef had its results influenced by transmission delays, as observed in the Bluetooth tests. Also, it manifested a higher magnitude at 1.0 m (98.12%) and smaller at 10.0 m. (93.33%). Furthermore, the proposed fuzzy inference system produced an excellent output in all tests, with a degree of membership of 1.00.

As previously described, Wi-Fi and Bluetooth standards influence each other. However, the IEEE 802.11 standard defines physical media (PHY) and media access control (MAC) layer protocols (Golmie et al., 2003). When a node using IEEE 802.11b as the wireless standard wants to send a packet over the network, it uses the carrier detection protocol running on the media access control (MAC) layer to determine whether the media is busy or idle and uses their knowledge of 802.11 and BT activity to predict collisions (Chiasserini & Rao, 2003; Mathew et al., 2010).

Challoo et al. (2012) studied the interference between Wi-Fi mainly as an aggressor in Bluetooth and ZigBee and concluded that IEEE 802.15.4 has a small impact on the IEEE 802.11 performance, while IEEE 802.11 can have great significance on the ZigBee and Bluetooth performances.

Furthermore, instabilities in connection and signal attenuation can influence performance. Correia et al. (2016) studied the potential and limitations of automation with a low-cost platform and developed a prototype for automatic irrigation monitoring and control, with remote activation via WEB application. The authors reported that the signal underwent attenuation in the distance between the router and the module of 10 m in a straight line due to the walls, furniture, and appliances arranged between the devices.

CONCLUSIONS

Low-cost platforms (hardware) enabled the development of the multiparametric probe with a total of US$ 42.82.

The experimental calibration in the evaluation of probe sensors must be very careful because it can harm statistical inferences regarding the adjustments and, therefore, in the monitoring of water quality parameters.

The applied fuzzy inference system was satisfactory and allowed the performance evaluations according to the conditions of the proposed study, managing to capture the distinctions between the investigated protocols from transmission quality (BER) and efficiency (Ef) parameters.

Among the communication technologies supported by the probe, the platform for Wi-Fi communication (IEEE 802.11 g/b/n) via ESP8266 version 01 presented the best performances compared to Bluetooth classic (IEEE 802.15.1) by the HC-06 V2.0 + EDR module.

REFERENCES

Allende A, Monaghan J, Irrigation (2015) Water Quality for Leafy Crops: A Perspective of Risks and Potential Solutions. International Journal of Environmental Research and Public Health 12 (1): 7457-7477. DOI: https://doi.org/10.3390/ijerph120707457.

Anderson G (2011) Bit-Error-Rate and Capacity Estimation in Wireless Networks. Tese, Philadelphia, Drexel University.

Baeza R, Contreras JI (2020) Evaluation of Thirty-Eight Models of Drippers Using Reclaimed Water: Effect on Distribution Uniformity and Emitter Clogging. Water 12(1463): 1-20. DOI: http://dx.doi.org/10.3390/w12051463.

Challoo R, Oladeinde A, Yilmazer N, Ozcelik S, Challoo L (2012) An Overview and Assessment of Wireless Technologies and Coexistence of ZigBee, Bluetooth and Wi-Fi Devices. Procedia Computer Science 12 (1):386 – 391. DOI: https://doi.org/10.1016/j.procs.2012.09.091.

Chen Y (2016) Spatial Autocorrelation Approaches to Testing Residuals from Least Squares Regression. PLoS ONE 11(1):1-19. DOI: https://doi.org/10.1371/journal.pone.0146865.

Chiasserini CF, Rao RR (2003) Coexistence Mechanisms for Interference Mitigation in the 2.4-GHz ISM Band. IEEE Transactions on Wireless Communications 2(5):964-975.

Cloete NA, Malekian R, Nair L (2016) Design of Smart Sensors for Real-Time Water Quality Monitoring. IEEE 4(1): 3975-3990. DOI: https://doi.org/10.1109/ACCESS.2016.2592958.

Conti A, Dardari D, Pasolini G, Andrisano O (2003) Bluetooth and IEEE 802.11b Coexistence: Analytical Performance Evaluation in Fading Channels. IEEE Journal on Selected Areas in Communications 21(2):259-269.

Correia GT, Rocha HRO, Rissino SD (2016) Automação de sistema de irrigação com monitoramento via aplicativo Web. Engenharia na agricultura 24(4):314-325.

Cunha RA, Silva AG, Nascimento LD, Lima LS, Bandeira MLFS (2020) POLYPUS: sonda de baixo custo para análise de parâmetros físico-químicos em águas superficiais. Revista gestão & sustentabilidade Ambiental 9(4):653-671. DOI: http://dx.doi.org/10.19177/rgsa.v9e42020653-671.

Faustine A, Mvuma AN, Mongi HJ, Gabriel MC, Tenge AJ, Kucel, SB (2014) Wireless sensor networks for water quality monitoring and control within lake victoria basin: Prototype development. Wireless Sensor Network 6(12):281-289. DOI: http://dx.doi.org/10.4236/wsn.2014.612027.

Figueiredo Filho D, Nunes F, Rocha EC, Santos ML, Batista M, Silva Júnior JÁ (2011) O que Fazer e o que Não Fazer com a Regressão: pressupostos e aplicações do modelo linear de Mínimos Quadrados Ordinários (MQO). Revista Política Hoje 20(1):44-99.

García L, Parra L, Jimenez JM, Lloret J, Lorenz P (2020) IoT-Based smart irrigation systems: an overview on the recent trends on sensors and iot systems for irrigation in precision agriculture. Sensors 20 (1042):1-48. DOI: https://doi.org/10.3390/s20041042.

Gillett D, Marchior A (2019) A Low-Cost continuous turbidity monitor. Sensors 19 (14):2-18. https://doi.org/10.3390/s19143039.

Golmie N, Van Dyck RE, Soltanian A, Tonnerre A, Rébala O (2003) Interference Evaluation of Bluetooth and IEEE 802.11b Systems. Wireless Networks 9(1):201–211.

Jorge Junior JA (2019) Desenvolvimento de sonda multiparâmetro dinâmica de qualidade de água (sonda aqua-x). Dissertação, Fernandópolis, Universidade Brasil.

Lia S, Xub LD, Zhao S (2018) 5G Internet of Things: A Survey. Journal of Industrial Information Integration 10(1):1-9. DOI: https://doi.org/10.1016/j.jii.2018.01.005.

Luiz SF (2013) Regionalização hidrológica de vazões mínimas por meio do Método dos Mínimos Quadrados Generalizados aplicada à bacia do Alto São Francisco. Dissertação, Belo Horizonte, Universidade Federal de Minas Gerais.

Mamdani EH (1973) Aplications of fuzzy algorithms for control of simple dynamic plant. IEE Proceedings… 121 (12):1585-1588.

Mathew A, Chandrababu N, Elleithy K, Rizvi S (2010) Interference of 802.11B WLAN and Bluetooth: analysis and performance evaluation. International journal of Computer Networks & Communications 2(3):140-150. DOI: http://dx.doi.org/10.5121/ijcnc.2010.2311.

Mendonça AM, Forgiarini FR, Silveira GL, Cruz JC (2017) Confiabilidade dos parâmetros monitorados em águas por sonda multiparâmetros. REGA 14 (3):1-11. DOI: http://dx.doi.org/10.21168/rega.v14n0.p3-3.

Mousa AK, Croock MS, Abdullah MN (2014) Fuzzy based Decision Support Model for Irrigation System Management. International Journal of Computer Applications 104(9):975-8887. DOI: http://dx.doi.org/10.5120/18230-9177.

Nikitha S, Nandhini T, Pavithra K, Kinol AMJ (2018) A BLUETOOTH BASED AUTOMATIC IRRIGATION SYSTEM. International Research Journal of Engineering and Technology 05(03):587-590.

Ogidan OK, Onile AE, Adegboro OG (2019) Smart irrigation system: a water management procedure. Agricultural Sciences 10 (1):25-31. DOI: https://doi.org/10.4236/as.2019.101003.

Pandit AP, Chavan SA, Bdgar PV, Sanap VR, Pawar MD (2019) IOT based smart irrigation system. IJARIIE 5(3):1025-1031.

Patil K, Patil S, Patil S, Patil V (2015). Monitoring of turbidity, PH & temperature of water based on GSM. International Journal for Research in Emerging Science and Technology 2(3):16-21.

Pessoa MAR, Souza FJ, Domingos P, Azevedo JPS (2020) Índice fuzzy de qualidade de água para ambiente lótico — IQAFAL. Engenharia Sanitária e Ambiental 25(1):21-30. DOI: https://doi.org/10.1590/S1413-41522020147587.

Pho K-H, Ly S, Ly S, T. Lukusa TM (2019) Comparison among Akaike Information Criterion, Bayesian Information Criterion and Vuong's test in Model Selection: A Case Study of Violated Speed Regulation in Taiwan. Journal of Advanced Engineering and Computation 3(1): 293-303. DOI: http://dx.doi.org/10.25073/jaec.201931.220.

Pinheiro, JC, Bates DM (2000) Mixed-effects models in S and S-PLUS. New York, Springer, p201-266.

Pinto MF, Camargo AP, Almeida ACS, Alves DG, Botrel TA (2011) Desenvolvimento de um sistema para o controle do pH da água para microirrigação. Revista Brasileira de Engenharia Agrícola e Ambiental 15(2):211–217.

Pratami LWD, Ariswati HG, Titisari D (2020) Effect of Temperature on pH Meter Based on Arduino Uno With Internal Calibration. Journal of Electronics, Electromedical, and Medical Informatics (JEEEMI) 2(1):23-27. DOI: https://dx.doi.org/10.35882/jeeemi.v2i1.5.

Rietmanm A, Kaplan ML, Cava RJ (1985) Lithium ion-poly (ethylene oxide) complexes. I. Effect of anion on conductivity. Solid State Ionics 17(1): 67-73.

Rizvi S, Mitchell J, Razaque A, Rizvi MR, Williams I (2020) A fuzzy inference system (FIS) to evaluate the security readiness of cloud service providers. Journal of Cloud Computing: Advances, Systems and Applications 9(42):1-17.

Sá YVA, Wen TC (2019) A inteligência artificial (Lógica Fuzzy) para cálculo de estoque de segurança local em empresas multinacionais. Gestão da Produção, Operações e Sistemas 14(4):1-10.

Shin SY (2002) Performance evaluation of interference Between Bluetooth Networks using Bit Error Rate. In: Winter Workshop. Kushu, Anais...

Sobral TEL, Barreto G (2016) Utilização dos critérios de informação na seleção de modelos de regressão linear. Proceeding Series of the Brazilian Society of Applied and Computational Mathematics 4(1): 1-7. DOI: https://doi.org/10.5540/03.2016.004.01.0220.

Ungureanu F, Lupu RG, Stan A, Craciun I, Teodosiu C (2010) Towards real time monitoring of water quality in river basins. Environmental Engineering and Management 9(9):1267-1274. DOI: http://dx.doi.org/10.30638/eemj.2010.164.

Veronez ACCS, Santo AF, Souza CC, Reis Neto JF (2019) Classificação de Frutas e Legumes Utilizando Lógica Fuzzy. Revista Ciências Gerenciais 23(37):52-56. DOI: https://doi.org/10.17921/1415-6571.2019v23n37p52-56.

Wel J, Yan M (2007) Bit error rate analysis of Wi-Fi and bluetooth under the interference of 2.45 GHz RFID. THE journal of China Universities of Posts and Telecommunications 14(1):89-93. DOI: http://dx.doi.org/10.1016/S1005-8885(08)60019-9.

Willmott CJ, Ackleson SG, Davis JJ, Feddema K, Klink DR (1985) Statistics for the evaluation and comparison of models. Journal of Geophysical Research 90(5):8995-9005.

Win TTN, Bogaard T, Giesen NV (2019) A Low-Cost water quality monitoring system for the Ayeyarwady river in myanmar using a participatory approach. Water 11(10):1-16. DOI: https://doi.org/10.3390/w11101984.

Yokoo MJ-I, Ortelan AA, Sarmento JLR, Rosa GJM, Cardoso FF, de Albuquerque LG (2014) Medidas repetidas no estudo de características de crescimento e carcaça avaliadas por ultrassom em novilhas de corte cruzadas. Boletim de Indústria Animal 71(3):200-210. DOI: http://dx.doi.org/10.17523/bia.v71n3p200.

Yuan C, Yang H (2019) Research on K-Value selection method of K-Means clustering algorithm. Multidisciplinary Scientific Journal 2(16):226-235. DOI: https://doi.org/10.3390/j2020016.

Zhoua B, Zhoua H, Puig-Barguésc J, Lia Y (2019) Using an anti-clogging relative index (CRI) to assess emitters rapidly for drip irrigation systems with multiple low-quality water sources. Agricultural Water Management 221(20): 270–278. DOI: https://doi.org/10.1016/j.agwat.2019.04.025.