A Bayesian Spatio-Temporal Model for the Climate-Conflict Nexus in Africa

Caterina Conigliani Valeria Costantini
Elena Paglialunga Andrea Tancredi

MEMOTEF 26-6-2023

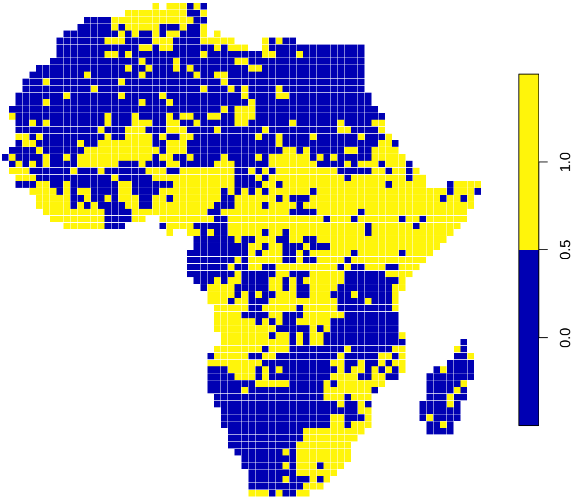
The data set

- Spatio-temporal data comprising the armed conflicts on the entire African continent based on a grid of 2,653 cells at 1 degree resolution (each cell covers an area of around 110x110 km).
- The yearly database covers the time span 1990-2016
- Information on violent events is extracted from the Uppsala Conflict Data Project - Georeferenced Event Dataset (UCDP). The database includes only events with at least one battle-related death

- Covariates: population, GPD, Gini index, forest (0/1), desert (0/1), (forest 0/1), city(0/1), ethnic group... and SPEI
- The Standardized Precipitation Evapotranspiration Index (SPEI) measures the onset, duration, and magnitude of drought/flood conditions with respect to normal conditions.
  - Positive values: excess of floods
  - Negative values excess of drought
- For the period 2017-2050 we have the SSP scenarios[1]

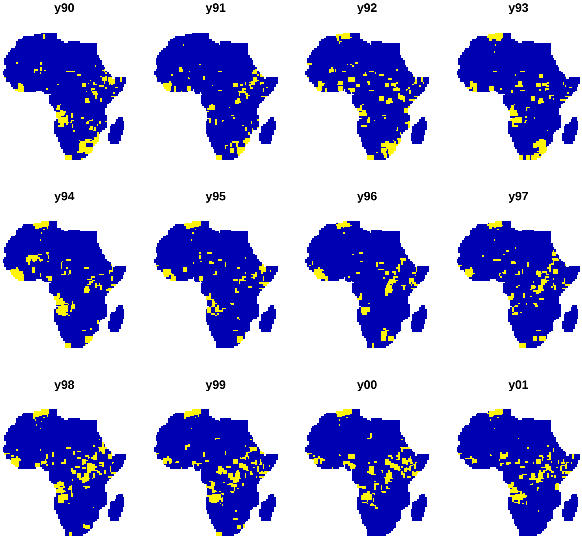
[1] Shared Socioeconomic Pathways (SSPs) are scenarios of projected socioeconomic global changes up to 2100

Cells with at least a conflict 1990−2016

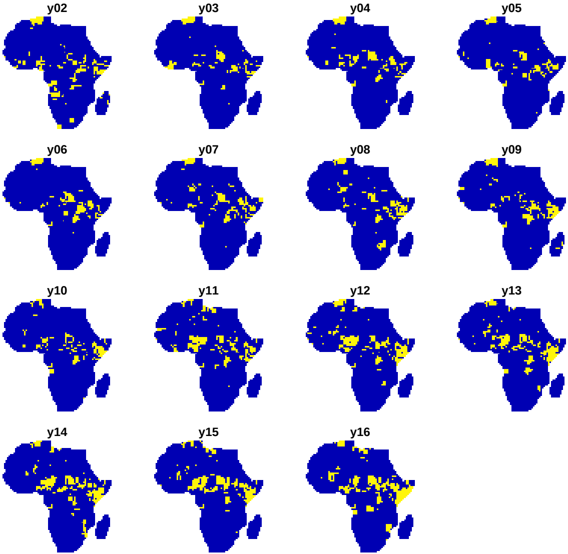


42.5% of the cells had at least a conflict during the period 1990-2016

Temporal evolution 1990-2001



Temporal Evolution 2002-2016



Aim of the study

- To investigate the implication that climate and socio-economic variables may have on violence
- Accounting for spatial and temporal dependence

Aim of the study

- To investigate the implication that climate and socio-economic variables may have on violence
- Accounting for spatial and temporal dependence

Besag model

Let η_i be a variable for cell i. A Gaussian Markov random field can be specified by the conditional distributions $\eta_i | \{\eta_j\}_{j \neq i}$

A common choice is to assume

$$\eta_i | \{\eta_j\}_{j \neq i} \; \sigma^2 \sim N\left(\tilde{\eta}_i, \; \frac{\sigma^2}{n_i}\right)$$

where

$$\tilde{\eta}_i = \frac{1}{n_i} \sum_{j \sim i} \eta_j$$

denotes the mean of the n_i spatially neighbouring cells of cel i, and σ^2 is an unknown variance parameter[2]

[2]Queen adjacency matrix for conflict data (a single point is enough to share a boundary)

- The Besag model is not proper
- There are linear combinations of the variables that have infinite variance or zero precision...this is not allowed in a proper distribution.
- In the Besag model it is caused by the fact that the conditional distributions give no information about the mean
- The problem is that it only accounts for similarities between regions
- The solution is to add an i.i.d. random effect in each region (a random intercept)[3]

[3] BYM model from Besag,York and Mollie'

- For instance, suppose that y_{it} is the indicator variable for the conflict in cell i, year t
- We can suppose that

$$y_{it} \sim \text{Bernoulli}(\lambda_{it})$$

  where

$$\eta_{it} = logit(\lambda_{it}) = \mu + u_i + v_i + f(c_{it})$$

  - Structured/spatial component u
  - Unstructured component v (i.i.d cell effect)
  - $f(c)$ is the non-linear effect of a covariate c

- The model in R-INLA

```
mgr1=inla(ncpos~1+eth_group+l1lnpop_ssp2+ d_city+d_desert+speiext+
               gdp_pcvar_ssp2+l1gini2_ssp2+l1t45mvy_mean+
               f(id,model="bym",graph=B),
               family="binomial",data=mydata)
```

```
Time used:
    Pre = 1.41, Running = 8.4, Post = 0.131, Total = 9.95
Fixed effects:
                  mean    sd 0.025quant 0.5quant 0.975quant   mode kld
(Intercept)     -7.082 0.261     -7.596   -7.081     -6.571 -7.080   0
eth_group        0.124 0.040      0.046    0.124      0.202  0.124   0
l1lnpop_ssp2     0.099 0.011      0.078    0.099      0.119  0.098   0
d_city           0.732 0.096      0.544    0.732      0.921  0.732   0
d_desert        -0.500 0.270     -1.028   -0.500      0.030 -0.500   0
speiextTRUE      0.173 0.043      0.088    0.173      0.258  0.173   0
gdp_pcvar_ssp2 -3.898 0.264     -4.415   -3.898     -3.382 -3.898   0
l1gini2_ssp2     2.297 0.232      1.842    2.296      2.752  2.296   0
l1t45mvy_mean    0.525 0.050      0.427    0.525      0.624  0.525   0

Random effects:
  Name    Model
    id BYM model

Model hyperparameters:
                                         mean       sd 0.025quant 0.5quant
Precision for id (iid component)     2221.390 1221.185    531.053 1930.154
Precision for id (spatial component)    0.096    0.004      0.088    0.096
                                     0.975quant     mode
Precision for id (iid component)       5210.870 1455.712
Precision for id (spatial component)      0.103    0.097

Marginal log-Likelihood:  -13581.99
```

Spatio-Temporal model

We can add a temporal effect by assuming that

$$\eta_{it} = logit(\lambda_{it}) = \mu + u_i + v_i + \gamma_t + \phi_t + f(c_{it})$$

where

- γ_t represents the temporally structured effect, modeled dynamically as a random walk
$$\gamma_t | \gamma_{t-1} \sim N(\gamma_{t-1}, \sigma^2_\gamma)$$
- ϕ_t represents the temporally unstructured effect (i.i.d.)

```
mgr2=inla(ncpos~1+eth_group+l1lnpop_ssp2+ d_city+d_desert+speiext+
       gdp_pcvar_ss<<<p2+l1gini2_ssp2+l1t45mvy_mean+
       f(id,model="bym",graph=B)+
       f(year,model="rw1")+
       f(year2,model="iid"),
     family="binomial",data=mydata)
```

```
    Pre = 1.46, Running = 31.6, Post = 0.131, Total = 33.2
Fixed effects:
                  mean    sd 0.025quant 0.5quant 0.975quant   mode kld
(Intercept)     -8.079 0.369     -8.819   -8.073     -7.372 -8.057   0
eth_group        0.119 0.039      0.043    0.119      0.196  0.119   0
l1lnpop_ssp2     0.217 0.029      0.163    0.215      0.276  0.211   0
d_city           0.665 0.096      0.477    0.665      0.855  0.665   0
d_desert        -0.060 0.278     -0.604   -0.061      0.485 -0.061   0
speiextTRUE      0.110 0.046      0.019    0.110      0.201  0.110   0
gdp_pcvar_ssp2 -3.439 0.295     -4.018   -3.439     -2.860 -3.439   0
l1gini2_ssp2     1.740 0.249      1.251    1.740      2.227  1.741   0
l1t45mvy_mean    0.579 0.056      0.469    0.579      0.689  0.579   0

Random effects:
  Name    Model
    id BYM model
   year RW1 model
   year2 IID model

Model hyperparameters:
                                        mean      sd 0.025quant 0.5quant
Precision for id (iid component)     653.302 777.305    131.615  421.153
Precision for id (spatial component)   0.106   0.007      0.094    0.105
Precision for year                    50.443  45.780      8.854   37.264
Precision for year2                   15.597  16.763      2.777   10.592
                                     0.975quant    mode
Precision for id (iid component)        2611.30 230.869
Precision for id (spatial component)       0.12   0.104
Precision for year                       171.79  21.295
Precision for year2                       59.05   5.908

Marginal log-Likelihood:  -13496.49
```

Time-space interaction

The time effect and the spatial effect can also interact

$$\eta_{it} = logit(\lambda_{it}) = \mu + u_i + v_i + \gamma_t + \phi_t + \delta_{it} + f(c_{it})$$

Four types of interactions

- Type 1: interaction between the unstructured effects v_i and ϕ_t. Time and space effects are still independent
- Type 2: interaction between the unstructured spatial effects u_i and the structured temporal effect γ_t. Each cell has a temporal correlation structure, but neighboring cells have independent temporal correlations

- Type 3: Interaction between the structured spatial effects v_i and the unstructured temporal effect ϕ_t. The spatial trends are different from year to year, but they are independent
- Type 4: Interaction between the structured spatial effects v_i and the structured temporal effect γ_t. The spatial trends are different from year to year, but they are dependent

Type 1 inetraction

```
mgr3=inla(ncpos~1+eth_group+l1lnpop_ssp2+ d_city+d_desert+speiext+
            gdp_pcvar_ssp2+l1gini2_ssp2+l1t45mvy_mean+
            f(id,model="bym",graph=B)+
  f(year,model="rw1")+
  f(year2,model="iid")+
  f(idt,model="iid"),
family="binomial",data=mydata)
```

```
Time used:
    Pre = 1.86, Running = 117, Post = 1.1, Total = 120
Fixed effects:
                  mean    sd 0.025quant 0.5quant 0.975quant   mode kld
(Intercept)     -8.095 0.369     -8.847   -8.084     -7.401 -8.059   0
eth_group        0.120 0.039      0.043    0.120      0.196  0.120   0
l1lnpop_ssp2     0.219 0.029      0.168    0.216      0.278  0.211   0
d_city           0.666 0.097      0.477    0.666      0.855  0.665   0
d_desert        -0.049 0.277     -0.592   -0.050      0.495 -0.050   0
speiextTRUE      0.110 0.046      0.019    0.110      0.200  0.110   0
gdp_pcvar_ssp2 -3.482 0.295     -4.060   -3.482     -2.904 -3.482   0
l1gini2_ssp2     1.726 0.249      1.237    1.726      2.212  1.727   0
l1t45mvy_mean    0.573 0.056      0.463    0.573      0.683  0.573   0

Random effects:
  Name    Model
    id BYM model
   year RW1 model
   year2 IID model
   idt IID model

Model hyperparameters:
                                            mean       sd 0.025quant 0.5quant
Precision for id (iid component)     9.63e+02 1.72e+03     36.474 4.63e+02
Precision for id (spatial component) 1.05e-01 6.00e-03      0.093 1.04e-01
Precision for year                   2.95e+03 9.52e+03     54.692 9.06e+02
Precision for year2                  5.58e+00 2.56e+00      1.702 5.24e+00
Precision for idt                    1.81e+04 1.66e+04   2459.975 1.34e+04
                                     0.975quant     mode
Precision for id (iid component)       5.03e+03   88.895
Precision for id (spatial component)   1.18e-01    0.103
Precision for year                     1.87e+04  119.038
Precision for year2                    1.14e+01    4.282
Precision for idt                      6.22e+04 6654.114
```

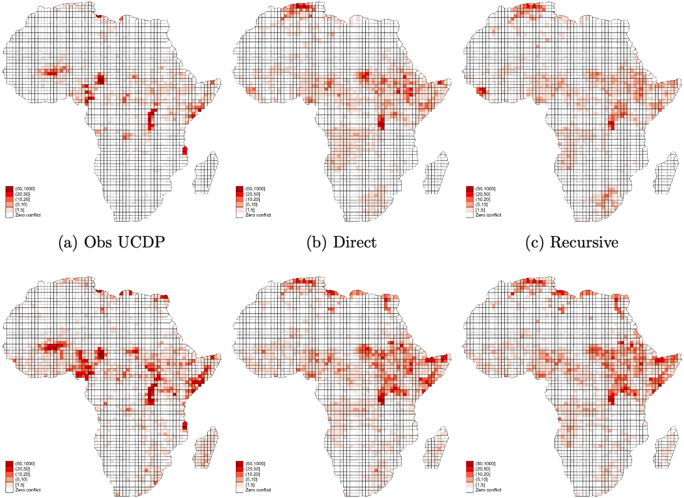
Discussion

Things done, to do, and to understand

- Zero-inflated negative binomial for the number of conflicts (similar model, done)
- Better modeling of the spei effect (extreme spei+sign spei+extreme:sign spei vs |spei|+sign spei+sign; |spei|) (under evaluation)
- Type 4 interaction (to do, computational problems!)
- Model comparison (DIC, WAIC...are provided by INLA, but still to do with once the type 4 interaction is obtained)
- Forecasting (to understand how to do with INLA...but we are there)

Predictions under a non-Bayesian model

Figure 2: Alternative forecasting methods (year 2020)



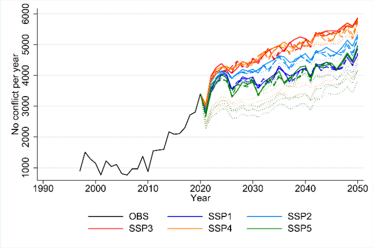
(a) Obs UCDP (b) Direct (c) Recursive

(d) Obs ACLED (e) Direct (f) Recursive

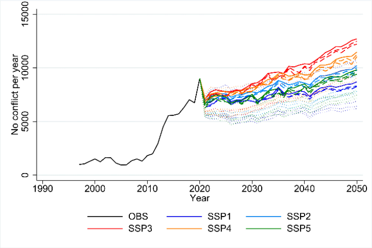
Note: Number of conflicts events are expressed as three-year average (2017-2020).

Predictions under a non-Bayesian model

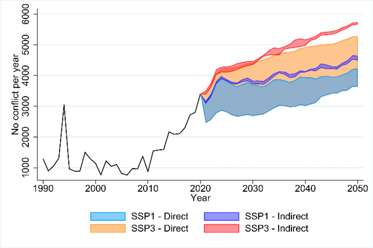
Figure 3: Conflicts projections by 2050 under SSPs (UCDP and ACLED)



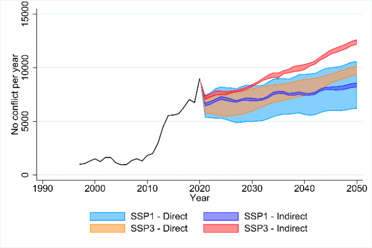
(a) SSPs with different models (UCDP)



(b) SSPs with different models (ACLED)



(c) SSP1 and SSP3 (UCDP)



(d) SSP1 and SSP3 (ACLED)