**Kalibrasi dan Analisis Karakteristik Tingkat Ketidakpastian**

**untuk Sensor Interfacing Voltage**

(Calibration and Analysis of Uncertainty Level Characteristics

for Sensor Interfacing Voltage)

Abstrak:

Penelitian melakukan kalibrasi regressi linear terhadap sensor dengan sinyal analog sebagai subjek pengelolaan. Kemudian sensor dimanfaatkan sebagai perangkat pengukuran dengan uncertainty level yang belum ditentukan. Kemudian dilakukan analisis terhadap hasil pengukuran untuk menentukan penyimpangan nilai pembacaan yang ditunjukkan sensor. Analisis menggunakan standar deviasi hasil pengukuran. 1000 sample dikumpulkan untuk menentukan confidence level dari representasi data. Pembacaan sensor berupa nilai tegangan nantinya diolah menjadi standar deviasi, Uncertainty Level error, Data Error Desity, and Analisys of Variance error distribution. Hasil Menunjukkan ....

Abstract:

1. Introduction

Sensors in general, are often used in various fields to help the process of measuring physical quantities in solving problems with the value of an object (Maulud and Abdulazeez, 2020). Sensors detect and respond directly to certain types of changes or inputs from the physical environment (Dinu and Apetrei, 2020). For example, when someone wants to know mass or weight of an object, they need a tool to determine the mass value of the object, such as a scale. Apart from sensors, researchers use manual measuring instruments as measuring media. However, each object has its level of accuracy and difficulty in the measurement process. The role of the sensor here is as a measuring medium with compact use (Moradi, Akbari and Wild, 2019). Sensors are connected to digital devices or software in their use. The sensor's work is determined by comparing the sensor's analog signal readings. The more stable the sensor's analog value, also better the level of reading the value on the sensor (Pahuja, 2022). The sensor requires an alignment circuit to manage the payoff value or uses a calibration process to manage the signal input. Calibration is done by looking for a mathematical equation of changes in analog sensor values (Magaski and Anwari, 2022). Dalam tulisan (Aroulanandam *et al.*, 2022) usually, sensor users can use microcontrollers to process analog signals. The sensor can also present a digital signal as a reading input. The difference between the reading of the analog signal input and the digital signal on the microcontroller is in the variation of signal changes where the analog is a voltage signal from 0 to 5 volts. Whereas digital signals only use 0 = LOW and 5 = HIGH as an input values, the range of differences used for sensor detection by digital read comes from wave frequencies that change with time. Then (Maulud and Abdulazeez, 2020) give statement, The sensor must be calibrated by comparison with other equipment to take advantage of the shift from the input signal. The calibration process can be carried out by adjusting the two x and y values. For example, x is the analog value then y is the measured value (based on international standards). Through these two values, intersecting (x, y) value variations are sought, then a linear regression formula is taken from the variations that have been collected. Linear regression is the simplest and most commonly used algorithm (Joebaedi *et al.*, 2019). This method uses a mathematical approach to perform predictive analysis. Linear regression allows projections of continuous/actual or mathematical variables to be used (Maulud and Abdulazeez, 2020). Sir Francis Galton first proposed the concept of linear regression in 1894, he used the method to measure the degree of trend (trend) in a population averange. Then in 1896, a review paper by Sir Francis Galton published "Family In Stature" he stated that there was a tendency (trend) for the average height of the population of children with parents who had a certain height to move or regress towards the average height of the entire population, (Maulud and Abdulazeez, 2020) after Sir Francis hypothesis the concept of Universal Regression Law was strengthened by Karl Pearson in the Journal of Biometrics, vol.2, 1903. In his research, the average height of the boys in the tall father group was less than their father's height, and the average height of the boys in the short father group was greater than their father's height. Regression analysis is concerned with the study of the dependence of one variable, namely the dependent variable or the dependent variable on one or more other variables, namely explanatory variables, to estimate or predict the mean or average value of the dependent variable population, considered constant (Purba and Purba, 2022). Therefore, partial correlation and regression are tests that allow scientists to understand the relationship between two variables to assess the confounding impact of the data. Evaluation methods and data modeling that build a linear relationship between these variables are dependent and independent (Tancev and Toro, 2022). The process will later draw the relationship between the dependent and independent variables from the test to predict a constant value for an increase/change.

The research process aims to calibrate the sensor used as a device voltage measurement. Then an analysis of the measurements is carried out and determines the deviation in the value indicated by the sensor with statistics data. After the sensor can do measurements, the next step is the data collection process to review the standard deviation level of each sample taken. Then determine the confidence level of the data representation. The measurement system was designed and structured with the Arduino Mega Pro microcontroller as the processing base. Data retrieval is done by wireless transmission using a Bluetooth sensor to record data using an application.

1. Methodology

The principle of using sensors in the process of measuring voltage and current depends on the formulation applied in writing the program (Junaldy, Sompie and Patras, 2019). On paper (Moradi, Akbari and Wild, 2019) confirms that the shift in the value in the inherent measurement process with the sensor reading input analog value. Every increase value in analog signal is followed by measurement value convergently due to the influence of the function of the resulting formula. However, concerns regarding the sensor's reliability are increased. So uncertainty level analysis is introduced to the sensor. In particular, raw sensor measurements that fall outside the process distribution are more difficult to identify because higher uncertainty is assigned to results that have not been observed during model training. Essentially it brings predictive maintenance integrated with the calibration without additional algorithms. Therefore, this uncertainty can be used as an indicator of the reliability of measurement results. (Tancev and Toro, 2022).

The formulation produces a constant measurement pattern to changes in sensor values. Certainly fluctuating value on readings still occurs due to the instability of the analog value (Aroulanandam *et al.*, 2022). An analytical approach using statistical methods is carried out to find out how well the uncertainty level sensor is in the reading process (Kapoor *et al.*, 2018). Read values are collected into data which will be processed in the analysis process. Elemental inputs are important in measuring the level of uncertainty to maintain acceptable levels of reliability of various configurations (An, Youn and Kim, 2022). This uncertainty factor is mapped in the value distribution graph, the percentage of error values for the data collected. The author's purpose is to find the sensor's level of accuracy by calculating and predicting the level of accuracy of the readings. The Bayesian Models approach with the Gaussian Regression Process (GPR) is also used in the data collection process. By using several approaches of course can obtained comparisons value (Allred *et al.*, 2021).

1. **Scope of Calibration Regression Analysis**

(Purba and Purba, 2022) in his writings said, in general, regression analysis is divided into three models: parametric regression analysis, non-parametric regression analysis, and semiparametric regression analysis (a combination of parametric and non-parametric regression). The main difference between parametric and non-parametric regression is that parametric regression analysis requires assumptions about both the functional form and the distribution of the residuals. However, if the assumptions cannot be met, non-parametric regression analysis can be used. Linear regression is included in the parametric model with the functional design process using the residual normality. Forecasting of the resulting function is said to be valid if the normality of the data is met (Maulud and Abdulazeez, 2020). The calibration process runs an optimization model of the parameter values to improve the fit between the observed response (measurement data) with the simulation process (Maria, Suhartanto and Fidari, 2022). The independent variable used is the analog value read by the sensor, which is susceptible to changes in the sensor value. The dependent variable is the voltage value of the measuring instrument and the independent variable is the value analog read. The process of compiling can be done with a simple equation of linear regression:

(1)

Where, = Dependent data set, = -intercept of the line, = Slope of the line, = Values of the independent data set.

(2)

(3)

Find,

(4)

The explanation behind the value is to determine the parameter estimate by taking each row from all data points (Pahuja, 2022). Residual normality plays a vital role in regression analysis. Linear residual values can be determined for y measurements and equation values, and then the residuals can be displayed (Maulud and Abdulazeez, 2020). Thus, fewer samples need to be discarded and shorter chains can be generated, resulting in faster inferences. Usually, several chains are run in parallel so that the quality of the sampling procedure can be evaluated (Tancev and Toro, 2022).

Finish with the process of forming equations that produce coefficients and constants from linear regression. The formula is applied to the program. (Amin *et al.*, 2022) explained to evaluate two probabilistic objectives, Monte Carlo Simulation (MCS) can be used to find the mean and standard deviation of the Fisher Information Matrix (FIM) determinants to model uncertainty and measurement noise. MCS can provide precise estimates of objectives when the number of sampling points to the random variable is large enough. However, the computation-intensive process is caused by a lot of deterministic function evaluation required, especially when embedded in a population-based algorithm. The use of surrogate modeling techniques is an alternative way to lighten the computational burden encountered in MCS. As an alternative (Van Der Veen, 2018) supervised learning algorithms in machine learning, the Gaussian Regression Process (GPR) is used, which is a flexible Bayesian model to make predictions in terms of average and variation over data, this makes the Bayesian model an approach the most efficient and accurate.

GPR is used as a replacement model for MCS to increase efficiency. Two types of random variables, namely flight models and noise measurements, were included in the study. Because the authors used sensors to measure the information modality of the observed structures, the dispersion can be propagated to a depth of several values, with measurements of more than 1000 samples. Considering that the FIM represents the design objectives, two models are available and can be selected when using the GPR as a model alternative to replacing the evaluation of actual functionality in the MCS (An, Youn and Kim, 2022).

First, in the initial step, build a modal linkage vector replacement model with the model carrier as a random variable, namely . After that, the noise measurement is then added to the ***q***-th, estimated modal capital, as shown below (An, Youn and Kim, 2022):

(5)

Where, denotes the vector with its elements consisting of 1, and represents the mean deviation from the actual value due to measurement noise in the ***q***-th modal vector. All model vectors approached and disturbed form a modal matrix of . At the end, the two probabilistic objectives are calculated based on the following formulation:

(6)

(7)

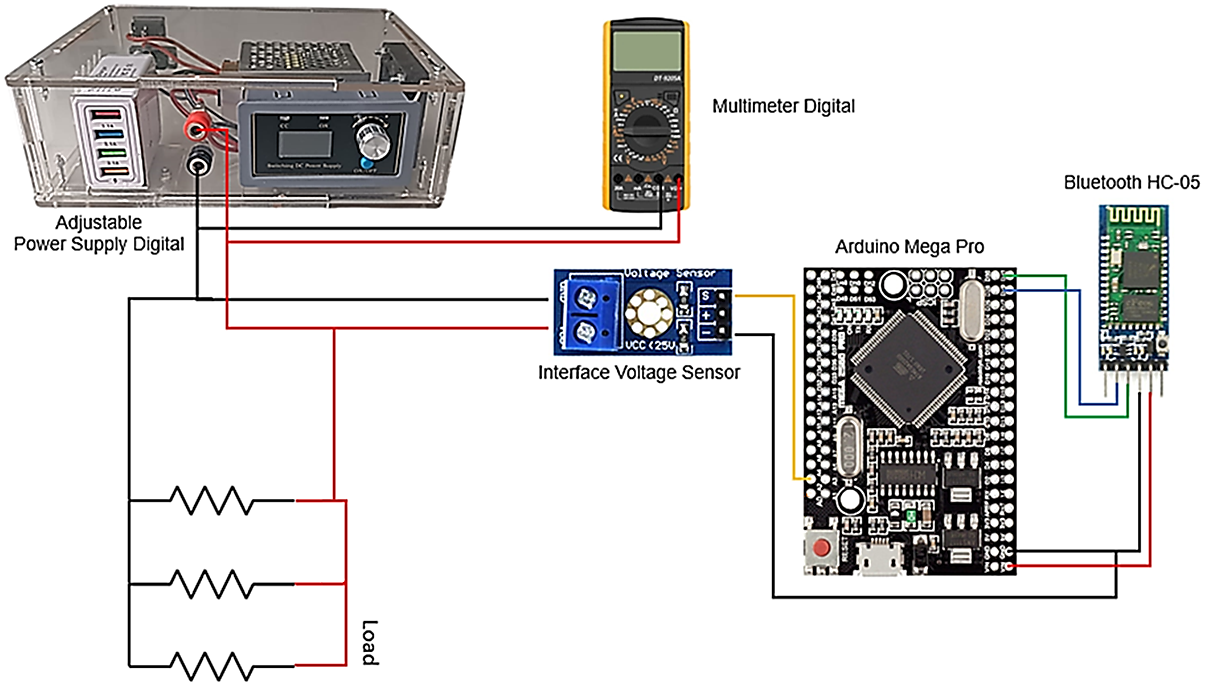
Where, is the total number of random samples used, and is the -th random sample. The second option is to directly construct a surrogate model for the determinants of FIM for measurement uncertainty and noise. states this way, based on a group of random samples for a vector of random variable , the mean and standard deviation of the determinant of FIM can be calculated using the expression:

(8)

(9)

Where, represents an estimate of the FILM determinant obtained with the GPR model.

Based on the two proposed replacement models. In the first option, the number of random variables used to construct the replacement model equals the number of model uncertainty parameters. In the second option, the number of random variables to build a replacement model is the amount of measurement noise parameters and uncertainty level (An, Youn and Kim, 2022). The decision is on the second option, where the author collects a sample of noise measurement parameters with more than 1000 samples to calculate the standard deviation and divide the level of uncertainty. Then the percentage error value is displayed for each sample of data taken.

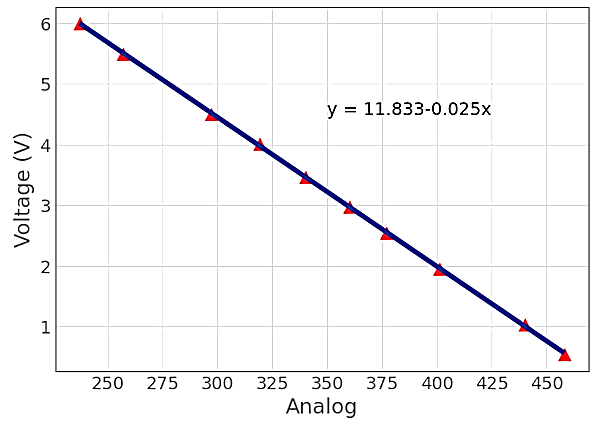


**Figure 1.** Setup Apparatus Measurement Device

1. Electrical Circuit

The work of the interfacing voltage sensor divides the voltage from the input into a voltage divider circuit of two resistors where R1 is 30K Ω and R2 is 7,5K Ω. The module can measure the voltage range between 0.02445 to 25 volts, as it is known that the measurement of the voltage value is carried out by connecting the measuring device in parallel with the voltage source. The voltage is divided into smaller quantities and then converted into analog signals (Suryawinata, Purwanti and Sunardiyo, 2017). The resistors are used in the setup device as a load where the variation is ignored, because basically the voltage can be measured without a load. The voltage is divided into smaller quantities and then converted into analog signals. The ratio of the actual voltage is limited to an analog signal.

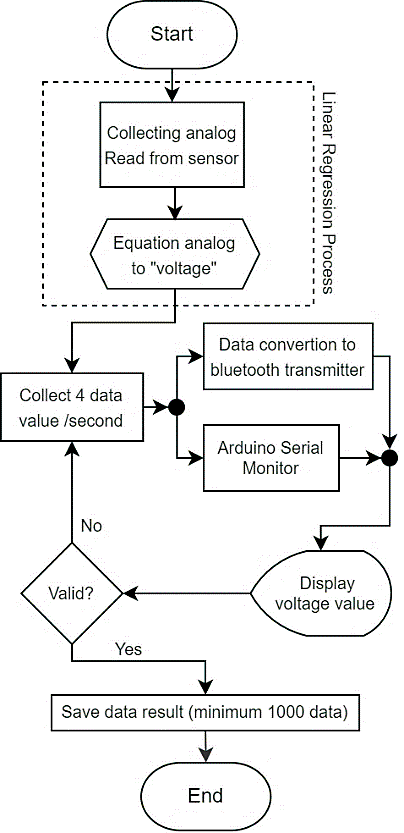
Components in the measurement system are voltage devices for power supply, sensors, and Arduino. The Bluetooth module is used for data transmission during the sampling process. However, before sampling, calibration is performed on the device. The device collects data and searches for coefficients and constants by comparing them with measuring instruments following the stipulations of measuring standards based on standard standards (Wishnu Pandu Prayudha, Fadhil and Novianto, 2022). refer purpose of regression analysis, which allows for the use of reliable results in value forecasting and prediction (Huriaty, 2015). After obtaining the two variable coefficients and the average increment constant, the value is entered with the estimated linear increase to obtain the average formulation of the analog sensor.



**Figure 2.** Sensors Linear Regression Result

The schematic in figure 1, shows the calibration device consists of the following elements: Power Supply, Potentiometer, Interfacing Voltage Sensor, and Arduino Mega Pro. The voltage measurement is taken from 6 to 0,5 volts, where the maximum voltage is limited. The calibration process does not require lots of data to get the formulation because finding coefficients and constants uses an even distribution so that the deviation value formed from the difference in data onwards is only 0,0001 (Pahuja, 2022). However, the more samples taken, the better the accuracy of the data obtained (Fachri, Sara and Away, 2015).

The comparison process begins after the equation value and obtained. Accuracy devices are tested to find error deviations and instrument accuracy is needed (Magaski and Anwari, 2022). Data processing looks for determinant standard parameters, deviation, mean, and error to the level of uncertainty. (Ceballes and Abdelkefi, 2021) explained to find precision measurement related to the concept of uncertainty. It is impossible to express the level of precision derived from a repeated measurement method using a conceptual error (error). Wrote on (Vurchio *et al.*, 2020) that possible precision quantification is using uncertainty concept. The payment method based on precision is the Standard Deviation (SD) value of the sample test used, then a minimum triplicate (3 repetitions) is performed. Using the Confident Level (CL) theory, the possible coverage obtained in the first approach is used to change the nominal interval of application.



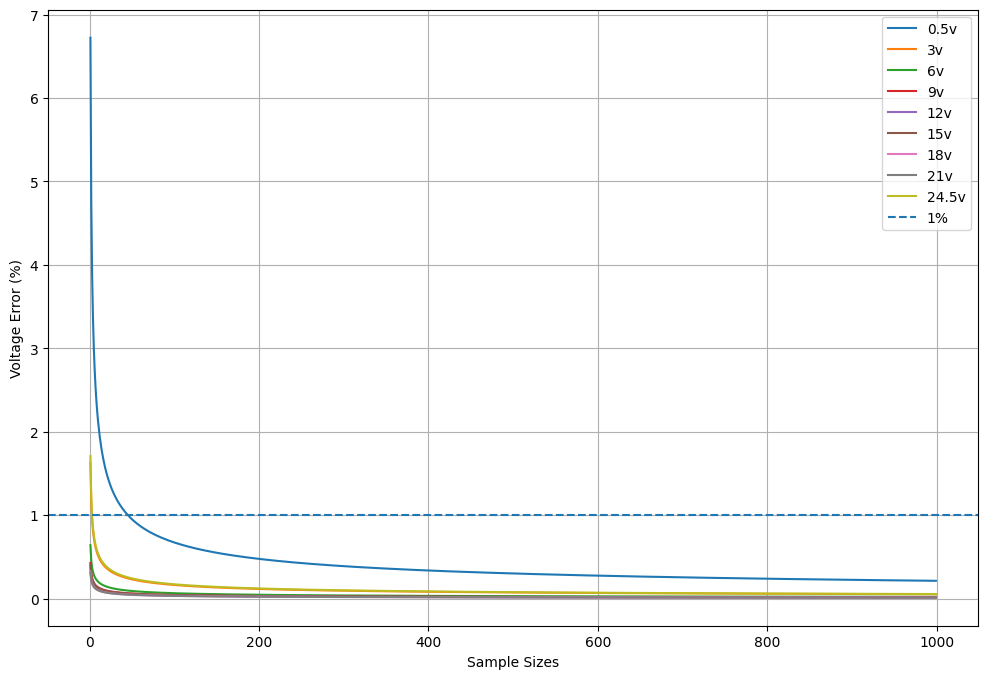
**Figure 3.** Device Algorithm Flow Chart

Measurements are made on the sensor by entering the calculation result equation. This analog data is still being converted into the equation to produce a measured voltage value. Then this value is compared with the results of direct measurements of the voltage value on the display and digital multimeter. The data from the voltage value is still experiencing fluctuations in value, but it is close to the actual measurement value. Then the reading results are collected from more than 1000 data. Uncertainty level error is used to determine how well the sensor measures. The standard deviation of 1000 data is searched to determine the uncertainty value. Readings are carried out at intervals of 4 data per second. The monitoring system directly uses the application display via a Bluetooth serial monitor application that is connected wirelessly. Sensor read form a voltage measurement is stored into a TXT file and processed as a standard deviation, Uncertainty Level error, Data Error density, and Analysis of Variance error distribution.

1. Result and Discussion

Collecting data in research is carried out using devices that give measurement real-time results. Variations of measurement value are collected to make data set. The nearest deviation is still within tolerance limits. However, if the value deviation is too far, the sensor may be experiencing interference, so it needs to be checked or repaired. After that, the sensor uncertainty level can be determined from the measurement process on the device.

To find how good sensor accuracy for instrument measurement, in the measuring process, the percentage of the standard error is calculated by varying the voltage reading value on the sensor from 0.5 V to 24.5 V with the sample. Then the measurement results are plotted in graph to see the comparison.



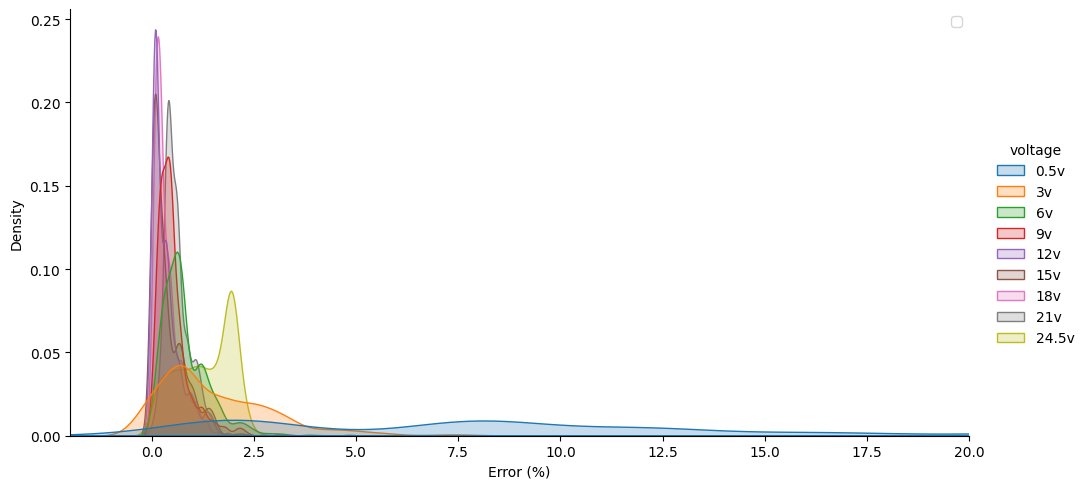
**Figure 4.** Result for Uncertainty Error Sample Measurement

Measurements were made by giving a measured voltage from the power supply. The voltage variation value was collected for about 5 minutes to produce more than 1000 samples for processing. Through calculations, the sample size is obtained when it reaches the standard error limit value of 1% for each voltage. The table below shows the standard deviation and data points intersecting with a 1% limit error.

**Table 3.** Uncertainty Error and Standard Deviation from 1000 Samples every Measurement

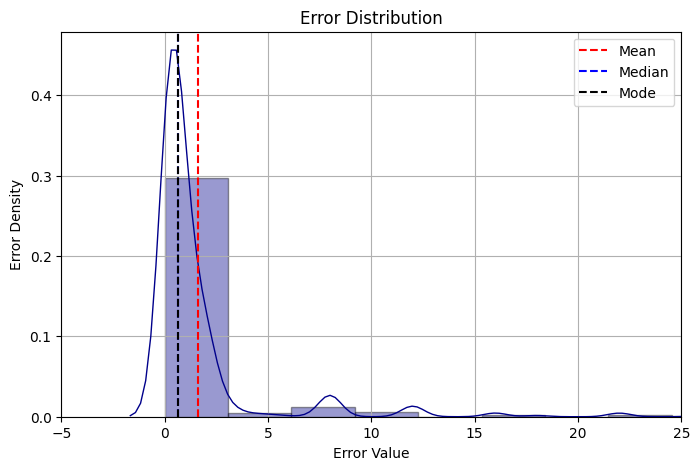
|  |  |  |
| --- | --- | --- |
| Voltage (V) | Standard  Deviation | N when  SE 1% |
| 0.5 | 6.724 | 45.208 |
| 3 | 1.626 | 2.645 |
| 6 | 0.640 | 0.409 |
| 9 | 0.427 | 0.182 |
| 12 | 0.370 | 0.136 |
| 15 | 0.407 | 0.165 |
| 18 | 0.322 | 0.104 |
| 21 | 0.312 | 0.097 |
| 24.5 | 1.708 | 2.918 |

Based on the error distribution data from each treatment given by treating voltage variations in the error measurement process, the largest iteration value required for the voltage value to reach an error of 1% is obtained in the measurement at a voltage of 0.5v, which is 45 iterations. The density distribution is carried out to see the distribution of each error at each voltage, and it is known that errors at a voltage of 0.5v have the largest standard deviation and the largest average value.



**Figure 4.** Standard Deviation Distribution from Sample Measurement

Data merging is done to get the statistical value of the population of each sample**.** Error distribution measurements were carried out for each voltage variation, and the average error value was 1.61331 and the median and mode value is 0.667 . Using the central tendency, it is known that the error distribution is a positive skewed distribution.

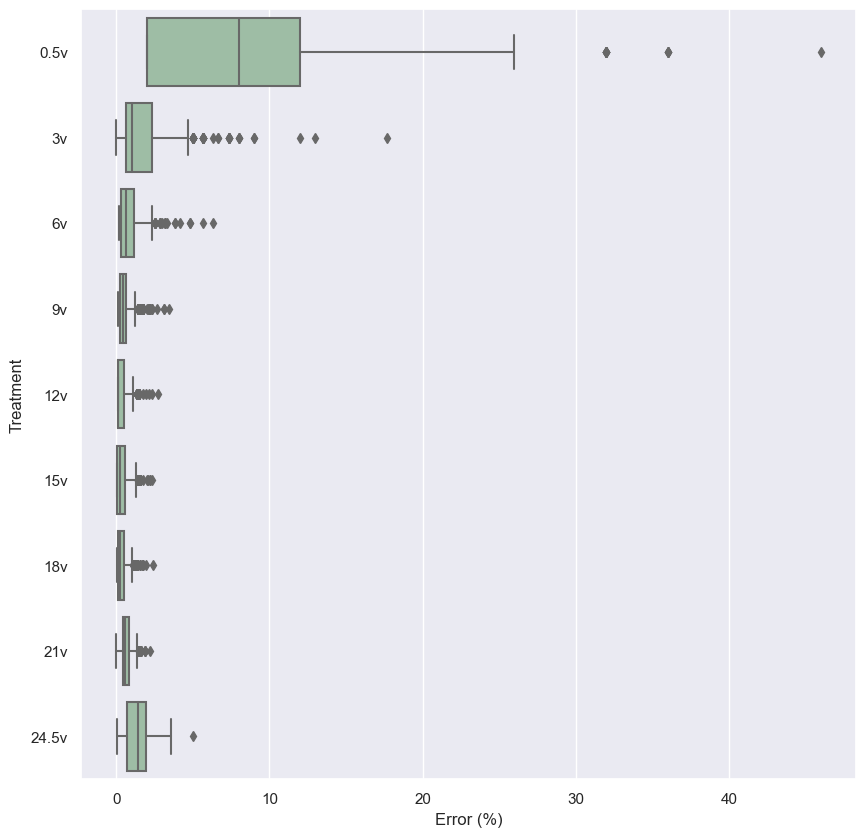


**Figure 5.** Overall Sample Error Distribution

To find out the distribution and variation and to test the differences in each data distribution, graphical visualization was carried out using boxplot charts and analysis using the Analysis of Variance (ANOVA) method. The type of ANOVA used is one-way ANOVA, and the following results are obtained. Based on the P-Value of the test, it is known that the test results reject the null hypothesis (H0) and indicate a significant difference between the treatment and accepting the alternate hypothesis (H1) at a significant value of 0.01.

**Table 3.** ANOVA result

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| DOF  Numerator | DOF  Denominator | Sum of Square | Mean of Square | F-Stat | P-Value | F-Critical |
| 8 | 8991.0 | 55711.196 | 6963.899 | 1268.461147 | 0 | 2.6416 |



Post Hoc analysis using the Tukey-HSD (Honestly Significant Difference) method was used to identify the differences between the two groups from the differences in the treatment of voltage values based on ANOVA results. Based on the calculation analysis table using Tukey-HSD, we get the difference between the treatment groups between the voltages that rejected null hypothesis and accept the alternate hypothesis is 13 pair from different treatment voltage, and the other 23 pair of voltage is accepting null hypothesis.

**Table 4.** Tukey-HSD result

|  |  |  |  |
| --- | --- | --- | --- |
| Voltage | | P-Value | Reject |
| 0.5v | 3v | 0,001 | True |
| 0.5v | 6v | 0,001 | True |
| 0.5v | 9v | 0,001 | True |
| 0.5v | 12v | 0,001 | True |
| 0.5v | 15v | 0,001 | True |
| 0.5v | 18v | 0,001 | True |
| 0.5v | 21v | 0,001 | True |
| 0.5v | 24.5v | 0,001 | True |
| 3v | 6v | 0,001 | True |
| 3v | 9v | 0,001 | True |
| 3v | 12v | 0,001 | True |
| 3v | 15v | 0,001 | True |
| 3v | 18v | 0,001 | True |
| 3v | 21v | 0,001 | True |
| 3v | 24.5v | 0,070083368 | False |
| 6v | 9v | 0,083068817 | False |
| 6v | 12v | 0,001 | True |
| 6v | 15v | 0,001 | True |
| 6v | 18v | 0,001 | True |
| 6v | 21v | 0,512962893 | False |
| 6v | 24.5v | 0,001 | True |
| 9v | 12v | 0,765369873 | False |
| 9v | 15v | 0,9 | False |
| 9v | 18v | 0,836611298 | False |
| 9v | 21v | 0,9 | False |
| 9v | 24.5v | 0,001 | True |
| 12v | 15v | 0,9 | False |
| 12v | 18v | 0,9 | False |
| 12v | 21v | 0,227134702 | False |
| 12v | 24.5v | 0,001 | True |
| 15v | 18v | 0,9 | False |
| 15v | 21v | 0,440291285 | False |
| 15v | 24.5v | 0,001 | True |
| 18v | 21v | 0,288171444 | False |
| 18v | 24.5v | 0,001 | True |
| 21v | 24.5v | 0,001 | True |

1. Conclusion

Based on the results of the design, testing, and data collection, the conclusions are as following :

1. Obtained the value of the equation from the calibration process for the current calculation of

I(x) = -7.526 + 0.015x

V(x) = 11.833 – 0.025x

The value x represents the reading value of 8-bit analog data from the microcontroller with correlation from each calibration is 0.99.

1. Melalui perhitungan standard error di ketahui error dapat mencapai nilai confidence 1% apabila dilakukan pengukuran data sebanyak 43 kali.
2. ANOVA analysis was performed to test whether there were differences in each treatment, and it was found that ANOVA rejected the null hypothesis (H0) and accepted the alternate hypothesis (H1) based on P-Value result.
3. Tukey-HSD analysis was performed to test whether there were differences in each treatment, and it was found that Tukey-HSD the result between the treatment groups voltages that rejected the null hypothesis and accepted the alternate hypothesis is 13 pairs from different treatment voltages. The other 23 pairs of voltage accept the null hypothesis. Indicates that there is some variation from several groups of voltages.
4. Awdawd