Sustainability Analytics and Modeling 6 (2026) 100052

Contents lists available at ScienceDirect

Sustainability Analytics and Modeling

journal homepage: www.journals.elsevier.com/sustainability-analytics-and-modeling

Providing context to conflict in the Lake Chad Basin: A quantitative assessment of socioeconomic and environmental influences

Patrick Sogno [a,*], Thorsten Hoeser [a], Penny Beames [b], Reeves Meli Fokeng [a], Claudia Kuenzer [a,c]

[a] Earth Observation Center (EOC) of the German Aerospace Center (DLR), Oberpfaffenhofen, 82234 Wessling, Germany
[b] Global Water Security Center, University of Alabama, Tuscaloosa, AL, United States
[c] Institute for Geography and Geology, University of Wuerzburg, 97074 Wuerzburg, Germany

ARTICLE INFO

Keywords:
Conflict
Africa
Lake Chad Basin
Sustainable development
Earth observation
Resource scarcity

ABSTRACT

Over the last two decades, the Lake Chad Basin (LCB) has experienced an increase in violent assaults by insurgency groups and militia clashes, with civilians often bearing the brunt. While numerous mappable factors are assumed to influence when and where conflicts arise, there is limited consensus in the choice of predictors and their meaningfulness for conflict research. This study explores the predictive relevance of socioeconomic and environmental variables for conflict occurrence in the LCB using a machine learning framework. We train and test a random forest regression model on conflict event data from the Armed Conflict Location & Event Data (2003–2020), combined with environmental indicators including air temperature, precipitation, evapotranspiration, soil moisture, surface water extent, and vegetation productivity, as well as socioeconomic variables such as population density, human development index, ethnic claims, and past conflict occurrence. The model achieves a Spearman rank correlation of $\rho = 0.45$ ($p < 0.001$) on the test data. Feature importance analysis indicates that anthropogenic and climate-related variables dominate the model's predictive performance. In particular, conflict dynamics exhibit strong path dependency, consistent with persistent structural conditions rather than isolated causal drivers. Seasonal variations in the predictive relevance of environmental factors align with key phases of the agricultural calendar, highlighting the importance of intra-annual dynamics. The presented approach identifies robust associations and temporal patterns consistent with theoretical accounts of conflict dynamics, while not permitting causal inference. While datasets capturing cooperation and de-escalatory interactions exist, their spatial and temporal resolution currently limit their integration into the present analysis. Incorporating such data at comparable resolution represents an important avenue for future research. Despite these limitations, the results suggest that considering environmental seasonality and socioeconomic context can contribute valuable insights for predictive modeling and inform research on conflict risk in vulnerable regions.

1. Introduction

Of the countries extending into the Lake Chad Basin (LCB), the respective regions situated in the basin are considered to be among the poorest and least developed. This is indicated by a lack of education, infrastructure, health care, and other services. The combination of lacking resources and economic perspective adds to the general instability of the LCB (Buhaug and von Uexkull, 2025). This is seen as one of the reasons why terrorist insurgency groups like Boko Haram have been able to recruit so successfully in the area (Riebe and Dressel, 2021). The current socio-political situation in the LCB is heavily shaped by the activities of these groups. The increasing insecurity in the area is also mirrored in the escalating number of conflicts and fatalities in the last two decades (Fig. 1). Over the entire observed timeframe (2003–2020), Armed Conflict Location & Event Data (ACLED) (Raleigh et al., 2023) has registered over 8800 conflicts in the LCB. In comparison to the rest of Africa, the region is among the most conflict-ridden, only being surpassed in the number of clashes by individual regions such as the coast of Somalia, Libya, and eastern Democratic Republic of the Congo. Many of those are attacks targeting civilians ($n = 4045$), actions by insurgency groups, namely Boko Haram and Islamic State West Africa (Lake Chad Faction), ($n = 2336$), or resource-related conflicts ($n = 684$) (Fig. 1(a)).

* Corresponding author.
E-mail address: patrick.sogno@dlr.de (P. Sogno).

https://doi.org/10.1016/j.samod.2026.100052
Received 27 September 2025; Received in revised form 17 January 2026; Accepted 9 February 2026
Available online 10 February 2026

Since 2003, which many consider as the founding year of Boko Haram, both the number of conflicts as well as the number of resulting fatalities have increased in all countries that are part of the LCB. With peaks in the years 2014 and 2015, conflicts in the basin have claimed over 10,000 lives annually (Fig. 1(*b*)). Over time, the spatial distribution of conflicts has changed. While conflicts in the early 2000s were concentrated in multiple hotspots (like Kano or Maiduguri in Nigeria, or Zalingei in Sudan), a more spread-out distribution can be seen for the recent years

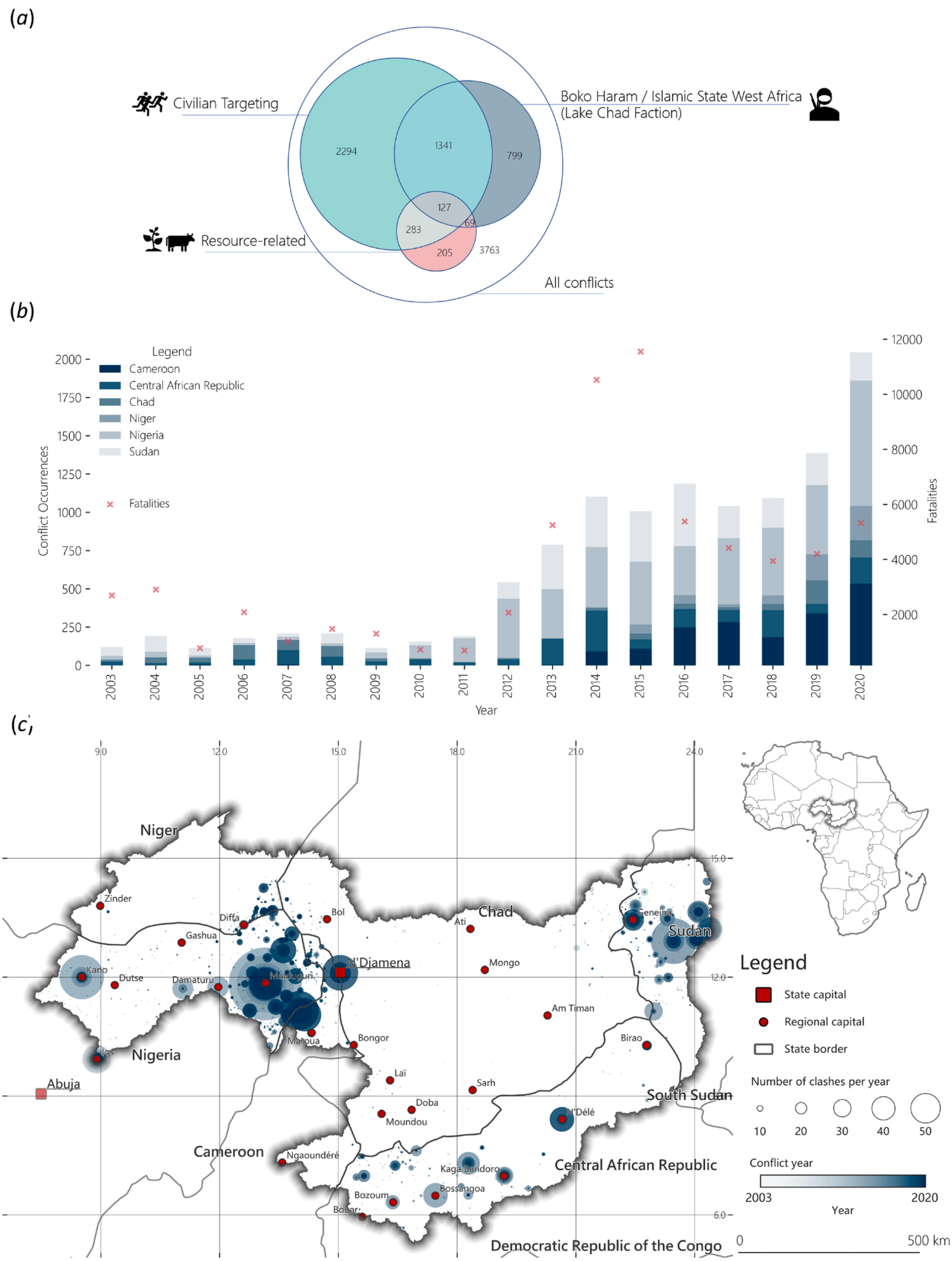

Fig. 1. Conflict in the Lake Chad Basin. Overview of a) important conflict subgroups, b) temporal development of conflict occurrences and fatalities, and c) spatial distribution of conflicts in the study area.

(Fig. 1(*c*)). At the same time, the original conflict hotspots still see significant conflict occurrences in recent years.

As Galli et al. (2022) point out, there is a high auto-correlation in conflict occurrences. As noted in ACLED, conflict actors are often motivated by grievances over previous confrontations. The tendency of conflict to perpetuate itself occurs through shifts in norms and a readiness to use violence, as well as through the availability of arms, and through indirect negative impacts on society. Specifically, negative impacts on the economy, governance, living standards, and increased vulnerability to natural hazards are identified as potential factors that may lock societies into a cycle of violence (Buhaug and von Uexkull, 2025).

The fear of attacks has led to the displacement of many rural actors in the basin, especially from the Nigerian side of the lake basin. Additionally, many pastoralists have changed their transhumance routes to avoid danger. Zieba et al. (2017) highlight that the Diamaré plain in the LCB has attracted hundreds of families and thousands of cattle in 2015 who were fleeing from violence in Nigerian border communities. They conclude that the increasing number of herders has led to an exacerbation of farmer-pastoralist conflicts (Zieba et al., 2017).

The link between conflict and the environment must be understood as one part of the complex dynamic interdependencies between natural resources, livelihoods, and socio-political processes (Galli et al., 2022). Explanations for conflicts lie not only with the natural environment or changes to it but also with the underlying vulnerabilities that characterize the region (Okpara et al., 2015). Whether conflict can be resolved before it turns into violent confrontations is also dependent on mitigating factors (Theisen et al., 2013), like functioning institutions that are able to de-escalate. When forums exist that allow for political participation and effective land dispute resolution, land disputes, which are among the main causes of conflict, can often be resolved peacefully (Deininger and Goyal, 2024). Traditionally, land disputes between small land owners can be resolved at the grassroots level by community leaders, while conflicts of large land owners ($\geq$ 5 ha) are settled at a sub-prefect level or by local law enforcement. However, in many places, established local institutions have been abolished without providing functioning substitutes (Zieba et al., 2017). This limits the effectiveness of customary usage claims. Legal documents that prove land ownership or usage rights do not exist in most cases. Fights can break out if land rights are perceived as contestable. This may be the case, especially if there is a perceived contradiction between customary and legal rulings of land ownership and usage rights (Deininger and Goyal, 2024). As Brottem (2016) point out, when such latent conflicts erupt in violent confrontations, it is often because of an underlying feeling of injustice or lack of respect. The actions taken by decision-makers in the region concentrate on confining the areas of political instability with military means. Meanwhile, increasing population pressure, and with it increasing overexploitation of the available natural resources, further deteriorate the situation in the area. Weak governance at the state level, symptomatic through the abolishment of established local institutions without providing functioning substitutes, further leads to available resources being overused (Zieba et al., 2017).

Natural resource scarcity, and especially water scarcity (Nkiaka et al., 2024), is a major concern in the LCB. Of the >47 million people living in the area, nearly a third (ca. 15 million people) is directly dependent on Lake Chad itself (Hassan et al., 2021). A lot of the resource scarcity in the region can be linked to rapid population growth, and subsequently, more people competing for the limited amount of available resources (Riebe and Dressel, 2021). While conflict can arise when scarce resources are further limited to a point where competition leads to a complete withdrawal of access to critical resources (Deininger and Goyal, 2024), resource scarcity does not automatically lead to conflict. Indeed, the majority of interactions on scarce water resources is cooperational (Bernauer and Böhmelt, 2020; Cong et al., 2025; Kåresdotter et al., 2025; Petersen-Perlman et al., 2017; Wolf et al., 2005). That being said, if conflict arises over scarce resources, the dispute can often be attributed to one of three issues: quantity, quality, and timing (Wolf et al., 2005). As the IPCC Report of 2023 outlines, mounting evidence suggests that much of the decreasing productivity in African agriculture comes from ongoing climate change. West Africa in particular faces increasing hot extremes since the 1950s. With that comes an increase in agricultural and ecological droughts (Calvin et al., 2023). As a consequence, livelihood strategies change and confrontations become more frequent, especially when goals are perceived as mutually exclusive, as often is the case for farmer-herder conflicts (Brottem, 2016).

Rather than investing time and energy into accessing water at intermediate stopping points where grazing resources are inadequate, transhumant herders choose to press on to where reliable water and feeding grounds can be expected. This can mean that zones that are off-limits for the herders are traversed anyway or that herders arrive early at their seasonal destinations, leading to resource competition for water and crop damage. Especially shortly after the rainy season has ended when farmers try to harvest their crops without damage by livestock and herders try to provide water and pasture for their animals, the two strategies become incompatible and livelihoods clash. The empirical relationship between climate and conflict is not conclusively discussed. Valid points are raised, especially critiquing the interpretation of found correlations between environmental variability or change and conflict as causal relationships (Selby, 2014). Often, quantitative studies base this notion on the idea that conflict is a consequence of competition for scarce resources. However, Theisen et al. (2013), for example, point out that pastoralist violence is seemingly more driven by tactical concerns than by resource-based grievances. One of the main points that is held against quantitative studies is that different analyses come to different conclusions on environmental factors that supposedly drive conflicts and whether they play a supporting or mitigating role (Selby, 2014; Theisen et al., 2013).

Amid the ongoing debate, calls are made to broaden the analyzed variables beyond rainfall and temperature and include factors that directly affect the livelihoods of the stakeholders (Brottem, 2016).

There is ongoing debate regarding the development of the environmental situation in the LCB. According to Pham-Duc et al. (2020), Lake Chad's surface water extent reduced slightly over the last two decades. This decrease is mostly confined to the northern pool and is due to increasing evapotranspiration (ET) and vegetation cover as well as a decrease in discharge from the Komadugu-Yobe, which is the western tributary subbasin (Pham-Duc et al., 2020; Sogno et al., 2024). These results contrast the findings of Adeyeri et al. (2019) who report an overall increase in temperature, precipitation, and river discharge in the Komadugu-Yobe subbasin. They conclude that increasing precipitation in the Sahel could revive the wetlands in the subbasin, but at the same time, excess water from heavy precipitation events could lead to flooding. Both of these dynamics demand for adequate measures to be put into place. Lemoalle et al. (2012) find that Lake Chad behaves as an amplifier of rainfall variability. They propose that its northern pools may run dry entirely for several years if severe and repeated droughts were to happen. The southern pool's extent is seen as stable, even slightly increasing as a consequence of stable local rainfall and increasing discharge from the Chari-Logone subbasin, which drains into Lake Chad from the south (Pham-Duc et al., 2020).

Identifying responses of Lake Chad following vegetation and rainfall fluctuations, Gbetkom et al. (2023) see increasing rainfall for 2000–2020 over the entire LCB, an increase in vegetation cover across the lake, and a positive correlation between rainfall, vegetation cover, and water extent and height. This assessment is at odds with that of Bennour et al. (2023), who report that the Sahel regreening did decrease surface runoff. In their eyes, the man-made expansion of cropland (seen as a consequence of the population increase) at the cost of forested areas has played an important role in limiting the decrease in surface runoff in the LCB. This aligns with the conclusion drawn by Li et al. (2007), who characterize the relationship between unsustainable land use, land cover change, and streamflow increase in West Africa via a hydrological model. They conclude that vegetation thinning from deforestation and overgrazing in savanna or grassland environments can lead to an increased water yield if it surpasses a certain threshold. Above this threshold, the reduction in vegetation activity leads to a direct reduction in net transpiration and increased surface runoff. The decreasing impact of human intervention on vegetation productivity in the area is also seen by Fu et al. (2023).

As outlined, several complex problems fuel the increasingly intense conflicts in the LCB. The ongoing decline of the basin's utility due to climate change and overexploitation of available resources, combined with increasing population pressure, poor governance, and additional insecurity introduced by insurgency groups characterize this humanitarian crisis (Hassan et al., 2021). A long-term perspective needs to be applied in the management of resources around Lake Chad. As Zieba

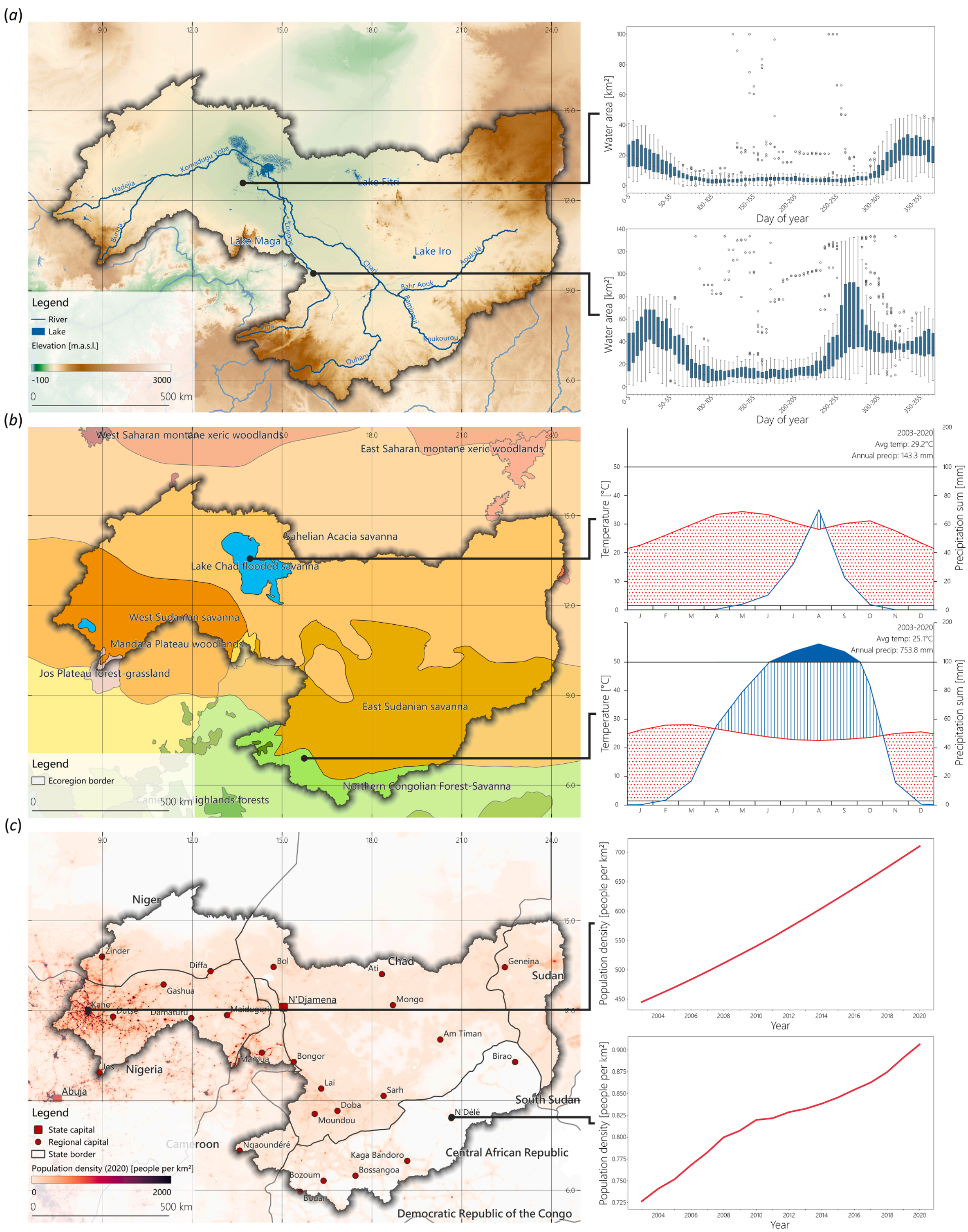

Fig. 2. Overview of the study area with a focus on (*a*) topography and surface water hydrology, (*b*) ecoregions and climate, and (*c*) demography and infrastructure.

et al. (2017) highlight, the groundwork for this has already been laid – in the form of the Lake Chad Basin Commission – but limited human, material, and financial resources constrain its effectiveness. They suggest that policy action may involve monitoring and assessment of the environmental resources of the lake, as well as the sources and dynamics of pressures that contribute to changes in these resources. This would directly address the problem that the capacity for climate adaptation in Africa is severely limited due to a lack of data, monitoring, and evaluation of climate change adaptation, as Calvin et al. (2023) outline. As it stands, environmental shifts due to climate change and usage change and a lack of institutional capacity to manage these shifts may lead to future conflict (Petersen-Perlman et al., 2017). While cooperative and de-escalatory interactions play a crucial role in resource (and specifically water) governance, this research's main objective is to provide insights into the factors statistically associated with conflict occurrence as observable manifestations of escalation. To identify these factors and provide these insights as contribution to the scientific discussion of resource related conflicts, two tasks are defined:

1) Consider a wide array of factors from multiple spheres, including intra-annual dynamics of environmental features that affect the livelihoods in the LCB, and predict conflict occurrences.
2) Untangle the socioeconomic and ecological factors statistically associated with conflicts based on their predictive power.

2. Materials and methods

2.1. Study area

The LCB is an endorheic basin, meaning all surface water flows lead into a terminal lake, the Lake Chad. There are various definitions of the basin, Neukum et al. (2023) for instance include a large swath of desert in Chad and Niger, which may be linked to the rest of the basin via groundwater flows but is largely unproductive concerning surface water (Mohamed et al., 2023). In this analysis, we work with the productive basin as proposed in the HydroBASINS dataset (Lehner and Grill, 2013). Using this definition, the basin covers 1.124.332 km^2. Being situated between longitudes 7.3° and 24.4°E and 5.3° and latitudes 16.2°N, the basin is shared amongst the countries Niger, Chad, Sudan, the Central African Republic, Cameroon, and Nigeria (Fig. 2). Its highest elevation reaches just above 3000 m in the Sudanese Marra Mountains in the basin's northeastern part of the basin. To the west, the basin is bordered by the Nigerian Jos Plateau. In the south, it is bounded by the Adamawa Plateau in Cameroon and the Central African Republic. To the southeast, the basin extends up to the Bongo Massif in the Central African Republic. About 90 % of Lake Chad's water stems from Chari-Logone discharge. The impact of climate variability on runoff is dominated by precipitation. Water loss due to human intervention has increased over time and the Chari-Logone runoff is mainly driven by human activities rather than climate variability (Zhu et al., 2019). Lake Chad itself has shrunk significantly between the early 1970s and the early 1990s. Over the last 30 years, the lake has seen a slow recovery, which is mainly linked to increasing precipitation rates in the basin (Pham-Duc et al., 2020), but may also be partially influenced by human impact in the form of deforestation and overgrazing, which may lead to decreasing transpiration rates and more runoff into the lake (Li et al., 2007). The subsurface contributes to ~70 % of the total water storage (TWS) at the lake, meaning most water is stored in soil moisture and groundwater (Pham-Duc et al., 2020). In part, the groundwater storage situation has improved between 2002 and 2021 due to increasing precipitation rates (Mohamed et al., 2023). However, there are reports of decreasing groundwater levels and drying up wells and boreholes as well, resulting from droughts and decreasing surface water productivity (GIZ, 2016). Adding to that, around Lake Chad the groundwater is partially brackish (Luxereau et al., 2012), limiting its use as drinking water and for irrigation. The LCB's subbasins differ substantially in surface water productivity, as an exemplary comparison of surface water areas for the estuary region south of Lake Chad and the middle reaches of the Logone River demonstrates (Fig. 2(*a*)). Apart from absolute differences in surface water area, Fig. 2(*a*) demonstrates a difference in the beginning and end of the flooding season for different subbasins in the LCB. The seasonal patterns of surface water area increases follow the seasonal pattern of precipitation. Depending on where they are situated in the basin, the water areas of lakes and rivers react to precipitation events with lags of up to five months (Fokeng et al., 2024).

The LCB almost exclusively falls into three Köppen-Geiger climates as illustrated by Beck et al. (2023). While the basin's south has a tropical savannah climate (Aw), the mid-latitudes have a hot and arid steppe climate (BSh). The northern part of the basin has a hot and arid desert climate (BWh). Precipitation sums range from under 80 mm per year to over 1600 mm per year with a clear North-South gradient. Mean air temperature ranges from ca. 23 °C to ca. 30 °C with a strong orographic dependency. Low-lying areas around Lake Chad are generally warmer than areas with higher elevations in the south and east of the basin. The seasonal temperature gradient is less pronounced the closer an area is to the equator. Comparing the temperatures of the Lake Chad Flooded Savanna to the Northern Congolian Forest Savanna for the timeframe from 2003 to 2020 visualizes this (Fig. 2(*b*)). In the former, the average temperature is 29.2 °C, with a minimum of 20.5 °C and a maximum of 35.5 °C. In the latter, the average temperature is 25.1 °C, with a maximum of 29.3 °C and a minimum of 22.2 °C. Maximum and minimum temperatures occur at different times for different areas in the LCB and are dependent on the changes in solar irradiation throughout the year (Fig. 2(*b*)). Due to the strong precipitation gradient, the onset and length of wet and dry seasons differ. In the northern part of the basin, the wet season only covers a few weeks between June and September. In the southern part of the basin, the wet season extends from April to October. Accordingly, the basin mostly consists of different types of savanna. While the basin's north is dominated by the Sahelian Acacia savanna, its middle reaches primarily consist of different types of Sudanian savanna. Different ecoregions come with different seasonal temperature and precipitation patterns, as well as hydrological and ecological differences, as Fig. 2(*b*) shows for the examples of Lake Chad flooded savanna versus Northern Congolian Forest-savanna.

As visible in Fig. 2(*c*), many of the major cities in the LCB are situated along rivers or lakes, ensuring reliable surface water access. Outgoing from these major cities, the concentration of urbanized areas, which coincides with high population density, follows road and rail infrastructure. Over the last two decades, the population density has increased in almost all parts of the LCB. However, there are stark contrasts in population density for different regions in the LCB. A high population concentration is visible, especially in the federal state of Kano in Nigeria. Population density increase in Kano is almost linear. Comparing this to one of the least densely populated regions in the LCB, Bamingui-Bangoran in the Central African Republic, we see that also here population density increases. However, the amount of people per area differs vastly between the two regions (Fig. 2(*c*)).

A majority of the population in the LCB depends on a subsistence economy (Okpara et al., 2016). As such, much of the basin is used for agriculture. Crops can generally be planted as unirrigated crops in alignment with the wet season, or as off-season crops, which often depend on irrigation. Major crops in the basin are maize, millet, rice, sorghum, and wheat (Olowoyeye and Kanwar, 2023; Reynolds et al., 2015). In areas with surface water access, particularly around Lake Chad, fishing is a prevalent occupation for the local population. In addition, the basin's pastures are used for livestock herding. Particularly sheep, goats, and cattle are commonly herded (Olowoyeye and Kanwar, 2023). Herds often move in a seasonal pattern, following the availability of water and vegetation for grazing (Zieba et al., 2017).

2.2. Datasets

All features included in the modeling effort can be categorized into one of four spheres of influence: the anthroposphere, climate, hydrosphere, and biosphere. In Section 2.2.1, we will present the environmental factors that represent the climate, hydrosphere, and biosphere. The anthroposphere is represented through socioeconomic factors, which are presented in Section 2.2.2, as well as previous conflict occurrences, covered in Section 2.2.3.

2.2.1. Environmental factors

All environmental factors included in the analysis are based on Earth observation (EO) datasets. The DLR Global WaterPack (GWP) is a dynamic surface water extent product based on MODIS Terra and Aqua observations. As of now, it covers the timeframe from 2003 to 2024, with data for 2025 and consecutive years becoming operational at the end of each year. Within this study, we observe the timeframe from 2003 up to the middle of 2020 (last included date is July 31, 2020). This timeframe was chosen as it is the last date for which all utilized datasets are available. The GWP has a spatial resolution of 250 m. It provides daily information on surface water coverage per pixel (Klein et al., 2024). The dataset is available at https://geoservice.dlr.de/web/datasets/globalwaterpack.

ERA5-Land is based on the ERA5 reanalysis but has been improved to make it more accurate for land applications. The dataset's spatial resolution is 9 km and its temporal resolution is 1 hour (Muñoz-Sabater et al., 2021). The dataset is available at https://doi.org/10.24381/cds.e2161bac. Depending on the parameter, the measurement unit varies. For temperature (T) it is kelvin. For potential evaporation (PET), total evaporation (ET), and total precipitation (P) it is meter of water equivalent. PET is open water evaporation (Pan evaporation), while ET is all evaporation plus an approximation of transpiration from vegetation. In ERA5-Land, downward fluxes are positive, which means that negative values indicate evaporation and positive values indicate condensation. P covers liquid and solid water but excludes fog, dew, and P that evaporates in the atmosphere before it arrives at the surface of the Earth.

Gross Primary Productivity (GPP) is the amount of CO_2 assimilated through photosynthesis. The measurement unit is kg C/m^2. We use the GPP dataset by Joiner and Yoshida (2020) to approximate vegetation activity and biomass. This dataset provides global gridded GPP estimates based on MODIS Terra and Aqua and FLUXNET 2015 eddy covariance tower sites. This dataset's spatial resolution is 5 km, its nominal temporal resolution is 1 day. This dataset has a temporal extent from March 2000 to the end of July 2020. The data is available at https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1835.

Soil moisture (SM) is the ratio of the volume of water per volume of soil (cm^3/cm^3). The utilized Global Surface Soil Moisture (GSSM1km) dataset provides surface soil moisture data (0–5 cm) on a daily temporal scale at 1 km spatial resolution for the years 2000 to 2020 (Han et al., 2023). The dataset is available at https://doi.org/10.6084/m9.figshare.21806457.v1.

2.2.2. Socioeconomic factors

We incorporate multiple socioeconomic factors into the regression model. This includes a yearly time series of population density data using the WorldPop Dataset (University of Southampton et al., 2020). This data is available at a spatial resolution of 1 km. Further, we include the Subnational Human Development Index (SHDI) in our analysis. The data is available in a yearly format for the years 1990 to 2022 and was retrieved from the Subnational HDI Database of the Global Data Lab, https://globaldatalab.org/shdi/. The SHDI provides information on key dimensions of the socioeconomic sphere (health, education, and income) on a subnational scale.

Lastly, previous studies have outlined the importance of including ethnic settlement outlines in conflict modeling efforts, as conflict occurrences seem to intensify along ethnic lines (e.g., (Galli et al., 2022; Ge et al., 2022; von Uexkull et al., 2016)). We therefore include a measure of competing ethnic claims per region into the analysis based on ethnic settlement outlines as defined in the latest version of the Geo-referencing Ethnic Power Relations dataset (GeoEPR, 2021) (Vogt et al., 2015).

2.2.3. Conflict occurrences

Conflict, and especially violent conflict, is only one of many interaction types between actors. Within the context of resource scarcity and foremost water-related confrontations, positive interactions (i.e., cooperation) or non-violent disputes are far more frequent (Kåresdotter et al., 2022; Kåresdotter et al., 2025; Petersen-Perlman et al., 2017). Ideally, both positive and negative interactions would be portrayed as different outcomes of the same investigated variable. However, datasets at the spatial and temporal resolution needed for local investigations are lacking (Bernauer and Böhmelt, 2020). Recently, a global dataset that features both cooperation and conflict in the context of water-related interactions has been published (Kåresdotter et al., 2022). This dataset covers interactions from 1951 until 2019, and provides rich, spatially and temporally aggregated information. For the modeling approach of this study, spatially and temporally fine-grained information for individual regions in the LCB is of major importance. Thus, conflict data from January 1, 2003, to December 31, 2020, has been gathered via the ACLED database (Raleigh et al., 2023), matching with our requirements. We consider all conflicts in the LCB within the investigated timeframe. This is done to maximize the number of conflict occurrences that are considered in the model, since initial tests showed that subsetting conflict occurrences by type or actors reduces the number of training and test cases, leading to lower model performance and less interpretable feature importance.

2.2.4. Static datasets

Our study area is the active part of the LCB, meaning the part that drains into Lake Chad via surface water inflow, according to the HydroBASINS dataset (Lehner and Grill, 2013). As proposed by Galli et al. (2022), thematically meaningful polygons are used in this investigation instead of a rectangular grid. These polygons are based on the intersection of subbasins, administrative borders, and ecoregions. Moving forward, we refer to these polygons as small hydrological units (SHUs). The subbasins we consider in the creation of the SHUs are also provided in the HydroBASINS dataset and reflect a Pfafstetter level of 7. The administrative boundaries included in the SHUs are on the federal state, region, or province level (admin level 1). We utilize the GADM dataset, version 4.1 (https://gadm.org/data.html). The ecoregions included in the SHUs are defined by Dinerstein et al. (2017).

To discern between irrigated and non-irrigated GPP, we mask with a global irrigated agricultural fields dataset (Meier et al., 2018) that is available here: https://doi.org/10.1594/PANGAEA.884744.

2.3. Methods

The workflow of our analysis (Fig. 3) is sectioned into five main steps. In the following sections, these main steps are described in detail. The backbone of our analysis is a random forest (RF) regression model that predicts conflict occurrences in the LCB. The main objective is, however, the assessment of the predictive relevance of individual features within this machine learning framework. This allows us to explore which features are used by the model and how strongly individual features contribute to the model's predictive performance. Feature importance is thus interpreted as a model-internal measure of relevance rather than evidence of causal influence. Based on a set of training data, the model learns the connection between given features (= independent variables) and labels (= dependent variable). It then predicts labels based on a secondary set of features. These predictions are compared to test labels to assess model performance. In our case, the features

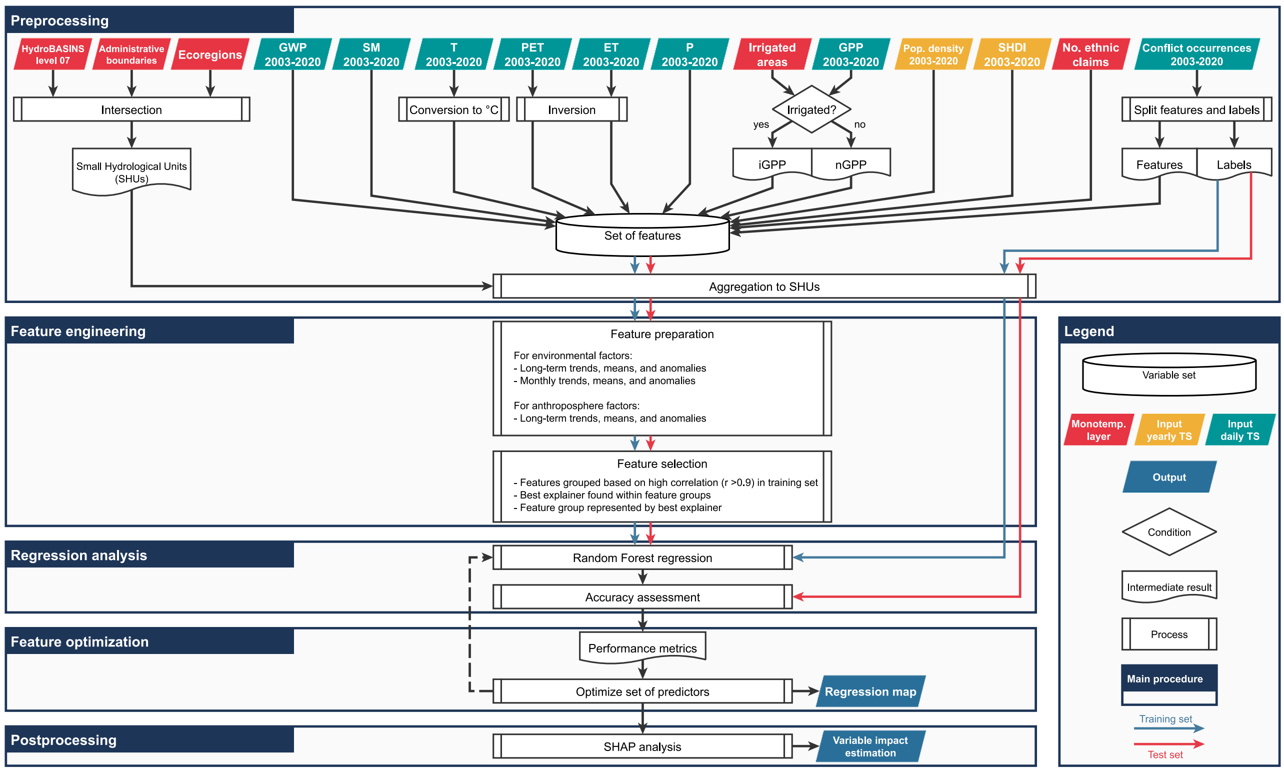

Fig. 3. Processing workflow for the regression analysis. Color-coding of the input datasets indicates their temporal resolution. This is important for the feature preparation step in the feature engineering section. For all time series the long-term trend, mean, and anomaly are calculated. For daily time series of environmental factors, monthly trends, means, and anomalies are calculated as well. Training set and test set are indicated as blue and red arrows, respectively. Postprocessing.

characterize the presented environmental and socioeconomic factors as well as previous conflicts. The training and test labels are conflict occurrences that happened in the training and test year, respectively. The split of conflict occurrences into features and labels for both training and testing the model is detailed in the following section.

2.3.1. Preprocessing

HydroBASINS, administrative boundaries, and ecoregions are intersected to obtain SHUs. All environmental factors for the regression analysis are preprocessed to a set of daily observations. T is converted from Kelvin to °C, while PET and ET are inverted from their ERA5-Land-based format to show higher (potential) evaporation as higher values. GPP is subdivided into GPP in irrigated areas (iGPP), and GPP in non-irrigated areas (nGPP) based on the global irrigated areas mask presented in Section 2.2.4. Of the socioeconomic factors, population density, and SHDI are available as yearly time series. The number of ethnic claims is a static dataset.

Conflict occurrences are split into labels and features. The labels are split by date into a training and a test set. The set of training labels includes all conflicts that happened between January 1, 2019, and December 31, 2019. The set of test labels includes all conflicts that happened between January 1, 2020, and December 31, 2020. An autocorrelation of conflict occurrences has been documented before (Galli et al., 2022) and is supported by research outlining the pathways of such positive feedback loops (Kamta et al., 2020). We therefore included conflict occurrence history into the set of features. For the creation of training features, we include all conflicts that happened between January 1, 2003, and December 31, 2018. For the creation of test features, we include all conflicts that happened between January 1, 2004, and December 31, 2019. All features are aggregated to SHUs before further feature engineering is started.

2.3.2. Feature engineering

The set of features is split by date into a training and a test set. The temporal extent of conflict occurrence feature time series is one year shorter than the temporal extent of all other features, to keep label and feature information separate. For all other features, the training set time series extend from January 1, 2003, to December 31, 2019, and the test set time series extend from January 1, 2004, to December 31, 2020. These time series are processed in the feature preparation step to extract long-term means, trends, and anomaly values, expressing the difference between the last year's value and the long-term mean. For time series with a daily temporal resolution, we additionally process mean values, trends, and anomaly values for each month. The following equations define the underlying calculations done for each considered factor and for each SHU to arrive at the respective mean, trend, and anomaly values.

$$\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t \tag{1}$$

Eq. (1) shows the calculation of $\bar{x}$, which is the long-term mean. Here, n is the number of observations in the time series, and x_t is the value of x at time t.

$$\beta = \frac{\sum_{t=1}^{n}(t-\bar{t})(x_t-\bar{x})}{\sum_{t=1}^{n}(t-\bar{t})^2} \tag{2}$$

Eq. (2) shows the calculation of the long-term linear trend (β). Again, n is the number of observations in the time series, of which t is a specific timestep, and x_t is the value of x at that timestep. Further, $\bar{x}$ is the mean value of x, and $\bar{t}$ is the mean value of t.

$$a = x_n - \bar{x} \tag{3}$$

Lastly, Eq. (3) shows the calculation of the anomaly values (a), which is the difference between x at the last observation (x_n) and the mean value of x, which is $\bar{x}$.

At this point, the training and the test set each hold $n = 322$ features. Some of those features are highly correlated and do not offer much additional informational value for the regressor. Therefore, to reduce the computational effort, we implemented a feature selection step. In this step, a correlation matrix is calculated for all features in the training set. Features with very high correlation coefficients ($r > 0.9$) are grouped. Within groups, the best explainer is selected, which is the feature with an $r > 0.9$ for the most features in the group. This feature is kept as a group representative, all other features are omitted from the analysis as they provide largely redundant information within the given feature set. This step reduces the number of features for the training set to $n = 207$. The feature selection for the training set is imposed onto the test set to ensure the same features are fed to the model during training and prediction.

2.3.3. Regression analysis

We employ a RF regressor (Breiman, 2001) with 200 trees and a maximum tree depth of 6. This parameter choice was made after initial tests, to optimize the number of trees and tree depth based on model performance. The model is fit using the training labels and all remaining $n = 207$ training features. In a first step, the RF model is trained on $n = 207$ features. Based on the premises in the feature engineering step, the employed feature set includes a broad set of environmental and socio-economic factors after removing highly redundant features. This is the baseline model. The model's predicting capability is tested against the test dataset. Performance is quantified using the root mean squared error (RMSE) of predicted conflicts on the conflicts in the test dataset. The RMSE has the same unit as the target variable (i.e., conflicts per SHU). Next, the list of features to be included in the model is optimized.

2.3.4. Feature optimization

Feature optimization is used here to explore whether comparable predictive performance can be reached with a smaller feature set. Starting with a list of all features, we iteratively leave individual features out of the model in a recursive feature elimination approach. The new model's predicting capability is tested and the RMSE of the new model with one feature left out is compared to the baseline model. If the new model performs better than or equally well as the baseline (meaning $RMSE_{baseline} \geq RMSE_{new}$), the dropped feature did not enhance model performance and is disregarded. The new model becomes the new baseline model, and the iterative process continues with the next feature. If the new model performs worse than the baseline (meaning $RMSE_{baseline} < RMSE_{new}$), the dropped feature did enhance model performance. It is added back to the model and cannot be dropped again within this elimination cycle. At the end of the elimination cycle, the RF model is trained using the remaining features.

From the list of dropped features, individual features are added back to the model if they enhance model performance in a second iterative process. This forward sequential feature selection works in the opposite direction of the feature elimination cycle. The feature optimization reduces and expands the feature list, alternating between the two cycles until the optimal set of features is found. To ensure reproducibility and to prevent the optimization routine from alternating between multiple optimal states, the order of the feature list is kept the same when adding features back in the expansion cycle.

We identify feature contributions to the model based on Shapely Additive Explanations (SHAP) (Lundberg and Lee, 2017). As the employed model relies on a tree-based regressor, we identify the features' SHAP values using a SHAP tree explainer (Lundberg et al., 2020). Apart from the directional SHAP values, we also analyze global feature importance, which is calculated as the global mean of absolute SHAP values of a feature.

3. Results

The main objective of this study is to provide insights into the contribution of factors to conflict occurrence in the LCB. We model and predict conflict occurrences in the LCB with the clear objective to investigate the importance of individual features used by the model and gain insights into what parameters are relevant for conflict occurrences based on that. Therefore, this study's modeling effort is a necessary step, but not the final result. As the aspiration is not to produce the perfect model to predict future conflicts, the performance of the final model needs to be considered for the later feature importance investigation but is not paramount.

3.1. Spatial patterns of conflict occurrence

Conflicts in the LCB are generally concentrated around Lake Chad, to its south, and in the far east of the basin in Sudan. The spatial patterns of conflicts used for training (2018–2019) and conflicts used for testing model performance (2019–2020) are similar. However, an increase in conflict occurrences can be seen for multiple SHUs. A concentration of conflict occurrences can be seen especially around Mora in Cameroon and Zalingei in Sudan (Fig. 4). This concentration exists both in the training and the test dataset. However, the number of conflict occurrences in the test dataset far exceeds that of the training dataset. In the training dataset, the area with the most conflicts lies in the border region between Nigeria and Cameroon. Here, 95 conflicts were reported between January 1, 2019, and December 31, 2019. In the testing dataset, the same area experienced the highest number of conflicts, with 132 documented incidents between January 1, 2020, and December 31, 2020. Model predictions for January 1, 2020, to December 31, 2020, largely mirror the spatial distribution of actual conflicts. The model's maximum estimate is about 59 conflicts in the mentioned border area between Nigeria and Chad.

Generally, model estimations are quite conservative, leading to underestimations, especially for outlier values in which conflict occurrence has increased extremely between the training year and the testing year. That being said, escalating numbers in conflict occurrences were correctly predicted for several SHUs, especially south of Lake Chad and in the border region between Chad and Sudan.

Model performance is visualized in Fig. 5. Using the final feature list, the RMSE is 3.83. The model shows a moderate but highly significant correlation to the test data (Fig. 5(*ai*)). As indicated, the model underestimates extreme outlier cases. However, for over 80 % of SHUs, the predicted number of conflict occurrences and the actual number of conflict occurrences in the test set are aligned (Fig. 5 (*aii*)). For 95 % of all SHUs, the absolute difference between the predicted and actual number of conflicts is ≤ 3.28. As visible in Fig. 5 (*aii*) and Fig. 5 (*aiii*), the model performance is affected negatively by the few extreme outlier values in the test set. The impact of outliers in the test dataset is visualized through a comparison of model performance for the full test set versus for a subset that includes values up to the 99th percentile in the test set (Fig. 5(*b*)). Not accounting for the most extreme outliers in the test set, the RMSE is at 2.11, correlation with the test set is still moderate, but highly significant. The linear regression between actual and predicted conflicts still indicates a general underestimation of modeled values, but less severe than shown in Fig. 5(*ai*). The model error in terms of the difference between predicted and actual conflicts is within the range of -0.5 to $+0.5$ for the majority of SHUs (Fig. 5 (*aii*)). This could be regarded as a rounding error since the values in the test set are integers and the predicted values are floating-point numbers. Looking at the absolute difference between predicted and actual conflicts, we see that model predictions are off by less than 0.25 from the actual number of conflicts for most SHUs.

In summary, we achieved a minor performance gain while decreasing model complexity through automatic feature optimization. Model performance is impacted by outlier cases, in which the number of conflicts has increased extremely between the training and testing year. Disregarding the top 1 % of conflict incidences, the model achieves an RMSE of 2.11.

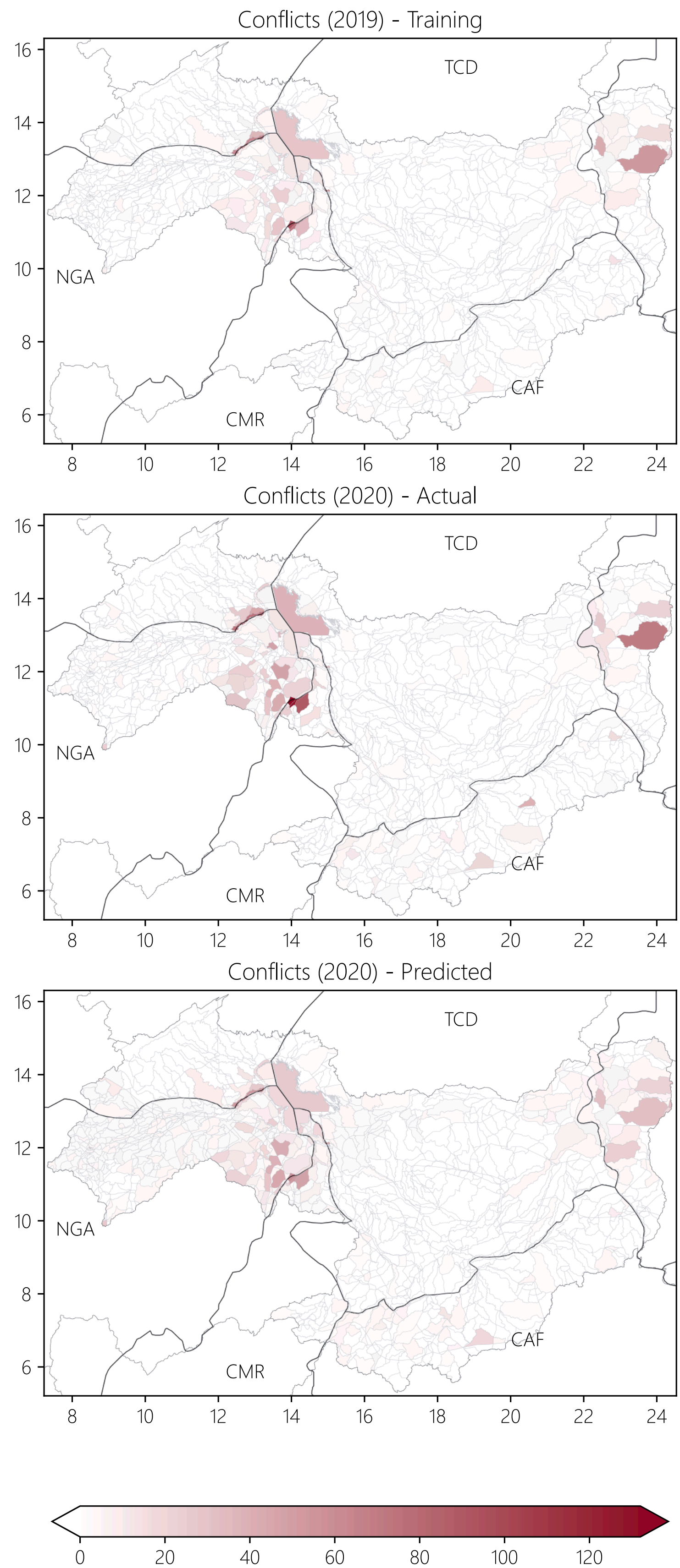

Fig. 4. Training and test (=actual) datasets versus model outcome (=predicted).

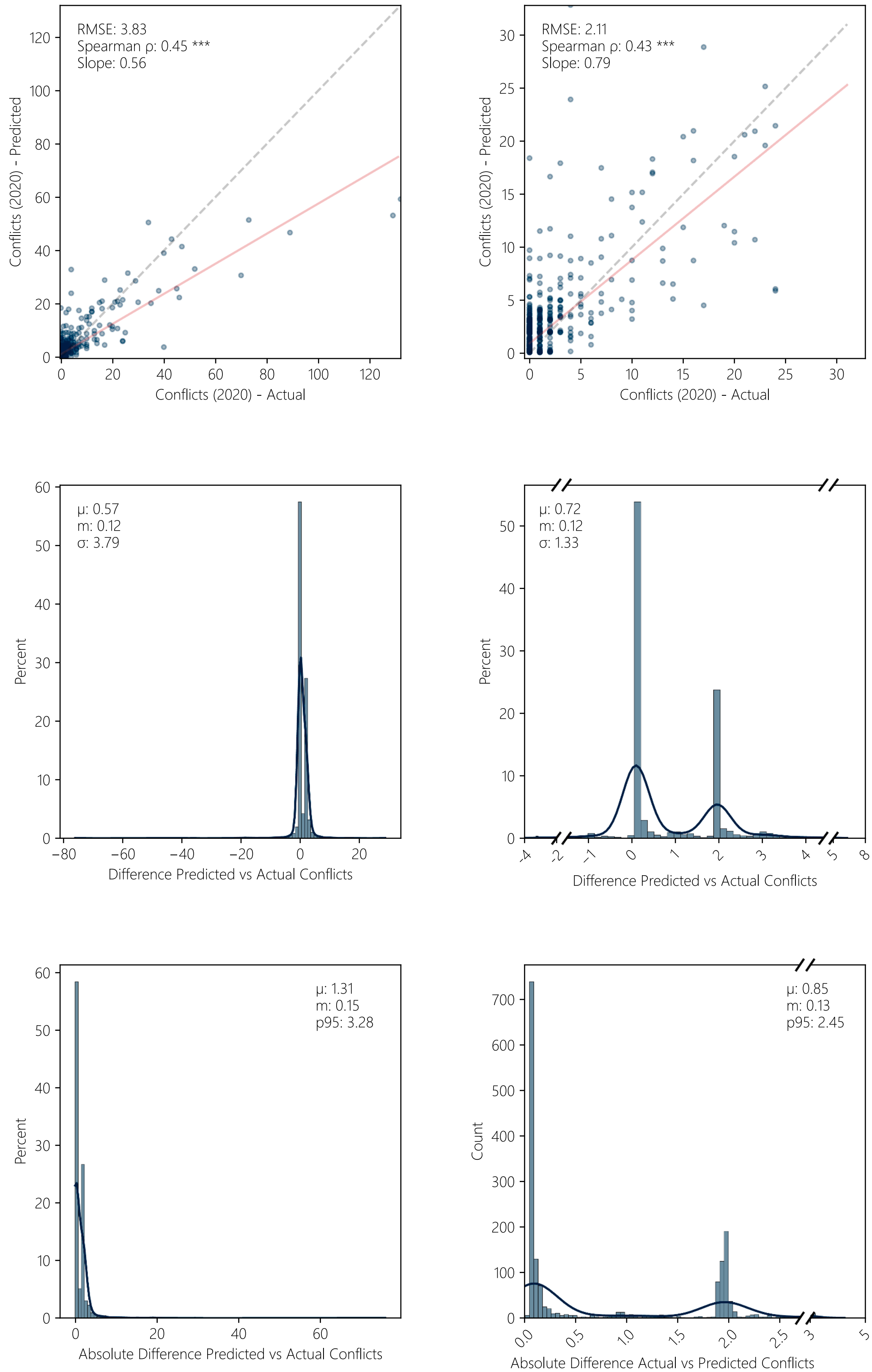

Fig. 5. Model performance indicators. (*a*) shows model performance as predicted versus actual conflicts for all SHUs, (*b*) is cropped to the 99th percentile to show model performance without the most extreme outliers. (*ai*) and (*bi*) show the distribution of predicted versus actual conflicts. Model performance is expressed through the RMSE, Spearman rho (ρ) and significance ('***' signifying a p-value of < 0.001), and the linear regression between actual and predicted values. To identify a systematic offset between actual and predicted conflicts, the regression line for actual conflicts ∼ predicted conflicts is shown in red, while the line of ideal fit is shown in gray. (*aii*) and (*bii*) show the difference between actual and predicted conflicts, including the mean difference (μ), mode of differences (m), and the standard deviation of differences (σ). (*aiii*) and (*biii*) show the absolute difference between predicted and actual conflicts. Along with the mean (μ) and mode (m) of absolute differences, the 95th percentile value is given, signifying the largest absolute difference between predicted and actual conflicts for non-outlier values.

3.2. Feature contribution

The contribution of individual features to the model is analyzed both as SHAP values and as global means of absolute SHAP values to signify feature importance. Fig. 6 gives an impression of the ten most dominant features contributing to the model. The features listed in this figure are sorted by importance for the model (meaning their weight in the model prediction). The colorscale of the individual points in the plot represents the feature value, while the point positioning along the x-axis shows the SHAP value. This value signifies the contribution of individual observations of features to the model prediction. Positive SHAP values increase the predicted value, while negative SHAP values decrease the predicted value.

The feature with the highest predictive power is the mean number of conflicts for the years before what the model sees as train or test labels (Fig. 6). The direction of feature values follows the direction of SHAP values. This means a low mean number of conflicts in previous years is a strong indicator of a low number of conflicts in the future. Conversely, a high number of previous conflicts leads the model to predict a high conflict occurrence. Consistent with previous work (Buhaug and von Uexkull, 2025; Galli et al., 2022; Kamta et al., 2020), previous conflicts show the strongest statistical impact in the model. At the same time, environmental factors dominate the remaining top-ranked features, accounting for eight of the ten most influential factors. This motivates a more detailed investigation of their spatial and temporal roles in conflict dynamics, providing a basis for hypothesis generation and future causal analysis. Apart from previous conflicts, particularly climate-related features have a high impact on the model. Three of the most important features relate to T, three to ET, and one to P. The most important environmental features are associated with specific months rather than the overall time series.

Of the ten most dominant predictors, seven show anomaly values, while another two show trends. Anomaly features show a reverse relationship with SHAP values; specifically, lower anomaly values correspond to a higher predicted number of conflicts, while higher anomaly values correspond to a lower predicted number of conflicts. For many of the most dominant predictors, there is a clear directional relationship between feature values and SHAP values. Fig. 7 illustrates the spatial distribution of feature values for the ten most dominant predictors. Notably, there is a high autocorrelation between the mean value of previous conflicts and the distribution of conflicts in both the training and test datasets. This finding aligns well with reports of an autocorrelative effect of previous conflicts on conflict risk (Galli et al., 2022; Kamta et al., 2020).

Looking at the feature importance (i.e., the global mean of absolute SHAP values) of the four considered spheres of influence, anthroposphere, climate, hydrosphere, and biosphere, we see that features that belong to the anthroposphere make up the majority of the total feature importance of the final model (52.9 %). Within this sphere, previous conflicts are the single most important factor, accounting for 99.6 % of the anthroposphere's total feature importance. All other factors that belong to the anthroposphere (i.e., population density, SHDI, and the number of ethnic claims) only contribute 0.4 % to the anthroposphere's feature importance (Fig. 8). The climate is the most important of the remaining spheres of influence, contributing ca. 37 % of total feature importance to the final model. Within this sphere, the most important factor is ET, contributing 69.4 % to the sphere's feature importance. T and P contribute 16.7 % and 7.7 %, respectively. PET is the least important factor, accounting for 6.1 % of the sphere's feature importance. Among the different spheres, the hydrosphere and biosphere have the lowest feature importance, contributing 6.4 % and 3.8 % to the total feature importance of the final model, respectively. In the hydrosphere, soil moisture is the primary contributor, making up 79 % of its feature importance, while the remaining 21 % is attributed to surface water features. For the biosphere, nearly 89.3 % of its feature importance is derived from nGPP, with the remainder coming from iGPP.

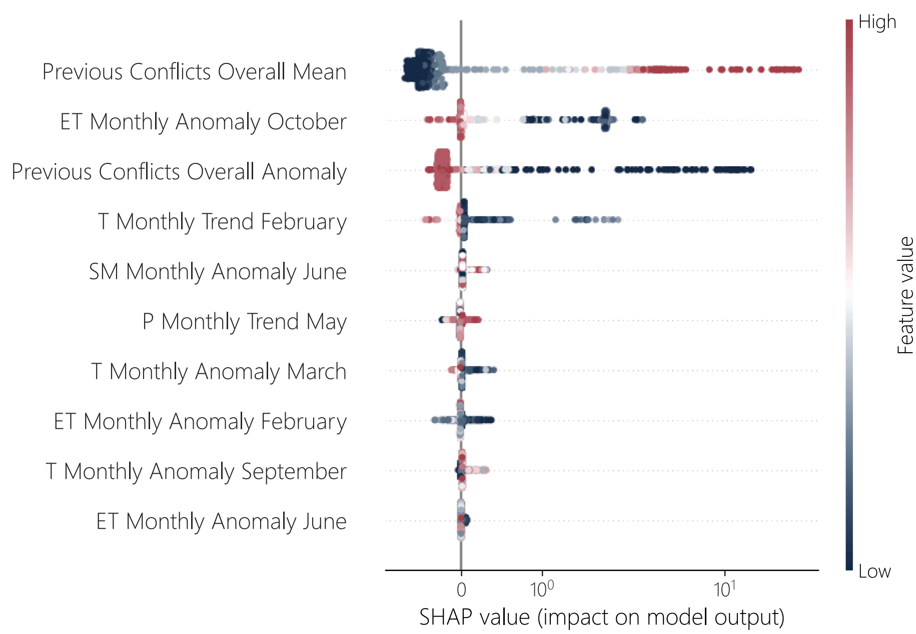

Fig. 6. Overview of SHAP feature importance of the ten most dominant predictors, sorted by their impact on the model. The abbreviations ET, T, SM, and P stand for total evapotranspiration, temperature, soil moisture, and precipitation, respectively. For each feature, a point cloud is shown, with each point standing for an observed feature value in one of the SHUs. The coloration and shading of the points per feature indicate whether a specific point has a higher (red) or lower (blue) value in comparison to the rest of the points for that feature. The SHAP value on the horizontal axis signifies the impact of observations on the model's prediction. Positive SHAP values increase the number of predicted conflicts, while negative SHAP values decrease the number of predicted conflicts. To give an example, the feature with the highest predictive power is the overall mean of previous conflicts. In SHUs where few conflicts have taken place in the past on average, the shown points are blue. The SHUs with moderate to high average numbers of conflicts in the past appear in pale to bright red. For the model decision, a low number of average previous conflicts decreases the expected number of future conflicts (the SHAP value is negative). On the other hand, moderate and especially high numbers of previous conflict occurrences elevate the expected number of conflicts in the future (the SHAP value is positive).

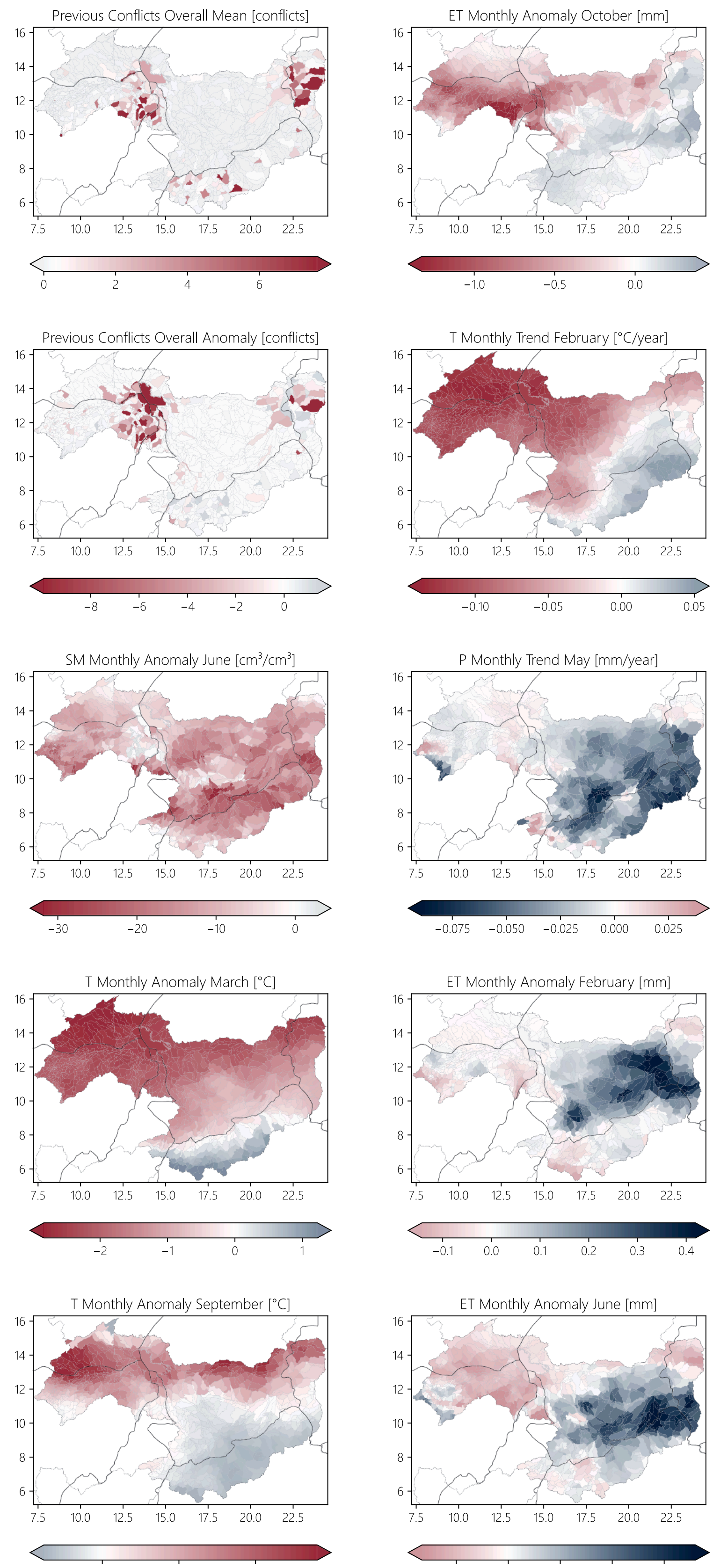

Fig. 7. Spatial distribution of observed values for the ten most dominant predictors. The direction of the colorscale relates to whether increasing feature values generally align with increasing SHAP values and thus increasing modeled conflicts. This directionality is indicated in Fig. 6. The abbreviations ET, T, SM, and P stand for total evapotranspiration, temperature, soil moisture, and precipitation, respectively.

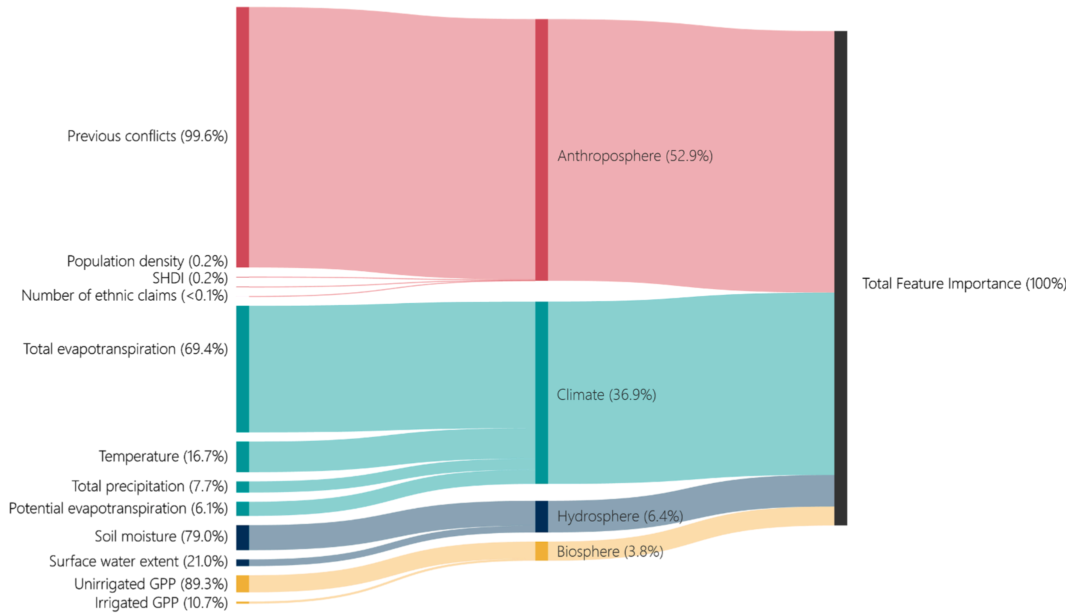

Fig. 8. Feature importance values for all twelve factors summarized per sphere of influence.

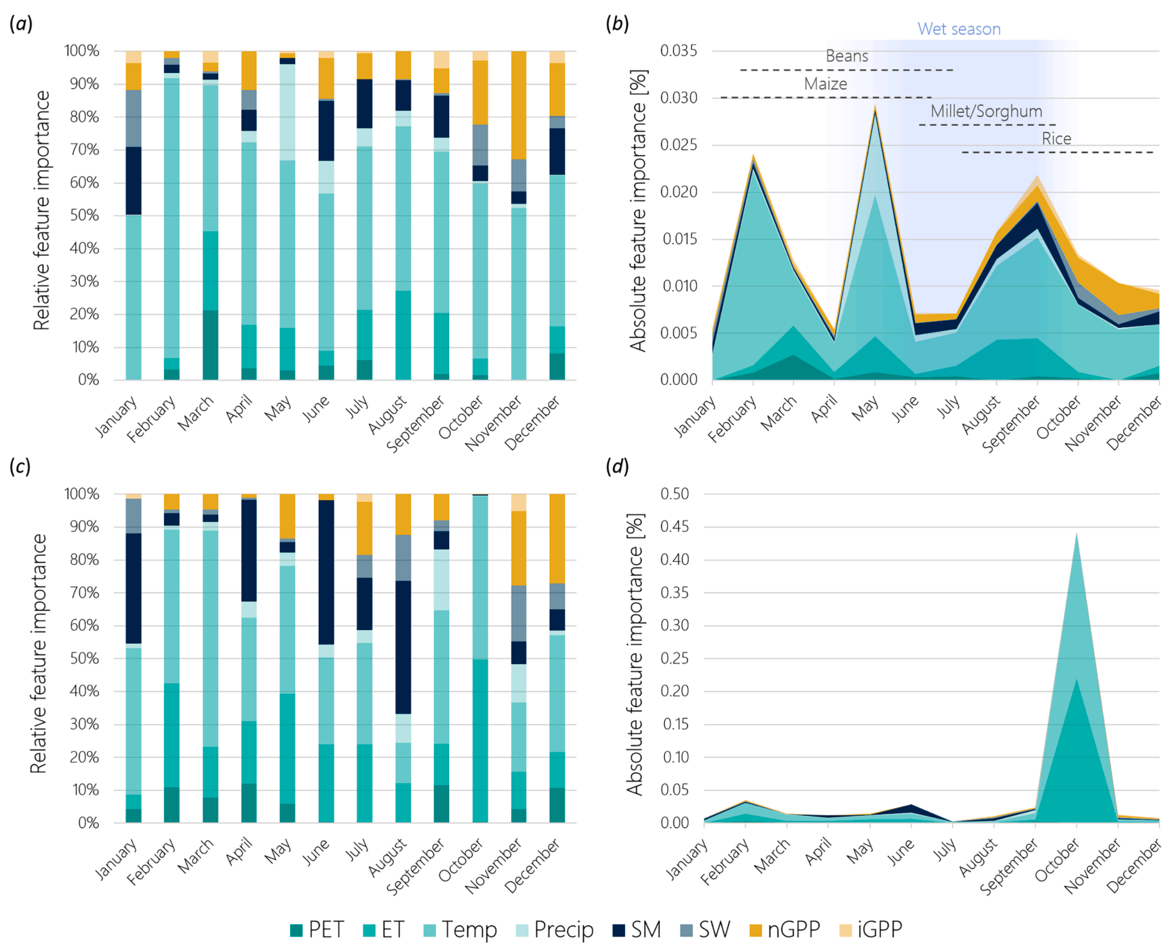

Fig. 9. Relative and absolute feature importance of monthly trends ((*a*), (*b*)) and anomalies ((*c*), (*d*)) of environmental factors. Values in (*a*) and (*c*) are percentages of the summed feature importance of all trend or anomaly features of a particular month. Subfigures (*b*) and (*d*) show the same expressed as percent values of absolute feature importance for the model. Additional information on the timing of the wet season and the cultivation periods of several important crops in the LCB is provided in (*b*) based on Batello et al. (2004). Features are aggregated to the factors potential evaporation (PET), total evaporation (including estimates on vegetation-based transpiration) (ET), air temperature (T), total precipitation (P), soil moisture (SM), surface water (SW), gross primary productivity in non-irrigated areas (nGPP), and gross primary productivity in irrigated areas (iGPP).

Overall, we see that the predictive power of the anthroposphere nearly exclusively hinges on the previous conflicts. The feature importance of previous conflicts outweighs all environmental factors. While environmental factors are not the main driver of conflict in the LCB, digging deeper into the feature importances of monthly environmental features allows for a detailed look into environmental stressors associated with conflicts.

The monthly features with the highest feature importance are monthly trends and anomalies. The relative feature importance of environmental factors is not distributed uniformly over the year (Fig. 9). Each sphere represented in Fig. 9 has a seasonality in its relative feature importance.

The trends of meteorological features hold the greatest relative significance among all monthly trend features throughout the year (Fig. 9 (a)). Their relative importance is highest in the early part of the year, specifically between February and May, with another notable peak in feature importance occurring between July and September. The initial peak aligns with the dry season and the onset of the wet season, while the latter peak takes place during the wet season. Notably, T remains the most significant meteorological factor year-round, while PET and ET trends show considerable relative significance in March and from July to September. P reaches its peak relative significance in May, coinciding with the beginning of the wet season, and maintains elevated values throughout this period. The feature importance of hydrosphere factors is higher from June to January and is at its lowest during the main dry season from February to May. The factors related to the biosphere display their highest relative importance in November, with relatively high values persisting from June to December.

In general, the trend characteristics observed in February, May, and from August to December hold the highest feature importance in the model (Fig. 9(b)). The seasonality of feature importances corresponds closely with agricultural periods in the LCB. The initial peak in February occurs during the dry season and aligns with the commencement of off-season maize and bean planting in areas that are seasonally flooded surrounding Lake Chad (Batello et al., 2004). The onset of the wet season, which typically falls between April and June depending on local climatic conditions, is characterized by a surge in feature importance, particularly relating to P trends. This increase in importance coincides with important phases in the agricultural calendar for maize, beans, as well as rainfed sorghum and cotton in various regions (Batello et al., 2004). During the wet season, SM and GPP become more significant (Fig. 9(b)).

In this period, several primary season crops are cultivated, with millet and sorghum being the most prominent. The cultivation of rainfed rice also occurs at this time, typically taking place in riverbeds and featuring both a rainfed and inundated phase (Batello et al., 2004). The inundated phase generally occurs in October and November, corresponding well with the heightened importance of surface water during these months.

Meteorological anomalies exhibit the greatest relative feature importance during the early months of the year and at the transition between the wet and dry seasons in October (Fig. 9(c)). Absolute importance of anomaly features peaks in October, largely influenced by the importance of ET and PET anomalies (Fig. 9(d)). Hydrological anomalies peak in importance during the height of the wet season (June-August). Anomalies related to the biosphere are most prominent during the wet season (July-September) and at the start of the dry season (November-December), primarily driven by nGPP anomalies. Anomalies in iGPP show limited relative importance throughout the year, peaking in July and November. While meteorological anomalies retain high relative significance all year long, hydrological anomalies surpass their importance in August, during the late wet season (Fig. 9(c)). Nevertheless, Fig. 9(d) clearly indicates that monthly anomalies generally have minimal influence on the model. The exception to this is the ET and PET anomalies in October.

4. Discussion

In the following section, we interpret the presented findings and put them into perspective with regard to the conflict situation in the LCB. We critically review the limitations of this study and the relevance of the presented findings.

4.1. Interpretation

Previous conflicts are the most dominant factor in the final model. This result is consistent with earlier studies in the LCB (Galli et al., 2022; Kamta et al., 2020) and with broader findings that conflict dynamics often exhibit strong temporal persistence (Buhaug and von Uexkull, 2025). In the final model, previous conflicts account for 52 % of the total feature importance. From a predictive perspective, this indicates a strong association between past and future conflict occurrences. The model thus captures the autocorrelative nature of conflict dynamics in the region, without explicitly representing the underlying feedback mechanisms.

The prominence of previous conflicts in the model is in line with qualitative and quantitative literature describing how sustained conflict activity constrains economic opportunities, disrupts livelihoods, and alters population movements, thereby shaping future conflict risk (Aniekwe, 2022; Ashindorbe et al., 2021; Buhaug and von Uexkull, 2025; Hassan et al., 2021; Kamta et al., 2020; Nkiaka et al., 2024; Olowoyeye and Kanwar, 2023; Zieba et al., 2017). These processes are not directly modeled here but may contribute to the strong predictive relevance of conflict history observed in the model.

Beyond conflict history, environmental factors, particularly climate-related factors, are weighed more heavily by the model than most socioeconomic factors. The patterns captured by the model are consistent with literature suggesting a link between environmental changes and conflict. While some suggest a causal relationship between environmental change and conflict (Hsiang et al., 2013; Leal Filho et al., 2022; Sharifi et al., 2021), this interpretation remains contested. Multiple studies have put forward points of critique regarding this interpretation, as well as the underlying methodological assumptions (Buhaug et al., 2014; Selby, 2014). Accordingly, the feature importance patterns identified here should be interpreted as indicators of predictive relevance rather than evidence of causality.

Several authors argue that relationships between environmental conditions and conflict are better understood as agential rather than causal (Meier et al., 2007; Selby, 2014; Witsenburg and Adano, 2009). In this view, environmental conditions form part of the situational context that conflict actors assess, and actions are planned accordingly. Nevertheless, even under this interpretation, environmental factors may carry meaningful information for understanding conflict dynamics.

Against this background, it is important to note that scarcity does not automatically translate into violent conflict. Scarce resources, especially in a water-related setting, more often than not lead to positive and cooperative interactions (Bernauer and Böhmelt, 2020; Kåresdotter et al., 2025; Petersen-Perlman et al., 2017; Wolf et al., 2005). In the context of water-related conflicts, Wolf et al. (2005) emphasize that disputes often occur within broader power struggles between actors, while water-specific clashes are often linked to quantity, quality, and timing. Translating this perspective to the LCB, climate and hydrological variables may reflect conditions under which existing tensions are more likely to materialize into observable conflict events. The inclusion of ET and PET in this study provides a more complete picture of the matter, which also features information on how much water evaporates and how much water would evaporate if there were an unlimited supply of water (i.e., the atmosphere's evaporative demand), thus giving an impression of green water and green water demand (Nkiaka et al., 2024). In this respect, green water means water resources that are tied to vegetation, and more specifically agricultural crops. Through this inclusion the current work extends earlier work that focused primarily on T and P

(Schleussner et al., 2016). There is a clear seasonality in the feature importance of ET and PET features. Considering previous studies, we interpret this seasonality as a hint towards intra-annual dynamics of green water and green water demand being especially critical during the agricultural main season, as well as at the start and end of off-season agricultural cycles. This aligns with the time when transhumant herders move their herds, and farmers sow and harvest their crops. Especially in these phases, vital access to water and agriculturally used areas may be blocked for pastoralists, as crop damage may occur from roaming livestock (Brottem, 2016). Especially when there are competing land claims and no recognized authority to settle such disputes, conflicts may arise. Over the last few years, these conflicts have become more and more violent (Aniekwe, 2022).

While population density and broader socioeconomic indicators, such as the SHDI, exhibit comparatively low feature importance in the mode, this does not imply that these factors are irrelevant. Instead, their predictive weight suggests that, within the present modeling framework and data structure, environmental variability and conflict history contain more information for predicting conflict occurrences. This finding is consistent with perspectives that view climate change as a stress multiplier that operates alongside governance failures and demographic pressures rather than a singular driver of conflict (Kåresdotter et al., 2025; Wakdok and Bleischwitz, 2021). Asah (2015) go even further and claim that linking climate change to deteriorating social and ecological systems defers agency and responsibility from local actors. In our eyes, this puts the focus on a perceived duality between human agency and climate change, and more directly, ecological change in the basin. Indeed, unsustainable resource use is part of the problem, but sustainable development cannot coexist with conflict and terror. Local actors are not passive victims of climate change. Of course, many try to secure their livelihood through diversification (Jellason et al., 2021). However, attacks that drain this resilience (e.g., cattle rustling, arson, extortion) become more and more frequent (Aniekwe, 2022) and make them more susceptible to environmental shocks (Okpara et al., 2016). Food security, which is a key component of sustainable development, is known to be negatively impacted by conflict. Conflict additionally forces migration and displacement (Martin-Shields and Stojetz, 2019), which increases resource competition in originally safer areas (Newman et al., 2023). Resource conflicts, such as transhumance-related inter-community conflicts in the LCB are directly impacted by climate change and the related uncertainty of seasonal patterns in T, P, and ET (McCarthy, 2020). We see this as a valid interpretation of the presented results, namely, the high feature importance of anomaly features at the start and end of the wet season.

Overall, the presented results support the interpretation that conflict dynamics in the LCB are shaped by impacts from all spheres, the anthroposphere, as well as the climate, hydrosphere, and biosphere. Previous conflicts dominate the model's predictions, but the impact of climate change and resource scarcity should not be dismissed as relevant contextual factors (Aniekwe, 2022). Krätli et al. (2018); McCarthy (2020); Okpara et al. (2017) propose that, aside shortcomings in governance, conflict in the LCB is driven by climate change and increased competition over scarce natural resources due to population growth. On that basis, we interpret the predictive impact of environmental change in the presented results as a relevant contextual factor that interacts with existing vulnerabilities and institutional constraints, contributing to observed patterns of conflict occurrence.

4.2. Limitations

The presented optimized features show what the model used in this study relies on for its predictions. While this can help understand correlations between features and outcomes, it does not inform about the causality of conflicts.

Especially findings like the low informational value of socioeconomic factors must be interpreted carefully. While their informational value for the presented model is low, others have found significant predictive capability in these factors (Galli et al., 2022; Ge et al., 2022; O'Loughlin et al., 2012; von Uexkull et al., 2016). We hypothesize that the low feature importance of these parameters in our analysis has to do with the low spatial and temporal resolution of the input layers. This shortcoming cannot be easily resolved with any currently available datasets and emphasizes the need for more research and reliable products with higher temporal and spatial resolution, especially in the socioeconomic domain.

There is inherent uncertainty in the choice of features we employ. While a constellation of SHDI and low previous conflict activity may serve as a rough approximation, our approach did not explicitly include mediating factors (e.g., governance capacity, local institutions, or conflict management mechanisms) that may impact conflict activity, as proposed by Theisen et al. (2013).

We account for both long-term trends and means and short-term disruptions in the form of anomalies and take into account seasonality in the form of monthly means, trends, and anomalies. However, this procedure to filter down time series to a manageable amount of features that can be fed to the model certainly entails a loss of information. Prospectively, approaching this topic from an angle of causal inference could allow leveraging the strengths of long EO data time series. However, we anticipate challenges that might complicate such an analysis, like possible hidden confounders and the low temporal resolution of demographic and socioeconomic datasets, especially in the LCB.

Many previous studies have tried focusing on a specific type of conflict, individual actors, or actor groups. In the presented research, we are considering all conflicts in the LCB within the investigated timeframe. This approach maximizes the number of conflict occurrences in training and test labels, as well as training and test features, and should therefore positively impact model learning. However, there is no way to attribute whether environmental features have, e.g., a more pronounced impact on resource-related conflicts than on insurgency attacks. Such an investigation may be an interesting next step. Newman et al. (2023) report a strong co-occurrence of agropastoral conflict and insurgency violence. They conclude that there may be significant intersection and overlap in their potential drivers.

Conflict, and especially violent conflict is a rare and extreme form of actor interaction. The chosen ACLED dataset presented a good fit for the study's requirements and has been successfully used in previous conflict modeling efforts (e.g., (Galli et al., 2022)). However, it has to be noted that dataset-specific uncertainties and limitations (Bernauer and Böhmelt, 2020) propagate through the modeling and subsequent interpretation. Most notably, the majority of non-violent negative interaction and all forms of positive (=cooperative) interaction between actors is not portrayed in this dataset. Alternative datasets exist that could be used to obtain a more complete picture of non-state conflicts in Africa (von Uexkull and Pettersson, 2018) or a global view on historic and recent water-related conflict and cooperation events (Kåresdotter et al., 2022). However, at the geographic scale and locality of this study and with the spatial and temporal resolution employed, leveraging these datasets was not viable. We propose that future research could address the limitations of the modeling framework in the choice of underlying EO data, the conflict and cooperation occurrence labels, and the choice of features. We advocate for assessing model performance through a cross-validation approach, longer time series of EO data and conflict data, a wider array of EO data and socioeconomic data, in combination with a wider range of regressors to balance the shortcomings of individual regressors.

4.3. Relevance of presented findings

Much of the research done in the field of conflict resolution and sustainable development focuses on qualitative studies. Often this is done through interviews, reviews, or conceptual frameworks. These approaches provide essential insights into local contexts, actor motivations, and institutional settings, but they rarely allow for a systematic assessment of the relative importance of different suspected conflict drivers across space and time. Quantitative approaches can complement this body of work by identifying patterns, regularities, and predictive signals that are difficult to detect through qualitative methods alone and help to formulate research questions and hypothesis for qualitative investigations of causalities.

The presented study builds on insights from previous research to inform the initial feature selection. Furthermore, we expand on frequently incorporated meteorological parameters to provide a broader representation of environmental change in the employed features as proposed by Schleussner et al. (2016). Our methodology attempts to limit the influence of our own biases by leaving the attribution of feature importance to a machine learning process. Through this, we obtain a list of features that reflect what impacts the model's predictive performance rather than subjective expectations.

This study offers a quantitative perspective on how changes in the anthroposphere and the environment are associated with conflicts in the LCB. Such information is relevant for monitoring and early-warning efforts, as well as for informing policy discussions that seek to address vulnerability in regions exposed to both conflict and environmental variability (Deininger and Goyal, 2024; Okpara et al., 2017). Importantly, our findings do not suggest that climate change is the only or even the main driver of conflict in the LCB. Instead, they support interpretations that view environmental change as a contextual factor or stress multiplier that interacts with existing socioeconomic and institutional vulnerabilities.

Historically, Africa's agricultural output growth was largely driven by area expansion, which increasingly conflicts with environmental goals and may trigger land degradation (Deininger and Goyal, 2024). In the context of escalating conflicts, it becomes clear that this unsustainable strategy has destabilizing effects on society (Asah, 2015). Over the past decade, many African countries have focused on large-scale land-based investments to increase productivity and diversify income sources, but the success of these transformation strategies was limited (Deininger and Goyal, 2024). Interventions were largely aimed at reducing the variability in the environment and lifestyles that are directed by seasonal patterns. Krätli et al. (2018) see this as a reason for increased social friction and undermined social and economic connectivity, resulting in less productive pastoralists and farmers who are more vulnerable to shocks, and a general increase in conflict incidence. These dynamics underscore the importance of understanding how environmental variability, land use, and conflict interact across time.

Whether the relationships identified between environmental factors and conflicts are causal or agential cannot be resolved within the present modeling framework. This requires more qualitative and participatory research. Nonetheless, quantifying the seasonal and interannual patterns through which environmental variables contribute to conflict prediction represents an important step toward integrating environmental information into conflict-sensitive planning and risk assessment. The presented results may therefore inform future research in conflict modeling and support discussions on early-warning mechanisms, relief, and recovery strategies aimed at managing the joint risks of drought, environmental stress, and violent conflict (Okpara et al., 2018).

5. Conclusion

The main objective of this research was to provide insights into the factors statistically associated with conflict occurrence, which may be valuable for future research in the field of sustainable development. To this end, two tasks were defined and tackled within this study:

1) Consider a wide array of factors from multiple spheres, including intra-annual dynamics of environmental features that affect the livelihoods in the LCB, and predict conflict occurrences.

We investigated how the anthroposphere, climate, hydrosphere, and biosphere impact the predictive performance in modeling conflict occurrences in the LCB. We considered long-term means, trends, and anomalies, as well as long-term monthly means, monthly trends, and monthly anomalies for factors with daily time series.

2) Untangle the socioeconomic and ecological factors that may be associated with conflicts based on their predictive power.

We employed a random forest model using conflict occurrences from 2019 as training labels. We tested model performance using conflict occurrences of 2020. Through feature optimization, we minimized the number of included features while increasing model performance. We performed a SHAP analysis to investigate the impact of previous conflicts, socioeconomic factors, climate factors, hydrological factors, and biosphere-related factors on the prediction of future conflict occurrences in the LCB.

- Previous conflicts are the single most dominant predictive factor to consider. They account for 52.9 % of total feature importance and 99.6 % of the anthroposphere's feature importance.
- The climate is the most dominant of all environmental factors in the model. It accounts for 36.9 % of the total feature importance.
- The hydrosphere accounts for 6.4 % of the total feature importance. Within this sphere, portrayed through soil moisture and surface water, soil moisture holds ca. 79 % of the total feature importance, and surface water contributes 21 % to the sphere's feature importance.
- The biosphere accounts for 3.8 % of the total feature importance. It is portrayed via the GPP in unirrigated and irrigated areas. Within this sphere, non-irrigated areas account for 89.3 % of the sphere's feature importance.
- Model performance is highest when all spheres are included.

These outcomes should be interpreted strictly as results of predictive modeling rather than as evidence of causal pathways. That being said, many of our model-based findings align well with quantitative and qualitative literature. There are some limitations to the final model. Especially outlier cases, in which conflict occurrences have spiked extremely in the testing year, are underestimated. We propose that future approaches can expand on the presented framework to tackle these limitations. The use of longer time series of EO data, the involvement of additional conflict and cooperation datasets, and a wider array of socioeconomic factors may be part of this. Still, the presented findings demonstrate how multi-spherical environmental indicators and their seasonal dynamics can enhance predictive conflict modeling, offering a machine-learning-based framework for integrating EO-based environmental information into conflict-sensitive monitoring and early-warning systems.

Data availability statement

This study utilized the following datasets. The DLR Global WaterPack, available for download at https://geoservice.dlr.de/web/datasets/globalwaterpack; ERA5-Land, available for download at https://doi.org/10.24381/cds.e2161bac; Global MODIS and FLUXNET-derived Daily Gross Primary Production V2, available for download at https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1835; Global Surface Soil Moisture, available for download at https://doi.org/10.6084/m9.figshare.21806457.v1; WorldPop population density, available for download at https://doi.org/10.5258/SOTON/WP00675; the Subnational Human Development Index, available for download at https://globaldatalab.org/shdi/; the Geo-referencing Ethnic Power Relations dataset, available for download at https://icr.ethz.ch/data/epr/geoepr/; Armed Conflict Location and Event Data, available for download at https://acleddata.com/data/; the HydroBASINS dataset, available for download at https://www.hydrosheds.org/products/hydrobasins#downloads; GADM administrative boundaries, available for download at https://gadm.org/data.html; Ecoregions, available for download at https://ecoregions.appspot.com/; Global Irrigated Areas, available for download at https://doi.org/10.1594/PANGAEA.884744.

Funding

This research has been supported by the DFG (Deutsche Forschungsgemeinschaft) within the framework of the Research Unit GlobalCDA (Global Calibration and Data Assimilation) (grant no. FOR 2630).

CRediT authorship contribution statement

Patrick Sogno: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Thorsten Hoeser:** Writing – review & editing. **Penny Beames:** Writing – review & editing. **Reeves Meli Fokeng:** Writing – review & editing. **Claudia Kuenzer:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We acknowledge the use of Grammarly v1.2.124.1571, which includes a generative AI assistant for English language editing. All AI-generated text suggestions have undergone rigorous revision by the authors. We acknowledge the use of Codeium v1.16 in the development of unit tests for the developed feature optimization functions. All AI-generated code suggestions have undergone rigorous revision by the authors.

References

Adeyeri, O.E., Laux, P., Lawin, A.E., Ige, S.O., Kunstmann, H., 2019. Analysis of hydrometeorological variables over the transboundary Komadugu-Yobe basin, West Africa. J. Water Clim. Change 11 (4), 1339–1354. https://doi.org/10.2166/wcc.2019.283.

Aniekwe, C.C., 2022. Conflict Analysis in the Lake Chad Basin 2020-2021. Trends, Developments and Implications For Peace and Stability. United Nations Development Programme Regional Hub West and Central Africa, N'Djamena, Chad.

Asah, S.T., 2015. Transboundary hydro-politics and climate change rhetoric: an emerging hydro-security complex in the lake chad basin. WIREs. Water. 2 (1), 37–45. https://doi.org/10.1002/wat2.1057.

Ashindorbe, K., Afatakpa, F., Owonikoko, S.B., 2021. Civilian Joint Task Force and Nigeria's Counter-Terrorism Operation: a critique of the community-based approach to insecurity. Afr. Secur. 14 (3), 286–305. https://doi.org/10.1080/19392206.2021.1998977.

Batello, C., Marzot, M., Touré, A.H., 2004. The Future is an Ancient lake. Traditional knowledge, Biodiversity and Genetic Resources For Food and Agriculture in Lake Chad Basin ecosystems. Food and Agriculture Organization of the United Nations, Rome.

Beck, H.E., McVicar, T.R., Vergopolan, N., Berg, A., Lutsko, N.J., Dufour, A., Zeng, Z., Jiang, X., van Dijk, A.I.J.M., Miralles, D.G., 2023. High-resolution (1 km) Köppen-Geiger maps for 1901–2099 based on constrained CMIP6 projections. Sci. Data 10 (1), 724. https://doi.org/10.1038/s41597-023-02549-6.

Bennour, A., Jia, L., Menenti, M., Zheng, C., Zeng, Y., Barnieh, B.A., Jiang, M., 2023. Assessing impacts of climate variability and land use/land cover change on the water balance components in the Sahel using Earth observations and hydrological modelling. J. Hydrol.: Reg. Stud. 47, 101370. https://doi.org/10.1016/j.ejrh.2023.101370.

Bernauer, T., Böhmelt, T., 2020. International conflict and cooperation over freshwater resources. Nat. Sustain. 3 (5), 350–356. https://doi.org/10.1038/s41893-020-0479-8.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32. https://doi.org/10.1023/A:1010933404324.

Brottem, L.V., 2016. Environmental change and farmer-herder conflict in agro-pastoral West Africa. Hum. Ecol. 44 (5), 547–563. https://doi.org/10.1007/s10745-016-9846-5.

Buhaug, H., Nordkvelle, J., Bernauer, T., Böhmelt, T., Brzoska, M., Busby, J.W., Ciccone, A., Fjelde, H., Gartzke, E., Gleditsch, N.P., et al., 2014. One effect to rule them all? A comment on climate and conflict. Clim. Change 127 (3), 391–397. https://doi.org/10.1007/s10584-014-1266-1.

Buhaug, H., von Uexkull, N., 2025. Strong rationale, weak evidence: why integrating research on sustainability and peacebuilding is needed. One Earth. 8 (9), 101452. https://doi.org/10.1016/j.oneear.2025.101452.

Calvin, K., Dasgupta, D., Krinner, G., Mukherji, A., Thorne, P.W., Trisos, C., Aldunce, P., Barrett, K., Blanco, G., et al., 2023. IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. In: Romero, J., Core Writing Team, Lee, H., Romero, J., et al. (Eds.). Intergovernmental Panel on Climate Change (IPCC), Geneva, Switzerland. Geneva, Switzerland. IPCC.

Cong C., Pan H., Yang Y., Cai Z., Kåresdotter E., Bahn J., Zhou H., Page J., Döring S., Krampe F., et al. 2025. Effective modes of transboundary water-related cooperation to turn future climate conflict risks into opportunities [Internet]. [accessed 2025 Dec 18]. https://doi.org/10.21203/rs.3.rs-6847954/v1.

Deininger, K., Goyal, A., 2024. Land Policies for Resilient and Equitable Growth in Africa. World Bank Publications, Washington, D. C., USA. https://doi.org/10.1596/978-1-4648-2024-3.

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N.D., Wikramanayake, E., Hahn, N., Palminteri, S., Hedao, P., Noss, R., et al., 2017. An ecoregion-based approach to protecting half the terrestrial realm. Bioscience 67 (6), 534–545. https://doi.org/10.1093/biosci/bix014.

Fokeng, R.M., Bachofer, F., Sogno, P., Klein, I., Uereyen, S., Kuenzer, C., 2024. Surface water dynamics of Lake Chad Basin (Sahelian Africa) based on daily temporal resolution Earth observation time series. J. Hydroinformatics 26 (9), 2325–2352. https://doi.org/10.2166/hydro.2024.130.

Fu, S., Zhou, Y., Lei, J., Zhou, N., 2023. Changes in the spatiotemporal of net primary productivity in the conventional Lake Chad Basin between 2001 and 2020 based on CASA model. Atmos. (Basel) 14 (2). https://doi.org/10.3390/atmos14020232.

Galli, N., Dell'Angelo, J., Epifani, I., Chiarelli, D.D., Rulli, M.C, 2022. Socio-hydrological features of armed conflicts in the Lake Chad Basin. Nat. Sustain. 5 (10), 843–852. https://doi.org/10.1038/s41893-022-00936-2.

Gbetkom, P.G., Crétaux, J.-F., Tchilibou, M., Carret, A., Delhoume, M., Bergé-Nguyen, M., Sylvestre, F., 2023. Lake Chad vegetation cover and surface water variations in response to rainfall fluctuations under recent climate conditions (2000-2020). Sci. Total Environ. 857, 159302. https://doi.org/10.1016/j.scitotenv.2022.159302.

Ge, Q., Hao, M., Ding, F., Jiang, D., Scheffran, J., Helman, D., Ide, T., 2022. Modelling armed conflict risk under climate change with machine learning and time-series data. Nat. Commun. 13 (1), 2839. https://doi.org/10.1038/s41467-022-30356-x.

GIZ, 2016. Report On the State of the Lake Chad Basin Ecosystem. Bonn and Eschborn. Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), Germany.

Han, Q., Zeng, Y., Zhang, L., Wang, C., Prikaziuk, E., Niu, Z., Su, B., 2023. Global long term daily 1 km surface soil moisture dataset with physics informed machine learning. Sci. Data 10 (1), 101. https://doi.org/10.1038/s41597-023-02011-7.

Hassan, S.I., Usman, A., Usman, G.U., 2021. Resource conflict and environmental challenges in the Lake Chad Basin: an analysis of the sedentary farmers and nomadic pastoralists conflict in Nigeria. Fuwukari J. Polit. Dev. 5 (1), 41–49.

Hsiang, S.M., Burke, M., Miguel, E., 2013. Quantifying the influence of climate on Human conflict. Science 341 (6151), 1235367. https://doi.org/10.1126/science.1235367.

Jellason, N.P., Conway, J.S., Baines, R.N., Ogbaga, C.C., 2021. A review of farming challenges and resilience management in the Sudano-Sahelian drylands of Nigeria in an era of climate change. J. Arid Env. 186, 104398. https://doi.org/10.1016/j.jaridenv.2020.104398.

Joiner, J., Yoshida, Y., 2020. Satellite-based reflectances capture large fraction of variability in global gross primary production (GPP) at weekly time scales. Agric Meteorol 291, 108092. https://doi.org/10.1016/j.agrformet.2020.108092.

Kamta, F.N., Schilling, J., Scheffran, J., 2020. Insecurity, resource scarcity, and migration to camps of internally displaced persons in Northeast Nigeria. Sustainability. 12 (17), 6830. https://doi.org/10.3390/su12176830.

Kåresdotter, E., Destouni, G., Lammers, R.B., Keskinen, M., Pan, H., Kalantari, Z., 2025. Water conflicts under climate change: research gaps and priorities. Ambio 54 (4), 618–631. https://doi.org/10.1007/s13280-024-02111-7.

Kåresdotter E., Skoog G., Pan H., Kalantari Z. 2022. New global dataset on historical water-related conflict and cooperation events [Internet]. [accessed 2025 Dec 19]. https://doi.org/10.5281/zenodo.7465153.

Klein, I., Uereyen, S., Sogno, P., Twele, A., Hirner, A., Kuenzer, C., 2024. Global WaterPack - the development of global surface water over the past 20 years at daily temporal resolution. Sci. Data 11 (1), 472. https://doi.org/10.1038/s41597-024-03328-7.

Krätli, S., Sougnabé, P., Staro, F., Young, H., 2018. Pastoral Systems in Dar Sila Chad: A Background Paper for Concern Worldwide. Feinstein International Center, Tufts University, Boston.

Leal Filho, W., Totin, E., Franke, J.A., Andrew, S.M., Abubakar, I.R., Azadi, H., Nunn, P.D., Ouweneel, B., Williams, P.A., Simpson, N.P., 2022. Understanding responses to climate-related water scarcity in Africa. Sci. Total Environ. 806, 150420. https://doi.org/10.1016/j.scitotenv.2021.150420.

Lehner, B., Grill, G., 2013. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. Hydrol. Process. 27 (15), 2171–2186. https://doi.org/10.1002/hyp.9740.

Lemoalle, J., Bader, J.-C., Leblanc, M., Sedick, A., 2012. Recent changes in Lake Chad: observations, simulations and management options (1973–2011). Glob. Planet. Change 80–81, 247–254. https://doi.org/10.1016/j.gloplacha.2011.07.004.

Li, K.Y., Coe, M.T., Ramankutty, N., Jong, R.D., 2007. Modeling the hydrological impact of land-use change in West Africa. J Hydrol 337 (3), 258–268. https://doi.org/10.1016/j.jhydrol.2007.01.038.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I., 2020. From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. 2 (1), 56–67. https://doi.org/10.1038/s42256-019-0138-9.

Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems, 30. Curran Associates, Inc., Los Angeles, USA, p. 9.

Luxereau, A., Genthon, P., Ambouta Karimou, J.-M, 2012. Fluctuations in the size of Lake Chad: consequences on the livelihoods of the riverain peoples in eastern Niger. Reg Env. Change 12 (3), 507–521. https://doi.org/10.1007/s10113-011-0267-0.

Martin-Shields, C.P., Stojetz, W., 2019. Food security and conflict: empirical challenges and future opportunities for research and policy making on food security and conflict. World Dev. 119, 150–164. https://doi.org/10.1016/j.worlddev.2018.07.011.

McCarthy, P., 2020. Preventing Transhumance-Related Intercommunity Conflict in Chad: Towards a climate-Sensitive Conflict Analysis. UNDP, Chad.

Meier, J., Zabel, F., Mauser, W., 2018. A global approach to estimate irrigated areas – a comparison between different data and statistics. Hydrol. Earth Syst. Sci. 22 (2), 1119–1133. https://doi.org/10.5194/hess-22-1119-2018.

Meier, P., Bond, D., Bond, J., 2007. Environmental influences on pastoral conflict in the Horn of Africa. Polit. Geogr. 26 (6), 716–735. https://doi.org/10.1016/j.polgeo.2007.06.001.

Mohamed, A., Abdelrady, A., Alarifi, S.S., Othman, A., 2023. Geophysical and remote sensing assessment of Chad's groundwater resources. Remote Sens. 15 (3), 560. https://doi.org/10.3390/rs15030560.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., et al., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data 13 (9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021.

Neukum, C., Morales-Santos, A., Ronelngar, M., Bala, A., Vassolo, S., 2023. Modelling groundwater recharge, actual evaporation, and transpiration in semi-arid sites of the Lake Chad basin: the role of soil and vegetation in groundwater recharge. Hydrol. Earth Syst. Sci. 27 (19), 3601–3619. https://doi.org/10.5194/hess-27-3601-2023.

Newman, E., Khiabani, P.H., Chandran, R., 2023. Intercommunal violence, insurgency, and agropastoral conflict in the Lake Chad Basin region. Small Wars Insur. 0 (0), 1–31. https://doi.org/10.1080/09592318.2023.2248868.

Nkiaka, E., Bryant, R.G., Kom, Z., 2024. Understanding links between water scarcity and violent conflicts in the Sahel and Lake Chad Basin using the water footprint concept. Earth's. Future 12 (2), e2023EF004013. https://doi.org/10.1029/2023EF004013.

Okpara, U.T., Stringer, L.C., Dougill, A.J., 2016. Lake drying and livelihood dynamics in Lake Chad: unravelling the mechanisms, contexts and responses. Ambio 45 (7), 781–795. https://doi.org/10.1007/s13280-016-0805-6.

Okpara, U.T., Stringer, L.C., Dougill, A.J., 2017. Using a novel climate–water conflict vulnerability index to capture double exposures in Lake Chad. Reg Env. Change 17 (2), 351–366. https://doi.org/10.1007/s10113-016-1003-6.

Okpara, U.T., Stringer, L.C., Dougill, A.J., 2018. Integrating climate adaptation, water governance and conflict management policies in lake riparian zones: insights from African drylands. Env. Sci Policy 79, 36–44. https://doi.org/10.1016/j.envsci.2017.10.002.

Okpara, U.T., Stringer, L.C., Dougill, A.J., Bila, M.D., 2015. Conflicts about water in Lake Chad: are environmental, vulnerability and security issues linked? Prog. Dev. Stud. 15 (4), 308–325. https://doi.org/10.1177/1464993415592738.

O'Loughlin, J., Witmer, F.D.W., Linke, A.M., Laing, A., Gettelman, A., Dudhia, J., 2012. Climate variability and conflict risk in East Africa, 1990–2009. Proc. Natl. Acad. Sci. u S. a 109 (45), 18344–18349. https://doi.org/10.1073/pnas.1205130109.

Olowoyeye, O.S., Kanwar, R.S., 2023. Water and food sustainability in the riparian countries of Lake Chad in Africa. Sustainability. 15 (13), 10009. https://doi.org/10.3390/su151310009.

Petersen-Perlman, J.D., Veilleux, J.C., Wolf, A.T., 2017. International water conflict and cooperation: challenges and opportunities. Water Int. 42 (2), 105–120. https://doi.org/10.1080/02508060.2017.1276041.

Pham-Duc, B., Sylvestre, F., Papa, F., Frappart, F., Bouchez, C., Crétaux, J.-F., 2020. The Lake Chad hydrology under current climate change. Sci. Rep. 10 (1), 5498. https://doi.org/10.1038/s41598-020-62417-w.

Raleigh, C., Kishi, R., Linke, A., 2023. Political instability patterns are obscured by conflict dataset scope conditions, sources, and coding choices. Humanit. Soc. Sci. Commun. 10 (1), 1–17. https://doi.org/10.1057/s41599-023-01559-4.

Reynolds, T.W., Waddington, S.R., Anderson, C.L., Chew, A., True, Z., Cullen, A., 2015. Environmental impacts and constraints associated with the production of major food crops in Sub-Saharan Africa and South Asia. Food Sec 7 (4), 795–822. https://doi.org/10.1007/s12571-015-0478-1.

Riebe, K., Dressel, A., 2021. The impact on food security of a shrinking Lake Chad. J. Arid Env. 189, 104486. https://doi.org/10.1016/j.jaridenv.2021.104486.

Schleussner, C.-F., Donges, J.F., Donner, R.V., Schellnhuber, H.J., 2016. Armed-conflict risks enhanced by climate-related disasters in ethnically fractionalized countries. Proc. Natl. Acad. Sci. 113 (33), 9216–9221. https://doi.org/10.1073/pnas.1601611113.

Selby, J., 2014. Positivist Climate conflict research: a critique. Geopolitics. 19 (4), 829–856. https://doi.org/10.1080/14650045.2014.964865.

Sharifi, A., Simangan, D., Lee, C.Y., Reyes, S.R., Katramiz, T., Josol, J.C., Muchangos, L. D., Virji, H., Kaneko, S., Tandog, T.K., et al., 2021. Climate-induced stressors to peace: a review of recent literature. Env. Res Lett 16 (7), 073006. https://doi.org/10.1088/1748-9326/abfc08.

Sogno, P., Klein, I., Uereyen, S., Bachofer, F., Kuenzer, C., 2024. Surface water dynamics of Africa: analysing continental trends and identifying drivers for major lakes and reservoirs. Int. J. Remote Sens. 45 (24), 9538–9568. https://doi.org/10.1080/01431161.2024.2412802.

Theisen, O.M., Gleditsch, N.P., Buhaug, H., 2013. Is climate change a driver of armed conflict? Clim. Change 117 (3), 613–625. https://doi.org/10.1007/s10584-012-0649-4.

von Uexkull, N., Croicu, M., Fjelde, H., Buhaug, H., 2016. Civil conflict sensitivity to growing-season drought. Proc. Natl. Acad. Sci. 113 (44), 12391–12396. https://doi.org/10.1073/pnas.1607542113.

von Uexkull, N., Pettersson, T., 2018. Issues and actors in African nonstate conflicts: a new data set. Int. Interact. 44 (5), 953–968. https://doi.org/10.1080/03050629.2018.1493478.

University of Southampton, University of Louisville, Universite de Namur, Columbia University. 2020. WorldPop. https://dx.doi.org/10.5258/SOTON/WP00675.

Vogt, M., Bormann, N.-C., Rüegger, S., Cederman, L-E, Hunziker, P., Girardin, L., 2015. Integrating data on ethnicity, geography, and conflict: the ethnic power relations data set Family. J. Confl. Resolut. 59 (7), 1327–1342. https://doi.org/10.1177/0022002715591215.

Wakdok, S.S., Bleischwitz, R., 2021. Climate Change, Security, and the resource nexus: case study of Northern Nigeria and Lake Chad. Sustainability. 13 (19), 10734. https://doi.org/10.3390/su131910734.

Witsenburg, K.M., Adano, W.R., 2009. Of rain and raids: violent livestock raiding in Northern Kenya. Civ. Wars 11 (4), 514–538. https://doi.org/10.1080/13698240903403915.

Wolf, A.T., Kramer, A., Carius, A., Dabelko, G.D., 2005. Managing water conflict and cooperation. State of the World 2005, 22nd ed. Routledge [place unknown].

Zhu, W., Jia, S., Lall, U., Cao, Q., Mahmood, R., 2019. Relative contribution of climate variability and human activities on the water loss of the Chari/Logone River discharge into Lake Chad: a conceptual and statistical approach. J Hydrol 569, 519–531. https://doi.org/10.1016/j.jhydrol.2018.12.015.

Zieba, F.W., Yengoh, G.T., Tom, A., 2017. Seasonal migration and settlement around Lake Chad: strategies for control of resources in an increasingly drying Lake. Resources 6 (3), 41. https://doi.org/10.3390/resources6030041.