THE UNIVERSITY *of* EDINBURGH

School of Geosciences

Dissertation
for the Degree of
MSc in Geographical Information Science

Sarah Katherine Wingrove

August 2025

The Statement of Copyright and Originality

I declare that this dissertation represents my own work, and that where the work of others has been used it has been duly accredited. I further declare that the length of the components of this dissertation is 4,723 words for the Research Paper and 3,750 words for the Technical Report.

I agree that this dissertation and associated electronic documents, web pages, data, files and computer programs can be retained by the University. ~~YES~~ / NO

I agree that, with the permission of my supervisor(s) or the Programme Director, these materials be made available for the purposes of preparing a publication. ~~YES~~ / NO

I agree that, with the permission of my supervisor(s) or the Programme Director, these materials can be used within the University of Edinburgh for continued research. ~~YES~~ / NO

Signature

x 

Date: 05 August 2025

Acknowledgements

I am deeply grateful to my project supervisor, Dr. Gary Watmough, for his unwavering guidance and support throughout the ideation, analysis, and writing process. I also thank the geosciences professors and my cohort for their encouragement and camaraderie throughout the year. To my family and friends, your faith in me never wavered, even when my own did, and for that, I am forever thankful.

PART I

Research Paper

Roads, Riches, and Residuals: A Multilevel Analysis of Subnational Determinants of Violent Conflict in Syria (2011-2019)

Abstract

This paper investigates the subnational determinants of violent conflict in Syria from 2011 to 2019. District-level, annual counts of total estimated fatalities, total conflict events, state-based violence, non-state violence, and one-sided violence are combined with two structural indices, (1) infrastructure (roads, population density, urban/built up land cover) and (2) land cover (croplands, open shrublands, barren/sparsely vegetated, savannas, cropland natural vegetation mosaics and grasslands), each derived via Principal Component Analysis (PCA). Three governorate-level human development indices (income, education, health) are residualised on the other two to isolate unique effects.

Mixed effects negative binomial models with governorate random intercepts, and a constant-only zero-inflation term for non-state violence, estimate how infrastructure, land cover, and human development imbalances predict conflict intensity, controlling for year fixed effects. Spatial clustering is assessed via Moran's I and Local Indicators of Spatial Association (LISA), confirming clustering of violence and identifying district-level hot spots. Results indicate that higher road density, urban cover, and population concentration are associated with significantly increased state-based and one-sided violence, while higher-than-expected education attainment may influence non-state violence against civilians.

By modelling district-level conflict with governorate-level heterogeneity, this analysis demonstrates how disaggregated, multilevel models combined with spatial autocorrelation diagnostics can uncover the structural drivers of violence and highlight the trade-off between within- and between-unit inference when human development data is available only at coarser scales.

1. Introduction

Since the start of the Syrian civil war in March 2011, the country has undergone tremendous amounts of suffering, widespread destruction, and mass displacement, reshaping Syria's social, political, and geographical realities. The conflict has displaced around half of Syria's population and has drawn in multi-sided foreign intervention (Blanchard, 2023). According to the United Nations Development Programme, the conflict "has undone nearly four decades of economic, social, and human process" (UNDP, 2025). Syria's Human Development Index has fallen below its 1990 level, plunging around 90% of Syrians into poverty and causing around 618,000 deaths, illustrating what literature refers to as development in reverse (UNDP, 2025; Collier et al., 2003).

Violence has been geographically uneven, clustering in hot spots at particular times, and this spatial structure is formally quantified via Moran's I and LISA. Night-time lights analysis shows urban centres lost nearly 80% of luminosity due to infrastructure damage and mass displacement (Li et al., 2017). Such localised patterns reflect broader trends in civil war studies, where violence tends to cluster in localised hot spots shaped by structural conditions (Nemeth et al., 2014). Even though civil wars often only directly affect a portion of a country's territory, humanitarian and economic consequences can radiate far beyond these areas (Buhaug et al., 2009; Schutte and Weidmann, 2011; O'Loughlin and Witmer, 2011). As Buhaug et al. (2009) argue, relying on national-level data can obscure localised dynamics, making subnational analysis important in understanding the spatial and temporal spread of violent civil conflict.

Research Objectives and Questions

This study combines district-level counts of conflict events with two PCA-derived indices (infrastructure; land cover) and three governorate-level human development indices (income, health, education), residualised on each other to isolate unique effects, to unpack the spatial and developmental drivers of violence. The research questions are:

- How do subnational determinants shape the intensity and type of violence across Syrian districts?
- Do different types of violence (state-based, non-state, one-sided) respond to different subnational predictors?
- How do spatial patterns of violence emerge and shift over the 2011-2019 period?

2. Background

Geography of Syria

Syria, officially the Syrian Arab Republic, occupies a strategic position in West Asia at the eastern end of the Mediterranean Sea, bounded by Turkey to the north, Iraq to the east, Jordan to the south, and Lebanon and Israel to the west and southwest (Figure 1). Syria's total area is around 187,437 km² and ranks 89th in the world by size (Central Intelligence Agency, 2025). The country's terrain transitions from a narrow, fertile coastal plain and western mountain belt to an arid plateau and the vast Syrian Desert in the east, interspersed by the Euphrates River valley (Central Intelligence Agency, 2025).

Context of the Syria Conflict (2011–2019)

The Syrian conflict began with peaceful protest in the south near Daraa which morphed into

violent repression by the Assad regime leading to the creation of the Free Syrian Army and armed-rebellion by mid-2011 (Ash and Obradovich, 2019). By 2013, the country had plunged into full-scale civil war. Insurgent groups took control of territory as they clashed with the regime (Blanchard, 2023). As the conflict progressed, foreign intervention became increasingly prominent. Iran, Hezbollah, and Russia supported the Assad regime, while the United States, Turkey, Gulf states, and several European countries backed opposition groups (Blanchard, 2023). The Islamic State's establishment of a caliphate across the Levant in 2013–14 further destabilised the situation. Syria's territory became fragmented into zones of control held by regime forces, rebel groups, ISIS, and Kurdish militias, transforming the conflict into a multi-front regional proxy war (Blanchard, 2023).

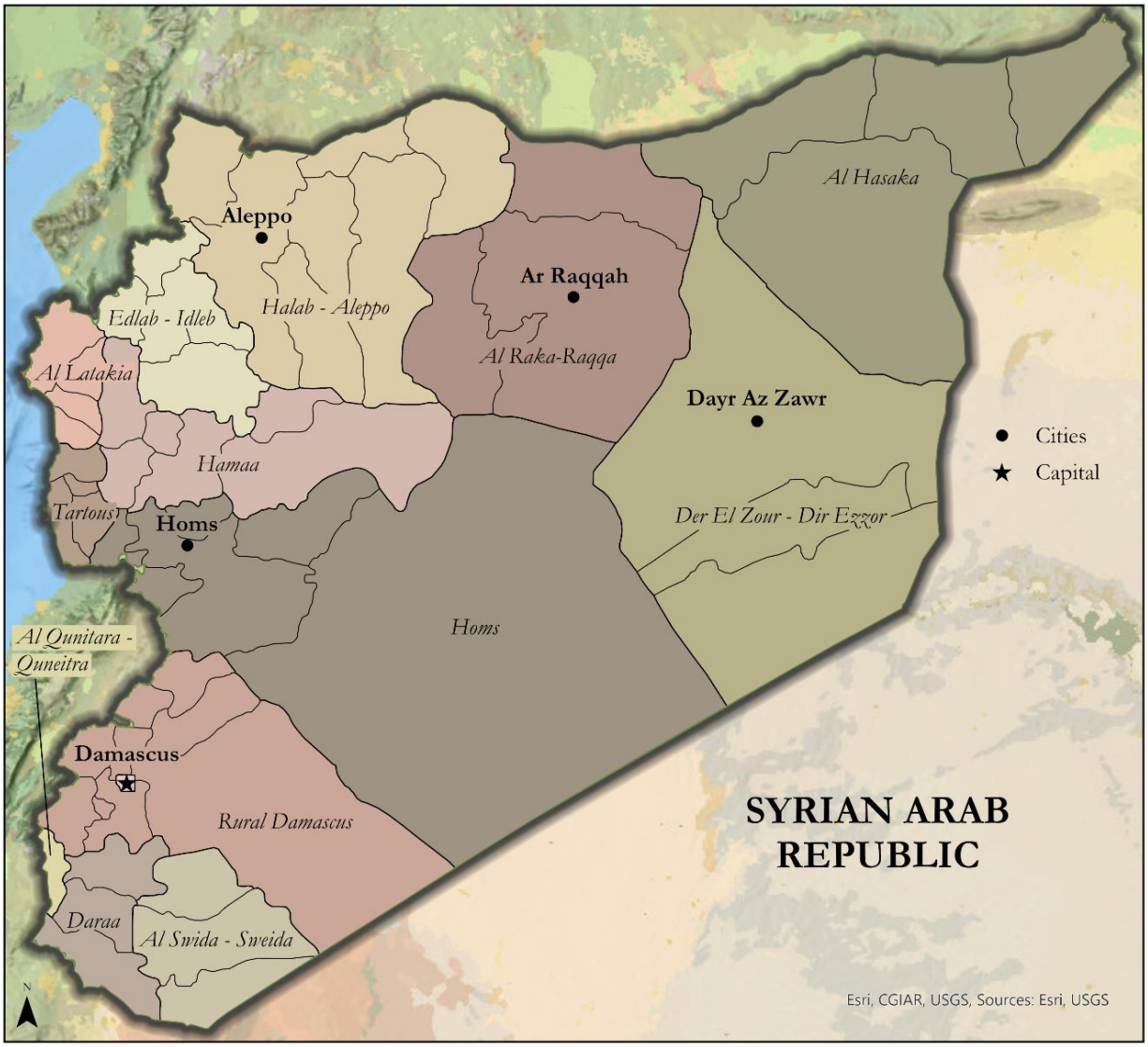

Figure 1. *Map of Syria showing governorate (colored, n = 14) and district boundaries (outlined, n = 60) and major cities. Administrative boundaries sourced from GADM (version 4.1).*

Structural Determinants and Disparities

Syria's pre-war development showed sharp regional disparities. Market-oriented reforms during the early 2000s, like cuts to fuel and fertiliser subsidies, tended to favour urban elites and deepened rural hardship (de Châtel, 2014). A severe drought from 2006 to 2010 also exacerbated

hardships by undercutting rain-fed agriculture and reducing yields and household incomes (Kelley et al., 2015; Gleick, 2014). As conditions worsened, an estimated one to 1.5 million people left farming communities for the outskirts of Syrian cities, and many fields were left uncultivated, although recent research finds that the extent of abandonment varied across regions (Dinç and Eklund, 2023).

Scholars disagree on how far climate stress contributed to Syria's uprising. Ash and Obradovich (2019) contend that the 2006-2010 drought undermined rain-fed farming, pushed many rural households to migrate internally and, by swelling populations in less-affected districts, intensified local grievances. By contrast, Selby et al. (2017, p. 233) argue there is "no clear and reliable evidence" that anthropogenic climate change drove the drought or that drought migration was on a scale sufficient to spark the conflict; instead, they point to corruption, authoritarian rule and long-standing urban-rural inequalities.

3. Literature Review

Subnational variation in civil war violence has become a growing area of interest in political science, geography, and development studies. While much of the early literature on civil conflict focused on national-level determinants such as regime type, ethnic fractionalisation, or economic growth, more recent research emphasises the importance of intra-state heterogeneity (Buhaug and Rød, 2006; Raleigh and Hegre, 2009).

3.1 Type of Violence and Conflict Diffusion

Sundberg and Melander (2013) explain that the Uppsala Conflict Data Program Georeferenced Data Event Dataset (UCDP GED) recognises three mutually exclusive forms of organised violence: state-based armed conflict, fought between at least one government and another organised actor; non-state conflict, involving clashes between armed groups with no state party; and one-sided violence, in which a state or organised group deliberately targets civilians. State-based armed conflict usually centres on incompatibilities over government or territory (UCDP, 2024), whereas non-state violence is most common in areas of limited or fragmented state authority that create governance vacuums (Müller-Crepon et al., 2021; Risse, 2013). One-sided violence is typically employed to suppress dissent, terrorise communities, or sever insurgent support networks (Valentino et al., 2004).

All three forms of violence occurred concurrently in the Syrian civil war. State-based violence between the Syrian government and various rebel groups accounted for most battle-related deaths and was marked by urban sieges and intense aerial bombardments in cities like Aleppo (2012 to 2016) and Homs (2011 to 2014) (Human Rights Watch, 2016). Syria also saw considerable non-state violence, including infighting among rebel factions and the territorial expansion of ISIS between 2013 and 2017 (Woodrow Wilson International Center for Scholars, n.d.). Both the Assad regime and armed opposition groups were responsible for widespread one-sided violence against civilians, including the use of barrel bombs, cluster munitions, indiscriminate shelling, and mass executions (UCDP, 2024; SNHR, 2024).

Pettersson and Wallensteen (2015) stress that reliable analysis of global conflict trends depends on treating state-based conflict, non-state conflict, and one-sided violence as separate categories, because each involves different actors. This study investigates whether infrastructure, development, and land use have distinct effects across different types of violence.

3.2 Human Development, Inequality, and Conflict

Empirical research finds mixed evidence on how human development and inequality relate to civil war. While cross-country studies consistently show that poverty and slow economic growth raise the risk of conflict onset (Collier et al., 2003), results for inequality are less uniform. Two broad explanations dominate: an opportunity perspective, which holds that rebellion is more likely when potential recruits face low opportunity costs or can profit from lootable resources (Collier and Hoeffler, 2004), and a grievance perspective, grounded in Gurr's theory of relative deprivation, which stresses perceived political, economic or cultural exclusion (Gurr, 2011).

Syria entered the conflict with marked regional disparities and high unemployment (de Châtel, 2014). Researchers point to horizontal inequalities, meaning persistent gaps in wealth, services and political influence between regions and social groups (Stewart, 2008), as an important source of grievance. Initial protests broke out in marginalised, rural Sunni-majority areas such as Daraa, Hama and Idlib (Leenders and Heydemann, 2012; Hinnebusch, 2012). The concept of a conflict trap further suggests that war and underdevelopment become mutually reinforcing; conflict destroys human capital and institutions, as seen in Syria's HDI collapse, which increases the likelihood of persistent violence (Collier et al., 2003, p. 53).

The relationship between development and conflict is complex and not necessarily linear. Classical modernisation theories suggest that higher development reduces violence by increasing opportunity costs, enhancing institutional capacity, and mitigating grievances (Collier and Hoeffler, 2004). However, partial modernisation, characterised by increases in education or urbanisation without political inclusion, may heighten instability by intensifying grievances and mobilisation (Goldstone et al., 2010; Urdal, 2006).

Education specifically illustrates these mixed effects. While education generally reduces conflict at the individual level (Østby, 2008), it may also mobilise marginalised groups (Cederman et al., 2013). In authoritarian contexts like Syria, educated regions perceived as political threats can face increased violence.

3.3 Land Use and Land Cover

Terrain and land cover have long been theorised as important environmental determinants of conflict. Fearon and Laitin (2003) argued that mountains and forests offer insurgents sanctuary, enabling protracted guerrilla warfare. Subsequent studies refined this view, showing that accessibility, land productivity, and drought sensitivity, ecological diversity all interact to shape conflict risk (Buhaug et al., 2009; von Uexkull et al., 2016).

The spatial distribution of urban areas, agricultural land, forests, and deserts can shape both the opportunities for conflict and the tactical behaviour of armed actors. Urbanisation tends to

concentrate population, infrastructure, and political authority, making cities strategic targets in civil wars. O'Loughlin and Witmer (2011) geocoded 14,177 violent events from 1999 to 2007 in the North Caucasus and, using SaTScan hot-spot detection plus geographically weighted regression, demonstrated that attacks clustered in forested rural settlements while communities located farther from primary roads experienced markedly lower levels of violence. In a recent study, Liu et al. (2025) used the Patch-generating Land Use Simulation (PLUS) model to simulate Syria's land use and land cover change during the civil war. They concluded that due to displacement between 2011 and 2015, crop and tree cover increased as crop cultivation moved to safer zones, decreasing bare ground (Liu et al., 2025). Between 2016 and 2020, crop growth slowed, bare ground spread in moderate war zones, and built-up land shrank sharply before modest reconstruction (Liu et al., 2025). To further test the impact of land cover on different types of conflict, this study summarises district-level land cover proportions via the first principal component, yielding a single PCA-derived land cover index for use in subsequent models.

3.4 Infrastructure, Accessibility, and Conflict Diffusion

Infrastructure represents a double-edged variable in conflict analysis. Roads and transportation networks affect how conflict spreads and how easily state and non-state actors operate. On the one hand, well-developed infrastructure enables governments to project power by facilitating the deployment of troops, supplies, and administrative reach. On the other hand, transport connectivity can also facilitate insurgent movement, conflict spread, and the targeting of more locations.

From an opportunity-based perspective, infrastructure tends to be the logistical backbone for insurgent activity. Fearon and Laitin (2003) argue that rebellions are more likely to emerge in regions where state authority is weak and insurgents can operate freely. Roads that link peripheral ethnic regions to the national capital tend to lower conflict risk by boosting state reach, whereas dense internal road networks within those same regions can facilitate rebel mobilisation and increase violence; hence the overall effect of roads depends on the local balance of control (Müller-Crepon et al., 2021).

Infrastructure also shapes patterns of civilian exposure to violence. Focusing on North and West Africa, Walther et al. (2025) blend Armed Conflict Location and Event Data (ACLED) event data with the Global Roads Inventory Project and show that about 70% of all recorded violent incidents between 2000 and 2024 took place within one kilometre of a road; the share was 86 percent in 2011 but fell to 61% by 2023 as jihadist groups shifted operations away from main transport corridors into more rural terrain.

4. Data and Methodology

This study combined five datasets to find relationships between predictors, conflict severity, and type: georeferenced conflict events, land cover, population density, road density, and socioeconomic variables. All spatial layers were harmonised to the same coordinate system (WGS 1984 UTM Zone 37N) and aggregated by district-year units to allow for consistent temporal and spatial comparison. Figure 4.1 gives an overview of the steps taken during preprocessing and analysis. Refer to Part II: Technical Report, Section 2 for an in-depth explanation of data source methodology.

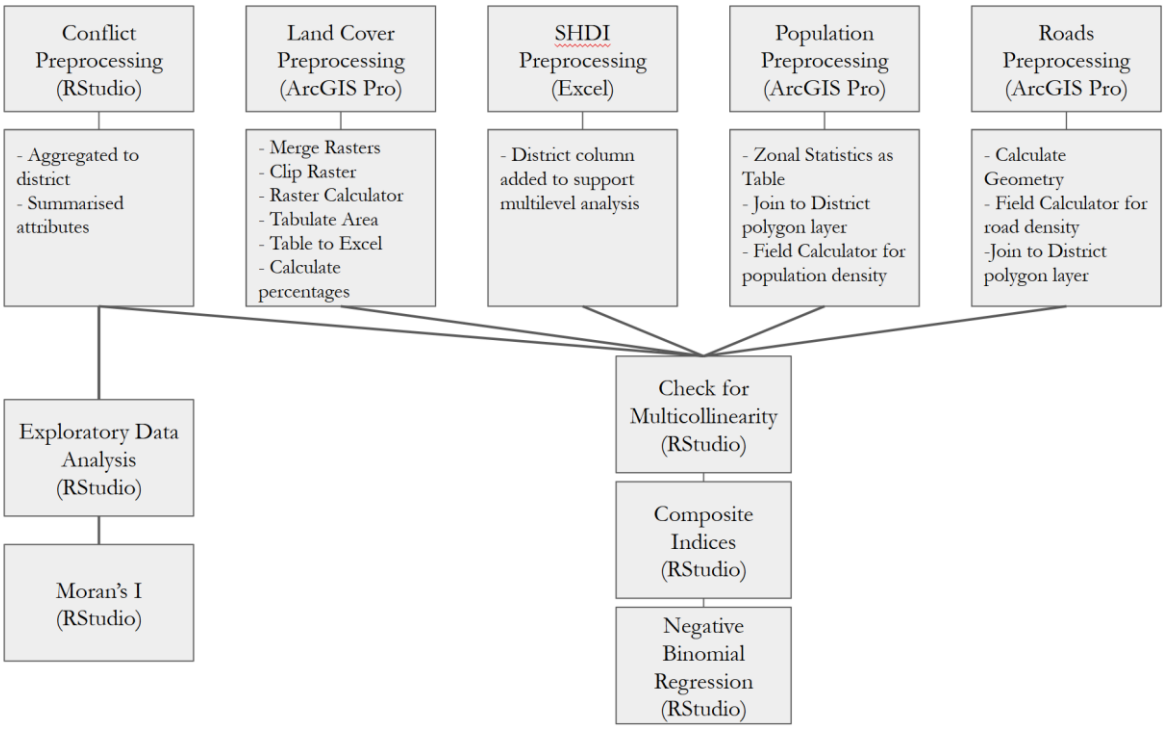

Figure 4: *Methodological workflow overview used in this study.*

4.1 Geospatial Data Acquisition and Preprocessing

4.1.1 Conflict Event Data

The violent conflict data for this study came from the Uppsala Conflict Data Program's Georeferenced Event Dataset (UCDP, 2024), which records each incident with latitude-longitude coordinates, violence type, actors, date, sources, and estimated fatalities. All events were spatially joined to second-level administrative boundaries (GADM, 2022, v4.1) in ArcGIS Pro (Esri, 2025, v3.5). Incidents located outside recognised district polygons (n = 405) were dropped from the dataset. A total of 82,989 events were imported into R (R Core Team, 2024, v4.2.2) and cleaned them with *dplyr* (Wickham et al., 2023, v1.1.4), *tidyr* (Wickham et al., 2024, v1.3.1), and *readr* (Wickham et al., 2024, v2.1.5). Event count data was aggregated by Syrian district (NAME_2) and year to generate five outcome variables: total conflict events (conflict_count), estimated fatalities (total_fatalities), state-based violence (state_based), non-state violence (non_state_based), one-sided violence (one_sided).

4.1.2 Land Cover and Infrastructure Indices

Land cover data was obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) MCD12Q1 annual raster product (Friedl and Sulla-Menashe, 2022), which classifies global land surface cover into 17 categories defined by the International Geosphere-Biosphere Programme (IGBP) at a spatial resolution of 500 metres. For each year, the raster datasets were clipped to Syria's border (GADM, 2022, v4.1). ArcGIS's Tabulate Area tool was used to compute pixel counts for each land cover class before calculating percentage values at a district level (Esri, 2025, v3.5).

Road density was calculated as total kilometres of road per square kilometre, using data from OpenStreetMap (OpenStreetMap Contributors, 2025). Population density was derived from WorldPop's annual UN-adjusted rasters, expressed as estimated persons per square kilometre. For detailed preprocessing procedures related to land cover, road density, and population density, refer to Part II: Technical Report, Section 3.

PCA was applied to z-standardised variables using R's *prcomp()* (R Core Team, 2024, v4.2.2) implementation to generate orthogonal indices that mitigate multicollinearity and reduce dimensionality (Jolliffe and Cadima, 2016). For the infrastructure variables (road density, population density and built-up land cover class) the first principal component (PC1) accounted for approximately 91% of total variance and was retained as the Infrastructure Index. A separate PCA on six MODIS land-cover classes (croplands, open shrublands, barren/sparsely vegetated, savannas, cropland natural vegetation mosaics and grasslands) yielded a PC1 that explained about 41% of variance, designated the land cover index. Only PC1 from each analysis was kept, as it accounted for the most significant proportion of shared variance among the input variables given its dominance in the variance structure (Filmer and Pritchett, 2001; Kassambara, 2017). The resulting component scores were mean-centred and standardized (z-score scaled) prior to inclusion as predictors in the multilevel negative binomial models.

4.1.3 Socioeconomic Indicators

Subnational development data came from the Global Data Lab's Subnational Human Development Index (SHDI), an adaptation of the UNDP Human Development Index (HDI) that recalculates income, education and health at the governorate level. Annual SHDI scores employ the same equal-weighted geometric-mean formula used in the UNDP HDI but substitute subnational (ADM-1) indicators for national ones, thereby enabling spatially disaggregated analysis (Smits and Permanyer, 2019; UNDP, 2024). To align the SHDI table with other datasets and permit multilevel modelling, a district-name column was appended so that each governorate-level record could be duplicated across its constituent districts.

Because the three SHDI components overlap conceptually and empirically, residualisation was used to reduce multicollinearity (García et al., 2019). Standardised education and health scores were each regressed on the standardised income score; the resulting residuals capture variation unique to education and health, respectively, thereby isolating their independent effects without inflating standard errors.

Alternative model specifications were also tested, including residualisation of income and health on education, as well as income and education on health. For a detailed discussion of these exploratory models and their outcomes, refer to Part II: Technical Report, Section 6.

4.2 Spatial Autocorrelation and LISA Cluster Analysis

The spatial structure of violence was evaluated using Global Moran's I and LISA. A district-level queen-contiguity weights matrix was constructed in R with the *spdep* package (Pebesma and Bivand, 2023, v1.3.10). Refer to Part II: Technical Report, Appendix 1 for all R packages and code used during this study's analysis.

Global Moran's I was computed for each violence outcome in each year (2011-2019) using 999 random-label permutations to obtain an exact, distribution-free p-value for the upper tail (Bivand and Wong, 2018). A significantly positive Moran's I indicates that high-violence districts tend to border other high-violence districts (and low with low), rather than exhibiting a random pattern (Moran, 1950).

LISA statistics were then calculated for every district, again with 999 permutations. Each local Moran's statistic yields a p-value testing whether the district-and its neighbours-form a hot spot (high-high), cold spot (low-low), or a spatial outlier (high-low or low-high) (Anselin, 1995). To control the false discovery rate arising from multiple local tests, p-values were adjusted via the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995), and only districts remaining significant after correction were displayed as clusters; non-significant districts are shown in grey (Part II: Technical Report, Section 4). All figures were produced using ggplot2 (Wickham, 2016, v3.5.1).

4.3 Multilevel Modelling

Negative binomial regression was selected to model annual district-level conflict outcomes (total fatalities, total events, state-based violence, non-state violence, and one-sided violence) because these outcomes exhibit overdispersion, where the variance exceeds the mean, which is typical with violent event data (Hilbe, 2011). To accommodate the hierarchical data structure of districts nested within governorates, each model incorporates a governorate-level random intercept, estimated via Laplace approximation in *glmmTMB* (Brooks et al., 2017, v1.1.7), which captures unobserved regional heterogeneity and avoids overstating precision (Gelman and Hill, 2007) Year fixed effects (2011 set as the reference) control for shocks common across all units.

A log link relates these predictors to the expected count μ, so that a one-standard-deviation increase in any z-score yields an incidence-rate ratio (IRR) of $\exp(\beta)$, interpretable as a $(\exp(\beta)-1) \times 100\%$ change in the outcome (Cameron and Trivedi, 2013). For non-state violence, which features excess zeros, a constant-only zero-inflation component (ziformula=~1) follows the mixture formulation of Lambert (1992).

Model adequacy and comparative fit were evaluated using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) (Burnham and Anderson, 2002) alongside simulated-quantile residual diagnostics for overdispersion and zero inflation from *DHARMa* (Hartig, 2024,

v0.4.7) and variance-inflation checks via the *performance* package (Lüdecke et al., 2021, v0.15.0) Although fixed-effects versions with governorate dummies sometimes achieved marginally lower AIC, mixed effects models consistently produced lower BIC and retained the ability to estimate time-invariant covariates effects (Wooldridge, 2021). All analyses were conducted in R 4.2.2 (R Core Team, 2024).

5. Results

This section presents findings from the statistical and spatial analyses of conflict patterns in Syria from 2011 to 2019. Key findings include:

- Conflict was spatially clustered, particularly in Aleppo, Idlib, and Damascus.
- Infrastructure was a key predictor of state-based and one-sided violence.
- Income consistently reduced all forms of violence, especially non-state violence.
- Ecological diversity deterred violence across all types.
- Education and health residuals had actor-specific effects.
- State-based violence remained steady over time, while non-state and one-sided violence declined after 2016.

5.1 Geographies of Violence: Spatial Clustering

In 2011, none of the conflict outcomes showed significant spatial clustering (Moran's I near zero). During 2012, all measures start to exhibit positive spatial autocorrelation: non-state violence rises from I = 0.18*** in 2012 to I = 0.54*** in 2019, and one-sided violence from I = 0.15** to I = 0.29*** over the same period (Table 5.1). For LISA cluster analysis and maps refer to Part II: Technical Report, Section 4. Figure 5.1 depicts the total number of conflict events by type of violence per district between 2011-2019.

Table 5.1. *Global Moran's I Statistics for Conflict Variables by Year (2011-2019)*

Year	Conflict Count	Estimated Fatalities	State-Based Violence	Non-State Violence	One-Sided Violence
2011	-0.03	-0.02	0.01	-0.04	-0.05
2012	0.17*	0.03	0.17*	0.18***	0.15**
2013	0.10	-0.02	0.10	0.46***	0.01
2014	0.05	-0.04	0.08	0.28***	0.15*
2015	0.07	-0.03	0.08	0.06	0.45***
2016	0.07	0.01	0.05	0.19**	0.37***
2017	0.16*	0.14*	0.17*	0.05*	0.26***
2018	0.18**	0.08	0.27***	0.11*	0.23***
2019	0.28***	0.24***	0.26***	0.54***	0.29***

Positive values indicate spatial clustering, while negative values suggest dispersion. Statistical significance is denoted by stars: $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***). Results demonstrate increasing spatial clustering over time, particularly for non-state and one-sided violence.

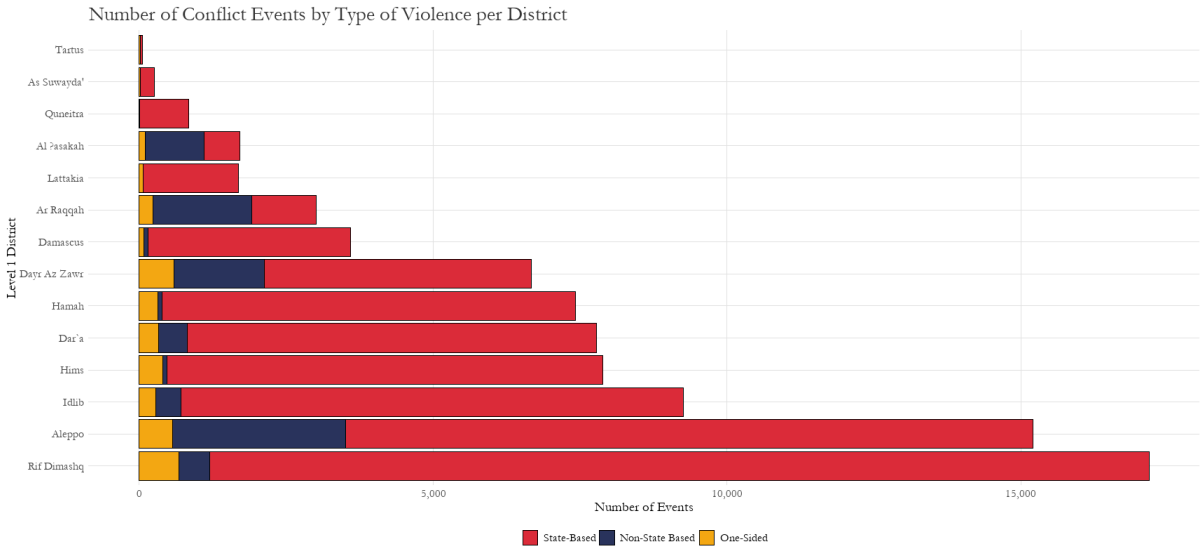

Figure 5.1. *Number of Conflict Events by Type of Violence Across Syrian Governorates, 2011–2019. Aleppo, Rif Dimashq, and Idlib experienced the highest overall event counts, reflecting their strategic importance and prolonged contestation. Stable regions like Tartus and As-Suwayda recorded minimal conflict. Conflict data from UCDP GED (version 25.1).*

5.2 Infrastructure Effects on Conflict

Infrastructure emerged as one of the most consistent predictors of conflict intensity. A one-standard-deviation increase in the infrastructure index raises expected fatalities by 28.9% (IRR = 1.289, β = 0.254, SE = 0.093, $p = 0.006$), total conflict events by 30.1% (IRR = 1.301, β = 0.263, SE = 0.081, $p = 0.001$), state-based violence by 33.9% (IRR = 1.339, β = 0.292, SE = 0.087, $p < 0.001$), and one-sided violence by 27.5% (IRR = 1.275, β = 0.243, SE = 0.099, $p = 0.014$) (Table 5.2). The association with non-state violence (IRR = 1.164, β = 0.152, SE = 0.111) does not reach statistical significance ($p = 0.171$), suggesting that infrastructure matters less for clashes exclusively between non-state actors. Better-connected and more densely settled districts experienced higher levels of lethal and state-linked violence.

Table 5.2. *Infrastructure Effects on Conflict Outcomes (2011-2019)*

Model	Infrastructure Estimate (β)	Incident Rate Ratio	Std. Error	Statistic	P-value	Significance
Estimated Fatalities	0.254	1.289	0.093	2.727	0.006	**
Conflict Count	0.263	1.301	0.081	3.230	0.001	**
State-Based	0.292	1.339	0.087	3.352	0.001	***
Non-State	0.152	1.164	0.111	1.369	0.171	
One-Sided	0.243	1.275	0.099	2.460	0.014	*

Negative binomial models show infrastructure significantly increases conflict count, fatalities, state-based, and one-sided violence, but not non-state violence. Coefficients, standard errors, and significance levels are reported: $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***). Percent changes are computed as (IRR−1)×100%.

5.3 Socioeconomic Inequality and Developmental Residuals

Higher income is strongly protective against violence: a one-standard-deviation increase in the income index reduces expected fatalities by 75.0% (IRR = 0.250, β = −1.387, SE = 0.510, p = 0.006), total conflict events by 75.1% (IRR = 0.249, β = −1.389, SE = 0.455, p = 0.002), and non-state violence by 98.8% (IRR = 0.012, β = −4.403, SE = 1.189, $p < 0.001$), with one-sided violence also falling by 63.7% (IRR = 0.363, β = −1.014, SE = 0.415, p = 0.015) (Table 5.3). The residualised education index has a marginally positive association with fatalities (IRR = 1.200, β = 0.693, SE = 0.358, p = 0.053) and significantly increases one-sided violence by 86.1% (IRR = 1.861, β = 0.621, SE = 0.281, p = 0.027) (Table 5.3). Residualised health shows no major effects on most outcomes but is associated with a 44.0% reduction in one-sided violence (IRR = 0.560, β = −0.581, SE = 0.211, p = 0.006) (Table 5.3). Overall, higher income and health residuals tend to suppress violence, whereas above-expected education is linked to elevated targeted civilian violence.

Table 5.3. *Effects of Income and Development Residuals on Conflict Outcomes*

Model	Term	Estimate	Incident Rate Ratio	Std. Error	Statistic	P-value	Signif.
Estimated Fatalities	Income	-1.387	0.250	0.510	-2.722	0.006	**
Estimated Fatalities	Education Residual	0.693	1.200	0.358	1.936	0.053	
Estimated Fatalities	Health Residual	-0.381	0.683	0.280	-1.361	0.173	
Conflict Count	Income	-1.389	0.249	0.455	-3.049	0.002	**
Conflict Count	Education Residual	0.539	1.714	0.337	1.600	0.110	
Conflict Count	Health Residual	-0.381	0.683	0.256	-1.490	0.136	
State-Based	Income	-0.967	0.380	0.499	-1.935	0.053	
State-Based	Education Residual	0.189	1.208	0.386	-0.490	0.624	
State-Based	Health Residual	-0.161	0.851	0.304	-0.529	0.597	
Non-State	Income	-4.403	0.012	1.189	-3.703	0.000	***
Non-State	Education Residual	0.105	1.111	0.931	0.112	0.911	
Non-State	Health Residual	-0.778	0.459	0.505	-1.539	0.124	
One-Sided	Income	-1.014	0.363	0.415	-2.442	0.015	*
One-Sided	Education Residual	0.621	1.861	0.281	2.214	0.027	*
One-Sided	Health Residual	-0.581	0.560	0.211	-2.759	0.006	**

Regression coefficients for income and education and health residuals. Higher income reduces most forms of violence, while positive education residuals increase one-sided violence. Percent changes are computed as (IRR−1)×100%.

5.4 Ecological Effects

The land cover index is positively and significantly associated with every form of violence. A one-standard-deviation increase in the land cover index (capturing greater proportions of croplands, shrublands, savannas, etc.) raises expected fatalities by 80.8% (IRR = 1.808, β = 0.592, SE = 0.151, $p < 0.001$), total conflict events by 35.8% (IRR = 1.358, β = 0.306, SE = 0.127, $p = 0.016$), and state-based violence by 41.5% (IRR = 1.415, β = 0.347, SE = 0.137, $p = 0.012$) (Table 5.4). The strongest effects appear for non-state violence (IRR = 1.839, a 83.9% increase) and one-sided violence (IRR = 1.704, a 70.4 % increase). These results suggest that districts dominated by mixed rural and semi-natural land covers experienced substantially higher rates of both state- and non-state forms of conflict.

Table 5.4. *Land Cover Effects on Conflict Outcomes*

Model	Land Cover Estimate	Incident Rate Ratio	Std. Error	Statistic	P-value	Signif.
Estimated Fatalities	0.592	1.808	0.151	3.917	0.000	***
Conflict Count	0.306	1.358	0.127	2.409	0.016	*
State-Based	0.347	1.415	0.137	2.527	0.012	*
Non-State	0.609	1.839	0.220	2.773	0.006	**
One-Sided	0.533	1.704	0.129	4.144	0.000	***

Land cover index is positively associated with all types of violence, with the strongest effects observed for non-state and one-sided violence. Percent changes are computed as (IRR−1)×100%.

5.5 Conflict Over Time: Trends and Transitions

The intensity and type of violence evolved significantly from 2011 to 2019 (Table 5.6; Figure 5.6). Conflict escalated sharply from 2012 to 2014, peaking during the height of opposition offensives and foreign intervention. After 2016, levels of violence gradually declined, particularly non-state and one-sided violence. Despite the broader de-escalation, state-based violence remained consistent throughout.

Table 5.6. *Temporal Trends in Conflict (2011-2019)*

Year	Trend Summary	State-Based	Non-State	One-Sided
2011	Baseline year	Low	Low	Minimal
2012	Sharp rise in conflict events	Increased	Rising	Rising
2013	Continued escalation	High	High	Moderate
2014	Peak of violence	Very High	Peak	High
2015	Start of decline	Very High	Declining	Peak
2016	Moderate decline	Sustained	Lower	Declining
2017	Acceleration of de-escalation	Sustained	Decreased	Lower
2018	Significant drop in non-state & one-sided	Sustained	Low	Minimal
2019	Lowest overall conflict levels	Sustained	Very Low	Minimal

Summary of conflict dynamics across time. Non-state and one-sided violence declined after 2016, while state-based violence remained more consistent.

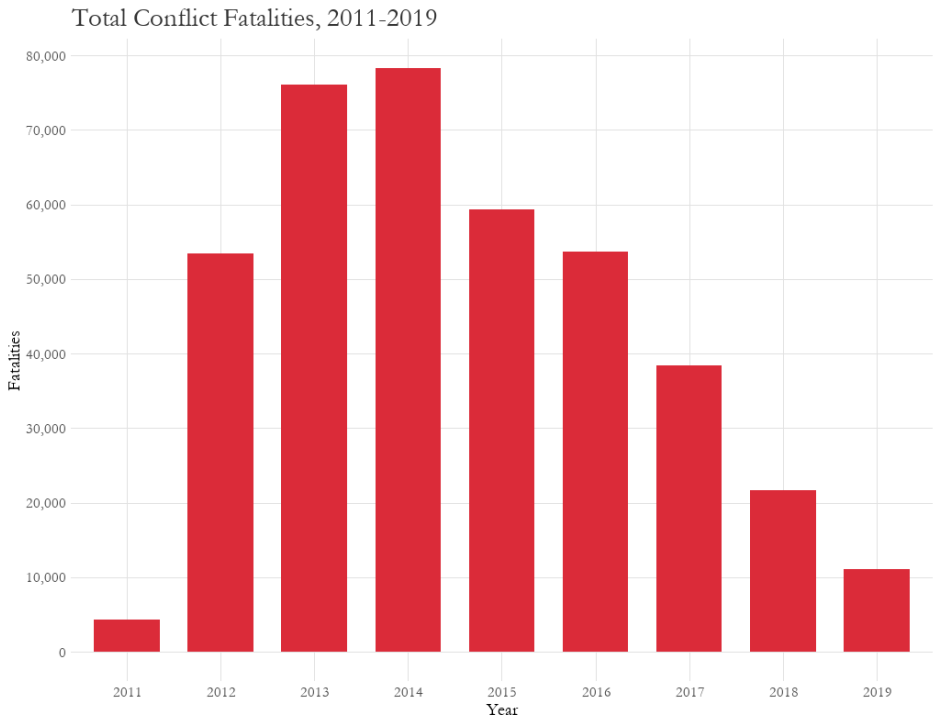

Figure 5.6. *Total Estimated Fatalities During Syrian Conflict, 2011–2019. Annual total conflict-related fatalities across all violence types in Syria from 2011 to 2019. Fatalities surged sharply from 2012, peaking in 2014 at nearly 80,000, followed by a steady decline as territorial consolidation and ceasefire agreements took effect. Conflict data from UCDP GED (version 25.1), symbolised by Estimated Fatalities (best_est).*

6. Discussion

The Syrian civil war's internal diversity serves as a natural experiment for examining how district-level structural determinants, conditional on governorate-level heterogeneity, influence patterns of violence. While this study corroborates many well-known findings from subnational conflict literature, it also complicates several prevailing assumptions, particularly those concerning infrastructure and education.

6.1 Infrastructure

Districts with denser road networks, higher population concentration, and greater urban land cover, experienced higher rates of state-based and one-sided violence, even after accounting for governorate-level baseline risk. This supports the view that dense road networks lengthen government reach and expose civilians to combat and repression (Fearon and Laitin, 2003; Müller-Crepon et al., 2021). Syria's own experience underscores the point: when government forces reopened the M5 highway between Damascus and Aleppo in February 2020, the entire corridor became a zone of checkpoints, shelling, and forced displacement (Reuters, 2020). Walther et al. (2025) find that 65 percent of violent incidents in North and West Africa between 2000 and 2024 occurred within one kilometre of a road. Crucially, that share fell from 87 percent in 2011 to around 60 percent by 2023, indicating that armed groups have progressively shifted operations into more remote terrain. In this study's district-level regressions, infrastructure had significant association with non-state violence.

Two mechanisms may help explain Syria's exception. Rebel infighting occurred primarily in contested urban-rural pockets, such as Idlib and western Aleppo, already saturated with regime troops (Asharq Al-Awsat, 2018). Second, at the height of the Islamic State's territorial reach (2014–17), regime forces focused their offensives on major highways, channelling hostilities toward regime-rebel front lines and largely preventing rebel-on-rebel battles in the well-connected regions (Reuters, 2017a; Reuters, 2017b; Asharq Al-Awsat, 2018).

6.2 Income and the Opportunity–Grievance Nexus

Higher governorate-level income is strongly protective across all violence types, consistent with opportunity-cost theories that view better economic prospects as a disincentive to rebel recruitment and a buffer for civilians (Collier and Hoeffler, 2004). Non-state violence is especially suppressed in wealthier governorates, suggesting insurgents tend to target poorer regions where local elites have fewer resources to mobilise self-defence (Østby, 2008). Yet the protective effect of income for one-sided violence is smaller, implying that the Syrian regime's targeting may not purely be income-based; rather, it may be moderated by perceived dissent potential, bringing education into focus.

6.3 Education Residuals

Governorates whose educational attainment exceeds what their income levels would predict experience significantly more one-sided violence. This pattern is consistent with theories of selective mass killing that depict educated communities as nodes of potential mobilisation, especially under autocratic rule (Cederman et al., 2013; Valentino et al., 2004). Importantly, the residualisation procedure used here removes income's confounding influence, isolating education's political rather than economic signal. The result therefore strengthens arguments that partial modernisation, rapid human-capital gains without corresponding political inclusion, can trigger coercive backlash (Goldstone et al., 2010).

6.4 Land Use and Land Cover

The district-level models reveal that mixed mosaics of cropland, shrubland, and barren terrain are linked to lower non-state violence but higher state-based and one-sided violence, underscoring within-governorate heterogeneity. This aligns with Buhaug et al. (2009) finding that inaccessible, rugged terrain impedes armed actors' mobility and surveillance, thereby shaping conflict dynamics.

6.5 Temporal Dynamics

Violent events in Syria declined markedly after 2016, but state-based operations stayed comparatively high, reflecting foreign intervention and a succession of ceasefires that froze front lines rather than ending them (Blanchard, 2023). This persistence of regime coercion alongside overall de-escalation illustrates how conflict can both wind down and become entrenched.

6.6 Methodological Reflections

The multilevel negative binomial framework, modelling district-level counts with governorate-level random intercepts, appropriately handles overdispersed data and clustered units, but involves trade-offs.

Residualising education and health on income handles multicollinearity but assumes linear relationships and carries any measurement error in governorate income into the derived residuals, which can bias coefficient estimates (García et al., 2019). Sensitivity checks with alternative residualisations, reported in the Part II: Technical Report, Section 6, partly mitigate this concern, but future work could test Bayesian measurement models that treat income, education, and health, as latent variables.

Although Moran's I and LISA confirmed spatial clustering of conflict outcomes, the model does not explicitly include spatially structure random effects; future work could integrate spatial-CAR or Gaussian terms to capture fine-scale dependence between adjacent districts (Dormann et al., 2007).

Conflict data derived from UCDP-GED, which, despite rigorous sourcing, is known to under-report events in remote deserts and over-report in media-dense areas like Aleppo (Sundberg and Melander, 2013). The positive infrastructure coefficient might thus partly reflect reporting bias. Triangulating with local NGO or satellite-inferred casualty datasets could clarify this point.

6.7 Limitations and Implications

Since conflict influences development just as development influences conflict, reverse causality cannot be entirely dismissed. By combining district-level infrastructure and land cover with governorate-level residualised development indices in a mixed effects model, this study bridges opportunity-cost and relative deprivation theories. Results indicate that economic deprivation may possibly fuel non-state conflict, whereas relative educational advantage attracts selective regime repression. For a more detailed discussion on limitations, refer to Part II: Technical Report, Section 9.

6.8 Future Directions

Incorporating dynamic, district-level territorial-control maps would test whether infrastructure effects invert under rebel governance. Agent-based simulations could model how road improvements alter contestation probability under different control scenarios. Comparative studies in Yemen or Libya, civil wars with similar urban-rural splits, could establish whether the infrastructure-violence link is Syria-specific or generalisable. A mixed-methods approach that combines geospatial mapping with qualitative fieldwork in areas like Idlib or Daraa could reveal the precise processes through which education intensifies regime elites' perceptions of threat.

References

Anselin, L. (1995) 'Local Indicators of Spatial Association—LISA', *Geographical Analysis*, 27(2), pp. 93–115. doi: 10.1111/j.1538-4632.1995.tb00338.x.

Ash, K. and Obradovich, N. (2019) 'Climatic Stress, Internal Migration, and Syrian Civil War Onset', Journal of Conflict Resolution, 64(1), pp. 3–31. doi: 10.1177/0022002719864140.

Asharq Al-Awsat (2018) 'Syria: Death Toll in Idlib Arms-Depot Blast Rises', *Asharq Al-Awsat*, 13 August. Available at: https://english.aawsat.com/home/article/1361906/syria-death-toll-idlib-arms-depot-blast-rises

Benjamini, Y. and Hochberg, Y. (1995). 'Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing', *Journal of the Royal Statistical Society: Series B (Methodological), 57*(1), 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x.

Bivand, R.S. and Wong, D.W.S. (2018). Comparing Implementations of Global and Local Indicators of Spatial Association. *Journal of Statistical Software, 83*(4), 1–36. doi: 10.18637/jss.v083.i04.

Blanchard, C.M. (2023) Syria Conflict Overview: 2011-2021. Available at: https://crsreports.congress.gov/product/pdf/IF/IF11080

Brooks, M.E., Kristensen, K., van Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Maechler, M. and Bolker, B.M. (2017) 'glmmTMB Balances Speed and Flexibility Among Packages for Zero-Inflated Generalized Linear Mixed Modeling', The R Journal, 9(2), pp. 378–400. doi: 10.32614/RJ-2017-066.

Buhaug, H. and Rød, J.K. (2006) 'Local Determinants of African Civil Wars, 1970–2001', *Political Geography*, 25(3), pp. 315–335. doi: 10.1016/j.polgeo.2006.02.005.

Buhaug, H., Gates, S. and Lujala, P. (2009) 'Geography, Rebel Capability, and the Duration of Civil Conflict', *Journal of Conflict Resolution*, 53(4), pp. 544–569. doi: 10.1177/0022002709336457.

Burnham, K.P. and Anderson, D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2nd edn. New York: Springer.

Cameron, A.C. and Trivedi, P.K. (1998) *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.

Cederman, L.-E., Gleditsch, K.S. and Buhaug, H. (2013) *Inequality, Grievances, and Civil War*. Cambridge: Cambridge University Press.

Central Intelligence Agency (2025) '*Syria*', in *The World Factbook*. Available at: https://www.cia.gov/the-world-factbook/countries/syria/ (Accessed: 23 June 2025).

Collier, P. & Hoeffler, A. (2004) 'Greed and Grievance in Civil War', *Oxford Economic Papers*, 56(4), pp. 563–595. doi: 10.1093/oep/gpf064.

Collier, P., Elliott, L.V., Hegre, H., Hoeffler, A., Reynal-Querol, M. and Sambanis, N. (2003) *Breaking the Conflict Trap: Civil War and Development Policy*. Washington, DC: World Bank & Oxford University Press.

de Châtel, F. (2014) 'The Role of Drought and Climate Change in the Syrian Uprising', *Middle Eastern Studies*, 50(4), pp. 521–535.

Dinç, P. & Eklund, L. (2023) 'Syrian Farmers in the Midst of Drought and Conflict: Causes, Patterns, and Aftermath of Land Abandonment and Migration', *Climate and Development*, 16(5), pp. 349–362. doi: 10.1080/17565529.2023.2223600

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M. and Wilson, R. (2007) 'Methods to Account for Spatial Autocorrelation in the Analysis of Species Distributional Data: A Review', *Ecography, 30*(5), 609–628. doi: 10.1111/j.2007.0906-7590.05171.x.

Esri (2025) *ArcGIS Pro* (Version 3.5) [computer software]. Redlands, CA: Environmental Systems Research Institute. Available at:

https://www.esri.com/en-us/arcgis/products/arcgis-pro/
Fearon, J.D. and Laitin, D.D. (2003) 'Ethnicity, Insurgency, and Civil War', *American Political Science Review*, 97(1), pp. 75–90. 10.1017/S0003055403000534.

Filmer, D., and Pritchett, L.H. (2001) 'Estimating Wealth Effects without Expenditure Data-or Tears: An Application to Educational Enrollments in States of India', *Demography*, *38*(1), 115–132. doi: 10.1353/dem.2001.0003.

Friedl M.A., & Sulla-Menashe D. (2022) MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V061. Sioux Falls, South Dakota, USA: NASA Land Processes Distributed Active Archive Center. doi: 10.5067/MODIS/MCD12Q1.061.

GADM (2022) *Global Administrative Areas Database* (Version 4.1). Available at: https://gadm.org/download_country.html.

García, C. B., Salmerón, R., García, C., and García, J. (2019) *Residualization: justification, properties and application*. Journal of Applied Statistics, 47(11), 1990–2010. doi: 10.1080/02664763.2019.1701638.

Gelman, A., and Hill, J. (2007) *Data Analysis Using Regression and Multilevel/ Hierarchical Models*. Cambridge: Cambridge University Press.

Gleick, P.H. (2014) 'Water, Drought, Climate Change, and Conflict in Syria', *Weather, Climate, and Society*, 6(3), pp. 331–340. doi: 10.1175/WCAS-D-13-00059.1.

Goldstone, J.A., Bates, R.H., Epstein, D.L., Gurr, T.R., Lustik, M.B., Marshall, M.G., Ulfelder, J. and Woodward, M. (2010) 'A Global Model for Forecasting Political Instability', *American Journal of Political Science*, 54(1), pp.190–208. doi: 10.1111/j.1540-5907.2009.00426.x.

Gujarati, D.N. & Porter, D.C. (2009) *Basic Econometrics*. 5th edn. New York: McGraw-Hill Education.

Gurr, T.R. (2011) *Why Men Rebel*. 1st edn. Abingdon: Routledge. doi: 10.4324/9781315631073.

Hartig, F. (2024) *DHARMa: Residual Diagnostics for Hierarchical (Multi-Level/Mixed) Regression Models*. R package version 0.4.7. Available at: https://CRAN.R-project.org/package=DHARMa.

Hilbe, J.M. (2011) *Negative Binomial Regression*. 2nd edn. Cambridge: Cambridge University Press.

Hinnebusch, R. (2012) 'Syria: From 'Authoritarian Upgrading' to Revolution?', *International Affairs*, 88(1), pp. 95–113. doi: 10.1111/j.1468-2346.2012.01059.x.

Human Rights Watch (2016) *Syria: War Crimes in a Month of Bombing Aleppo*. Available at: https://www.hrw.org/news/2016/12/01/russia/syria-war-crimes-month-bombing-aleppo

Jolliffe, I.T. & Cadima, J. (2016) 'Principal Component Analysis: A Review and Recent Developments', *Philosophical Transactions of the Royal Society A*, 374(2065), 20150202. doi: 10.1098/rsta.2015.0202.

Kassambara, A. (2017) *Practical Guide to Cluster Analysis in R: Unsupervised Machine Learning*. STHDA.com Edition 1. Available at: https://xsliulab.github.io/Workshop/2021/week10/r-cluster-book.pdf

Kelley, C.P., Mohtadi, S., Cane, M.A., Seager, R. and Kushnir, Y. (2015) 'Climate Change in the Fertile Crescent and Implications of the Recent Syrian Drought', *Proceedings of the National Academy of Sciences*, 112(11), pp. 3241–3246. doi: 10.1073/pnas.1421533112.

Lambert, D. (1992) 'Zero-Inflated Poisson Regression, With an Application to Defects in Manufacturing', *Technometrics*, 34(1), 1–14. doi: 10.2307/1269547.

Leenders, R. and Heydemann, S. (2012) 'Popular Mobilization in Syria: Opportunity and Threat, and the Social Networks of the Early Risers', *Mediterranean Politics*, 17(2), pp. 139–159. doi: 10.1080/13629395.2012.694041.

Li, X., Xu, H.M. and Wu, C.Q. (2017) 'Intercalibration Between DMSP/OLS and VIIRS Night-Time Light Images to Evaluate City Light Dynamics of Syria's Major Human Settlement During Syrian Civil War', International Journal of Remote Sensing, 38(21),

pp. 5934–5951. doi: 10.1080/01431161.2017.1331476.

Liu, Q., Zhang, X., Tian, H., Yang, M. Zhang, Z. and Li, H. (2025) 'The Impacts of War on Land Use and Land Cover in Syria Based on Simulation', *Remote Sensing in Earth Systems Sciences* [Preprint]. doi: 10.1007/s41976-025-00205-8.

Lüdecke, D., Ben-Shacher, M.S., Patil, I., Waggoner, P., and Makowski, D. (2021). 'performance: An R Package for Assessment, Comparison and Testing of Statistical Models', *Journal of Open Source Software*, 6(60), 3139. doi: 10.21105/joss.03139.

Moran, P.A.P. (1950). 'Notes on Continuous Stochastic Phenomena', *Biometrika, 37*(1/2), 17–23. doi: 10.1093/biomet/37.1-2.17.

Müller-Crepon, C., Hunziker, P. and Cederman, L.E. (2021) 'Roads to Rule, Roads to Rebel: Relational State Capacity and Conflict in Africa', *Journal of Conflict Resolution*, 65(2–3), pp. 563–590. doi: 10.1177/0022002720963674.

Nemeth, S.C., Mauslein, J.A. and Stapley, C. (2014) 'The Primacy of the Local: Identifying Terrorist Hot Spots Using Geographic Information Systems', *The Journal of Politics,* 76(2), pp. 304–317. doi: 10.1017/S0022381613001333.

O'Loughlin, J. and Witmer, F.D.W. (2011) 'The Localized gGographies Violence in the North Caucasus of Russia, 1999–2007', *Annals of the Association of American Geographers*, 101(1), pp. 178–201. doi: 10.1080/00045608.2010.534713.

OpenStreetMap Contributors (2025) *OpenStreetMap* [June 2025]. Available at: https://www.openstreetmap.org

Østby, G. (2008) 'Polarization, Horizontal Inequalities and Violent Civil Conflict'. *Journal of Peace Research, 45*(2), pp. 143–162. doi: 10.1177/0022343307087169.

Pebesma, E. & Bivand, R. (2023) *Spatial Data Science: With Applications in R*. 1st edn. Boca Raton: Chapman & Hall/CRC. doi: 10.1201/9780429459016.

Pettersson, T. & Wallensteen, P. (2015) 'Armed Conflicts, 1946–2014', *Journal of Peace Research*, 52(4), pp. 536–550. doi: 10.1177/0022343315595927.

R Core Team (2024) *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available at: https://www.R-project.org/.

Raleigh, C., and Hegre, H. (2009) 'Population Size, Concentration, and Civil War: A Geographically Disaggregated Analysis', *Political Geography, 28*(4), pp. 224–238. doi: 10.1016/j.polgeo.2009.05.007.

Reuters (2017a) 'Syrian Army Advances Along Vital Highway in Rebel-Held Hama Province', Reuters, 20 April. Available at: https://www.reuters.com/article/world/syrian-army-advances-along-vital-highway-in-rebel-held-hama-province-idUSKBN17M1T8/

Reuters (2017b) 'Syrian Army Advance Helps Secure Aleppo Road', Reuters, 3 June. Available at: https://www.reuters.com/article/world/syrian-army-advance-helps-secure-aleppo-road-idUSKBN18U0EP/

Reuters (2020) 'Syria Announces Damascus-Aleppo Highway Open to Traffic', Reuters, 22 February. Available at: https://www.reuters.com/article/world/syria-announces-damascus-aleppo-highway-open-to-traffic-idUSKCN20G0KM/

Risse, T. (ed.) (2013) *Governance Without a State? Policies and Politics in Areas of Limited Statehood*. New York: Columbia University Press.

Schutte, S. and Weidmann, N. (2011) 'Diffusion Patterns of Violence in Civil Wars', *Political Geography*, 30(3), pp. 143–152. doi: 10.1016/j.polgeo.2011.03.005.

Selby, J., Dahi, O., Fröhlich, C. (2017) 'Climate Change and the Syrian Civil War Revisited', *Political Geography*, 60, pp. 232–244. doi: 10.1016/j.polgeo.2017.05.007.

Smits, J. and Permanyer, I. (2019) 'The Subnational Human Development Database', *Scientific Data*, 6, 190038. doi: 10.1038/sdata.2019.38.

Stewart, F. (2008) 'Horizontal Inequalities and Conflict: An Introduction and Some

Hypotheses', in Stewart, F. (ed.) *Horizontal Inequalities and Conflict*. London: Palgrave Macmillan, pp. 3–24. doi: 10.1057/9780230582729_1.

Sundberg, R. and Melander, E. (2013) 'Introducing the UCDP Georeferenced Event Dataset', Journal of Peace Research, 50(4), pp. 523–532. doi: 10.1177/0022343313484347.

Syrian Network for Human Rights (SNHR) (2024) *Daraa Governorate: Barrel Bomb Attacks by the Syrian Regime—September 2024 Statistical Report*. Available at: https://snhr.org/wp-content/uploads/2024/09/R240519E.pdf

UCDP (2024) 'Definitions: State Based Conflict, Non-State Conflict and One-Sided Violence', Uppsala: Department of Peace and Conflict Research, Uppsala University. Available at: https://www.uu.se/en/department/peace-and-conflict-research/research/ucdp/ucdp-definitions/

UNDP (2025) *Accelerating Economic Recovery Critical to Reversing Syria's Decline and Restoring Stability* [online]. Available at: https://www.undp.org/syria/press-releases/accelerating-economic-recovery-critical-reversing-syrias-decline-and-restoring-stability

Urdal, H. (2006) 'Clash of Generations? Youth Bulges and Political Violence', *International Studies Quarterly*, 50(3), pp. 607–629. doi: 10.1111/j.1468-2478.2006.00416.x.

Valentino, B., Huth, P. and Balch-Lindsay, D. (2004) '"Draining the Sea": Mass Killing and Guerrilla Warfare', *International Organization*, 58(2), pp. 375–407. doi: 10.1017/S0020818304582061.

von Uexkull, N., Croicu, M., Fjelde, H. and Buhaug, H. (2016) 'Civil Conflict Sensitivity to Growing-Season Drought', *Proceedings of the National Academy of Sciences*, 113(44), pp. 12391–12396. doi: 10.1073/pnas.1607542113.

Walther, O.J., Radil, S.M. and Thurston, A. (2025) 'Political Violence and Transport Infrastructure in West Africa', *African Security Review*, pp. 1–22. doi:10.1080/10246029.2025.2502017.

Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Wickham, H., François, R., Henry, L., Müller, K. and Vaughan, D. (2023) *dplyr: A Grammar of Data Manipulation* (v1.1.4). Available at: https://CRAN.R-project.org/package=dplyr.

Wickham, H., Hester, J. and Bryan, J. (2024) *readr: Read Rectangular Text Data* (v2.1.5). Available at: https://CRAN.R-project.org/package=readr.

Wickham, H., Vaughan, D. and Girlich, M. (2024) *tidyr: Tidy Messy Data* (v1.3.1). Available at: https://CRAN.R-project.org/package=tidyr.

Woodrow Wilson International Center for Scholars (n.d.) *Syria*. Available at: https://www.wilsoncenter.org/syria (Accessed: 10 June 2025).

Wooldridge, J.M. (2021) *Introductory Econometrics: A Modern Approach*, 8th edn. Boston: Cengage.

PART II

Technical Report

Roads, Riches, and Residuals: A Multilevel Analysis of Subnational Determinants of Violent Conflict in Syria (2011-2019)

Table of Contents

Table of Figures

Table of Tables

1. Introduction

This technical report supplements the main research paper and provides more details into data acquisition and source methodology, data preprocessing, LISA cluster map results, health and education random effects models, comparative analysis between residualised human development variables, and *DHARMa* model diagnostics. The Appendices include R code and other pertinent information.

2. Data Source Methodology

2.1 Uppsala Conflict Data Program's Georeferenced Event Dataset (UCDP GED)

UCDP GED is a global, event-level inventory of organized violence covering 1 January 1989–31 December 2024 in its latest public release, version 25.1 (Stina, 2025). A GED "event" is a single instance in which an organised actor uses armed force against another organised actor or against civilians, producing at least one direct death at a clearly identified place and date (Sundberg and Melander, 2013). Once a conflict dyad exceeds the programme-wide 25-battle-deaths threshold, every lethal incident involving those parties, whether before or after the threshold year, is included, ensuring continuous temporal coverage for that dyad (Croicu and Sundberg, 2017). Each observation carries latitude-and-longitude coordinates and low, best, and high fatality estimates, giving users the ability to trade granularity for confidence when aggregating or modelling (Pettersson et al., 2023).

Data is assembled through a deliberately redundant two-pass workflow. In the first pass, coders collect candidate reports each month from global newswire feeds in Dow Jones Factiva, using broad keyword strings, and from BBC Monitoring translations, which together generate most initial events (Stina, 2025). The second pass targets gaps revealed in that sweep, drawing on local media, NGO and UN documents, and selected social media accounts to refine borderline cases and enrich under-reported conflicts (Croicu and Sundberg, 2017). All entries are produced by full-time regional specialists and then funnelled through a multilayer quality-assurance regime: coders follow detailed protocols, meet weekly to resolve anomalies, and run more than fifty automated consistency tests (e.g., coordinate-density and ID checks) before a project manager approves public release (Pettersson et al., 2023).

2.2 Global Data Lab's Subnational Human Development Index (SHDI)

SHDI adapts the United Nations Development Programme's (UNDP) national HDI to first-level administrative units (Global Data Lab, 2025). Like the original index, it derives three dimension scores, health, education, and standard of living (Smits and Permanyer, 2019). Each indicator is normalised against fixed goalposts and the three dimension indices are then combined with a second geometric mean to produce the SHDI (Global Data Lab, 2025; Smits and Permanyer, 2019). The resulting regional values are proportionally rescaled so that their population-weighted national average equals the UNDP-reported HDI for the same year, ensuring strict comparability between subnational and national figures (Smits and Permanyer, 2019). Where direct observations are missing for a particular region-year, short-range

interpolation or extrapolation fills the gaps (Global Data Lab, 2025; Smits and Permanyer, 2019). The most recent public release, version 8.1 dated 23 October 2024, delivers annual SHDI series spanning 1990-2022 (Global Data Lab, 2024).

2.3 MODIS Land Cover Data

The MCD12Q1 Collection 6.1 land-cover product uses year-long, 500 m reflectance composites from Terra and Aqua MODIS sensors to generate global land cover maps via an ensemble of decision-tree classifiers trained on thousands of reference sites (Friedl and Sulla-Menashe, 2022). Collection 6.1 retains the original classifier but incorporates better-calibrated reflectance data, updated view-angle correction tables, a Terra polarization adjustment, and a small UMD land-cover class fix, resulting in more temporally stable maps (NASA Earthdata, 2022).

2.4 WorldPop Gridded Population Dataset

WorldPop's unconstrained top-down gridded-population workflow begins by taking the latest census counts for every administrative unit and redistributing them over a global 100 × 100 m grid (WorldPop, n.d.). A random-forest model relates population density in thousands of training polygons to covariates (night-time lights, land-cover, building footprints, road and river networks, elevation, etc.) (Stevens et al., 2015). The trained model produces a relative-density surface and cell-level weights are scaled so the total within each source unit exactly matches its reported population, yielding internally consistent 100 m (and aggregated 1 km) layers for every year (WorldPop, n.d.).

Because the algorithm does not presuppose any uninhabitable areas, every land pixel receives at least a small value. This reduces omission of genuine but unmapped settlements that "constrained" products may miss (Stevens et al., 2015). Evaluations in urban Namibia found that unconstrained maps can "bleed" some population into truly uninhabited pixels and slightly underestimate dense city cores (Thomson et al., 2022). Even so, median absolute errors were comparable to other global datasets and accuracy improved wherever up-to-date covariates were available (WorldPop, n.d.).

2.5 OpenStreetMap Road Data

OpenStreetMap (OSM) builds its global map through a community-sourced, verifiable-data workflow: volunteers record GPS traces, field surveys, open government datasets, and satellite imagery, then upload edits that are tracked in a public, version-controlled database (Haklay and Weber, 2008). Each feature is stored as a simple geometry accompanied by descriptive tags so contributors can capture both physical objects and semantic attributes (Ramm et al., 2010). Continuous peer review, automated validators, and quality metrics help spot conflicts and steer refinements (Barron et al., 2013).

3. Preprocessing Data

3.1 Conflict Data

All conflict outcomes (conflict_count, total_fatalities, state_based, non_state_based, and one_sided) were aggregated at the district level using the *sum()* function, excluding missing values with *na.rm = TRUE*, from the *dplyr* package in R (Wickham et al., 2023, v1.1.4). For detailed code, see Appendix 1.2. When the conflict data were joined with the Syria district polygon shapefile for spatial analysis, 405 records were excluded because they fell outside the country's administrative boundaries. This resulted in the dataset *conflict_summary_district* for multilevel modelling. During exploratory data analysis, state-based, non-state, and one-sided violence for each year were mapped in ArcGIS Pro (Esri, 2025, v3.5) to observe any spatial or temporal trends (Figure 3.1 and 3.1.2).

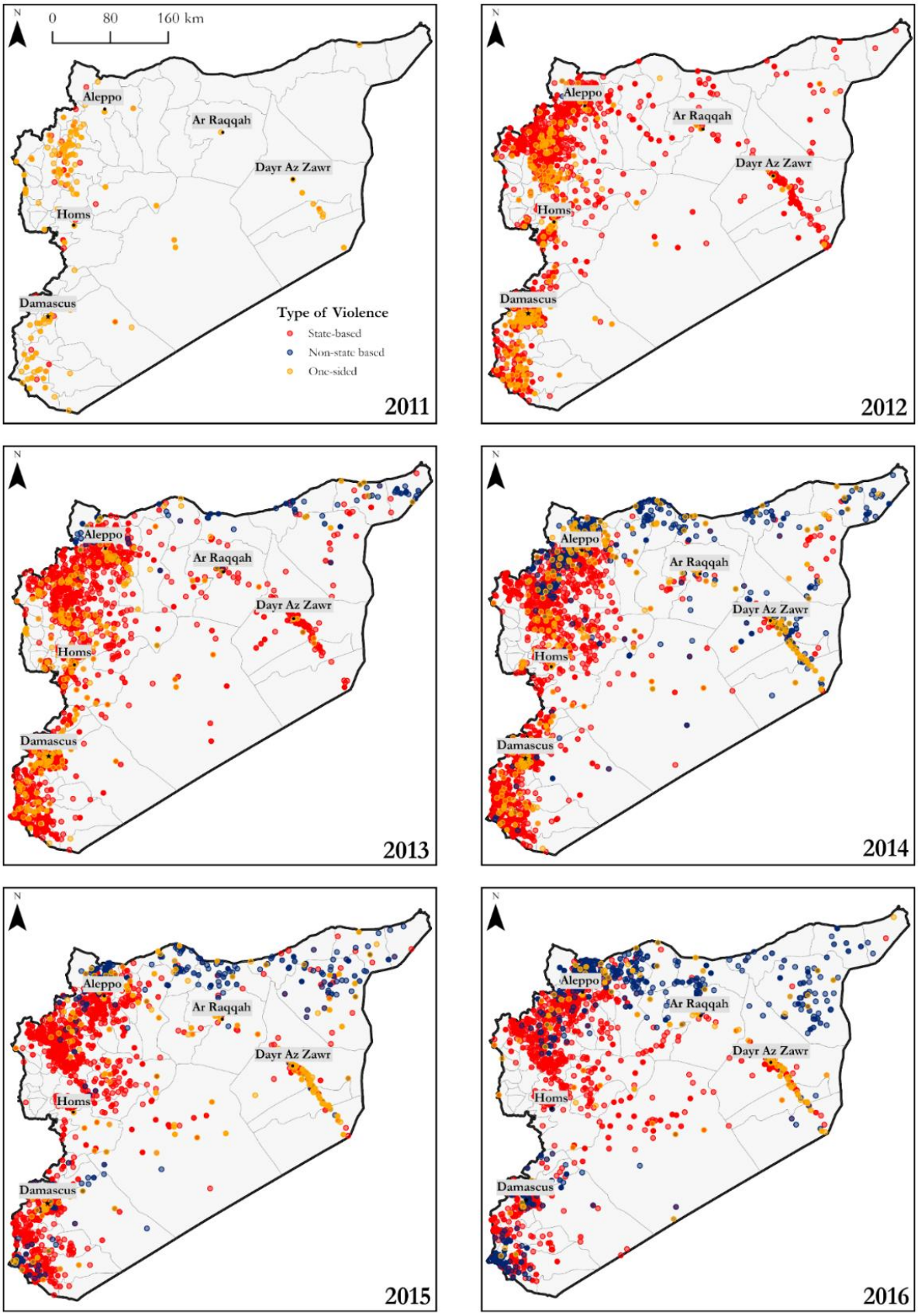

Figure 3.1. *Total Number of Violent Conflicts by Type of Violence in Syria by Year. Conflict data from UCDP GED (version 25.1), symbolised by Type of Violence. Administrative boundaries data sourced from GADM (version 4.1).*

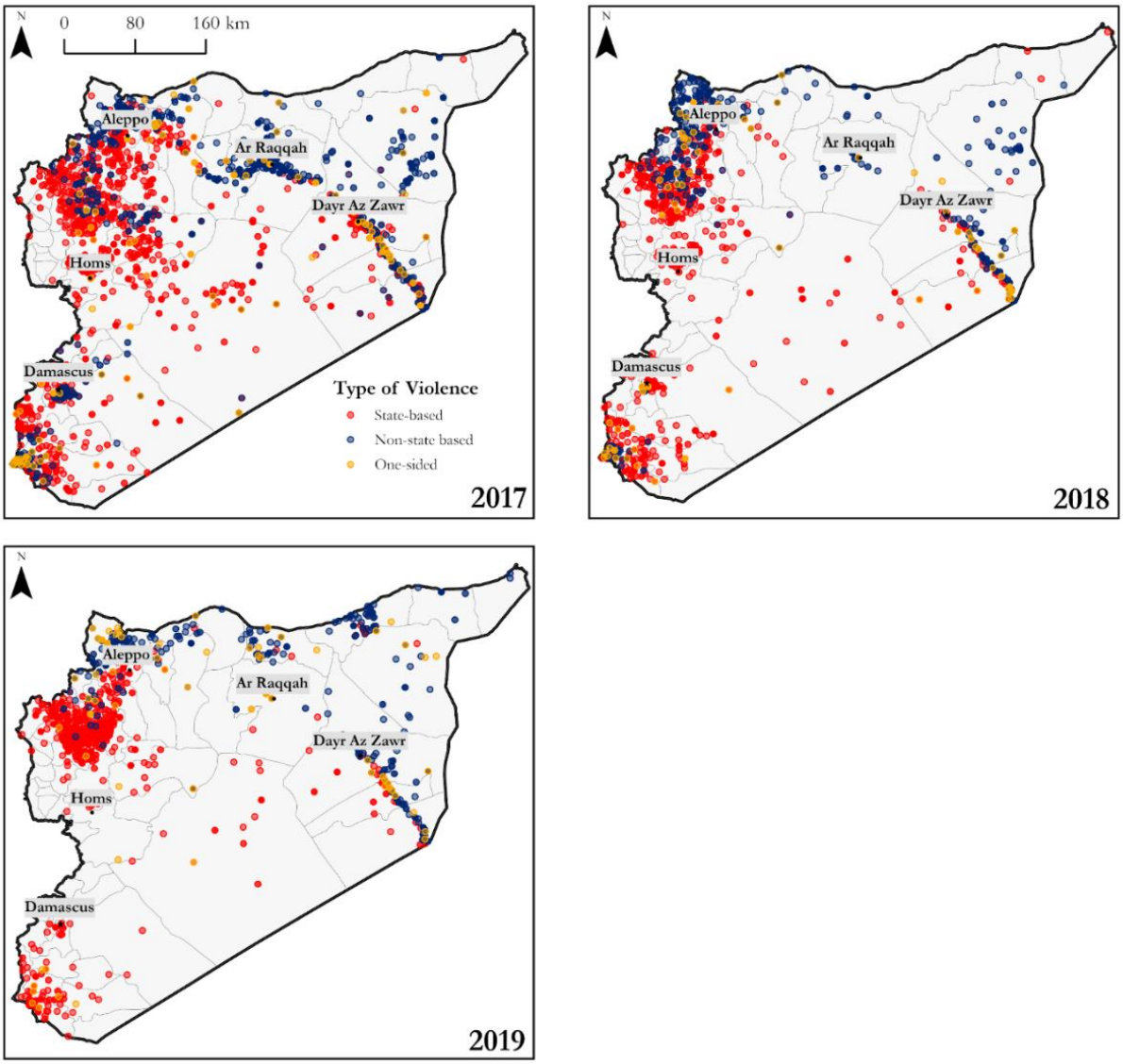

Figure 3.1.2. *Total Number of Violent Conflicts by Type of Violence in Syria by Year. Conflict data from UCDP GED (version 25.1), symbolised by Type of Violence. Administrative boundaries data sourced from GADM (version 4.1).*

3.2 Land Cover

Eighteen MODIS rasters were utilised to calculate land cover percentages at the district level in ArcGIS Pro (Esri, 2025, v3.5). Since the rasters were sinusoidal, the *Define Projection* tool was employed to standardise the geographic coordinate system across datasets and was set to the projection GCS_WGS_1984. Each raster was then projected to WGS_1984_UTM_Zone_37N using the *Project Raster* tool with the Nearest Neighbour resampling method to preserve the categorical nature of the land cover values. As Syria is covered by two separate raster tiles for each year, the *Merge Rasters* tool was applied to combine the corresponding rasters, resulting in a single raster per year and a total of nine merged rasters. These were then clipped to Syria's national border using the *Clip Raster* tool.

Following clipping, the *Raster Calculator* tool was used to reclassify the land cover categories. A nested conditional (Con) statement was applied to retain only the relevant land cover classes (values 1–14, 16, and 17) since other land cover classes are not represented in Syria. Symbology

was imported from a reference layer to apply consistent colour schemes to land cover classes that matched the legend for MODIS-derived land cover data (Figure 3.2 and 3.2.1).

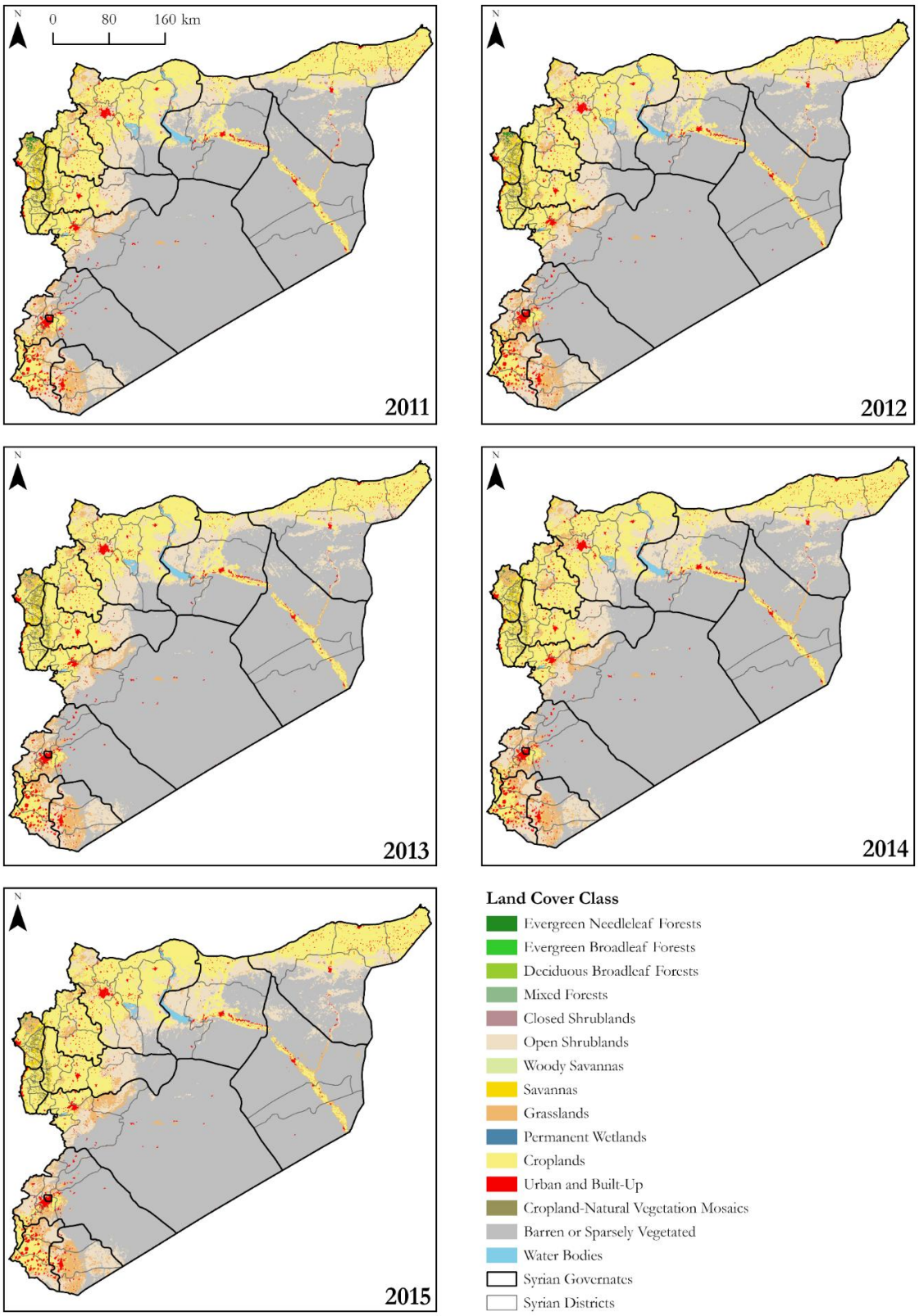

Figure 3.2. *Land Cover in Syria by Year. Land cover data sourced from NASA Earthdata, MODIS/Terra + Aqua Land Cover Type Yearly L3 Global 500 m SIN Grid (Friedl & Sulla-Menashe, 2022, version 6.1). Administrative boundaries from GADM (version 4.1).*

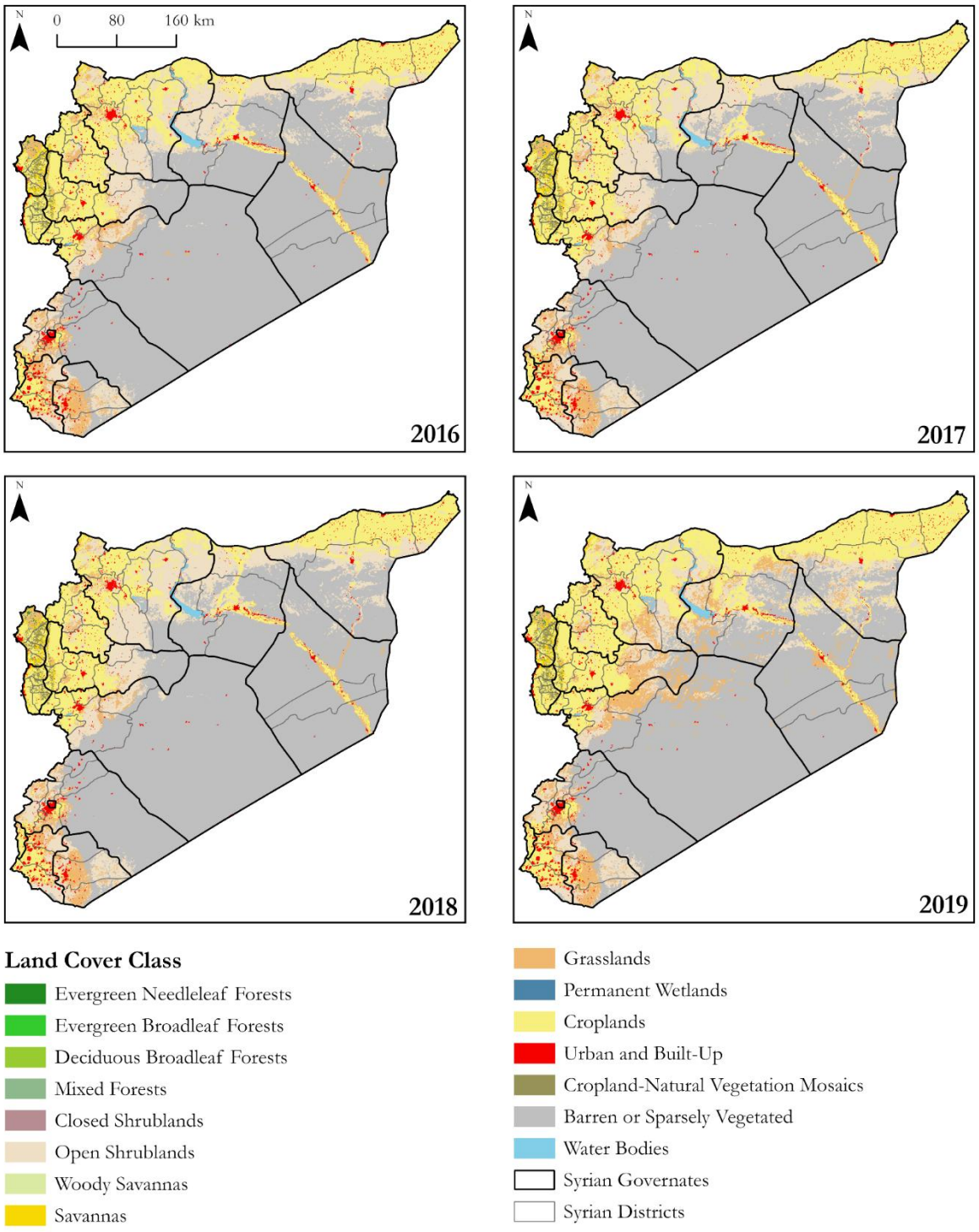

Figure 3.2.1. *Land Cover in Syria by Year. Land cover data sourced from NASA Earthdata, MODIS/Terra + Aqua Land Cover Type Yearly L3 Global 500 m SIN Grid (Friedl & Sulla-Menashe, 2022, version 6.1). Administrative boundaries from GADM (version 4.1).*

The *Tabulate Area* tool was then used to calculate the number of pixels for each land cover class within Syria's 60 district boundaries. The resulting tables were exported to Excel (using the Table to Excel tool) for further processing. In cases where some land cover classes were not present in certain districts, columns were manually added back to maintain a consistent classification structure.

The area for each land cover class was calculated using the known pixel resolution. Each pixel measured 463.31 metres × 463.31 metres, resulting in an area of 214,654.6 square metres per pixel. To convert this to square kilometres, the following formula was applied:

Area (km²) = (Number of Pixels × 214,654.6) / 1,000,000

Finally, the percentage of each land cover class within a district was calculated using the formula:

(Area of Land Cover Class / Total District Area) × 100,

yielding the percentage of each land cover class relative to the district's total area.

3.3 Population Density

Nine WorldPop unconstrained, UN-adjusted annual population raster datasets (100 m resolution) were used for the years 2011–2019 to calculate yearly population and population density at the district level.

Each raster was first projected to WGS_1984_UTM_Zone_37N using *Project Raster* in ArcGIS Pro (Esri, 2025, v3.5). The *Zonal Statistics as Table* tool was applied to summarise the total population within each administrative district boundary, producing a table containing the summed population values for each year. These resulting zonal tables were joined to the district layer using NAME_2. Two fields were calculated for each year: pop_20XX, representing the summed population for that year, and dens_20XX, indicating population density. The population density was calculated using the formula:

dens_20XX = (SUM / (Shape_Area / 1,000,000))

where the population sum (SUM) was divided by the district area converted from square metres to square kilometres, providing population density expressed as people per square kilometre. This process was repeated individually for each of the nine WorldPop raster datasets corresponding to different years. After calculating the respective population and density fields for each year, each zonal table join was removed before repeating the process for the next year's dataset to reduce the risk of overwriting earlier years' values.

3.4 Road Density

After importing OSM data into ArcGIS Pro and reprojecting to WGS_1984_UTM_Zone_37N, a field named road_length_m was added to the road layer to store segment lengths. The length of each road segment was computed in metres using the *Calculate Geometry* tool. A spatial join was performed between the road layer and the Syrian district polygon layer, which applied a merge rule of sum to aggregate the total length of all road segments within each district. After the

spatial join, a new field named area_km2 was added to the district polygon layer, and district areas were calculated in square kilometres using the *Calculate Geometry* tool. The third and final field created was road_density_km_per_km2. Using the *Field Calculator*, road density was computed with the formula:

road_density_km_per_km2 = road_length_m / 1000 / area_km2

This converted road lengths from metres to kilometres and divided by district area in square kilometres, producing the final road density measure (km/km²). The resulting road density values were joined back to the original district layer using NAME_2 (Figure 3.4).

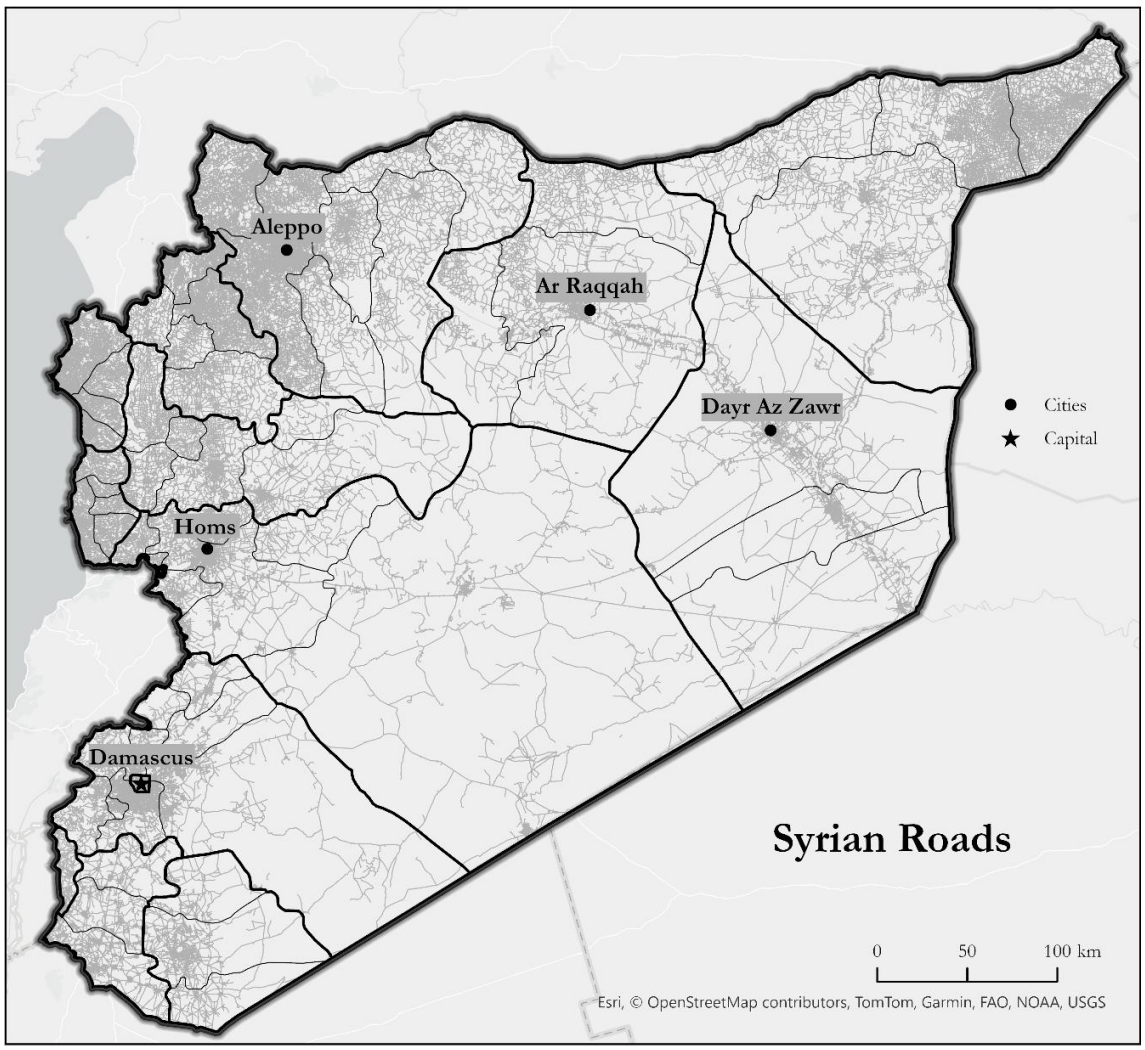

Figure 3.4. *Roads in Syria. Roads data from OpenStreetMap and administrative boundaries sourced from GADM (version 4.1).*

4. LISA Cluster Analysis

This Local Indicators of Spatial Association (LISA) cluster analysis recognizes spatial patterns of types of violence (state-based, non-state, and one-sided) across Syria from 2011 to 2019. By measuring the degree of spatial autocorrelation in conflict intensity, LISA maps show where similar levels of violence are geographically clustered. This method distinguishes statistically significant hot spots (e.g., high-high clusters) and cold spots (e.g., low-low clusters), as well as spatial outliers (Anselin, 1995).

In 2011, a hot spot of state violence shows up in Salamiyah, while the surrounding districts emerge as stable cold spots (Figure 4.1). In 2012, state-based violence jumps south to a hot spot in Damascus. By 2013–14, patterns diffuse: 2013 yields a cold spot cluster in Daraa and neighboring At-Tall, and 2014 a lone hot spot reappears in Damascus. No district reaches significance in 2015 or 2017, reflecting a more dispersed fighting pattern. In 2016, new cold spots materialise in the north, Harem, As-Safirah, Afrin, and Azaz districts, before a brief low-low cluster surfaces in At-Tall district in 2018 (Figure 4.2). Finally, in 2019 a pronounced hot spot re-emerges in the north near Idlib and Arihah.

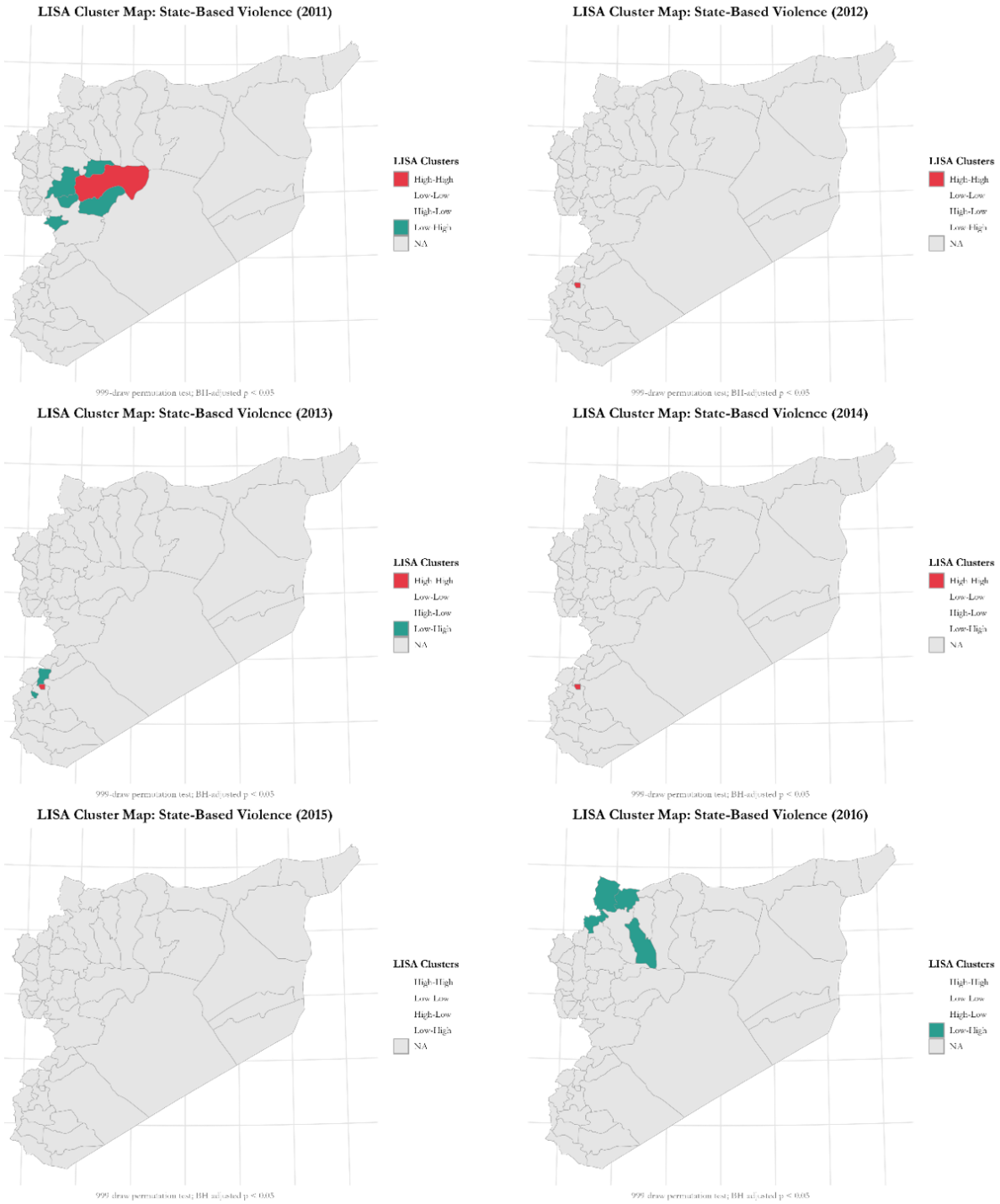

Figure 4.1. *LISA Cluster Map Sequence for State-Based Violence (2011-2016). Conflict data from UCDP GED (version 25.1) summarised by district-year state-based violence. Administrative boundaries data sourced from GADM (version 4.1).*

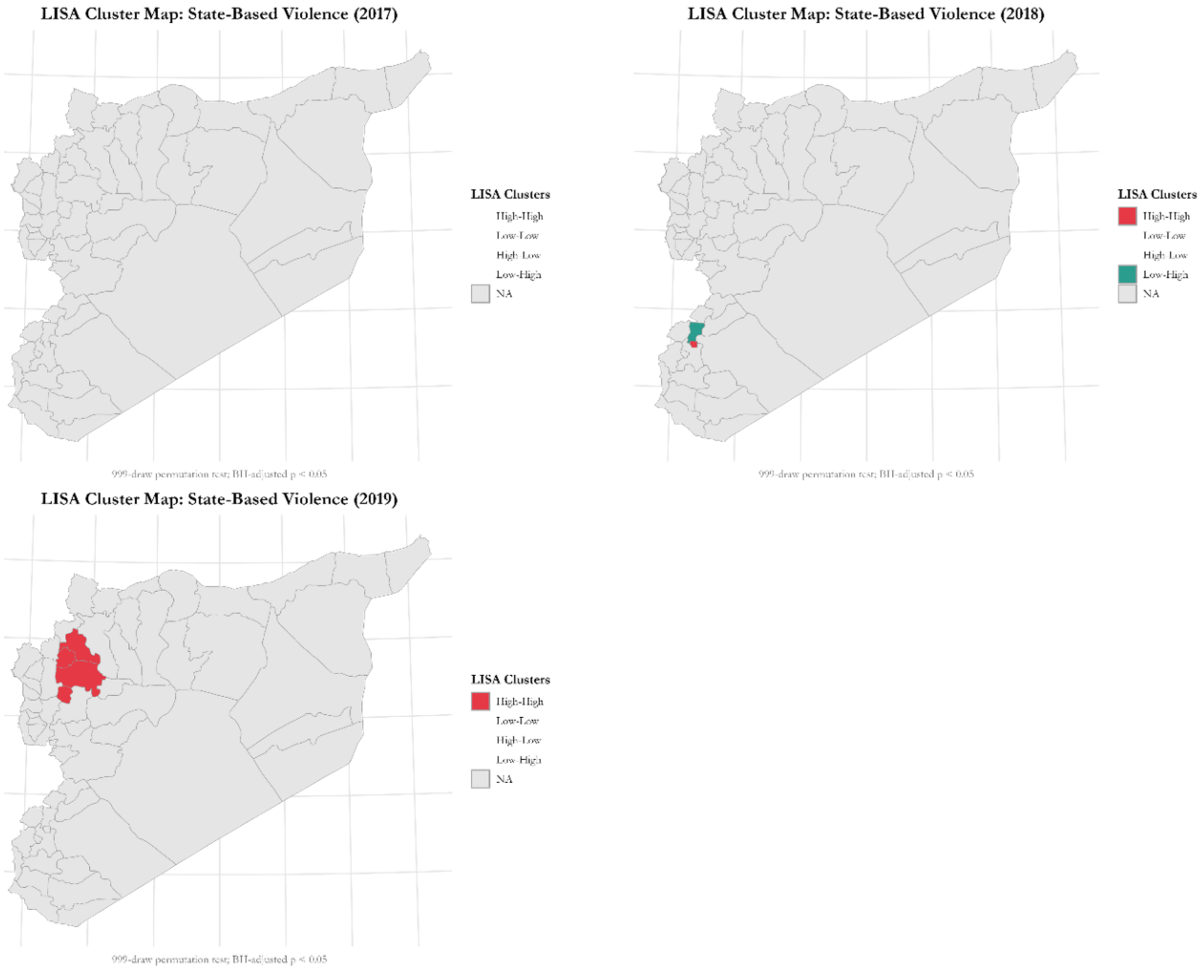

Figure 4.2. *LISA Cluster Map Sequence for State-Based Violence (2017-2019). Conflict data from UCDP GED (version 25.1) summarised by district-year state-based violence. Administrative boundaries data sourced from GADM (version 4.1).*

Throughout the non-state violence LISA panel (Figure 4.3), there is a shift in insurgency movement over the decade. In 2011 there are only low–low cold spots around central districts (e.g. An-Nabk, Al-Qusayr, Ar-Rastan, and Al-Mukharram), reflecting minimal rebel activity. By 2012 a nascent hot spot appears in northern Aleppo (Afrin), and in 2013–14 the violence intensifies in Ra's al-'Ayn and Al Qamishli and then spilling east into Ar-Raqqah. In 2017, a cold spot emerges in Tal Abyad area (Figure 4.4). Finally, 2019's map is dominated by a broad high–high cluster stretching from Ar-Raqqah through Mayadin down to Abu Kamal, marking the fiercest non-state violence of the conflict.

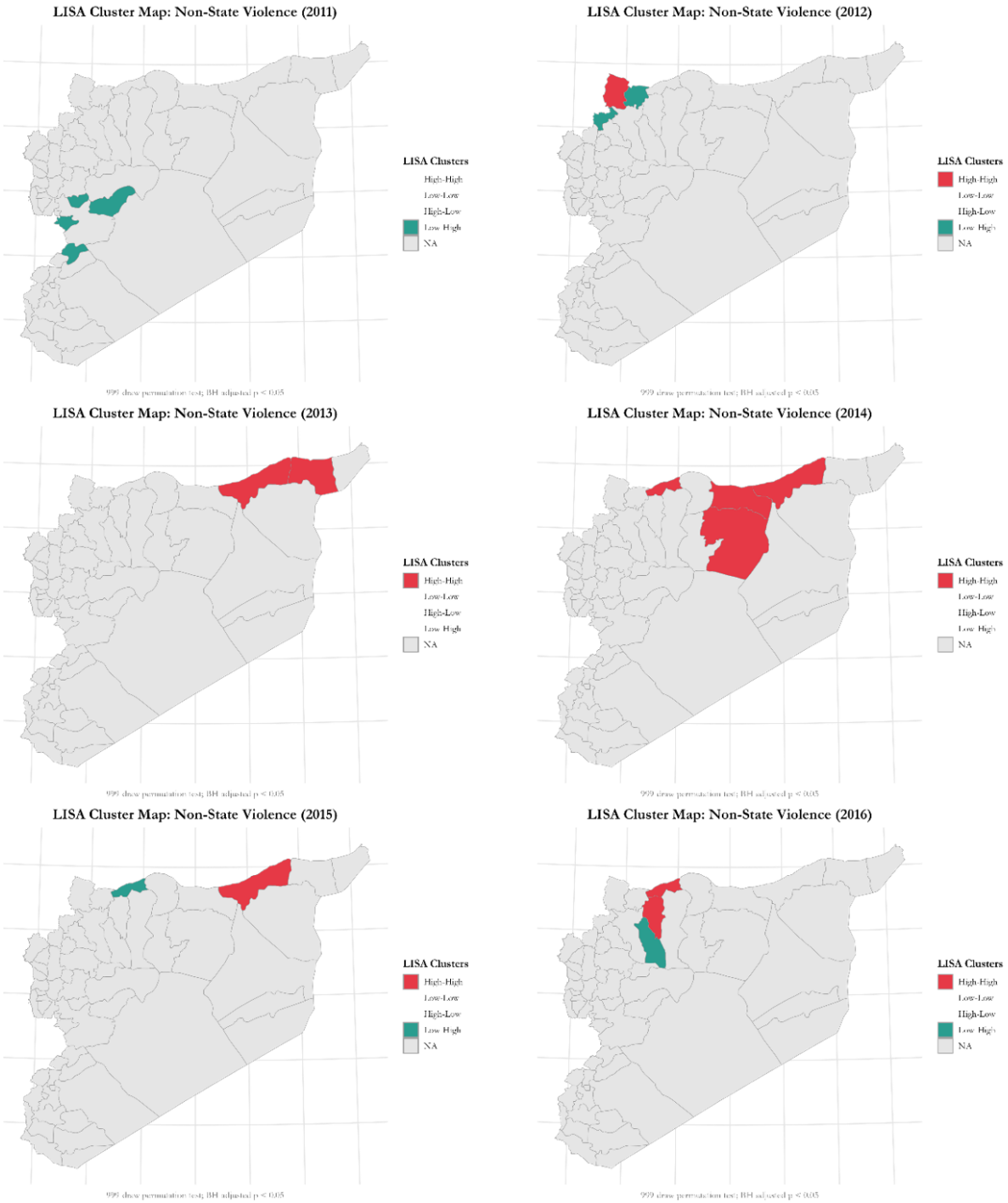

Figure 4.3. *LISA Cluster Map Sequence for Non-State Based Violence (2011-2016). Conflict data from UCDP GED (version 25.1) summarised by district-year non-state based violence. Administrative boundaries data sourced from GADM (version 4.1).*

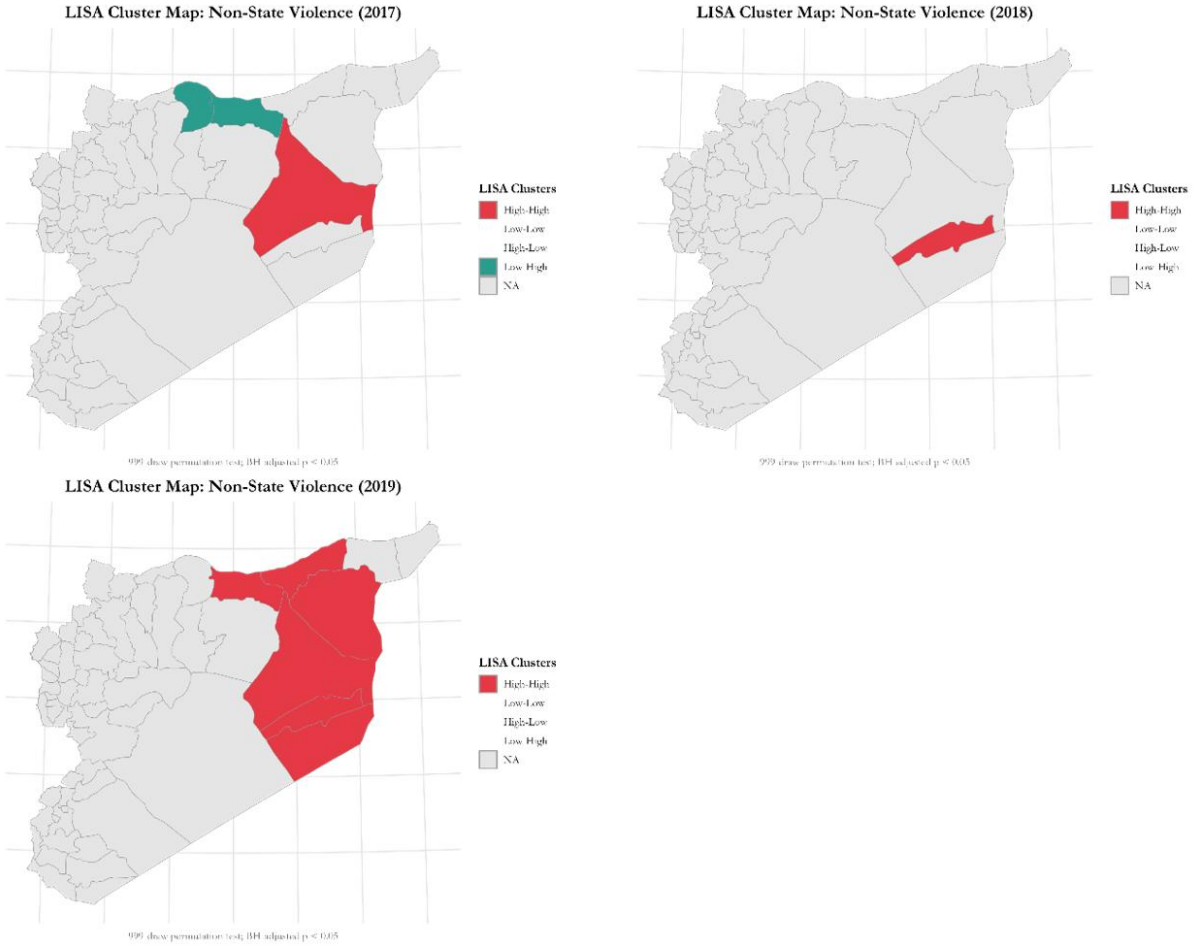

Figure 4.4. *LISA Cluster Map Sequence for Non-State Based Violence (2017-2019). Conflict data from UCDP GED (version 25.1) summarised by district-year non-state based violence. Administrative boundaries data sourced from GADM (version 4.1).*

In 2011, low–low clusters of one-sided killings appear near Al-Qusayr, Al-Mukharram, and Ar-Rastan (Figure 4.5). A mixed hot spot/cold spot pops up in the southwest in 2012. Significant hot spots occur in the southeast between 2015-17, and no district is significant in 2019, suggesting violence subsided (Figure 4.6).

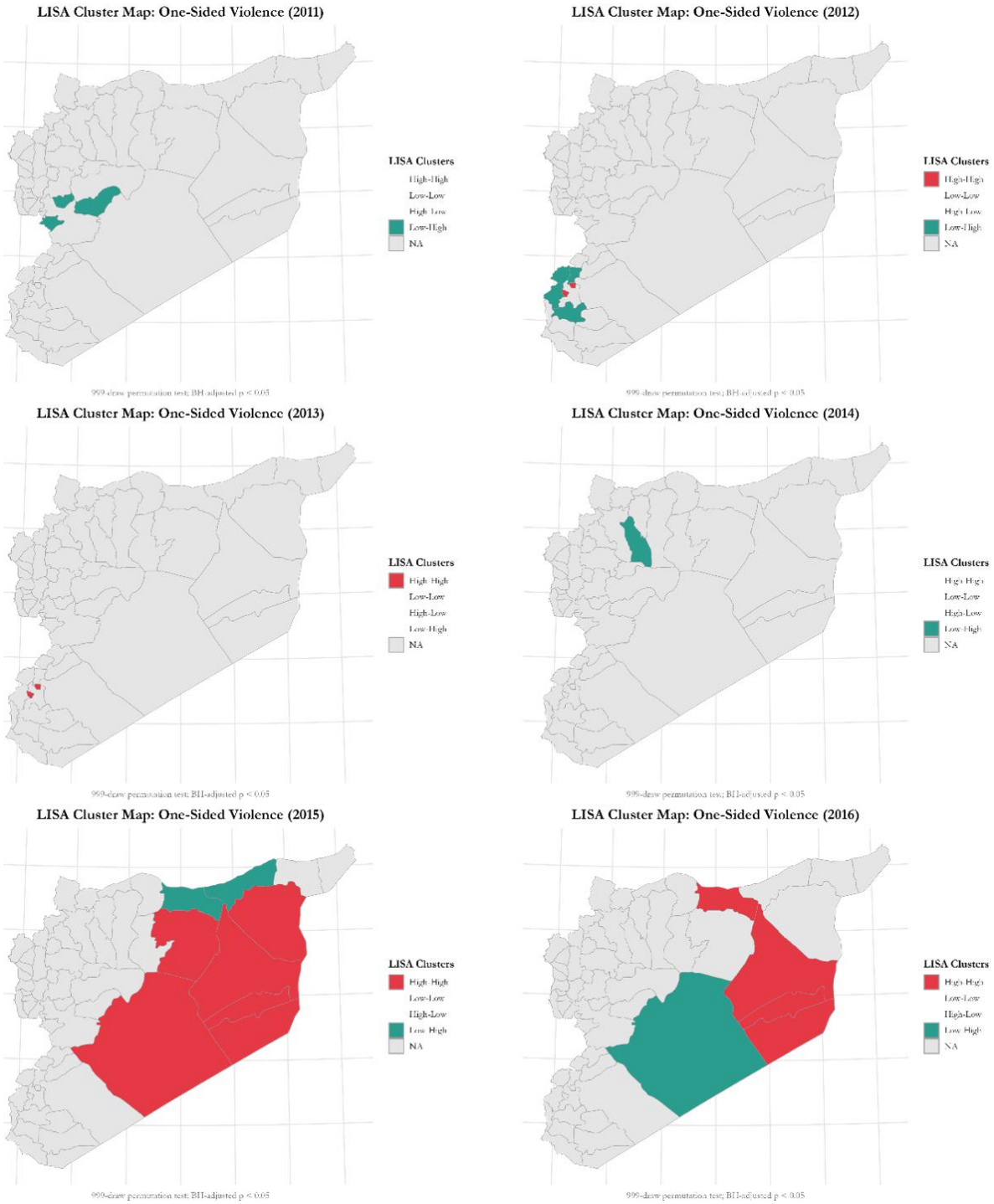

Figure 4.5. *LISA Cluster Map Sequence for One-Sided Violence (2011-2016). Conflict data from UCDP GED (version 25.1) summarised by district-year one-sided violence. Administrative boundaries data sourced from GADM (version 4.1).*

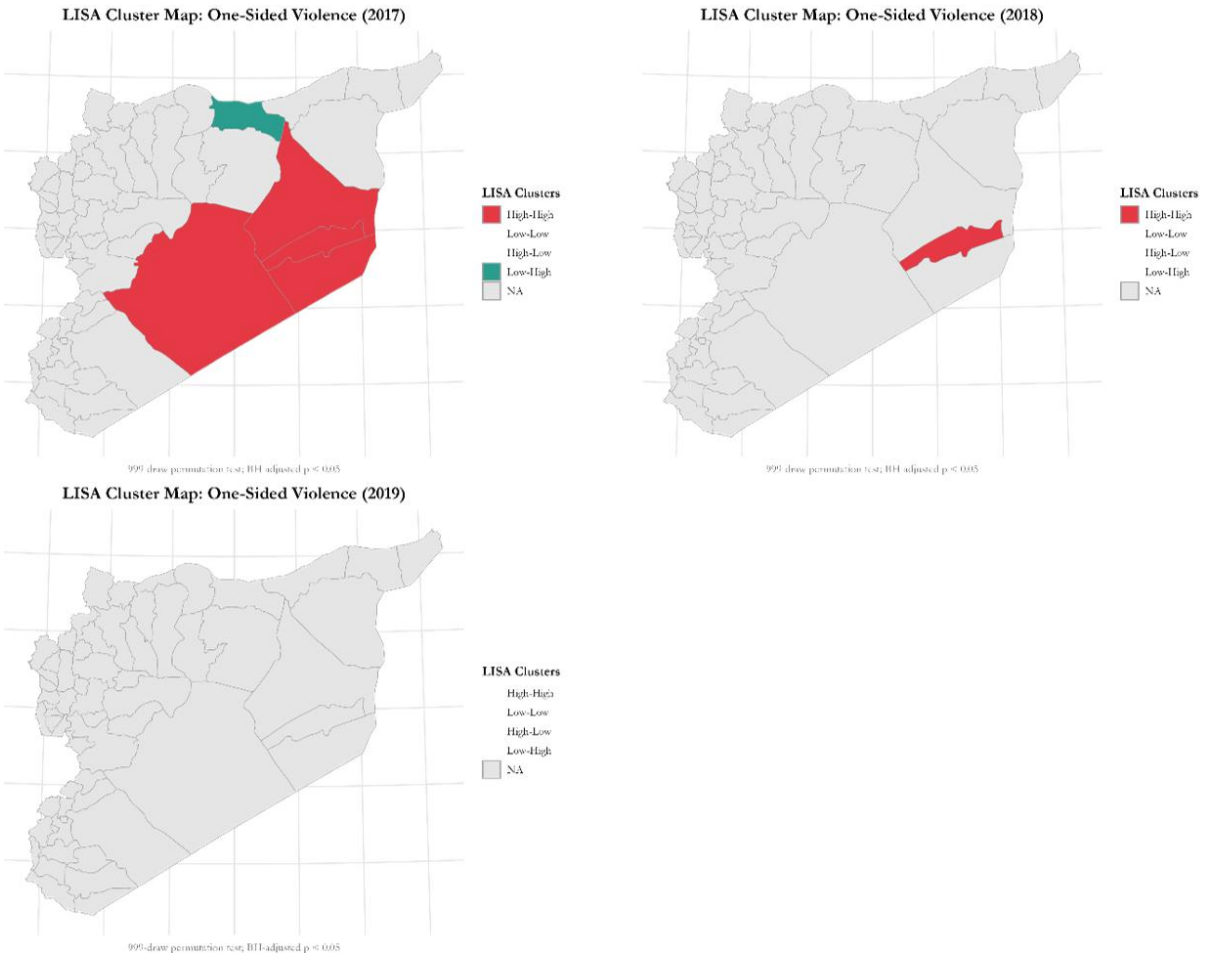

Figure 4.6. *LISA Cluster Map Sequence for One-Sided Violence (2017-2019). Conflict data from UCDP GED (version 25.1) summarised by district-year one-sided violence. Administrative boundaries data sourced from GADM (version 4.1).*

5. Expanded Methodology

5.1 Principal Component Analysis

To reduce dimensionality and collinearity among related predictors, two separate principal components analyses (PCA) were conducted using R's *prcomp(scale=TRUE)* (Jolliffe and Cadima, 2016). First, an infrastructure index was derived from the vector

$$\mathbf{X}_{\mathbf{infra},it} = (\mathbf{road\_dens}_{it}, \mathbf{pop\_dens}_{it}, \mathbf{urban\_built}_{it}),$$

where district is *i* and year is *t*, from which the first principal component

$$\mathbf{infra\_inx}_{it} = \mathbf{w}_{\mathbf{infra}}\mathbf{x}_{\mathbf{infra},it}$$

was extracted, explaining 91% of variance. A second index for land cover was obtained from six MODIS classes (croplands, shrublands, grasslands, savannas, mosaics, and barren/sparsely vegetated) yielding PC1 that accounted for 41% of variance. Each PC1 score was then mean-centred and standardised to a z-score via *scale()*, producing infra_idx_z and land_idx_z for modelling.

5.2 Residualisation of Human Development Indices

Governorate-level indices of income (I_{it}), education (E_{it}), and health (H_{it}) from SHDI first z-scored (Smits and Permanyer, 2019). Then, by regressing each on the other two, we obtained three residuals that capture the unique variance of each dimension (García et al., 2019):

$$\mathbf{E}_{\mathbf{it}} = \beta_0 + \beta_1\mathbf{I}_{it} + \beta_2\mathbf{H}_{it} + \varepsilon_{it}^{\mathbf{E\sim I,H}},$$

With $\text{edu\_resid\_inc}_{it} = \varepsilon_{it}^{E\sim I,H}$. Similar regressions produced six residual variables (edu_resid_inc, health_resid_inc, inc_resid_edu, health_resid_edu, inc_resid_hlt, edu_resid_hlt). This residualisation removes shared variance among SHDI dimensions, preventing inflated standard errors in the multivariate models.

5.3 Spatial Weights and Autocorrelation Diagnostics

Prior to regression, spatial clustering of conflict outcomes was quantified using both Global Moran's I and LISA. A queen-contiguity weights matrix $W=(\omega_{ij})$ was constructed with sped with *spdep (poly2nb + nb2listw)* (Bivand and Pebesma, 2023). Global Moran's I was calculated for each outcome in each year and tested via 999 random-label permutations *(moran.mc)*, yielding an exact upper-tail p-value (Bivand and Wong, 2018). Significant positive Moran's I indicated clustering of similar violence intensities. LISA statistics I_i were computed with 999 permutations and resulting p-values were adjusted using the Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995). Only districts with significant local clustering (high–high or low–low) after correction were mapped in *ggplot2* (Wickham, 2016).

5.4 Negative Binomial Mixed Effects Modeling

To accommodate overdispersed count outcomes and the nested structure of districts within governorates, mixed effects NB2 models were estimated in R with *glmmTMB* (Brooks et al., 2017). Let Y_{it} denote the count of a given type of violence (conflict count, estimated fatalities, state-based, non-state, and one-sided) in district i and year t. The model assumes

$$\mathbf{Y_{it} \sim NegBin(\mu_{it}, \kappa),\ Var(Y_{it}) = \mu_{it} + \kappa\mu_{it}^2,}$$

Consistent with the NB2 formulation (Hilbe, 2011).

With log-link,

$$\mathbf{\log(\mu_{it}) = \beta_0 + \beta_1 infra\_idx_{it} + \beta_2 land\_idx_{it} + \beta baseSHDI_{it} + \beta_4 res1_{it} + \beta_5 res2_{it} + \Sigma_{t=2012}^{(2019)} \gamma_t D_t + u_{g(i)},}$$

where:
- $infra\_idx_{it}$ and $land\_idx_{it}$ are the district-level infrastructure and land cover indices;
- $baseSHDI_{it}$, $res1_{it}$, $res2_{it}$ cycle through income, education, health, and their residuals;
- D_t are year dummy variables (2011 omitted as reference) with coefficients γ_t;
- $u_{g(i)} \sim N(0,\sigma^2)$ is a random intercept for governorate g containing district i (Gelman and Hill, 2007).

For non-state violence, characterised by excess zeros, a zero-inflated component was added:

$$\mathbf{Pr(Y_{it} = 0) = \pi + (1-\pi)\ (1+\kappa\mu_{it})^{-1/\kappa},}$$

where π was estimated via *(ziformula=~1)* (Lambert, 1992).

All models were fit using the Laplace approximation in glmmTMB (Brooks et al., 2017), and coefficients are reported as incidence-rate-ratios, $\exp(\beta)$ (Cameron and Trivedi, 2013).

5.5 Model Evaluation and Comparison

Model fit was assessed using AIC and BIC (Burnham and Anderson, 2002). Residual diagnostics employed DHARMa's simulation-based tests for overdispersion and zero inflation (Hartig, 2022) as well as performance's VIF check for collinearity (Lüdecke et al., 2024). Fixed effects counterparts, substituting factor(NAME_1) for (1|NAME_1), were run to gauge robustness; although fixed effects occasionally lowered AIC, mixed effects consistently yielded better BIC and permitted inclusion of between-governorate covariates (Wooldridge, 2021).

Detailed tables of AIC/BIC and fixed effects models results appear in Appendix 2 and 3. All analyses were conducted in R 4.4.2 (R Core Team, 2022) using key packages: *glmmTMB* v1.1.7 (Brooks et al., 2017), *spdep* v1.3-10 (Bivand & Pebesma, 2023), *DHARMa* v0.4.5 (Hartig, 2022), and *performance* v0.8.1 (Lüdecke et al., 2024).

6. Mixed Effects Model Residualised on Education and Health

The education-centred model reveals a strong and consistent relationship between education residuals and multiple forms of violence. Table 6.1 shows that districts whose educational attainment exceeds what their governorate's income and health levels would predict face substantially higher conflict rates. A one–standard–deviation increase in the education index corresponds to more than a three-fold rise in total conflict events (IRR = 3.10, $p = 0.004$) and a 3.38-fold increase in fatalities (IRR = 3.38, $p = 0.006$), with the strongest effects on non-state violence (IRR = 16.09, $p = 0.022$) and one-sided violence (IRR = 2.42, $p = 0.021$). These findings align with relative deprivation theory: when educational attainment rises faster than improvements in income or health, unmet political expectations and enhanced mobilisation capacity can fuel collective violence (Gurr, 1970).

By contrast, the health-centred specification positions the SHDI health index as the key predictor, with education and income residuals included as controls. Table 6.2 shows that a one-SD increase in the health index predicts a 90.7% reduction in non-state violence (IRR = 0.09, $p < .001$), a 52.5% decline in one-sided attacks (IRR = 0.48, $p < .001$), and significant decreases in both overall conflict counts and fatalities. This may suggest that resilient health systems signal effective governance and can help protect civilians.

Because these are observational regressions, reverse causality cannot be fully excluded; conflict may influence service provision just as much as service provision shapes conflict onset. Further research could include exploring different interaction terms (e.g., Education × Infrastructure) which may uncover thresholds where human capital only becomes conflict-inducing once mobility constraints disappear.

Table 6.1. *Effects of Education and Development Residuals on Conflict Outcomes*

Model	Term	Estimate (β)	Incident Rate Ratio	Std. Error	Statistic	P-value	Signif.
Estimated Fatalities	Education	1.217	3.377	0.446	2.726	0.006	**
Estimated Fatalities	Income Residual	-0.781	0.458	0.582	-1.341	0.180	
Estimated Fatalities	Health Residual	-0.381	0.683	0.280	-1.361	0.173	
Conflict Count	Education	1.131	3.099	0.391	2.890	0.004	**
Conflict Count	Income Residual	-0.883	0.414	0.524	-1.685	0.092	
Conflict Count	Health Residual	-0.381	0.683	0.256	-1.490	0.136	
State-Based	Education	0.706	2.203	0.421	1.679	0.093	
State-Based	Income Residual	-0.779	0.459	0.592	-1.315	0.188	
State-Based	Health Residual	-0.161	0.851	0.304	-0.529	0.597	
Non-State	Education	2.778	16.087	1.215	2.287	0.022	*
Non-State	Income Residual	-4.027	0.018	1.120	-3.594	0.000	***
Non-State	Health Residual	-0.778	0.460	0.505	-1.539	0.124	
One-Sided	Education	0.884	2.421	0.382	2.312	0.021	*
One-Sided	Income Residual	-0.376	0.687	0.444	-0.846	0.398	

One-Sided	Health Residual	-0.581	0.560	0.211	-2.759	0.006	**

Regression coefficients and exponentiated coefficients (IRRs) for education, income, and health residuals. Estimates are from count regression models; coefficients are on the log scale, and IRRs are obtained by exponentiating them.

Table 6.2. *Effects of Health and Development Residuals on Conflict Outcomes*

Model	Term	Estimate (β)	Incident Rate Ratio	Std. Error	Statistic	P-value	Signif.
Estimated Fatalities	Health	-0.707	0.493	0.255	-2.776	0.006	**
Estimated Fatalities	Income Residual	-0.781	0.458	0.582	-1.341	0.180	
Estimated Fatalities	Education Residual	0.693	1.200	0.358	1.936	0.053	
Conflict Count	Health	-0.742	0.476	0.268	-2.770	0.006	**
Conflict Count	Income Residual	-0.883	0.414	0.524	-1.685	0.092	
Conflict Count	Education Residual	0.539	1.714	0.337	1.600	0.110	
State-Based	Health	-0.472	0.624	0.324	-1.460	0.144	
State-Based	Income Residual	-0.779	0.459	0.592	-1.315	0.188	
State-Based	Education Residual	0.189	1.208	0.386	0.490	0.624	
Non-State	Health	-2.368	0.093	0.374	-6.329	0.000	***
Non-State	Income Residual	-4.027	0.018	1.120	-3.594	0.000	***
Non-State	Education Residual	0.105	1.111	0.931	0.112	0.911	
One-Sided	Health	-0.745	0.475	0.179	-4.169	0.000	***
One-Sided	Income Residual	-0.376	0.687	0.444	-0.846	0.398	
One-Sided	Education Residual	0.622	1.863	0.281	2.214	0.027	*

Regression coefficients and exponentiated coefficients (IRRs) for health, income, and education residuals. Health residuals consistently predict lower levels of violence, especially in non-state and one-sided conflict. Income residuals are strongly protective, while education residuals are positively associated with one-sided violence. Estimates are from count regression models using log links; IRRs are exponentiated coefficients.

7. Comparing Random vs. Fixed Effects for Residualised SHDI Indices

This section shows how RE and FE models treat residualised human development predictors differently. Allowing governorate intercepts to vary under RE preserves between-governate contrasts in income, education, and health while controlling for overdispersion (Cameron and Trivedi, 2013). By contrast, including a full set of governorate dummies under FE absorbs cross-sectional variation, inflates standard errors, and biases coefficient estimates toward zero (Greene, 2011). As a result, RE and FE yield different narratives about how the within-dimension deviations of education, health and income interact with patterns of violence in Syria over 2011-2019.

7.1 Base-Income Models

When income is entered at its observed value, the RE model (Research Paper, Table 5.3) reveals strong residual effects on one-sided violence. A one-standard-deviation positive education residual is associated with an 86% increase in one-sided events (IRR = 1.861, $p = 0.027$), while a one-standard-deviation positive health residual corresponds to a 44% decrease (IRR = 0.560, $p = 0.006$). Under FE, the education residual becomes non-significant (IRR = 0.687, $p = 0.398$), the the health residual remains important (IRR = 0.560, $p = 0.006$) (Appendix 3, Table 3.1).

7.2 Base-Education Models

Using education as the anchor, RE estimates in Table 6.1 indicate that higher-than-predicted education raises conflict across multiple outcomes: estimated fatalities (IRR = 3.377, $p = 0.006$), total conflict events (IRR = 3.099, $p = 0.004$), non-state violence (IRR = 16.087, $p = 0.022$), and one-sided violence (IRR = 2.421, $p = 0.021$). Under FE, these education residual effects are much smaller and mostly insignificant, except for non-state violence (IRR = 16.087, $p = 0.022$) and one-sided violence (IRR = 1.863, $p = 0.027$), illustrating how HE absorbs cross-governorate education contrasts schooling contrasts in four of the five outcomes (Appendix 3, Table 3.2).

7.3 Base-Health Models

When health is focal, a positive health residual strongly reduces violence under RE (Table 6.2): non-state violence (IRR = 0.093, $p < 0.001$), estimated fatalities (IRR = 0.493, $p = 0.006$, and total conflict events (IRR = 0.476, $p = 0.006$). In contrast, higher-than-expected income cuts non-state violence by 98% (IRR = 0.018, $p < 0.001$) under RE, but most of these effects disappear under FE.

Model-comparison metrics (AIC, BIC) are nearly identical for RE and FE runs on the same outcome (Appendix 2), implying that residualised indices offer interchangeable explanatory power once infrastructure, land cover, and year/governorate effects are controlled.

8. Model Diagnostics

8.1 Multicollinearity and Zero-Inflation

Table 8.1 reports two model-adequacy checks calculated for every candidate ZINB specification. Zero-inflation parameter (π) is the share of observations the model's logit gate classifies as structural zeros. The values are taken directly from the fitted *glmmTMB* objects (Brooks et al., 2017, v1.1.7) and correspond to the likelihood-ratio test implemented in *DHARMa::testZeroInflation()* (Hartig, 2024, v0.4.7). The condition number (κ), obtained via *performance::check_collinearity()* (Lüdecke et al., 2021, v0.15.0), gauges overall multicollinearity. Together, π pinpoints outcomes that truly need a zero-inflated link, while κ shows how multicollinearity escalates when governorate effects are treated as fixed rather than random.

High π values flag outcomes that genuinely need a zero-inflated link, for example, fixed-effects version of one-sided violence shows $\pi = 0.94$, implying that 94% of governorate–year cells are predicted as structural zeros. Conversely, fatalities have a negligible $\pi = 0.05$. Large κ values signal problematic multicollinearity; the fixed-effects version of the non-state model reaches $\kappa \approx 1.4 \times 10^6$, whereas its random-effects counterpart remains below 630, illustrating how fixed governorate dummies inflate collinearity.

The random-intercept ZINB models display moderate multicollinearity ($\kappa \approx 430$–$1{,}420$), which correspond to condition indices ($\sqrt{\kappa}$) of roughly 20-38, clustering around the threshold of 30 for moderate-high collinearity (Belsley et al., 1980). Nevertheless, the models converge cleanly, coefficient signs are sensible, and z-statistics remain in the 2–4 range, suggesting that parameter estimates are still reliable (Kutner et al., 2005).

Table 8.1: *Model Diagnostics for All Mixed-Effects and Fixed-Effects Models*

Model	Zero Inflation (π)	Collinearity (κ)	Condition Index ($\sqrt{\kappa}$)
fatalities_base_income_ran	0.048	766.56	27.69
conflicts_base_income_ran	0.176	1047.32	32.36
state_based_base_income_ran	0.48	1420.08	37.68
non_state_base_income_ran	0.712	625.76	25.01
one_sided_base_income_ran	0.712	428.81	20.71
fatalities_base_education_ran	0.024	766.57	27.69
conflicts_base_education_ran	0.184	1047.32	32.36
state_based_base_education_ran	0.544	1420.11	37.68
non_state_base_education_ran	0.688	625.76	25.01
one_sided_base_education_ran	0.712	428.82	20.71
fatalities_base_health_ran	0.056	766.57	27.69
conflicts_base_health_ran	0.232	1047.32	32.36
state_based_base_health_ran	0.584	1420.10	37.68
non_state_base_health_ran	0.688	625.76	25.01
one_sided_base_health_ran	0.712	428.81	20.71
fatalities_base_income_fix	0	38877.64	197.17

conflicts_base_income_fix	0.048	22961.19	151.53
state_based_base_income_fix	0.08	17214.95	131.21
non_state_base_income_fix	0.232	1382345.95	1175.73
one_sided_base_income_fix	0.936	23671.69	153.86
fatalities_base_education_fix	0	38875.53	197.17
conflicts_base_education_fix	0.048	22961.27	151.53
state_based_base_education_fix	0.08	17214.80	131.21
non_state_base_education_fix	0.24	1382297.40	1175.71
one_sided_base_education_fix	0.936	23671.89	153.86
fatalities_base_health_fix	0	38875.70	197.17
conflicts_base_health_fix	0.048	22961.43	151.53
state_based_base_health_fix	0.072	17214.62	131.20
non_state_base_health_fix	0.256	1382340.22	1175.73
one_sided_base_health_fix	0.936	23671.82	153.86

9. Limitations and Interpretative Caution

Reverse causality cannot be fully ruled out, as conflict may shape development as much as vice versa. The fitted models provide insight into conflict dynamics across health, education, and income-based baselines; however, interpretative caution is warranted by several diagnostic indicators. Extreme variance inflation factors (VIFs) in the fixed-effects specifications, which can reach into the tens or even hundreds of thousands, are evident in nearly all models, evidence of substantial collinearity. In addition to distorting coefficient estimates, this multicollinearity may also obscure the identification of independent predictor effects and inflate standard errors. The non-state and one-sided conflict outcomes are particularly characterized by moderate to high zero inflation in a number of models. These issues expose the necessity of a cautious interpretation of model outputs, particularly in relation to individual covariate effects. They may necessitate the use of more flexible modelling strategies in future work or re-specification (e.g., variable reduction, alternative distributions, interaction terms).

References

Anselin, L. (1995) 'Local Indicators of Spatial Association—LISA', *Geographical Analysis*, 27(2), pp. 93–115. doi: 10.1111/j.1538-4632.1995.tb00338.x.

Barron, C., Neis, P. and Zipf, A. (2013) 'A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis', *Transactions in GIS*, 18(6), pp. 877–895. doi: 10.1111/tgis.12073

Belsley, D.A., Kuh, E. and Welsch, R.E. (1980) *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: Wiley.

Benjamini, Y., and Hochberg, Y. (1995) 'Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*(1), 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x.

Brooks, M.E., Kristensen, K., van Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Maechler, M. and Bolker, B.M. (2017) 'glmmTMB Balances Speed and Flexibility Among Packages for Zero-Inflated Generalized Linear Mixed Modeling', *The R Journal*, 9(2), pp. 378–400. doi: 10.32614/RJ-2017-066.

Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2nd edn. New York: Springer.

Cameron, A.C. and Trivedi, P.K. (2013) *Regression Analysis of Count Data*. 2nd edn. Cambridge: Cambridge University Press.

Croicu, M. and Sundberg, R. (2017) *UCDP GED Codebook version 17.1*. Uppsala: Department of Peace and Conflict Research, Uppsala University.

Esri (2025) *ArcGIS Pro* [computer software], Version 3.5, Redlands, CA: Environmental Systems Research Institute. Available at: https://www.esri.com/en-us/arcgis/products/arcgis-pro/

Friedl, M.A. and Sulla-Menashe, D. (2022) *MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500 m SIN Grid V061*. Sioux Falls, South Dakota, USA: NASA Land Processes Distributed Active Archive Center. doi: 10.5067/MODIS/MCD12Q1.061.

GADM (2022) *Global Administrative Areas Database* (Version 4.1). Available at: https://gadm.org/download_country.html

García, C.B., Salmerón, R., García, C., and García, J. (2019) 'Residualization: Justification, Properties and Application', *Journal of Applied Statistics*, 47(11), 1990–2010. doi: 10.1080/02664763.2019.1701638.

Gelman, A., and Hill, J. (2007) *Data Analysis Using Regression and Multilevel/ Hierarchical Models*. Cambridge: Cambridge University Press.

Global Data Lab (2024) *Subnational Human Development Index (SHDI) Database (Version 8.1)*. Nijmegen: Institute for Management Research, Radboud University. Available at: https://globaldatalab.org/shdi/archive/

Global Data Lab (2025) *Subnational Human Development Index (SHDI): Methods*. Nijmegen: Institute for Management Research, Radboud University. Available at: https://globaldatalab.org/methods/

Greene, W.H. (2011) *Econometric Analysis*. 7th edn. Harlow: Pearson.

Gurr, T.R. (2011) *Why Men Rebel*. 1st edn. Abingdon: Routledge. doi: 10.4324/9781315631073.

Haklay, M. and Weber, P. (2008) 'OpenStreetMap: User-Generated Street Maps', *IEEE Pervasive Computing*, 7(4), pp. 12–18. doi: 10.1109/MPRV.2008.80.

Hilbe, J.M. (2011) *Negative Binomial Regression*. 2nd edn. Cambridge: Cambridge University Press.

Jolliffe, I.T. and Cadima, J. (2016) 'Principal Component Analysis: A Review and Recent Developments', *Philosophical Transactions of the Royal Society A*, 374(2065), 20150202. doi: 10.1098/rst.2015.0202.

Kutner, M.H., Nachtsheim, C.J., and Neter, J. (2005) *Applied Linear Regression Models*. 4th ed. New York: McGraw-Hill.

Lambert, D. (1992) 'Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. *Technometrics*, 34(1), pp.1–14. doi: 10.2307/1269547.

NASA Earthdata / LP DAAC (2022) *MODIS/Terra + Aqua Land Cover Type Yearly L3 Global 500 m SIN Grid, Version 6.1*. Available at: https://lpdaac.usgs.gov/products/mcd12q1v006/

OpenStreetMap Contributors (2025) *OpenStreetMap* [June 2025]. Available at: https://www.openstreetmap.org

Pettersson, T., Davies, S. and Öberg, M. (2023) 'Organized Violence 1989–2022 and the Return of Conflict Between States', *Journal of Peace Research*, 60(4), pp. 691–708. doi: 10.1177/00223433231185169.

Ramm, F., Topf, J. and Chilton, S. (2010) *OpenStreetMap: Using and Enhancing the Free Map of the World*. Cambridge: UIT Cambridge.

Smits, J. and Permanyer, I. (2019) 'The Subnational Human Development Database', *Scientific Data*, 6(1), 190038. doi: 10.1038/sdata.2019.38.

Stevens, F.R., Gaughan, A.E., Linard, C. and Tatem, A.J. (2015) 'Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data', *PLOS ONE*, 10(2), e0107042. doi: 10.1371/journal.pone.0107042.

Stina, H. (2025) *UCDP GED Codebook version 25.1*. Uppsala: Department of Peace and Conflict Research, Uppsala University.

Sundberg, R. and Melander, E. (2013) 'Introducing the UCDP Georeferenced Event Dataset (UCDP GED)', *Journal of Peace Research*, 50(4), pp. 523–532. doi: 10.1177/0022343313484347.

Thomson, D.R., Leasure, D.R., Bird, T., Tzavidis, N. and Tatem, A.J. (2022) 'How Accurate are WorldPop-Global-Unconstrained Gridded Population Data at the Cell-Level?: A Simulation Analysis in Urban Namibia', *PLOS ONE*, 17(7), e0271504. doi: 10.1371/journal.pone.0271504.

Wickham, H., François, R., Henry, L., Müller, K. & Vaughan, D. (2023) *dplyr: A Grammar of Data Manipulation* (v1.1.4) [online]. Available at: https://CRAN.R-project.org/package=dplyr.

Wooldridge, J.M. (2021) *Introductory Econometrics: A Modern Approach*, 8th edn. Boston: Cengage.

WorldPop (no date) *Gridded Population Estimate Datasets and Tools: Top-Down Constrained vs Unconstrained Modelling*. Southampton: University of Southampton. Available at: https://www.worldpop.org/methods/

Appendices

Appendix 1: R Code

Appendix 1.1: R Package, Version, and Reference

Package	Version	Source
dplyr	1.1.4	Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). dplyr: A Grammar of Data Manipulation. R package version 1.1.4, <https://CRAN.R-project.org/package=dplyr>.
tidyr	1.3.1	Wickham, H., Vaughan, D., & Girlich, M. (2024). tidyr: Tidy Messy Data. R package version 1.3.1, <https://CRAN.R-project.org/package=tidyr>.
readr	2.1.5	Wickham, H., & Hester, B.J. (2024). readr: Read Rectangular Text Data_. R package version 2.1.5, <https://CRAN.R-project.org/package=readr>.
sf	1.0.21	Pebesma, E., & Bivand, R. (2023). Spatial Data Science: With Applications in R. Chapman and Hall/CRC. https://doi.org/10.1201/9780429459016
spdep	1.3.13	Pebesma, E., & Bivand, R. (2023). Spatial Data Science: With Applications in R. Chapman and Hall/CRC. https://doi.org/10.1201/9780429459016
ggplot2	3.5.2	Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.
glue	1.8.0	Hester, B., & Bryan, J. (2024). glue: Interpreted String Literals. R package version 1.8.0, <https://CRAN.R-project.org/package=glue>.
stringr	1.5.1	Wickham, H. (2023). stringr: Simple, Consistent Wrappers for Common String Operations. R package version 1.5.1, <https://CRAN.R-project.org/package=stringr>.
magick	2.8.7	Ooms, J. (2025). magick: Advanced Graphics and Image-Processing in R. R package version 2.8.7, <https://CRAN.R-project.org/package=magick>.
purrr	1.1.0	Wickham, H., & Henry, L. (2025). purrr: Functional Programming Tools. R package version 1.1.0, <https://CRAN.R-project.org/package=purrr>.
tibble	3.3.0	Müller, K., & Wickham, H. (2025). tibble: Simple Data Frames. R package version 3.3.0, <https://CRAN.R-project.org/package=tibble>.
tidyverse	2.0.0	Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H. (2019). "Welcome to the tidyverse." Journal of Open Source Software, *4*(43), 1686. doi:10.21105/joss.01686 <https://doi.org/10.21105/joss.01686>.
broom	1.0.9	Robinson, D., Hayes, A., & Couch, S. (2025). broom: Convert Statistical Objects into Tidy Tibbles. R package version 1.0.9, <https://CRAN.R-project.org/package=broom>.
broom.mixed	0.2.9.6	Bolker, B., & Robinson, D. (2024). broom.mixed: Tidying Methods for Mixed Models. R package version 0.2.9.6, <https://CRAN.R-project.org/package=broom.mixed>.
lme4	1.1.37	Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.
glmmTMB	1.1.11	Brooks, M.E., Kristensen, K., van Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Maechler, M., & Bolker, B.M. (2017). glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. The R Journal, 9(2), 378-400. doi:10.32614/RJ-2017-066.
MASS	7.3.61	Venables, W.N.. & Ripley, B.D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0
performance	0.15.0	Lüdecke, D., Ben-Shachar, M.S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. Journal of Open Source Software, 6(60), 3139. https://doi.org/10.21105/joss.03139
DHARMa	0.4.7	Hartig, F. (2024). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.4.7, <https://CRAN.R-project.org/package=DHARMa>.

writexl	1.5.4	Ooms, J. (2025). writexl: Export Data Frames to Excel 'xlsx' Format. R package version 1.5.4, <https://CRAN.R-project.org/package=writexl>.
showtext	0.9.7	Qiu, Y. (2024). showtext: Using Fonts More Easily in R Graphs. R package version 0.9-7, <https://CRAN.R-project.org/package=showtext>.
ggrepel	0.9.6	Slowikowski, K. (2024). ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'. R package version 0.9.6, <https://CRAN.R-project.org/package=ggrepel>.

Appendix 1.2: Conflict Summary District

```
# Load Libraries
pkgs <- c("dplyr", "tidyr", "readr")
invisible(lapply(pkgs, require, character.only = TRUE))

# Define Directories for Input and Output
data_dir <- "M:/dissfinal/Data/Data Acquisition/Uppsala Conflict Data"
out_dir  <- "M:/dissfinal/Data"
dir.create(out_dir, showWarnings = FALSE, recursive = TRUE)
p <- function(...) file.path(data_dir, ...)

# Read in Data
conflicts <- read_csv(p("cleanconflictdata.csv")) %>%
  filter(year >= 2011 & year <= 2019)

# Aggregate Summary Statistics by District-Year
# Conflict Counts
conflict_counts <- conflicts %>%
  group_by(NAME_2, year) %>%
  summarise(conflict_count = n(), .groups = "drop")

# Total Estimated Fatalities
fatalities <- conflicts %>%
  group_by(NAME_2, year) %>%
  summarise(total_fatalities = sum(best_est, na.rm = TRUE), .groups =
"drop")

# Violence Type Counts
violence_type_counts <- conflicts %>%
  group_by(NAME_2, year, type_of_violence) %>%
  summarise(count = n(), .groups = "drop") %>%
  pivot_wider(
    names_from = type_of_violence,
    values_from = count,
    values_fill = 0,
    names_prefix = "violence_type_"
  )

# Merge Datasets
district_year_summary <- conflict_counts %>%
  left_join(fatalities, by = c("NAME_2", "year")) %>%
  left_join(violence_type_counts, by = c("NAME_2", "year"))

# Save CSV
write_csv(district_year_summary, file.path(out_dir,
"conflict_district_summary.csv"))
```

Appendix 1.3: Global Moran's I and LISA Cluster

```
# Load Libraries
pkgs <- c(
  "sf", "spdep", "dplyr", "ggplot2", "glue", "readr", "stringr",
  "magick", "purrr", "tibble"
)
invisible(lapply(pkgs, require, character.only = TRUE))

# Read in Data
shp_path <- "M:/dissfinal/Data/Data Acquisition/Syria
Shapefile/SyriaLevel2.shp"
shp <- sf::st_read(shp_path) %>% sf::st_make_valid()

panel <- readr::read_csv(
  "M:/Data/conflict_district_summary.csv"
)

# Join Shapefile and Conflict Summary Data
standardize <- function(x) {
  x %>%
    stringr::str_to_lower() %>%
    stringr::str_replace_all("-", " ") %>%
    stringr::str_squish()
}
shp   <- shp   %>% mutate(NAME_2 = standardize(NAME_2))
panel <- panel %>% mutate(NAME_2 = standardize(NAME_2))

# Define Variables
vars <- c(
  one_sided       = "One-Sided Violence",
  non_state_based = "Non-State Violence",
  state_based     = "State-Based Violence",
  conflict_count = "Conflict Count",
  total_fatalities = "Estimated Fatalities"

)

# Global Moran's I
moran_results <- map_dfr(
  .x = names(vars),
  .f = function(var) {
    map_dfr(
      .x = sort(unique(panel$year)),
      .f = function(yr) {
        # Subset & join
        dat_y <- panel %>% filter(year == yr)
        shp_y <- left_join(shp, dat_y, by = "NAME_2") %>%
          filter(!is.na(.data[[var]]))

        # Build neighbours & weights
        nb <- spdep::poly2nb(shp_y, queen = TRUE)
        lw <- spdep::nb2listw(nb, style = "W", zero.policy = TRUE)

        # Analytical Moran's I
        ana <- spdep::moran.test(shp_y[[var]], lw, zero.policy = TRUE)
        # Permutation Moran's I
```

```
        mc  <- spdep::moran.mc(  shp_y[[var]], lw, nsim = 999, zero.policy
= TRUE)

        # Tidy results
        tibble(
          variable      = var,
          year          = yr,
          I_statistic   = as.numeric(ana$estimate["Moran I statistic"]),
          I_exp         = as.numeric(ana$estimate["Expectation"]),
          I_var         = as.numeric(ana$estimate["Variance"]),
          p_value_anal  = ana$p.value,
          p_value_perm  = mc$p.value
        )
      }
    )
  }
)

# Global Moran's I Table
write_csv(moran_results, file.path("M:/dissfinal/Outputs",
"global_moransI.csv"))

# LISA Maps
cluster_pal <- c(
  "High-High" = "#e63946",
  "Low-Low"   = "#1d3557",
  "High-Low"  = "#f4a261",
  "Low-High"  = "#2a9d8f"
)
out_dir <- "C:/Users/skwin/Dissertation/Outputs/LISA_Maps"
dir.create(out_dir, showWarnings = FALSE, recursive = TRUE)

run_lisa <- function(var, label, year) {
  dat_y <- panel %>% filter(year == !!year)
  shp_y <- left_join(shp, dat_y, by = "NAME_2") %>%
    filter(!is.na(.data[[var]]))

  nb <- poly2nb(shp_y, queen = TRUE)
  lw <- nb2listw(nb, style = "W", zero.policy = TRUE)

  lisa <- localmoran_perm(shp_y[[var]], listw = lw, nsim = 999, zero.policy
= TRUE)

  shp_y <- shp_y %>%
    mutate(
      lisa_p     = lisa[, "Pr(z != E(Ii))"],
      lisa_p_adj = p.adjust(lisa_p, method = "BH"),
      z          = scale(.data[[var]]),
      z_lag      = scale(lag.listw(lw, .data[[var]], zero.policy = TRUE)),
      cluster    = case_when(
        z > 0 & z_lag > 0 & lisa_p_adj < 0.05 ~ "High-High",
        z < 0 & z_lag < 0 & lisa_p_adj < 0.05 ~ "Low-Low",
        z > 0 & z_lag < 0 & lisa_p_adj < 0.05 ~ "High-Low",
        z < 0 & z_lag > 0 & lisa_p_adj < 0.05 ~ "Low-High",
        TRUE                                  ~ NA_character_
      ),
      cluster = factor(cluster, levels = names(cluster_pal))
```

```
    )

  p <- ggplot(shp_y) +
    geom_sf(aes(fill = cluster), color = "grey60", size = 0.2) +
    scale_fill_manual(values = cluster_pal, na.value = "grey90", drop =
FALSE, name = "LISA Clusters") +
    labs(
      title   = glue("LISA Cluster Map: {label} ({year})"),
      caption = "999-draw permutation test; BH-adjusted p < 0.05"
    ) +
    theme_minimal(base_family = "Garamond") +
    theme(
      plot.title   = element_text(size = 16, face = "bold", hjust = 0.5),
      plot.caption = element_text(size = 10, hjust = 0.5, colour =
"grey40"),
      legend.title = element_text(size = 12, face = "bold"),
      legend.text  = element_text(size = 10),
      panel.grid.major = element_line(colour = "grey90"),
      axis.text  = element_blank(),
      axis.title = element_blank(),
      axis.ticks = element_blank()
    )

  fname <- file.path(out_dir, glue("lisa_{var}_{year}.png"))
  ggsave(fname, p, width = 8, height = 6, dpi = 300)
  message("Saved: ", basename(fname))
}

# Generate Maps
years <- 2011:2019
for (v in names(vars)) {
  for (yr in years) run_lisa(v, vars[[v]], yr)
}

# Panel Maps
stitch_maps <- function(var) {
  files <- glue("{out_dir}/lisa_{var}_{years}.png")
  imgs  <- lapply(files, function(f) if (file.exists(f))
magick::image_read(f) else NULL) %>%
    Filter(Negate(is.null), .)
  if (length(imgs) == 0) stop(glue("No images for {var}"))

  rows <- lapply(seq(1, length(imgs), by = 2), function(i) {
    pair <- imgs[i:min(i+1, length(imgs))]
    if (length(pair) == 1) {
      blank <- magick::image_blank(
        width  = magick::image_info(pair[[1]])$width,
        height = magick::image_info(pair[[1]])$height,
        color  = "none"
      )
      pair <- c(pair, blank)
    }
    magick::image_append(magick::image_join(pair), stack = FALSE)
  })
  panel <- magick::image_append(magick::image_join(rows), stack = TRUE)

  outp <- file.path(out_dir, glue("lisa_{var}_panel.png"))
```

```
  magick::image_write(panel, outp)
  message("Panel written: ", basename(outp))
}
for (v in names(vars)) stitch_maps(v)
```

Appendix 1.4: Negative Binomial Models and Diagnostics

```
# Load Libraries
pkgs <- c(
  "tidyverse", "broom", "broom.mixed", "lme4", "glmmTMB", "MASS",
  "performance", "DHARMa", "sf", "spdep", "glue", "writexl",
  "showtext", "ggrepel", "ggplot2"
)
invisible(lapply(pkgs, require, character.only = TRUE))

# Set theme with Garamond
ggplot2::theme_set(ggplot2::theme_minimal(base_family = "Garamond"))
set.seed(42)

# Define Directories for Input and Output
data_dir <- "M:/dissfinal/Data"
out_dir  <- "M:/dissfinal/Outputs"
dir.create(out_dir, showWarnings = FALSE, recursive = TRUE)
p <- function(...) file.path(data_dir, ...)

# Read in Clean Datasets
land_cover <- readr::read_csv(p("clean_land_cover.csv"))
socio      <- readr::read_csv(p("cleansocioeconomicdata.csv"))
conflict   <- readr::read_csv(p("conflict_district_summary.csv"))
pop_road   <- readr::read_csv(p("pop_dens_road_dens.csv"))

# Land Cover to Wide
land_long <- land_cover %>%
  pivot_longer(matches("^\\d{4}_pct_"),
               names_to  = c("year", "class"),
               names_sep = "_pct_",
               values_to = "pct") %>%
  mutate(year = as.integer(year)) %>%
  pivot_wider(names_from = class, values_from = pct)

# Population and Roads to Long
pop_long <- pop_road %>%
  pivot_longer(matches("^(pop|dens)_\\d{4}$"),
               names_to      = c(".value", "year"),
               names_pattern = "(.*)_(\\d{4})") %>%
  mutate(year = as.integer(year))

# SHDI to Long
socio_long <- socio %>%
  pivot_longer(-NAME_1,
               names_to      = c("year", ".value"),
               names_pattern = "(\\d{4})_(.*)") %>%
  mutate(year = as.integer(year))

# Merge All Datasets
panel <- land_long %>%
  left_join(pop_long,  by = c("NAME_1", "NAME_2", "year")) %>%
  left_join(conflict,  by = c("NAME_1", "NAME_2", "year")) %>%
  left_join(socio_long, by = c("NAME_1", "year"))

rm(land_cover, pop_road, socio, conflict)
```

```
# Construct Infrastructure and Land Cover Indices via PCA
safe_pca <- function(df){
  ok <- complete.cases(df)
  if (sum(ok) < 2) return(rep(NA_real_, nrow(df)))
  pr <- prcomp(df[ok, ], scale. = TRUE)
  res <- rep(NA_real_, nrow(df)); res[ok] <- pr$x[, 1]
  scale(res)[, 1]
}

# Columns to Include in PCA
infra_cols <- intersect(names(panel), c("dens", "road_densi", "road_dens",
"urban_built_up"))
land_cols  <- intersect(names(panel), c(
  "croplands", "barren_sparsely_vegetated", "open_shrublands",
"grasslands",
  "cropland_natural_vegetation_mosaics", "savannas"))

# Add PCA Indices to Panel
panel <- panel %>%
  mutate(
    infra_idx = safe_pca(dplyr::pick(all_of(infra_cols))),
    land_idx  = safe_pca(dplyr::pick(all_of(land_cols)))
  )

# Standardise SHDI, Infrastructure, and Land Cover Indices

z_vars <- c("infra_idx", "land_idx", "income_index", "educational_index",
"health_index")
panel  <- panel %>%
  mutate(across(all_of(z_vars), ~ as.numeric(scale(.x)), .names =
"{.col}_z"))

# Residualisation to Remove Multicollinearity
resid_on <- function(y, x, d) resid(lm(reformulate(x, y), data = d))

# Compute Residuals for Each Component
panel$edu_resid_inc    <- resid_on("educational_index_z", "income_index_z",
panel)
panel$health_resid_inc <- resid_on("health_index_z",      "income_index_z",
panel)
panel$inc_resid_edu    <- resid_on("income_index_z",
"educational_index_z", panel)
panel$health_resid_edu <- resid_on("health_index_z",
"educational_index_z", panel)
panel$inc_resid_hlt    <- resid_on("income_index_z",      "health_index_z",
panel)
panel$edu_resid_hlt    <- resid_on("educational_index_z", "health_index_z",
panel)

# Outcome Variables
outcomes <- c(
  fatalities  = "total_fatalities",
  conflicts   = "conflict_count",
  state_based = "state_based",
  non_state   = "non_state_based",
  one_sided   = "one_sided"
)
```

```
# SHDI-Based Model Specifications
base_specs <- list(
  income    = list(base = "income_index_z",        res1 = "edu_resid_inc",
res2 = "health_resid_inc"),
  education = list(base = "educational_index_z",   res1 = "inc_resid_edu",
res2 = "health_resid_edu"),
  health    = list(base = "health_index_z",        res1 = "inc_resid_hlt",
res2 = "edu_resid_hlt")
)

#Negative Binomial Model Using glmmTMB
fit_nb <- function(outcome, spec, random = TRUE, zi = FALSE){
  rhs <- paste(
    "infra_idx_z + land_idx_z +",
    spec$base, "+", spec$res1, "+", spec$res2,
    "+ factor(year)", if (random) "+ (1|NAME_1)" else "+ factor(NAME_1)"
  )
  form <- as.formula(paste(outcome, rhs, sep = " ~ "))

  glmmTMB::glmmTMB(
    formula   = form,
    data      = panel,
    family    = glmmTMB::nbinom2(link = "log"),
    ziformula = if (zi) ~1 else ~0
  )
}

# Mixed Effects Models
models_random <- list()
for (b in names(base_specs)) {
  ms <- purrr::imap(outcomes, ~ fit_nb(.x, base_specs[[b]], random = TRUE,
zi = (.y == "non_state")))
  names(ms) <- paste(names(ms), "base", b, "ran", sep = "_")
  models_random <- c(models_random, ms)
}

#Fixed Effects Models
models_fixed <- list()
for (b in names(base_specs)) {
  ms <- purrr::imap(outcomes, ~ fit_nb(.x, base_specs[[b]], random = FALSE,
zi = (.y == "non_state")))
  names(ms) <- paste(names(ms), "base", b, "fix", sep = "_")
  models_fixed <- c(models_fixed, ms)
}

# Tidy Model Results
tidy_tbl_random <- purrr::map_dfr(models_random, broom.mixed::tidy, effects
= "fixed", .id = "model")
tidy_tbl_fixed  <- purrr::map_dfr(models_fixed,  broom.mixed::tidy, effects
= "fixed", .id = "model")

# Annotate Significance Levels
auto_sig <- function(df) df %>%
  mutate(signif = dplyr::case_when(
    p.value < .001 ~ "***",
    p.value < .01  ~ "**",
```

```
    p.value < .05  ~ "*",
    p.value < .1   ~ ".",
    TRUE           ~ ""
  ))

tidy_tbl_random <- auto_sig(tidy_tbl_random)
tidy_tbl_fixed  <- auto_sig(tidy_tbl_fixed)

# Export to Excel
writexl::write_xlsx(
  list(
    random_effects = tidy_tbl_random,
    fixed_effects  = tidy_tbl_fixed
  ),
  file.path(out_dir, "nb_models_random_vs_fixed.xlsx")
)

# Model Diagnostics
diag_summary <- function(m){
  sim <- DHARMa::simulateResiduals(m)
  tibble::tibble(
    overdispersion = performance::check_overdispersion(m)$p.value,
    zero_inflation = DHARMa::testZeroInflation(sim)$p.value,
    collinearity   = performance::check_collinearity(m)$VIF %>% max(na.rm =
TRUE)
  )
}

diag_tbl <- dplyr::bind_rows(
  purrr::imap_dfr(models_random, ~ diag_summary(.x) %>% dplyr::mutate(model
= .y, spec = "random")),
  purrr::imap_dfr(models_fixed,  ~ diag_summary(.x) %>% dplyr::mutate(model
= .y, spec = "fixed"))
)

# Export to Excel
writexl::write_xlsx(
  list(diagnostics = diag_tbl),
  file.path(out_dir, "nb_model_diagnostics.xlsx")
)

# Null Models to Compare AIC and BIC
fit_null_nb <- function(outcome, random = TRUE, zi = FALSE){
  rhs <- paste("1 + factor(year)", if (random) "+ (1|NAME_1)" else "+
factor(NAME_1)")
  form <- as.formula(paste(outcome, rhs, sep = " ~ "))

  glmmTMB::glmmTMB(
    formula   = form,
    data      = panel,
    family    = glmmTMB::nbinom2(link = "log"),
    ziformula = if (zi) ~1 else ~0
  )
}

null_models <- list()
for (b in names(base_specs)) {
```

```
  for (y in names(outcomes)) {
    outcome_var <- outcomes[[y]]
    is_zi <- y == "non_state"
    name_r <- paste(y, "null", b, "ran", sep = "_")
    name_f <- paste(y, "null", b, "fix", sep = "_")
    null_models[[name_r]] <- fit_null_nb(outcome_var, TRUE, is_zi)
    null_models[[name_f]] <- fit_null_nb(outcome_var, FALSE, is_zi)
  }
}

# Extract AIC and BIC for Comparison
build_ic_table <- function(models_subset, type_label) {
  imap_dfr(models_subset, function(m, name) {
    parts <- strsplit(name, "_")[[1]]
    outcome <- parts[1]
    spec <- ifelse("base" %in% parts, parts[which(parts == "base") + 1],
                   ifelse("null" %in% parts, parts[which(parts == "null") +
1], "unknown"))
    model_type <- ifelse("null" %in% parts, "null_model", type_label)

    tibble(
      model_name = name,
      outcome = outcome,
      spec = spec,
      model_type = model_type,
      AIC = AIC(m),
      BIC = BIC(m)
    )
  }) %>%
    pivot_wider(
      names_from = model_type,
      values_from = c(AIC, BIC),
      names_glue = "{.value}_{model_type}"
    )
}

# Extract Random and Fixed Models and Null Counterparts
random_and_null_models <- c(models_random, null_models)[grepl("_ran$",
names(c(models_random, null_models)))]
fixed_and_null_models  <- c(models_fixed,  null_models)[grepl("_fix$",
names(c(models_fixed,  null_models)))]

# Build Tables
ic_tbl_random <- build_ic_table(random_and_null_models, "random_effects")
ic_tbl_fixed  <- build_ic_table(fixed_and_null_models,  "fixed_effects")

# Export to Excel
writexl::write_xlsx(
  list(
    random_effects_ic = ic_tbl_random,
    fixed_effects_ic  = ic_tbl_fixed
  ),
  file.path(out_dir, "nb_model_info_criteria.xlsx")
)
```

Appendix 2: AIC and BIC Values

AIC and BIC values for each model and its corresponding null model.

Table 2.1. Mixed Effects AIC and BIC Values Against AIC and BIC Null Models

Outcome	Specification	AIC	AIC Null	BIC	BIC Null
Fatalities	Income	6968.00	7001.32	7036.67	7048.53
Conflict	Income	5651.68	5682.62	5720.35	5729.82
State	Income	5165.94	5182.34	5234.60	5229.55
Non-State	Income	2263.27	2285.10	2336.23	2336.60
One-Sided	Income	2660.57	2700.35	2729.23	2747.56
Fatalities	Education	6968.00	7001.32	7036.67	7048.53
Conflict	Education	5651.68	5682.62	5720.35	5729.82
State	Education	5165.94	5182.34	5234.60	5229.55
Non-State	Education	2263.27	2285.10	2336.23	2336.60
One-Sided	Education	2660.57	2700.35	2729.23	2747.56
Fatalities	Health	6968.00	7001.32	7036.67	7048.53
Conflict	Health	5651.68	5682.62	5720.35	5729.82
State	Health	5165.94	5182.34	5234.60	5229.55
Non-State	Health	2263.27	2285.10	2336.23	2336.60
One-Sided	Health	2660.57	2700.35	2729.23	2747.56

Table 2.2. Fixed Effects AIC and BIC Values Against AIC and BIC Null Models

Outcome	Specification	AIC	AIC Null	BIC	BIC Null
Fatalities	Income	6943.17	6974.39	7059.04	7068.80
Conflict	Income	5620.08	5649.80	5735.95	5744.22
State	Income	5131.14	5147.68	5247.01	5242.10
Non-State	Income	2242.00	2240.90	2362.16	2339.61
One-Sided	Income	2642.96	2672.84	2758.83	2767.25
Fatalities	Education	6943.17	6974.39	7059.04	7068.80
Conflict	Education	5620.08	5649.80	5735.95	5744.22
State	Education	5131.14	5147.68	5247.01	5242.10
Non-State	Education	2242.00	2240.90	2362.16	2339.61
One-Sided	Education	2642.96	2672.84	2758.83	2767.25
Fatalities	Health	6943.17	6974.39	7059.04	7068.80
Conflict	Health	5620.08	5649.80	5735.95	5744.22
State	Health	5131.14	5147.68	5247.01	5242.10
Non-State	Health	2242.00	2240.90	2362.16	2339.61
One-Sided	Health	2642.96	2672.84	2758.83	2767.25

Appendix 3: Fixed Effects Model Results

Fixed effects model outcomes including estimate coefficients, standard error, statistic, p-value, and significance information.

Table 3.1. *Effects of Income and Development Residuals on Conflict Outcomes*

Model	Term	Estimate	Std. Error	Statistic	P-value	Signif.
Estimated Fatalities	Income	-1.633	1.081	-1.511	0.131	
Estimated Fatalities	Education Residual	0.979	1.164	0.841	0.400	
Estimated Fatalities	Health Residual	-0.314	0.724	-0.433	0.665	
Conflict Count	Income	1.131	0.391	2.890	0.004	**
Conflict Count	Education Residual	-0.883	0.524	-1.685	0.092	
Conflict Count	Health Residual	-0.381	0.256	-1.490	0.136	
State-Based	Income	0.706	0.421	1.679	0.093	
State-Based	Education Residual	-0.779	0.592	-1.315	0.188	
State-Based	Health Residual	-0.161	0.304	-0.529	0.597	
Non-State	Income	2.778	1.215	2.287	0.022	*
Non-State	Education Residual	-4.027	1.120	-3.594	0.000	***
Non-State	Health Residual	-0.778	0.505	-1.539	0.124	
One-Sided	Income	0.884	0.382	2.312	0.021	*
One-Sided	Education Residual	-0.376	0.444	-0.846	0.398	
One-Sided	Health Residual	-0.581	0.211	-2.759	0.006	**

Regression coefficients for income and education and health residuals. Higher income reduces most forms of violence, while positive education residuals increase one-sided violence.

Table 3.2. *Effects of Education and Development Residuals on Conflict Outcomes*

Model	Term	Estimate	Std. Error	Statistic	P-value	Signif.
Estimated Fatalities	Education	1.217	0.446	2.726	0.006	**
Estimated Fatalities	Income Residual	-0.781	0.582	-1.341	0.180	
Estimated Fatalities	Health Residual	-0.381	0.280	-1.361	0.173	
Conflict Count	Education	1.131	0.391	2.890	0.004	**
Conflict Count	Income Residual	-0.883	0.524	-1.685	0.092	
Conflict Count	Health Residual	-0.381	0.256	-1.490	0.136	
State-Based	Education	0.706	0.421	1.679	0.093	
State-Based	Income Residual	-0.779	0.592	-1.315	0.188	
State-Based	Health Residual	-0.161	0.304	-0.529	0.597	
Non-State	Education	2.778	1.215	2.287	0.022	*
Non-State	Income Residual	-4.027	1.120	-3.594	0.000	***
Non-State	Health Residual	-0.778	0.505	-1.539	0.124	
One-Sided	Education	0.884	0.382	2.312	0.021	*
One-Sided	Income Residual	-0.376	0.444	-0.846	0.398	

One-Sided	Health Residual	-0.581	0.211	-2.759	0.006	**

Regression coefficients for income and education and health residuals. Higher income reduces most forms of violence, while positive education residuals increase one-sided violence.

Table 3.3. *Effects of Health and Development Residuals on Conflict Outcomes*

Model	Term	Estimate	Std. Error	Statistic	P-value	Signif.
Estimated Fatalities	Health	-0.707	0.255	-2.776	0.006	**
Estimated Fatalities	Income Residual	-0.781	0.582	-1.341	0.180	
Estimated Fatalities	Education Residual	0.693	0.358	1.936	0.053	
Conflict Count	Health	-0.742	0.268	-2.770	0.006	**
Conflict Count	Income Residual	-0.883	0.524	-1.685	0.092	
Conflict Count	Education Residual	0.539	0.337	1.600	0.110	
State-Based	Health	-0.472	0.324	-1.460	0.144	
State-Based	Income Residual	-0.779	0.592	-1.315	0.188	
State-Based	Education Residual	0.189	0.386	0.490	0.624	
Non-State	Health	-2.368	0.374	-6.329	0.000	***
Non-State	Income Residual	-4.027	1.120	-3.594	0.000	***
Non-State	Education Residual	0.105	0.931	0.112	0.911	
One-Sided	Health	-0.745	0.179	-4.169	0.000	***
One-Sided	Income Residual	-0.376	0.444	-0.846	0.398	
One-Sided	Education Residual	0.622	0.281	2.214	0.027	*

Regression coefficients for income and education and health residuals. Higher income reduces most forms of violence, while positive education residuals increase one-sided violence.

Appendix 4: Data Structure

The dissfinal folder can be found on the M drive (M:/dissfinal). It is a compressed .zip file due to storage constraints. If unzipped, the size of the folder is approximately 1.2 GB.

-dissfinal Folder
- -Data Folder
  - -cleansocioeconomicdata.csv
  - -clean_land_cover.csv
  - -pop_dens_road_dens.csv
  - -Data Acquisition Folder
    - -GDL Shapefiles Folder
      - -GDL Shapefiles V6.4 large.cpg
      - -GDL Shapefiles V6.4 large.dbf
      - -GDL Shapefiles V6.4 large.prj
      - -GDL Shapefiles V6.4 large.shp
      - -GDL Shapefiles V6.4 large.shx
    - -MODIS Rasters Folder
      - -2011MCD12Q1.A2011001.h20v05.061.2022161182608.hdf
      - -2011MCD12Q1.A2011001.h21v05.061.2022161190538.hdf
      - -2012MCD12Q1.A2012001.h20v05.061.2022162070608.hdf
      - -2012MCD12Q1.A2012001.h21v05.061.2022162082835.hdf
      - -2013MCD12Q1.A2013001.h20v05.061.2022164221506.hdf
      - -2013MCD12Q1.A2013001.h21v05.061.2022164231444.hdf
      - -2014MCD12Q1.A2014001.h20v05.061.2022165112941.hdf
      - -2014MCD12Q1.A2014001.h21v05.061.2022165131212.hdf
      - -2015MCD12Q1.A2015001.h20v05.061.2022166005310.hdf
      - -2015MCD12Q1.A2015001.h21v05.061.2022166031836.hdf
      - -2016MCD12Q1.A2016001.h20v05.061.2022166215233.hdf
      - -2016MCD12Q1.A2016001.h21v05.061.2022166220401.hdf
      - -2017MCD12Q1.A2017001.h20v05.061.2022168072854.hdf
      - -2017MCD12Q1.A2017001.h21v05.061.2022168090051.hdf
      - -2018MCD12Q1.A2018001.h20v05.061.2022169151345.hdf
      - -2018MCD12Q1.A2018001.h20v05.061.2022169151345.hdf
      - -2019MCD12Q1.A2019001.h20v05.061.2022169180558.hdf
      - -2019MCD12Q1.A2019001.h21v05.061.2022169193259.hdf
    - -Socioeconomic Data Folder
      - -GDL-Custom-set-of-indicators-(2011)-data.csv
      - -GDL-Custom-set-of-indicators-(2012)-data.csv
      - -GDL-Custom-set-of-indicators-(2013)-data.csv
      - -GDL-Custom-set-of-indicators-(2014)-data.csv
      - -GDL-Custom-set-of-indicators-(2015)-data.csv
      - -GDL-Custom-set-of-indicators-(2016)-data.csv
      - -GDL-Custom-set-of-indicators-(2017)-data.csv
      - -GDL-Custom-set-of-indicators-(2018)-data.csv
      - -GDL-Custom-set-of-indicators-(2019)-data.csv
    - -Syria Shapefile Folder
      - -SyriaLevel2.cpg

-SyriaLevel2.dbf
-SyriaLevel2.prj
-SyriaLevel2.sbn
-SyriaLevel2.sbx
-SyriaLevel2.shp
-SyriaLevel2.shx
-Uppsala Conflict Data Folder
-gedevents-2025-02-27.csv
-cleanconflictdata.csv
-Land Cover Percentages Data Folder
-LandCover2011.csv
-LandCover2012.csv
-LandCover2013.csv
-LandCover2014.csv
-LandCover2015.csv
-LandCover2016.csv
-LandCover2017.csv
-LandCover2018.csv
-LandCover2019.csv
-LandCoverMerged_2011_2019.csv
-Outputs Folder
-LISA_Maps Folder
lisa_non_state_based_panel.png
lisa_one_sided_panel.png
lisa_state_based_panel.png
-global_moransI.csv
-nb_model_diagnostics.xlsx
-nb_model_info_criteria.xlsx
-nb_model_info_criteria.xlsx
-nb_models_random_vs_fixed.xlsx
-R Code Folder
-conflict_district_summary.R
-moransi_lisa_maps.R
-nb_diagnostics.R

Appendix 5: Download Data

To download Syria datasets, visit links below (Accessed 4 August 2025):

OpenStreetMaps: https://data.humdata.org/dataset/hotosm_syr_roads
WorldPop: https://hub.worldpop.org/geodata/listing?id=69
The Data Lab: https://globaldatalab.org/shdi/table/
Uppsala Conflict Data Program: https://globaldatalab.org/shdi/table/
Land Cover: https://www.earthdata.nasa.gov/data/catalog/lpcloud-mcd12q1-061
Syria GADM Shapefile: https://gadm.org/download_country.html