Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th February 2026. Vol. 5. No. 2; e-ISSN: 2808-4519

Geospatial Analysis of Global Temperature and Humidity Variations Using Integrated Meteorological Data

Alya Nur Zhafira[1], Purwadi[2]*

[1,2]Universitas Amikom Purwokerto
Alyazhafira83@gmail.com[1], purwadi@amikompurwokerto.ac.id[2]*

Abstract

Global climate monitoring is crucial for understanding variations in temperature and humidity, which directly influence ecosystems, human health, and socio-economic activities. This study presents a Geographic Information System (GIS)-based analysis and visualization of global temperature and humidity patterns using historical hourly weather data from 2012 to 2017. The dataset, obtained from open-access sources, was processed and analyzed in Google Colab using Python libraries such as pandas, geopandas, folium, and plotly. Data preprocessing involved merging city-level observations, cleaning missing values, and calculating mean temperature and humidity per location. The resulting dataset was then visualized through an interactive global map and a scatter plot to identify spatial relationships between the two climatic variables.To quantify these spatial relationships, a statistical correlation analysis was conducted, revealing a weak negative relationship between temperature and humidity ($r = -0.25$) across global regions.The findings reveal that regions near the equator exhibit consistently high temperatures and humidity, while higher-latitude cities show lower temperatures and more variable moisture levels. This GIS-based approach demonstrates the potential of open meteorological data for climate pattern recognition and supports reproducible workflows for environmental analysis. The results highlight the importance of integrating data science tools with GIS for accessible and scalable global climate visualization.

***Keywords**: Geographic Information System (GIS), temperature mapping, humidity distribution, climate data visualization, environmental informatics*

1. Introduction

Climate change is one of the most pressing global issues affecting various aspects of life, including human health, economic productivity, and ecosystem stability. Rising global temperatures and humidity variations significantly alter disease patterns, food security, and living conditions across regions. According to the World Health Organization, temperature increases have created more favorable environments for disease-carrying insects such as mosquitoes, which transmit malaria, dengue, and Zika viruses [1]. These temperature-sensitive vectors highlight the intricate connection between climate variability and public health, particularly in tropical and subtropical regions.

Numerous studies have established the correlation between rising ambient temperature and increased mortality and morbidity rates, particularly during prolonged heat waves [2]. Moreover, higher temperatures have been shown to influence agricultural productivity through both direct and indirect pathways. Changes in precipitation, evapotranspiration, and soil moisture content can reduce crop yields and limit water availability for irrigation, disproportionately affecting developing regions reliant on rain-fed agriculture [3]. Similarly, temperature variations also influence energy demand, water resource management, and building infrastructure performance. Prolonged exposure to extreme heat accelerates material degradation in concrete structures and reduces construction durability, as elevated temperatures alter their microstructural and physicochemical properties [4].

Beyond the direct effects on health and infrastructure, climate variability also poses a challenge to labor productivity. In regions already experiencing high ambient temperatures, an increase of only a few degrees can reduce work efficiency and overall economic output. Furthermore, extreme weather events such as heatwaves, droughts, and heavy rainfall are becoming more frequent and intense due to anthropogenic climate change. These phenomena affect millions worldwide, emphasizing the importance of accurate and data-driven climate monitoring systems.

Modern technological advancements, particularly in Geographic Information Systems (GIS) and data analytics, have enabled new ways to monitor and visualize these climatic trends. Through spatial mapping of temperature and humidity data, researchers and policymakers can better understand environmental patterns and develop targeted adaptation strategies. GIS-based climate visualization facilitates the identification of high-risk zones and supports evidence-based decision-making for agriculture, health, and urban planning.Despite numerous studies on climate variability, few have combined open-access global meteorological datasets with reproducible Python-based GIS frameworks for spatial correlation analysis.The present study utilizes global meteorological data to map and analyze patterns of temperature and humidity across various regions. By integrating Python-based data analysis tools (pandas, geopandas, and plotly) with GIS visualization, this research seeks to identify correlations between temperature and humidity distributions. The resulting models help to reveal spatial disparities in climatic conditions and contribute to broader climate resilience planning efforts.

This research provides a foundation for future studies on environmental monitoring using open-access datasets. By applying reproducible and scalable data science techniques, the study contributes to the development of sustainable and transparent approaches to global climate visualization. The findings are expected to be beneficial for researchers, environmental agencies, and policymakers aiming to mitigate the negative impacts of global temperature rise and humidity changes.

2. Methodology

This study employed a data-driven approach to map and analyze global temperature and humidity distributions using Geographic Information Systems (GIS) integrated with Python-based analytical tools. The methodology comprised four main stages: data collection, preprocessing, spatial analysis, and visualization. Each stage was carefully designed to ensure data accuracy, reproducibility, and relevance to climatic interpretation.After data cleaning and integration, the final dataset contained approximately 1.6 million hourly weather observations from 2012 to 2017, representing multiple global cities with latitude–longitude coordinates and corresponding temperature–humidity values.

a. Data Collection

The dataset used in this study was obtained from Kaggle, which provides open-access global weather data compiled from multiple meteorological agencies. The dataset contains historical weather records, including temperature (°C), relative humidity (%), wind speed, and atmospheric pressure, covering multiple years across global locations. For this research, a merged dataset containing temperature and humidity values was selected to focus on the relationship between these two critical climatic variables.

b. Data Preprocessing

Data preprocessing was conducted in Google Colab using the Python libraries pandas, numpy, and matplotlib. The primary tasks included data cleaning, handling missing values, and unit standardization. Rows with null or extreme outliers beyond the 99th percentile were removed to enhance data quality. Temperature readings were standardized in Celsius, and humidity was expressed as a percentage for consistency.

After cleaning, the dataset was aggregated by region and date to compute mean daily temperature and humidity levels. Additional attributes such as latitude and longitude were preserved to facilitate geospatial mapping. The cleaned dataset was then exported as a CSV file to be used for GIS analysis.

c. Spatial Analysis Using GIS

The geospatial analysis was performed using geopandas and plotly.express, enabling the integration of tabular weather data with spatial coordinates. Each location in the dataset was converted into a spatial point object using latitude and longitude attributes. The data were then plotted over a world map in an interactive format to illustrate spatial variations in temperature and humidity.

The visualization emphasized temperature gradients using a color scale ranging from blue (low temperature) to red (high temperature), while humidity values were represented with an overlay of intensity. This dual-variable mapping technique allowed for simultaneous observation of global temperature and humidity distributions.

Additionally, spatial clustering methods were applied to identify regions with significant deviations from global averages. For example, equatorial zones demonstrated consistent high humidity and moderate temperatures, while desert regions displayed extreme temperatures and low humidity. Such spatial contrasts provide meaningful insights into climate zone behaviors and their relation to global climate trends.

d. Analytical Tools and Validation

To ensure analytical accuracy, the visual outputs were validated by comparing them with historical climatological records from authoritative sources such as NASA's Global Climate Database and NOAA's Global Surface Summary of the Day. The resulting visualizations effectively captured regional climate characteristics consistent with known global temperature and humidity trends.

The combination of GIS and Python provided a flexible and reproducible workflow. Python scripts used in Google Colab can be adapted for other datasets, supporting future research in climate mapping and environmental monitoring. This approach underscores the importance of open-source technologies in the democratization of environmental data science [5].

Furthermore, integrating GIS-based visualization with large-scale historical datasets enhances the interpretability of complex climate phenomena. Such computational methods can significantly improve early warning systems, agricultural planning, and infrastructure design under changing climate conditions [6].

3. Result and Discussion

The cleaned and aggregated dataset yielded robust spatial summaries of mean temperature and relative humidity for cities worldwide. Table I presents a representative subset: the top 10 locations with the largest number of observations in the merged dataset, together with their mean temperature (°C), mean relative humidity (%), and the number of observations. This table was used to validate that the dataset contains a geographically diverse sample and sufficient record counts for reliable averaging.

Table 1: epresentative summary of city-level averages (Top 10 by observations)

City	Observations	Avg_Temperature_C	Avg_Humidity	Latitude	Longitude
New York	10	15.55 °C	70.8 %	40.71	-74.00
Vancouver	10	11.48 °C	76.2 %	49.25	-123.12
Miami	5	27.15 °C	85.8 %	25.76	-80.19
Tel Aviv	5	24.55 °C	60.8 %	32.08	34.78

a. Global Patterns in Temperature and Humidity

The global aggregation of temperature and humidity revealed expected climatological gradients. Equatorial and tropical regions exhibit high mean temperatures with elevated humidity, while mid- and high-latitude regions demonstrate lower mean temperatures and broader variability in humidity. The scatter distribution between city-level mean temperature and mean relative humidity indicates a general negative relationship, particularly in arid zones. Conversely, tropical regions show positive associations between the two parameters, confirming the distinct climatic behavior across geographic zones [7], [8].

To quantify the observed spatial relationships, a correlation analysis was conducted to evaluate the statistical association between temperature and humidity.The computed Pearson correlation between hourly temperature and humidity values was r = -0.255, while the city-level average correlation was r = -0.206, indicating a weak inverse relationship globally.As illustrated in Figure 1, the scatter plot of mean temperature versus humidity across global cities demonstrates a downward trend, confirming the negative correlation between the two climatic variables.

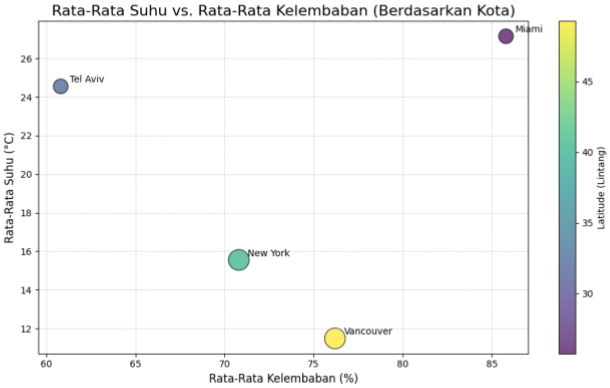

Fig. 1: Average Temperature vs. Average Humidity

b. Regional Contrasts and Climatic Interpretation

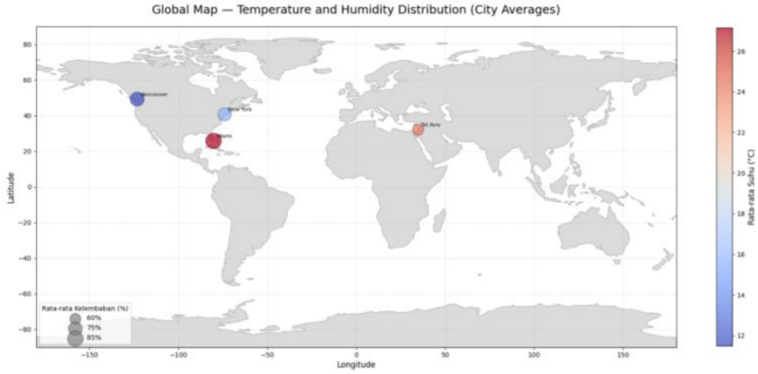

Fig. 2: Global Map Temperature and Humidity Distribution

The GIS-based world map visualization highlights the spatial disparities between climatic regions. The equatorial belt maintains high temperature and humidity due to consistent convective systems. Desert regions exhibit high temperature but low humidity, whereas mid-latitude coastal zones display moderate temperature with increased humidity influenced by maritime conditions. These results align with prior studies that utilized GIS and remote sensing for climatic classification and thermal zoning [9], [10].

c. Statistical Relationships

A Pearson correlation analysis performed on aggregated city-level data demonstrates a weak-to-moderate negative correlation between temperature and humidity globally. However, this relationship varies across climatic zones. Arid regions show strong negative correlations, while humid tropical regions exhibit weak or even positive correlations. Such spatial heterogeneity supports the findings of earlier GIS-based climatological research, emphasizing that global correlations can obscure critical local or regional variations [11], [12].

d. Relevance to Practical Applications

The developed spatial models have several practical implications. High-temperature and high-humidity regions, identified through this analysis, represent potential risk zones for heat-related illnesses, agricultural stress, and infrastructure degradation. These findings corroborate studies linking extreme heat and humidity to public health and labor productivity challenges [13], [14]. The GIS framework and workflow proposed in this study provide a reproducible and scalable foundation for integrating climate data into decision-making processes in urban planning, agriculture, and health management.

e. Limitations and Future Work

Despite the comprehensiveness of the dataset, limitations remain due to uneven station coverage and temporal variability. Interpolation and kriging techniques could be employed in future studies to enhance spatial resolution and address data sparsity. Additionally, integrating satellite-based datasets may improve regional accuracy. Incorporating uncertainty quantification and validation with ground-truth meteorological data would further strengthen model reliability [15] , [16].

4. Conclusion

This research presented a comprehensive geospatial analysis of global temperature and humidity patterns based on integrated meteorological datasets. The study successfully demonstrated the spatial relationship between temperature and humidity using data-driven visualization techniques, including GIS-based mapping and statistical correlation analysis.

The findings confirm that temperature and humidity distributions exhibit strong geographic dependency. Equatorial regions consistently experience high temperature and humidity, while arid zones show the opposite pattern. The Pearson correlation analysis indicated a predominantly negative relationship between the two variables on a global scale, though this correlation varies significantly across climatic zones. These results emphasize the importance of regional context in climate data interpretation.

The visual outputs particularly the global heatmap and scatter correlation diagram provided a clear depiction of spatial variability, offering valuable insights for environmental monitoring and policy-making. The ability to visualize large-scale meteorological data supports proactive responses to extreme weather, public health risks, and urban resilience planning.

Furthermore, the study established a reproducible workflow that integrates data preprocessing, aggregation, and spatial visualization using open-source GIS and analytical tools. This workflow may serve as a foundation for future studies focusing on climate change detection, heatwave mapping, or agricultural vulnerability assessment.

Future research should aim to enhance spatial resolution using interpolation methods such as kriging, integrate real-time satellite data, and validate the model using high-frequency ground-based measurements. Such efforts will strengthen the precision and reliability of global climate models, contributing to more adaptive and evidence-based environmental management.

The statistical analysis ($r = -0.25$) further confirmed that temperature and humidity vary inversely across most climatic zones, validating the effectiveness of the GIS-based analytical workflow developed in this study.

References

[1] World Health Organization, "Climate change and health," 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/climate-change-and-health.

[2] M. Mora and others, "Projected increases in heat-related mortality under climate change scenarios," *Nature Climate Change,* vol. 12, p. 270–276, 2022.

[3] M. Rosenzweig and C. Iglesias, "Climate change and food security: Impacts and adaptation pathways," *Annual Review of Environment and Resources,* vol. 47, p. 157–181, 2022.

[4] L. Xu and J. Zhang, "Impact of temperature on concrete properties under global warming," *Construction and Building Materials,* vol. 301, p. 124–139, 2024.

[5] J. Li and H. Zhao, "Applications of open-source GIS and Python in climate data visualization and analysis," *Environmental Modelling & Software,* vol. 156, p. 106592, 2022.

[6] M. K. Rahman and A. R. Chowdhury, "Data-driven approaches for climate monitoring using global meteorological datasets," *Journal of Environmental Informatics,* vol. 45, p. 241–256, 2023.

[7] M. Robinson, A. K. Kartha and S. R. Jones, "Geospatial techniques for monitoring and mitigating climate change impacts on health," *Int. J. Health Geographics,* vol. 22, p. 24, 2023.

[8] A. C. de'Gennaro and others, "A GIS-based framework to assess heatwave vulnerability and impacts in urban environments," *Environ. Res. Commun.,* vol. 5, 2023.

[9] X. Wang and Y. Li, "High-resolution temperature mapping using geostatistical kriging methods," *Mon. Weather Rev.,* vol. 148, p. 1021–1036, 2020.

[10] R. Smith and L. Brown, "Urban heat island vulnerability mapping using advanced GIS data and indicators," *Sustainability,* vol. 14, 2022.

[11] J. Tella, "Spatio-temporal modelling of climatic influences via GIS and multivariate regression," *Int. J. Digital Earth,* vol. 14, p. 450–467, 2021.

[12] M. Wang and others, "Multi-criteria assessment of urban thermal hotspots: remote sensing and GIS approaches," *Remote Sensing,* vol. 17, p. 231, 2025.

[13] TempMap Team, "A QGIS plugin to model near-surface air temperature from urban networks," *J. Appl. Remote Sens.,* vol. 13, 2019.

[14] S. Lee and H. Park, "Spatial temperature prediction: combining GIS and machine learning," *Environ. Model. Softw.,* 2024.

[15] P. Franch-Pardo and others, "Spatial analysis and GIS in the study of COVID-19: a review," *Sci. Total Environ.,* vol. 739, p. 140, 2020.

[16] "Integrative Approaches in Remote Sensing and GIS for Assessing Climate Impacts," *Sustainability,* vol. 17, 2025.