Journal of Science, Management and Technology
ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

Advances and Challenges in IoT Sensors Data Handling and Processing in Environmental Monitoring Networks

Sruthi Sathyamoorthy[1], Ugochukwu Okwudili Matthew[2], Temitope Samson Adekunle[3], Nwamaka U. Okafor[4]

[1]School of Computing and Augmented Intelligence, Arizona State University, United States (ssathy10@asu.edu).

[2]Computer Science Department, Hussaini Adamu Federal Polytechnic, Nigeria (macdon4ru2003@gmail.com)

[3]Computer Science Department, Colorado State University, USA. (Temitope.adekunle@colostate.edu)

[4]School of Electrical and Electronic Engineering, University College Dublin, Ireland, (nwamaka.okafor@ucdconnect.ie)

Abstract

Advances in IoT technologies provide new epoch in ecological sensing, leading to the deployment of millions of sensor devices to sense and monitor the environment. IoT sensors have the capacity to provide high spatial and temporal resolution data to supplement traditional data-gathering methods, thereby filling the gaps that exist within current environmental data-gathering methods. Applications of IoT sensors in environmental monitoring are broad ranging from monitoring air quality, to monitoring biodiverse regions including forests and peatlands to protecting endangered species. The use of IoT sensor devices in environmental monitoring, however, has raised several questions, especially pertaining to the quality of sensor data, reliability, accuracy, and in-field performance. IoT sensors are prone to failures and errors especially when deployed for medium to longer-term. A common question within the IoT research domain is how to handle IoT sensor data, especially in terms of processing, fusion with other data sources and analysis to glean useful insights from the data in support of effective decision-making. Several authors have proposed different data handling methods for IoT sensor data and proposed techniques have led to improvement in overall data quality and field performance. Methods for addressing IoT sensor data analysis integration with emerging technologies, such as cloud computing, fog computing, and edge computing along with methods to make Data storage choices have also been proposed. This paper surveys the various methods for handling and processing IoT sensor data in environmental monitoring networks, the prospects, challenges, and limitations of these methods are examined.

Index Terms: IoT Sensors, Environmental Monitoring, Data Analysis, Machine Learning, Cloud Computing.

ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

I. Introduction

The proliferation of the Internet of Things (IoT) devices has brought about transformative advancements in environmental monitoring, revolutionizing our ability to monitor and understand the complexities of the natural ecosystem(Besson et al., 2022). IoT-based sensor networks have emerged as a powerful tool, enabling the collection of robust datasets from diverse environmental settings, ranging from remote forests and aquatic ecosystems to bustling urban areas(Marcu et al., 2019). This unprecedented capability has provided researchers, policymakers, and stakeholders with valuable insights into environmental dynamics, enabling evidence-based decision making to address pressing environmental challenges. The ability to monitor environmental phenomena with high spatial and temporal resolution has opened up new avenues for studying climate change, natural disasters, pollution, and the impacts of human activities on the planet(Saeed, Al-Naffouri, & Alouini, 2020).

By deploying a multitude of sensors, each tailored to measure specific environmental attributes, IoT technology has enabled the creation of comprehensive monitoring systems, capable of capturing intricate environmental interactions and feedback loops. This level of data granularity and coverage has not only enhanced our understanding of environmental processes but has also catalyzed the development of innovative strategies for conservation, sustainable resource management, and disaster preparedness. This review paper aims to provide a comprehensive examination of the methodologies, issues, and emerging trends in IoT sensor data handling and processing within the context of environmental monitoring. By synthesizing the current state of the technologies and processes in the field, we aim to shed light on the multifaceted issues that underpin the reliability and effectiveness of IoT-based environmental monitoring systems. The paper explores the diverse array of IoT sensor deployment across terrestrial, aquatic, and atmospheric environments considering various issues faced by IoT sensors, especially with sensor selection, deployment, and integration of data from multiple sensors as well as integration of sensors with other technologies.

An essential aspect of IoT-based environmental monitoring is in handling the vast amount of data generated from IoT sensor devices distributed across vast geographic areas. We evaluated the issues encountered with IoT data processing and analysis which play a significant role in converting raw sensor data into useful and actionable insights, more specifically, we reviewed

ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

research and trends in machine learning applications for data processing, statistical modeling, and data fusion approaches for extracting patterns, anomalies, and trends from IoT data streams. We also examined the issues surrounding data quality(Okafor, Alghorani, & Delaney, 2020), data sparsity, sensor calibration, energy efficiency, and sensor lifespan,all of which impact the reliability and sustainability of IoT-based monitoring systems(Suryavansh, Benna, Guest, & Chaterji, 2021).

In overall, the contributions of this paper include:

1) Presentation of a comprehensive and in-depth review of the various methods used for data handling and processing in IoT-based environmental monitoring.

2) Identification of key challenges and issues faced by IoT sensors in diverse areas of environmental monitoring, directing attention to critical challenges that need to be addressed to ensure the effectiveness and reliability of IoT-based systems.

3) By exploring the various applications of IoT-based environmental monitoring, the paper offers valuable insights for decision-making processes, such as climate change mitigation, disaster response, pollution control, and sustainable resource management. These insights can inform policymakers and stakeholders in making evidence-based decisions to address environmental challenges effectively.

This work not only inspires innovative research in the environmental monitoring domain to drive the development of effective and scalable environmental monitoring strategies that safeguard our planet's natural resources but it also enhances our understanding of the current IoT monitoring landscape. The rest of the paper is organized as follows; section II presents the state of the art, detailing advances and diverse issues faced by IoT sensors in environmental monitoring systems, particularly in terms of sensor deployment, data collection and analysis, outlining issues pertaining to data quality, data sparsity, sensor calibration, and energy efficiency and the various methods proposed to tackle these challenges. Section III presents the challenges in IoT sensor deployment and data collection. Section IV presents data processing and analysis, in section V, we discussed the integration of IoT sensors with cloud, fog, and edge computing. This section also deals with data storage choices in environmental monitoring networks detailing the issues of data retention, compression, and accessibility. Finally, the other subsections presented the conclusion and recommendations for future work.

ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

II. State of the Art

Environmental monitoring processes have undergone massive transformation through the integration of IoT technology, enabling a new era of insights and solutions(Dhingra, Madda, Gandomi, Patan, & Daneshmand, 2019). IoT technologies provide real-time data streams that unravel intricate connections in forest monitoring, track vital parameters in aquatic life monitoring, and uncover complex relationships in ecological sensing(Okafor & Delaney, 2019). The impact extends to diverse ecosystems, from peatlands to coral reefs, where IoT sensors offer insights into carbon dynamics and coral health. Moreover, these systems revolutionize plant health monitoring for optimized agriculture and elevate animal health monitoring to ensure thriving ecosystems. Beyond individual ecosystems, IoT technology plays a pivotal role in monitoring climate change indicators, safeguarding water quality, managing wildfires, and advancing the vision of smart cities(Gaur, Scotney, Parr, & McClean, 2015). The integration of IoT into environmental monitoring presents a transformative paradigm, offering the granularity and scope needed to comprehend the complexities of our planet's delicate balance.

One of the fundamental aspects of IoT-based environmental monitoring is data collection through a network of sensors. The developments in sensor miniaturization, energy efficiency, and wireless communication have expanded the scope and scalability of monitoring systems, allowing for more extensive coverage and higher granularity of data. The integration of edge and cloud computing has become a crucial strategy for managing the immense data generated by IoT-based environmental monitoring systems, enabling a balance between real-time responsiveness and efficient data management(Roostaei, 2018). While IoT-based environmental monitoring systems hold immense potential, they also face significant challenges. Data quality assurance, sensor calibration, energy efficiency, data security, and privacy preservation are critical concerns that require ongoing research and innovation. Moreover, the integration of AI techniques, such as explainable AI and generative models, holds promise for enhancing the capabilities of monitoring systems.

III. IoT Sensor Deployment and Data Collection: Challenges and Prospects

IoT-based deployments have revolutionized data collection, enabling insights across diverse domains. However, realizing the full potential of these deployments involves addressing multifaceted challenges while embracing innovative advancements.

ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

Location and Connectivity Challenges: location and connectivity challenges constitute a pivotal consideration, requiring meticulous selection of sensor placement to ensure the capturing of representative and accurate data(Azhaguramyaa & Srinivasan, 2021). In fact, connectivity issues are an essential aspect of sensor deployment, demanding solutions that overcome network coverage gaps to ensure reliable data transmission(Gluhak et al., 2011).

Cost Considerations: Achieving equilibrium between cost considerations is crucial, as the expense of sensor deployment, maintenance, and replacement must align with budget constraints(Metallidou, Psannis, & Egyptiadou, 2020). The advancements in remote calibration and predictive maintenance techniques, coupled with the integration of machine learning for efficient sensor deployment planning, contribute to overcoming these challenges(Catarinucci et al., 2015). While challenges remain, the strides made in technology and methodologies offer promising solutions to enhance the reliability, efficiency, and sustainability of IoT-based deployments. Also, cost effective maintenance and upkeep strategies are imperative to sustain deployment accuracy and data integrity, necessitating timely calibration and maintenance of sensors. However, cost implications may be more pronounced in specific domain where requirements for sensor accuracy is high i.e in air quality monitoring.

The deployment of CO_2 sensors in IoT-based environmental monitoring is instrumental in tracking air quality and carbon emissions. However, the expensive nature of deploying reliable CO_2 sensors presents a critical challenge that requires careful consideration. Understanding the key cost factors associated with sensor deployment is essential in managing budget constraints and optimizing resources(Y. Li et al., 2022). To alleviate the financial burden, exploring alternative sensor options is crucial. Research into more cost-effective sensor technologies that offer accurate measurements while reducing cost can be a potential solution(Cunin, McGrath, & MacNamee, 2018). Also, considering the investment required, the longevity and maintenance of CO_2 sensors play a pivotal role. A comprehensive assessment of the return on investment involves evaluating the sensor lifespan and associated maintenance costs against the benefits derived from accurate and continuous data collection(Chandramouli et al., 2021). Furthermore, achieving scalability in large-scale sensor deployments is another facet that necessitates attention. The challenge lies in ensuring accurate data collection while balancing the costs of acquiring, installing, and maintaining a multitude of sensors(Marques & Pitarma, 2019).

Journal of Science, Management and Technology
ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

Sensor calibration: The accuracy of lower-cost sensors warrants attention. Ensuring that these sensors yield reliable measurements entails a robust process of calibration and validation to uphold data quality(Venkatraman Jagatha et al., 2021). Data quality plays a significant role in the adoption of IoT sensors for various monitoring and sensing purposes. Data quality refers to the accuracy, completeness, consistency, and reliability of the data collected from various sensors and devices within the IoT ecosystem(Palaniswami, Rao, & Bainbridge, 2017).

As IoT systems generate massive amounts of data from diverse sources, ensuring high data quality is essential to extract meaningful insights, make informed decisions, and enable accurate predictions(Sicari, Rizzardi, Miorandi, Cappiello, & Coen-Porisini, 2016). Addressing data quality issues requires advanced data preprocessing techniques that filter out erroneous readings and interpolate missing data points. Effective and accurate sensor calibration has been identified as an essential tool for reliable data collection. Sensor drift, sensitivity variations, and environmental changes can lead to inaccuracies in measurements. Researchers are investigating calibration techniques that involve continuous monitoring and adjustment to maintain data accuracy over time(Martins, Fonseca, Farinha, Reis, & Cardoso, 2023). These techniques often leverage machine learning algorithms to adapt to changing sensor behavior.

TABLE I: Review of Existing IoT-Based Environmental Monitoring Systems

IoT Monitoring System	Application and Challenges	Method and Device Used
Forest Monitoring(Marcu et al., 2019).	Smart system for forest monitoring energy efficient, weather-proofed, capable of withstanding harsh environmental condition and able to communicate efficiently when placed in remote areas.	Raspberry Pi 3, IoT, M2M orientated communication protocol (MQTT), Python programs, Android
Aquatic and Aquaculture(Raju & Varma, 2017).	Deployed in creeks, lakes, rivers, and dam catchments; Features: simplicity of design, low power consumption, renewable power supply, plug and play, near real-time delivery of reliable and accurate sensor data, system stability/reliability over time, affordability, flexibility.	GPSB, Arduino Mega 2560, flash memory storage, Wireless communication.
Ecological sensing(Okafor & Delaney, 2019),(Matthew et al., 2021).	A framework for autonomous ecological sensing system through Peatland Monitoring using low cost sensors and wireless systems.	IoT, LCS, NB-IoT, LoRaWAN, 4G, ZigBee , Databases, APIs

ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

Smart city(Gaur et al., 2015).	Multi-Level Smart City architecture is proposed based on semantic web technologies and Dempster-Shafer uncertainty theory.	Wireless sensor network, Dempster Shafer theory, context-aware reasoning.
Climate monitoring(Saeed et al., 2020).	X-IoT Framework which consist of the Internet of underwater things for smart oceans, Internet of underground things for smart agriculture, seismic monitoring, and Oil/Gas fields, Internet of space things for outer space exploration, to provide global coverage, and to enable inter-satellite communications.	IoT
IoT system for peatland management(Liew et al., 2021).	IoT system with cyber physical system to reduce labor cost, enhance reliability, wireless communication, cloud server for data analytics, prediction of groundwater level of the peatland.	ML, IOT
Wildfire detection(H.-H. Liu, Chang, Chen, & Fu, 2021).	Wildfire model with multiple ignition points, carbon emission model based on real biomass information, satellite link budget analysis. Simulation results based on real environmental data of California in 2020 demonstrated that deploying as few as one sensor per km2 could reduce the annual carbon emission by more than ten times.	IoT, wildfire model, carbon emission model, satellite link budget analysis.
Great Barrier Coral Reef(Palaniswami et al., 2017).	WSN/IoT to monitor the complex marine environments, including the GBR. AI algorithms, we were able to detect Cyclone.	WSN/IoT based Monitoring, AI algorithms.
Plant Health(Pavel, Kamruzzaman, Hasan, & Sabuj, 2019).	Classify plant diseases based on images from raspberry pi, temperature, pH, humidity.	Raspberry Pi based IoT device, SVM, K Means.
Dairy Cow Health(Unold et al., 2020).	Aggregate behavior indicator technique for data measurement to precisely discriminate cow activities. Challenge: use of a magnetometer for low-energy estimation of cow mobility, and their correlation with the diagnosis of diseases manifested by walking changes; extend system to other diseases.	IoT, Cloud, User Application
Air Pollution(Dhingra et al., 2019).	Computational Complexity issues while dealing with large data.	IoT, Mobile application, Gas sensor.
Water quality	Dolphin Swarm Algorithm (DSA-ELM) was used as the evaluation model of the water quality.	Dolphin Swarm Algorithm.

ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

Carbon Dioxide Monitoring(Ming, Habeeb, Md Nasaruddin, & Gani, 2019).	Fully functional Carbon Dioxide Monitoring System integrating MQ135 carbon dioxide sensor with Node MCU ESP8266; sensor nodes implementation, use of ARIMA for forecast improvement.	MQ135 carbon dioxide sensor, ESP8266 Wi-Fi module, Firebase Cloud Storage Service and Android mobile application
Smart Cities for Healthcare(Alromaihi, Elmedany, & Balakrishna, 2018).	Computational attacks, Listening attacks, Broadcasting attacks, Integrity, Authentication, Secure localization, Data freshness, Privacy.	IoT
Environmental Sustainability and Climate Change(Fraser & Fraser, 2019).	A framework for interpretation and use of climate change projections and for innovations in climate and environmental science driven by key societal and economic stakeholders; IoT cyberinfrastructure.	IoT, cyberinfrastructure
Humidity and Dust monitoring(Beebi, 2018).	Low cost, small size, wide coverage and efficiency.	IoT
Radiation monitoring (X. Li et al., 2016).	Higher cost, low stability temperature variation.	HPXe chamber
Crop quality(Pathak et al., 2019).	SVM based remotely sensed synthetic aperture radar for paddy rice monitoring.	SVM, Back scattering features, regression tree, IoT
Pesticides, Fertilizer, Irrigation monitoring(Dimitriadis & Goumopoulos, 2008).	IoT-based system that uses ML-based prediction on sensor data.	Naıve Bayes, IoT
Pest Control(L. Liu et al., 2019).	IoT and deep learning using global, local features for pest monitoring.	CNN, IoT
peatland monitoring(Suhaili, Yusop, & Wan, 2020).	System for water level detection on peatland using visual IoT.	visual Internet of things (IoT), Raspberry Pi, camera, Arduino
Waste Management(Aleyadeh & Taha, 2018).	IoT-based architecture that targets (i) monitoring the waste volume and content in a waste bin, as well as the bin's surroundings (ii) dynamic scheduling and routing of waste collection vehicles based on the relayed information from the bins. waste bin design detects any obstacles around the bin and monitors illegal dumping in the vicinity of the bin.	IoT, Routing protocols

Deforestation(Andreadis, Giambene, & Zambon, 2021).	Framework for automatically detecting illegal tree-cutting using ultralow-power tiny devices, embedding edge-computing microcontrollers and long-range wireless communication to cover vast areas in the forest; audio classification solution based on convolutional neural networks is proposed, designed specifically for resource-constrained wireless edge devices.	IoT, CNN, edge-computing, ultra-low power tiny device

Data Fusion: Data fusion has been identified as a vital component of IoT data collection process. Data fusion involves the amalgamation of data from various sensors and sources to provide a comprehensive understanding of the monitored environment. The heterogeneity of sensors, each with distinct measurement units and formats pose difficulties in standardization and synchronization(Hall & Llinas, 1997). Uncertainties and inaccuracies introduced by sensors necessitate the incorporation of probabilistic models for robust fusion(Abdulhafiz & Khamis, 2013). The authors detailed how inaccuracies in the data collected from various data fusion techniques can be solved by networked smart objects and machine learning models(Rizzardi, Miorandi, Sicari, Cappiello, & Coen-Porisini, 2016).

Data Sparsity: Data sparsity is another common issue faced by IoT sensors. It is a situation where sensors might not uniformly cover the entire monitoring area, leading to sparse data streams. Sparse data can hinder the performance of traditional statistical models and machine learning algorithms, making it difficult to extract meaningful patterns(Du et al., 2020). Methods such as data imputation, spatial interpolation, and multi-agent deep reinforcement learning-based method have been proposed for mitigating the effects of data sparsity.

Energy Efficiency: Many IoT devices operate on battery power and energy efficiency is a critical concern for prolonged system operation. Data transmission, processing, and analysis require energy, and optimizing these processes to reduce power consumption is paramount. Research in this area includes low-power communication protocols, duty-cycling strategies, and energy-efficient data processing algorithms(Soleymani et al., 2020). Energy efficiency is essential for prolonging the lifespan of IoT sensors. The lifespan of sensors impacts the sustainability of monitoring systems. Frequent sensor replacements can be costly and disruptive. Extending sensor lifespan involves managing power consumption, implementing predictive maintenance strategies, and using robust materials to withstand environmental conditions. Machine learning algorithms

Journal of Science, Management and Technology
ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

are being applied to predict sensor failures and optimize maintenance schedules (Guo, Qu, Huang, & Yao, 2011).

Furthermore, there are approaches that use algorithms like window-based time series forecasting mechanisms for data reduction protocol for reducing the amount of data transmissions from IoT sensors to the server resulting in an increase in the battery lifetime of the sensor nodes.

Privacy and Security: Privacy and Security has often generated a lot of questions within the IoT sensor research community. As IoT devices gather sensitive and personal information from users, organizations, and environments, the risk of unauthorized access, data breaches, and privacy infringements becomes pronounced. Unauthorized access to IoT data can lead to significant consequences, ranging from identity theft to unauthorized surveillance. Ensuring confidentiality and integrity throughout the data lifecycle is paramount to protect against breaches during collection, transmission, and storage. Cryptographic techniques, including encryption and digital signatures, are widely adopted to secure IoT data against eavesdropping and tampering. Access control mechanisms play a crucial role in ensuring that only authorized entities can access and manipulate data. Secure communication protocols, such as HTTPS and MQTT with TLS/SSL, help safeguard data during transit(Atzori, Iera, & Morabito, 2010). However, challenges arise in managing these measures across the heterogeneous landscape of IoT devices. Moreover, the emergence of edge, fog, and cloud computing paradigms adds layers of complexity to IoT security.

While edge computing minimizes data transmission, it demands robust mechanisms to secure data at the device level. Fog computing introduces intermediary processing points, requiring secure communication and data validation between sensors and fog nodes. Cloud computing necessitates encryption and access controls to protect data stored remotely. The urgency of addressing IoT privacy and security concerns has led to calls for industry standards and regulatory frameworks. Collaborative efforts are essential to developing guidelines that balance innovation with data protection. The General Data Protection Regulation (GDPR) in Europe, for instance, includes provisions that emphasize users' right to consent and control over their data. Such regulations provide a foundation for protecting individual privacy rights while promoting the benefits of IoT data collection.

ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

IV. IoT Sensor Data Analytics

Data Fusion: Analyzing IoT sensor data can be a crucial step to uncover any major insights and for prediction. Most often, the data that we get from one particular sensor source might not be sufficient or might be sporadic. To overcome this process we use a process called Data fusion(Ding, Jing, Yan, & Yang, 2019). In this process, we try to combine the data from different sources and run. Data fusion methods use probabilistic, statistical, knowledge-based, and inference and reasoning methods. The probabilistic methods include Bayesian networks, maxi-mum likelihood estimation methods, inference theory, Kalman filtering, etc. Statistical methods include covariance, cross-variance, and other statistical analyses(De Paola, Ferraro, Gaglio, Re, & Das, 2016). Knowledge-based methods include artificial neural networks, fuzzy logic, genetic algorithms, etc. Sensor fusion can be used for increases in the quality of data, increase reliability, measure unmeasured states, and increase the coverage area. Many satellites, aircraft, and underground or underwater machines use seismic, EM radiation, chemical or biological sensors to collect accurate information or identify natural phenomena in environment monitoring from very long distances; Military and defense services use this technique in ocean surveillance, air-to-air or ground-to-air defense, battlefield intelligence, data acquisition, warning, defense systems, etc., using EM radiation from large distances

Denoising: Noise in sensor data signals can result in erroneous analysis results. Data Denoising is a process of re-moving unwanted and unnecessary information from the data. Major techniques used are Wavelet transformation techniques to reduce the Signal Noise Ratio (SNR) and to achieve the desired Bit Error Rate by decreasing the wavelet transformation coefficient(Berkner & Wells, 1998). For the majority of Environmental monitoring systems, real-time denoising is required for real-time predictions and data analysis.

Missing Data Imputation: Missing data is a common issue faced in IoT-based environmental monitoring systems that use technologies like GPS tracking, smart transportation, smart cities, etc. With suitable data analysis techniques, the missing data can be imputed. Most often, a predictive analysis using Machine learning algorithms like Naive Bayes, SVM, Linear regression, and logistic regression is executed to predict the missing data and add it to the input data set. Algorithms like Temporal and Spatial Correlation Algorithms, the Gaussian Mixture model can also be used (Yan, Xiong, Hu, Wang, & Zhao, 2015).

ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

Outlier Detection: Outliers in data can result in erroneous prediction or data analysis. Methods like DBSCAN has been found effective for outlier detection, also, Principal Component Analysis can be performed to select the most appropriate features for prediction(Al-khatib, Mohammed, & Abdelmajid, 2020).

V. Fundamental Design of IoT Sensors Processing and Handling of Data in Environmental Monitoring

The essential design for the layers of IoT sensor data processing, fusion, and analysis is shown in **Fig1**. The main components of the IoT sensor data layer are several IoT sensors that are able to measure the physical surroundings and record changes in the environment in real time. Infrared (IR) and radiofrequency identification (RFID) sensors, accelerometers, gas, gyroscopes, motion sensors, optical sensors, and temperature, pressure, humidity, and level are among the frequently utilized IoT sensors. IoT sensors are primarily connected to wireless communication interfaces, microprocessor units, storage units, control units, and power systems. Size, processing power, memory, networking capabilities, and storage capacity are all limited in IoT sensor devices. For IoT sensor device connectivity, wireless communication protocols like Bluetooth, Wi-Fi, Zig Bee, Near Frequency connectivity (NFC), and LTE/4G mobile technologies are frequently utilized(Poongodi, Gopal, & Saini, 2021).

For use in industrial, scientific, and environmental applications, the majority of IoT sensor data are processed in real time. It is vital to process these sensed data to remove ambiguities before further evaluating them in order to increase knowledge and decision-making. Given this, the data processing layer concentrates on a number of activities, such as data denoising, data aggregation, data outlier identification, and data repair for missing data. Different sensor data challenges are brought about by multiple heterogeneous sensor devices, and the data fusion layer is responsible for handling them. Combining real sensor data from multiple IoT sensor devices is the aim of data fusion. The combined data from many sources is then sent to the data analysis layer for efficient knowledge development and decision-making(Olaniyi, Okunleye, & Olabanji, 2023).

Journal of Science, Management and Technology
ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

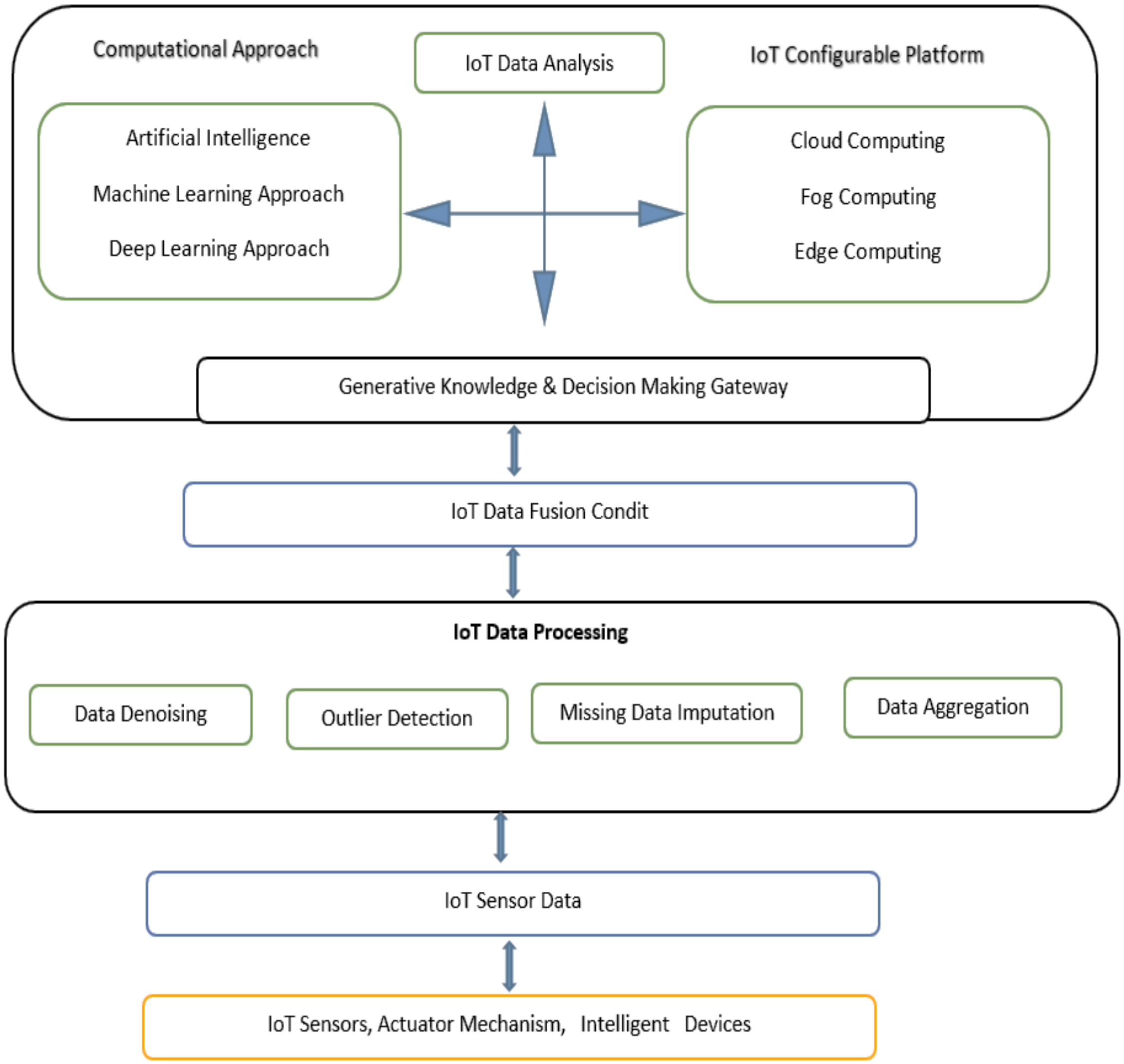

Fig 1: The framework for processing, fusing, and analyzing data from IoT sensors(Krishnamurthi, Kumar, Gopinathan, Nayyar, & Qureshi, 2020).

The core technique of data fusion is the direct fusing of sensor data from several sensor devices. In the improved method, data fusion comes after the initial feature extraction, and identity declaration comes after the feature extraction. Highly accurate information inference and decision-making are made possible by this method. IoT sensor data analysis is the result of the current transformation of edge, cloud, and fog computing through the use of cutting-edge technology(Chakraborty et al., 2023). These supporting technologies provide a widely available, robust, and intuitive foundation for handling the varied and dynamic nature of sensor data from

ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

IoT devices. Thus, generating intelligent capabilities that can manage a wide variety of Internet of Things applications is the aim of the data analysis layer. Reducing computing and storage costs, enhancing IoT network security and privacy, lowering network latency, guaranteeing scalability, and enabling failsafe and risk-free IoT solutions are the goals of these platforms.

TABLE II: DATA ANALYSIS IN IOT-BASED ENVIRONMENTAL MONITORING

Research	**Proposed Solution**
Data Preprocessing	Proposed Solution data quality, denoising, missing data imputation
Data Fusion	Triple Phases Resource Utilized Data Fusion (TPRUDF)
Selection of Algorithms	supervised and unsupervised ML
Real-time Processing	RTID for real-time IoT stream processing
Data Analytics Challenges	large volume of data, processing offline, real-time processing, parallel processing
Data Preprocessing	Data Denoising, Missing Data Imputation
Data Analysis	Data Outlier detection

VI. Integration of IoT Sensors with Cloud, Fog and Edge Computing and Data Storage Choices

The expansion of IoT-based deployments has resulted in a paradigm shift in computational models, characterized by a spectrum of processing strategies that cater to diverse requirements. Fog computing emerges as a promising approach, mitigating the challenges of distributing computation closer to the data source. The placement of resources at the network edge addresses latency and bandwidth concerns while enhancing responsiveness(Nandhakumar, Baranwal, Choudhary, Golec, & Gill, 2024). Cloud computing In contrast, cloud computing remains an essential component for managing vast amounts of data. This centralized model offers scalability and extensive processing capabilities but necessitates efficient data transmission, storage, and

ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

security mechanisms. Edge computing, a complementary paradigm, leverages de-vices near data sources to perform initial processing, reducing the need for data transmission to remote locations. Database selection stands as a crucial decision, with a comparison between SQL and NoSQL databases for IoT data storage being essential. While SQL databases offer structured storage and strong data consistency, NoSQL databases provide flexibility and scalability for handling diverse data types. Scalability and performance challenge emerges when dealing with large volumes of time-series data, demanding storage solutions capable of accommodating rapid growth and high-throughput data processing. Data retention policies establishing effective data retention policies is equally critical. Defining appropriate data retention periods and implementing archiving strategies are essential to managing the influx of historical data while maintaining storage efficiency. Data Compression techniques play a vital role in mitigating storage requirements. Implementing compression mechanisms can significantly reduce the storage footprint, making efficient use of available resources(Amir et al., 2023).

TABLE IV: Data Storage in IoT-Based Environmental Monitoring.

Research	Proposed Solution
Fog computing	Placement of resources at the network edge, addresses latency and bandwidth concerns while enhancing responsiveness.
Cloud computing	centralized model offers scalability and extensive processing capabilities but necessitates efficient data transmission, storage, and security mechanisms.
Edge computing	Devices near data sources to perform initial processing, reducing the need for data transmission.
Database Selection	Comparison between SQL and NoSQL databases for IoT data storage.
Scalability	Databases for IoT data storage case study of Noise pollution prediction in urban area with a large volume of data using long short-term memory. (LSTM).

ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

Data retention policy	Defining appropriate data retention periods and implementing archiving strategies.
Data Compression	Implementing compression mechanisms can significantly reduce the storage footprint.
Data Accessibility	Cooperative File System (CFS) with a distributed hash table (DHash) for block storage.
Database Management Difficulties	Size, Scale and Indexing, Query Languages, Process Modeling and Transactions, Heterogeneity and Integration, Time Series Aggregation , Archiving, Data Protection; SQL, XML, GUI ; Relaxation of ACID properties; modular software development; context-aware pervasive computing.
Blockchain-based database	Blockchain Bitcoin Backbone Protocol BBP) based Databases for IoT; limited computation power, use edge or fog proxy.

As the volume of data increases, ensuring data accessibility becomes paramount. A chosen storage solution should enable efficient querying, retrieval of relevant data to extract actionable insights in a timely manner.

VII. Conclusion

In conclusion, the rapid evolution of IoT technology has ushered in a new era of environmental monitoring, offering unprecedented capabilities to observe and analyze our planet's intricate ecosystems. This review paper has meticulously examined the methods, challenges, and prospects of IoT sensors data handling and processing in environmental monitoring, highlighting the multifaceted nature of this dynamic field. As evidenced by the diverse applications discussed, IoT sensor networks have become instrumental in addressing pressing environmental issues. The ability to collect high-resolution, real-time data from remote and challenging environments has empowered researchers and decision-makers with invaluable insights into climate dynamics, natural disasters, habitat preservation, and pollution management.

ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

The deployment of tailored sensors in terrestrial, aquatic, and atmospheric settings has facilitated comprehensive monitoring, enabling a holistic understanding of interconnected ecological processes. The paper's exploration of data handling techniques emphasizes the critical role of efficient data aggregation, transmission, and storage mechanisms. This is particularly significant in environments with limited connectivity and energy resources. Equally crucial is the data processing phase, where sophisticated algorithms and analytics are harnessed to extract actionable insights from the deluge of raw sensor data. The integration of machine learning, statistical modeling, and data fusion techniques has enabled the detection of trends, anomalies, and correlations that inform effective decision-making.

Nevertheless, this review has also illuminated the challenges that underscore the journey toward harnessing IoT technology for environmental monitoring. Issues like data accuracy, sensor calibration, energy efficiency, and security demand innovative solutions. Interoperability among disparate sensor systems and standardization of data formats remain crucial for collaborative efforts across different monitoring initiatives. In envisioning the road ahead, the paper serves as a guidepost for future research and innovation in this domain. As technology advances, opportunities abound to refine sensor networks, enhance data processing algorithms, and integrate emerging technologies like edge computing and 5G connectivity. The lessons learned from addressing challenges in energy efficiency, sustainability, and data privacy can be applied to foster the resilience and longevity of monitoring systems. By elucidating the intricacies of IoT sensors data handling and processing in environmental monitoring, this review paper equips researchers, practitioners, and policymakers with the knowledge needed to drive trans-formative change. As our world faces increasingly complex environmental challenges, the insights presented here pave the way for the development of innovative solutions that harness the power of IoT technology to create a more sustainable and harmonious future for all.

Conflict Of Interest

There is no conflict of interest regarding this research work

Funding :

There is no approved funding for this research work

Journal of Science, Management and Technology
ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

References

Abdulhafiz, W. A., & Khamis, A. (2013). Handling data uncertainty and inconsistency using multisensor data fusion. *Advances in Artificial Intelligence*, *2013*, 11-11.

Al-khatib, A. A., Mohammed, B., & Abdelmajid, K. (2020). A survey on outlier detection in Internet of Things big data. *Big Data-Enabled Internet of Things; IET: London, UK*, 265-272.

Aleyadeh, S., & Taha, A.-E. M. (2018). *An IoT-Based architecture for waste management.* Paper presented at the 2018 IEEE International Conference on Communications Workshops (ICC Workshops).

Alromaihi, S., Elmedany, W., & Balakrishna, C. (2018). *Cyber security challenges of deploying IoT in smart cities for healthcare applications.* Paper presented at the 2018 6th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW).

Amir, M., Deshmukh, R. G., Khalid, H. M., Said, Z., Raza, A., Muyeen, S., . . . Sopian, K. (2023). Energy storage technologies: An integrated survey of developments, global economical/environmental effects, optimal scheduling model, and sustainable adaption policies. *Journal of Energy Storage, 72*, 108694.

Andreadis, A., Giambene, G., & Zambon, R. (2021). Monitoring illegal tree cutting through ultra-low-power smart iot devices. *Sensors, 21*(22), 7593.

Atzori, L., Iera, A., & Morabito, G. (2010). The internet of things: A survey. *Computer Networks, 54*(15), 2787-2805.

Azhaguramyaa, V., & Srinivasan, K. (2021). *Location-aware security system for smart cities using iot.* Paper presented at the Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2020.

Beebi, F. (2018). Environmental monitoring system using IoT. *India Res. Tech. Organiz, 5*, 64-68.

Berkner, K., & Wells, R. (1998). *Wavelet transforms and denoising algorithms.* Paper presented at the Conference Record of Thirty-Second Asilomar Conference on Signals, Systems and Computers (Cat. No. 98CH36284).

Besson, M., Alison, J., Bjerge, K., Gorochowski, T. E., Høye, T. T., Jucker, T., . . . Clements, C. F. (2022). Towards the fully automated monitoring of ecological communities. *Ecology Letters, 25*(12), 2753-2775.

Catarinucci, L., De Donno, D., Mainetti, L., Palano, L., Patrono, L., Stefanizzi, M. L., & Tarricone, L. (2015). An IoT-aware architecture for smart healthcare systems. *IEEE Internet of Things Journal, 2*(6), 515-526.

Chakraborty, K., Sharma, M., Kommuri, K., Subrahmanyam, V., Patil, P., & Yadav, M. S. (2023). Resource Management Techniques for the Internet of Things, Edge, and Fog Computing Environments. *Engineering Proceedings, 59*(1), 12.

Chandramouli, M., Boraiah, V. K., Shivalingappa, R. P., Basavanna, V., Doddamani, S., & Ningaiah, S. (2021). Based Carbon Dioxide Sensors: Past, Present, and Future Perspectives. *Biointerface Res. Appl. Chem., 12*, 2353-2360.

Cunin, T., McGrath, S., & MacNamee, C. (2018). *Environmental monitoring based on Internet of Things Technology.* Paper presented at the 2018 12th International Conference on Sensing Technology (ICST).

ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

De Paola, A., Ferraro, P., Gaglio, S., Re, G. L., & Das, S. K. (2016). An adaptive bayesian system for context-aware data fusion in smart environments. *IEEE Transactions on Mobile Computing, 16*(6), 1502-1515.

Dhingra, S., Madda, R. B., Gandomi, A. H., Patan, R., & Daneshmand, M. (2019). Internet of Things mobile–air pollution monitoring system (IoT-Mobair). *IEEE Internet of Things Journal, 6*(3), 5577-5584.

Dimitriadis, S., & Goumopoulos, C. (2008). *Applying machine learning to extract new knowledge in precision agriculture applications.* Paper presented at the 2008 Panhellenic Conference on Informatics.

Ding, W., Jing, X., Yan, Z., & Yang, L. T. (2019). A survey on data fusion in internet of things: Towards secure and privacy-preserving fusion. *Information Fusion, 51*, 129-144.

Du, P., Bai, X., Tan, K., Xue, Z., Samat, A., Xia, J., . . . Liu, W. (2020). Advances of four machine learning methods for spatial data handling: A review. *Journal of Geovisualization and Spatial Analysis, 4*, 1-25.

Fraser, A., & Fraser, A. (2019). Environmental sustainability and climate change. *Achieving the Sustainable Management of Forests*, 75-92.

Gaur, A., Scotney, B., Parr, G., & McClean, S. (2015). Smart city architecture and its applications based on IoT. *Procedia Computer Science, 52*, 1089-1094.

Gluhak, A., Krco, S., Nati, M., Pfisterer, D., Mitton, N., & Razafindralambo, T. (2011). A survey on facilities for experimental internet of things research. *IEEE Communications Magazine, 49*(11), 58-67.

Guo, D., Qu, X., Huang, L., & Yao, Y. (2011). Sparsity-based spatial interpolation in wireless sensor networks. *Sensors, 11*(3), 2385-2407.

Hall, D. L., & Llinas, J. (1997). An introduction to multisensor data fusion. *Proceedings of the IEEE, 85*(1), 6-23.

Krishnamurthi, R., Kumar, A., Gopinathan, D., Nayyar, A., & Qureshi, B. (2020). An overview of IoT sensor data processing, fusion, and analysis techniques. *Sensors, 20*(21), 6076.

Li, X., Liu, Q., Yang, R., Wen, J., Zhang, J., Cai, E., & Zhang, H. (2016). The combination of ground-sensing network and satellite remote sensing in Huailai county. *IEEE Sensors Journal, 16*(10), 3819-3826.

Li, Y., Yu, L., Zheng, C., Ma, Z., Yang, S., Song, F., . . . Wang, Y. (2022). Development and field deployment of a mid-infrared CO and CO2 dual-gas sensor system for early fire detection and location. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 270*, 120834.

Liew, J. T., Sali, A., Noordin, N. K., Mohd. Ali, B., Hashim, F., Mashohor, S., . . . Sekhari, A. (2021). *Sustainable peatland management with IoT and data analytics.* Paper presented at the Smart and Sustainable Collaborative Networks 4.0: 22nd IFIP WG 5.5 Working Conference on Virtual Enterprises, PRO-VE 2021, Saint-Étienne, France, November 22–24, 2021, Proceedings 22.

Liu, H.-H., Chang, R. Y., Chen, Y.-Y., & Fu, I.-K. (2021). *Sensor-based satellite IoT for early wildfire detection.* Paper presented at the 2021 IEEE Globecom Workshops (GC Wkshps).

Liu, L., Wang, R., Xie, C., Yang, P., Sudirman, S., Wang, F., & Li, R. (2019). *Deep learning based automatic approach using hybrid global and local activated features towards large-scale multi-class pest monitoring.* Paper presented at the 2019 IEEE 17th International Conference on Industrial Informatics (INDIN).

ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

Marcu, A.-E., Suciu, G., Olteanu, E., Miu, D., Drosu, A., & Marcu, I. (2019). *IoT system for forest monitoring.* Paper presented at the 2019 42nd International Conference on Telecommunications and Signal Processing (TSP).

Marques, G., & Pitarma, R. (2019). A cost-effective air quality supervision solution for enhanced living environments through the internet of things. *Electronics, 8*(2), 170.

Martins, A., Fonseca, I., Farinha, J. T., Reis, J., & Cardoso, A. J. M. (2023). Online Monitoring of Sensor Calibration Status to Support Condition-Based Maintenance. *Sensors, 23*(5), 2402.

Matthew, U. O., Kazaure, J. S., Onyebuchi, A., Daniel, O. O., Muhammed, I. H., & Okafor, N. U. (2021). *Artificial intelligence autonomous unmanned aerial vehicle (UAV) system for remote sensing in security surveillance.* Paper presented at the 2020 IEEE 2nd International Conference on Cyberspac (CYBER NIGERIA).

Metallidou, C. K., Psannis, K. E., & Egyptiadou, E. A. (2020). Energy efficiency in smart buildings: IoT approaches. *IEEE Access, 8*, 63679-63699.

Ming, F. X., Habeeb, R. A. A., Md Nasaruddin, F. H. B., & Gani, A. B. (2019). *Real-time carbon dioxide monitoring based on iot & cloud technologies.* Paper presented at the Proceedings of the 2019 8th International Conference on Software and Computer Applications.

Nandhakumar, A. R., Baranwal, A., Choudhary, P., Golec, M., & Gill, S. S. (2024). Edgeaisim: A toolkit for simulation and modelling of ai models in edge computing environments. *Measurement: Sensors*, *31*, 100939.

Okafor, N. U., Alghorani, Y., & Delaney, D. T. (2020). Improving data quality of low-cost IoT sensors in environmental monitoring networks using data fusion and machine learning approach. *ICT Express*, *6*(3), 220-228.

Okafor, N. U., & Delaney, D. (2019). Considerations for system design in IoT-based autonomous ecological sensing. *Procedia Computer Science, 155*, 258-267.

Olaniyi, O., Okunleye, O. J., & Olabanji, S. O. (2023). Advancing data-driven decision-making in smart cities through big data analytics: A comprehensive review of existing literature. *Current Journal of Applied Science and Technology, 42*(25), 10-18.

Palaniswami, M., Rao, A. S., & Bainbridge, S. (2017). Real-time monitoring of the great barrier reef using internet of things with big data analytics. *ITU J. ICT Discov, 1*(1).

Pathak, A., AmazUddin, M., Abedin, M. J., Andersson, K., Mustafa, R., & Hossain, M. S. (2019). IoT based smart system to support agricultural parameters: A case study. *Procedia Computer Science, 155*, 648-653.

Pavel, M. I., Kamruzzaman, S. M., Hasan, S. S., & Sabuj, S. R. (2019). *An IoT based plant health monitoring system implementing image processing.* Paper presented at the 2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS).

Poongodi, T., Gopal, R., & Saini, A. (2021). IoT Architecture, Communication Technologies, and Its Applications. *Internet of Things, Artificial Intelligence and Blockchain Technology*, 1-24.

Raju, K. R. S. R., & Varma, G. H. K. (2017). *Knowledge based real time monitoring system for aquaculture using IoT.* Paper presented at the 2017 IEEE 7th international advance computing conference (IACC).

Rizzardi, A., Miorandi, D., Sicari, S., Cappiello, C., & Coen-Porisini, A. (2016). *Networked smart objects: moving data processing closer to the source.* Paper presented at the Internet of Things. IoT Infrastructures: Second International Summit, IoT 360° 2015, Rome, Italy, October 27-29, 2015, Revised Selected Papers, Part II.

ISSN (Print): 1597-7463; eISSN (Online): 2811-2598, Volume 5, Issue 2 https://dx.doi.org/10.4314/hpjsmt.v5i2.3

Roostaei, J. (2018). *IoT Based Edge and Cloud Computing for Smart Environmental Engineering Applications*: Wayne State University.

Saeed, N., Al-Naffouri, T. Y., & Alouini, M.-S. (2020). Around the world of IoT/climate monitoring using internet of X-things. *IEEE Internet of Things Magazine, 3*(2), 82-83.

Sicari, S., Rizzardi, A., Miorandi, D., Cappiello, C., & Coen-Porisini, A. (2016). A secure and quality-aware prototypical architecture for the Internet of Things. *Information Systems, 58*, 43-55.

Soleymani, S. A., Goudarzi, S., Kama, N., Adli Ismail, S., Ali, M., MD Zainal, Z., & Zareei, M. (2020). A hybrid prediction model for energy-efficient data collection in wireless sensor networks. *Symmetry, 12*(12), 2024.

Suhaili, W., Yusop, Y., & Wan, A. (2020). *Opportunities in using visual IOT in the mitigation for peatland area.* Paper presented at the IOP Conference Series: Earth and Environmental Science.

Suryavansh, S., Benna, A., Guest, C., & Chaterji, S. (2021). A data-driven approach to increasing the lifetime of IoT sensor nodes. *Scientific Reports, 11*(1), 22459.

Unold, O., Nikodem, M., Piasecki, M., Szyc, K., Maciejewski, H., Bawiec, M., . . . Zdunek, M. (2020). *IoT-based cow health monitoring system.* Paper presented at the International Conference on Computational Science.

Venkatraman Jagatha, J., Klausnitzer, A., Chacón-Mateos, M., Laquai, B., Nieuwkoop, E., van der Mark, P., . . . Schneider, C. (2021). Calibration method for particulate matter low-cost sensors used in ambient air quality monitoring and research. *Sensors, 21*(12), 3960.

Yan, X., Xiong, W., Hu, L., Wang, F., & Zhao, K. (2015). Missing value imputation based on gaussian mixture model for the internet of things. *Mathematical Problems in Engineering, 2015*.