Ingeniería y Competitividad
Vol 27 (2) 2025 doi: 10.25100/iyc.v27i2.14876

Vol 27 (2)

ISSN
0123-3033
e- 2027-8284

Review Article

Hybrid Models and Technological Configurations for Forecasting Soil-Meteorological Variables: An Exhaustive Review

Modelos híbridos y configuraciones tecnológicas para pronóstico de variables suelo-meteorológicas: revisión exhaustiva

Giovanni Alexander Cuaran Páez[1] Luis Gabriel Lafaurie Ponce[1-2] Jenny Lucía Huertas Delgado[1-2] Daniel Alejandro Molina Cuaichar[1]

[1] Programa de Ingeniería Ambiental, Universidad Mariana, Pasto, Nariño, Colombia.
[2] Grupo de investigación Ambiental (GIA).

How cite?

Cuaran GA, Lafaurie LG, Huertas JL, Molina DA. Hybrid Models and Technological Configurations for Forecasting Soil-Meteorological Variables: An Exhaustive Review. Ingeniería y Competitividad, 2025, 27(2)e-30514876.

https://doi.org/10.25100/iyc.v27i2.14876

Received: 05/06/25
Revised: 06/10/25
Accepted: 08/15/25
Online: 08/26/25

Correspondence

jlhuertas@umariana.edu.co

Spanish version

Abstract

Introduction: Homemade designs of climate and soil monitoring stations have shown significant advances in recording, processing, calibration, and forecasting applications, particularly in sustainable agriculture. Technological development has promoted the use of low-cost air and soil sensors, enabling the creation of useful databases to improve measurements and spatio-temporal predictions.
Objective: To provide an exhaustive analysis of technological configurations, applications, trends, quality, and research needs in soil–meteorological monitoring and forecasting.
Materials and Methods: A review of published scientific articles was conducted, complemented with statistical analyses and data mining techniques, particularly hierarchical clustering, to identify patterns and relevant approaches in the field.
Results: Studies indicate that the calibration of air and soil sensors through mathematical, statistical, and artificial intelligence methods has enabled the generation of reliable records. These data have been used to predict weather and soil conditions, as well as future levels of nutrients and contaminants. Moreover, there is a growing trend toward the use of hybrid models that combine two or more forecasting methods, together with meteorological and edaphic stations equipped with diverse sensors.
Conclusions: The integration of low-cost technologies and hybrid models strengthens the accuracy of soil–meteorological records and forecasts. However, challenges remain in improving data quality and consolidating research trends to guide future advancements in the field.

Keywords: Environmental analysis, precision agriculture, sensor, big data, artificial intelligence

Resumen

Introducción: Los diseños de estaciones caseras de monitoreo del clima y del suelo han mostrado avances significativos en aplicaciones de registro, procesamiento, calibración y pronóstico, especialmente en agricultura sustentable. El desarrollo tecnológico ha impulsado el uso de sensores de aire y suelo de bajo costo, lo que ha permitido conformar bases de datos útiles para mejorar las mediciones y predicciones espacio-temporales.
Objetivo: Analizar de manera exhaustiva las configuraciones tecnológicas, aplicaciones, tendencias, calidad y necesidades de investigación en el monitoreo y pronóstico de variables suelo-meteorológicas.
Materiales y Métodos: Se revisaron artículos científicos publicados y se aplicaron herramientas de análisis estadístico y minería de datos, particularmente clustering jerárquico, para identificar patrones y enfoques relevantes en el campo.
Resultados: Los estudios muestran que la calibración de sensores de aire y suelo mediante métodos matemáticos, estadísticos y de inteligencia artificial ha permitido obtener registros confiables. Estos datos se emplean en la predicción de condiciones meteorológicas, del suelo y de niveles futuros de nutrientes y contaminantes. Asimismo, se observa un crecimiento en la utilización de modelos híbridos que combinan dos o más métodos de pronóstico, junto con estaciones meteorológicas y edáficas equipadas con diferentes sensores.
Conclusiones: La integración de tecnologías de bajo costo y modelos híbridos fortalece la precisión de los registros y pronósticos suelo-meteorológicos. Sin embargo, se identifican desafíos en la mejora de la calidad de los datos y en la consolidación de tendencias que orienten la investigación futura.

Palabras clave: Análisis ambiental, Agricultura de precisión, Sensores, Macrodatos, Inteligencia Artificial.

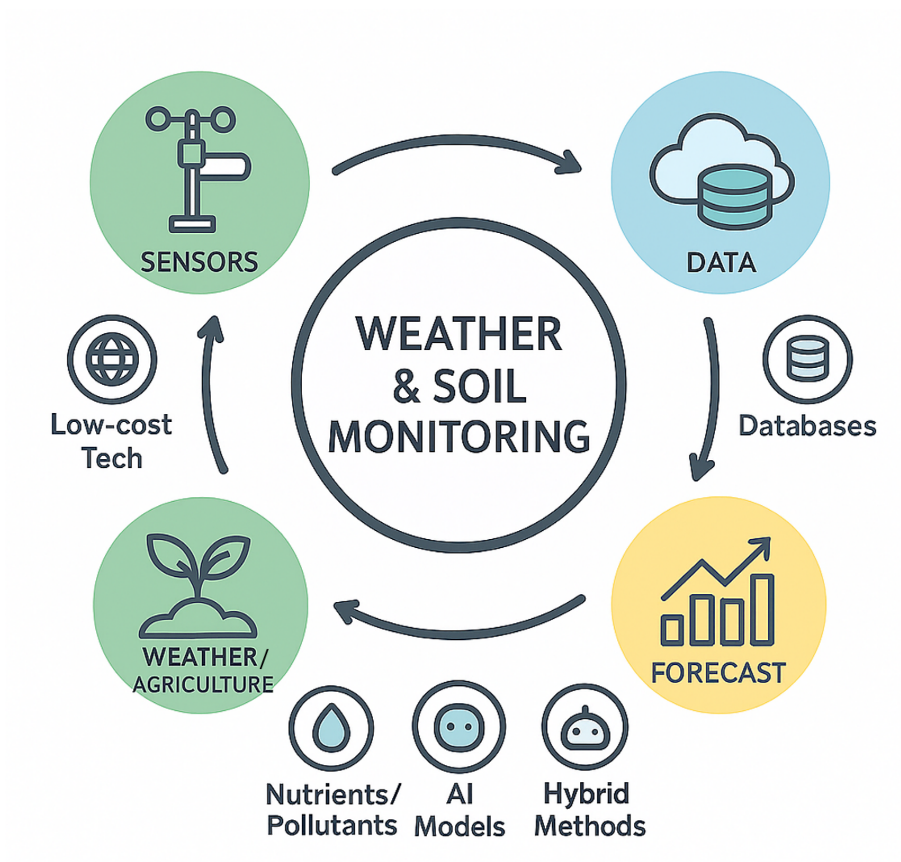
Contribution to the Literature

Why was it conducted?

Recording pollutants like particle matter in air or heavy metals in soil is essential to either estimate contamination levels or manage mitigation policies. Many current researches have proposed monitoring devices using one or more sensors types to correlating and predicting pollutant concentrations. Nevertheless, there are a lack of information about the monitoring stations with standard weather or soil variables that support efficiently the assessment of the contamination dynamic. Therefore, we conducted this research to identify the main technological configurations of the air and soil pollutants monitoring stations grouping statistically similar measurement protocols and devices, utilized by different works.

What were the most significant results? What do these results contribute?

The main technological configurations of air/soil sensors in monitoring stations were identified for pollution, weather and soil recording, and weather/soil forecasting applications. The main hybrid forecasting models used to predict weather/soil conditions and pollutants have been identified. The research trends and gaps regarding the year, location, and uses of the main weather/soil monitoring stations and forecasting models have been assessed and proposed for future works. The authors introduce the concept of holistic monitoring stations, grouping measurements of weather conditions and pollutants, nutrients, and physical-chemical properties in air and soil, presenting a novel approach to comprehensive environmental insights. By incorporating works about spatial and temporal forecasting of pollutants and variables, this research contributes to informed decision-making in sustainable agriculture and environmental management, fostering a healthier future for our planet.



doi: 10.25100/iyc.v27i2.14876

Introduction

Air and soil variable measurements have been used in diverse applications, including weather forecasting, agriculture, disaster prediction, air/soil diagnosis and pollution forecasting. These records are typically obtained by independent solid-state sensors or robust systems like weather and soil stations. Rapid technological development and the available information on forums or websites have enabled the development of low-cost homemade systems capable of measuring various weather conditions such as air temperature, air humidity, atmospheric pressure, wind direction, wind speed, precipitation, or irradiance. Additionally, these systems can measure soil variables like soil temperature, soil moisture and pH. Researchers have reported that the mechanical-electronic design of these measurement systems ensures their accuracy through a calibration and adjustment process. Alternatively, some systems process data using embedded mathematical models, statistical techniques, artificial intelligence (AI), and hybrid methods in microcontrollers. For example, greenhouses and small farmers have integrated local automated measurement of pH, moisture, precipitations and temperature sensors, integrated with coupled feedback control systems, to manage plant growth and irrigation effectively. Moreover, some farms include anemometers and small weather stations to record local weather.

Although these advantages are affordable, many researchers still rely on air/soil data from databases and robust weather/soil stations. This choice is because these long-time and spatially distributed signals provide enough information about local weather and soil conditions. However, it is important to note that their storage and management are expensive. Furthermore, the processing of data using specialized software often requires time to enable a deep understanding of the time/spatial evolution of the air/soil physical and chemical properties. It is noticeable that both weather stations and data platforms are usually publicly available, facilitating widespread access and utilization unlike soil monitoring stations and their data, which are primarily used for specific research purposes and are not frequently found in public data platforms. This limited accessibility to soil data can pose challenges for researchers seeking comprehensive and integrated datasets for their environmental studies.

Climatic change is currently a pressing issue because it leads to global temperature rise, variations in rainfall patterns, decreased snow cover and glacier shrinking (1). Extensive research has been conducted to assess the correlation between these impacts, air/soil pollution and weather/soil conditions. Temperature changes, for example show a high positive correlation with particulate matter concentrations of diameter size lower than $2.5\ \mu m$ and $10\ \mu m$ ($PM_{2.5}$ and PM_{10} respectively) and diverse pollutants such as sulfur dioxide (SO_2) or nitrogen dioxide (NO_2) emissions. Conversely, precipitations exhibit a low negative correlation with these pollutants (2). Weather stations and climatic sensors are used to measure and record temperature, humidity, atmospheric pressure, wind speed and wind direction data. These datasets are then utilized in forecast models to predict the dynamic behavior of these conditions and assess the behavior of pollutants (3–6).

On the other hand, air and soil pollution has become a significant global public health concern due to the high levels of pollutants released by anthropogenic activities including modern industries, vehicular emissions and urbanization (7,8). Pollutants can have negative health impacts even at low concentrations, increasing the risk of cardiovascular diseases, and respiratory and pulmonary

doi: 10.25100/iyc.v27i2.14876

infections, and weakening the immune system (9–11). Regarding air pollution, in 2015, the World Health Organization (WHO) recognized air pollution as a risk factor for noncommunicable diseases such as ischemic heart disease, stroke, chronic obstructive pulmonary disease, asthma and cancer (12). Moreover, correlations between air pollution and diseases have been found, with overlapping time-series of particulate matter with a diameter less than $2.5 \mu m$ ($PM_{2.5}$) and mortality due to Covid-19 in Europe (13). On the other hand, ingestion of polluted vegetables and soil poses a risk factor for gastric and gastrointestinal disorders due to high metal concentrations such as lead (Pb), zinc (Zn), copper (Cu), cadmium (Cd), arsenic (As), nickel (Ni) and chromium (Cr) in certain crops (14,15).

Furthermore, air and soil pollutants have detrimental effects on the environment, regarding to climatic change. These effects include phenomena like acid rain, haze, the greenhouse effect, the ozone hole, photochemical smog, and a decrease in soil microbiota (biomass and microbial activity) (16–20).

Due to the reasons, the implementation of a pollutant release control management system becomes imperative. Rapid technological advancements in the Fourth Industrial Revolution have prompted researchers to consider digital technologies. These technologies have led the development and evaluation of forecasting models that employ time and spatial series records of weather and soil variables to predict the dynamic behavior of air and soil, including contamination levels. To achieve this, these models incorporate mathematical, statistical, and AI methods to process weather and soil databases, describing the past, present, and future (prediction) of air/ soil dynamics behavior. The findings from these models hold significant potential for supporting the application of environmental policies, reducing contamination levels in line with local environmental laws, and mitigating socio-environmental negative impacts. Thus, understanding the correlations between air/ soil record technologies, weather/soil variables, air/soil dynamics, air/ soil pollution, as well as evaluation: forecasting models is crucial in assessing the impact of both traditional and novel approaches on weather/soil quality improvement, environmental pollution reduction, and disasters prevention.

Researchers can greatly benefit from the results of this study to advance future research methodologies related to the development of both air/soil variables measurement systems and the embedded processing information techniques for weather or pollution forecasting models.

Motivation of the review

During the first two decades of the 21st century, the importance of predicting weather and soil conditions has significantly increased, meanly due to the strong correlations observed between measurements of air/soil physicochemical properties and pollutant levels, this knowledge has wide-ranging applications, including agriculture applications, disaster prevention, and ecological management (6,18,21). Researchers have dedicated their efforts to developing forecasting methods that allow them to obtain accurate weather, soil and pollutant records. Armed with this valuable information, environmental management policies can be developed to effectively reduce pollutant levels in compliance with the upper thresholds of local environmental laws (20,22–25). Additionally, these advancements have empowered farms to enhance crop quality by optimizing planting and

doi: 10.25100/iyc.v27i2.14876

harvesting periods, determining appropriate plant irrigation schedules, and managing pesticide concentrations, these improvements are achieved through the application of control and forecast systems that utilize weather/soil records as predictors (21,26).

On the other hand, numerous reviews have been conducted to both enumerate and assess the advantages and limitations of different forecasting methods (3,27–30). In a recent study, Alexander Baklanov and Zhang (31) recognized the existing information gap regarding Air Quality Forecast (AQF) mathematical models in previous reviews. Their study evaluates the advantages and accuracy of utilizing multiscale forecasting systems, weather process models, and atmospheric emission models such as Seamless Environmental Prediction Systems (SEPS), Coupled Chemistry-Meteorology Modeling (CCMM) and Ensemble Forecast of Analyses (EFA). This approach involves the integration of various data sources and mathematical models of different types with alternative assumptions, using distinct pattern recognition methods (32). Their contribution improves the AQF by incorporating online coupling of atmospheric dynamics and chemistry, improved data assimilation and fusion, machine learning, multi-scale prediction approaches and sub-seasonal to seasonal forecasting.

Forecast models are also used in agriculture. For example, Thorat et.al (21) identified suitable insecticides and fertilizers to be applied to crops through the processing of pest images using the Transition Probability Function (TPF) and Convolutional Neural Network (CNN) methods. This approach allowed them to identify suitable insecticides and fertilizers for crops. The system integrated Nitrogen, Potassium and Phosphorus concentration measures (these are individual and synchronous measures, usually named by their chemical symbols NKP), pest images and information on the main insecticides used, enabling it to predict and recommend the best insecticide for pest control. Khaydukova et.al (33) employed a similar methodology, but they also considered pH, electrical conductivity (EC) and organic carbon as additional factors to evaluate. These values were correlated using multivariable regression techniques.

On the other hand, Lu Bai et.al (34,35) reviewed the theory and applications of hybrid forecasting models, compassing statistical, AI and numerical methods, along with their theory and applications. Furthermore, they compared and evaluated the advantages and disadvantages of different combinations of these methods. Statistical models have a wide range of environmental applications and require less processing time; however, a large amount of time series data is needed. AI models, on the other hand, improve forecasting with relatively smaller amount of data; but, their algorithm can be unstable. Hybrid models, which combine statistical and AI approaches, offer robustness, high adaptability and a low risk of failure; however, its design is relatively complex.

Other researchers have focused on describing specific methods to address information gaps in environmental forecasting models (17–19). Hui Liu et.al (19) assessed the accuracy, advantages and limitations of data spectral decomposition methods that involve frequency components. They explored techniques such as Wavelet transformation, and variational and empirical decomposition, in conjunction with forecasting models such as Artificial Neural Networks (ANN), Support Vector Machines (SVM) and Extreme Learning Machines (ELM).

doi: 10.25100/iyc.v27i2.14876

Despite, some reviews collect and assess different weather/soil variables and their forecast methods, there remains an information gap concerning the correlations and grouping between predictors, methods, forecast models, quality and other features. This review aims to address this gap by summarizing and correlating weather/soil measurement methods, applications and forecasting weather/soil methods used in different research, aiming to answer the question:

What are the main technological configurations of weather/soil stations, and their main hybrid forecast models used in different air/soil applications?

Organization of the Review

In this review, we first compile and organize the essential information from fifty research documents. Next, we discuss the applications of weather/soil records including devices, communications methods, storage techniques and forecasting methods. Also, we provide detailed descriptions, geographical locations, years of study, and commonly employed communications approaches.

Subsequently, we present a critical description of the primary electronic configurations and devices, applications, embedded systems and hybrid forecast models used in different research. This information is supported by results from multivariate analysis including assessment of normality, principal components analysis, correlation and hierarchical clustering. Finally, conclusions and recommendations are provided.

Materials and Methods

Literature search and selection strategy

For our study, we have built a search equation using the following structure: “Station type” AND “Weather or soil condition” AND “Programming method” AND “Communication system” AND “Forecast method”. Each classifier contains the following combination of keywords:

“Station type”: weather station OR soil station

“Weather or soil condition”: temperature OR humidity OR precipitation OR pressure OR moisture OR pH OR NPK OR sensors OR record

“Programming method”: Microcontroller OR Python OR SQL OR MySQL

“Communication system”: WiFi OR IoT OR MicroSD OR Bluetooth OR Xbee OR zigbee

“Forecast method”: Autonomous learning OR forecasting OR air quality OR soil quality

Databases like Google Scholar, Scopus, Scielo, Science Direct, IEEE Xplore, and digital research institute repositories were used as data sources, where the search equation was applied to obtain a set of research documents relating to the measurement of weather and soil conditions, which includes technologies such as soil-weather stations or embedded systems. Furthermore, we have included papers with applications of traditional (calibration) and forecasting models to assess or predict air and soil quality, using environmental time and spatial records as predictors. A statistical sample of 50 random research documents from the largest set obtained was considered.

doi: 10.25100/iyc.v27i2.14876

Table 1. Classification of different research

Station type	Weather or soil condition	Programming Method		Communication system		Forecast method		Ref.
		One	Two or more	One	Two or more	One	Two or more	
Weather and soil	Air temperature, moisture		✓		✓			(36)
		✓		✓		✓		(37)
		✓			✓			(26,38)
				✓				(39,40)
					✓			(41,42)
					✓	✓		(43)
		✓			✓	✓		(44,45)
	Air temperature, Humidity			✓				(46)
					✓	✓		(47)
Weather	Air temperature, Humidity, gases	✓		✓		✓		(48)
						✓		(49,50)
		✓		✓				(51)
							✓	(30,52,53)
		✓			✓		✓	(5)
	Air temperature, Humidity, Precipitation		✓	✓		✓		(54)
				✓		✓		(55)
		✓		✓		✓		(56–58)
				✓			✓	(59)
							✓	(60)
		✓					✓	(61)
	Gases				✓		✓	(62)
							✓	(23,63,64)
		✓					✓	(65)
	Air temperature, Humidity, Pressure	✓		✓				(66–68)
						✓		(69)
		✓				✓		(70)
	Humidity				✓			(71)
Soil	NKP			✓				(72)
		✓		✓		✓		(21)
	Soil pH, NKP			✓		✓		(73)
		✓				✓		(33)
				✓				(74)
		✓			✓			(75)
	Moisture, soil pH	✓				✓		(76)
				✓				(77)
		✓		✓				(78)
	Soil temperature	✓		✓		✓		(79)

The research documents were categorized into the following classifiers: type of station (weather or soil), weather and soil conditions, type of software, type of communication, forecasting method,

doi: 10.25100/iyc.v27i2.14876

year, and place (Table 1). We organized the information into a matrix with each classification as columns and the number of documents as rows.

Statistical processing

Frequency was calculated for each feature to identify similar characteristics among the researches. A Shapiro-Wilk test and Kendall correlation were performed to determine and evaluate the normal distribution level and correlations, respectively, for weather/soil conditions, programming method, communication system and forecast method. Subsequently, a Principal Component Analysis (PCA) was applied to weather/soil and forecast model data to group different air/soil sensors into five main weather/soil stations based on the principal components obtained. Additionally, different forecasting methods were clustered into eleven hybrid forecast models, considering the trend toward achieving accuracy with lower time processing and storage requirements through complementary methods.

Finally, Kendall correlation and Hierarchical type ward clustering were calculated between weather/soil stations, hybrid forecast models, publishing year and continent data to assess the main configuration uses.

Main weather/soil measurements and monitoring stations

In Figure 1 a), the frequency of weather records used in papers is shown. Among the weather variables, temperature exhibits the highest variability, followed by relative humidity, irradiance, wind speed and pressure. These variables readily measured and recorded due to there are many inexpensive, affordable and accurate sensors available in both robust weather stations and homemade designs. Temperature and relative humidity are frequently measured together, and they exhibit a strong Kendall correlation (0.69 with $p < 0.05$) (see Appendix). Commercial sensors often integrate devices that generate electric potential in response to changes in temperature and humidity. Consequently, these two variables are often observed together. As a result, these variables are often found together since relative humidity depends directly and closely on the ambient temperature and the amount of water vapor holding in the surrounding air (80,81). Similarly, wind speed shows a strong correlation with wind direction (0.74). Additionally, the temperature has a considerable correlation with precipitation, pressure, irradiance and speed of wind (0.34, 0.41, 0.47, 0.47 and 0.38, respectively). Thus, temperature is commonly measured jointly with other weather conditions. However, while weather conditions are mainly correlated, measures of Ultraviolet Light (UV) exhibit week correlations with other weather conditions.

doi: 10.25100/iyc.v27i2.14876

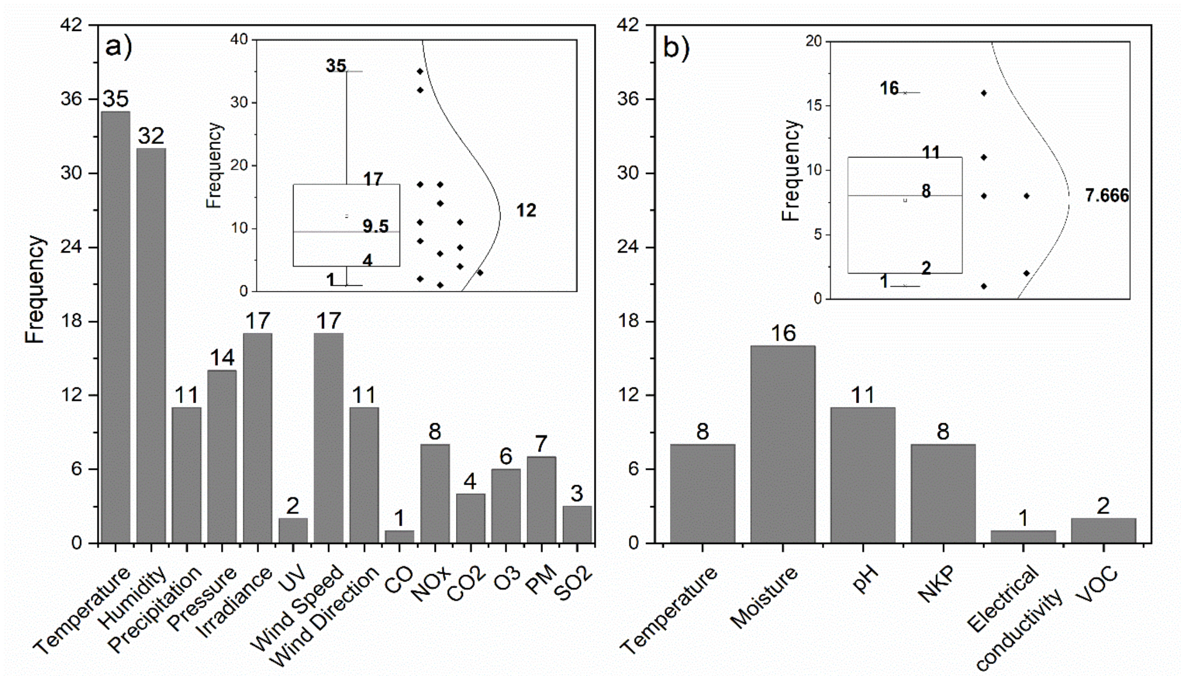


Figure 1. Histogram, box plot and normal curve of a) weather and b) soil variables.

The box plot and normal curve analysis indicate that gas pollutants such as NO_x (Nitrogen oxides with $x = \{1,2\}$), CO_x (Carbon oxides with $x = \{1,2\}$), $PM_{2.5}$ or PM_{10} have lower frequency. Pollutants can vary depending on their location; thus, the pollutants that are the object of various research are mainly different and are grouped close to a frequency of 1. Furthermore, data distribution for these pollutants is non-normal, as evident from the Shapiro-Wilk test ($p < 0.05$). The geometric mean of the data is 8.0 with a standard deviation value of 10.9, indicating that between $-2.9 \approx 0$ and $18.9 \approx 19$ documents out of the total considered, account for 95% of weather variables. Interestingly, approximately 24% of the papers considered in the analysis account for 85.71% of the analyzed weather variables. This suggests that a significant portion of the research focused on a relatively small number of weather variables, temperature and humidity measures stood out as the main variables in approximately 80% of the papers. These variables play a crucial role in assessing air quality and predicting atmospheric conditions due to their direct impact on various environmental processes and human health.

Soil variables are less frequently measured, as shown in Figure 1 b); however, moisture has significant research interest in determining soil quality. Even, in systems that include both weather and soil sensors, surface air quality is influenced by moisture and surface temperature (71,73–76,79). The distribution data for soil variables is found to be normal, as indicated by the Shapiro-Wilk test ($p > 0.05$), with a mean value of 7.7 and standard deviation value of 5.6, indicating that between $2.1 \approx 2$ and $13.3 \approx 13$ documents out of the total considered, account for 95% of soil variables analyzed. This result contrasts with the Kendall correlation test (see Appendix), as only soil EC with Volatile Organic Compounds (VOC) shows a high correlation (0.70).

Therefore, applications or devices that jointly record different soil variables are not common. Soil temperature and moisture exhibit a correlation of 0.40 ($p < 0.05$), indicating a low use of hybrid

doi: 10.25100/iyc.v27i2.14876

soil temperature-moisture sensors by researchers, in contrast to hybrid air temperature-humidity sensors. It is important to highlight that soil moisture, or the method by which this variable is determined within the soil, is directly related to the dielectric constant of water and is not influenced by temperature as in the case with environmental relative humidity. In this regard, Topp et. al (82–84) indicate that through the use of appropriate calibration curves, the measurement of the dielectric constant can be directly linked to soil moisture.

In specific case, EC is recorded in only one paper, mainly due to the unaffordable cost, expensive and complexity of designing homemade sensors. Despite this, commercial electronic devices that include pH, NKP and EC measurements, are also expensive. Therefore, current efforts are focused on describing soil conditions through the processing of pH and NKP records using multi-measurement soil sensors, as shown in Table B.1, where the NKP and soil pH measures in papers have a correlation of 0.56 ($p < 0.05$), which is higher than the correlation between NKP and soil EC (0.33).

EC serves as an indirect indicator that allows for estimating the concentration of dissolved salts in the soil the value of this variable is significantly influenced by soil moisture, where low moisture contents indicate higher EC values due to the amount of salts remaining constant, but with lower water content in the solution, salts often result from chemical fertilization, being typically composed by macronutrients such as nitrogen, phosphorus and potassium. Additionally, chemical fertilizers can influence the increase or decrease of the pH values in soils, depending on the nature of the applied fertilizers.

Programming and communication protocols in weather/soil applications

The acquisition of temporal or spatial records of weather or soil variables depends on different sources, whether they are commercial or homemade designs. Additionally, these records rely on the embedded programming software used for storing and processing weather and soil data. Due to the availability of private and public weather stations, such as the World Air Quality Index project (85), obtaining records without adding electronic devices has become straightforward. As a result, data is often processed using specialized software such as Python, Java, QGIs, JSON, NCEP CFSv2 or MYQL (Figure 2 a)). Specialized software offers the advantage of reducing the programming and processing time while enhancing data accuracy. Likewise, embedded systems provide data accuracy at a lower cost but come with longer programming and processing time. Therefore, researchers must choose between cost efficiency and processing time to achieve successful results.

doi: 10.25100/iyc.v27i2.14876

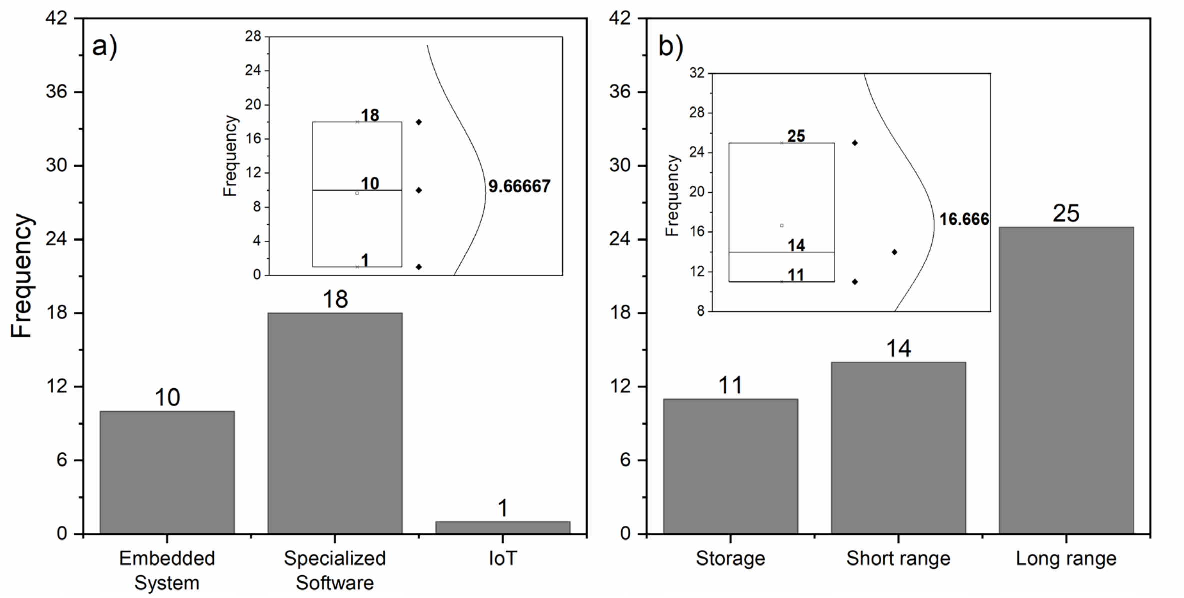


Figure 2. Histogram, box plot and normal curve of a) programming and b) communication.

Research focusing on weather station development often involves embedded sensors and microcontrollers to obtain raw and processed weather records generally (58,86,87). These designs vary from traditional weather stations to Unmanned Aerial Vehicle (UAV) applications, such as weather drone incorporation (88). Low-cost technology integration is a common feature in these applications. For example, supply systems like photovoltaic cells or Li-ion batteries are included to improve device performance used in crop monitoring, intelligent agriculture, pollution detection and nutrient analysis (5,36,42,45). In contrast, in researches where the processing information for forecast weather conditions is the main aim, specialized software is predominantly employed (89,90). IoT applications for programming are less frequent, as data processing is typically done with software embedded in available hardware. Programming types exhibit a normal distribution (Shapiro-Wilk test with $p > 0.05$), with a mean value of 9.7 ± 8.5 indicating that approximately 20% of the works utilize embedded systems, specialized software and IoT for data recording and processing. A Kendall correlation test conducted on programming types (see Appendix) does not reveal any significant correlation among them, Thus, only one method is applied for recording and processing data.

Weather and soil data are typically presented in tables, graphics or figures through graphical user interfaces (GUIs), enabling the assessment and analysis of the dynamic behavior of variables. This information needs to be accessible and processable by software; hence, the storage and forwarding of databases and spatial-time signals are essential. Microcontrollers such as Arduino or Microchip devices have limited internal memory that can be used to store weather, soil and pollutant records directly. However, as the memory fills up, the processing power decreases. Although a greater memory capacity can be found in some microcontrollers, incorporating external devices is a better choice due to their larger memory capacity, easy decoupling and robustness. For example, Chacón and García (36), utilized SD cards to store temperature, humidity, atmospheric pressure, brightness, UV, precipitation and wind speed to prevent information loss. Similarly, other authors (36,72) have used EEPROM devices and databases (Table 2) to ensure the availability of weather and soil spatial-

doi: 10.25100/iyc.v27i2.14876

time signals, which can be suitable for assessing historical air/soil quality or can be employed AQF and Soil Quality Forecast (SQF).

Similar to the programming method, there are no significative correlations (see Appendix) between different range types of communication (Figure 2 b)). The communication methods exhibit a normal distribution (Shapiro-Wilk test $p > 0.05$) with a mean value of 16.7 ± 7.4. Thus, 33.33% of the research documents utilize both short and long rage communications jointly with storage devices.

Table 2. Communication protocols

Storage	Short range	Long range
Micro SD	Radio Frequency	GSM/GPRS
EEPROM	Zigbee	WiFi
Database	Xbee	Web site
	Bluetooth	

While data storage in different electronic devices preserves the information, efficient and secure communication between embedded software on microcontrollers, GUIs and websites is crucial. Wireless communication protocols have been implemented to facilitate the transfer of data from storage devices to nearby or distant data processing devices. Currently, wireless communication protocols find wide application in biotechnology, agriculture, ecology and military operations due to their easy setup and relatively unlimited communication range. Remote applications with GUIs display real-time data from weather/soil monitoring stations as a graphical time series, allowing users to observe the behavior of the weather/soil records. Fast wireless systems are commonly integrated to enable this functionality.

Table 3. Features of the communication protocols (91,92)

	Communication protocols		
	Zigbee/Xbee	**Bluetooth**	**WiFi**
Operation bandwidth	2.4 GHz	2.402 – 2.480 GHz	2.4 – 5 GHz
Range	70- 400 meters	1 - 100 meters	1 – 90 meters
Network topology	Mesh network	Adhoc piconets	Unlimited nodes (devices)
Number of devices per network	2 - 65000	2 - 8	Unlimited
Maximum transmission speed	250 Kbps	1 Mbps	9.6 Gbps

Zigbee systems have been used to send data to generate short-time local alerts utilizing mathematical models based on weather parameters recorded from nearby monitoring devices, taking advantage of its frequency of approximately $250\ kbps$ (Table 3). Zigbee enables rapid data

doi: 10.25100/iyc.v27i2.14876

transmission facilitating real-time management of air/soil quality, improving parameter settings in precision agriculture, pollutant assessment, forecasting methods or weather conditions analysis (33,38,40,50).

Xbee is another communication protocol with features similar to Zigbee; however, their architecture and trademark are different. Xbee applications primarily focus on weather/ soil monitoring using cost-effective wireless stations in farms with communication distances up to 400 meters. For example, Devaraju et.al. (58) designed a weather station with Modbus communication protocol using Xbee modules. Data is transmitted to a software application on PC/laptop; and uploaded data to an online server. On the other hand, greenhouse and crop management have been improved through data transmission between sensors and control systems. Xbee modules are employed to transmit temperature, moisture and nutrient measurements, enabling the assessment of optimal real-time irrigation conditions for plantations like coffee crops (39,71).

Low-cost and portable weather/soil stations are frequently installed in rural areas, typically covering an area up to 100 meters; therefore, Bluetooth devices are used to establish wireless connections between the monitoring stations and smartphones or PC/laptops. This setup improves water management, particularly in irrigation practices, and enables yield predictions in precision and intelligent agriculture applications (26,43).

The aforementioned applications are addressed by other researchers using long-rage communication protocols. Currently, data transmission through WiFi and storing the data in the cloud or web pages have emerged as affordable and inexpensive options for transmitting and processing weather/soil information. Installation of devices developed within the framework of the 4th industrial revolution such as IoT, Big Data and AI; enables the rapid information exchange across geographically distant devices. In this way, weather/soil data sent can be processed by specialized laboratory software, allowing optimized agriculture and atmospheric management (57,68,74).

Forecasting methods in air/soil quality and pollutants applications

The primary aim of different research works is to describe weather or soil conditions using measurements or databases. Furthermore, most authors have focused their works on predicting the behavior of these conditions. The use of the different forecasting methods does not follow a normal distribution (Shapiro-Wilk $p < 0.05$) because there are numerous prediction techniques without a clear dominant approach. Therefore, the frequency of use for each forecasting method is generally low. On average, 5.15% of the research applied 92.31% of the forecasting methods (Figure 3), meaning that works could use one or more prediction models together. Performance of AI algorithms could be improved by integrating statistical and mathematical models to achieve better accuracy in the forecast.

From Figure 3, forecast models are grouped into three main sets: mathematical, statistical and AI models. Knowledge-based or mathematical models are commonly applied due to their high analysis potential based on physical and chemical axioms and concepts. For example, Ramadan et.al (79) utilized Frequency Domain Analysis (FDA) to improve the measurement of soil dielectric constants in value-added soil stations. Zeng et.al (3) estimated $PM_{2.5}$ concentrations between spatially

doi: 10.25100/iyc.v27i2.14876

separated monitoring points, using Spatial Interpolation (SI) combined with machine learning algorithms. Espinosa et.al (52) predicted NO_2 concentrations in southeastern Spain by identifying and optimizing multiple linear regression models through a Multi-Objective Optimization (MOO) method, utilizing a learning algorithm to forecast pollution level and minimize the error of each linear regression. In these works, successful accuracy results were obtained, but generally, the processing time is noticeable, and they do not consider some variables or noise factors, which could decrease the forecast accuracy. Therefore, only 6% of the works employed mathematical models (Figure 3).

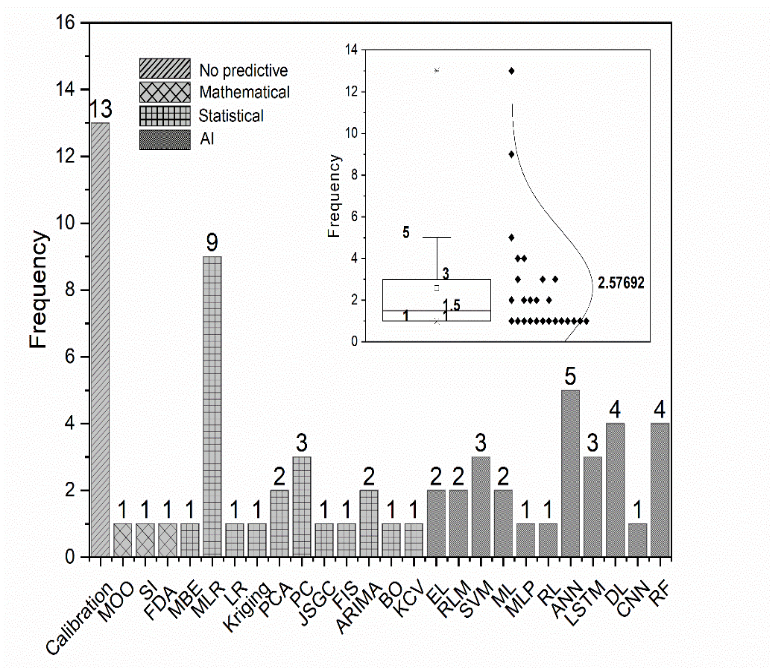


Figure 3. Histogram, box plot and normal curve of forecasting methods

Statistical models are more widely used, accounting for 46% of the research, due to the shorter processing time while maintaining reasonable accuracy, but the physical/chemical axioms are lost. It is noticeable that one common approach in different works is the use of one or more linear regressions, named Multilinear Regression (MLR). For example, Jeong et.al (70) employed MLR to correlate PM_{10} concentrations with climatic variables in wintertime achieving skillful seasonal forecasting. Furthermore, researchers have utilized various correlation techniques such as Jammalamadaka-SenGupta Correlation (JSGC) and Pearson Correlation (PC), as well as spatial estimation techniques like Kriging, Fuzzy Interference System (FIS), time series analysis like Autoregressive Integrates Moving Average (ARIMA) and optimized systems like Bayesian Optimization (BO). Of particular interest, many researchers have evaluated and validated AI applications through the use of K-fold Cross-Validation (KCV). The method guarantees that the results are independent of the training and probe data (59,60).

Regarding AI models, different documents and papers explicitly list the AI algorithm used such as multilayer perceptron (MLP), Long Short-Term Memory (LSTM), Ensemble Learning (EL) and Support Vector Machines (SVM). However, other works only mention Machine Learning (ML) types

doi: 10.25100/iyc.v27i2.14876

grouping similar algorithms. In fact, the term "ML" is rarely used on its own. For example, Bernardes et.al (59) and Yamamoto et.al (60) probe different regression methods including ML models to calibrate low-cost weather sensors as a group of non-specific methods applied. In the same context, ML paradigms are sometimes mentioned as a single type of AI algorithm; but the paradigms group different algorithms with similar characteristics. Supervised learning algorithms search for the best approximate function that fits a specific dataset of input and output values. Unsupervised learning algorithms, on the other hand, use only input data to determine patterns or trends without knowledge of the output data. These patterns are assessed in clusters to identify the best approximations. Reinforcement Learning (RL) algorithms are used when input and output values are unknown; instead, a set of rules is considered to assess a set of random events for the algorithm to learn about multiple ways to achieve an aim.

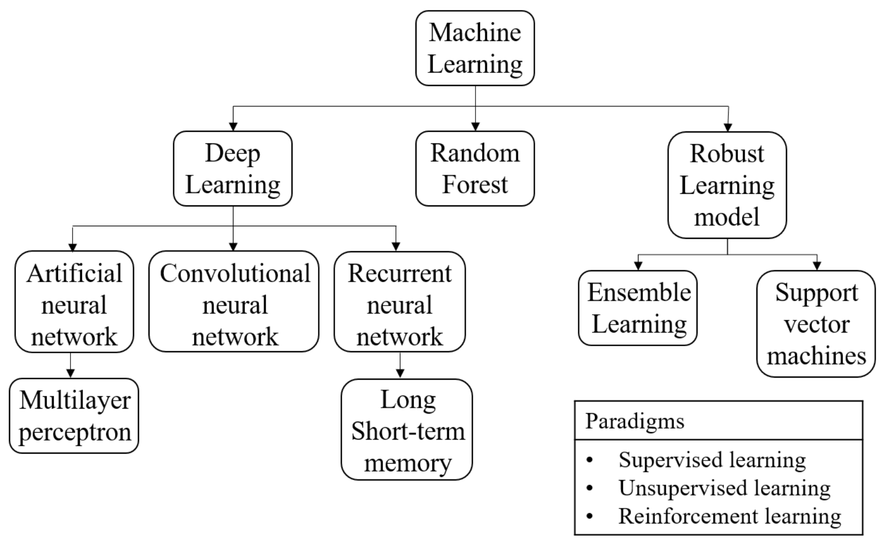


Figure 4. Types of machine learning algorithms and their paradigms

Research works without forecasting applications are labeled as "Calibration". Mainly, the validation of a monitoring station design is determined by the record of reliable measurement in different sensor devices (26% of the works). Reliability is achieved through calibration and adjustment processes. Nevertheless, there are some works where AI methods are applied to evaluate the calibration level in weather/soil stations, but they are not considered forecasting methods because calibration is the aim. Therefore, in this work, the term "calibration method" is used solely to refer to research works without forecasting application.

In Figure 4, the hierarchical order of the ML algorithms considered is shown. Deep Learning (DL) groups multilayer algorithms based on human brain functions, commonly called neural networks; it is often used to identify patterns, and its algorithms are designed under the unsupervised and reinforcement learning paradigms. The Random Forest (RF) method combines the output of multiple trained decision trees to generate a single result. Robust Learning Models (RLM) are a set of AI methods used jointly to obtain better predictive performance. Specifically, EL groups learning algorithms and SVM categorizes unlabeled data using a set of given training examples, based on a set of statistical learning frameworks.

doi: 10.25100/iyc.v27i2.14876

Main technological configurations and trends of the weather/soil stations

To identify the main weather/soil station configurations used by researchers, a PCA was conducted on the weather/soil sensor types (Figure 5). The variables were grouped into five principal water/soil stations. Weather/soil station 1 has the greatest eigenvalue (4.23), therefore, it is expected that common variables such as temperature, humidity, precipitation, atmospheric pressure, wind speed and wind direction are projected or grouped on this component. However, pH soil and NPK are included because there are applications where the soil's acidity level is tested and correlated with weather variables data, such as in crop irrigation management or intelligent agriculture (33,40,42,72). This result is supported by the correlation between the variables in Table A.1 and Table B.1.

On the other hand, weather/soil station 2 includes the measurement of the CO, CO_2 and $PM_{2.5} - PM_{10}$ concentrations. These gases are associated with pollution and the increasing prevalence of pulmonary diseases; thus, their applications are focused on assessing the AQF mainly. However, this result is an outlier because usually, there are not many researches that measure these variables jointly; instead, they are recorded as complementary and necessary (eigenvalue 1.65) along with other weather conditions to predict pollution levels and calibrate low-cost monitoring stations (23,62,64).

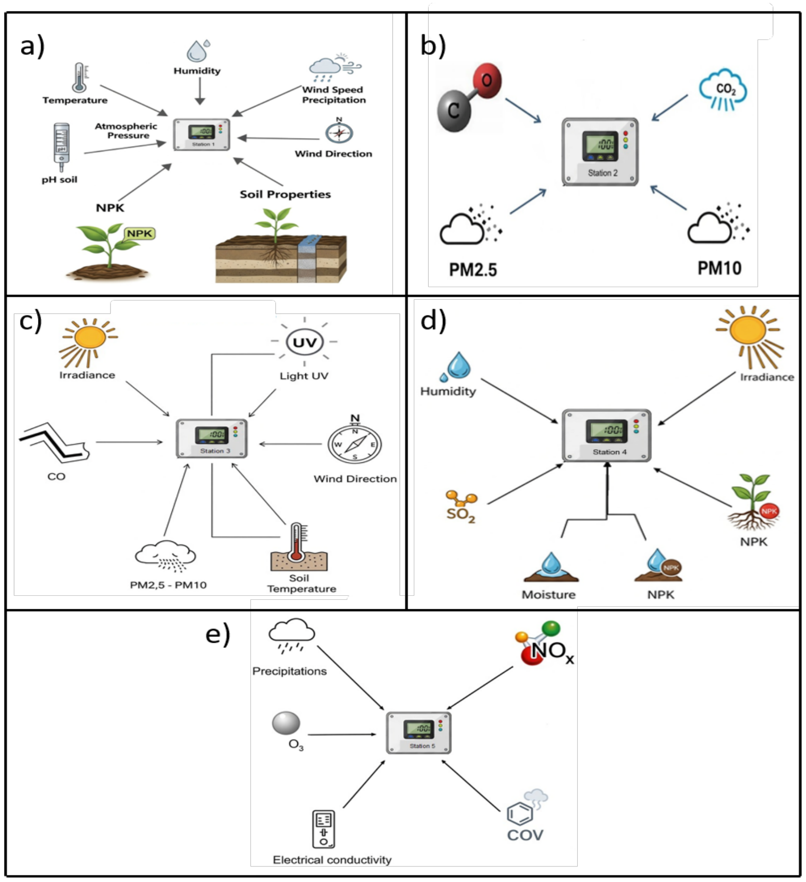


Figure 5. Technological configurations to weather/soil station a) 1, b) 2, c) 3, d) 4 and e) 5.

Likewise, weather/soil station 3 measures CO and $PM_{2.5} - PM_{10}$ concentrations while also adding records for irradiance, UV light, wind direction, and soil temperature records. Solar radiation

doi: 10.25100/iyc.v27i2.14876

features on the Earth's surface depend on different environmental factors and both incident energy and wide spectrum of radiation are attenuated by absorption and scattering processes in the atmosphere. Furthermore, changes in solar radiation have correlations with pollutant concentrations, such as $PM_{2.5} - PM_{10}$, as spectral transmittance decreases in the presence of haze and particulate matter in the air, due to photons collisions, with pollutants, ionized molecules and ions (93). Interactions like absorption, Compton scattering and photoelectric effect decrease the photon energy and increase the internal energy of atmospheric gases, which potentially can affect soil temperature. Additionally, wind direction contributes to energy transfer by convection, acting as an external factor in the thermodynamic behavior of the polluted air (45,63). The phenomenological interactions observed and the PCA results for weather/soil station 3 suggest an interesting research focus on assessing the correlation between solar radiation and pollutants.

Regarding Figure 5, weather/soil station 4 could be valuable in agriculture applications. Both humidity and moisture contents significantly affect crop sprout and growth. Low or high humidity levels can decrease the growth rate or increase the risk of diseases, respectively (40). Furthermore, monitoring soil nutrient concentrations, such as nitrogen, phosphorus and potassium is essential for controlling and improving the production and quality of crops in greenhouses or farms. PM, SO_2 However, these measurements are no enough (21,72). Pollutants in the air and soil, such as and heavy metal particles can negatively impact crops; in this case, SO_2 monitoring is considered, because insecticides used in fumigation processes may contain trace amounts of this gas which can negatively affect crop yield as seen in vineyards with excessive application (46). Irradian0.36 pollutants correlation was previously mentioned; also, irradiance has a direct correlation of with humidity (see Appendix); therefore, recording of irradiance is expected (52,63). Therefore, weather/soil station 4 has the potential to record, predict and control variables that affect soil quality and crop health.

Weather/soil station 5 stands out as a noteworthy result since it considers the measurement of both precursor elements of tropospheric ozone (or ground-level ozone), namely NO_x gases, and VOC produced by combustion or pesticide use. NO_x gases react with VOC in the presence of sunlight to create tropospheric ozone (O_3), which can lead to respiratory diseases or affect photosynthesis, growth rate, gene expression, and cell membrane function in plants (94,95). Therefore, recording and controlling these variables in SQF applications can be crucial to improving crop yield in farms or greenhouses. Precipitations and EC are additional indicators due to their correlation with pH and NKP (33,47).

doi: 10.25100/iyc.v27i2.14876

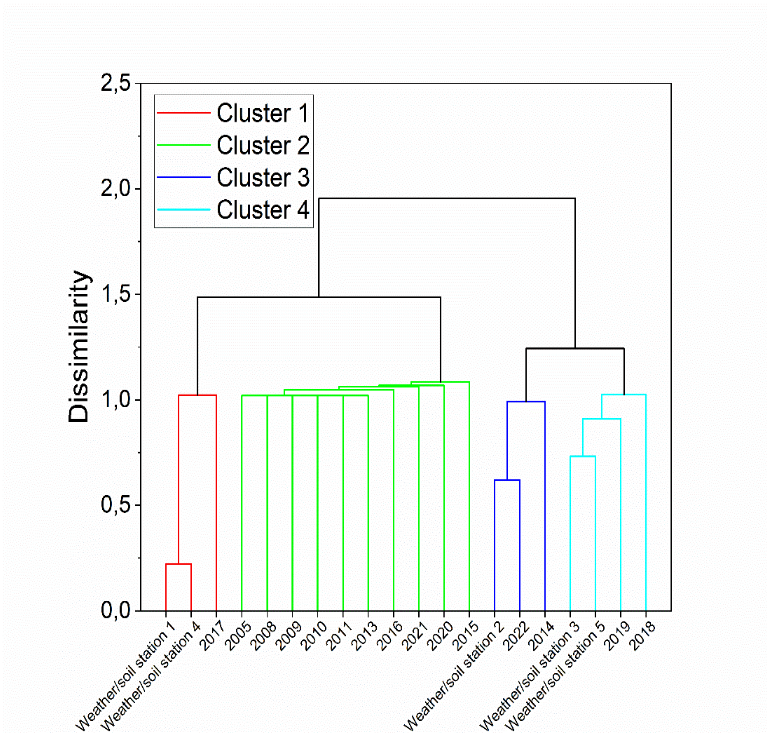


Figure 6. (Dendrogram) Grouping of weather/soil stations and publishing year

Hierarchical type ward clustering was calculated to group the weather/soil stations and the years into four clusters based on their predominant uses in different works. Ward clustering cuantify a dissimilarity coefficient, such that, lower values of dissimilarity means that two or more elements are significantly close or have common features.   Figure 6 shows that weather/soil stations 1 and 4 are usually used together (cluster 1: red line). This result is expected because there is a direct correlation between temperature-humidity and soil pH-NKP (see Appendix) measured in AQF and SQF research. Additionally, in 2017, there were 11 documents focusing in soil management applications mainly; therefore, works with recording of common weather/soil variables are represented by publishing year 2017 and weather/soil stations 1 and 4 in the same cluster. However, 2017 is not the only significant publishing year for this type of work, because raw signals have been recorded since 2005. It is noteworthy that cluster 2 (green line) does not include any weather/soil station and includes 10 publishing years, from 2005 to 2021. Nevertheless, cluster 2 is close to cluster 1 (Figure 6), suggesting that signals recorded with weather/soil stations 1 and 4 were made between 2005 and 2021, even during the COVID-19 pandemic period in 2020. Furthermore, devices in cluster 1 have a trend of being used in future works, despite the correlation between weather/soil stations 1 and 4, and the publishing year 2022 shows a significantly inverse trend (see Appendix).

Cluster 3 showed in Figure 6 (blue line) includes pollutant measurements obtained through weather/soil station 2, spanning the publishing years 2014 and 2022. This result aligns with expectations because weather/soil station 2 exhibits a significantly direct correlation with the year 2022 and a non-significant correlation with the year 2014 (see Appendix). Works in 2022 tend to predict the pollutant behavior using statistical and AI applications. For example, Gladkova and Saychenko (65,96) utilize Long Short-Term Memory (LSTM) AI and AutoRegresive Integrated

doi: 10.25100/iyc.v27i2.14876

Moving Average (ARIMA) statistical models to forecast the $PM_{2.5} - PM_{10}$ concentrations in 7 cities in Russia, taking advantage of the ability to interrelate events with an indefinite time lag. Similarly, other investigations incorporate AI, statistical and mathematical models to predict pollutant dynamics and concentrations in recent year (52,70). As a result, these types of works have garnered research attention due to their potential applications in public policy management focused on safeguarding the environment, food quality and public health.

In contrast, cluster 4 groups (cyan line) of monitoring stations used in soil or agriculture applications during the years 2018 and 2019. This observation aligns with the fact that there are fewer works using soil data for monitoring, controlling, or predicting soil conditions compared to the number of AQF research documents. It is noticeable that cluster 4 is in close proximity to cluster 3, indicating a research interest in recording and assessing soil data for farm and greenhouse applications at present. Furthermore, soil data could be complemented with pollutant measurements in future SQF works.

In terms of publishing continent, the hierarchical clustering results (Figure 7) show that researchers in Asia focus their works on the technological design of weather/soil stations for recording or controlling system applications (cluster 1: red line). On the other hand, works in South America tend to focus on AQF assessment through pollutants and their correlated variables, such as irradiance and wind direction. Similarly, Europe oriented toward SQF research, but is also closely associated with cluster 2 (green line), indicating a likelihood of finding AQF research in this region as well.

On the other hand, there are no cluster associated with any specific weather/soil station and North American and African continents (cluster 4: cyan line), despite there are work with air/soil data applications. The proximity of cluster 4 to cluster 3 (blue line) suggests a potential gap in AQF and SQF research in these continents. In this context, there is potential research focuses on recording and processing air/soil and pollutants data in specific locations within North America and Africa, even with the application of forecast models to examine the correlation with monitoring stations (see Appendix). In contrast, Asian researchers may utilize the robust information obtained from weather stations 1 and 4 to apply novel prediction methods. Finally, works in South America and Europe are updated with AQF and SQF topics, but it is expected that there will be a growing number of research endeavors in these areas. The identification of these research trends and gaps can serve as a valuable guide for future studies and contribute to a more comprehensive understanding of weather, soil, and air quality dynamics worldwide.

doi: 10.25100/iyc.v27i2.14876

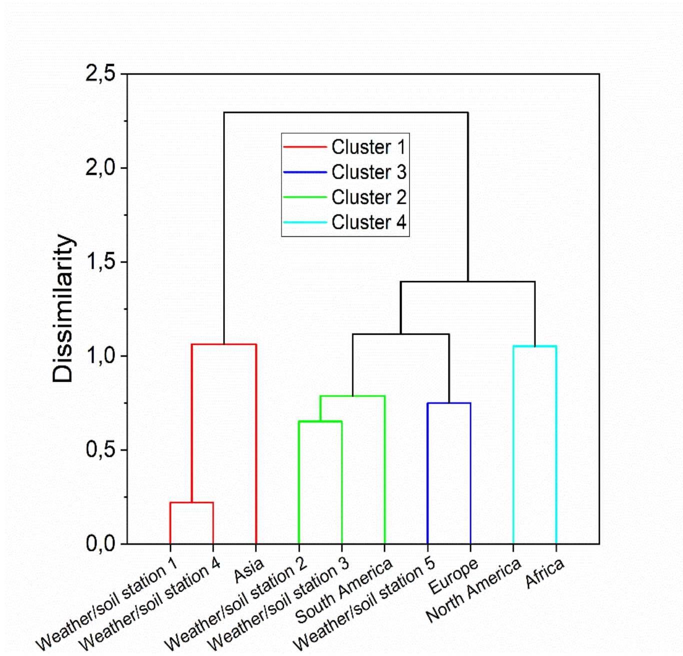


Figure 7. (Dendrogram) Grouping of weather/soil stations and publishing continent

Main configurations and trends of the hybrid forecasting models

Throughout the 21st century, researchers have evaluated different methods to predict the behavior of air/soil conditions and pollutant dynamics. These methods include mathematical, statistical and AI models. However, it is increasingly common to adopt a hybrid approach combining two or more methods to organize, filter and process data, improving the overall performance process or prediction accuracy. Generally, mathematical and statistical models are often employed as preprocessing tools, facilitating the removal of irrelevant data and grouping significant data.

To identify the prevailing combinations of these hybrid forecasting models, simultaneously used in different forecast or record works, a PCA was conducted on a diverse range of models, including those works with only calibration processes reported (Figure 8), this analysis resulted in the identification of 11 principal components, collectively referred to as hybrid forecasting models (HFM). Thus, HFM are models that apply two or more mathematical, statistical and forecast techniques to improve the processing data performance.

HFM1 (with the highest eigenvalue) primarily includes ML forecasting models, specifically RL algorithms like LSTM. The LSTM and other RL models are complemented by the application of the KCV statistical method, which is often used to assess the skill of an ML model by evaluating its performance on data not used during training. Furthermore, LSTM is especially effective when accurate forecasts are required between two events that are long-time apart, such as health damages or risk associated with increased $PM_{2.5}$ concentration. The LSTM algorithm can predict the trend of these types of pollutants to prevent heath risky (63,65). Therefore, HFM1 is predominantly used in long-term AQF applications with successful accuracy. Similarly, HFM11 shares comparable characteristics, as it also uses LSTM algorithms but includes the estimation of time-series based indicators (ARIMA), this additional information improves the pollutant levels prediction (65).

doi: 10.25100/iyc.v27i2.14876

Spatial values estimated by different methods have shown significant results in research

Specifically, HFM2 (Figure 8) groups a Spatial Inference method based on Geostatistical Interpolation using Gaussian Regression (kriging) and the CNN method, which is generally the first layer of the neural network used to extract spatial features. These methods find applications in agriculture (45) or AQF assessment (63). Additionally, MBE is included as a measure of the model's skill. Moreover, the "calibration method" or works without forecasting method are included as potential topics to be assessed by forecasting spatial methods. Therefore, HFM2 is mainly used to predict air/soil conditions separately between different spatial points.

HFM3 includes statistical methods characterized by ignoring the target function that connects events or abstract concepts. FIS is a method that utilizes fuzzy logic to map given input to output data, as demonstrated by Pruthi and Liu (53), who combine FIS with a neural network to achieve high-resolution predictions of $PM_{2.5}$ concentrations for the next three days based on $PM_{2.5}$ time-series data. Similarly, BO serves as mathematical support to validate the suitability of predictive analysis conducted with DL methods in weather stations (5). While the works reviewed did not explicitly combine SI and FDA mathematical methods with FIS and BO, there is potential for their joint application in estimating spatial values.

doi: 10.25100/iyc.v27i2.14876

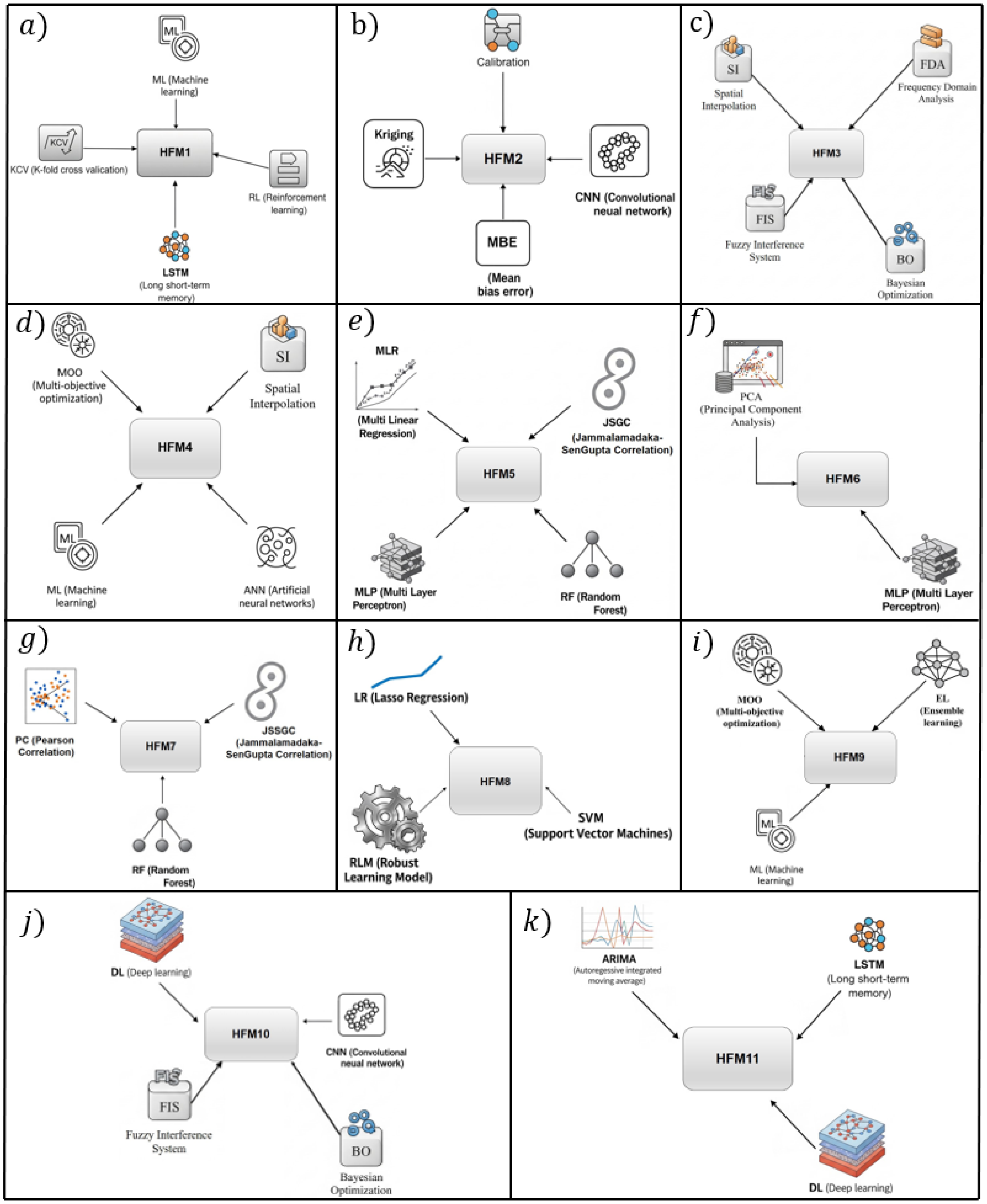


Figure 8. Configurations to Hybrid Forecasting Models a) HFM1, b) HFM2, c) HFM3, d) HFM4, e) HFM5, f) HFM6, g) HFM7, h) HFM8, i) HFM9, j) HFM10, k) HFM11

On the other hand, HFM10 employs a combination of FIS and BO with the CNN method, generating a better prediction accuracy for spatial values. The CNN method is frequently employed in spatial applications to facilitate the convolution of multilayer information, contributing to better

doi: 10.25100/iyc.v27i2.14876

forecasts. Despite, HFM11 having a lower eigenvalue (1.08) compared to HFM3 (2.46), it is expected that HFM11 will be applied more frequently in both AQF and SQF applications due to the greater computational processing capacity of the CNN method (5,53).

NO_2 gas is a toxic pollutant and a precursor of nitrates and $PM_{2.5}$. Its presence in the atmosphere is correlated with harmful effects on lungs and respiratory tracts. Thus, different research have focused on predicting NO_2 and $PM_{2.5}$ concentrations to avoid pollution levels that pose risk to both human health and the environment (3,50,52). In these studies, the use of HFM4 and HFM9 serves a similar purpose. For example, Zeng et.al (3) and Agarwal et.al (50) applied a model based on ANN trained with weather conditions, $PM_{2.5}$, PM_{10}, O_3 and NO_2 records to generate real-time AQF. On the other hand, Espinosa et.al (52) used MOO and EL methods to predict NO_2 concentrations over one week at different monitoring points. The methodology describes the optimization by minimizing the root mean squared error of multiple linear regression models obtained from the training method (EL) incorporating weather and geographical data. Furthermore, ANN was utilized to calibrate low-cost AQmesh (air quality sensors) and temperature sensors by filtering the negative effects of multiple environmental factors on measurements, thus improving the estimation accuracy of NO_2 and temperature values (49,60). In this context, both HFM4 and HFM9 present promising opportunities for both forecasting pollutant levels, mainly time-spatial NO_2 concentrations, and calibrating sensor devices.

HFM6 is a special case, as there is only one research study that uses PCA and MLP models jointly to predict $PM_{2.5}$ and PM_{10} concentrations (23). In this study, measurements of particle matter and different weather conditions are subject to PCA analysis to maximize their correlations and generate new variables that are mutually orthogonal. These new variables serve as input variables for the MLP model. Data are then processed through non-linear transformations across different layers (perceptron), identifying patterns with the smallest error, such as mean squared error. As a result, HFM6 demonstrates promising potential for utilization in future similar research studies and could be explored in SQF applications, particularly in the realm of intelligent agriculture.

HFM5, HFM7 and HFM8 are mainly employed for calibrating weather/soil stations using multiple decision tree analysis (RF) or maximizing the width of the gap between different categories (SVM) formed by air/soil measures (30,49,59). The validation of results is obtained through the application of different statistical methods. Specifically, HFM5 is utilized to calculate correlations on angular data, such as wind direction, using the JSGC method; therefore, HFM5 holds potential for calibrating the measurement of angular variables obtained by sensors in weather/soil stations. Conversely, HFM7 is employed for calibrating linear variables sensors using the PC method. On the other hand, HFM8 uses LR method on each category formed by SVM application to achieve the best accuracy. Unlike HFM5, HFM8 does not use correlation calculations but serves as an alternative way to explore calibration applications, even in AQF and SQF applications.

doi: 10.25100/iyc.v27i2.14876

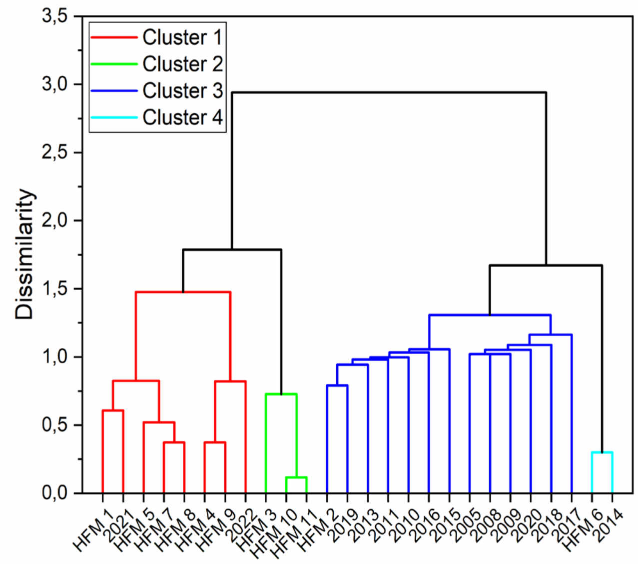


Figure 9. (Dendrogram) Grouping of hybrids forecasting models and publishing year

Hierarchical ward clustering was calculated for HFMs and publishing years (Figure 9), grouping them into four clusters. Researches on air quality long-time predictions (HFM1) were carried out in 2021, and there is a significant positive correlation between HFM1 and the year 2021 (see Appendix). HFM4 and HFM9 are more closely associated with the year 2022 than other methods, indicating that current interest of researchers in forecasting weather stations to predict NO_2 concentrations using ANN and EL methods (cluster 1: red line). Nevertheless, only HFM9 exhibits a significant correlation with 2022, whereas HFM4 has a low positive correlation with 2020; thus, the use of the EL method is more common in recent works. It is expected that HFMs with AI methods such as ANN, RF, SVM and EL will continue to receive attention from researchers due to the increasing application of these types of forecasting models in the calibration of weather/soil stations and the prediction of pollutant levels (HFM5, HFM7 and HFM8). HFM3, HFM10 and HFM11 are close to cluster 1, indicating the cutting-edge nature of spatial forecast applications.

doi: 10.25100/iyc.v27i2.14876

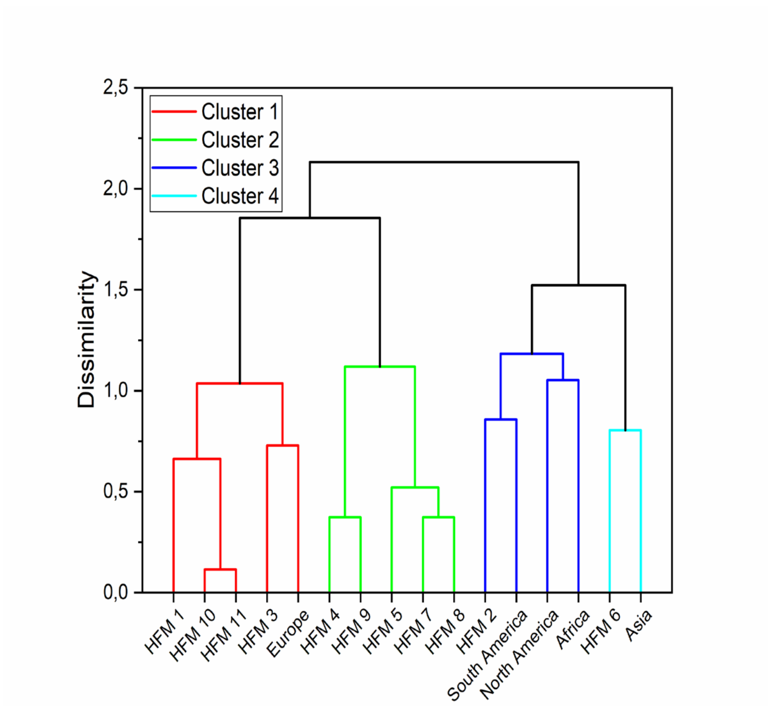


Figure 10. (Dendrogram) Grouping of hybrid forecasting models and publishing continent

On the other hand, researches focused on recording air/soil stations and designing accurate weather/soil stations, especially for agriculture applications (HFM2), have been conducted over the last two decades (cluster 3: blue line). However, despite the existence of such works until 2019, their trend is decreasing in comparison to the increasing use of AI applications. Cluster 4 (cyan line) includes only the application of PCA and MLP methods in 2014; this result highlights a gap in research where these models could be proposed to evaluate AQF, SQF and sensor calibration in current studies.

Hierarchical clustering was performed considering publishing continent and HFMs resulting in four clusters (Figure 10). Cluster 1 and 2 (red and green line respectively) group all HMFs used to forecast and evaluate air quality, including the calibration of weather/soil stations through AI, in Europe. Therefore, most novel researches are concentrated in Europe, highlighting the significant funding in research aimed at proposing and applying environmental public policies that allow for the reduction of greenhouse gas emissions' negative impact. Furthermore, there is a strong trend to use HFM11 in this continent, with a significant correlation of 0.38 (see Appendix).

On the other hand, researches focused on the design of weather or soil stations for agriculture applications is more common in South America, North America and Africa (cluster 3: blue line). In Asia, HFM6 is the principal hybrid forecast method; thus, another potential research focus is disclosed regarding the forecasting of pollutants, nutrients, or physical variables in Asian soils.

Main links between weather/soil stations and hybrid forecasting models

Finally, hierarchical ward clustering analysis was calculated, considering both weather/soil stations and HFMs (Figure 11). Weather/soil stations 1 and 4 are general-purpose monitoring stations equipped with air and soil sensors to capture climatic and soil conditions including temperature, humidity, wind speed and direction, NKP, pH, and others. On the other hand, weather/soil

doi: 10.25100/iyc.v27i2.14876

stations specifically are used as sensing electrochemical devices, specifically tailored to obtain measurements of the aforementioned variables. In fact, these monitoring stations are clustered in group 1, exhibiting the greatest hierarchical distance compared to other clusters (Figure 11). Hence, there is a lack of information about the use of forecasting methods to calibrate or predict the spatial-temporal behavior of the variables obtained by these monitoring stations. However, it is crucial to recognize that measures obtained by weather/soil stations 1 and 4 are not solely employed as input variables for training algorithms aimed at predicting other variables, but also, they are not typically the target for prediction or calibration in research works. Moreover, there is a significant negative correlation between weather/soil station 4 and the application of RF for calibrating circular variable sensors (see Appendix), indicating a trend of avoiding their joint application. In brief, while weather/soil station variables might not have a clear interest among researchers, their utilization remains pivotal for advancing AQF and SQF works.

Air pollutants like CO, CO_2 and particle matter ($PM_{2.5} - PM_{10}$) can be effectively recorded by low-cost monitoring stations, such as weather/soil stations 2 and 3. This advantage lies in the ability to obtain a useful and reliable database to train DL models like CNN and MLP, enabling accurate time and spatial predictions of pollutant levels. Indeed, HFM2 and HFM6 are clustered with weather/soil stations 2 and 3. It is noticeable that these research works predominantly originate from South America and Asia, suggesting an opportunity to explore these issues in other continents. Furthermore, methods for long term and spatial prediction are closely associated with cluster 2 (green line), indicating that LSTM, CNN, KCV, FIS, BO and ARIMA methods are usually applied to forecast particle matter and carbon gas concentrations in Europe. The prevalence of these

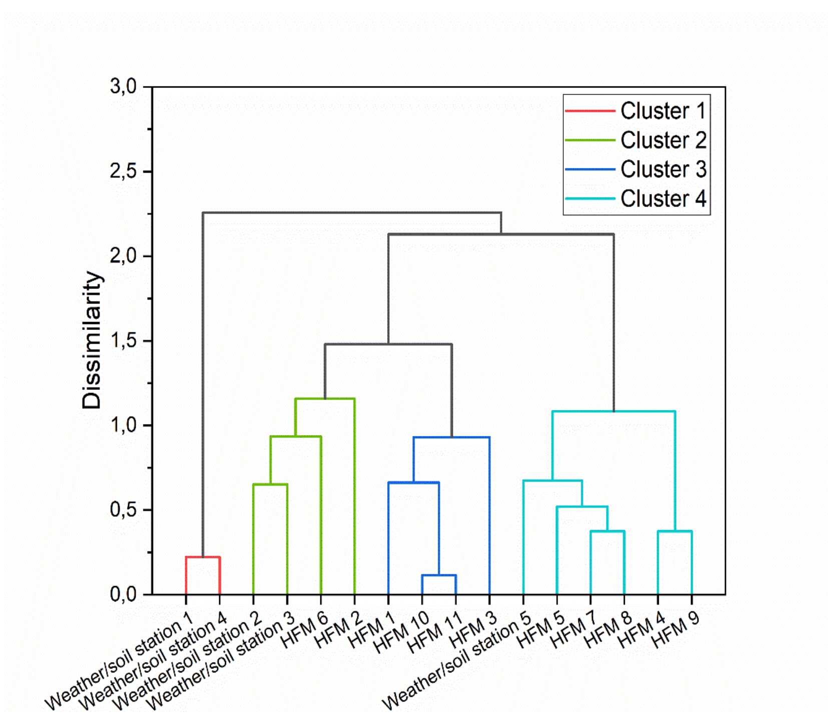


Figure 11. (Dendrogram) Grouping of hybrid forecasting models and weather/soil stations

doi: 10.25100/iyc.v27i2.14876

methods in European research highlights their potential to address air quality forecasting challenges in this region.

Weather/soil station 5 demonstrates a strong relation with soil pollution records as it includes measurement of NO_x, , O_3 and COV gases, which are known to have negative effects on plants. Thus, the prediction of these variables has current significant research interest. In fact, ANN and EL forecast models have been successfully applied to predict NO_2 concentrations. Besides, the sensors could be calibrated by RF and SVM models to achieve the best accuracy and improve prediction quality. Then, weather/soil station 5 serves a dual purpose in predicting soil variables and calibrating soil sensors through the applications of HFM4, HFM5, HFM7, HFM8 and HFM9. Notably, weather/soil station 5 has a significant correlation with HFM4 and HFM5 (see Appendix); indicating a clear trend in calibrating and forecasting NO_2 and circular variables time signals such as wind direction, through the use of ANN, RF and JSGC methods.

On the whole, it is recommended that future research focused on AQF evaluation in continents other than Europe, using weather/soil stations 2 and 3 along with LSTM, CNN and MLP forecast models. Nevertheless, there is an opportunity to apply these methods to calibrate sensors measuring CO, CO_2 and particle matter. On the other hand, for SQF evaluation, researchers should use weather/soil station 5 and utilize ANN, RF, SVM and EL forecast methods to calibrate soil sensors and predict concentrations of gases like NO_x, O_3 and COV. Weather/soil stations 1 and 4 could be effectively utilized to record essential variables, which can serve as predictors or inputs in forecasting models. It should be noted that any combinations not currently represented by the clusters in Figure 11 remain unexplored research areas with the potential to be considered in future works. These unexplored combinations offer promising avenues for advancing the field of environmental engineering and enhancing air and soil quality forecasting applications.

Conclusions

Currently, the use of AI technologies covers most knowledge topics, including the forecast of weather and soil conditions, especially the reliable prediction of air and soil pollutant levels. Forecast models based on these cutting-edge AI technologies have demonstrated the ability to achieve accurate predictions using weather/soil measurements as both input variables and target variables. Furthermore, research studies focus the design of measurement devices to send, receive, store and process data using embedded software and long-range communication protocols like Wi-Fi.

There are three main technological configurations used along of the researches. The first, is a holistic monitoring station which allows the recording of both weather conditions and pollutants in the air, as well as nutrients, physical-chemical conditions and pollutants in the soil. This configuration is not supported by AI applications, since its aim is the assuring of high-quality data recording and storing.

The second technological configuration, in addition to greenhouse gases monitoring station, includes AI methods to assess AQF, with a specific focus on long-short time records and forecast of carbon gases and particle matter, applying AI forecast methods such as LSTM, CNN and MLP which

doi: 10.25100/iyc.v27i2.14876

are supported by statistical and mathematical models like KCV, kriging, PCA, FIS, BO, ARIMA, SI and FDA.

The third configuration groups SQF and calibration devices, with particular attention to recording spatial and temporal signal data of soil pollutants, correlated with harmful effects in plants attributed to the presence of NO_x, O_3 and COV gases. Additionally, the spatial and temporal forecasting of these pollutants is evaluated using ANN, RF, SVM methods, which are complemented by MLR, JSGC, PC and LR statistical methods, and MOO mathematical method. Furthermore, this type of technology includes calibration based on AI methods as a preprocessing of recording data in to improve prediction accuracy.

These findings provide crucial insights into unexplored technological configurations, which hold substantial potential for future research endeavors. Additionally, they present a comprehensive overview of the current trends in AQF and SQF research works. Leveraging the power of advanced AI technologies, environmental engineering can achieve enhanced precision in weather and soil condition forecasting, along with effective mitigation of pollutants, ultimately contributing to a sustainable and healthier environment for all.

Appendix

https://hdl.handle.net/20.500.14112/30954

Statements and Declarations

The authors declare that they have no conflict or competing interests.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-profit sectors.

Acknowledgments

The authors gratefully acknowledge the support provide by the Environmental Engineering program at Universidad Mariana for their generous funding of this project, as approved by agreement 252 of December 7, 2021. Additionally, we extend our appreciation to the Program for Young Researches and Innovators, as well as the internal announcement of Universidad Mariana for the year 2022, for their valuable contributions to this research.

Author contributions (CRediT)

Conceptualization, Investigation: Giovanni Alexander Cuaran Paez. Conceptualization, Methodology, Formal analysis, Investigation, Writing – Original Draft: Luis Gabriel Lafaurie Ponce. Conceptualization, Writing – Reviewing and Editing, Visualization: Jenny Lucia Huertas Delgado. Writing – Reviewing and Editing: Daniel Alejandro Molina Cuaichar.

Financing: does not declare. Ethical aspect: does not declare.

References

1. Balogun AL, Tella A, Baloo L, Adebisi N. A review of the inter-correlation of climate change, air pollution and urban sustainability using novel machine learning algorithms and spatial information science. Urban Climate. 2021 Dec;40:100989. https://doi.org/10.1016/j.uclim.2021.100989

2. De Vries W, Dobbertin MH, Solberg S, Van Dobben HF, Schaub M. Impacts of acid deposition, ozone exposure and weather conditions on forest ecosystems in Europe: an overview. Plant Soil. 2014 Jul;380(1-2):1-45. https://doi.org/10.1007/s11104-014-2056-2

3. Zeng Z, Gui K, Wang Z, Luo M, Geng H, Ge E, et al. Estimating hourly surface PM2.5 concentrations across China from high-density meteorological observations by machine learning. Atmospheric Research. 2021 Jun;254:105516. https://doi.org/10.1016/j.atmosres.2021.105516

4. Espinosa R, Palma J, Jiménez F, Kamińska J, Sciavicco G, Lucena-Sánchez E. A time series forecasting based multi-criteria methodology for air quality prediction. Applied Soft Computing. 2021 Dec;113:107850. https://doi.org/10.1016/j.asoc.2021.107850

5. Akinosho TD, Oyedele LO, Bilal M, Barrera-Animas AY, Gbadamosi AQ, Olawale OA. A scalable deep learning system for monitoring and forecasting pollutant concentration levels on UK highways. Ecological Informatics. 2022 Jul;69:101609. https://doi.org/10.1016/j.ecoinf.2022.101609

6. Wang J, Li H, Yang H, Wang Y. Intelligent multivariable air-quality forecasting system based on feature selection and modified evolving interval type-2 quantum fuzzy neural network. Environmental Pollution. 2021 Apr;274:116429. https://doi.org/10.1016/j.envpol.2021.116429

7. Organization WH. Ambient air pollution: a global assessment of exposure and burden of disease. World Health Organization; 2016. https://www.cleanairjournal.org.za/article/view/7001

8. Koushal S, Arya D, Anbarasan S, Haloi D, M R, Rahman T, et al. Soil Pollution: Sources, Effects, and Mitigation Strategies. AJSSPN. 2025 Feb 15;11(1):280-90. https://doi.org/10.9734/ajsspn/2025/v11i1480

9. Kumar A, Attri AK. Correlating respiratory disease incidences with corresponding trends in ambient particulate matter and relative humidity. Atmospheric Pollution Research. 2016 Sep;7(5):858-64. https://doi.org/10.1016/j.apr.2016.05.005

10. Combes A, Franchineau G. Fine particle environmental pollution and cardiovascular diseases. Metabolism. 2019 Nov;100:153944. https://doi.org/10.1016/j.metabol.2019.07.008

11. Keikhosravi G, Fadavi SF. Impact of the inversion and air pollution on the number of patients with Covid-19 in the metropolitan city of Tehran. Urban Climate. 2021 May;37:100867. https://doi.org/10.1016/j.uclim.2021.100867

12. Organization WH. WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide [Internet]. World Health Organization; 2021. 290 p. Available from: https://www.who.int/publications/i/item/9789240034228

13. Renard JB, Surcin J, Annesi-Maesano I, Delaunay G, Poincelet E, Dixsaut G. Relation between PM2.5 pollution and Covid-19 mortality in Western Europe for the 2020-2022 period. Science of The Total Environment. 2022 Nov;848:157579. https://doi.org/10.1016/j.scitotenv.2022.157579

14. Różański SŁ, Castejón JMP, McGahan DG. Child risk assessment of selected metal(loid)s from urban soils using in vitro UBM procedure. Ecological Indicators. 2021 Aug;127:107726. https://doi.org/10.1016/j.ecolind.2021.107726

15. Khan ZI, Ahmad K, Yasmeen S, Akram NA, Ashraf M, Mehmood N. Potential health risk assessment of potato (Solanum tuberosum L.) grown on metal contaminated soils in the central zone of Punjab, Pakistan. Chemosphere. 2017 Jan;166:157-62. https://doi.org/10.1016/j.chemosphere.2016.09.064

16. da Silva Júnior AH, Mulinari J, de Oliveira PV, de Oliveira CRS, Reichert Júnior FW. Impacts of metallic nanoparticles application on the agricultural soils microbiota. Journal of Hazardous Materials Advances. 2022 Aug;7:100103. https://doi.org/10.1016/j.hazadv.2022.100103

doi: 10.25100/iyc.v27i2.14876

17. Masood A, Ahmad K. A review on emerging artificial intelligence (AI) techniques for air pollution forecasting: Fundamentals, application and performance. Journal of Cleaner Production. 2021 Nov;322:129072. https://doi.org/10.1016/j.jclepro.2021.129072

18. Liu H, Yan G, Duan Z, Chen C. Intelligent modeling strategies for forecasting air quality time series: A review. Applied Soft Computing. 2021 Apr;102:106957. https://doi.org/10.1016/j.asoc.2020.106957

19. Liu H, Yin S, Chen C, Duan Z. Data multi-scale decomposition strategies for air pollution forecasting: A comprehensive review. Journal of Cleaner Production. 2020 Dec;277:124023. https://doi.org/10.1016/j.jclepro.2020.124023

20. Zhou C, Li S, Wang S. Examining the Impacts of Urban Form on Air Pollution in Developing Countries: A Case Study of China's Megacities. IJERPH. 2018 Jul 24;15(8):1565. https://doi.org/10.3390/ijerph15081565

21. Thorat T, Patle BK, Kashyap SK. Intelligent insecticide and fertilizer recommendation system based on TPF-CNN for smart farming. Smart Agricultural Technology. 2023 Feb;3:100114. https://doi.org/10.1016/j.atech.2022.100114

22. Özdemir U, Taner S. Impacts of Meteorological Factors on PM 10 : Artificial Neural Networks (ANN) and Multiple Linear Regression (MLR) Approaches. Environmental Forensics. 2014 Oct 2;15(4):329-36. https://doi.org/10.1080/15275922.2014.950774

23. He HD, Lu WZ, Xue Y. Prediction of particulate matters at urban intersection by using multilayer perceptron model based on principal components. Stoch Environ Res Risk Assess. 2015 Dec;29(8):2107-14. https://doi.org/10.1007/s00477-014-0989-x

24. Tian J, Chen D. A semi-empirical model for predicting hourly ground-level fine particulate matter (PM2.5) concentration in southern Ontario from satellite remote sensing and ground-based meteorological measurements. Remote Sensing of Environment. 2010 Feb;114(2):221-9. https://doi.org/10.1016/j.rse.2009.09.011

25. Awewomom J, Dzeble F, Takyi YD, Ashie WB, Ettey ENYO, Afua PE, et al. Addressing global environmental pollution using environmental control techniques: a focus on environmental policy and preventive environmental management. Discov Environ. 2024 Feb 6;2(1):8. https://doi.org/10.1007/s44274-024-00048-y

26. Kim Y, Evans RG, Iversen WM. Remote sensing and control of an irrigation system using a distributed wireless sensor network. IEEE Transactions on Instrumentation and Measurement. 2008;57(7):1379-87. https://doi.org/10.1109/TIM.2008.917198

27. Ben Ishak A, Moslah Z, Trabelsi A. Analysis and prediction of PM10 concentration levels in Tunisia using statistical learning approaches. Environmental and Ecological Statistics. 2016 Sep 1;23(3):469-90. https://doi.org/10.1007/s10651-016-0349-8

28. Boznar M, Lesjak M, Mlakar P. A neural network-based method for short-term predictions of ambient SO2 concentrations in highly polluted industrial areas of complex terrain. Atmospheric Environment Part B Urban Atmosphere. 1993 Jun;27(2):221-30. https://doi.org/10.1016/0957-1272(93)90007-S

29. Chaloulakou A, Grivas G, Spyrellis N. Neural Network and Multiple Regression Models for PM 10 Prediction in Athens: A Comparative Assessment. Journal of the Air & Waste Management Association. 2003 Oct;53(10):1183-90. https://doi.org/10.1080/10473289.2003.10466276

30. Xiong J, Yao R, Wang W, Yu W, Li B. A spatial-and-temporal-based method for rapid particle concentration estimations in an urban environment. Journal of Cleaner Production. 2020 May;256:120331. https://doi.org/10.1016/j.jclepro.2020.120331

doi: 10.25100/iyc.v27i2.14876

31. Baklanov A, Zhang Y. Advances in air quality modeling and forecasting. Global Transitions. 2020;2:261-70. https://doi.org/10.1016/j.glt.2020.11.001

32. Wu H, Levinson D. The ensemble approach to forecasting: A review and synthesis. Transportation Research Part C: Emerging Technologies. 2021;132:103357. https://doi.org/10.1016/j.trc.2021.103357

33. Khaydukova M, Kirsanov D, Sarkar S, Mukherjee S, Ashina J, Bhattacharyya N, et al. One shot evaluation of NPK in soils by “electronic tongue.” Computers and Electronics in Agriculture. 2021 Jul;186:106208. https://doi.org/10.1016/j.compag.2021.106208

34. Bai L, Wang J, Ma X, Lu H. Air Pollution Forecasts: An Overview. Int J Environ Res Public Health. 2018;44. https://doi.org/10.3390/ijerph15040780

35. Slater LJ, Arnal L, Boucher MA, Chang AYY, Moulds S, Murphy C, et al. Hybrid forecasting: blending climate predictions with AI models. Hydrol Earth Syst Sci. 2023 May 15;27(9):1865-89. https://doi.org/10.5194/hess-27-1865-2023

36. Chacón RL, García HM. Sistema de bajo coste para la medida y monitorización en agricultura inteligente. SAAEI 2018: 25 Seminario Anual de Automática, Electrónica Industrial e Instrumentación 2018: Barcelona, España: 4-6 Julio, 2018: proceedings book. 2018;250-6. https://acortar.link/t74tVZ

37. Fisher DK, Kebede H. A low-cost microcontroller-based system to monitor crop temperature and water status. Computers and Electronics in Agriculture. 2010 Oct;74(1):168-73. https://doi.org/10.1016/j.compag.2010.07.006

38. Piamonte M, Huerta M, Clotet R, Padilla J, Vargas T, Rivas D. WSN prototype for African oil palm bud rot monitoring. Advances in Intelligent Systems and Computing. 2018;687:170-81. https://doi.org/10.1007/978-3-319-70187-5_13

39. Abad J, Farez J, Chasi P, Guillermo JC, García-Cedeño A, Clotet R, et al. Coffee Crops Variables Monitoring: A Case of Study in Ecuadorian Andes. Advances in Intelligent Systems and Computing. 2019;893:202-17. https://doi.org/10.1007/978-3-030-04447-3_14

40. SV G M, Galande SG. Measurement of NPK, Temperature, Moisture, Humidity using WSN [Internet]. 2015 p. 84-9. Available from: https://www.ijera.com/papers/Vol5_issue8/Part%20-%203/M58038489.pdf

41. Sahitya G, Balaji N, Naidu CD, Abinaya S. Designing a wireless sensor network for precision agriculture using zigbee. Proceedings - 7th IEEE International Advanced Computing Conference, IACC 2017. 2017 Jul;287-91. https://doi.org/10.1109/IACC.2017.0069

42. Kapse S, Kale S, Bhongade S, Sangamnerkar S, Gotmare Y. IOT Enable Soil Testing & NPK Nutrient Detection. Jac: a Journal of Composition Theory. 2020;XIII:310-8.

43. Mesas-Carrascosa FJ, Santano DV, Meroño JE, Orden MS de la, García-Ferrer A. Open source hardware to monitor environmental parameters in precision agriculture. Biosystems Engineering. 2015 Sep;137:73-83. https://doi.org/10.1016/j.biosystemseng.2015.07.005

44. Fahmi N, Huda S, Prayitno E, Rasyid MUHA, Roziqin MC, Pamenang MU. A prototype of monitoring precision agriculture system based on WSN. 2017 International Seminar on Intelligent Technology and Its Application: Strengthening the Link Between University Research and Industry to Support ASEAN Energy Sector, ISITIA 2017 - Proceeding. 2017 Nov;2017-January:323-8. https://doi.org/10.1109/ISITIA.2017.8124103

45. Juan MNV, Faruk FR, Quezada YML. Design and implementation of WSN for precision agriculture in white cabbage crops. Proceedings of the 2017 IEEE 24th International Congress on Electronics, Electrical Engineering and Computing, INTERCON 2017. 2017 Oct; https://ieeexplore.ieee.org/abstract/document/8079671

doi: 10.25100/iyc.v27i2.14876

46. Karimi N, Arabhosseini A, Karimi M, Kianmehr MH. Web-based monitoring system using Wireless Sensor Networks for traditional vineyards and grape drying buildings. Computers and Electronics in Agriculture. 2018 Jan;144:269-83. https://doi.org/10.1016/j.compag.2017.12.018

47. Dorji U, Pobkrut T, Kerdcharoen T. Electronic nose based wireless sensor network for soil monitoring in precision farming system. 2017 9th International Conference on Knowledge and Smart Technology: Crunching Information of Everything, KST 2017. 2017 Mar;182-6. https://doi.org/10.1109/KST.2017.7886087

48. Martínez VDV, García FGF, Cervantes GG, Medina M de JF, Casillas HAM. Desarrollo y validación de una estación meteorológica automatizada de bajo costo dirigida a agricultura. Revista mexicana de ciencias agrícolas. 2015;6(6):1253-64. https://www.scielo.org.mx/scielo.php?pid=S2007-09342015000600009&script=sci_arttext

49. Cordero JM, Borge R, Narros A. Using statistical methods to carry out in field calibrations of low cost air quality sensors. Sensors and Actuators B: Chemical. 2018 Aug;267:245-54. https://doi.org/10.1016/j.snb.2018.04.021

50. Agarwal S, Sharma S, R. S, Rahman MH, Vranckx S, Maiheu B, et al. Air quality forecasting using artificial neural networks with real time dynamic error correction in highly polluted regions. Science of The Total Environment. 2020 Sep;735:139454. https://doi.org/10.1016/j.scitotenv.2020.139454

51. Fernandez L, Huerta M, Sagbay G, Clotet R, Soto A. Sensing climatic variables in a orchid greenhouse. 2017 International Caribbean Conference on Devices, Circuits and Systems, ICCDCS 2017. 2017 Jun;101-4. https://doi.org/10.1109/ICCDCS.2017.7959719

52. Espinosa R, Jiménez F, Palma J. Multi-objective evolutionary spatio-temporal forecasting of air pollution. Future Generation Computer Systems. 2022 Nov;136:15-33. https://doi.org/10.1016/j.future.2022.05.020

53. Pruthi D, Liu Y. Low-cost nature-inspired deep learning system for PM2.5 forecast over Delhi, India. Environment International. 2022 Aug;166:107373. https://doi.org/10.1016/j.envint.2022.107373

54. Rodríguez Á, Figueredo J. Selección e implementación de un prototipo de estación meteorológica. 2016;341-52. https://acortar.link/MawXhB

55. Strigaro D, Cannata M, Antonovic M. Boosting a Weather Monitoring System in Low Income Economies Using Open and Non-Conventional Systems: Data Quality Analysis. Sensors 2019, Vol 19, Page 1185. 2019 Mar;19(5):1185. https://doi.org/10.3390/s19051185

56. Saini H, Thakur A, Ahuja S, Sabharwal N, Kumar N. Arduino based automatic wireless weather station with remote graphical application and alerts. 3rd International Conference on Signal Processing and Integrated Networks, SPIN 2016. 2016 Sep;605-9. https://doi.org/10.1109/SPIN.2016.7566768

57. Botero-Valencia JS, Mejia-Herrera M, Pearce JM. Low cost climate station for smart agriculture applications with photovoltaic energy and wireless communication. HardwareX. 2022 Apr;11:e00296. https://doi.org/10.1016/j.ohx.2022.e00296

58. Devaraju JT, Suhas KR, Mohana HK, Patil VA. Wireless Portable Microcontroller based Weather Monitoring Station. Measurement. 2015 Dec;76:189-200. https://doi.org/10.1016/j.measurement.2015.08.027

59. Bernardes GFLR, Ishibashi R, Ivo AAS, Rosset V, Kimura BYL. Prototyping low-cost automatic weather stations for natural disaster monitoring. Digital Communications and Networks. 2022 May; https://doi.org/10.1016/j.dcan.2022.05.002

60. Yamamoto K, Togami T, Yamaguchi N, Ninomiya S. Machine Learning-Based Calibration of Low-Cost Air Temperature Sensors Using Environmental Data. Sensors 2017, Vol 17, Page 1290. 2017 Jun;17(6):1290. https://doi.org/10.3390/s17061290

61. Perez P, Reyes J. Prediction of Particulate Air Pollution using Neural Techniques. :7. https://acortar.link/1eaA01

62. Zimmerman N, Presto AA, Kumar SPN, Gu J, Hauryliuk A, Robinson ES, et al. A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring. Atmospheric Measurement Techniques. 2018;11(1):291-313. https://doi.org/10.5194/amt-11-291-2018

63. Lin CY, Chang YS, Abimannan S. Ensemble multifeatured deep learning models for air quality forecasting. Atmospheric Pollution Research. 2021 May;12(5):101045. https://doi.org/10.1016/j.apr.2021.03.008

64. Hofman J, Nikolaou M, Shantharam SP, Stroobants C, Weijs S, Manna VPL. Distant calibration of low-cost PM and NO2 sensors; evidence from multiple sensor testbeds. Atmospheric Pollution Research. 2022 Jan;13(1):101246. https://doi.org/10.1016/j.apr.2021.101246

65. Gladkova E, Saychenko L. Applying machine learning techniques in air quality prediction. Transportation Research Procedia. 2022;63:1999-2006. https://doi.org/10.1016/j.trpro.2022.06.222

66. Smith JR. Weather Station and Data Logger. Programming the PIC Microcontroller with MBASIC. 2005 Jan;691-728. https://doi.org/10.1016/B978-075067946-6/50030-4

67. Benghanem M. Measurement of meteorological data based on wireless data acquisition system monitoring. Applied Energy. 2009 Dec;86(12):2651-60. https://doi.org/10.1016/j.apenergy.2009.03.026

68. Cao-Hoang T, Duy CN. Environment monitoring system for agricultural application based on wireless sensor network. 7th International Conference on Information Science and Technology, ICIST 2017 - Proceedings. 2017 May;99-102. https://doi.org/10.1109/ICIST.2017.7926499

69. Ma RH, Wang YH, Lee CY. Wireless Remote Weather Monitoring System Based on MEMS Technologies. Sensors (Basel, Switzerland). 2011 Mar;11(3):2715. https://doi.org/10.3390/s110302715

70. Jeong JH, Choi J, Jeong JY, Woo SH, Kim SW, Lee D, et al. A novel statistical-dynamical method for a seasonal forecast of particular matter in South Korea. Science of The Total Environment. 2022 Nov;848:157699. https://doi.org/10.1016/j.scitotenv.2022.157699

71. Mat I, Kassim MRM, Harun AN, Yusoff IM. IoT in Precision Agriculture applications using Wireless Moisture Sensor Network. ICOS 2016 - 2016 IEEE Conference on Open Systems. 2017 Mar;24-9. https://doi.org/10.1109/ICOS.2016.7881983

72. Pratama H, Yunan A, Candra RA. Design and Build a Soil Nutrient Measurement Tool for Citrus Plants Using NPK Soil Sensors Based on the Internet of Things. Brilliance: Research of Artificial Intelligence. 2021 Dec;1(2):67-74. https://doi.org/10.47709/brilliance.v1i2.1300

73. Suresh DS, V JKP, J RC. Automated Soil Testing Device. ITSI Transac-tions on Electrical and Electronics Engineering (ITSI-TEEE) ISSN (PRINT) 2013;(1):2320-8945.

74. Madhumathi R, Arumuganathan T, Shruthi R. Soil NPK and Moisture analysis using Wireless Sensor Networks. 2020 11th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2020. 2020 Jul; https://doi.org/10.1109/ICCCNT49239.2020.9225547

75. Shylaja SN, Veena MB. Real-time monitoring of soil nutrient analysis using WSN. 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing, ICECDS 2017. 2018 Jun;3059-62. https://doi.org/10.1109/ICECDS.2017.8390018

76. Flores-Medina M, Flores-García F, Velasco-Martínez V, González-Cervantes G, Jurado-Zamarripa F. Monitoreo de humedad en suelo a través de red inalámbrica de sensores. Tecnología y ciencias del agua. 2015 Oct;6(5):75-88. https://www.scielo.org.mx/scielo.php?pid=S2007-24222015000500006&script=sci_arttext

77. Ananthi N, Divya J, Divya M, Janani V. IoT based smart soil monitoring system for agricultural production. Proceedings - 2017 IEEE Technological Innovations in ICT for Agriculture and Rural Development, TIAR 2017. 2018 Jan;2018-January:209-14. https://doi.org/10.1109/TIAR.2017.8273717

78. Sindhu P, Indirani G. IoT Enabled Soil Testing. Asian Journal of Computer Science and Technology. 2018;7(S1):54-7. https://doi.org/10.51983/ajcst-2018.7.S1.1805

79. Ramadan KM, Oates MJ, Molina-Martinez JM, Ruiz-Canales A. Design and implementation of a low cost photovoltaic soil moisture monitoring station for irrigation scheduling with different frequency domain analysis probe structures. Computers and Electronics in Agriculture. 2018 May;148:148-59. https://doi.org/10.1016/j.compag.2017.12.038

80. Lawrence MG. The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air: A Simple Conversion and Applications. Bull Amer Meteor Soc. 2005 Feb;86(2):225-34. https://doi.org/10.1175/BAMS-86-2-225

81. Tehrani NA, Esfahani IC, Sun H. Simultaneous humidity and temperature measurement with micropillar enhanced QCM sensors. Sensors and Actuators A: Physical. 2024 Feb;366:115039. https://doi.org/10.1016/j.sna.2024.115039

82. Topp GC, Davis JL, Annan AP. Electromagnetic determination of soil water content: Measurements in coaxial transmission lines. Water Resour Res. 1980 Jun;16(3):574-82. https://doi.org/10.1029/WR016i003p00574

83. Mane S, Das N, Singh G, Cosh M, Dong Y. Advancements in dielectric soil moisture sensor Calibration: A comprehensive review of methods and techniques. Computers and Electronics in Agriculture. 2024;218:108686. https://doi.org/10.1016/j.compag.2024.108686

84. Amankwah SK, Ireson AM, Brannen R. An improved model and field calibration technique for measuring liquid water content in unfrozen and frozen soils with dielectric probes. Vadose Zone Journal. https://acsess.onlinelibrary.wiley.com/doi/pdf/10.1002/vzj2.20225

85. The World Air Quality Index project. aqicn.org. 2023 [cited 2023 Jun 27]. Air Quality Open Data Platform. Available from: https://aqicn.org/data-platform/tos/

86. Rosiek S, Batlles FJ. A microcontroller-based data-acquisition system for meteorological station monitoring. Energy Conversion and Management. 2008 Dec;49(12):3746-54. https://doi.org/10.1016/j.enconman.2008.05.029

87. Babalola TE, Babalola AD, Olokun MS. Development of an ESP-32 Microcontroller Based Weather Reporting Device. JERR. 2022 Jul 12;27-38. https://doi.org/10.9734/jerr/2022/v22i1117577

88. Sziroczak D, Rohacs D, Rohacs J. Review of using small UAV based meteorological measurements for road weather management. Progress in Aerospace Sciences. 2022 Oct;134:100859. https://doi.org/10.1016/j.paerosci.2022.100859

89. Outay F, Galland S, Gaud N, Abbas-Turki A. Simulation of connected driving in hazardous weather conditions: General and extensible multiagent architecture and models. Engineering Applications of Artificial Intelligence. 2021 Sep;104:104412. https://doi.org/10.1016/j.engappai.2021.104412

doi: 10.25100/iyc.v27i2.14876

90. Han E, Ines AVM, Baethgen WE. Climate-Agriculture-Modeling and Decision Tool (CAMDT): A software framework for climate risk management in agriculture. Environmental Modelling & Software. 2017 Sep;95:102-14. https://doi.org/10.1016/j.envsoft.2017.06.024

91. Danbatta SJ, Varol A. Comparison of Zigbee, Z-Wave, Wi-Fi, and Bluetooth Wireless Technologies Used in Home Automation. In: 2019 7th International Symposium on Digital Forensics and Security (ISDFS) [Internet]. Barcelos, Portugal: IEEE https://doi.org/10.1109/ISDFS.2019.8757472

92. Kumar, Bhuvana C, Anushya S. Comparison of ZigBee and Bluetooth wireless technologies-survey. In: 2017 International Conference on Information Communication and Embedded Systems (ICICES) [Internet]. Chennai, India: IEEE; 2017 [cited 2023 May 10]. p. 1-4. Available from: https://doi.org/10.1109/ICICES.2017.8070716

93. Ye S, Xue P, Fang W, Dai Q, Peng J, Sun Y, et al. Quantitative effects of PM concentrations on spectral distribution of global normal irradiance. Solar Energy. 2021 May;220:1099-108. https://doi.org/10.1016/j.solener.2020.08.070

94. Zhang C, Stevenson D. Characteristic changes of ozone and its precursors in London during COVID-19 lockdown and the ozone surge reason analysis. Atmospheric Environment. 2022 Mar;273:118980. https://doi.org/10.1016/j.atmosenv.2022.118980

95. Vazquez Santiago J, Inoue K, Tonokura K. Diagnosis of ozone formation sensitivity in the Mexico City Metropolitan Area using HCHO/NO2 column ratios from the ozone monitoring instrument. Environmental Advances. 2021 Dec;6:100138. https://doi.org/10.1016/j.envadv.2021.100138

96. Wongrin W, Chaisee K, Suphawan K. Comparison of Statistical and Deep LearningMethods for Forecasting PM2.5 Concentrationin Northern Thailand. Pol J Environ Stud. 2023 Feb 23;32(2):1419-31. https://doi.org/10.15244/pjoes/157072

doi: 10.25100/iyc.v27i2.14876