# A Zero-Inflated Poisson Generalized Additive Model for Forecasting Conflict Fatalities

**David Muchlinski**[1] **and Chandler Thornhill**[2]

[1]**Georgia Tech**
[2]**Georgia Tech**

## ABSTRACT

Conflict data often face a seeming limitation: outbreaks of major episodes of political violence are, thankfully, rare. This means that the modal value of the dependent variables is often zero, but the tails of conflict distributions can also be strongly right skewed. This poses forecasting problems for both practitioners, who want to know whether any conflict might occur, and for methodologists, who are concerned with accurately predicting how much conflict there may be, as predictions may be strongly biased downwards by this zero-inflation. We propose a semi-parametric zero-inflated constrained generalized additive model for forecasting conflict. This model proceeds by first utilizing a binomial distribution to predict whether any conflict will be observed, then, given a positive prediction of any conflict, utilizes a Poisson distribution to predict how much conflict there may be. Our model achieves predictive accuracy as measured by the continuous ranked probability score comparable or better than the VIEWS benchmarks for nearly all years in the test set. We generate predictions for the VIEWS forecasting window of June 2024-July 2025 by simulating from the model's predictive distribution.

Keywords: VIEWS, Prediction, Zero-Inflated Poisson

Federal Foreign Office PREVIEW Prediction Visualisation Early Warning VIEWS PREDICTING CONFLICT & HUMANITARIAN IMPACTS

## INTRODUCTION

Major[1] episodes of political violence are, thankfully, rare. Most countries in the world do not experience any battle-related fatalities from international or domestic conflict most of the time, though trends in civil and international conflict have been increasing since the 2010 (Von Einsiedel et al., 2017). Importantly, most conflict deaths from organized political violence fall outside of countries involved in ongoing civil or international conflicts (Krause, 2016), leaving hope that such conflicts may be mitigated before they erupt into open war. Once a state finds itself in a war, however, the probability of peaceful resolution declines with the length of time a country is at war (Mueller et al., 2022).

The rare, but destructive, nature of political violence thus poses problems for the scholarly forecasting community and practitioners. Both communities would like to know two related problems: first, whether there will be any (new) conflict or not, and second, how intense such conflict may be. A forecasting model that misses the outbreak of a major war between Iran and Israel, for instance, is poorly calibrated from an academic standpoint, and provides little in the way of policy-relevant advice to practitioners from a use-case perspective. Thus, both communities have a need for quantitative forecasting models that are accurate in space, i.e. correctly forecasting *where* conflict will be, and accurate in *intensity* and are generally correct in estimating the number of possible fatalities.

[1]This paper documents a contribution to the VIEWS Prediction Challenge 2023/2024. Financial support for the Prediction Challenge was provided by the German Ministry for Foreign Affairs. For more information on the Prediction Challenge please see (Hegre, ming) and `https://viewsforecasting.org/research/prediction-challenge-2023`

We propose a conflict forecasting model based on a zero-inflated Poisson generalized additive model, or GAM. GAMs are ideal forecasting models as they are able to incorporate non-linearities in the independent variables with ease (Goldsmith and Butcher, 2018). The model works in two stages. It first estimates a hurdle parameter to estimate a binary prediction of whether or not conflict will be present in a given observation, then, assuming a positive prediction of conflict, it utilizes a Poisson distribution to estimate the count of the number of fatalities. We explain our model in further detail below. We then provide some accuracy metrics in the test set against the VIEWS benchmark forecasts. We find our zero-inflated Poisson (views_zip) GAM does better than the benchmark models in nearly all years. We finally discuss our predictions for the true future prediction window of 2024-2025.

## A ZERO-INFLATED POISSON MODEL

As stated previously, we utilize a generalized additive model with a zero-inflated Poisson distribution to forecast conflict fatalities. We utilize a relatively simple set of covariates to model the Poisson distribution. Given the result from the first VIEWS prediction competition that conflict is generally autoregressive (D'Orazio and Lin, 2022), we calculate two, four, and six month rolling averages of conflict fatalities using the *ged_sb* dependent variable as well as utilizing, one, two, three, four, five, and six month temporal lags of fatalities. To ensure that predictions stay in a reasonable bound, we also utilize a state's total population, as well as fatalities as a percentage of population. To estimate the zero-inflation probabilities, we utilize decay functions of state-based (*decay_ged_sb_5*, *decay_ged_sb_500*, and *decay_ged_sb_100*) non-state (*decay_acled_ns_5*), and one-sided violence (*decay_acled_os_5*) taken from the VIEWS feature set. The smoothness of the splines in the GAM was held constant for all features at a value of $k = 5$. All `R` code for replicating model results and true future predictions accompany this submission.

The model is trained initially on historical conflict data from 1980-2017, then predictions are made for data beginning in 2018. For conflict predictions for 2019, the 2018 historical data is added to the training set and the model is estimated again with predictions made for 2019. This sliding training/test window continues for every year up to 2023.

## RESULTS IN THE TEST SET

Our model achieves forecast accuracy comparable to, or superior to, the best performing VIEWS benchmark model for each year in the testing window. Table 1 below provides the mean CRPS score for our model as well as the best performing VIEWS benchmark model for that year. As is clearly seen in the table, our zero-inflated GAM outperforms the VIEWS benchmark models for every year except 2019 and 2022. However, the CRPS of our ZIP model for 2019 is very close to the VIEWS benchmark for that year. Thus we can confidently say that our model's performance is better than or comparable to the VIEWS benchmark models for all years except 2022, which, for unique reasons (most likely Russia's invasion of Ukraine or revisions to Ethiopia's fatality counts in the UCDP data) exhibits more unexpected conflict compared to all years.

**Table 1.** CRPS Scores in Test Set: VIEWS Benchmarks and Zero-Inflated GAM

| CRPS | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 |
|---|---|---|---|---|---|---|
| Best VIEWS Benchmark | 14.48288 | **9.146306** | 21.33933 | 76.84948 | **120.2492** | 50.35671 |
| ZIP GAM | **9.0377** | 9.7993 | **14.631** | **72.05** | 161.45 | **28.754** |

## TRUE FUTURE PREDICTIONS

We simulate 1000 draws from the predictive distribution of our model using a zero-inflated Poisson distribution. Our predictions for the true future are submitted along with this write-up. Our predictions for the average number of monthly fatalities for the true future window of July 2024-June 2025 for the top ten countries appear in Table 2 below.

**Table 2.** Top 10 Countries with Highest Average Monthly Fatality Counts

| Country | Avg. Monthly Fatality Count |
|---|---|
| Ethiopia | 2535 |
| Ukraine | 2096 |
| Afghanistan | 1468 |
| Israel | 467 |
| Syria | 324 |
| Azerbaijan | 132 |
| Somalia | 109 |
| Nigeria | 92 |
| Burkina Faso | 74 |

## REFERENCES


D'Orazio, V. and Lin, Y. (2022). Forecasting conflict in africa with automated machine learning systems. *International Interactions*, 48(4):714–738.

Goldsmith, B. E. and Butcher, C. (2018). Genocide forecasting: Past accuracy and new forecasts to 2020. *Journal of Genocide Research*, 20(1):90–107.

Hegre, H. (Forthcoming). The 2023/24 views prediction competition. Journal of Peace Research.

Krause, K. (2016). From armed conflict to political violence: mapping & explaining conflict trends. *Daedalus*, 145(4):113–126.

Mueller, H., Rauh, C., and Ruggieri, A. (2022). Dynamic early warning and action model.

Von Einsiedel, S., Bosetti, L., Cockayne, J., Salih, C., and Wan, W. (2017). Civil war trends and the changing nature of armed conflict. *Occasional paper*, 10:1–10.