**UNCERTAINTY ANALYSIS OF VOLTAGE MEASUREMENT**

**USING ATMEGA328P MICROCONTROLLER:**

**AN ANOVA TEST APPROACH**

**James Julian1\*, Ade Fikri Fauzi2, Annastya Bagas Dewantara2, Faiz Daffa Ulhaq2 and Fitri Wahyuni1**

1 Mechanical Engineering, Universitas Pembangunan Nasional Veteran Jakarta

2 Electrical Engineering, Universitas Pembangunan Nasional Veteran Jakarta

**e-mail**: *zames@upnvj.ac.id*

**DOI: 10.14710/medstat.XX.X.XX-XX**

|  |  |
| --- | --- |
| **Article Info:**  Received:  Accepted:  Available Online:  **Keywords:**  *ATmega328p; Chauvenet; linear regression; voltage sensor; Tukey-HSD.* | **Abstract:** The application of sensors in the industry requires a high level of accuracy, and the need increases as sensors are combined with Machine Learning. This study references other works and focuses on developing formulations to reduce errors in measurement. The 0–25-volt DC interfacing voltage sensor is a choice compared to other sensors because it is low price. The sensor measurement results are then analyzed using the Analysis of Variance (ANOVA) test to demonstrate the normality of the data. Then Tukey Honestly Significant Difference (HSD) method is used to test whether voltage variations are accepted or rejected. One thousand samples of voltage data were taken for each voltage measurement of 0.5v, 3v, 6v, 9v, 12v, 15v, 18v, 21v, and 24.5v were measured using the ADC readings of the calibrated ATmega328p with a multimeter. Chauvenet criterion was used to eliminate noise-induced outlier data, with 834 outlier data out of 9000 samples calculated to have a probability above 0.5%. The relative error of all voltage variations was analyzed using ANOVA, and it was found that the mean of the voltage variations was different and rejected H0. The post hoc test using the Tukey-HSD method was used to determine the significant differences between the voltage groups. It resulted in 9 pairs of comparisons accepting the null hypothesis (H0) and 27 pairs accepting the alternate hypothesis (H1). The data was found to be normally distributed through the calculation of residual ANOVA and visualization of data with the Q-Q plot. |

1. **INTRODUCTION**

The development of electronic devices in the world industry continues to enter a new phase to achieve perfection. Electronic devices are applied in the industry mainly in measuring and automatization. Changes in the measurement process are seen when previously measurements used an analog system, now most of the tools are working digitally (Davis & Clowers, 2023). In recent years, with the development of the industrial era 5.0, sensors have been integrated with Machine Learning and even Artificial Intelligence. Sensor measurement data is taken and then processed with an embedded program for analysis or making-decision. The application of sensors in the industry requires a high level of accuracy, and the need increases as sensors are combined with Machine Learning. The higher the measurement accuracy, the higher the price of the sensor (Abubakar et al., 2017). That makes electronic device developers continue to compete in designing a sound measurement system with high accuracy and precision.

Many studies have discussed the workings and accuracy of various types of sensors. In summary, sensor performance is determined by comparing analog signal readings more stable the analog value readings, the better the sensor's measurement results (Pahuja, 2022). Using the sensor must go through calibration first to set the measurement parameters. Use accompanied by calibration is carried out using Polynomial Regression with the Arduino Mega based ZMPT101B sensor (Abubakar et al., 2017). The results indicate that sensor measurements after calibration have increased accuracy, with error values calculated at 0.9% to 2.4%. The study shows the importance of the calibration process as the resulting measurement error changes with higher accuracy.

The accuracy of measuring voltage is essential because electronic devices support most human activities. Identification of the voltage is required to provide a response ranging from extra high voltage devices to different low voltage devices. Roman in his research (Hrbac et al., 2020), built measurement devices to measure high-voltage electricity, where its use can be implemented in transformers, high-voltage equipment, and electricity distribution networks. Contrary to this, Agustina (Lascano et al., 2017) in the scope of neurology, uses evoked potential devices utilizing extra-low voltage measurements to diagnose patients.

The role of voltage measurement is essential, so further discussion must be carried out regarding the calibration process and analysis of measurement results from measuring devices or sensors. Follow-up research was carried out using statistics to determine the measurement accuracy level from variations in data on sensors by applying the Analysis of Variance (ANOVA) test (Chen et al., 2022). The purpose of using the ANOVA test is to test differences in sensor measurement variations, with normally distributed data or not. The data obtained represents that the results have significant differences, so further tests such as Tukey or Scheffe must be carried out.

This paper focuses on developing calibration to reduce measurement errors. The ANOVA test was introduced to the data obtained to evaluate the normality results of the data plots. This research evaluates the DC voltage sensor 0-25 volts using the ATmega328P microcontroller for processing. DC 0-25 volts is low-cost if compared to other sensors. Therefore the authors are interested in researching the accuracy and precision of sensor readings.

1. **LITERATURE REVIEW**
   1. **Linear Regression**

The calibration process in the regression analysis involves using an optimization model to match the linearity between the observed data of the independent variable, which is the value read by the sensor, and the dependent variable, which is the voltage from the measuring instrument (Maria et al., 2022). A linear line is drawn based on the measurement results of the digital multimeter and then intersects with the reading of the analog value on the sensor. Raw sensor measurements that fall outside the process distribution are more difficult to identify, as unobserved results during model training result in higher deviations (Tancev & Toro, 2022).

|  |  |
| --- | --- |
|  | (1) |

|  |  |
| --- | --- |
|  | (2) |

|  |  |
| --- | --- |
|  | (3) |

with

= intercept of the line

= slope of the line

= values of the independent data set

= total number of values

|  |  |
| --- | --- |
|  | (4) |

After formulation, a constant variable is obtained from the equation on the increase in the linear line. Then the equation is embedded into the Arduino ATmega328p program as the basis for processing the voltage sensor. The low error rate of measurement is obtained after the calibration process. However, it is necessary to conduct further analysis of the data to determine whether the measurements form normal variations or have significant differences in the data distribution.

Before the ANOVA test is applied in analyzing data from sensor readings, the decision to remove outliers is made to reduce noise in the data. The difference between the actual and measured values in percentage form is a relative measurement error. It is used to evaluate the accuracy of a measurement or calculation. When the measurement deviation is too far, the data cannot be classified into relative error. Chauvenet's criterion is introduced to identify data that have a probability.

* 1. **Outlier Data Classification**

Being outside of the measurement. Chauvenet controls for errors by using un-paired data type (abnormal) (Christensen, 2015), where data deviate too far from others. That can be due to noise from rapid, repeated measurement processes. Noise data falls under the category of random errors. Outlier noise data is categorized as invalid, as the values produced do not fall within the measurement distribution (Wang et al., 2018).

|  |  |
| --- | --- |
|  | (5) |

|  |  |
| --- | --- |
|  | (6) |

with

= maximum allowable deviation

= Z-score calculation

= quantile distribution

= probability represented by one tail of the distribution

= value of suspected outlier

= sample mean

= sample standard deviation

Data is classified as an outlier when ≥ is rejected or inappropriate. Where it turns out that > so that the data is classified as an outlier. If the conditions fulfill the equation ≥ accepted, the data still enter the distribution. Searching for outliers one by one will undoubtedly take quite a long time. Dmax dan zscore calculations use programming to remove applied stress data outliers so that distribution results are obtained that are ready to be used in the ANOVA test.

* 1. **Standard Deviation and Uncertainty Level**

Adjustment of the distribution by removing data that is not needed is followed by looking for the standard deviation. All standard deviations of the nine measurement variations are calculated. After getting the standard deviation and mean, the analysis continues to find the sensor's uncertainty level. This level of uncertainty can be used as an indicator of measurement reliability (Aroulanandam et al., 2022). The 95% confidence level is determined by referring to several studies. The 95% confidence model with a statistical significance of 5% is often used for data representation (Van Der Veen, 2018).

|  |  |
| --- | --- |
|  | (7) |

|  |  |
| --- | --- |
|  | (8) |

|  |  |
| --- | --- |
|  | (9) |

with

= standard deviation

= individual value

= mean value

= standard error

* 1. **Data Classification with ANOVA Test**

The use of ANOVA assumes that the data is normally distributed, and the groups being compared have the same variance. If this assumption is not met, the ANOVA results may be biased (Christensen, 2015). The type of equation used in the calculations is one-way ANOVA, with stress data used to compare variances between groups. Then from the data it can be seen the distribution between groups, whether there is a significant difference or not. The data represented on the ANOVA distribution chart has an upper and lower bound, with continuous lines to represent the abundance of the distribution in the boxed region and little data in the striped portion. The equation used to calculate ANOVA F-ratio is shown in equation (13).

|  |  |
| --- | --- |
|  | (10) |

|  |  |
| --- | --- |
|  | (11) |

|  |  |
| --- | --- |
|  | (12) |

|  |  |
| --- | --- |
|  | (13) |

with

= sum of square between group

= group average of i

= number of groups

= sum of square within groups

= variance of each group

= ratio

= mean square between group

= mean square due to error within group

* 1. **Tukey’s Honestly Significant Difference (HSD)**

Tukey's method is then used as a companion to the ANOVA distribution plot. Tukey's HSD is applied to mathematically identify the plot against the results of the ANOVA distribution. This method is used to determine the hypothesis for nine variations of the data, classifying them as accepted or rejected. The equation for Tukey's HSD is shown in equation (14).

|  |  |
| --- | --- |
|  | (14) |

with

= mean square for within group from the ANOVA

= standardize range statistic

= number in each group

With is the average value of the data distribution is not significantly different, is the average value of the data distribution is significantly different. and describe conditions that can be seen from the data representation. To ensure the validity of the data, comparisons were made between data with each set. Determined using a p-value > 0.05.

* 1. **Quantile-quantile (Q-Q) Plot Graph**

The Q-Q plot is used to assess the normality of the deviation level of the data set by plotting the data being tested against the quantiles of the normal distribution (Christensen, 2015). In creating a Q-Q plot, the data is first sorted from the lowest to the highest value. The data are represented as points in actual data, and these points will remain on the line formed by the z-score of each data. If the data is normally distributed, the points representing the data will follow a linear line and not deviate significantly. Conversely, if the data is distributed chaotically or randomly, the data could be better in terms of order and distribution.

1. **MATERIAL AND METHOD**

Research equipment was prepared and designed as a measurement setup to support data collection with Arduino devices as shown in Figure 1. Voltage readings are taken using an adjustable power supply as the measurement source, with the supply terminals connected in parallel to the measuring instrument and the device under test. Sensor measurements are also displayed on the OLED display in real time, providing a clear and live view of any changes in measurement values.

|  |
| --- |
|  |
| **Figure 1**. Setup Apparatus Voltage Measurement |

The research method was carried out in several steps as shown in Figure 2. The measurement value is taken with a multimeter as a readable instrument, with an assumed error of ±1-3%. The processing base uses an Arduino Nano ATmega328p 10-bit microcontroller. As stated in Moradi's paper (Moradi et al., 2019), it is confirmed that the sensor's value can shift during the measurement process as it reads the analog input value. A converging measurement value follows each increase in the analog signal value due to the influence of the formula. The use of a customized formula can increase the sensor's reliability. Then, Linear Regression is introduced in the sensor calibration process. Data collection was performed by dividing the data into different voltage values. To eliminate outliers, a filtering process using the Chauvenet criterion was implemented. A one-way ANOVA was conducted to investigate the differences in variations among the different voltage values. If the ANOVA test rejected the null hypothesis, post-hoc analysis using Tukey-HSD was performed. The final analysis included computation of descriptive statistics such as the mean, standard deviation, standard error, and uncertainty value.

* 1. **Microcontroller**

The microcontroller device processes the interfacing voltage sensor data and works with a voltage divider circuit. The circuit consists of two resistors: Resistor number one (R1) is 30K Ω, and resistor number two (R2) is 7.5K Ω. The interfacing voltage module can measure a voltage range of 0-25 volts. It is known that the measurement of the voltage value is carried out by connecting the measuring instrument in parallel with the voltage source. The voltage is divided into smaller quantities, then converted into analog signals (Junaldy et al., 2019). Then, the rheostat can be neglected because the voltage still can be measured without a load.

The processing speed of the Arduino ATmega328p is 16 MHz, and it has 20 operating pins, including 6 analog pins, 6 Pulse with Modulation (PWM) outputs, and 8 input and output (I/O) pins. Equipped with a 10-bit ADC that can convert analog to digital signals with 10-bit resolution (Debnath et al., 2022). Measurements can be made using a 10-bit ATmega328p by converting the analog signal to be measured into a digital signal. The analog signal come from various sources. After the analog signal is converted into a digital signal, it can be processed by the ATmega328p microcontroller using its program (Zhang et al., 2018). The ATmega328p divides the input voltage into 10-bit or 1023 decimal resolution, with 5 volts representing 1023 and 0 volts representing 0 decimal at V-reference 5V. The data collection technique was carried out by collecting nine sample variations ranging from 0-25 volts, with 1000 data from each variation used to calculate the standard deviation and divide the level of uncertainty.

|  |
| --- |
|  |
| **Figure 2**. Research Process Flowchart |

The research process on sensor measurement results is divided into two. The first stage is the pre-process, which includes sensor calibration with linear regression to find a digital value multiplier formula, then converted into a voltage quantity. The second stage is the analysis process, managing data that has been measured or retrieved by sensors to be processed into diagrams representing data. The final result, as a process of analysis and discussion, describes the data management process and determines whether the quality of the sensor is feasible or not from the data produced.

1. **RESULTS AND DISCUSSION**

Linear regression is applied using equation (4) to determine the coefficient and constant. In the context of calibrating a voltage sensor with the ATmega328p microcontroller's ADC, linear regression is used to find the best straight-line fit to the data points obtained by measuring the sensor's output voltage and comparing it to the ADC reading. The result obtained after calculation using linear regression is shown in table 1, shows more accurate measurements using oversampling and linear regression techniques with lower error results, namely 0.32% higher than measurements without linear regression calibration and oversampling techniques.

Accurate data collection from sensors is essential in order to obtain accurate results. It can be done by appropriately setting the measurement frequency, sample size, and the type of sensor used based on the need. In this case, data sampling is performed by varying the voltage on the sensor with each data collection of 1000 samples. The voltage variations given are 3v, 6v, 9v, 12v, 15v, 18v, 21v, and 24.5v, with the following description of the data is showed in table 2.

|  |
| --- |
| Chart, line chart  Description automatically generated |
| **Figure 3**. Linear Regression between ATmega328p Analog Readings to Voltage Reading |

|  |  |
| --- | --- |
|  |  |
| (a) | (b) |
| **Figure 4**. (a) Relative error (%) distribution for each sample,  (b) Relative error (%) distribution for population | |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 1**. Data Summary | | | | | | | | | |
|  | 0.5v | 3v | 6v | 9v | 12v | 15v | 18v | 21v | 24.5v |
| Sample size | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 |
| Mean | 0,50 | 3,02 | 6,00 | 9,00 | 12,00 | 14,99 | 18,09 | 21,02 | 24,52 |
| Std | 0,023 | 0,016 | 0,017 | 0,016 | 0,018 | 0,018 | 0,018 | 0,02 | 0,018 |
| Maximum data | 0,38 | 2,94 | 5,92 | 8,94 | 11,92 | 14,87 | 17,94 | 20,96 | 24,45 |
| Q1 | 0,5 | 3,02 | 5,99 | 8,99 | 11,99 | 14,99 | 18,01 | 21,01 | 24,5 |
| Q2 | 0,5 | 3,02 | 6,01 | 9,01 | 11,99 | 14,99 | 18,01 | 21,01 | 24,52 |
| Q3 | 0,53 | 3,04 | 6,01 | 9,01 | 12,01 | 15,01 | 18,04 | 21,04 | 24,52 |
| Minimum data | 0,6 | 3,06 | 6,06 | 9,06 | 12,06 | 15,06 | 18,09 | 21,09 | 24,57 |

|  |
| --- |
| Diagram, engineering drawing  Description automatically generated |
| **Figure 5**. Standardized Relative Error (%) of each Samples |

In the above boxplot graph in Figure 4(a), it is known that the voltage variation of 0.5v has the largest spread compared to other data, affecting the overall population distribution result as shown in Figure 4(b). The variation and deviation of this 0.5v voltage are caused by the floating pin effect of the ATmega328p, which reads electromagnetic signal noise inference inducing voltage and causing readings on the microcontroller.

The Gaussian distribution consists of a set of data that forms a normal curve with the mean as the peak and the standard deviation as the curve's width. Each data is analysed for distribution and variation to determine the sensor's characteristics. A histogram of the data volume is also displayed to represent the volume of the distribution fraction. They are using equation (8) to find the standard deviation. The distribution of the data that has been standardized using the Gauss distribution and histogram is visualized in Figure 5.

Data filtering is performed using data that is converted into standardized population data. The Chauvenet Criterion is used to eliminate data probability of less than 0.5%, which is considered an outlier and must be removed. Calculate this probability. The mean and standard deviation of the data must first be determined. Then, the probability of each data is calculated using the normal distribution formula. Data with a probability less than 0.5% is considered an outlier and must be removed from the data refer to equation (5), with the result shown in table 2.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 2**. Chauvenet Criterion Result | | | | |
|  | Iteration - 1 | Iteration - 2 | Iteration - 3 | Iteration - 4 |
| Mean | 0.32 | 0.09 | 0.13 | 0.08 |
| Std | 1.67 | 0.56 | 0.34 | 0.21 |
| Outlier | 386 | 127 | 403 | 18 |

A one-way ANOVA, using equation (12), can be performed to determine if there is a significant difference between groups in the relative error of the treatment data for the dependent variable. If significant differences are found, it can be concluded that the independent variable influences the dependent variable. Conversely, if no significant differences are found, it can be concluded that the independent variable does not affect the dependent variable.

In a one-way ANOVA, the grand mean represents the overall mean of all the data points in all the groups being compared. It is calculated by taking the average of all the data points in all the groups. The mean of each group, on the other hand, represents the average of all the data points in that specific group. The grand mean is used to compare the means of each group to the overall mean of all the data points. The difference between the mean of each group and the grand mean can give an indication of how the groups differ from the overall population. If the means of the groups are similar to the grand mean, it suggests that the groups are similar to the overall population, whereas if the means of the groups are significantly different from the grand mean, it suggests that the groups are different from the overall population, the mean of each group is shown by boxplot graph with grand mean is shown by red line in Figure 6.

|  |
| --- |
|  |
| **Figure 6**. Relative error (%) from each sample |

The value of α = 0.05 is used in determining the p-value threshold of ANOVA. The calculation results shown in table 4 show that the ANOVA test rejects H0 and accepts H1. That explains a significant difference between the treatment data groups on the dependent variable.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 3**. ANOVA One Way Result | | | | |
|  | Sum of Square | Degree of Freedom | F | PR(>F) |
| Treatment | 977,53 | 8 | 152,04 | 0 |
| Residual | 6548,50 | 8148 |  |  |

The Tukey HSD method compares each treatment by making 36 pairwise data comparisons using equation (13). Of these, nine paired comparisons reject H1 and accept H0, and 27 paired comparisons accept the H1 and reject H0. Thus, the determination of using H1 is more dominant as shown in table 4.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 4**. Tukey’s HSD Result | | | | | | | |
| Voltage (V) | | Diff | Lower | Upper | q-value | p-value | H0 |
| 0.5 | 3 | 0,55 | 0,52 | 0,58 | 81,74 | 0,00 | Rejected |
| 0.5 | 6 | 0,02 | -0,01 | 0,04 | 3,00 | 0,46 | Accepted |
| 0.5 | 9 | 0,01 | -0,01 | 0,04 | 2,48 | 0,68 | Accepted |
| 0.5 | 12 | 0,00 | -0,02 | 0,03 | 0,38 | 0,90 | Accepted |
| 0.5 | 15 | 0,01 | -0,01 | 0,04 | 1,98 | 0,90 | Accepted |
| 0.5 | 18 | 0,01 | 0,07 | 0,12 | 16,16 | 0,00 | Rejected |
| 0.5 | 21 | 0,11 | 0,08 | 0,13 | 17,43 | 0,00 | Rejected |
| 0.5 | 24,5 | 0,07 | 0,04 | 0,01 | 11,69 | 0,00 | Rejected |
| 3 | 6 | 0,54 | 0,51 | 0,56 | 91,33 | 0,00 | Rejected |
| 3 | 9 | 0,54 | 0,51 | 0,56 | 92,03 | 0,00 | Rejected |
| 3 | 12 | 0,55 | 0,53 | 0,58 | 94,30 | 0,00 | Rejected |
| 3 | 15 | 0,57 | 0,54 | 0,59 | 96,74 | 0,00 | Rejected |
| 3 | 18 | 0,46 | 0,43 | 0,48 | 77,96 | 0,00 | Rejected |
| 3 | 21 | 0,45 | 0,42 | 0,47 | 76,65 | 0,00 | Rejected |
| 3 | 24,5 | 0,48 | 0,46 | 0,51 | 82,61 | 0,00 | Rejected |
| 6 | 9 | 0,00 | -0,02 | 0,02 | 0,63 | 0,9 | Accepted |
| 6 | 12 | 0,01 | -0,01 | 0,04 | 3,16 | 0,38 | Accepted |
| 6 | 15 | 0,03 | 0,01 | 0,05 | 6,02 | 0,00 | Rejected |
| 6 | 18 | 0,08 | 0,06 | 0,10 | 15,86 | 0,00 | Rejected |
| 6 | 21 | 0,09 | 0,06 | 0,11 | 17,38 | 0,00 | Rejected |
| 6 | 24,5 | 0,05 | 0,03 | 0,07 | 10,47 | 0,00 | Rejected |
| 9 | 12 | 0,01 | -0,01 | 0,03 | 2,54 | 0,66 | Accepted |
| 9 | 15 | 0,03 | 0,00 | 0,05 | 5,40 | 0,00 | Rejected |
| 9 | 18 | 0,08 | 0,06 | 0,10 | 16,53 | 0,00 | Rejected |
| 9 | 21 | 0,09 | 0,07 | 0,11 | 18,06 | 0,00 | Rejected |
| 9 | 24,5 | 0,06 | 0,03 | 0,08 | 11,13 | 0,00 | Rejected |
| 12 | 15 | 0,01 | -0,01 | 0,04 | 2,86 | 0,52 | Accepted |
| 12 | 18 | 0,10 | 0,07 | 0,12 | 19,10 | 0,00 | Rejected |
| 12 | 21 | 0,10 | 0,08 | 0,13 | 20,62 | 0,00 | Rejected |
| 12 | 24,5 | 0,07 | 0,05 | 0,09 | 13,69 | 0,00 | Rejected |
| 15 | 18 | 0,11 | 0,09 | 0,13 | 21,95 | 0,00 | Rejected |
| 15 | 21 | 0,12 | 0,10 | 0,14 | 23,48 | 0,00 | Rejected |
| 15 | 24,5 | 0,08 | 0,06 | 0,10 | 16,55 | 0,00 | Rejected |
| 18 | 21 | 0,01 | -0,01 | 0,03 | 1,53 | 0,9 | Accepted |
| 18 | 24,5 | 0,03 | 0,00 | 0,05 | 5,41 | 0,00 | Rejected |
| 21 | 24,5 | 0,03 | 0,01 | 0,06 | 6,94 | 0,00 | Rejected |

Through filtering using Chauvenet, four iterations were obtained with a total of 843 data discarded. Analysis using Q-Q plot is performed to determine whether the data distribution is normally distributed or not by comparing the data distribution to the diagonal line. Through Q-Q plot and probability density chart, the normal distribution of residual ANOVA data is obtained as shown in Figure 7(a) and 7(b).

|  |  |
| --- | --- |
|  |  |
| (a) | (b) |
| **Figure 7**. (a) Q-Q plot from residual ANOVA,  (b) Probability density of residual ANOVA | |

The outlier spread on the Q-Q plot is caused by the 0.5v variation that affects the Grand Mean value of the ANOVA residual and the standard deviation of the Chauvenet filtering process. This variation causes non-uniform outlier removal, resulting in outliers on the Q-Q plot. Standard error is used to describe the extent of the spread of voltage data relative to the population mean and to estimate the likelihood of the sample distribution relative to the population distribution, and uncertainty describes the extent of uncertainty in a measurement or statistical test of voltage sensor data acquisition with the data shown in table 5.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 5**. Final Result | | | | | | | | | |
|  | **0.5V** | **3V** | **6V** | **9V** | **12V** | **15V** | **18V** | **21V** | **24.5v** |
| Mean | 0.50 | 3.01 | 6.00 | 9.00 | 12.00 | 15.00 | 18.01 | 21.02 | 24.52 |
| Standard Deviation | 0.00 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| Standard Error | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Voltage | 0.50  ±  0.00 | 3.01  ±  0.01 | 6.00  ±  0.01 | 9.00  ±  0.01 | 12.00  ±  0.01 | 15.00  ±  0.01 | 18.01  ±  0.02 | 21.02  ±  0.02 | 24.52  ±  0.01 |

Based on the final results of the nine variations, the range of error variations at a value of 0.5 volts is 0. That is due to the excessive data loss caused by the filtering, resulting in the final results only considering data classification at the precise value of 0.50 volts.

1. **CONCLUSION**

The R2 value obtained from the linear regression analysis shows that the measurement results have a high enough correlation with a value greater than 0.99. This shows that the linear regression model can explain the existing data variations well. The filtering process using the Chauvenet Criterion shows 9.3% or 843 data outliers caused by noise. This outlier data is dominated by the 0.5v treatment data, which has the greatest variation from the other treatments. The largest data variation occurred at a voltage of 0.5v, which was caused by the floating pin effect of the ATmega328p, which can read electromagnetic signal noise inference inducing voltage and resulting in readings on the microcontroller. Variation analysis was performed using the ANOVA one way method and Tukey HSD, and the results showed that the data rejected H0 and accepted the H1 with a total of 36 pairs of Tukey HSD data comparisons. Of these, 9 pairs of comparisons rejected H1 and accepted H0, while the other 27 pairs of comparisons accepted H1 and rejected H0. Based on the research that has been done, it can be concluded that measurements using a voltage sensor with an ATmega328p microcontroller can be used to measure voltages from 3v to 24.5v.

**REFERENCES**

Abubakar, I., Khalid, S. N., Mustafa, M. W., Shareef, H., & Mustapha, M. (2017). Calibration of ZMPT101B voltage sensor module using polynomial regression for accurate load monitoring. *ARPN Journal of Engineering and Applied Sciences*, *12*(4), 1076–1084.

Aroulanandam, V. V., Sherubha, P., Lalitha, K., Hymavathi, P., & Thiagarajan, R. (2022). Sensor data fusion for optimal robotic navigation using regression based on an IOT system. *Measurement: Sensors*, 100598.

Chen, W. H., Carrera Uribe, M., Kwon, E. E., Lin, K. Y. A., Park, Y. K., Ding, L., & Saw, L. H. (2022). A comprehensive review of thermoelectric generation optimization by statistical approach: Taguchi method, analysis of variance (ANOVA), and response surface methodology (RSM). *Renewable and Sustainable Energy Reviews*, *169*.

Christensen, R. (2015). Analysis of variance, design, and regression: Linear modeling for unbalanced data, second edition. In *Analysis of Variance, Design, and Regression: Linear Modeling for Unbalanced Data, Second Edition*.

Davis, E. J., & Clowers, B. H. (2023). Low-cost Arduino controlled dual-polarity high voltage power supply. *HardwareX*, *13*.

Debnath, D., Baidya, B., & Ghosh, S. (2022). Performance analysis of Arduino processor based hybrid power generation system (HPGS). *Materials Today: Proceedings*.

Hrbac, R., Kolar, V., Bartlomiejczyk, M., Mlcak, T., Orsag, P., & Vanc, J. (2020). A development of a capacitive voltage divider for high voltage measurement as part of a combined current and voltage sensor. *Elektronika Ir Elektrotechnika*, *26*(4), 25–31.

Junaldy, M., Sompie, S. R., & Patras, L. S. (2019). Rancang Bangun Alat Pemantau Arus Dan Tegangan Di Sistem Panel Surya Berbasis Arduino Uno. *Jurnal Teknik Elektro Dan Komputer*, *8*(1), 9–14.

Lascano, A. M., Lalive, P. H., Hardmeier, M., Fuhr, P., & Seeck, M. (2017). Clinical evoked potentials in neurology: A review of techniques and indications. *Journal of Neurology, Neurosurgery and Psychiatry*, *88*(8), 688–696.

Maria, A., Suhartanto, E., & Fidari, J. S. (2022). Validasi Data Curah Hujan Satelit dengan Data Stasiun Hujan di DAS Sadar, Kabupaten Mojokerto. *Jurnal Teknologi Dan Rekayasa Sumber Daya Air*, *2*(2), 367–375.

Moradi, V., Akbari, M., & Wild, P. (2019). A fluorescence-based pH sensor with microfluidic mixing and fiber optic detection for wide range pH measurements. *Sensors and Actuators, A: Physical*, *297*, 111507.

Pahuja, R. (2022). Development of semi-automatic recalibration sytem and curve-fit models for smart soil moisture sensor. *Measurement: Journal of the International Measurement Confederation*, *203*(August), 111907.

Tancev, G., & Toro, F. G. (2022). Variational Bayesian calibration of low-cost gas sensor systems in air quality monitoring. *Measurement: Sensors*, *19*, 100365.

Van Der Veen, A. M. H. (2018). Bayesian methods for type A evaluation of standard uncertainty. *Metrologia*, *55*(5), 670–684.

Wang, C., Caja, J., & Gómez, E. (2018). Comparison of methods for outlier identi fi cation in surface characterization. *Measurement*, *117*(July 2017), 312–325.

Zhang, C. G., Chen, C. J., Settu, K., & Liu, J. T. (2018). Angle-Scanning surface plasmon resonance system with 3D Printed components for biorecognition investigation. *Advances in Condensed Matter Physics*, *2018*.