Predicting Rebel Activity in Sub-Saharan Africa:

An Evaluation of Different Time-Series Methods.*

Marco Morucci

m.morucci@ucl.ac.uk

*The analysis presented here was conducted using the R language and various external libraries. Code and data for full reproducibility are available at https://github.com/marcomorucci/predicting-rebels.

Introduction

Recent advances in machine-learning and computational statistics have prompted an increasing number of political scientists [Blair et al., 2014, Yonamine, 2013, Hirose et al., 2014], to build statistical models for prediction, rather than explanation. These attempts are definitely laudable for at least two reasons: first, they make easily testable and falsifiable prediction basing on theory, thus improving the overall methodological quality of research in political science. Second, they offer a practical instrument for policy-makers to adopt the theories developed by political scientists and put them to practice.

One particular subfield in which predictive models have been employed somewhat more systematically is conflict prediction. Recently, increasingly successful attempts at predicting the beginning and onset of conflicts at the international [Hegre et al., 2009], [Ulfelder, 2013], national [Gleditsch and Ward, 2013] and local [Blair et al., 2014] level have been made using various statistical techniques with varying degrees of complexity.

The past and existing studies in the field focus on prediction of conflict at a specific geographic level, either international, national or local. Here I present an actor-centered approach: I look at a specific category of players at the state-level, rebel groups, I try to predict increases and decreases in their overall activity and in how this activity will be distributed between conflict and cooperation.

I employ auto-regressive properties of series of interactions between rebel groups and other state actors. I train three different models, one based on univariate properties of overall rebel activity, another based on both univariate properties and ensemble patterns of different kinds of interaction and a third model that mixes the two approaches.

I build the models based on two specific approaches, the ARIMA model for univariate time-series employed to predict conflict at a local level by Shellman [2006] and Vector Auto Regression (VAR) for multi-variate time series of rebel behaviour as employed by Pevehouse and Goldstein [1999].

Using data from the GDELT database, I forecast aggregate rebel group behaviour in 26 African countries, chosen on the basis of greater data availability and greater activity of rebel and separatist groups in these countries compared to others.

The three models employed reflect different understandings of what make good predictors of violent and non-violent activity by rebel groups: the univariate approach tries to predict future activity basing only on past patters of the specific kind of behaviour analysed, whereas multivariate models allow for interaction between different kind of behaviours. By testing these models separately, I hope to highlight which of these patterns seems to work best for prediction.

My focus here is not explaining why a notion should work better than another, instead it is detailing in which ways each model succeeds and fails, ultimately presenting an overall evaluations of different predictors of future rebel behaviour. In the next sections I will give a brief overview of the existing models of conflict prediction, and then move on to illustrate the methodology behind the models, reporting their performance and concluding with some general remarks on actor-based behaviour prediction.

Approaches to conflict Prediction

Using past event sequences both violent and non-violent to predict future conflict is a rather old idea in itself. Simple Markov-Chain sequence models had already been in use throughout the 70s and 80s. Hidden Markov Models came in to supplant these in the mid 90s [Schrodt and Gerner, 1997], and have been widely used to predict intra-national conflict ever since [Schrodt, 2000, Schrodt and Gerner, 2004, Shellman, 2006, Shearer, 2007, Petroff et al., 2013]. Other approaches at sequence-based prediction include zero-inflated count model [Bagozzi, 2011], sequence similarity analysis [D'Orazio et al., 2011] and ARFIMA models [Yonamine, 2013].

Recently, some more predictive models basing on exogenous variables, rather than event sequences have surfaced: starting with Weidmann and Ward [2010], who employ several demographic and geographic variables to predict conflict in Bosnian municipalities through MCMC logistic regression. Blair et al. [2014] employ survey data from several Liberian communities to predict onset of conflict in the following year, in each of those communities; while Hirose et al. rely on civilian attitudes as a predictor for future civil war. Lastly, Ward et al. [2013] employ hierarchical models to predict civil conflict at the national and cross-country level; combining several exogenous variables such as levels of democracy and infant mortality with data on interaction between governments and their opponents. Ulfelder [2013] and Montgomery et al. [2012] propose combining several predictive models so to maximise prediction accuracy. My approach takes cues from theirs and employs a combination of univariate and multivariate time series to produce a better forecast.

Most of these studies' aim is to predict conflict onsets at a specific geographical level, be it international, national, regional or local. Here instead I focus on the actor-level; trying to forecast the aggregate behaviour of rebel organizations at the state-level.

Shellman [2006] employs ARIMA models to predict interactions between governments and dissidents, using event data from Latin America. I try to employ differences between kinds of actions previously performed by the rebels as well, divided into four broad categories. To do so, I employ Vector Auto Regression, an approach already introduced by Pevehouse and Goldstein [1999] to predict onsets of conflict between Serbia and Kosovo. The authors employ counts of interactions between several actors aggregated into a measure of "net cooperation" (e.g. cooperation events - aggression events) and use daily values of this measure to forecast future behaviour from the Serbian regime.

Here I employ a similar approach, however I don't aggregate cooperative and non-cooperative behaviour; instead, I use VAR on the singular event counts for each typology to estimate them by taking into account interaction effects that a type of behaviour might have on another.

It should be noted that more advanced time-series methodologies have been applied for forecasting conflict events, both for the univariate [Yonamine, 2013] and multivariate [Brandt and Freeman, 2006, Brandt et al., 2011] models. My purpose here is to highlight how different approaches contribute to predictive performance, rather than building the best-performing model right away. Using simpler ARIMA and VAR models, I hope to retain some explanatory ability for the features I employ, while still outputting relevant predictions.

Data

The data employed here is from the Global Data on Events, Location, and Tone (GDELT) project [Leetaru and Schrodt, 2013]. The subset of the data I obtain consists of around 400.000 dyadic interaction events between various rebel and separatist groups and other state-level actors, such as government, military or civilians[1].

The data is aggregated on a monthly basis, following the example of several authors [Yonamine, 2013, Bagozzi, 2011, Weidmann and Ward, 2010, DOrazio and Yonamine]; and event counts for each category are considered, plus an overall sum of the events, to account for overall rebel activity.

The data is coded according to the CAMEO ontology [Schrodt et al., 2008], which codes events into increasingly specific categories; here I adopt the broadest level of event code aggregation which employs four categories:

- Verbal Cooperation (VC), which encompasses events such as statements, comments and appeals.
- Material Cooperation (MC), for economic and military aid.
- Verbal Violence (VV) for different kinds of threats and ultimatums.
- Material Violence (MV) which accounts for all the acts related to the use of force.

Event counts for each category are aggregated monthly and forecast for the coming month; as such they constitute the principal object of the prediction. Note that this is somewhat

[1] The data was obtained through the Google BigQuery API for GDELT

unusual for conflict prediction efforts, as they usually focus on predicting a 0/1 chance of conflict onset in the next time period [Hegre et al., 2009, Weidmann and Ward, 2010, Gleditsch and Ward, 2013, Schrodt, 2006, Shearer, 2007, Shellman, 2006]. This makes the task at hand somewhat more difficult as the prediction attempted has a higher degree of specificity. D'Orazio et al. [2011] and Yonamine [2013] attempt the same kind of real-valued conflict level prediction, albeit using somewhat different methods and in different geographical locations. The time period considered ranges from Jan 1990 to Nov 2014, thus accounting for an overall 299 months of activity. The last 23 months, corresponding to years 2013-2014 are separated from the training data and employed as an out-of-sample evaluation dataset.

The 23 countries analyzed are all from the sub-Saharan Africa region. Figure shows them plotted alongside with the main locations of violent rebel activity within each of them. This is possible thanks to GDELT's addition to geolocation data to its event reports. The countries where selected for mainly two reasons: 1 multiple countries in the same region allow for the analysis to be based on a broader sample, while still controlling for regional differences, and two, these countries were the ones that reported the most significant amounts of rebel activity in the decades onto which the analysis is based. This allows, in order, to minimize the number of missing data points.

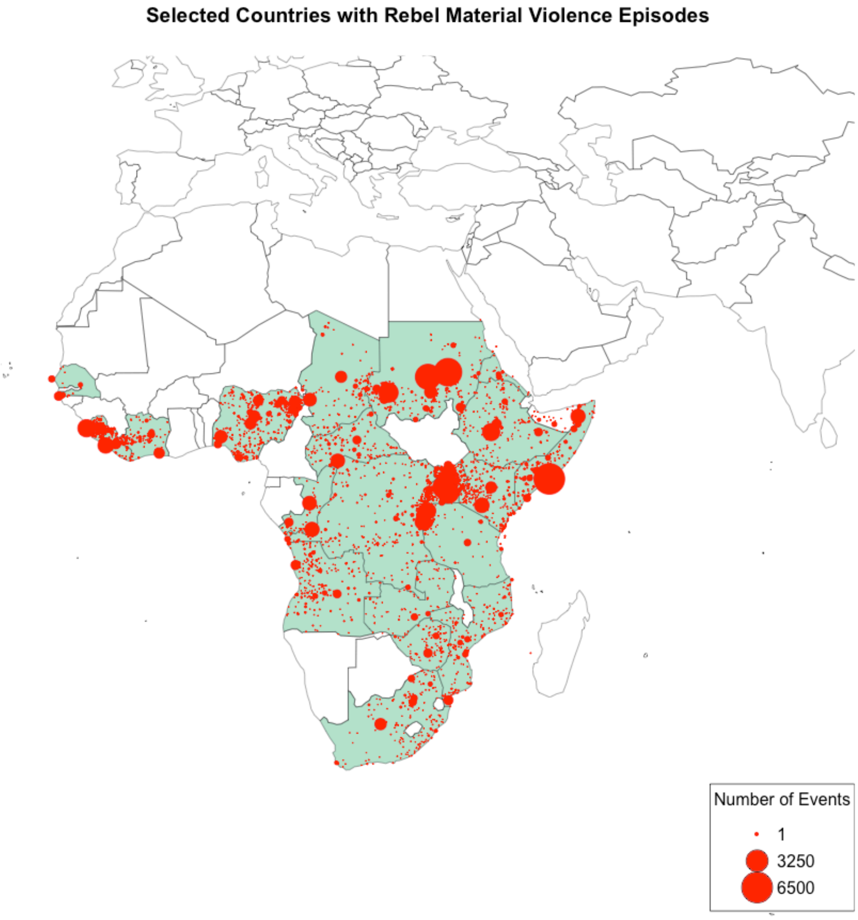

Figure 1:

Theoretical Background for The Models

Each one of the models presented here is based on a somewhat different understanding of how different patterns of behaviour of rebel groups influence future ones.

Model 1 posits that the specific kinds of actions taken by rebels one month will influence which actions are taken next month, I.E. a greater level of verbal violence the month before could lead to greater levels of material violence the month after. This type of escalation has already been used to predict conflict periods especially with HMMs. Schrodt [2006] designs a probabilistic model able to transition between peaceful, low-conflict and high-conflict phases, thus managing to account for the escalation sequence that a model of this kind might rely on.

Model 1 also includes actions from other actors (mostly governmental) targeting the rebels, this doesn't seem to strengthen the predictive accuracy of the model, as shown in figure . Shellman [2006] shows that government-rebels interactions do influence each other, using sequence analysis. The reason why government behaviour toward the rebels might not help predictive accuracy in this case, might lie in the monthly aggregation of data. As Shellman shows, aggregating event counts monthly might cause a loss in the predictive accuracy of event sequences. Nonetheless, D' Orazio and Yonamine [] employ the same methodology and find that event sequences leading to conflict are rather different between each other and as such it is hard to find generalizable patterns within them.

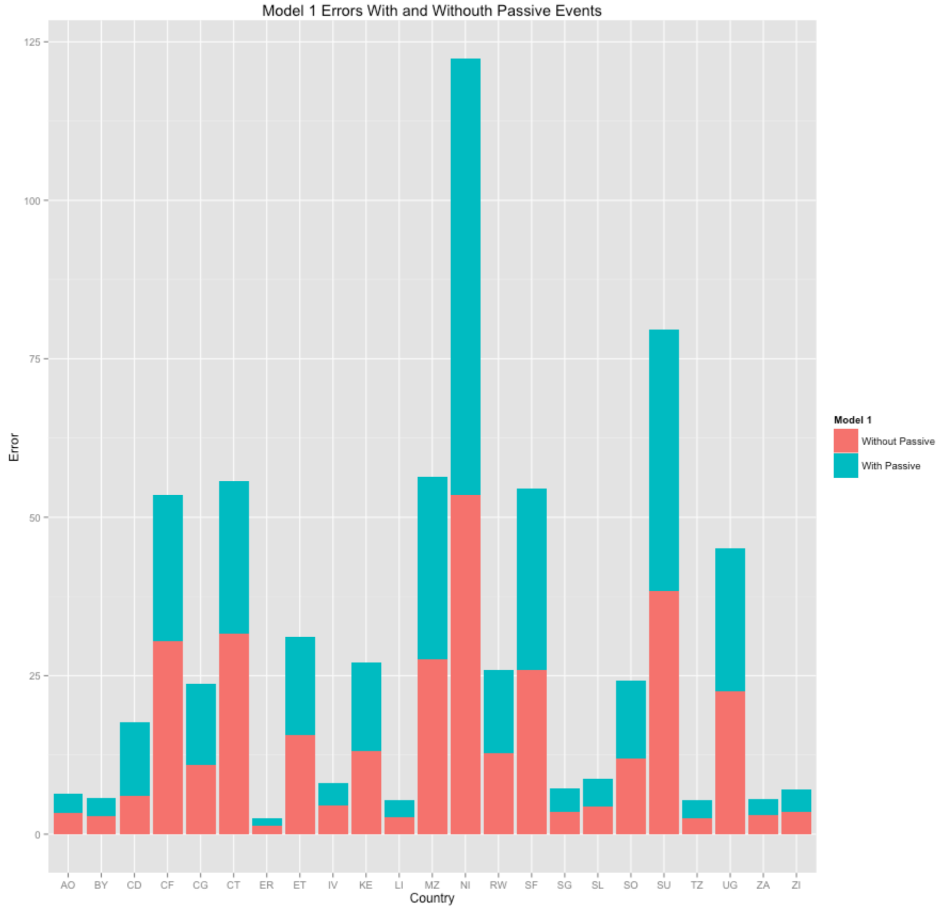

Figure 2:

Model 2 relies on individual patterns for each different type of action, using only past levels of that behaviour itself. This same approach has been employed several times to study conflict outsets [Yonamine, 2013, Shearer, 2007, Weidmann and Ward, 2010] and tends to assume conflict to have visible autoregressive properties.

Model 3 aims at combining these two approaches and is based on the simple observation that individual event counts for each typology are related to the overall count of actions recorded for the rebels in the country of interest at the time of interest. This implies that vector autoregression models such as model 1 will capture much of the overall variance as explanatory, failing to recognize the overall trend in all the different action types and thus seeing them as more correlated than they actually are. To account for this, model 3 trains a VAR on the ratios of each kind of action relative to their total activity, and an ARIMA on the total itself. After doing this, predicted ratios are multiplied with the predicted total to obtain final estimates.

Model Specification

Each model is trained separately on each different country's corresponding time-series, as such 26 different models are built and tested; each one with different parameters and coefficients. For this purpose, the choice of input parameters (I. E. p, d, q) has to be automated. From now on, I will refer to T as the total number of months (299) in the dataset, E as the total number of event typologies considered (4) and N as the total number of countries in the sample (26).

Model 1

The first model consists of a Vector Auto Regression (VAR) estimator for the four categories of actions considered. The predicted counts for each category are then summed up to predict the total amount of actions for each group, thus providing an estimate for an overall increase/decrease in activity.
The final form of the estimator is:

$$y_t = c + \phi_1 y_{t-1} \dots \phi_p y_{t-p} + \epsilon_t$$

Where y is a matrix $\in \mathbb{Z}^{ExT}$ of event counts for each category, for each month from Jan 1990 to Nov 2014 e.g:

$$\begin{pmatrix} VC_{2014-11}, VC_{2014-10}, \ldots, VC_{1990-01} \\ MC_{2014-11}, MC_{2014-10} \ldots, MC_{1990-01} \\ VV_{2014-11}, VV_{2014-10}, \ldots, VV_{1990-01} \\ MV_{2014-11}, MV_{2014-10}, \ldots, MV_{1990-01} \end{pmatrix} \tag{1}$$

y_t is then a $E \times 1$ vector of event counts at time t (E.G. a column of the matrix) ϕ_t is a $E \times E$ matrix of real-valued coefficients for time t and ϵ_t is a $E \times 1$ vector of error terms at time $t.\epsilon_t$ is assumed to be white noise.

The model allows each parameter to influence the estimate for the other, thus taking into account the influences that having acted in a certain way in the past might have on future actions. Effectively this allows all the variables within the model to be treated as endogenous, regressing on both lags of the chosen dependent variables and lags of all the included predictors [Freeman et al., 1989].

Here I include monthly counts for each of the four different categories of actions that rebel groups can perform each month. I also train an additional model in which actions from other actors targeting rebel groups are also included as regressors, but their influence on the prediction accuracy for this model is vastly negative, thus suggesting that the most appropriate version of this model is the more parsimonious one.

The optimal number of lags to include in the model in order to maximise predictive accuracy is estimated automatically using the Bayesian Information Criterion [Pfaff, 2008].

The VAR model allows for the inclusion of seasonal or exogenous variables. Here I don't include exogenous factors, in order to obtain a better evaluation of the performance of different action typologies as predictors, as such the c coefficient matrix represents constants for each event typology model.

Model 2

The second model I train considers action counts for each typology independently and uses ARIMA to predict increases and decreases in these counts.

Its inclusion here derives from the wish to highlight the predictive power of each single event pattern, while disregarding interaction between the different event typologies.

An ARIMA(p, d, q) process is defined by a number p of lags on the dependent variables, a number q of lags on the error term. The final prediction for an event count of type e at time t will take the following form:

$$\Delta y_t = c + \sum_{i=1}^{p} \phi_i \Delta y_{t-1} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t \tag{2}$$

Each time-series is differentiated in order to guarantee stationarity and Δy_t is the d-order difference for y_t. c represents a constant term and ϵ_t represents the error at time t, which is assumed to be a white noise process.

In this case, the final sum of actions is computed by summing up predicted counts for each single type.

The choice of p, d, q is automated for each country using the Hyndman-Khandakar algorithm [Hyndman and Khandakar, 2008]; which selects the d values basing on a series of KPSS tests and the p and q values which minimize the Akaike Information Criterion (AIC). This automatic choice allows an optimal model to be trained for each country considered.

Model 3

This model aims at mixing both the patterns in the total amount of activity of the rebel groups analysed and the individual counts of each typology of action performed. In order to do this at best, single event counts at time t are considered as proportions of the total activity at time t, I.E. they are divided by the total count of events at that time. Doing this allows us to get rid of the correlation between the event counts at each time period due to raise in the overall activity of the groups; and to focus, instead, only on event-specific correlation between typologies of behaviours.

This model is composed of two parts, first, the proportion of the total activity of a group in the coming month is estimated through a VAR of the following form:

$$\frac{y_t}{\sum_{e=1}^{E} y_{e,t}} = c + \phi_1 \frac{y_{t-1}}{\sum_{e=1}^{E} y_{e,t-1}} + \ldots + \phi_p \frac{y_{t-p}}{\sum_{e=1}^{E} y_{e,t-p}} + \epsilon_t \tag{3}$$

Where $\sum_{e=1}^{E} y_{e,t}$ is a scalar value consisting of the sum of events of each kind at time t and is separately estimated through and ARIMA(p,d,q) process in the second step of the model:

$$\Delta \sum_{e=1}^{E} y_{e,t} = c + \sum_{i=1}^{p} \phi_i \Delta \sum_{e=1}^{E} y_{e,t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t \tag{4}$$

Finally,event ratios estimated with the VAR are then multiplied with the estimated sum to obtain a real event-count for each category. Parameter choices detailed for the VAR and ARIMA above apply also to this model.

Model Evaluation Metrics and Procedure

I evaluate the performance of all three models three different mean error metrics, respectively collapsing model errors by month, country and event type. The first month-by-month error metric was already employed by Yonamine [2013]. Here I modify it slightly to average the errors by event as well:

$$error-by-month(m) = \frac{1}{NE} \sum_{c=1}^{N} \sum_{e=1}^{E} |prediction(m,t,e,c) - count(t,e,c)| \tag{5}$$

The error metrics by country and by event follow the same logic, collapsing two different variables each time:

$$error-by-country(m) = \frac{1}{TE} \sum_{t=1}^{T} \sum_{e=1}^{E} |prediction(m,t,e,c) - count(t,e,c)| \tag{6}$$

$$error-by-event(m) = \frac{1}{TN}\sum_{t=1}^{T}\sum_{c=1}^{N}|prediction(m,t,e,c) - count(t,e,c)| \tag{7}$$

I present all the different metrics to provide the best possible overview of the performance of the three models tested.

The models are tested within the out-of-sample framework that is common for machine learning applications: out of the 299 months in the total 1990-2014 data, the last 23, corresponding to years 2013-2014, are left out to serve as test-set onto which the model trained on the previous months is assessed. This framework has already been employed a number of times in conflict prediction studies [Weidmann and Ward, 2010, Yonamine, 2013, D'Orazio et al., 2011]; however Shellman [2004, 2006] doesn't use this framework in his previous assessments of government-rebel interactions.

Procedurally, I adopt the same approach of Weidmann and Ward [2010]:

For each country

1. train each model on data from 1990 to 2012
2. forecast 1 month ahead
3. incorporate that month into the training data
4. repeat for all the 23 months until Nov. 2014.

Results

Here I present aggregate error metrics for the three models tested: table 1 shows mean errors by month, and model 3 outperforms the other two in 15 out of 23 cases, showing that there is a clear advantage in aggregating model predictions on a monthly basis. The same can be said for country-level errors as well, as shown in table 2. Here model 3 has the lowest error in 16 cases out of 26, showing constant performance both spatially and temporally. Lastly, the errors-by-event summary shows that, on average, model 3 is better at predicting all four typologies of events, excluding, somewhat paradoxically, the total activity count itself.

Another interesting trend emerging from the different error counts is the bad performance of the VAR model (M1), which ends up performing the worst on both a by-country and by-month basis. Why this might be is somewhat uncertain, however, Shellman [2004] suggests that aggregating events monthly (and even more so yearly) detracts from their ability to produce the patterns needed for successful forecast. In this case, the results shown here could reinforce this theory.

Month	M1	M2	M3	Best	Worst
2013-01	40.35	18.81	16.50	M3	M1
2013-02	13.35	14.72	15.00	M1	M3
2013-03	22.65	20.18	20.01	M3	M1
2013-04	18.48	21.64	19.53	M1	M2
2013-05	11.09	14.48	13.02	M1	M2
2013-06	11.62	14.47	13.85	M1	M2
2013-07	28.29	19.68	19.54	M3	M1
2013-08	18.64	13.66	12.84	M3	M1
2013-09	21.88	16.24	15.34	M3	M1
2013-10	27.96	19.08	18.76	M3	M1
2013-11	34.78	22.64	22.68	M2	M1
2013-12	31.68	21.05	21.12	M2	M1
2014-01	26.93	20.65	20.18	M3	M1
2014-02	15.47	15.32	13.08	M3	M1
2014-03	15.23	17.10	15.07	M3	M2
2014-04	89.44	19.62	16.02	M3	M1
2014-05	26.66	20.91	18.81	M3	M1
2014-06	36.26	18.14	17.23	M3	M1
2014-07	82.53	37.04	37.48	M2	M1
2014-08	59.60	24.02	22.47	M3	M1
2014-09	33.83	27.11	26.01	M3	M1
2014-10	20.47	17.64	17.99	M2	M1
2014-11	38.59	26.01	24.72	M3	M1
				M1: 4	M1: 18
			Totals:	M2: 4	M2: 4
				M3: 15	M3: 1

Table 1: Model errors by month

Country	M1	M2	M3	Best	Worst
Angola	3.35	2.07	1.87	M3	M1
Burundi	2.83	1.57	1.50	M3	M1
Chad	6.06	5.32	5.20	M3	M1
Congo	30.48	27.47	26.15	M3	M1
DRC	10.88	8.91	8.18	M3	M1
CAR	31.55	43.88	39.90	M1	M2
Eritrea	1.30	1.25	1.11	M3	M1
Ethiopia	15.63	15.67	14.03	M3	M2
Cote d'Ivoire	4.50	3.16	2.45	M3	M1
Kenya	13.05	12.35	13.08	M2	M3
Liberia	2.68	2.83	2.54	M3	M2
Mozambique	27.67	29.23	30.84	M1	M3
Nigeria	53.51	59.26	64.50	M1	M3
Rwanda	12.83	17.02	15.98	M1	M2
South Africa	25.84	24.34	21.55	M3	M1
Senegal	3.54	3.80	3.55	M1	M2
Sierra Leone	4.37	2.97	2.73	M3	M1
Somalia	11.88	13.72	12.70	M1	M2
Sudan	38.39	40.43	40.07	M1	M2
Tanzania	2.56	2.81	2.42	M3	M2
Uganda	22.59	23.68	22.58	M3	M2
Zambia	3.06	2.11	2.18	M2	M1
Zimbabwe	3.50	3.58	3.48	M3	M2
				M1: 7	M1: 10
			Totals:	M2: 2	M2: 11
				M3: 14	M3: 3

Table 2: Model errors by country

Turning to results for specific instances of violent conflict (MV) prediction, the error patterns are somewhat different and surprising. First of all we see that the mixed model still performs fairly well and following to the overall trend; second we can notice a surprising difference in mean error in month-by-month and country-by-country forecasts between model 2 and model 1. In the first case, M1 performs the worst overall, while in the country case, M2 performs the worst. This might suggest that VAR could work best when few cases are aggregated across a broader temporal spectrum, while univariate ARIMA works best when prediction should remain consistent across a greater number of cases, but not as much through

Event Type	M1	M2	M3	Best	Worst
VC	49.20	31.08	30.48	M3	M1
MC	12.12	8.07	6.78	M3	M1
VV	20.50	13.41	13.09	M3	M1
MV	44.39	27.47	25.69	M3	M1
Total	118.65	67.92	69.79	M2	M1
				M1: 0	M1: 5
			Totals:	M2: 1	M2: 0
				M3: 4	M3: 0

Table 3: Model errors by event

time.

	M1	M2	M3
MAE	31.55	20.00	19.01

Table 6: Mean Average Error by Model

Month	M1	M2	M3	Best	Worst
2013-01	56.28	27.13	24.83	M3	M1
2013-02	21.34	16.15	18.79	M2	M1
2013-03	36.36	28.41	33.01	M2	M1
2013-04	18.47	30.06	27.05	M1	M2
2013-05	16.86	23.52	14.40	M3	M2
2013-06	13.69	21.08	18.76	M1	M2
2013-07	42.08	32.30	29.42	M3	M1
2013-08	23.27	24.07	19.32	M3	M2
2013-09	32.56	31.14	27.02	M3	M1
2013-10	46.05	34.54	33.48	M3	M1
2013-11	41.08	32.30	29.08	M3	M1
2013-12	49.12	25.57	29.28	M2	M1
2014-01	31.42	22.70	18.30	M3	M1
2014-02	19.65	23.45	18.05	M3	M2
2014-03	20.42	21.54	24.72	M1	M3
2014-04	140.81	28.55	28.04	M3	M1
2014-05	38.04	24.74	25.66	M2	M1
2014-06	52.02	31.23	22.57	M3	M1
2014-07	98.52	29.67	28.43	M3	M1
2014-08	92.92	28.50	16.12	M3	M1
2014-09	37.92	33.38	36.88	M2	M1
2014-10	30.92	21.86	17.41	M3	M1
2014-11	61.24	40.01	50.13	M2	M1
				M1: 4	M1: 17
			Totals:	M2: 4	M2: 5
				M3: 15	M3: 1

Table 4: Material Violence forecasting error by Month

Country	M1	M2	M3	Best	Worst
Angola	4.07	5.48	2.66	M3	M2
Burundi	5.19	4.88	2.11	M3	M1
Chad	9.67	16.59	7.61	M3	M2
Congo	47.17	88.27	49.78	M1	M2
DRC	14.32	22.91	11.90	M3	M2
CAR	46.86	152.17	59.86	M1	M2
Eritrea	1.03	3.43	1.05	M1	M2
Ethiopia	16.69	51.65	15.53	M3	M2
Cote d'Ivoire	6.01	8.80	3.70	M3	M2
Kenya	25.81	40.63	25.05	M3	M2
Liberia	4.14	8.75	3.11	M3	M2
Mozambique	29.43	109.13	31.43	M1	M2
Nigeria	108.44	219.06	135.97	M1	M2
Rwanda	16.23	60.47	23.54	M1	M2
South Africa	26.09	75.28	26.60	M1	M2
Senegal	4.25	9.60	4.17	M3	M2
Sierra Leone	6.54	9.55	3.83	M3	M2
Somalia	25.31	39.54	25.31	M1	M2
Sudan	57.84	130.42	56.41	M3	M2
Tanzania	3.25	9.86	3.34	M1	M2
Uganda	30.70	83.81	32.95	M1	M2
Zambia	2.12	6.34	1.41	M3	M2
Zimbabwe	4.07	9.52	4.31	M1	M2
				M1: 12	M1: 2
			Totals:	M2: 0	M2: 24
				M3: 14	M3: 0

Table 5: Material Violence forecasting error by country

Conclusion

The findings highlighted by contrasting these different models suggest specific cautions should be taken when dealing with event data for prediction. First of all, the remarks made by Goldstein [1992] that daily or even single event data represents the best possible level of inference for understanding actor behaviour patterns. The lack of performance of model 1 can support Shellman's hypothesis that, indeed, monthly aggregation of interaction patterns masks much of the predictive ability of these patterns. As such, sequence-based approaches such as Hidden Markov Models or Sequence Analysis might be the best way of dealing with events as predictors of future actor behaviour.

Second, the better performance of model 3 strengthens the idea that multiple approaches have to be combined, in order to obtain the best possible prediction. In this vein, the inclusion of exogenous variables as predictors of future actor behaviour is also advisable.

Lastly, the better overall performance of the third model, suggests that using simple count patterns to predict coming events might not be the best performing solution. Indeed, looking at single action typologies within the context of broader increases and decreases in actor activity might lead to better prediction.

References

Benjamin Bagozzi. Forecasting civil conflict with zero-inflated count models. *Manuscript. Pennsylvania State University*, 2011.

Robert Blair, Christopher Blattman, and Alexandra Hartman. Predicting local-level violence. In *New York: presented at the Northeast Political Methodology Meeting*, 2014.

Patrick T Brandt and John R Freeman. Advances in bayesian time series modeling and the study of politics: Theory testing, forecasting, and policy analysis. *Political Analysis*, 14(1): 1–36, 2006.

Patrick T Brandt, John R Freeman, and Philip A Schrodt. Real time, time series forecasting of inter-and intra-state political conflict. *Conflict Management and Peace Science*, 28(1): 41–64, 2011.

Vito D'Orazio, James E Yonamine, and Philip A Schrodt. Predicting intra-state conflict onset: An event data approach using euclidean and levenshtein distance measures. In *69th Annual Meeting of the Midwest Political Science Association, Chicago, Il, USA*, 2011.

Vito DOrazio and James Yonamine. Kickoff to conflict: A sequence analysis of intra-state conflict-preceding event structures.

John R Freeman, John T Williams, and Tse-min Lin. Vector autoregression and the study of politics. *American Journal of Political Science*, pages 842–877, 1989.

Kristian Skrede Gleditsch and Michael D Ward. Forecasting is difficult, especially about the future: Using contentious issues to forecast interstate disputes. *Journal of Peace Research*, 50(1):17–31, 2013. doi: 10.1177/0022343312449033. URL `http://jpr.sagepub.com/content/50/1/17.abstract`.

Joshua S Goldstein. A conflict-cooperation scale for weis events data. *Journal of Conflict Resolution*, 36(2):369–385, 1992.

Håvard Hegre, Joakim Karlsen, Håvard Mokleiv Nygård, Håvard Strand, and Henrik Urdal. Predicting armed conflict 2010–2050. *Typescript, University of Oslo*, 2009.

Kentaro Hirose, Kosuke Imai, and Jason Lyall. Can civilian attitudes predict civil war violence? *Available at SSRN 2446168*, 2014.

Rob J. Hyndman and Yeasmin Khandakar. Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software*, 27(3):1–22, 7 2008. ISSN 1548-7660. URL `http://www.jstatsoft.org/v27/i03`.

Kalev Leetaru and Philip A Schrodt. Gdelt: Global data on events, location, and tone, 1979–2012. In *of: Paper presented at the ISA Annual Convention*, volume 2, page 4, 2013.

Jacob M Montgomery, Florian M Hollenbach, and Michael D Ward. Improving predictions using ensemble bayesian model averaging. *Political Analysis*, 20(3):271–291, 2012.

Vladimir B Petroff, Joe H Bond, and Doug H Bond. Using hidden markov models to predict terror before it hits (again). In *Handbook of Computational Approaches to Counterterrorism*, pages 163–180. Springer, 2013.

Jon C. Pevehouse and Joshua S. Goldstein. Serbian compliance or defiance in kosovo? statistical analysis and real-time predictions. *The Journal of Conflict Resolution*, 43(4):pp. 538–546, 1999. ISSN 00220027. URL `http://www.jstor.org/stable/174616`.

Bernhard Pfaff. Var, svar and svec models: Implementation within R, package vars. *Journal of Statistical Software*, 27(4), 2008. URL `http://www.jstatsoft.org/v27/i04/`.

Philip A Schrodt. Pattern recognition of international crises using hidden markov models. *Political complexity: Nonlinear models of politics*, pages 296–328, 2000.

Philip A Schrodt. Forecasting conflict in the balkans using hidden markov models. In *Programming for Peace*, pages 161–184. Springer Netherlands, 2006.

Philip A Schrodt and Deborah J Gerner. Empirical indicators of crisis phase in the middle east, 1979-1995. *Journal of Conflict Resolution*, 41(4):529–552, 1997.

Philip A Schrodt and Deborah J Gerner. An event data analysis of third-party mediation in the middle east and balkans. *Journal of Conflict Resolution*, 48(3):310–330, 2004.

Philip A Schrodt, Ömür Yilmaz, Deborah J Gerner, and Dennis Hermreck. The cameo (conflict and mediation event observations) actor coding framework. In *2008 Annual Meeting of the International Studies Association*, 2008.

Robert Shearer. Forecasting israeli-palestinian conflict with hidden markov models. *Military Operations Research*, 12(4):5–15, 2007.

Stephen M Shellman. Time series intervals and statistical inference: The effects of temporal aggregation on event data analysis. *Political Analysis*, 12(1):97–104, 2004.

Stephen M Shellman. Process matters: Conflict and cooperation in sequential government-dissident interactions. *Security Studies*, 15(4):563–599, 2006.

Jay Ulfelder. A multimodel ensemble for forecasting onsets of state-sponsored mass killing. In *APSA 2013 Annual Meeting Paper*, 2013.

Michael D Ward, Nils W Metternich, Cassy L Dorff, Max Gallop, Florian M Hollenbach, Anna Schultz, and Simon Weschle. Learning from the past and stepping into the future: Toward a new generation of conflict prediction. *International Studies Review*, 15(4):473–490, 2013.

Nils B Weidmann and Michael D Ward. Predicting conflict in space and time. *Journal of Conflict Resolution*, 54(6):883–901, 2010.

James E Yonamine. Predicting future levels of violence in afghanistan districts using gdelt. *Unpublished Manuscript*, 2013.