Modelling Fatalities Caused by Conflict Events in Bangladesh

Feargus Jamieson Ball [1] Daniel Moss [2] Raiha Browning [3] Judith Rousseau [2]

[1]University of Edinburgh [2]University of Oxford [3]University of Warwick

Research Aims

- Develop different methods to model fatalities as a response to conflict events in Bangladesh.
- Use and evaluate MCMC algorithms to sample model parameters.
- Evaluate the performance of the models against each other.

Data Collection and Manipulation

- The data have been collected by ACLED, the Armed Conflicts Events and Location Data Project, and cover the period from January 2010 to December 2021.
- Data include the spatial co-ordinate location and date of each event, as well as a record of the number of associated fatalities.
- Spatial and temporal aggregation has been conducted to provide information about the number of events and fatalities on each day in each district.
- I have then aggregated the data weekly to reduce noise and improve modelling.

Observed Data

Of the six conflict events, four of them are **violent**, so commonly result in fatalities:

- Battles Violence between two armed groups.
- Explosions/Remote violence One-sided violence conducted from afar.
- Riots Demonstrators engaging in violence against property or people.
- Violence against civilians Armed groups inflicting violence on unarmed civilians.

The other two **nonviolent** conflict event types are:

- Protests Nonviolent action by demonstrators, usually against a political entity.
- Strategic Developments Nonviolent action by entities otherwise involved in conflict.

The chart below illustrates how the different event types have different proportions of fatalities:

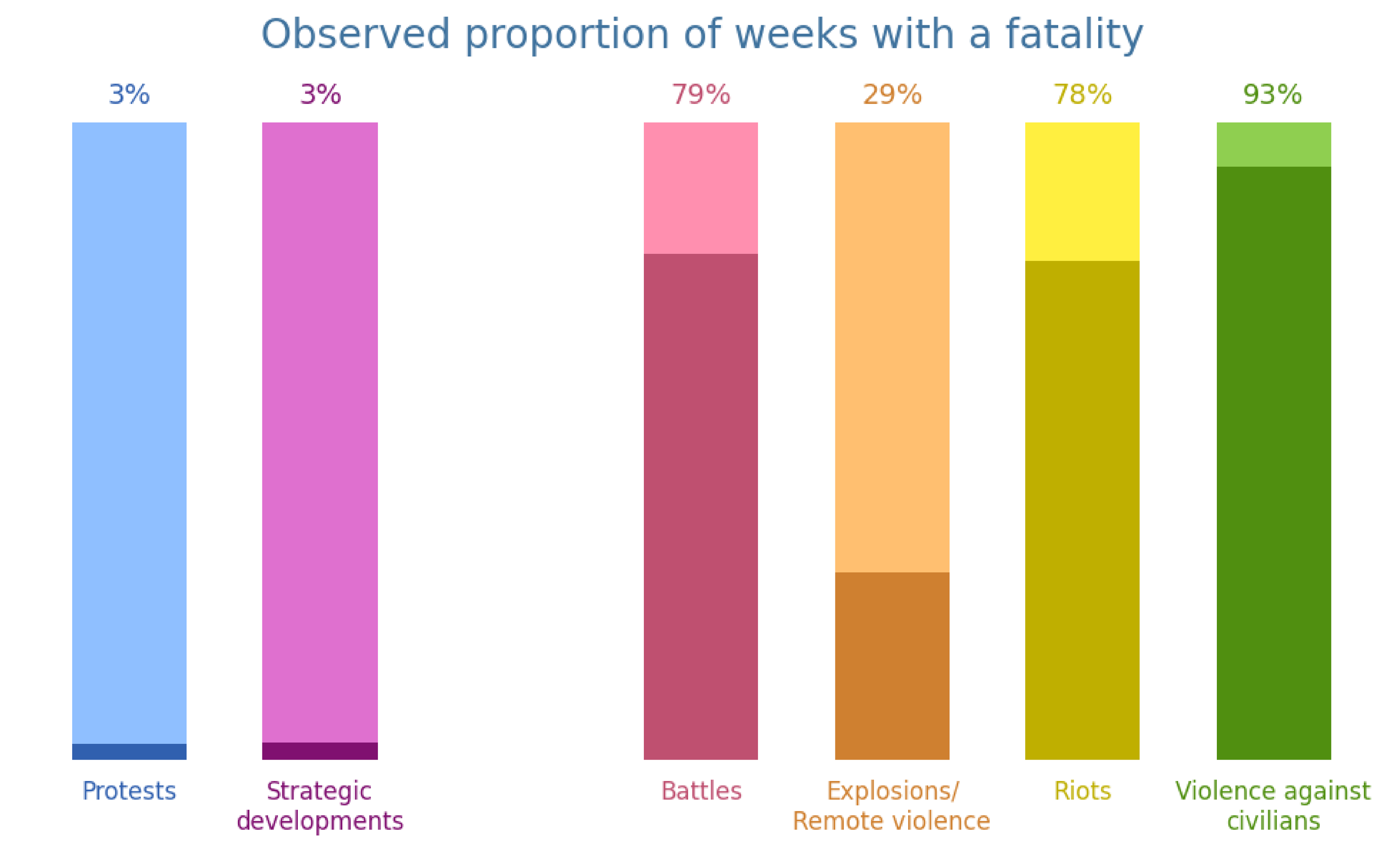

Figure 1. Proportion of weeks with at least one event and a fatality, by event type.

References

[1] Michael A. Newton and Adrian E. Raftery. Approximate bayesian inference with the weighted likelihood bootstrap. *Journal of the Royal Statistical Society Series B (Methodological)*, 56(1):3–48, 1994.

Graphical Results: Events in Dhaka

The figure below shows the posterior distributions of the probability of a fatality in Dhaka given one event of a given type in a week:

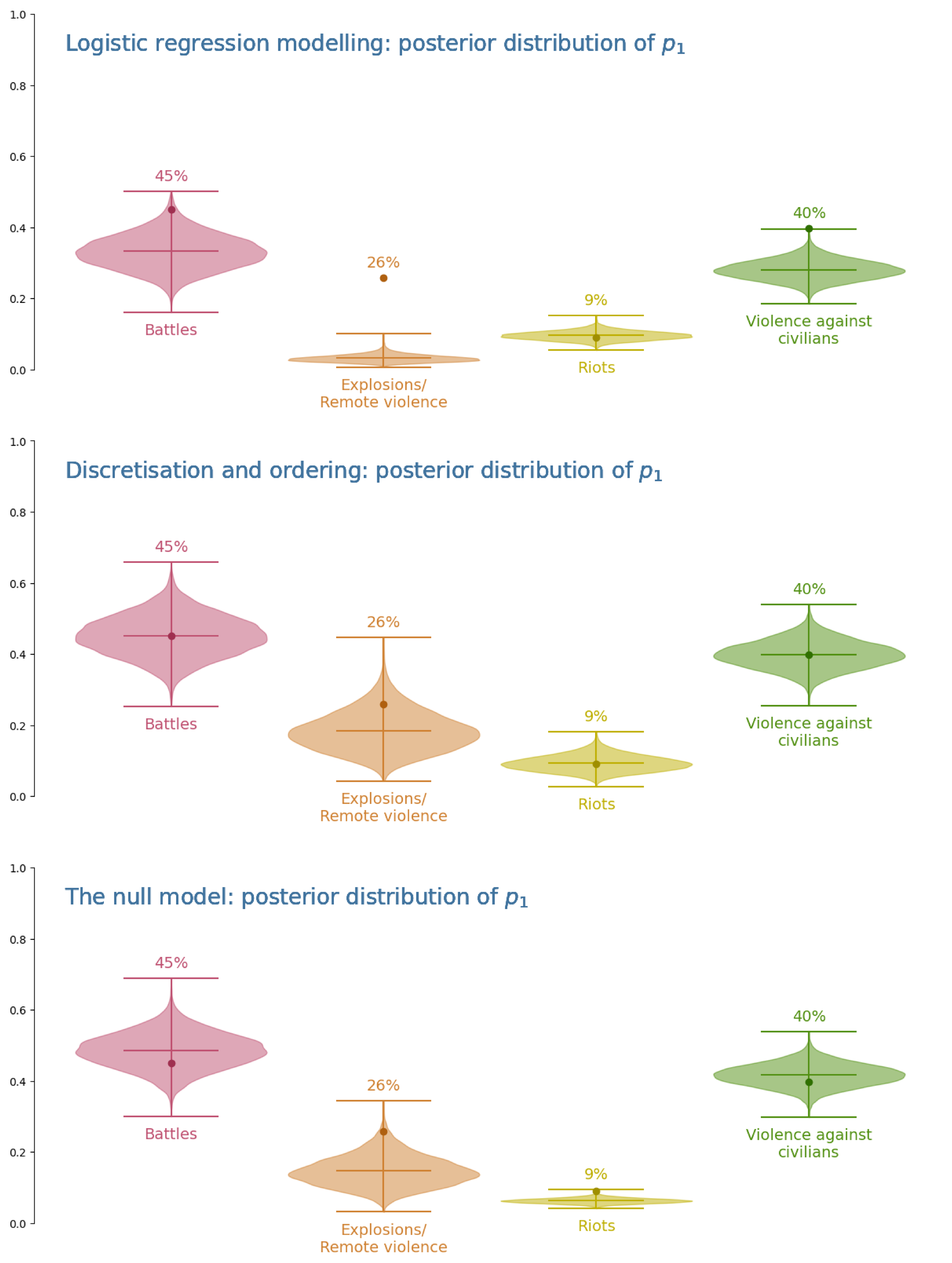

Figure 2. Posterior distributions of p_1 under the different models.

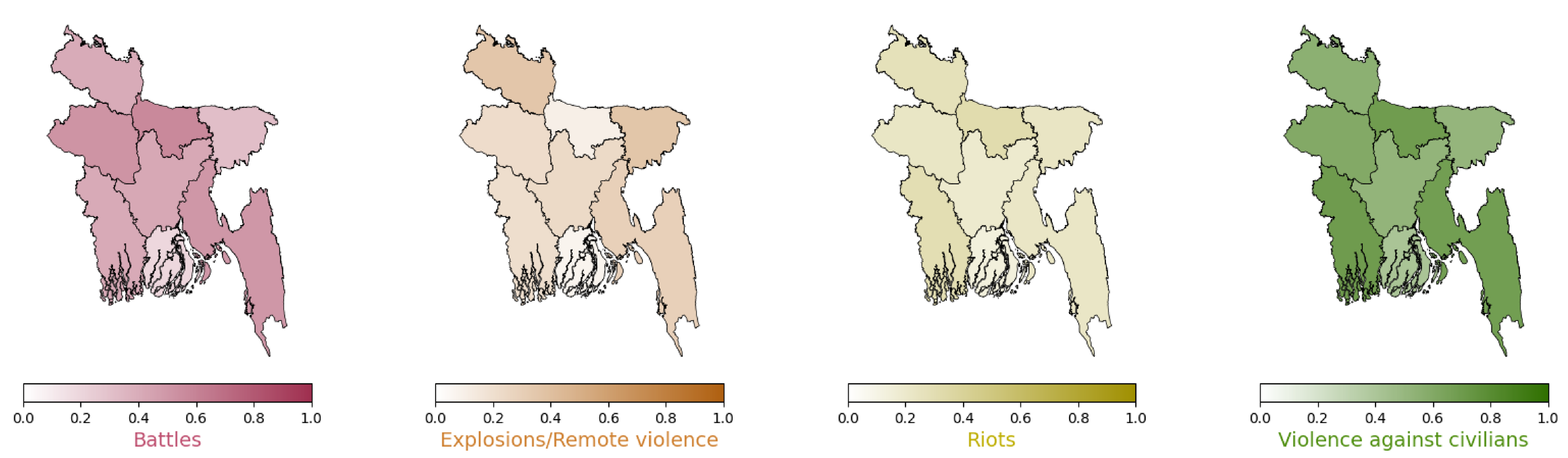

Figure 3. Posterior mean of p_1 for the Discretisation and Ordering model.

Logistic Regression Modelling

The first model models the proportion of weeks with a fatality as a logistic regression on the number of events, with $p_j = \exp(\alpha + \beta j)/(1 + \exp(\alpha + \beta j))$ for j events in a week.

The table below shows the effective sample sizes of 10000 parameter samples sampled both by the Hamiltonian algorithm in Stan, and my implementation of Metropolis-Hastings:

		Battles	Explosions/ Remote violence	Riots	Violence against civilians
Stan	α	9673	9772	10084	10110
	β	9729	9754	9796	9902
M-H	α	514	1900	501	355
	β	514	1750	199	384

Table 1. Effective Sample Sizes from the Stan and Metropolis-Hastings output.

It is clear that Stan is much better at sampling than the Metropolis Hastings algorithm. All the plots to the left were created with samples from Stan.
Issues with this model include it describing weeks with noninteger conflict events and its nonzero posterior probability of a fatality in a week with no events.

Discretisation and Ordering

This model categorises the observations by event count into categories $k = 1, \dots, K$. Parameters $p_K, q_{K-1}, q_{K-2}, \dots, q_1 \in (0, 1)$ are sampled and p_j is defined as

$$p_j = \mathbb{P}(F_i = 1 | S_i \in \text{category } j) = p_K \prod_{k=j}^{K-1} q_k. \tag{1}$$

This forces the ordering $p_1 < p_2 < \cdots p_K$ on each sample and the posterior distributions overall.

Some progress was made sampling with Metropolis-Hastings, which involved transforming proposed parameters between the proposal space $\mathbb{R}$ and the parameter space $(0, 1)$.

The discretisation built into this model resolves the issues of the logistic regression model.

The issue here is that it is difficult to decide how to categorise the data. Further analysis could investigate and implement methods to categorise the data systematically.

The Null Model

The third model considers an intuitive relationship between the number of events and the probability of a fatality, by scaling the probability of a fatality given one event:

$$p_j = \mathbb{P}(F_i = 1 | S_i = j \text{ events}) = 1 - (1 - p_1)^j \quad j = 2, 3, \dots. \tag{2}$$

Since there is only one parameter to sample, the Metropolis-Hastings algorithm performed best in this model. The samples on the left were found using Stan, as this was again more effective.

Model Evaluation

The table below shows the log-marginal likelihoods for each model as estimated using an iterative method provided by Newton and Raftery [1]:

	Battles	Explosions/ Remote violence	Riots	Violence against civilians
Logistic Regression	-67.58	-35.32	-197.28	-151.57
Discretisation	-69.47	-45.37	-232.18	-161.55
Null Model	-60.06	-26.26	-189.88	-139.99

Table 2. Log Marginal Likelihood for each model and event.

The **Null Model** is preferred for all event types. The low likelihoods of the discretisation model for some event types show that a helpful direction of future analysis would be to develop a systematic binning procedure to improve the capabilities of this model.

s2196099@ed.ac.uk