Evaluating Sensor-Derived Data Quality for IoT-based Temperature Monitoring

Aissa Bensattalah[1,*] and Youcef Belhadji[2]

[1] Department of Science and Technology, University of Tiaret, Tiaret, 14000, Algeria. Laboratoire des Méthodes de Conception des Systèmes (LMCS), 16000 Oued-Smar, Algiers, Algeria a_bensattalah@esi.dz

[2] Department of Eelctrical Engineering, University of Tiaret ,Tiaret, 14000, Algeria youcef.belhadji@uni-tiaret.dz

[*] Corresponding author

Abstract. IoT sensors undergo substantial fluctuations in their conditions, encompassing events of connectivity, disconnection, and alterations in environmental parameters. Within the scope of this paper, we introduce an experimental methodology to optimize the data quality of a temperature measurement and control system. To achieve the aim of the study, we employed a set of essential hardware components for data acquisition and processing. The integration comprised two types of temperature sensors of heterogeneous technologies: Dallas DS18B20, operating as a digital sensor, and LM35, used as an analog sensor. The measurement procedure encompasses two scenarios: simple tests involving individual sensor measurements and multiple tests entailing concurrent measurement using a group of three sensors of the same technology. The tests are made under ambient temperature and under heat source then cold environment (refrigerator). Applying a descriptive statistical approach, we computed the mean, variance, and standard deviation to assess the data quality of the system. This assessment aimed to gauge accuracy and completeness, identify variations, and comprehend implications. We also extract critical insights regarding the error and performance of both sensors within the examined operational conditions. The results show that DS18B20 present more accuracy and completeness than LM35.

Keywords: IoT · statistical descriptive · data quality · monitoring.

1 Introduction

Technological advancement involves utilizing scientific knowledge to develop tools and applications that enhance various aspects of our daily lives. One significant progression is the emergence of communication technologies, notably exemplified by the Internet of Things (IoT). The IoT field has seen substantial development, becoming a pivotal area with the potential to greatly improve everyday life. IoT enables the control of homes and associated devices through a single interface, a capability that was once unimaginable. However, a large segment of the population still lacks a clear understanding of IoT's various aspects.

C. A. Kerrache et al. (eds.), *Proceedings of the International Conference on Emerging Intelligent Systems for Sustainable Development (ICEIS 2024)*, Advances in Intelligent Systems Research 184,
https://doi.org/10.2991/978-94-6463-496-9_9

Many users remain unaware of its underlying mechanisms, characteristics, and the diverse fields of application it encompasses.

Fundamentally, IoT systems are built upon a framework comprising essential components that enable connectivity and operational functionalities among interconnected devices. As discussed in the works of Jara, Zamora, and Skarmeta [1], as well as Perera and Zaslavsky [2], sensors play a critical role in acquiring physical data. With the exponential growth in data collection and generation driven by IoT, the importance of data quality has become increasingly pronounced. As the number of interconnected devices increases, the volume of data exchange in IoT reaches unprecedented levels. The significance of accumulating big data in IoT is closely tied to data quality, which is intrinsically linked to data quantity. As data volumes expand, ensuring reliable, accurate, and pertinent information is paramount for effective applications and decision-making [3]. Data quality substantially influences model outcomes and business decisions [4], as poor data quality can lead to errors, misguided decisions, and a loss of trust in IoT systems [5], [6]. Therefore, evaluating, ensuring, and enhancing IoT data quality is critical, with key dimensions including accuracy, completeness, and security [7].

In the context of environmental monitoring, dealing with missing data due to sensor issues or network congestion is common [8]. Various methods, including statistical approaches, are employed to address missing data through interpolation [9]-[13]. This study focuses on assessing data quality in an IoT temperature monitoring system, acknowledging its significance across various domains such as environmental and industrial monitoring [14]-[18].

2 Data Quality Assessment Methods

Data Quality Assessment (DQA) is a scientific and statistical procedure aimed at determining whether data adheres to the established quality standards for projects or business operations. It involves the systematic evaluation of data against predefined quality dimensions. In the context of IoT systems, data quality is of paramount importance. The vast volume and diversity of data generated by interconnected devices pose significant challenges in terms of management and quality. IoT data originates from various sources, such as sensors, mobile devices, and smart devices, which differ in format, structure, and accuracy [19]. The distributed nature of IoT, involving numerous interconnected devices and sensors, further complicates the task of ensuring data quality.

Sensors can experience technical issues, interferences, or malfunctions, resulting in inaccurate or compromised data. Additionally, the diversity of data sources can complicate efforts to maintain the consistency and integrity of collected data. Therefore, DQ in IoT is a critical concern for businesses, researchers, and end-users. Ensuring high data quality is essential for the reliability and accuracy of information used in applications and decision-making processes [3]. The importance lies in developing methodologies and technological frameworks to systematically assess, guarantee, and enhance data quality within IoT systems. This requires careful consideration of various dimensions and complexities inherent to the domain, including accuracy, consistency, integrity, relevance, reliability, and availability [7].

However, several challenges can impact DQ, such as the volume of data, heterogeneity, temporal variability, security, and privacy. Various strategies are employed to identify and quantify errors in sensor data within IoT contexts. These strategies include: Calibration [20], Provenance verification [4], Continuous monitoring [21], User feedback [22] Statistical methods [23], Data fusion Machine learning [24], Fuzzy expert systems [25]. These strategies can be used individually or in combination, depending on the requirements of each IoT application. This study pioneers the comparative evaluation of digital and analog temperature sensors within an IoT framework, employing a dual-technology approach rarely explored in existing literature. By implementing real-time data collection using Arduino and ESP32 platforms and assessing the quality of the collected data, this research offers a novel, practical, and scalable solution for continuous temperature monitoring. The detailed statistical analysis applied in this study provides deeper insights into sensor performance, setting a new standard for data quality evaluation in IoT-based systems.

2.1 Statistical Approaches for Data Quality Assessment

The chosen method for assessing the quality of the studied IoT system is descriptive statistical approach. This approach serves a dual purpose: On the one hand extract valuable insights from the data collected and on the other hand identify potential quality problems. The approach enables informed decision-making and the implementation of measures to enhance IoT data quality. Statistical analysis involves approaches that assess data coherence, distribution, correlation. It also encompasses conformity with specific models or expectations, and detection of outliers, inconsistencies, and unexpected trends. Various statistical analysis approaches are employed for evaluating IoT data quality. These include descriptive analysis, which calculates statistical metrics like mean, median, variance and standard deviation [26]-[28]. Others approaches encompass: correlation analysis[29], regression analysis [30], estimation methods [31], hypothesis testing [32], Student's t-test [33], chi-square test [34], Fisher's exact test [35]. In our work, we applied statistical descriptive approach. It is considered as a powerful approach for analyzing and summarizing IoT data, enabling deeper insights into patterns and trends [36], [37]. It also has been used for data exploration and understanding in IoT integrated with cloud computing [38]. Recent studies in smart farming and network intrusion detection in IoT have highlighted the effectiveness of descriptive analysis [39]-[41]. In the context of our study, the descriptive statistical approach was chosen over alternative methods for several reasons. Firstly, it provides immediate and easily interpretable insights into the data, which are crucial for initial assessments of accuracy and reliability. Secondly, it serves as a foundational step, ensuring that basic assumptions about the data are validated before proceeding to more complex analyses. Thirdly, its direct measurement of key data quality dimensions aligns well with the objectives of this study. While advanced methods like machine learning or data fusion offer deeper insights, they require larger datasets, more computational resources, and sophisticated expertise, which may not be feasible in the initial stages of data quality assessment. Indeed, variance, mean, and standard deviation are fundamental statistical measures used in descriptive statistical approach for IoT data quality assessment. The mean of a dataset provides an indication of the central tendency of the data where the variance measures the spread or dispersion of data points around the mean. For the standard deviation is the square root of the variance that provides a measure of the typical amount of deviation from the mean. Thus, it gives insight into the magnitude of the accidental or environmental error.

$$Varianace = \frac{1}{n}\sum_{i=1}^{n}(m_i - \bar{m})^2 \tag{1}$$

$$Standarddeviation = \sqrt{Variance} \tag{2}$$

Where $n_{\text{ß}}$ are the value of the ith measure, $\bar{m}$:is the mean of the data set of measure and n: represents the number of the measure.

2.2 Case Study: IoT System for Temperature Monitoring

Experimentation plays an essential role in evaluating different scenarios and measuring the impact of parameters on outcomes. In recent years, the integration of IoT technology into environmental monitoring has revolutionized data collection and analysis methods. The use of microcontrollers such as ESP32 is one of the devices to help develop IoT systems and improve their performance. This approach allows real-world testing on embedded devices, data collection in real conditions, and evaluation of proposed solutions. The benefits provided by IoT technology significantly enhance the efficiency, accuracy, and scalability of such experiments. Utilizing IoT technology not only boosts the quality and reliability of data collection but also equips researchers with powerful tools for remote monitoring, making it a crucial asset for modern temperature monitoring experiments. In the given experiment, IoT technology facilitates the collection temperature data from both digital and analog sensors. This collection allows for a thorough assessment of sensor performance under various conditions. In this study, we propose an experiment using Arduino and ESP32 boards to control and collect data for a temperature measurement and monitoring system. The selection of the ESP32 for concurrent sensor operations is motivated by its cost-effectiveness, ease of use, versatility, and adequate performance for standard IoT applications. Although FPGAs and other SoCs present distinct advantages in terms of high performance, complexity, and suitability for specialized applications, the ESP32 offers a more practical and straightforward solution for experiments centered on temperature monitoring and similar IoT use cases. This involved utilizing essential hardware components for data acquisition and processing. To achieve this, two temperature sensors with distinct technologies were employed: the DS18B20, a digital sensor, and the LM35, an analog sensor. Both sensors operate within the same accuracy range of ±0.5°C, enabling continuous real-time temperature monitoring. The selection of DS18B20 and LM35 sensors in this experiment is justified by their distinct technologies, facilitating a comprehensive performance comparison between digital and analog sensors. Employing both types enables an assessment of data quality across diverse sensor technologies, ensuring a robust analysis. Additionally, their compatibility with common microcontrollers and ease of integration facilitate straightforward data acquisition and processing, making them relevant and strategic choices for this study. To exchange data with Arduino and connected Sensors, the CoolTerm app was exploited as a simple serial port communication. The use of ESP32 devkit board facilitated internet connectivity via http protocol using CoolTerm software. The incorporation of web server into the Arduino code for ESP32, collected data could be displayed, stored in flash memory, and interacted with through a web interface.

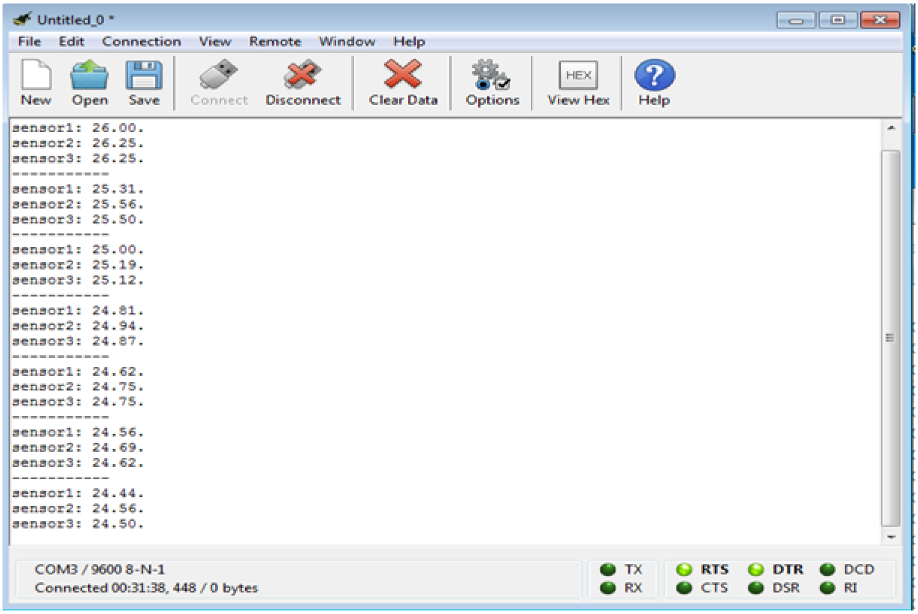


Fig. 1. Collected data via Coolterm tool in the case of three sensors

The measurement and data collection process involved two distinct scenarios, each with specific methodologies, as illustrated in Fig. 2:

a. **Simple Measurements Scenario**:

Individual readings were obtained sequentially using the DS18B20 sensor followed by the LM35 sensor.

b. **Multiple Measurements Scenario**:

Two groups of sensors were employed, each comprising three sensors: group A consisted of DS18B20 sensors, while group B featured LM35 sensors.

Throughout the experimentation period spanning 2 months, the following structured approach was adopted:

1. **Timeframe**: The experiment was divided into intervals of 5 days, providing a comprehensive assessment over time.

2. **Data collection phases**:

Within each day, three one-hour phases were designated for temperature readings.

Phase 1: Ambient conditions (first hour) Temperature readings were taken under normal, ambient conditions.

Phase 2: Introduction of external heat source (second hour)

An external heat source was introduced to observe its impact on the recorded temperatures.

Phase 3: Monitoring in cold environment (third hour)

Temperatures were monitored in a controlled cold environment to assess sensor responsiveness.

- Measurement frequency: During each one-hour phase, readings were taken at a rate of 20 measurements per hour. - Interval between measurements: A 3-minute interval was maintained between successive measurements to capture nuanced temperature variations.

This protocol facilitated a comprehensive evaluation of sensor performance across various conditions and scenarios, allowing for detailed analysis and comparison of temperature data collected by different sensors.

3 Analyse and Results Discussion

The data obtained from the measurement experiments is essential for assessing how well the sensors perform, especially in terms of accuracy and completeness. This section focuses on analyzing the results obtained from each test, paying particular attention to the fluctuations in temperature observed during different parts of the day. Statistical calculations, such as calculating the mean, are crucial in summarizing data, identifying trends, facilitating comparisons, and evaluating the precision of the measurements.

3.1 Simple Measurement

The figures depict the computed means, variances, and standard deviations, as illustrated in Fig. 3, Fig. 4, and Fig. 5, respectively. These graphical representations reveal consistent patterns with noticeable discrepancies between different phases. Initially, the mean values showed close proximity over the initial three days, followed by a deviation of 2 to 3°C. This observed discrepancy likely stems from environmental influences affecting the measurement conditions. Other examination of the recorded measurements allows for a comparative analysis between the two sensors. Further review of the recorded measurements enables

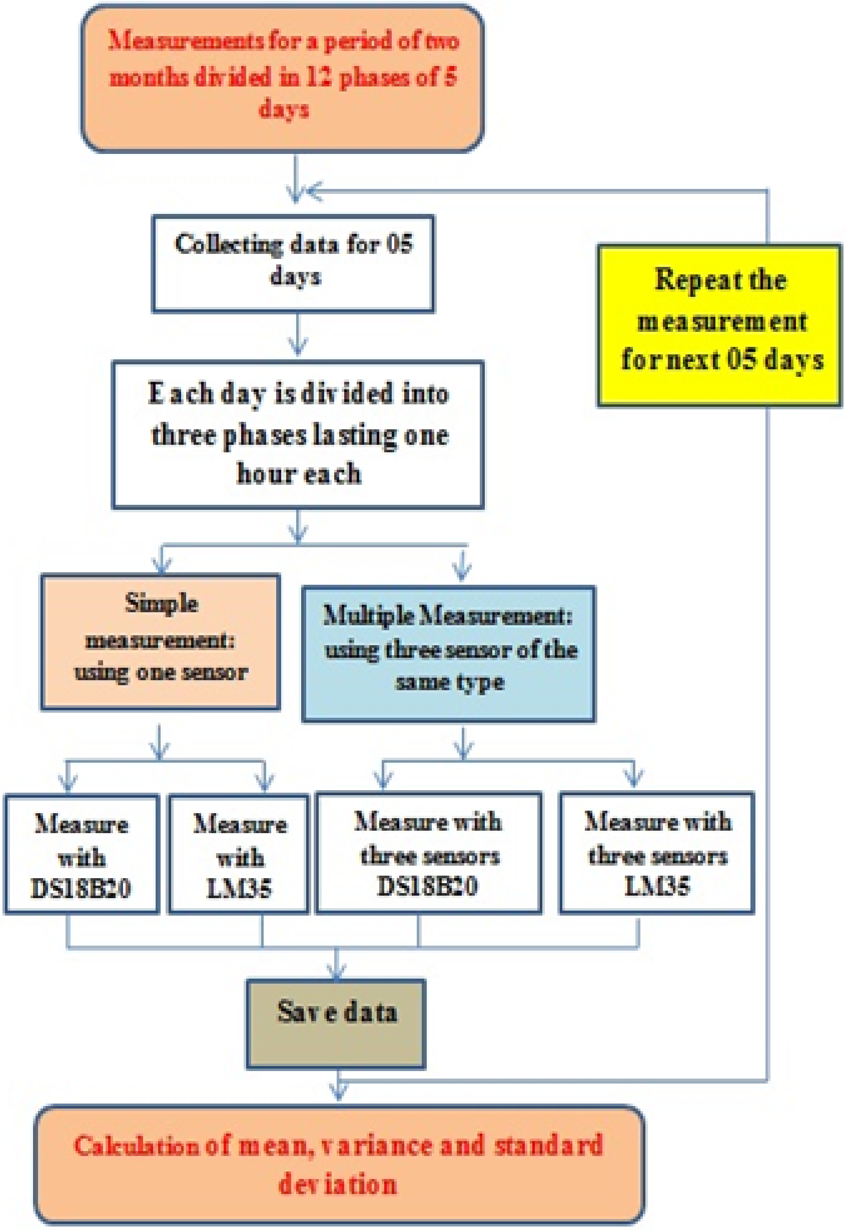


Fig. 2. Presentation of the scenario of measure

a comparative analysis between the two sensors. Upon examination, it became apparent that the LM35 sensor registered a considerable number of zero values throughout the 5-day period, notably during the initial hour.

When encountering instances of zero values or non-readings, it's crucial to take into account various specific factors pertaining to the operation of analog sensors:

- Faulty or insecure connections between the sensor and its associated circuitry during measurement can result in inaccurate readings or complete absence of readings.

- Ensuring a stable voltage supply within a defined range is imperative for the accurate operation of analog sensors.

- Analog sensors often generate signals that require scaling and conversion into temperature values, necessitating attention to signal range and appropriate scaling methods.

Moreover, factors such as sensor degradation or aging over time, reference voltage stability, sensor calibration, sensitivity and the presence of noise and interference should also be considered as potential limitations.

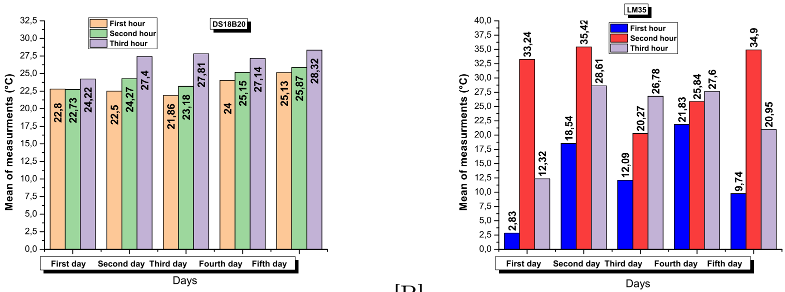
Consequently, the recorded count fell below the anticipated 20 readings, leading to the conclusion that the first sensor is more suitable for the task. In terms of the calculated means, it is evident that the Dallas DS18B20 sensor is significantly affected by environmental conditions, especially during the initial hour. The calculations provide valuable insights into data central tendency, dispersion, and variability. Regarding the DS18B20 digital sensor, the outcomes can be affected by factors such as the communication protocol, signal integrity, firmware, code, and others. The variance and standard deviation observed in the LM35 sensor dataset indicate a level of variability in temperature readings around the average value. Similarly, the variance and standard deviation in the DS18B20 sensor dataset show variability in temperature measurements, although these values may differ. A higher variance and standard deviation imply a greater spread of data points from the mean, potentially indicating more pronounced temperature fluctuations over time.

3.2 Multiple Measurements

As previously highlighted, the measurements were also conducted across multiple test scenarios involved the utilization of three sensors (identified as a, b, and c) referred to DS18B20 sensor. Subsequently, measurements were obtained using an additional trio of sensors (labeled d, e, and f) belonging to the second sensor type: LM35. The aim of this measurement is to assess data collected by each sensor, both in comparison to sensors of the same technology and sensors of different technologies to evaluate and ensure that the reading are correcte. The graphical representation of mean values for ambient conditions, heat source and cold conditions can be observed in the accompanying figures. Notably, a discernible distinction emerges upon contrasting the measurements derived from the two distinct sensor types, a distinction that becomes particularly pronounced between those for ambient conditions and for collected for cold condition. However, during the third trial conducted under the influence of heat source, these measurements exhibit similar magnitudes. As a result, it becomes evident that the DS18B20 sensor exhibits a superior level of accuracy for our specific measurement context. Furthermore, upon meticulous analysis of the recorded outcomes, an average occurrence of 70% null values was observed for the LM35 sensor during both the first and second trials. The results demonstrated that the DS18B20 sensor consistently outperformed the LM35 sensor in terms of accuracy and reliability, as indicated by lower variance and standard deviation. In contrast, the LM35 sensor showed higher variability and occasional zero values, particularly under ambient and cold conditions, likely due to its susceptibility to noise, voltage fluctuations, and environmental interference. Furthermore, the observation that three DS18B20 sensors showed weak variation when used concurrently highlights their robustness and reliability, with closely clustered mean temperature readings indicating high precision and consistency. This suggests that DS18B20 sensors perform reliably under identical conditions, reinforcing their suitability for accurate temperature monitoring. Conversely, the LM35 sensors exhibited greater variability and higher instances of zero values, indicating inconsistencies and potential reliability issues attributed to their analog nature and sensitivity to environmental factors and noise.

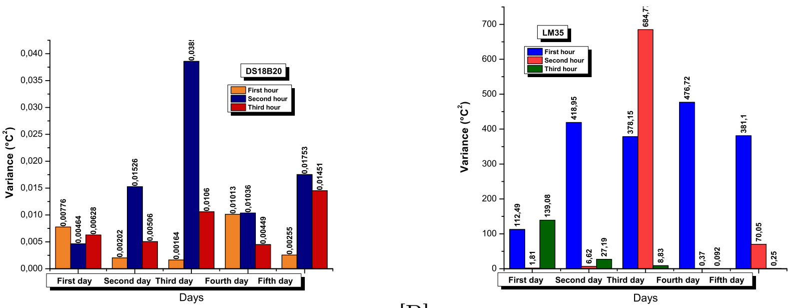
4 Conclusion

In conclusion, this study offers valuable insights into the performance of digital (DS18B20) and analog (LM35) temperature sensors within an IoT-based monitoring framework. It underscores the effectiveness of descriptive statistical methods in evaluating sensor data quality. While the DS18B20 sensor exhibits superior accuracy and reliability, limitations such as a small sample size, controlled conditions, technical constraints, and reliance on descriptive statistics need consideration. Future research should focus on extending the duration of data collection, increasing sample sizes, experimenting in diverse environments, exploring various hardware platforms, and employing advanced data analysis techniques to enhance the robustness and applicability of the findings. These efforts are crucial for developing more dependable IoT-based environmental monitoring systems. Moving forward, we aim to explore advanced methodologies such as regression analysis, integrate additional protocols like MQTT, test a broader range of sensors, and establish a user-friendly platform for assessing data quality, applicable to both individual and industrial settings.



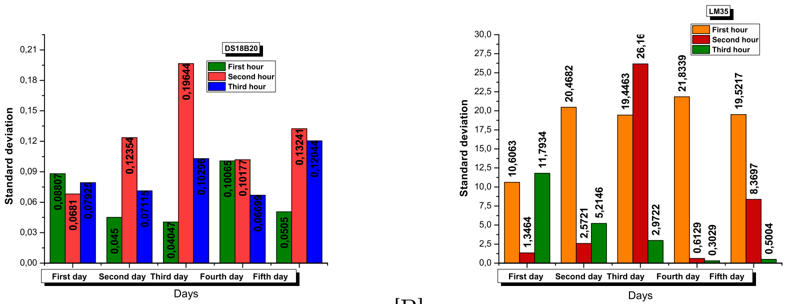
[A] [B]

Fig. 3. Illustration of the calculated Mean during the sampling phases for both DS18B20 (A) and LM35 (B) sensors.



[A] [B]

Fig. 4. Illustration of the calculated variance during the sampling phases for both DS18B20 (A) and LM35 (B) sensors.



[A] [B]

Fig. 5. Illustration of the calculated standard deviation during the sampling phases for both DS18B20 (A) and LM35 (B) sensors.

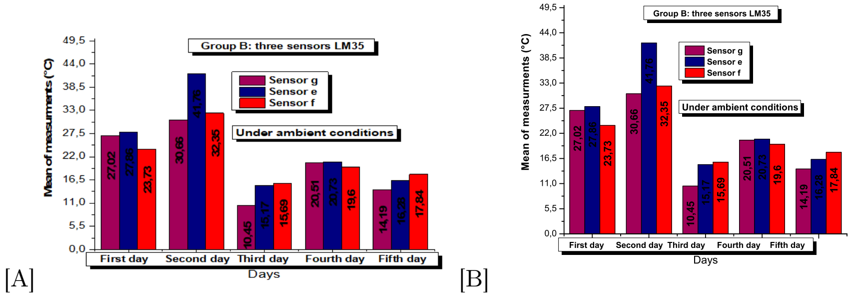


Fig. 6. Mean temperature changes for multiple tests under ambient conditions: A:DS18B20 and B:LM35.

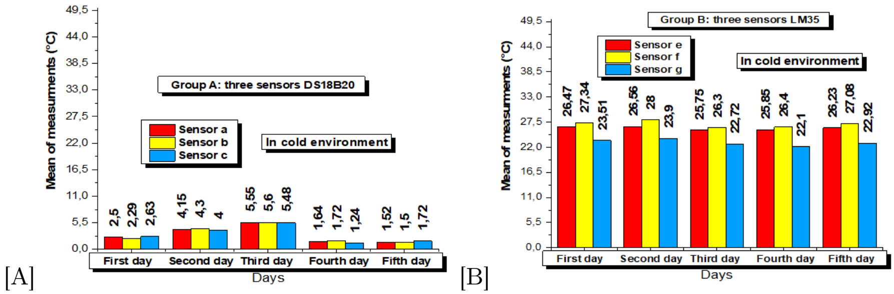


Fig. 7. Mean temperature changes for multiple tests in cold environment: A:DS18B20 and B:LM35.

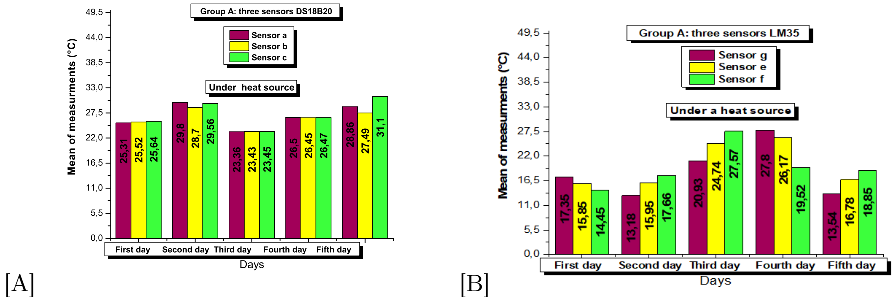


Fig. 8. Mean temperature changes for multiple tests under heat source: A:DS18B20 and B:LM35.

References

1. C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, 'Sensing as a service model for smart cities supported by Internet of Things', Trans. Emerging Tel. Tech., vol. 25, no. 1, pp. 81–93, (2014).
2. D. Bandyopadhyay and J. Sen, 'Internet of Things: Applications and Challenges in Technology and Standardization', Wireless Pers Commun, vol. 58, no. 1, pp. 49–69, (2011).
3. T. J. Saleem and M. A. Chishti, Eds., Big data analytics for Internet of things, First edition. Hoboken, NJ: Wiley, (2021).
4. J. Byabazaire, G. O'Hare, and D. Delaney, 'Data Quality and Trust: Review of Challenges and Opportunities for Data Sharing in IoT', Electronics, vol. 9, no. 12, pp. 2083, (2020).
5. P. Zikopoulos, Understanding big data: analytics for enterprise class Hadoop and streaming data. New York: McGraw-Hill, (2012).
6. R. Matheus, M. Janssen, and D. Maheshwari, 'Data science empowering the public: Data-driven dashboards for transparent and accountable decision-making in smart cities', Government Information Quarterly, vol. 37, no. 3, p. 101284,(2020).
7. R.Y.Wang, and D. M.Strong, Beyond Accuracy: What Data Quality Means to Data Consumers, Journal of Management Information System, 12 (4), pp.5-28, (1996).
8. F. A. Alaba, M. Othman, I. A. T. Hashem, and F. Alotaibi, 'Internet of Things security: A survey', Journal of Network and Computer Applications, vol. 88, pp. 10–28, (2017).
9. Gubbi, J., et al. Internet of Things (IoT): A vision, architectural elements, and future directions. Future Generation Computer Systems, 29(7), pp.1645-1660, (2013).
10. S. M. Riazul Islam, Daehan Kwak, M. Humaun Kabir, M. Hossain, and Kyung-Sup Kwak, 'The Internet of Things for Health Care: A Comprehensive Survey', IEEE Access, vol. 3, pp. 678–708, (2015).
11. D. Corrales, A. Ledezma, and J. Corrales, 'From Theory to Practice: A Data Quality Framework for Classification Tasks', Symmetry, vol. 10, no. 7, p. 248, (2018).
12. G. Rathee, R. Sandhu, H. Saini, M. Sivaram, and V. Dhasarathan, 'A trust computed framework for IoT devices and fog computing environment', Wireless Netw, vol. 26, no. 4, pp. 2339–2351, (2020).
13. L. K. Ramasamy and S. Kadry, Blockchain in the Industrial Internet of Things. IOP Publishing, (2021).
14. N. Khan, K. S. Khattak, S. Ullah, and Z. Khan, 'A Low-Cost IoT Based System for Environmental Monitoring', in 2019 International Conference on Frontiers of Information Technology (FIT), Islamabad, Pakistan: IEEE, pp. 173–1735, (2019).
15. S. L. Ullo and G. R. Sinha, 'Advances in Smart Environment Monitoring Systems Using IoT and Sensors', Sensors, vol. 20, no. 11, p. 3113, (2020).
16. S. Jain and K. Chandrasekaran, 'Industrial Automation Using Internet of Things':, in Research Anthology on Cross-Disciplinary Designs and Applications of Automation, I. R. Management Association, Ed., IGI Global, pp. 355–383, (2022).
17. T. Popović, N. Latinović, A. Pešić, Ž. Zečević, B. Krstajić, and S. Djukanović, 'Architecting an IoT-enabled platform for precision agriculture and ecological monitoring: A case study', Computers and Electronics in Agriculture, vol. 140, pp. 255–265, (2017).
18. J. Cabra, D. Castro, J. Colorado, D. Mendez, and L. Trujillo, 'An IoT Approach for Wireless Sensor Networks Applied to e-Health Environmental Monitoring', in 2017 IEEE International Conference on Internet of Things (iThings) and IEEE

Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Exeter: IEEE, Jun. 2017, pp. 578–583.
19. M. Chen, S. Mao, and Y. Liu, ‘Big Data: A Survey’, Mobile Netw Appl, vol. 19, no. 2, pp. 171–209, (2014).
20. J. Buelvas, D. Múnera, D. P. Tobón V., J. Aguirre, and N. Gaviria, ‘Data Quality in IoT-Based Air Quality Monitoring Systems: a Systematic Mapping Study’, Water Air Soil Pollut, vol. 234, no. 4, p. 248, Apr. (2023).
21. I. Taleb, M. A. Serhani, C. Bouhaddioui, and R. Dssouli, ‘Big data quality framework: a holistic approach to continuous quality management’, J Big Data, vol. 8, no. 1, p. 76, Dec. (2021).
22. I. E. Alaoui and Y. Gahi, ‘The Impact of Big Data Quality on Sentiment Analysis Approaches’, Procedia Computer Science, vol. 160, pp. 803–810, (2019).
23. K. Fizza, P. P. Jayaraman, A. Banerjee, D. Georgakopoulos, and R. Ranjan, ‘Evaluating Sensor Data Quality in Internet of Things Smart Agriculture Applications’, IEEE Micro, vol. 42, no. 1, pp. 51–60, Jan. (2022).
24. N. U. Okafor, Y. Alghorani, and D. T. Delaney, ‘Improving Data Quality of Low-cost IoT Sensors in Environmental Monitoring Networks Using Data Fusion and Machine Learning Approach’, ICT Express, vol. 6, no. 3, pp. 220–228, (2020).
25. J. Qi, P. Yang, L. Newcombe, X. Peng, Y. Yang, and Z. Zhao, ‘An overview of data fusion techniques for Internet of Things enabled physical activity recognition and measure’, Information Fusion, vol. 55, pp. 269–280,(2020).
26. C. Cichy, and R. Stefan. Fuzzy Expert Systems for Automated Data Quality Assessment and Improvement Processes. EKAW (Posters and Demos), pp. 7-21, (2020).
27. L. Ott and M. Longnecker, An introduction to statistical methods and data analysis, Seventh edition. Australia: Cengage Learning, (2016).
28. A. Field, Discovering statistics using IBM SPSS statistics, 5th edition. Thousand Oaks, CA: SAGE Publications, (2017).
29. W. Jia, R. M. Shukla, and S. Sengupta, ‘Anomaly Detection using Supervised Learning and Multiple Statistical Methods’, in 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), Boca Raton, FL, USA: IEEE, pp. 1291–1297,(2019).
30. L. Li, M. Rong, and G. Zhang, An Internet of Things QoE evaluation method based on multiple linear regression analysis,10th International Conference on Computer Science and Education (ICCSE), Cambridge, United Kingdom: IEEE, pp. 925–928, (2015).
31. A. John, B. Cardiff, and D. John, ‘A Generalized Signal Quality Estimation Method for IoT Sensors’, in 2020 IEEE International Symposium on Circuits and Systems (ISCAS), Seville, Spain: IEEE, Oct., pp. 1–5, (2020).
32. F. Emmert-Streib and M. Dehmer, ‘Understanding Statistical Hypothesis Testing: The Logic of Statistical Inference’, MAKE, vol. 1, no. 3, pp. 945–961,2019.
33. T. K. Kim, T test as a parametric statistic, Korean J Anesthesiol, vol. 68, no. 6, pp. 540, (2015).
34. M. Mayr and J. Himmelbauer, ‘Detecting Anomalies in Production Quality Data Using a Method Based on the Chi-Square Test Statistic’, in Big Data Analytics and Knowledge Discovery, M. Song, I.-Y. Song, G. Kotsis, A. M. Tjoa, and I. Khalil, Eds., in Lecture Notes in Computer Science, vol. 12393. Cham: Springer International Publishing, pp. 348–363, (2020).
35. M. Sarhan, S. Layeghy, and M. Portmann, ‘Feature Analysis for Machine Learning-based IoT Intrusion Detection’, 2021.

36. S. Rajeswari, K. Suthendran, Feature Selection Method based on Fisher's Exact Test for Agricultural Data, IJRTE, vol. 8, no. 4S2, pp. 558–564, (2019).
37. A. Botta, W. De Donato, V. Persico, and A. Pescapé, 'Integration of Cloud computing and Internet of Things: A survey', Future Generation Computer Systems, vol. 56, pp. 684–700, (2016).
38. A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, 'Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications', IEEE Commun. Surv. Tutorials, vol. 17, no. 4, pp. 2347–2376, (2015).
39. S. Li, L. D. Xu, and S. Zhao, 'The internet of things: a survey', Inf Syst Front, vol. 17, no. 2, pp. 243–259, (2015).
40. T. M. Booij, I. Chiscop, E. Meeuwissen, N. Moustafa, and F. T. H. D. Hartog, 'ToN_IoT: The Role of Heterogeneity and the Need for Standardization of Features and Attack Types in IoT Network Intrusion Data Sets', IEEE Internet Things J., vol. 9, no. 1, pp. 485–496, (2022).
41. L.George, and A. Harrison, Internet of Things Sensing Infrastructures and Data-Driven Planning Technologies in Smart Sustainable City Governance and Management. Geopolitics, History, and International Relations, vol. 13, no. 2, pp. 23–36, (2021).