Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

Environmental Analysis of Hemorrhagic Fever in Iraq Using Machine Learning

Ammar Abdul-Sattar Al-Ramahi [a, *], ***Ali Hasan Taresh*** [b]

[a] *Iraqi commission for computers and Informatics, Informatics Institute for Postgraduate Studies, Iraq.Email: ms202130658@iips.icci.edu.iq*

[b] *Iraqi commission for computers and Informatics, Informatics Institute for Postgraduate Studies, Iraq.Email: alihtaresh@uoitc.edu.iq*

ARTICLE INFO

Article history:
Received: 25 /05/2023
Rrevised form: 14 /07/2023
Accepted : 18 /07/2023
Available online: 30 /09/2023

Keywords:

Iraq hemorrhagic fever

Satellite imagery

Data Analysis

Risk Assessment

Spatio-temporal Patterns

ABSTRACT

Crimean-Congo Hemorrhagic Fever (CCHF) is a viral disease with rising prevalence in Iraq. This research focused on investigating factors influencing CCHF spread in Dhi Qar Governorate, which has seen a substantial surge in cases. The problem was to uncover which environmental and climatic factors correlate with CCHF outbreaks. Meteorological data, rodent populations, vegetation indices, and epidemiological records were analyzed using data preprocessing techniques like interpolation and ARIMA modeling. Pearson correlation analysis was applied to quantify associations between CCHF cases and factors like temperature, humidity, rodent prevalence, and vegetation lushness. Results showed strong positive correlations of CCHF with rodent populations, temperature, solar radiation, and evapotranspiration. Negative correlations were found with humidity and vegetation health. The conclusions indicate environmental factors significantly influence CCHF outbreaks in Dhi Qar. This can inform prevention strategies targeting ecological and climatic drivers of the disease.

MSC..

https://doi.org/10.29304/jqcm.2023.15.3.1264

1. Introduction

Hemorrhagic fevers are serious viral illnesses; Crimean-Congo Hemorrhagic Fever (CCHF) spreads via infected ticks or contact with contaminated animal substances [1][2][3]. This study explores the correlation between Hemorrhagic Fever and regional factors in Iraq to understand the disease's spread and prevention strategies[4][5]. The World Health Organization reported 1,085 suspected, 287 confirmed cases, and 52 deaths due to CCHF in Iraq up to August 2022 [6]. A rising CCHF incidence highlights the need for early alert systems and prevention tactics[7]. Our research focuses on the Dhi Qar Governorate, the most impacted by CCHF, investigating factors influencing its spread[7][8][9]. We study relationships between CCHF occurrences and factors such as climate, rodent populations, vegetation, and epidemiological data[10][11]. Our results indicate a positive correlation between CCHF cases and factors like rodent prevalence, temperature, solar radiation, and evapotranspiration and a negative correlation with humidity and vegetation health [12][7][13]. Future research

---

*Corresponding author

Email addresses: ms202130658@iips.icci.edu.iq

Communicated by ‘sub etitor’

should consider social, economic, and demographic variables to understand CCHF prevalence (refer to Figure 1 for a visual summary).

Several studies have been conducted on dengue prediction and surveillance using data-driven methods. Aimrun Wayayok et al. [14], 2023, developed six different Long Short-Term Memory (LSTM) models for dengue prediction in Malaysia, achieving effective results. Emmanuelle Sylvestre et al.[2], 2022, conducted a systematic review and found that combining real-world data and Big Data with machine learning methods improves dengue management. Kirstin Roster et al. [15], 2022, compared machine learning algorithms for predicting monthly dengue cases in Brazilian cities, with random forests performing best. Jinwei Dong [16], 2022, emphasized the importance of big geospatial data, cloud computing, and deep learning models in dengue risk prediction. Samrat Kumar Dey et al. [17], 2022, developed machine learning models to predict dengue cases in various districts of Bangladesh, achieving significant accuracy.

Limited literature exists on CCHF dynamics in Iraq. This study analyzes high-resolution data in Dhi Qar to identify factors contributing to CCHF outbreaks. It enhances understanding of the increasing cases in southern Iraq and contributes to our knowledge of CCHF prevalence while showcasing data analytics methods for studying epidemiological associations.

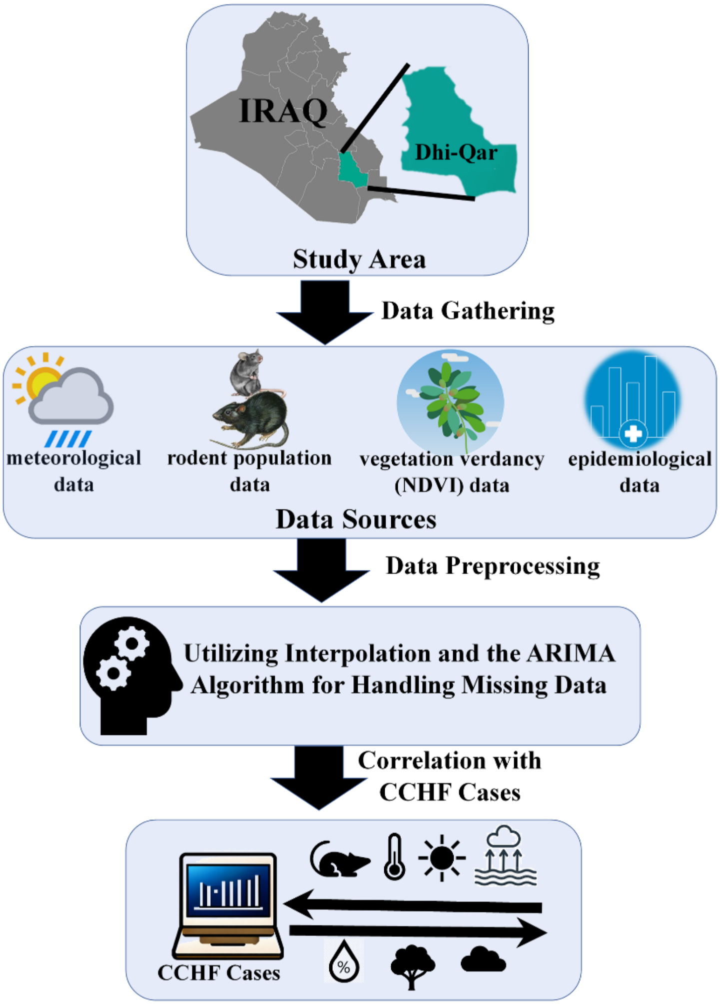


Figure 1 Environmental and Climatic Influences on Crimean-Congo Hemorrhagic Fever Prevalence in Dhi Qar

2. Materials and Methods

2.1. Study Area

The growing number of CCHF cases in Iraq underscores the significance of examining the factors influencing the disease. World Health Organization (WHO) records[15] show that a significant outbreak occurred in Iraq. Moreover, data reveals an unprecedented surge in Crimean-Congo Hemorrhagic Fever cases in Dhi Qar Governorate[7][8], making it the most impacted region regarding CCHF infections. Consequently, this research chose Dhi Qar Governorate as the study area to investigate the factors affecting CCHF. Situated in southern Iraq[18][19]. Dhi Qar is distinguished by its desert climate [20], characterized by scarce rainfall and high summer temperatures that can reach up to 50 degrees[20], along with a mild winter climate[21]. The governorate lies between 32-33.30 degrees north latitude and 45.37-47.12 degrees east longitude [30]. The area's location and climate are essential factors that may contribute to CCHF's spread and prevalence, and this research seeks to examine these factors and their influence on the disease. Studying the factors affecting CCHF in Dhi Qar Governorate can aid in developing effective early warning, prevention, and control measures to minimize potential disease risk[22]

2.2. Data Gathering

Data was compiled from the Dhi Qar Governorate to address the research question. The investigation employed several data sources, comprising:

- **Meteorological data**

  Weather conditions are regarded as a primary influence on Hemorrhagic Fever. Thus, examining climatic factors and conducting thorough analyses are crucial to detecting outbreaks, identifying climatic factors that aid in the spread, and potentially reducing or eliminating the disease within the study area[1][23][24]. As a result, data from the Ministry of Agriculture's Iraqi Agrometeorological Center was collected, encompassing daily records spanning 14 years. The collected data includes weather parameters such as rain amount (Rain), maximum temperature (AT Max), minimum temperature (AT Min), average temperature (AT Avg), maximum relative humidity (RH Max), minimum relative humidity (RH Min), total solar radiation (SLR Total Mj/m2), average wind speed (WS Avg m/s), and evapotranspiration (ET).

- **The Rodents Abundance**

  Rodents serve as significant reservoirs for numerous zoonotic pathogens [25][26][27][28][29] The number of burrows determines the frequency (Runways count) per dunam and the number of affected dunams in each governorate. In Dhi Qar Governorate, this totaled 21,860 dunams. Annual statistics were obtained from the Ministry of Agriculture's Directorate of Plant Protection.

- **Vegetation Greenness Indices**

  Vegetation is also considered an essential environmental factor associated with the prevalence of Crimean-Congo Hemorrhagic Fever [30] [31] dataset incorporates NDVI (Normalized Difference Vegetation Index) values extracted from satellite images, which were processed using GEE-based geospatial extensive data analysis.

- **Epidemiological Data**

  This category includes information on Crimean-Congo Hemorrhagic Fever cases obtained from the Veterinary Directorate.

2.3. Pre-processing

re-processing Data preprocessing is essential for successful data-driven projects, as it helps address inconsistencies, noise, and redundancies in the raw data [32]Proper data preparation enables analysts and data scientists to gain more accurate and valuable insights from their models and algorithms. One critical step in data preprocessing is cleaning, which includes handling missing values. Various techniques, such as interpolation and the ARIMA algorithm, can be used to impute missing values [33][34][35]

In our study's data preprocessing phase, we followed the steps outlined below:

- **Climate data**

  To handle missing values within the climate dataset, the Autoregressive Integrated Moving Average (ARIMA) algorithm was employed [36][37][38][39]. The climate data collected from provincial stations in the study area spanned from 2008 to 2022. Although the research primarily focused on 2022, a 14-year dataset was provided by the Ministry of Agriculture's Iraqi Agrometeorological Center, As shown in Figure 2. For handling missing data in the 2022 timeframe, the ARIMA algorithm was applied to the time series from 2008 to the date of the missing data for the relevant climatic parameters. And it produced outstanding results with an R-squared value of 1.00 and a Mean Absolute Error (MAE) of 0.00. These ideal outcomes can be attributed to the algorithm's reliance on a comprehensive dataset spanning 14 years. This extensive data coverage allows the ARIMA model to accurately capture the underlying patterns and trends in the data, ensuring highly reliable predictions. Consequently, the ARIMA algorithm proves to be an effective technique for handling missing data in this context. The collected climate data exhibited rare instances of missing values.

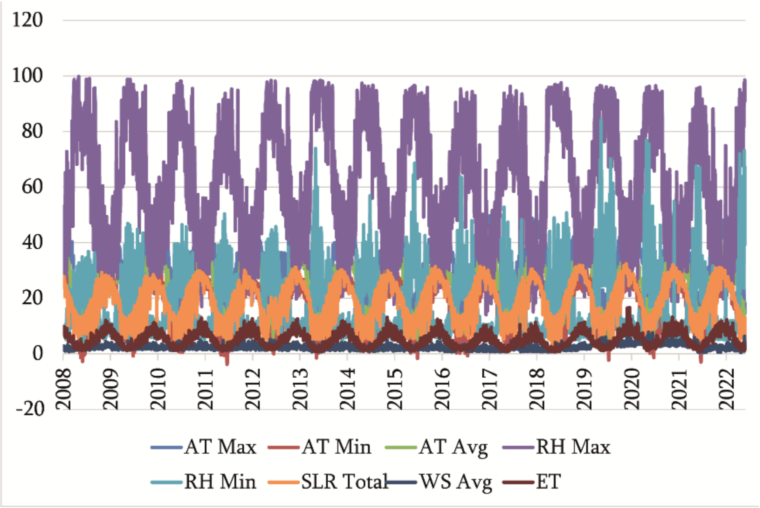


Figure 2 Climate Data features for 14 years

- **Rodent Data**

  To preprocess rodent data, annual statistics were initially converted into daily statistics by applying the annual figures to the abundance curve observed in previous rodent studies [40][41] As shown in Figure 3[40] the plan for distributing annual statistics and converting them into daily statistics represented the abundance of rodents. Origin Lab's digitizer tool was used to apply our study's dates and yearly statistics to a graphical representation of the daily abundance curve. This process facilitated the extraction of daily statistics for the number of burrows, representing daily rodent abundance [42] [43], enabling the identification of trends and patterns that were not evident in the original annual data. The interpolation algorithm was then applied to fill in days without points on the rodent abundance curve, converting the data from a sparsely spaced period to daily accuracy [44][45] have also achieved an R-squared value of 1.00, indicating a perfect fit, and a Mean Absolute Error (MAE) of 0.00. These results emphasize the suitability of the interpolation algorithm for this data type. Interpolation is a technique that estimates a function's value at a certain point based on known values at nearby points, which can help enhance the accuracy of machine learning models or enable data visualization and analysis [35]

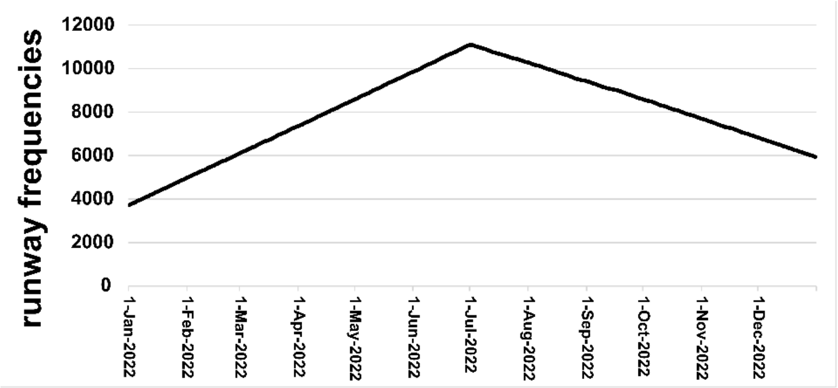
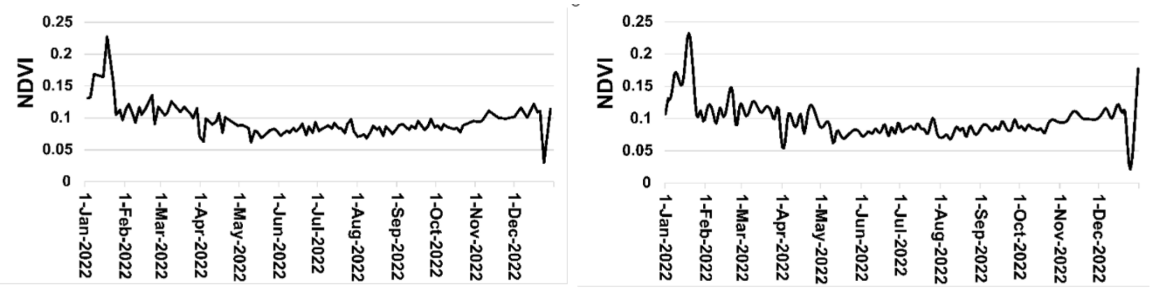


Figure 3 Indicators of Daily rodents Abundance [40]

- **Vegetation Greenness**

Obtaining daily NDVI values for a specific location is challenging due to satellite revisit time, cloud cover, atmospheric interference, and seasonal variability. The combined revisit time of Sentinel-2 satellites is around five days, and clouds and atmospheric conditions can impact image quality and accuracy. These factors and seasonal changes impede the consistent provision of precise daily NDVI values [46][47]. As a result, the interpolation feature in Origin Lab was utilized to fill in missing data and derive a more accurate time series. Applying interpolation to the vegetation greenness data produced a refined dataset for subsequent analysis. Refer to Figure 4.



(A) Before applying interpolation to the vegetation greenness data, The number of points in the curve is 119 for the year 2022

(B) After applying interpolation to the vegetation greenness data, Become the number of points on the curve is 365 for the year 2022

Figure 4 Indicator of Vegetation

This enables drawing more precise and reliable inferences from the data[48]. Refer to Table 1 and Table 2.

Table 1 sample dataset

Date	Rain	AT Max	AT Min	AT Avg	RH Max	RH Min	SLR Total Mj/m2	WS Avg m/s	ET	NDVI	rodents abundance	total cases
01/01/2022	0	16.84	4.17	10.505	92.2	47.22	9.4	0.59	1.2	0.106922	3723	0

02/01/2022	0.6	16.3	7.11	11.705	86.8	49.49	5.81	0.63	1.1	0.118961	3763.872618	0
03/01/2022	1	18.56	4.47	11.515	91.5	28.40	11.66	1.62	2.1	0.131	3804.745237	0
04/01/2022	0.1	15.28	3.48	9.38	95.6	41.42	13.45	1.2	1.5	0.12921	3845.617855	0
05/01/2022	0	17.78	0.89	9.335	94.6	28.23	13.07	0.93	1.6	0.13233	3886.490474	0
06/01/2022	0	18.73	3.66	11.195	86.2	27.01	13.51	1.79	2.3	0.14336	3927.363092	0
07/01/2022	0	20.2	4.74	12.47	86.3	32.01	13.44	1.24	2	0.15774	3968.235711	0
08/01/2022	0	21.01	3.12	12.065	90.4	28.76	13.95	0.9	1.8	0.169	4009.108329	0
09/01/2022	0	21.55	3.77	12.66	88.6	25.71	12.81	0.6	1.6	0.17227	4049.980947	0
10/01/2022	2.5	20.79	6.64	13.715	89.8	36.26	8.85	1.96	2.3	0.16907	4090.853566	0
11/01/2022	0.3	20.64	12.44	16.54	89.3	53.70	8.83	1.73	1.9	0.16248	4131.726184	0
12/01/2022	0	22.16	10.01	16.085	91	47.19	11.75	1.15	1.9	0.15561	4172.598803	0
13/01/2022	2.4	19.72	10.71	15.215	94.6	67.10	5.08	1.47	1.3	0.15155	4213.471421	0
14/01/2022	0	19.61	11.49	15.55	90	66.72	8.32	1.29	1.5	0.15339	4254.344039	0
15/01/2022	0.1	19.62	5.88	12.75	95.9	34.18	13.77	0.93	1.8	0.16425	4295.216658	0
16/01/2022	0	19.47	6.5	12.985	88.8	23.14	14.43	1.87	2.6	0.18519	4336.089276	0

Table 2 Statistical summary of the dataset

	Rain	AT Max	AT Min	AT Avg	RH Max	RH Min	SLR Total Mj/m2	WS Avg m/s	ET	NDVI	rodents abundance	total cases
count	365.00	365.00	365.00	365.00	365.00	365.00	365.00	365.00	365.00	365.00	365.00	365.00
mean	0.16	33.49	16.93	25.08	56.54	16.70	20.22	1.92	5.67	0.10	7972.42	0.44
std	0.88	10.27	8.38	9.24	22.34	13.71	6.89	0.77	2.78	0.03	1930.53	0.85
min	0.00	11.64	-3.02	5.57	19.78	5.00	1.67	0.44	1.00	0.02	3723.00	0.00
25%	0.00	23.87	10.01	16.54	36.79	7.04	14.43	1.42	3.00	0.08	6547.55	0.00
50%	0.00	34.79	17.00	25.76	54.35	10.92	21.04	1.82	5.60	0.09	8092.37	0.00
75%	0.00	43.06	24.59	34.15	77.87	22.76	26.00	2.36	8.00	0.11	9582.93	1.00
max	10.80	49.93	35.12	42.26	98.61	73.21	31.02	6.26	12.20	0.23	11083.58	5.00

2.4. Experimental Configuration and Software

The analysis and processing of data in this research were conducted using various software tools and programming languages. Python 3.10.9 was the principal programming language, facilitating efficient data handling and processing[49]. Furthermore, specialized applications for data analysis, like Origin Lab, were utilized for data visualization and interpretation[50]. Data mining instruments, comprising Weka, KNIME Analytics Platform, and Orange Data Mining, were also employed to investigate the data and discern patterns and associations[51]. The data analysis and processing were carried out on a personal computer equipped with an 11th Gen Intel(R) Core(TM) i7-11800H @ 2.30GHz processor, 32GB of installed memory (RAM), and a 64-bit Microsoft Windows 11 Pro operating system, providing adequate computational power for the complex analyses required in the study. The Google Earth Engine (GEE) code editor platform was employed to obtain the NDVI (Normalized Difference Vegetation Index) from satellite imagery, contributing a crucial element to the dataset[52]. Several Python libraries, such as Pandas, scikit-learn, seaborn, and Matplotlib, were used for data processing, visualization, and presenting the findings[53] As shown in Figure 5

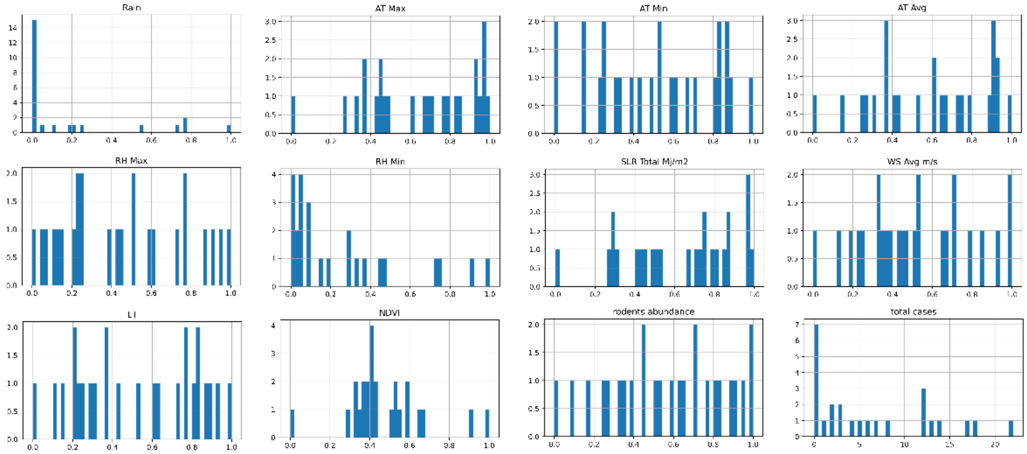
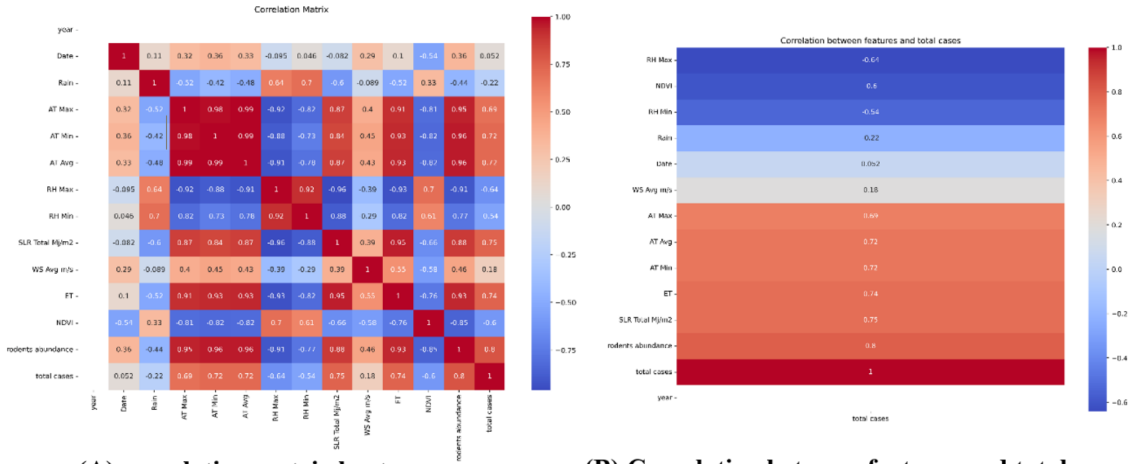


Figure 5 Histogram of numerical features

3. Results

The main goal of this research was to investigate the Pearson correlation coefficients between diverse factors, such as meteorological data, rodent abundance, vegetation greenness (NDVI), and the total instances of Crimean-Congo Hemorrhagic Fever (CCHF) in Iraq's Dhi Qar Governorate. These correlation coefficients shed light on the associations between various features and the emergence of CCHF in the area. The most substantial positive correlation was discovered between the target variable (total cases of CCHF) and rodent abundance, exhibiting a correlation coefficient of 0.7954. This finding implies that a rise in rodent populations is strongly connected to increased CCHF cases. Additional positive correlations were identified between the target variable and maximum temperature (AT Max) at 0.6901, average temperature (AT Avg) at 0.7194, minimum temperature (AT Min) at 0.7226, total solar radiation (SLR Total Mj/m^2) at 0.7463, and evapotranspiration (ET) at 0.7442. These positive correlations signify increased temperatures, solar radiation, and evapotranspiration are linked to more CCHF cases. In contrast, the most significant negative correlation was observed between the target variable and maximum relative humidity (RH Max) at -0.6398, followed by vegetation greenness (NDVI) at -0.6036 and minimum relative humidity (RH Min) at -0.5428. These negative correlations suggest an increase in relative humidity and vegetation greenness corresponds to a decline in total CCHF cases. Other negative correlations were detected between the target variable and rain amount (Rain) at -0.2201. This outcome indicates that more significant rainfall correlates with reduced CCHF cases. The average wind speed (WS Avg m/s) showed a weak positive correlation with the target variable at 0.1827, suggesting that higher wind speeds might be slightly correlated with increased CCHF cases. The date variable had a very weak positive correlation with the total number of CCHF cases at 0.0519, indicating no significant relationship between the date and the actual cases of the disease.

Refer to Figure 6 (A), which provides visual representations for a comprehensive overview of all correlations, and to Figure 6 (B) for their relationships with the total cases of CCHF.



Correlation Matrix

	year	Date	Rain	AT Max	AT Min	AT Avg	RH Max	RH Min	SLR Total Mj/m2	WS Avg m/s	ET	NDVI	rodents abundance	total cases
year														
Date		1	0.11	0.32	0.36	0.33	-0.095	0.046	-0.082	0.29	0.1	-0.54	0.36	0.052
Rain		0.11	1	-0.52	-0.42	-0.48	0.64	0.7	-0.6	-0.089	-0.52	0.33	-0.44	-0.22
AT Max		0.32	-0.52	1	0.98	0.99	-0.92	-0.82	0.87	0.4	0.91	-0.81	0.95	0.69
AT Min		0.36	-0.42	0.98	1	0.99	-0.88	-0.73	0.84	0.45	0.93	-0.82	0.96	0.72
AT Avg		0.33	-0.48	0.99	0.99	1	-0.91	-0.78	0.87	0.43	0.93	-0.82	0.96	0.72
RH Max		-0.095	0.64	-0.92	-0.88	-0.91	1	0.92	-0.96	-0.39	-0.93	0.7	-0.91	-0.64
RH Min		0.046	0.7	0.82	0.73	0.78	0.92	1	0.88	0.29	0.82	0.61	0.77	0.54
SLR Total Mj/m2		-0.082	-0.6	0.87	0.84	0.87	-0.96	-0.88	1	0.39	0.95	-0.66	0.88	0.75
WS Avg m/s		0.29	-0.089	0.4	0.45	0.43	-0.39	-0.29	0.39	1	0.55	-0.58	0.46	0.18
ET		0.1	-0.52	0.91	0.93	0.93	-0.93	-0.82	0.95	0.55	1	-0.76	0.93	0.74
NDVI		-0.54	0.33	-0.81	-0.82	-0.82	0.7	0.61	-0.66	-0.58	-0.76	1	-0.85	-0.6
rodents abundance		0.36	-0.44	0.95	0.96	0.96	-0.93	-0.77	0.88	0.46	0.93	-0.85	1	0.8
total cases		0.052	-0.22	0.69	0.72	0.72	-0.64	-0.54	0.75	0.18	0.74	-0.6	0.8	1

(A) correlation matrix heatmap

Correlation between features and total cases

	total cases
RH Max	-0.64
NDVI	0.6
RH Min	-0.54
Rain	0.22
Date	0.052
WS Avg m/s	0.18
AT Max	0.69
AT Avg	0.72
AT Min	0.72
ET	0.74
SLR Total Mj/m2	0.75
rodents abundance	0.8
total cases	1
year	

(B) Correlation between features and total cases

Figure 6 Visual representations of the Pearson correlation coefficient

4. Conclusion

This study enhances understanding of factors affecting Crimean-Congo Hemorrhagic Fever's (CCHF) prevalence in Dhi Qar, Iraq, identifying correlations between ecological and climatic factors and CCHF instances. This can guide targeted disease prevention measures. Challenges persist in data quality, accessibility, model comprehension, and data source integration. Also, ethical and privacy issues related to personal data usage need addressing. Future research should utilize innovative data collection, processing, and advanced machine learning techniques to understand disease dynamics better and predict accuracy. Ensuring ethical and privacy protection in data use is crucial for public trust and significant health outcome improvements.

5. Acknowledgments

We sincerely thank all contributors for their valuable input in improving disease surveillance and global public health. Key acknowledgments include:

- Dr. Mostafa Kamel of Al-Qadisiyah's veterinary hospital for his guidance.
- The Iraqi Agrometeorological Center (Ministry of Agriculture) for meteorological data.
- Dr. Enas Jawad from the Veterinary Department of the Ministry of Agriculture for providing us with the daily case count data for the study area in 2022.
- The Directorate of Plant Protection (Ministry of Agriculture) for data on rodent abundance in Dhi Qar;
- Dr. Amer Ibrahim of Al-Qadisiyah University / College of Veterinary Medicine for providing crucial references and studies;
- Dr. Omar Hussein, Dr. Ali Malik, and Ali Fayez for their consistent encouragement throughout the research..

References

[1] P. Guo *et al.*, “Developing a dengue forecast model using machine learning: A case study in China,” *PLoS Negl Trop Dis*, vol. 11, no. 10, p. e0005973, Oct. 2017, doi: 10.1371/journal.pntd.0005973.

[2] E. Sylvestre *et al.*, “Data-driven methods for dengue prediction and surveillance using real-world and Big Data: A systematic review,” *PLoS Neglected Tropical Diseases*, vol. 16, no. 1. Public Library of Science, p. e0010056, Jan. 01, 2022. doi: 10.1371/journal.pntd.0010056.

[3] Minnesota Department of Health, “Viral Hemorrhagic Fevers (VHFs),” *Minnesota Department of Health*. https://www.health.state.mn.us/diseases/vhf/vhf.html (accessed Mar. 06, 2023).

[4] K. Tallam and M. P. Quang, “Applications of artificial intelligence in predicting dengue outbreaks in the face of climate change: a case study along coastal India,” *medRxiv*, p. 2023.01.18.23284134, Jan. 2023, doi: 10.1101/2023.01.18.23284134.

[5] A. Fanelli and D. Buonavoglia, “Risk of Crimean Congo haemorrhagic fever virus (CCHFV) introduction and spread in CCHF-free countries in southern and Western Europe: A semi-quantitative risk assessment,” *One Health*, vol. 13, p. 100290, Dec. 2021, doi: 10.1016/j.onehlt.2021.100290.

[6] “WHO | Situation Report Iraq | Week 31 | United Nations in Iraq.” https://iraq.un.org/en/194395-who-situation-report-iraq-week-31 (accessed Mar. 09, 2023).

[7] “Crimean-Congo Hemorrhagic Fever - Iraq,” *World Health Organization (WHO)*. https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON386 (accessed Mar. 13, 2023).

[8] R. A. Alhilfi, H. A. Khaleel, B. M. Raheem, S. G. Mahdi, C. Tabche, and S. Rawaf, “Large outbreak of Crimean-Congo haemorrhagic fever in Iraq, 2022,” *IJID Regions*, vol. 6, pp. 76–79, Mar. 2023, doi: 10.1016/J.IJREGI.2023.01.007.

[9] “Situation Report Iraq Week 31 (ending 7 August 2022),” *World Health Organization*, Accessed: Mar. 11, 2023. [Online]. Available: https://iraq.un.org/sites/default/files/2022-08/WHO-Iraq-SitRep_Week-31.pdf

[10] V. Andreo, M. Belgiu, D. B. Hoyos, F. Osei, C. Provensal, and A. Stein, “Rodents and satellites: Predicting mice abundance and distribution with Sentinel-2 data,” *Ecol Inform*, vol. 51, pp. 157–167, May 2019, doi: 10.1016/j.ecoinf.2019.03.001.

[11] H. Jamil *et al.*, “Knowledge, attitudes, and practices regarding Crimean-Congo hemorrhagic fever among general people: A cross-sectional study in Pakistan,” *PLoS Negl Trop Dis*, vol. 16, no. 12, p. e0010988, Dec. 2022, doi: 10.1371/journal.pntd.0010988.

[12] F. Duygu, T. Sari, T. Kaya, O. Tavsan, and M. Naci, “The relationship between crimean-congo hemorrhagic fever and climate: Does climate affect the number of patients?,” *Acta Clin Croat*, vol. 57, no. 3, pp. 443–448, Sep. 2018, doi: 10.20471/acc.2018.57.03.06.

[13] M. Okely, R. Anan, S. Gad-Allah, and A. M. Samy, “Mapping the environmental suitability of etiological agent and tick vectors of Crimean-Congo hemorrhagic fever,” *Acta Trop*, vol. 203, p. 105319, Mar. 2020, doi: 10.1016/J.ACTATROPICA.2019.105319.

[14] M. A. ; Majeed *et al.*, “A Deep Learning Approach for Dengue Fever Prediction in Malaysia Using LSTM with Spatial Attention,” *International Journal of Environmental Research and Public Health 2023, Vol. 20, Page 4130*, vol. 20, no. 5, p. 4130, Feb. 2023, doi: 10.3390/IJERPH20054130.

[15] K. Roster, C. Connaughton, and F. A. Rodrigues, “Machine-Learning–Based Forecasting of Dengue Fever in Brazilian Cities Using Epidemiologic and Meteorological Variables,” *Am J Epidemiol*, vol. 191, no. 10, pp. 1803–1812, Sep. 2022, doi: 10.1093/AJE/KWAC090.

[16] Z. Li and J. Dong, “Big Geospatial Data and Data-Driven Methods for Urban Dengue Risk Forecasting: A Review,” *Remote Sensing 2022, Vol. 14, Page 5052*, vol. 14, no. 19, p. 5052, Oct. 2022, doi: 10.3390/RS14195052.

[17] S. K. Dey *et al.*, “Prediction of dengue incidents using hospitalized patients, metrological and socioeconomic data in Bangladesh: A machine learning approach,” *PLoS One*, vol. 17, no. 7 July, p. e0270933, Jul. 2022, doi: 10.1371/journal.pone.0270933.

[18] “Dhi Qar Map - State - Iraq - Mapcarta.” https://mapcarta.com/12539716 (accessed Mar. 10, 2023).

[19] “Al-Nāṣiriyyah | Iraq | Britannica.” https://www.britannica.com/place/Al-Nasiriyyah (accessed Mar. 10, 2023).

[20] “Climate Dhi Qar: Temperature, climate graph, Climate table for Dhi Qar - Climate-Data.org.” https://en.climate-data.org/asia/iraq/dhi-qar-2055/ (accessed Mar. 10, 2023).

[21] “**الموجز الاحصائي ذي قار** 2018 - **الجهاز المركزي للاحصاء**.” https://cosit.gov.iq/ar/1216-16-6-2021 (accessed Mar. 10, 2023).

[22] N. Agarwal, S. R. Koti, S. Saran, and A. Senthil Kumar, “Data mining techniques for predicting dengue outbreak in geospatial domain using weather parameters for New Delhi, India,” *Curr Sci*, vol. 114, no. 11, pp. 2281–2291, Jun. 2018, doi: 10.18520/cs/v114/i11/2281-2291.

[23] S. Chae, S. Kwon, and D. Lee, “Predicting infectious disease using deep learning and big data,” *Int J Environ Res Public Health*, vol. 15, no. 8, p. 1596, Jul. 2018, doi: 10.3390/ijerph15081596.

[24] Y. Ince *et al.*, “Crimean-Congo hemorrhagic fever infections reported by ProMED,” *International Journal of Infectious Diseases*, vol. 26, pp. 44–46, Sep. 2014, doi: 10.1016/j.ijid.2014.04.005.

[25] P. Weidinger *et al.*, “Potentially Zoonotic Viruses in Wild Rodents, United Arab Emirates, 2019—A Pilot Study,” *Viruses 2023, Vol. 15, Page 695*, vol. 15, no. 3, p. 695, Mar. 2023, doi: 10.3390/V15030695.

[26] M. Aminikhah *et al.*, “Rodent host population dynamics drive zoonotic Lyme Borreliosis and Orthohantavirus infections in humans in Northern Europe,” *Sci Rep*, vol. 11, no. 1, pp. 1–11, Aug. 2021, doi: 10.1038/s41598-021-95000-y.

[27] H. Dahmana, L. Granjon, C. Diagne, B. Davoust, F. Fenollar, and O. Mediannikov, “Rodents as hosts of pathogens and related zoonotic disease risk,” *Pathogens*, vol. 9, no. 3, p. 202, Mar. 2020, doi: 10.3390/pathogens9030202.

[28] X. Deng *et al.*, “Distinct Genotype of Hantavirus Infection in Rodents in Jiangxi Province, China, in 2020–2021,” *Zoonoses*, vol. 2, no. 1, Sep. 2022, doi: 10.15212/zoonoses-2022-0034.

[29] J. F. Vesgan *et al.*, “Transmission dynamics and vaccination strategies for Crimean-Congo haemorrhagic fever virus in Afghanistan: A modelling study,” *PLoS Negl Trop Dis*, vol. 16, no. 5, p. e0010454, May 2022, doi: 10.1371/JOURNAL.PNTD.0010454.

[30] Q. Dong and Q. Zhang, “The Estimation of a Remote Sensing Model of Three-Dimensional Green Space Quantity and Research into Its Cooling Effect in Hohhot, China,” *Land 2022, Vol. 11, Page 1437*, vol. 11, no. 9, p. 1437, Aug. 2022, doi: 10.3390/LAND11091437.

[31] S. Li, L. Zhu, L. Zhang, G. Zhang, H. Ren, and L. Lu, “Urbanization-Related Environmental Factors and Hemorrhagic Fever with Renal Syndrome: A Review Based on Studies Taken in China,” *Int J Environ Res Public Health*, vol. 20, no. 4, p. 3328, Feb. 2023, doi: 10.3390/IJERPH20043328/S1.

[32] K. K. Al-jabery, T. Obafemi-Ajayi, G. R. Olbricht, and D. C. Wunsch II, “Data preprocessing,” *Computational Learning Approaches to Data Analytics in Biomedical Applications*, pp. 7–27, Jan. 2020, doi: 10.1016/B978-0-12-814482-4.00002-4.

[33] N. Bokde, M. W. Beck, F. Martínez Álvarez, and K. Kulat, “A novel imputation methodology for time series based on pattern sequence forecasting,” *Pattern Recognit Lett*, vol. 116, p. 88, Dec. 2018, doi: 10.1016/J.PATREC.2018.09.020.

[34] Z. Zhang, “Missing data imputation: focusing on single imputation,” *Ann Transl Med*, vol. 4, no. 1, p. 9, Jan. 2016, doi: 10.3978/J.ISSN.2305-5839.2015.12.38.

[35] R. Kohn and C. F. Ansley, “Estimation, prediction, and interpolation for ARIMA models with missing data,” *J Am Stat Assoc*, vol. 81, no. 395, pp. 751–761, 1986, doi: 10.1080/01621459.1986.10478332.

[36] L. Kong *et al.*, “Time-Aware Missing Healthcare Data Prediction Based on ARIMA Model,” *IEEE/ACM Trans Comput Biol Bioinform*, 2022, doi: 10.1109/TCBB.2022.3205064.

[37] A. Choudhary, S. Kumar, M. Sharma, and K. P. Sharma, “A Framework for Data Prediction and Forecasting in WSN with Auto ARIMA,” *Wirel Pers Commun*, vol. 123, no. 3, pp. 2245–2259, Apr. 2022, doi: 10.1007/S11277-021-09237-X/METRICS.

[38] S. Mancini, A. B. Francavilla, A. Longobardi, G. Viccione, and C. Guarnaccia, “Predicting daily water tank level fluctuations by using ARIMA model. A case study,” *J Phys Conf Ser*, vol. 2162, no. 1, p. 012007, Jan. 2022, doi: 10.1088/1742-6596/2162/1/012007.

[39] S. Sivaramakrishnan, T. F. Fernandez, R. G. Babukarthik, and S. Premalatha, “Forecasting Time Series Data Using ARIMA and Facebook Prophet Models,” *Big Data Management in Sensing: Applications in AI and IoT*, pp. 75–87, Jan. 2021, doi: 10.1201/9781003337355-4/FORECASTING-TIME-SERIES-DATA-USING-ARIMA-FACEBOOK-PROPHET-MODELS-SIVARAMAKRISHNAN-TERRANCE-FREDERICK-FERNANDEZ-BABUKARTHIK-PREMALATHA.

[40] L. Hansson, "Field Signs as Indicators of Vole Abundance," *J Appl Ecol*, vol. 16, no. 2, p. 339, Aug. 1979, doi: 10.2307/2402512.

[41] B. Takele, M. Yihune, and A. Bekele, "Composition, abundance and distribution of rodent species in Alemsaga Priority State Forest and farmlands, northwestern Ethiopia," *Afr J Ecol*, vol. 60, no. 3, pp. 447–455, Sep. 2022, doi: 10.1111/AJE.12976.

[42] "Help Online - Origin Help - Digitizer." https://www.originlab.com/doc/Origin-Help/Tool-Digitizer (accessed Mar. 16, 2023).

[43] "Help Online - Tutorials - Digitizer Tool." https://www.originlab.com/doc/Tutorials/Digitizer-Tool (accessed Mar. 16, 2023).

[44] "Interpolation." https://www.originlab.com/videos/details.aspx?pid=1889 (accessed Mar. 16, 2023).

[45] "Help Online - Origin Help - Interpolate/Extrapolate Y from X." https://www.originlab.com/doc/origin-help/math-inter-extrapoltate-yfromx (accessed Mar. 16, 2023).

[46] M. Drusch *et al.*, "Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services," *Remote Sens Environ*, vol. 120, pp. 25–36, May 2012, doi: 10.1016/J.RSE.2011.11.026.

[47] B. E. Smith, A. Gardner, A. Schneider, and M. Flanner, "Modeling biases in laser-altimetry measurements caused by scattering of green light in snow," *Remote Sens Environ*, vol. 215, pp. 398–410, Sep. 2018, doi: 10.1016/J.RSE.2018.06.012.

[48] M. Imran and A. Ahmad, "Enhancing data quality to mine credible patterns," *https://doi.org/10.1177/01655515211013693*, Jun. 2021, doi: 10.1177/01655515211013693.

[49] "The Python Language Reference — Python 3.10.10 documentation." https://docs.python.org/3.10/reference/ (accessed Mar. 16, 2023).

[50] "OriginLab - Origin and OriginPro - Data Analysis and Graphing Software." https://www.originlab.com/ (accessed Mar. 16, 2023).

[51] "KNIME | Open for Innovation." https://www.knime.com/ (accessed Mar. 16, 2023).

[52] "Google Earth Engine." https://earthengine.google.com/ (accessed Mar. 16, 2023).

[53] "pandas - Python Data Analysis Library." https://pandas.pydata.org/ (accessed Mar. 16, 2023).