IJID One Health 11 (2026) 100126

Contents lists available at ScienceDirect

IJID One Health

journal homepage: www.sciencedirect.com/journal/ijid-one-health

Exploring the drivers of Crimean-Congo hemorrhagic fever in Southern Iraq

Gina E.C. Charnley[a,b,*], Celine Tabche[a], Salman Rawaf[a], Mohammed H. Waheed[c], Hiam R. Sadkhan[c], Ahmed Fuoad[c], Tahseen S.F. Al-Nazal[c], Ali R. Hashim[d], Alaa K. Mousa[d], Hamzah Abbas Oudah[c], Zeenah Atwan[a,e]

[a] School of Public Health, Imperial College London, London, United Kingdom
[b] Institute for Global Health, University College London, London, United Kingdom
[c] Public Health Department, Basrah Health Directorate, Basrah, Iraq
[d] Department of Internal Medicine, College of Medicine, University of Basrah, Basrah, Iraq
[e] Central Laboratory, Department of Microbiology, College of Medicine, University of Basrah, Basrah, Iraq

ARTICLE INFO

Keywords:
Crimean-Congo hemorrhagic fever
Iraq
Displacement
Drought
Climate change
Hyalomma

ABSTRACT

Objectives: Crimean-Congo hemorrhagic fever (CCHF) is a zoonotic viral tick-borne disease that has seen a resurgence in Iraq in recent years, and this study aimed to explore environmental and socio-economic drivers of this resurgence in southeastern Iraq.

Methods: We used weekly CCHF case data by province and demographic strata (sex, children, adults, older adults) and performed initial covariate exploration using Pearson correlation and penalized feature screening (Least Absolute Shrinkage and Selection Operator (LASSO), elastic net). We fitted generalized additive mixed models (negative binomial) with cyclic seasonality and region-level random effects to a subset of the best performing covariates. Residual temporal autocorrelation (AR) was assessed and AR(1) mixed-effects refits were attempted until a best-fit model was identified.

Results: As anticipated, land use and many socio-economic covariates showed strong correlations in initial covariate exploration, and similar patterns were observed among humidity, temperature, and precipitation. The main model (total cases) explained ~80% deviance, showing strong seasonality and spatial heterogeneity. The self-calibrated Palmer Drought Severity Index and counts of incoming internally displaced persons were the most stable predictors.

Conclusion: In southeastern Iraq, CCHF risk was seasonal and spatially heterogeneous and environmental and socio-economic contexts likely modulate outbreaks and resurgence.

Introduction

Crimean-Congo hemorrhagic fever (CCHF) is a zoonotic viral tick-borne disease (genus *Orthonairovirus*) associated with severe hemorrhagic manifestations (fever, hepatitis, organ failure, and potentially extended periods of disability). Mortality varies by context, with high-fatality outbreaks documented in humans, worsened by a lack of specific treatment options or vaccines [1–3]. Transmission is largely through bites from infected *Hyalomma* ticks (mainly, *Hyalomma marginatum* spp.) or through contact with infected blood or tissues of viremic individuals. The virus is maintained in a wide range of wild and domestic animals, creating multiple amplifying and reservoir hosts. Furthermore, there is evidence of sexual transmission and vertical transmission during pregnancy [4,5].

CCHF was first described in Crimea in 1944 among soldiers and agricultural workers and later recognized as identical to a virus isolated from a child in the Congo in 1956. Today, CCHF is endemic in Africa, western and southeastern Asia, and the Balkans, with 3 billion people considered at risk [3,6]. CCHF was first documented in Iraq in the late 1970s, and early cases were sporadic and often underreported due to limited surveillance and diagnostic capacity [4,7,8]. The 2000s saw intermittent rural cases, and increased surveillance and reporting efforts in the 2010s highlighted recurrent outbreaks in southern and central provinces (Dhi-Qar, Wasit, and Babylon), with some nosocomial transmission [4,7,8]. From 2021 to 2022, there has been a significant surge and in 2022, Iraq experienced one of its largest-ever outbreaks (> 200 cases), particularly, in the southern provinces, resulting in

* Corresponding author.
E-mail address: georgina.charnley@imperial.ac.uk (G.E.C. Charnley).

https://doi.org/10.1016/j.ijidoh.2026.100126
Received 12 February 2026; Received in revised form 31 March 2026; Accepted 7 April 2026
Available online 10 April 2026

dozens of deaths (case fatality rate ~30%) [2,4,8,9]. Health system constraints (limited personal protective equipment and disrupted veterinary and tick control services) and cultural/occupational exposures (e.g. animal slaughtering during religious practices) may have amplified the risk [2,4,9].

Factors such as increased temperatures, uncontrolled animal movement, and vector behavior have been identified as contributors to the 2022 outbreak in Iraq. Environmental changes are thought to have played a role in the resurgence of CCHF in Iraq [10]. The country has experienced increased temperatures and droughts, creating favorable conditions for *Hyalomma* proliferation, which thrive in high temperatures [10,11]. These climatic conditions have also led to rural displacement, with over 12,000 families affected as of March 2023. Displaced populations often reside in areas with limited access to health care and are more likely to engage in high-risk activities, such as livestock handling without proper protective measures. Furthermore, the transboundary movement of animals, often unregulated, has facilitated the spread of ticks and the CCHF virus across regions [12].

CCHF is the most widely distributed viral tick-borne disease in the world, and the second most distributed viral hemorrhagic fever. It is one of the seven highest-priority epidemic-prone diseases by the World Health Organization (WHO) and included in the WHO R&D Blueprint priorities for research due to its potential for high mortality outbreaks [2,3]. Calls for more research and funding into CCHF are now long-standing, but the disease continues to receive scant attention scientifically and politically. Given the heterogeneity and socio-economic and cultural divergences among endemic regions, potential reasons for sporadic cases of CCHF cannot be explained by a single mechanism. The resurgence of CCHF in Iraq underscores the complex interplay between socio-economic and environmental factors. To further explore these complex interactions, we used recent CCHF data for southeastern Iraq and generalized additive mixed models to explore the links between several known CCHF drivers.

Materials and methods

Data processing and harmonization

We used daily CCHF line list data provided by the Public Health Department, Basrah Health Directorate, Basrah, Iraq, reported by district and province. The data were aggregated to ISO-8601 epidemiologic weeks and to the Global Administrative Areas unit 1 (provinces) scale [13]. These spatial and temporal units were chosen because they best represented the spread of the data across the period (e.g. to ensure sufficient cases fell within the spatial/temporal unit). Weekly case counts were disaggregated by males, females, children (less than 16 years of age), adults (16 to less than 60 years old), and older adults (60 years or older); these six outcomes formed the basis of all modeling analyses. The age of 60 years or older was used to define “older adults” to reflect the population structure in Iraq, where the average life expectancy is 72 years [14] and ensure sufficient sample sizes within each subgroup for stable model estimations.

Climatic and socio-economic covariates were harmonized to the common temporal/spatial scales of the CCHF outcome data (administrative unit 1 and weekly). Records containing missing values were removed using complete case exclusion to ensure consistent inputs across all models. Data which were only available to a courser resolution than required were repeated and those on a finer scale were averaged. Full information, including original granularity and sources, are available in Table 1. All analyses were conducted in R (version 4.3.2), using the packages tidyverse, glmnet, mgcv, and glmmTMB [15–18].

Exploratory analysis

We first generated descriptive time series plots of weekly case counts stratified by administrative unit (GID1) to visualize seasonal and

Table 1
Full details of the environmental and socio-economic covariates considered in the best-fit generalized additive mixed-effect models.

Covariate group	Covariates	Original spatial granularity	Original temporal granularity	Source
Land Cover	Coverage of trees, scrubs, grass, built-up, bare, seasonal/permeant water, moss/lichens, snow	% coverage to 100 m × 100 m grid cells	Static, 2020	Copernicus Land Monitoring Service [34]
Elevation	Elevation in meters above sea level	30 m × 30 m grid cell	Static, 2020	NASA SRTM [35]
Conflict	Number of reported conflict events	X,Y coordinate	Daily	ACLED [36]
Meteorology	Temperature (°C), relative humidity (%) and precipitation (mm)	0.25° × 0.25° grid cell	Daily	Copernicus Climate Data Store, ERA [5,37]
Soil moisture	Surface soil moisture (%)	0.25° × 0.25° grid cell	Daily	Copernicus Climate Data Store [38]
Drought	Self-calibrated palmers drought severity index	0.5° × 0.5° grid cell	Monthly	University of East Anglia Climate Research Unit [39]
Islamic Holidays	Binary presence of Eid al-Fitr or Eid al-Adha in current or previous week	National	Weekly	Republic of Iraq [40]
Poverty	Multidimensional poverty index value, headcount ratio, intensity, vulnerability, and severe poverty	Administrative Unit 1	Annually	United Nations Development Programme [41]
Human Displacement	Number of incoming internally displaced people	Administrative Unit 1	Quarterly	International Organization for Migration [42].

inter-regional heterogeneity across all six outcomes. We then evaluated multicollinearity among covariates by computing Pearson correlation coefficients across all continuous predictors and plotting the resulting matrix as a labeled heatmap. This assessment informed the decision to apply penalized regression methods during feature selection.

Preliminary feature selection using penalized regression

Due to the large number of potential covariates and their high degree of correlation, we used penalized regression to perform outcome-specific feature screening. Each of the six outcomes were modeled independently, reflecting potential differences in demographic risk structures and exposure pathways. For each outcome, we constructed a model matrix of all candidate covariates and fitted Poisson models using the *glmnet* framework [16], applying two complementary penalization schemes:

1. Least Absolute Shrinkage and Selection Operator (LASSO) regression ($\alpha = 1$): to induce sparsity and eliminate uninformative predictors by shrinking coefficients exactly to zero.
2. Elastic net regression ($\alpha = 0.5$): to stabilize selection in the presence of multicollinearity, distributing weight across groups of correlated predictors rather than arbitrarily selecting a single variable.

In both models, the regularization parameter (λ) was selected using K-fold cross-validation. After fitting, we extracted all covariates whose coefficients were non-zero at the optimal penalty parameter. LASSO and elastic net impose different shrinkage structures; therefore, agreement between them provides a stringent indicator of a covariate with a genuine and reproducible association [19]. We defined the "top covariates" for each outcome as those jointly selected by both penalized models, that is, the intersection of the non-zero coefficient sets. This stability-based criterion naturally yields a different number of covariates for each outcome, reflecting underlying variation in signal strength and demographic-specific drivers.

This two-step screening process reduces overfitting risk, constrains the model search space, and limits the multiple-comparison burden before fitting more flexible models. Variable importance metrics were derived from the absolute values of the non-zero coefficients, and the stable predictors were visualized to facilitate interpretability. The resulting covariate sets represent outcome-specific, data-driven subsets of the original candidate variables and were exported for consideration in subsequent best-fit model development.

Main modeling framework

Weekly case counts for each of the six outcomes (total, male, female, child, adult, older) were independently modeled using generalized additive mixed models (GAMs), with a negative binomial likelihood to accommodate overdispersion observed in exploratory analyses [20,21]. For each outcome, the base model included the following:

1. A cyclic cubic spline for epidemiologic week, s(epiweek, bs = "cc," k = 12), which flexibly captures seasonal patterns while ensuring continuity between the final and first weeks of the year.
2. A region-specific random effect, s(GID1, bs = "re"), to account for unobserved spatial heterogeneity in baseline risk.

The selection of a negative binomial family was motivated by the presence of variance exceeding the mean in all outcomes, which would produce biased standard errors under a Poisson model. Allowing a dispersion parameter yields more reliable inference for over-dispersed surveillance counts [17,22]. Each of the six outcomes were modeled separately to allow outcome-specific seasonality, spatial heterogeneity, and covariate effects to emerge.

Forward Akaike Information Criterion (AIC)–based covariate selection

For each outcome, the set of covariates identified during penalized regression was entered into a forward stepwise model selection procedure. Beginning from the base seasonal + random effects model, each candidate covariate was added individually, and the corresponding AIC was computed. Covariates producing a reduction in AIC greater than 2 were retained, and the process iterated until no further improvement greater than the threshold was detected.

AIC was used because it balances model fit with parsimony and provides a likelihood-based metric appropriate for comparing non-nested GAMs with differing fixed-effect structures. Conducting this procedure separately for each of the six outcomes allowed the final covariate sets to differ across demographic subgroups [23].

Importantly, this forward stepwise procedure is inherently path-dependent: the first covariate selected is the one that gives the greatest immediate reduction in AIC, and this initial choice influences which covariates appear beneficial at subsequent steps. Covariates that are correlated with the first-selected variable may no longer offer sufficient additional improvement to be retained, even if they would have been selected had the search begun from a different starting point. Thus, each outcome's final model represents the best incremental improvement under an AIC-minimization strategy applied to the pre-filtered covariate set rather than a guarantee of the globally optimal combination of predictors.

Assessment of temporal autocorrelation and AR(1) mixed-effects refinement

After fitting the final GAM for each outcome, we examined Pearson residuals to assess within-region temporal autocorrelation. For each administrative unit, the lag-1 Autocorrelation Function (ACF1) was computed, and the median ACF1 across all regions was used as a robust overall diagnostic. A threshold of median ACF1 > 0.25 indicated that serial dependence was not adequately captured.

In such cases, the model was re-estimated using a generalized linear mixed model with an AutoRegressive (AR)(1) correlation structure (glmmTMB). These models were specified with the same fixed-effect covariates identified by the GAM, a random intercept for GID1 and an AR(1) correlation term over epidemiologic week within each region. This approach distinguishes long-term spatial heterogeneity (random intercepts) from short-term epidemic dynamics (serial correlation), improving inference when autocorrelation is non-negligible. Convergence issues were noted where present. This refinement was performed independently for each of the six outcomes, reflecting their distinct residual structures [18,24].

Sensitivity analysis

We ran several sensitivity analyses to provide further clarification and strengthening around methodological decisions; these included analyses related to temporal aggregation, structured covariates, temporal autocorrelation, and leave-one-year-out (LOYO) cross-validation.

To assess the potential impact of aggregating covariates from differing temporal resolutions to weekly scale, we conducted a sensitivity analysis using data aggregated to monthly resolution. Weekly case counts and covariates were aggregated to calendar months at the administrative unit 1 level. The primary model structure was then refit for the main outcome (total cases), retaining the same covariates identified in the weekly analysis (self-calibrated Palmer Drought Severity Index [scPDSI] and incoming internally displaced persons [IDPs]), alongside a cyclic seasonal smooth for month and region-level random effects. Model coefficients and overall fit were compared with the original weekly model to evaluate the robustness of the identified associations.

To assess whether covariate selection was sensitive to the omission of seasonal and spatial structure during penalized screening, we conducted an additional analysis in which candidate covariates were

evaluated within the full modeling framework. A baseline GAM, including only a cyclic seasonal smooth and region-level random effects, was first fitted. Candidate covariates were then assessed based on their incremental contribution to model fit within this framework.

To evaluate residual temporal dependence beyond lag 1, we computed within-region Pearson residual autocorrelation up to lag 6. Median autocorrelation across regions was used as a summary diagnostic.

To assess whether model results were driven by any single year, we conducted a LOYO analysis. The model was refit iteratively excluding each year in turn, and predictive performance was evaluated on the held-out data. Covariate selection and coefficient estimates were compared across folds to assess stability.

Results

Spatial and temporal structure of the CCHF data

Between April 19, 2022 to August 10, 2025, there were 497 unique cases of CCHF reported in the line list data collected by the Public Health Laboratory in Basrah, Iraq; these ranged in age from 1 to 90 years and were mainly reported in Al-Basrah (n = 441) and Maysan (n = 46), followed by Dhi-Qar (n = 6), Al-Qadisiyah (n = 1), Karbala' (n = 1), and Wasit (n = 1). Demographically, males exceeded females (308 vs 189), and adults accounted for most cases (n = 399), with 62 older adults and 30 children.

The main model outcome (total cases) is shown in Figure 1 in the temporal and spatial resolution of the analysis (administrative unit 1 and epidemiologic week). The year 2023 reported the highest number of cases across the observation period with 317 cases, although it should be noted that at the time of the analysis, the year of 2025 was incomplete.

Preliminary analysis

Pearson correlation coefficient values for all the continuous variables included in the analysis for the full data set used for model fitting are presented in Figure 2. As anticipated, land use and many socioeconomic covariates showed strong correlations. Similar patterns were observed among humidity, temperature, and precipitation, reflecting their intrinsic coupling within the global climate system. This high degree of collinearity among many covariates confirmed the need for a preliminary covariates selection process before identifying covariates for the best-fit models Figure 2.

For the feature selection using elastic net and LASSO, precipitation variables produced particularly high values across all outcomes, followed by tree cover, permanent and seasonal water, scPDSI, and holidays in the previous or current week. Using older adult cases as the outcome gave the most distinct results, with land cover and holidays having a much stronger impact, with the proportion urban land (BuiltUp) also having particularly high variable importance. Figure 3 shows the variable importance values for the main outcome (total cases); the importance values for the demographic outcomes and the full table of importance values for the top selected covariates for each outcome are shown in Appendix 1–2.

Best-fit model according to the main modeling framework

Across the potential covariates identified during penalized feature screening, the stepwise AIC procedure selected two predictors for the best-fit GAM for the main outcome (total cases): scPDSI and the number of IDPs. These covariates produced the largest incremental improvements in model fit when added to the baseline seasonal spline and regional random effect's structure.

The final GAM explained a substantial proportion of the variation in weekly incidence (deviance explained: 80.1%), indicating strong overall model performance. The model retained the cyclic seasonal spline for epidemiologic week (edf $\approx$ 6.26, F = 10.0, $P < 0.001$), describing a pronounced non-linear seasonal pattern and a random intercept for GID1 regions (edf $\approx$ 2.77, F = 4.0, $P < 0.001$), confirming meaningful spatial heterogeneity in baseline risk.

Both selected covariates showed statistically significant associations with weekly case counts.

- scPDSI was positively associated with incidence (β = 0.426, SE = 0.075, $P = 1.8 \times 10^{-8}$), indicating higher risk during wetter conditions (positive scPDSI denotes wetter conditions, negative scPDSI denotes drier).
- IDP numbers also demonstrated a positive association (β = 0.000945, SE = 0.000330, P = 0.004), consistent with elevated transmission potential or increased vulnerability in areas experiencing an increase in IDPs.

The estimated random-effect variance for GID1 was modest (0.036), but its significance suggests structural regional differences remained even after accounting for environmental and displacement-related predictors. Temporal autocorrelation was negligible; the median within-region lag-1 ACF was −0.014, below the prespecified threshold for AR(1) correction, and AR(1) mixed-effects refitting was not required. No warnings related to dispersion or convergence were encountered. The full output for the model is shown below in Table 2.

Sex-specific models

For male and female cases, the best-fit model retained two covariates: scPDSI and IDP numbers, the same as the total cases. Higher scPDSI values were associated with increased male case counts (β = 0.436, SE = 0.080, $P = 7.2 \times 10^{-8}$) and female counts (β = 0.523, SE = 0.105, $P = 7.6 \times 10^{-7}$). IDP numbers also showed a positive association with male (β = 0.00104, SE = 0.00036, P = 0.0035) and female cases (β = 0.00154, SE = 0.00045, P = 0.00074). The male cases model performed slightly better (probably due to a larger number of cases), explaining 80.5% of deviance, compared with 70.1% in the female case model. There was a strong seasonal pattern (male: s[epiweek], edf $\approx$ 6.03, F = 10, female: edf $\approx$ 4.96, F = 10) alongside significant spatial heterogeneity (male: s[GID1], edf $\approx$ 2.76, F = 3, female: s[GID1], edf $\approx$ 2.77, F = 4). Residual temporal autocorrelation was minimal in both models (median ACF1 = 0.11 and −0.006 for males and females, respectively), remaining below the threshold for AR(1) correction, and the models showed no convergence or dispersion issues.

Age-specific models

Adults formed the largest demographic group in the line list data and, therefore, as expected, the best-fit model for adults closely aligned with the model for all cases, selecting the same covariates (scPDSI β = 0.465, SE = 0.073, $P < 1 \times 10^{-9}$ and IDP numbers β = 0.00117, SE = 0.00033, P = 0.00039) and performed similarly (79.97% deviance explained, median ACF1 −0.012). In contrast, for child cases, the best-fit model included two covariates: tree cover and soil moisture. Soil moisture showed a significant negative association with weekly child case incidence (β = −0.521, SE = 0.217, P = 0.017), indicating lower risk during wetter soil conditions. Tree cover had a large negative coefficient but with very high uncertainty (β = −173.44, SE = 5105.50, P = 0.97), suggesting an unstable or weak effect. Seasonal structure was captured by the weekly smooth term (edf $\approx$ 1.84, $P \approx 0.75$), and substantial spatial variation remained (s[GID1], edf $\approx$ 1.19, $P \approx 1$). The model explained 72.6% of deviance; however, unlike other outcomes, residual temporal autocorrelation was notable (median ACF1 = 0.47), exceeding the threshold for AR(1) correction, and the

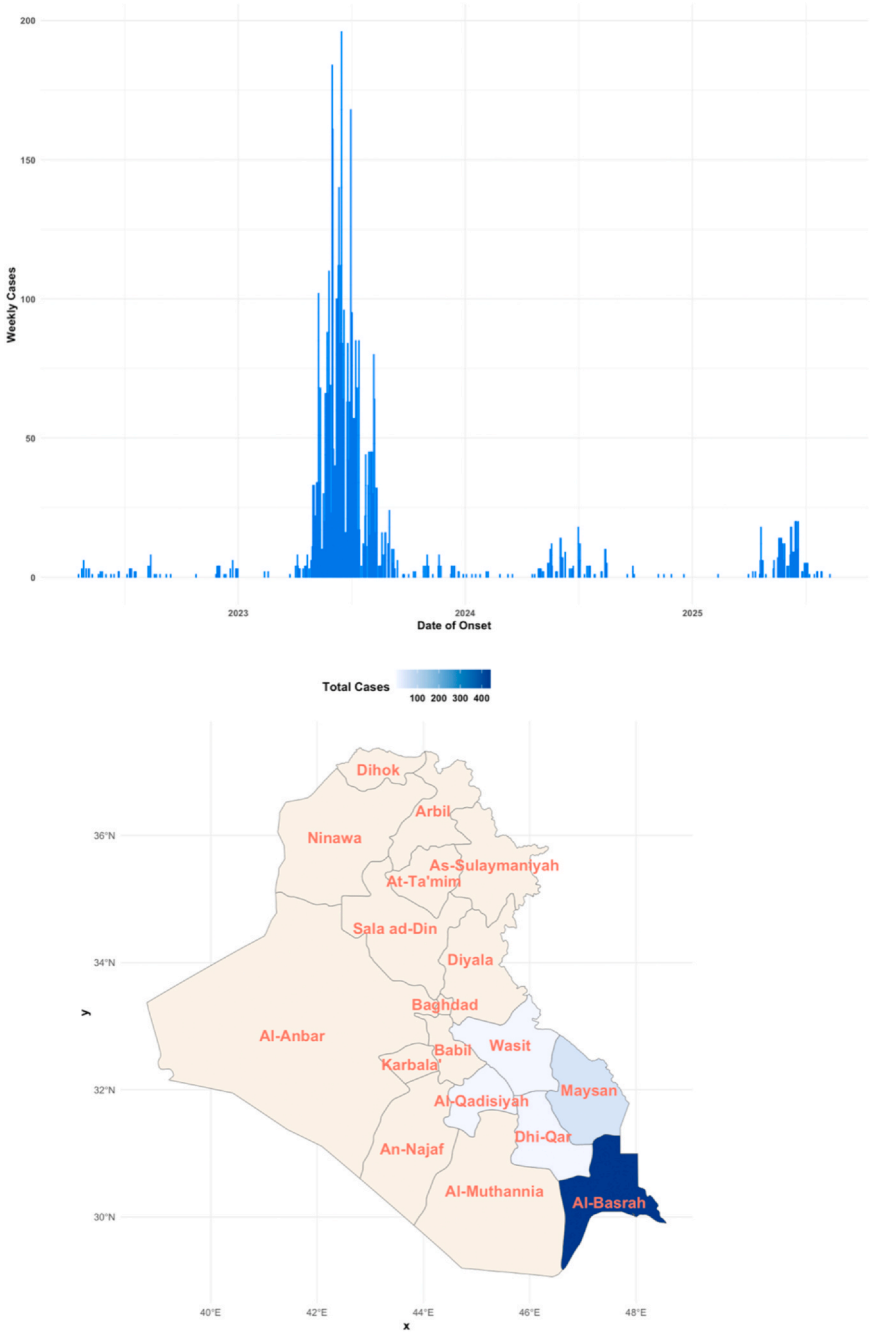

Figure 1. The temporal (top) and spatial (bottom) resolution of the data fit to the models here. The temporal scale is by epidemiologic week and the spatial scale is administrative unit 1.

AR(1)-refitted glmmTMB model did not converge, potentially due to the lower number of child cases over the study period.

For older adult cases, the AIC procedure selected three covariates: tree cover, scPDSI, and the presence of a holiday in the current week. Among these, only scPDSI was statistically significant ($\beta = 0.254$, SE = 0.113, $P = 0.025$), indicating higher older incidence during wetter-than-normal weeks. Tree cover showed a borderline negative association ($\beta = -30.30$, SE = 16.03, $P = 0.059$), whereas the holiday indicator was non-significant ($\beta = -16.7$, SE = 1377.4, $P = 0.99$). The seasonal smooth (edf $\approx$ 3.86, $F \approx 0.33$) and spatial term (edf $\approx$ 1.40, $F \approx 1$) showed comparatively weak structure relative to other demographic groups. Residual temporal autocorrelation was substantial (median ACF1 = 0.39), suggesting that within-region temporal dependence was not fully captured; the AR(1) glmmTMB refit did not converge. Random-effect variance was the highest among all outcomes ($\approx$1.52), indicating notable unexplained spatial heterogeneity and only 58.7% of deviance explained. A full table for all model outputs and performance metrics for the demographic outcomes are shown in Appendix 3.

Sensitivity analysis

When the data were aggregated to monthly resolution and the primary model refit, the association between scPDSI and CCHF incidence remained positive and statistically significant ($\beta = 0.234$, SE = 0.112, $P = 0.03$), consistent with the weekly analysis. In contrast, the effect of incoming IDPs was attenuated and no longer statistically significant at monthly resolution. Overall model performance was comparable, with slightly higher deviance explained at monthly scale (82.7%). These

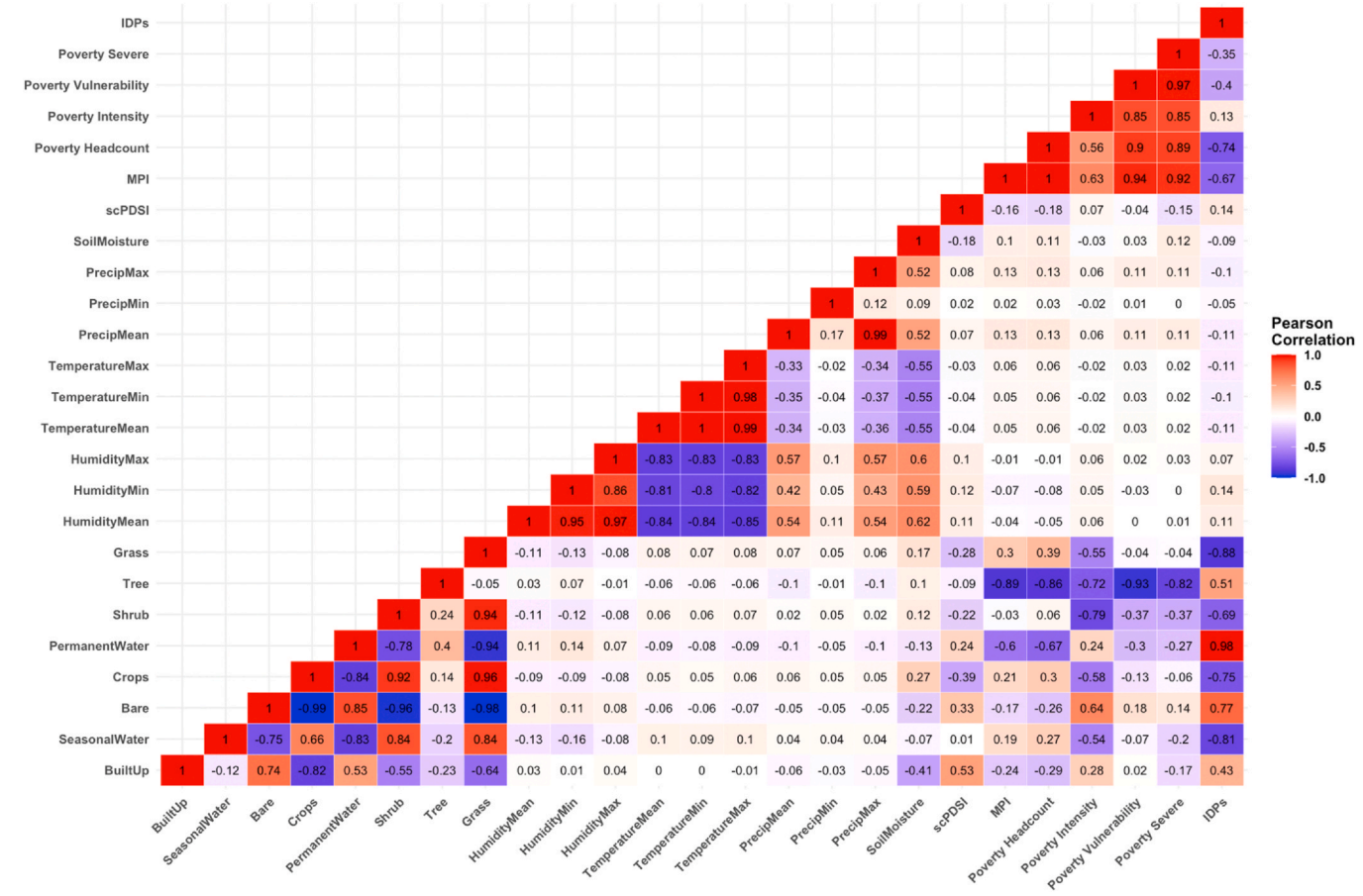

Figure 2. Pearson correlation coefficient values for all continuous variables used in the analysis. IDPs, internally displaced persons; MPI, multidimensional poverty index; Precip, precipitation; scPDSI, self-controlled palmer drought severity index.

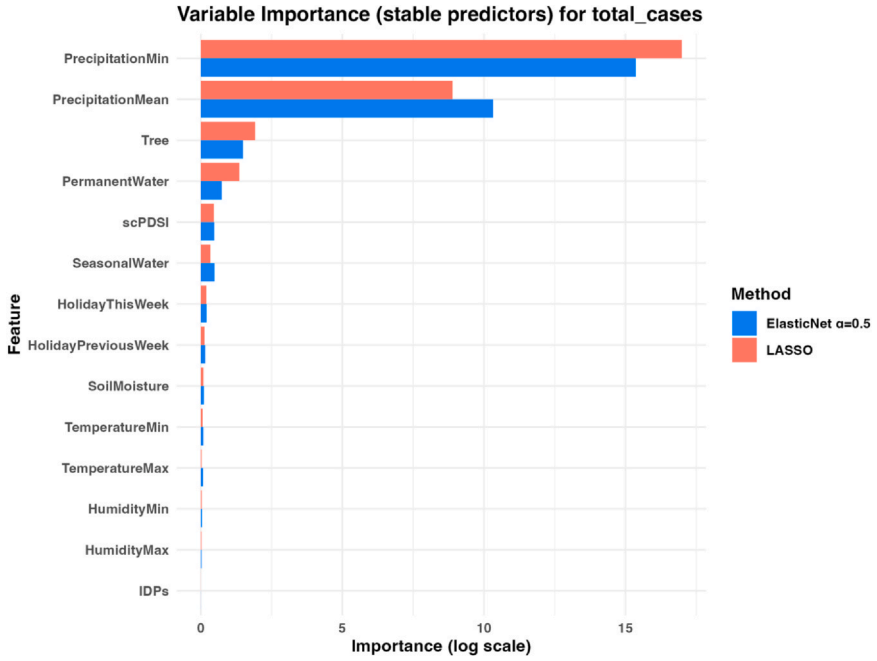

Figure 3. Variable importance values for the selected covariates for the main outcome (total cases) using the two methods: elastic net and LASSO. The importance values are on a logarithmic scale. IDPs, internally displaced persons; scPDSI, self-controlled palmer drought severity index.

Table 2
Model output and performance metrics for the best-fit generalized additive mixed model when considering the best performing subset of covariates against total numbers of weekly cases.

Selected covariates	Akaike Information Criterion	Deviance explained	Random effect	RE variance	AR1 recommended	Median ACF1
scPDSI, IDPs	658.5423	0.8011	TRUE	0.0364	FALSE	−0.0135
Parametric terms		**Estimate**	**Std_Error**		***P*-value**	
(Intercept)		−7.0073	3.0282		0.0209	
scPDSI		0.4255	0.0746		1.8401×10^{-8}	
IDPs		0.0009	0.0003		0.0043	
Smooth terms			**EDF**		**F**	
s(epiweek)			6.2589		10	
s(GID1)			2.7718		4	

IDPs, internally displaced persons; scPDSI, self-controlled palmer drought severity index.

findings indicate that the main hydro-climatic signal is robust to temporal aggregation, whereas the displacement-related effect appears more sensitive to temporal resolution.

The baseline model including only seasonal and spatial structure explained 76.8% of the deviance. When candidate covariates were evaluated within this framework, scPDSI and incoming IDPs remained the strongest predictors and were retained in the final model, increasing the deviance explained to 80.1%. This indicates that the primary covariates identified during penalized screening are robust to the inclusion of seasonal and spatial effects.

Residual temporal autocorrelation was minimal across all examined lags. Median within-region lag-1 autocorrelation was close to zero (−0.014), and higher-order lags showed similarly low values (e.g. lag-2 ≈0.08, with subsequent lags near zero), indicating no substantial remaining temporal dependence.

In the LOYO analysis, the same covariates (scPDSI and IDPs) were consistently selected across all folds, indicating stable model structure. Coefficient estimates varied between folds, particularly, for scPDSI, reflecting inter-annual variability, but the overall importance of hydro-climatic conditions was maintained. Predictive performance on held-out years was reasonable, with the model capturing the overall structure of observed incidence. These findings suggest that the main results are not driven by any single outbreak year.

Discussion

We analyzed 4 recent years (2022–2025) of CCHF line list data for southeastern Iraq, aggregated to a weekly and administrative unit one temporal and spatial scale, respectively. The cases were heavily concentrated in the province of Al-Basrah, and the largest number of cases were reported in 2023. We explored a range of environmental and socio-economic covariates against weekly cases of CCHF and against subsets of the cases by demographic groups to see whether the relationships between CCHF risk factors altered by age or gender. Using GAMs, we generally found similar results across the demographic groups, with the main model (total cases) explaining 80.5% of deviance and showing a strong seasonal pattern and significant spatial heterogeneity. scPDSI and numbers of incoming IDPs were consistently found to be the most stable predictors, except in the child and older adult outcomes; however, these models performed poorer, potentially due to the smaller number of cases than the adults and gender demographic groups.

In our models, higher scPDSI values, indicating wetter-than-normal conditions, were consistently associated with increased CCHF incidence, in contrast to precious work which states drier and hotter conditions being more suited for *Hyalomma* ticks [25]. Wetter periods in otherwise relatively dry landscapes promote vegetation growth and improve microclimatic conditions for off-host tick stages, promoting survival, and simultaneously supporting higher densities and movement of livestock and wildlife hosts [26]. These finding additionally support the inclusion in some models and the importance found in the preliminary analysis of tree cover, which, as an ecosystem, often supports animals and ticks. Together with seasonally intensified agricultural and animal-handling activities, these conditions create a landscape of elevated human exposure to infected ticks and animal blood. Similar associations between CCHF incidence and rainfall, humidity, or vegetation indexes have been reported in Iran, Turkey, India, and other endemic settings, and risk-mapping studies highlight precipitation and vegetation as key environmental correlates of CCHF risk rather than extreme aridity [27–29].

Interestingly, in support of previous studies that found wetter periods leading to higher CCHF incidence, precipitation and vegetation were found to be the most important predictors during the preliminary elastic net and LASSO analysis. However, the variable importance from LASSO/elastic net differs from the GAM AIC-selected covariates because the two methods solve different modeling problems. LASSO ranks predictors for linear predictive performance, with no seasonality or spatial structure, whereas the GAM evaluates incremental explanatory value after accounting for non-linear seasonality, spatial random effects, negative binomial error, and correlations between covariates [19]. As a result, many variables that appear important in LASSO become redundant once key structural components were included in the GAM. scPDSI perhaps dominated in the final models, compared with precipitation, because it captures moisture, soil water, vegetation and livestock and human agriculture behavior all in one index.

The second most stable predictor across the six outcomes was a higher number of IDPs coming into the reporting province. The effect was positive but to a much smaller magnitude than for scPDSI and less stable across the sensitivity analyses. Suggested mechanisms for why this positive relationship may have emerged here may be due to higher IDPs also leading to more animal movements, which may be infected and/or be carrying infected *Hyalomma* ticks, especially from marsh areas that are affected the most from droughts in Iraq [30,31]. Once these IDPs arrive and settle temporarily, this may then lead to extra pressure on health care systems, potentially reducing early testing and reporting and leading to higher nosocomial transmission. The attenuation of the IDP effect at monthly resolution likely reflects the coarser temporal granularity of displacement data and reduced statistical power after aggregation, which may limit the detection of shorter-term dynamics.

On the contrary to the finding that higher incoming IDPs increased CCHF transmission, the Multidimensional Poverty Index (MPI) nor the number of conflict events were found to be significant covariate. For MPI, this may be due to the coarseness of the data, only being reported as a static value through the analysis for each province and, therefore, unlikely to capture more subtle nuances in poverty and inequity in the region. More cases are reported from the districts outside Basrah and few other spots in central areas where poverty is relatively high and overcrowding are more abundant. For example, farmers who were affected by the dry season, impacting their animals and lands, left the marshes and the agricultural lands toward the city center [31]. In addition, if the risk factor is a higher number of incoming IDPs rather than

outgoing, this may reflect why conflict events were not found to be significant, with potentially conflict events in neighboring provinces being a better performing covariate. Although southeastern Iraq has been comparatively more politically stable after 2003; therefore, drought-related displacement and agricultural stressors likely predominate local displacement [10,11].

Despite the model for older adults not performing as well, likely due to a lower number of cases in that demographic group, the model did select the presence of a holiday week (namely, Eid al-Fitr and Eid al-Adha) in the best-fit model, and these predictors were also found to be relatively important during the elastic net and LASSO preliminary analysis. The presence of a holiday or religious festival is consistent with known risk factors for CCHF transmission [11], along with other infectious diseases, such as Rift Valley fever and Middle-East Respiratory Syndrome (MERS) [32,33]. During religious festivals, there is often increased animal movement, uncontrolled slaughtering, and an increase in consumption of animal products which may be infected with the pathogen. The findings here suggest that in southeastern Iraq in recent years, the timing of Eid may have led to an increase in CCHF transmission and an increase in education around this time may help to reduce its spread.

The study has several limitations, the first being the strength of scPDSI as a predictor, which is a coarse hydro-climatic index and may have captured a range of different factors such as moisture, vegetation, host productivity, and human seasonality, which cannot be untangled. It should be noted that our results do not mean drought is irrelevant to CCHF in Iraq (drier conditions has been linked to CCHF in previous studies [10,25]) but rather that the risk signal in the data set sits on the "wetter" side of the curve. In addition, the mechanisms identified here could differ by region and timescale because only a relatively constrained data set was used in terms of its spatial scale and granularity and the temporal length of the data set. Droughts occur over long-time scales, not weeks, which were used here; therefore, it may be interesting to see whether the relationship with CCHF and wetter-than-normal conditions hold over months/years, where drier conditions seeing higher CCHF incidence is perhaps more likely to proliferate. A greater range of covariates, particularly, density of livestock, access to health care, and tick presence may yield different results and find different stable predictors; however, in this instance, data sets on these variables were not available. For example, local knowledge and experience have found that calves rather than older cattle carry a higher burden of *Hyalomma* ticks; therefore, the age structure of the livestock population could be driver CCHF in the area. Future work should incorporate livestock/tick surveillance, health care access, and population denominators/exposure within a One Health framework.

Conclusion

In southeastern Iraq (2022–2025), weekly CCHF variation was best explained by seasonality, spatial heterogeneity, and two key predictors: wetter-than-normal hydro-climatic conditions (scPDSI) and incoming IDPs. Extra vigilance is likely needed during both extremes; ensuring adequate education to reduce the spread and sustainable health care, which possess the ability to test and reduce onward transmission, is essential going forward. Strengthening seasonal Risk Communication and Community Engagement (RCCE), vector/animal movement and slaughtering oversight, and health system infection prevention and control, particularly, around festival periods, are essential to mitigate transmission and future outbreaks. Sustained international and multi-sector support and coordination is crucial to mitigate the impact of CCHF and prevent future outbreaks in Iraq and beyond.

Author contributions

GECC, CT, and ZA conceptualized the study; GECC designed the study design and ran the analysis; SR, MHW, HRSH, AO, TSAN, AF, ARH, AAM, and ZA collected and stored the CCHF data; GECC wrote the initial draft of the manuscript; all authors were involved in interpretation of the results and gave feedback to the initial draft; GECC finalized the final manuscript for submission.

Ethical approval

There was no direct human or animal involvement in this work and all data used in the analysis here were fully anonymized or relied on open-sources which are freely available; therefore, no ethics review was needed.

Code availability

All codes are available via an open-access GitHub repository (https://github.com/GinaCharnley/CCHF-IRQ), and covariate data were extracted via open sources which are cited throughout the manuscript.

Funding

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

Data availability

For more information regarding the CCHF data, please contact the Public Health Department, Basrah Health Directorate, Basrah, Iraq.

Declaration of competing interest

The authors have no competing interests to declare.

Acknowledgments

We would like to acknowledge our colleagues at the Molecular Laboratory and communicable diseases and surveillance departments for their great support and enabling us to obtain all the necessary approvals. GECC, CT and ZA acknowledge joint center funding from the UK Medical Research Council and the Department for International Development [MR/R0156600/1].

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ijidoh.2026.100126.

References

[1] World Health Organization. Crimean-Congo haemorrhagic fever, 〈https://www.who.int/news-room/fact-sheets/detail/crimean-congo-haemorrhagic-fever〉; 2025 [accessed 18 December 2025].
[2] World Health Organization. Investing in risk communication and community engagement to battle Crimean-Congo haemorrhagic fever in Iraq, 〈https://www.who.int/about/accountability/results/who-results-report-2020-mtr/country-story/2022/investing-in-risk-communication-and-community-engagement-to-battle-crimean-congo-haemorrhagic-fever-in-iraq〉; 2022 [accessed 08 May 2025].
[3] Mostafavi E, Salehi-Vaziri M, Keyvani H, Aghamohammadi A, Tazmini M, Mokhtari-Azad T, et al. Emerging and re-emerging infectious diseases in the WHO eastern Mediterranean Region, 2001–2018. Int J Health Policy Manag 2022;11:1286–300. https://doi.org/10.34172/ijhpm.2021.13.
[4] Sah R, Mohanty A, Abdelaal A, Reda A, Rabaan AA, Rodriguez-Morales AJ. Crimean-Congo haemorrhagic fever (CCHF) outbreak in Iraq: currently emerging situation and mitigation strategies. Int J Surg 2022;106:106916. https://doi.org/10.1016/j.jisu.2022.106916.
[5] Pshenichnaya NY, Sydenko IS, Klinovaya EP, Romanova EB, Zhuravlev AS. Possible sexual transmission of Crimean-Congo hemorrhagic fever. Int J Infect Dis 2016;45:109–11. https://doi.org/10.1016/j.ijid.2016.03.005.
[6] Dreshaj S, Ahmeti S, Ramadani N, Dreshaj G, Humolli I, Dedushaj I. Current situation of Crimean-Congo hemorrhagic fever in southeastern Europe and

neighboring countries: a public health risk for the European Union? Travel Med Infect Dis 2016;14:81–91. https://doi.org/10.1016/j.tmaid.2015.12.003.
[7] Al-Obaidi S, Kadhim M, Al-Shadeedi SM, Abdulkareem S, Al-Kaisy S. Epidemiological and clinical features of Crimean–Congo hemorrhagic fever in Iraq: a retrospective study. Infect Dis 2023;55:128–36.
[8] Jasim H, Al-Dabbagh S, Ghadban J, Ismail S. Outbreak of Crimean–Congo haemorrhagic fever in Iraq: an emerging threat. BMJ Glob Health 2023;8:e012670.
[9] Médecins Sans Frontières. Iraq: bridging the gaps to curb the spread of Crimean Congo haemorrhagic fever, ⟨https://msf-lebanon.org/news/iraq-bridging-the-gaps-to-curb-the-spread-of-crimean-congo-haemorrhagic-fever/⟩; 2023 [accessed 08 May 2025].
[10] Khazaal S, Abdulrahman Z, Abdalla I. Climate change and the re-emergence of Crimean–Congo hemorrhagic fever in Iraq: a climate–health crisis. One Health 2024;17:100600.
[11] Atwan Z, Alhilfi R, Mousa AK, Rawaf S, Torre JDL, Hashim AR, et al. Alarming update on incidence of Crimean-Congo hemorrhagic fever in Iraq in 2023. IJID Reg 2023;10:75–9. https://doi.org/10.1016/j.ijregi.2023.11.018.
[12] Alwan KH, Al-Jawari SM, Mohammed AM, Obaid YH. The impact of climate change on population migration and displacement within the southern governorates of Iraq. Int. J. Environ. Impacts 2025;8:1267–77.
[13] Global administrative areas (GADM). GADM maps and data, ⟨https://gadm.org⟩; 2025 [accessed 18 December 2025].
[14] World Bank. Life expectancy at birth, total (years) – Iraq, ⟨https://data.worldbank.org/indicator/SP.DYN.LE00.IN?locations=IQ⟩; 2023 [accessed 18 December 2025].
[15] Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the tidyverse. J Open Source Softw 2019;4:1686. https://doi.org/10.21105/joss.01686.
[16] Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010;33:1–22. https://doi.org/10.18637/jss.v033.i01.
[17] Wood SN. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. J R Stat Soc B 2011;73:3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x.
[18] Brooks ME, Kristensen K, Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R J 2017;9:378–400. https://doi.org/10.32614/RJ-2017-066.
[19] Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc B 2005;67:301–20. https://doi.org/10.1111/j.1467-9868.2005.00503.x.
[20] Wood SN. Generalized additive models: an introduction with R. 2nd ed. Boca Raton: Chapman & Hall/CRC; 2017.
[21] Hilbe JM. Negative binomial regression. 2nd ed. Cambridge: Cambridge University Press; 2011.
[22] Wood SN. Low-rank scale-invariant smoothers for generalized additive mixed models. J R Stat Soc B 2006;68:95–114.
[23] Akaike H. A new look at the statistical model identification. IEEE Trans Autom Control 1974;19:716–23. https://doi.org/10.1109/TAC.1974.1100705.
[24] Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM. Mixed effects models and extensions in ecology with R. New York: Springer; 2009.
[25] Chitimia-Dobler L, Springer A, Lang D, Lindau A, Fachet K, Dobler G, et al. Molting incidents of Hyalomma spp. carrying human pathogens in Germany under different weather conditions. Parasit Vectors 2024;17:70. https://doi.org/10.1186/s13071-024-06175-y.
[26] Estrada-Peña A, Martínez-Avilés M, Muñoz-Reoyo MJ. A population model to describe the distribution and seasonal dynamics of the tick *Hyalomma marginatum* in the Mediterranean Basin. Transbound Emerg Dis 2011;58:213–23. https://doi.org/10.1111/j.1865-1682.2010.01198.x.
[27] Yılmaz S, İba Yılmaz S, Alay H, Koşan Z, Eren Z. Temporal tendency, seasonality and relationship with climatic factors of Crimean-Congo hemorrhagic fever cases (East of Turkey: 2012–2021). Heliyon 2023;9:e19593. https://doi.org/10.1016/j.heliyon.2023.e19593.
[28] Chanda MM, Kharkwal P, Dhuria M, Prajapathi A, Yogisharadhya R, Shome BR, et al. Quantifying the influence of climate, host and change in land-use patterns on occurrence of Crimean Congo Hemorrhagic Fever (CCHF) and development of spatial risk map for India. One Health 2023;17:100609. https://doi.org/10.1016/j.onehlt.2023.100609.
[29] Ansari H, Shahbaz B, Izadi S, Zeinali M, Tabatabaee SM, Mahmoodi M, et al. Crimean-Congo hemorrhagic fever and its relationship with climate factors in southeast Iran: a 13-year experience. J Infect Dev Ctries 2014;8:749–57. https://doi.org/10.3855/jidc.4020.
[30] Nama AH, Alwan IA, Pham QB. Climate change and future challenges to the sustainable management of the Iraqi marshlands. Environ Monit Assess 2023;196:35. https://doi.org/10.1007/s10661-023-12168-8.
[31] Al-mudaffar fawzi N, Goodwin KP, Mahdi BA, Stevens ML. Effects of Mesopotamian Marsh (Iraq) desiccation on the cultural knowledge and livelihood of Marsh Arab women. Ecosyst Health Sustain 2016;2:e01207. https://doi.org/10.1002/ehs2.1207.
[32] Xiao Y, Beier JC, Cantrell RS, Cosner C, DeAngelis DL, Ruan S. Modelling the effects of seasonality and socioeconomic impact on the transmission of Rift Valley fever virus. PLOS Negl Trop Dis 2015;9:e3388. https://doi.org/10.1371/journal.pntd.0003388.
[33] Khan K, Sears J, Hu VW, Brownstein JS, Hay S, Kossowsky D, et al. Potential for the international spread of Middle East respiratory syndrome in association with mass gatherings in Saudi Arabia. PLoS Curr 2013;5. https://doi.org/10.1371/currents.outbreaks.a7b70897ac2fa4f79b59f90d24c860b8. ecurrents.outbreaks.a7b70897ac2fa4f79b59f90d24c860b8.
[34] Copernicus Land Monitoring Service. Land cover, (raster 10 m), global, annual. version 1, ⟨https://land.copernicus.eu/en/products/global-dynamic-land-cover/land-cover-2020-raster-10-m-global-annual⟩; 2020 [accessed 18 December 2025].
[35] NASA Earth Data. NASADEM SRTM-only height and height precision mosaic global 1 arc second V001, ⟨https://www.earthdata.nasa.gov/data/catalog/lpcloud-nasadem-shhp-001⟩; 2020 [accessed 18 December 2025].
[36] ACLED. Conflict data. ⟨https://acleddata.com/conflict-data⟩; 2025 [accessed 18 December 2025].
[37] Copernicus climate data Store. ERA5 post-processed daily statistics on single levels from 1940 to present. Copernicus, ⟨https://cds.climate.copernicus.eu/datasets/derived-era5-single-levels-daily-statistics?tab=overview⟩; 2025 [accessed 18 December 2025].
[38] Copernicus climate data Store. Soil moisture gridded data from 1978 to present. Copernicus, ⟨https://cds.climate.copernicus.eu/datasets/satellite-soil-moisture?tab=overview⟩; 2025 [accessed 18 December 2025].
[39] University of East Anglia Climate Research Unit. Datasets, ⟨https://www.uea.ac.uk/groups-and-centres/climatic-research-unit/data⟩; 2025 [accessed 18 December 2025].
[40] al-Adha E. 10–13 Dhu aal-Hijjah) [in Arabic]. Law No. 12 of public holidays, Article 1 — Eid al-Fitr (1–3 Shawwal). Republic of Iraq: Ministry of Justice; 2024.
[41] United Nations Development Programme. Global Multidimensional Poverty Index (MPI), ⟨https://hdr.undp.org/content/2025-global-multidimensional-poverty-index-mpi#/indicies/MPI⟩; 2025 [accessed 18 December 2025].
[42] UN Migration International Organization for Migration. Datasets, ⟨https://dtm.iom.int/datasets⟩; 2025 [accessed 18 December 2025].