nature portfolio

https://doi.org/10.1038/s41893-022-00936-2

Supplementary information

Socio-hydrological features of armed conflicts in the Lake Chad Basin

In the format provided by the authors and unedited

Supplementary Methods

Balance and water indicators

The water indicators are calculated starting from a soil water balance performed monthly and at a 5 arc minutes resolution. Equation (1) is the water balance for a single cell:

$$\frac{\Delta S}{\Delta t} = P_t + I_t - ET_{act,t} - Dp_t - R_t \qquad \text{Equation (1)}$$

where $\Delta S[mm]$ is the daily change in water storage in the cell, Δt is the timestep of one day used in the simulation, $P_t[\frac{mm}{day}]$ is the daily effective precipitation, $I_t[\frac{mm}{day}]$ is the irrigation supply (only for irrigated crops), $ET_{act,t}[\frac{mm}{day}]$ is the actual evapotranspiration, $Dp_t[\frac{mm}{day}]$ is the deep percolation and $R_t[\frac{mm}{day}]$ is the surface runoff. According to the guidelines of the FAO paper 56[77], the evapotranspiration is given by the product of the reference evapotranspiration and a crop coefficient, accounting also for stress of the plant in the case of rainfed crops. Crop specific cultivated areas are retrieved from the MIRCA dataset[78]. Non-harvested areas are retrieved from the GlobCover 2009 project[79] and their evaporation coefficients are taken in their calibrated form from[80]. Precipitation data are retrieved as rainfall from the University of East Anglia's Climate Research Unit CRU TS 2.0 dataset[81] and converted to effective by using a Dunnian approach. Although the balance is performed on a daily timescale, the results are aggregated by month.
From the balance in Equation (1) a set of water indicators is calculated, the first of which is water scarcity $WS[-]$, computed as the ratio between water withdrawal and freshwater availability:

$$WS_i[m^3] = \frac{Dom_i + Ind_i + Agr_i}{0.2R_i + \langle [0.2R - (D + I + A)]_{up(i)} \rangle} \qquad \text{Equation (2)}$$

where $Dom_i[m^3]$, $Ind_i[m^3]$ and $Agr_i[m^3]$ are the domestic, industrial, and agricultural blue water footprints in the cell i, respectively and $R_i[m^3]$ is the runoff. The $up(i)$ pedix means that its argument is taken as sum of the upstream contributions to the cell i, whereas the pointy brackets $\langle \quad \rangle$ mean that the argument is set to zero when negative. In this way, we consider water availability in a cell as the sum of water directly available locally and the surplus (non-used water availability) coming from upstream areas, when present. Domestic and industrial water footprint are taken from Mekonnen & Hoekstra[69], whereas the agricultural blue water footprint and the runoff are outputs of the hydrological balance model. Millimetric fluxes are converted into volumes using pixel-specific areas. The 0.2 factor in Equations (1) to (6) accounts for the presence of environmental flows[81].
Equation (3) is analogous to the denominator of Equation 2, and it is used to calculate the total amount of water available for withdrawal:

$$WA_i[m^3] = 0.2R_i + \langle [0.2R - (Dom + Ind + Agr)]_{up(i)} \rangle \qquad \text{Equation (3)}$$

In order to have a measure of water availability that is more representative of its importance for human livelihoods, we compute per capita water availability indicator WAPC, dividing Equation (3) by the number of inhabitants in the cell. The cell population is calculated converting the population density data retrieved from WorldPop[70]. The resulting formula is Equation (4).

$$WAPC_i\left[\frac{m^3}{cap}\right] = \frac{0.2R_i + \langle [0.2R - (Dom + Ind + Agr)]_{up(i)} \rangle}{Population_i} \qquad \text{Equation (4)}$$

The yearly cumulate of $WAPC$ can be compared to the thresholds proposed by Falkenmark et al.[82], representing the minimum amount of water a single person requires yearly to avoid conditions of water stress and scarcity. In Figure S2, these bounds are used for representation, as their original yearly values in the yearly map and as rescaled monthly values in the monthly map.
To better consider the importance of water for livelihoods, evapotranspiration from cultivated areas is calculated as the green water flux for the millet and sorghum cultivated areas during their growing season (Figure S1). These crops are selected as the main crops in the area, basing on FAOSTAT data on crop production for the six countries intersecting the study area (Niger, Nigeria, Cameroon, Chad, Central African Republic, Sudan)[83]. Then, the crop calendars are interpolated to select a unique range going from July to October. Here $GW_i[m^3]$ stands for the total flux, whereas its per capita flux is denoted by $GWPC_i\left[\frac{m^3}{cap}\right]$.
The third couple of indicators ($WAG, WAPCG$) accounts for all the water resources sustaining human essential needs. From an operational point of view, the sum of water availability and green water for food production is computed, obtaining the formulas in Equations (5) and (6):

$$WAG_i[m^3] = 0.2R_i + GW_{cult_i} + \langle [0.2R - (D + I + A)]_{up(i)} \rangle \qquad \text{Equation (5)}$$

$$WAGPC_i\left[\frac{m^3}{cap}\right] = \frac{0.2R_i + GW_{cult_i} + \langle [0.2R - (D + I + A)]_{up(i)} \rangle}{Population_i} \qquad \text{Equation (6)}$$

The indicators are put into relation to conflict through two different analyses: spatial econometrics and conflict points analysis.

Spatial econometrics

In Zero-Inflated Poisson (ZIP) regression models[60], the Poisson distribution is conditioned by a non-zero outcome of a binomial distribution. ZIP models assume that the dependent variable is Poisson-distributed with a probability $1 - \pi$, and is concentrated on zero with probability π. This means that a zero in the data may originate from a null outcome of the binomial distribution (the Poisson counting process did not set off) or from a zero counts outcome of the Poisson counting process, as in Equation (7):

$$\begin{cases} P(y = 0) = \pi + (1 - \pi)e^{-\lambda} \\ P(y = k) = (1 - \pi)\dfrac{e^{-\lambda}\lambda^k}{k!} \quad k = 1,2,\dots \end{cases} \qquad \text{Equation (7)}$$

We build a Bayesian hierarchical model structure that includes both the spatial components and the zero-inflated component[74]. The model parameters are estimated in a Bayesian approach. The unknown parameters are understood as random variables with a prior joint distribution and the statistical problem consists of updating this distribution by computing a posterior joint conditional probability of the parameters given the data.
To account for spatial interactions in the outcome variables and in the covariates, we set up four model specifications on the regressive component for the estimation of the Poisson parameter λ, with a logarithmic link function. The four models are defined with increasing complexity, following the spatial model taxonomy by Elhorst[84]. The first is the baseline (BSL) model in Equation (8), with no spatial interaction:

$$\log(\lambda) = \beta X + \varepsilon \qquad \text{Equation (8)}$$

We assign partially informative zero-centered gaussian independent priors to the regression coefficients β. The error term is a priori normally distributed with zero mean and standard deviation σ_ε, whose prior distribution is inverse-gamma with uninformative parameters. The second model is the Spatially Lagged Explanatory Variables X (SLX) model in Equation (9), which accounts for the effects of covariates in the neighbouring cells. The neighbouring cells are identified via a square contiguity matrix, whose elements $w_{i,j}$ are equal to 1 when cell i is adjacent to cell j and zero otherwise. By convention $w_{i,i} = 0$. The contiguity matrix is then row-normalized to obtain the spatial weights matrix W, which is then used in the models:

$$\log(\lambda) = X\beta + WX\theta + \varepsilon \qquad \text{Equation (9)}$$

In the SLX case, to avoid shadow effects, we selected only the spatial lag of covariates having a negligible covariance between a cell and its neighbor cell values. Parameter θ shares the same prior with β. The spatial lag of conflict (SLC) model in Equation (10) accounts for the spatial interaction of conflicts through a spatially structured random effects component u, i.e. a random variable centered in the average of its neighbors' values, and multiplied by the coefficient ρ; ρ can therefore be thought of as the spatial autocorrelation of the outcome between each cell and its neighbours[85]:

$$\begin{gathered} \log(\lambda) = X\beta + u + \varepsilon \\ u \sim N(\rho W u, \tau_u) \;\; 0 < \rho < 1 \end{gathered} \qquad \text{Equation (10)}$$

Following Elhorst[84], the spatial autoregressive parameter ρ is a priori logit-beta-distributed, with uninformative parameters. The last competing model given in Equation (11) is a spatial lag of conflict and covariates (SLCX) model. It includes spatially lagged values of both the random effect u and the independent variables X:

$$\log(\lambda) = X\beta + WX\theta + u + \varepsilon \qquad \text{Equation (11)}$$

As far as the binomial parameter π is concerned, it is thought of as a transformed parameter of a hyperparameter α:

$$\pi = \frac{\exp \alpha}{1 + \exp \alpha} \qquad \text{Equation (12)}$$

where a priori α has zero-centered normal distribution with high standard deviation. In this way, the prior on π is diffuse on the $(0,1)$ range.

The fit is performed via integrated nested Laplace approximation, using the R-package "INLA"[75] and the model selection criterion is the deviance information criterion (DIC). Where the DIC results are too close to call, we choose the model whose posterior density best fits the observed data histograms.

Supplementary Results

Balance and water indicators

The water indicators calculated through the soil water balance model (Table 1) show how climate and human activities interact in shaping water availability. Green water consumption (GW) presents an uneven distribution in the study area. Darfur shows much higher values of GW than the neighboring Chadian areas (Figure S 1), resulting rather from greater extent of harvested area than from different potential evapotranspiration. By comparing yearly and monthly values of water availability (WA), total water availability (WAG), their per capita values (WAPC, WAGPC) and water scarcity (WS), the yearly analysis does not capture the seasonality of water stress that is well highlighted by the monthly analysis. In fact, the yearly WAPC, WAGPC and WS show that most of the area is not under any stress or shortage of water (see the first map in Figure S 2). Instead, the in-depth analysis of monthly water balance demonstrates that this is mostly due to the extremely high values of water availability during the rainy season, in particular in the months of July and August. By contrast, the months outside of the rainy season are characterized, especially for Darfur and for the Komadugu-Yobe river basin, by low water availability and absolute water scarcity, i.e. values lower than 500m^3/capita (Figure S 2). Roughly half of the study area is in water scarcity conditions for at least half of the year (Figure 1). The GW presents no particular seasonality, since it is defined for the sole growing season of the area's main food crops. These indicators are investigated in relation to conflict via two different statistical analyses. WS is used as a covariate in the spatial econometrics model, while the other six indicators in Table 1 are clustered and combined to build up patterns of environmental stress in the conflict points analysis.

Environmental patterns as support to case study analyses: the case of the Boko Haram territorialization in 2015

In 2015, Boko Haram related conflicts are mostly located in Nigeria but represent also high shares of conflict events in Cameroon, Niger, and Chad. They mostly follow the abundance-scarcity pattern ('Boko Haram 2015' graph in Figure S 3), a pattern that puts the focus on human interdependencies from water resources rather than on resource availability. All considered water availability indicators are high when taken as volumes and low when taken as per capita values, showing how the most proximate factor in the dynamic is population density variation. In a dynamic way, areas that have high biophysical availability of water resources attract increased human pressure (and competition) which, given the low governance capability, eventually leads to low per capita water availability values. This mechanism characterized the violent stage of the Boko Haram insurgency in Nigeria and its spatial spillover towards areas politically not exposed to conflict, testifying the strategy shift from ruthless violence to territorial control[48].Yet, while Cameroon conflicts also follows the abundance-scarcity pattern, no pattern is detected in Chad and Niger conflicts (see 'Niger 2015' and 'Cameroon 2015' graphs in Figure S 3 for comparison). Cameroon is politically the most stable country in the area: only 12% of the events taking place in Cameroon involved Cameroonian actors. Conflicts connected to Boko Haram in Cameroon are in an area with high agricultural potential and very similar hydrological pattern as the Nigerian area where Boko Haram conflicts started. For instance, the Maga Dam and the fertile Waza-Logone floodplains are located in the Cameroonian part of the study area[61]. This suggests that the Boko Haram conflict spillover to Cameroon is one uncommon, but significant, case where a specific pattern of water availability played a role. Instead, in the case of Chad and Niger, the expansion of Boko Haram from Nigeria, affecting mostly the lake shores, happened for predominantly strategic reasons[86], a consequence of Boko Haram looking for shelter in the rural areas and islands around Lake Chad, which are beyond the reach of military forces [48]. Therefore, the absence of a pattern is consistent both with the history of the conflict dynamic and with the spatial econometric analysis. Moreover, it may testify that even conflict dynamics related in some way to particular processes involving water resources can spill over to neighbouring areas in which different processes take place, or, once again, it may show that apparently weak spatial econometric model correlations stand for relevant, but not universal, socio-environmental mechanisms. Yet, land seizure and control, together with the presence of refugees in the same region, may put additional pressure on natural resources on the Lake Chad shores, in particular on land[38,87].

Supplementary Figures

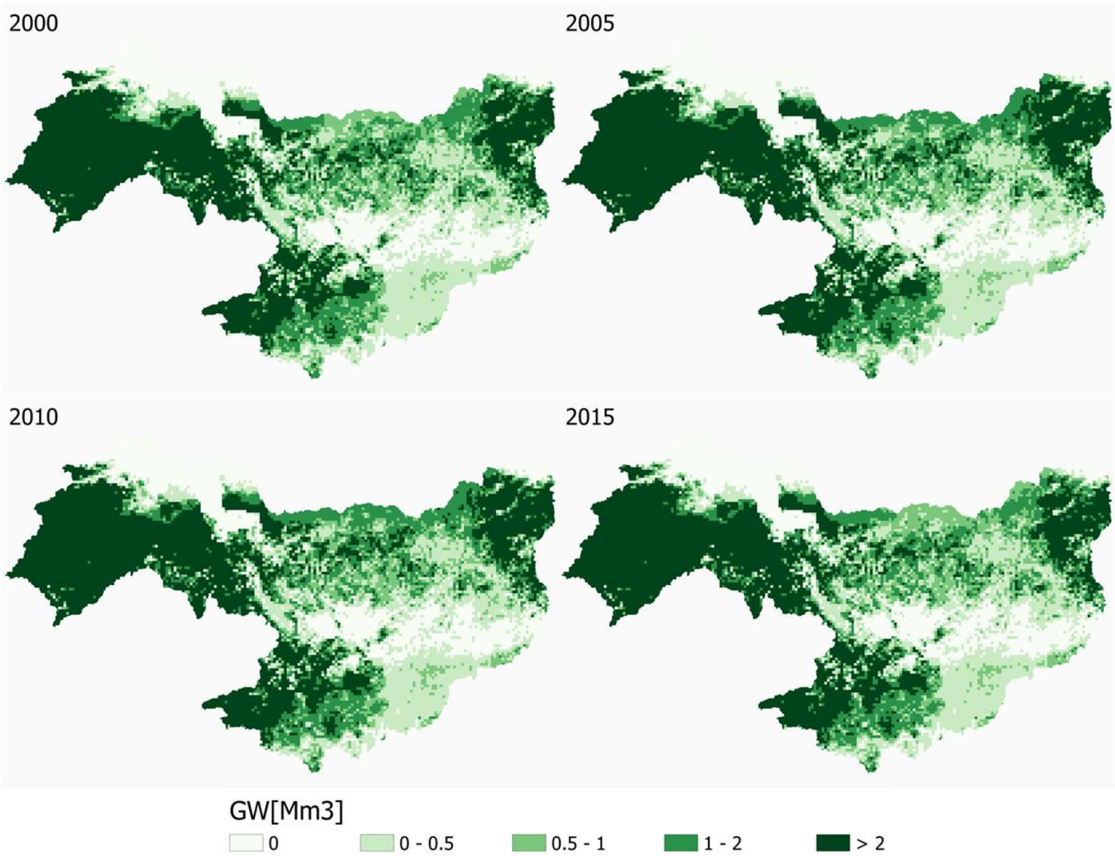

Figure S 1. Water demand for main crops during growing season in 2000, 2005, 2010, 2015. Basin outlines retrieved from [77]

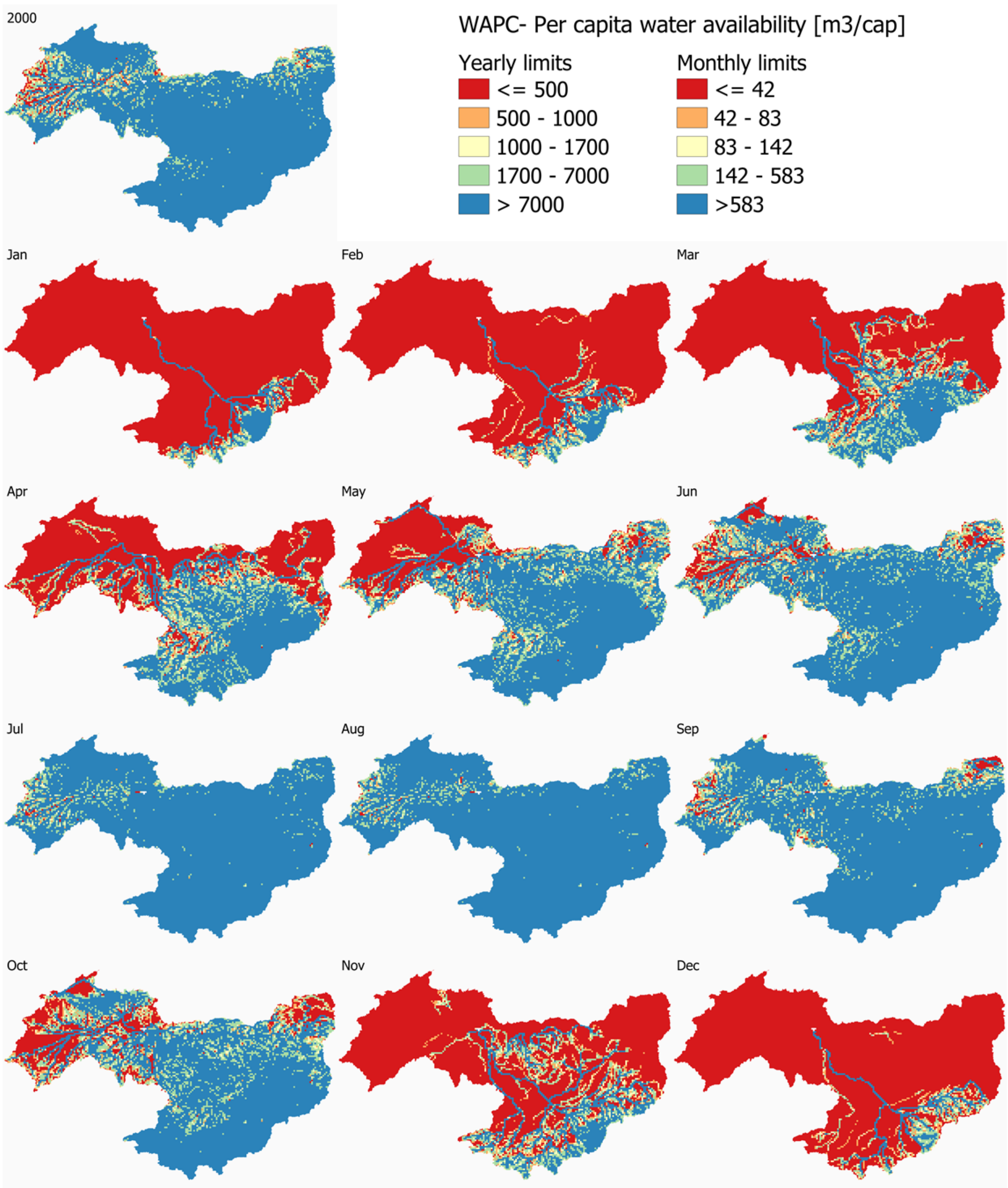

Figure S 2. Yearly and monthly WAPC values for the year 2000. Basin outlines retrieved from [77]

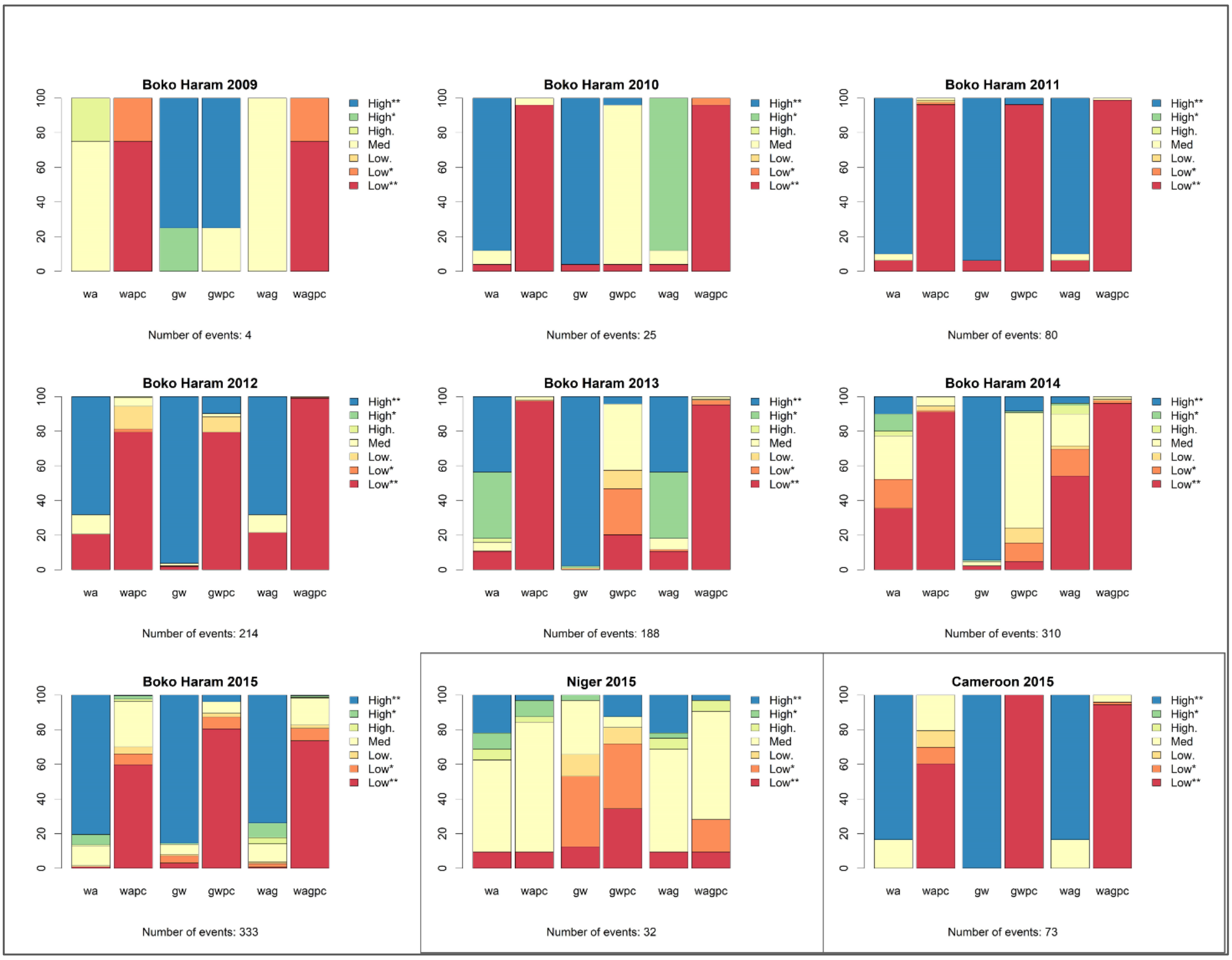

Figure S 3. Indicator clustering combinations for different conflict subsets. Each graph represents a year-country subset of conflict event. The year and the country are reported in the title, whereas the subset size is reported in the graph subtitle, below the graph. Each vertical bar represents the distribution among hotspot and coldspot classes of one indicator. The indicator acronym is reported below the bar. The color scheme is built on different significance levels in the same way as in Figure 4 (**=99% significance, *=95% significance, .=90% significance).

Supplementary tables

Excel file Supplementary.xlsx.