International Interactions
Empirical and Theoretical Research in International Relations

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gini20

Employing local peacekeeping data to forecast changes in violence

Lisa Hultman, Maxine Leis & Desirée Nilsson

To cite this article: Lisa Hultman, Maxine Leis & Desirée Nilsson (2022) Employing local peacekeeping data to forecast changes in violence, International Interactions, 48:4, 823-840, DOI: 10.1080/03050629.2022.2055010

To link to this article: https://doi.org/10.1080/03050629.2022.2055010

INTERNATIONAL INTERACTIONS
2022, VOL. 48, NO. 4, 823–840
https://doi.org/10.1080/03050629.2022.2055010

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Employing local peacekeeping data to forecast changes in violence

Lisa Hultman, Maxine Leis, and Desirée Nilsson

Uppsala University

ABSTRACT

One way of improving forecasts is through better data. We explore how much we can improve predictions of conflict violence by introducing data reflecting third-party efforts to manage violence. By leveraging new sub-national data on all UN peacekeeping deployments in Africa, 1994–2020, from the Geocoded Peacekeeping (Geo-PKO) dataset, we predict changes in violence at the local level. The advantage of data on peacekeeping deployments is that these vary over time and space, as opposed to many structural variables commonly used. We present two peacekeeping models that contain several local peacekeeping features, each with a separate set of additional variables that form the respective benchmark. The mean errors of our predictions only improve marginally. However, comparing observed and predicted changes in violence, the peacekeeping features improve our ability to identify the correct sign of the change. These results are particularly strong when we limit the sample to countries that have seen peacekeeping deployments. For an ambitious forecasting project, like ViEWS, it may thus be highly relevant to incorporate fine-grained and frequently updated data on peacekeeping troops.

Una forma de mejorar las predicciones es a través de la incorporación de mejores datos. Exploramos hasta qué punto podemos mejorar las predicciones sobre la violencia en los conflictos introduciendo datos que reflejen los esfuerzos de terceros para manejar la violencia. Del conjunto de datos de Geocoded Pacekeeping (Geo-PKO), y aprovechando los nuevos datos subnacionales sobre todos los despliegues de fuerzas para el mantenimiento de la paz de la ONU en África, de 1994 a 2020, predecimos los cambios en la violencia a nivel local. La ventaja de los datos sobre los despliegues de las fuerzas para el mantenimiento de la paz es que estos varían a lo largo del tiempo y del espacio, a diferencia de muchas variables estructurales utilizadas habitualmente. Presentamos dos modelos para el mantenimiento de la paz que contienen varias características locales para dicha actividad, cada uno de ellos con un

KEYWORDS
Civil war; forecasting; peacekeeping; sub-national

CONTACT Lisa Hultman lisa.hultman@pcr.uu.se Department of Peace and Conflict Research, Uppsala University, Box 514, 751 20 Uppsala, Sweden

Supplemental data for this article can be accessed here.

conjunto separado de variables adicionales que forman el respectivo punto de referencia. Los errores medios de nuestras predicciones mejoran solo marginalmente. No obstante, al comparar los cambios observados y predichos en la violencia, las características del mantenimiento de la paz mejoran nuestra capacidad para identificar la señal correcta de cambio. Estos resultados son particularmente sólidos cuando limitamos la muestra a países que han visto desplegar fuerzas de paz. Por lo tanto, para un proyecto ambicioso de predicciones, como el Sistema de Alerta Temprana de Violencia (Violence Early-Warning System, ViEWS), puede ser de suma importancia incorporar datos de grano fino y actualizados frecuentemente sobre las tropas para el mantenimiento de la paz.

L'un des moyens d'améliorer les prévisions est d'utiliser de meilleures données. Nous étudions à quel point nous pouvons améliorer les prédictions de la violence des conflits en introduisant des données reflétant les efforts de tiers de gérer la violence. Nous tirons profit de nouvelles données subnationales sur l'ensemble des déploiements de maintien de la paix de l'ONU en Afrique de 1994 à 2020 qui sont issues du jeu de données Geocoded Peacekeeping (Geo-PKO) pour prédire les évolutions de la violence au niveau local. L'avantage des données sur les déploiements de maintien de la paix est qu'elles varient dans le temps et l'espace contrairement à de nombreuses variables structurelles couramment utilisées. Nous présentons deux modèles de maintien de la paix qui comprennent plusieurs caractéristiques locales de maintien de la paix, chacun avec un ensemble distinct de variables supplémentaires qui constituent leurs références respectives. Les erreurs moyennes de nos prédictions ne s'améliorent que marginalement. Cependant, si nous comparons les évolutions observées et prédites de la violence, il s'avère que les caractéristiques de maintien de la paix améliorent notre capacité à identifier le signe approprié d'une évolution. Ces résultats sont particulièrement solides lorsque nous limitons l'échantillon aux pays qui ont connu des déploiements de maintien de la paix. Pour un projet de prévision ambitieux, comme le projet ViEWS (Violence early-warning system, système d'alerte précoce sur la violence), il peut donc être très pertinent d'intégrer des données détaillées et fréquemment mises à jour sur les troupes de maintien de la paix.

Introduction

There is a wealth of studies showing that peacekeeping operations are successful in reducing armed conflict (Walter, Howard, and Fortna 2021). Yet, only a few studies have evaluated this key policy tool in a prediction framework

(Hegre, Hultman, and Nygård 2019) and no study has so far explored if peacekeeping can improve our predictions of armed conflict at the local level. Predicting the dynamics of political violence at the local level is difficult, as demonstrated by Hegre, et al. (2019). Most spatially disaggregated data that we have are structural variables that do not change much over time, such as geographic or demographic variables. Therefore, in order to predict changes in violence, we have to rely on the history of violence. Any previous hotspot of violence is likely to continue to be a hotspot of violence. However, such a model disregards policy attempts to mitigate violence, and the result is that we risk overpredicting violence in times of de-escalation. One important policy response that varies across both time and space is the deployment of UN peacekeepers. Our contribution is therefore to introduce a model that includes new fine-grained data on peacekeeping deployments to forecast changes in armed conflict at the sub-national level.

In this paper, we present two peacekeeping models (referred to as the pko_slim and the pko_views models). The first is a slightly revised version of the model that we have contributed to the ViEWS (Violence Early-Warning System) prediction competition (Hegre, Vesco, and Colaresi 2022).[1] This contains our local peacekeeping features and an additional 10 variables that we use as the benchmark. The second model contains the same peacekeeping features but instead added to the ViEWS benchmark model (introduced in more detail in the introductory article of the special issue), which contains a larger set of features. In order to evaluate the performance, we compare the pko models to their respective benchmark models.

We find that our peacekeeping models beat their benchmark models in both MSE and TADDA scores for most of the predicted steps, although with varying strength. We furthermore map our predictions from the pko_slim model overlaid with the presence of peacekeeping troops. Those results show that our model that takes both peacekeeping troops and the history of violence into account does fairly well in predicting locations of both escalation and de-escalation of violence in countries with peacekeeping troops. This is furthermore reflected when we compared observed and predicted values in DRC; we find that our model does fairly well in predicting the correct sign of change in violence. As our models have most to offer in countries with peacekeeping operations, we explore their full potential by extending the analysis and running our models on a sub-sample of all countries with peacekeeping deployments. Our model that adds the peacekeeping features to the ViEWS benchmark shows a significant

[1]The revised version is the result of the competition process. The results from the original model are presented in Supplementary Appendix A.1.

improvement over the benchmark in this sample, especially in TADDA scores that reflect the direction of change in violence.

Forecasting is an important tool for humanitarian organizations and other actors operating in conflict areas (e.g. Altay and Narayanan 2020). Predicting dynamics of violence is particularly useful for identifying areas of concern and planning operations. Our analysis demonstrates the relevance and challenges of incorporating data on international responses to violence in forecasting models.

Local Peacekeeping Data

A number of studies have shown that the deployment of peacekeepers reduces the risk of violence in the area close to the base (e.g. Fjelde, Hultman, and Nilsson 2019; Phayal and Prins 2020; Cil et al. 2020; Ruggeri, Dorussen, and Gizelis 2017; Hunnicutt and Nomikos 2020). Since peacekeepers are deployed in response to violence (Ruggeri, Dorussen, and Gizelis 2018; Fjelde, Hultman, and Nilsson 2019), initial peacekeeping deployments may also convey the missions' intelligence about likely future hotspots of conflict escalation. Our data display significant spatial variation over time. The number of troops change from t-1 to t in approximately 10 percent of all observations with peacekeeping presence. Half of those are increases and half are decreases in the number of troops, indicating that this could be a useful variable to include when forecasting changes in conflict violence. We therefore include the number of troops deployed in the cell, a spatial lag that measures the total number of troops in the first order queen neighboring cells and two variables with varying weights attributed to the relationship between events that are distant in space and time. These variables jointly reflect the presence, size, and spatial reach of peacekeepers that can improve predictions by identifying new hotspots of violence and capturing processes of conflict mitigation.

The peacekeeping data come from the Geocoded Peacekeeping Operations (Geo-PKO) dataset (Cil et al. 2020), which is updated as of August 2020. Geo-PKO extracts from deployment maps the exact location, as well as an estimate of the size, of deployments. Since the longitude and latitude of all deployments are coded from maps that also include names of cities and towns, the spatial precision of the data is high. Size estimates are based on the reported number of battalions, companies, and platoons in each location, using standardized numbers. While the actual numbers may vary slightly by deployment, these variations are not likely to be systematic. Mission deployment maps are usually updated with intervals of a few months. We extrapolate the data in between observations, thereby assuming that deployment patterns remain unchanged until there is an updated map.

The last map for each mission is extrapolated until the month that the mission officially closes, or until our last month of observation (i.e. August 2020) if the mission is still ongoing.

Model Specification

We rely on the ViEWS infrastructure (Hegre et al. 2019) to extract all variables except the peacekeeping data, and to generate the predictions for the dependent variable of interest. Given the high level of spatial disaggregation of the peacekeeping data, the level of analysis is the PRIO-GRID-month (pgm), limited to Africa. The data on violence, which is used to construct the dependent variable of changes in violence, is based on the Georeferenced Event Dataset from the UCDP (Uppsala Conflict Data Program) (Sundberg and Melander 2013) and monthly releases of the UCDP Candidate list (Hegre et al. 2020).

Following the standard modelling approach implemented by ViEWS, we split the data into a training, calibration, and test partition (Hegre et al. 2019). This is described in detail in the introductory article of this special issue (Hegre, Vesco, and Colaresi, 2022). Note, that the training period starts in January 1994 due to the availability of the peacekeeping data. Furthermore, the forecasts are estimated with a random forest regressor, see the scikit-learn package (Pedregosa et al. 2011). The defined hyperparameters are in line with the standard ViEWS practise and presented in more detail in the benchmark guidelines (Jansen et al. 2020). The model was tuned by increasing the number of trees to 500. Addressing the computational challenges on the pgm level and considering the unbalanced nature of the outcome, we downsample the data based on a non-stratified random sampling process to only include one percent of all observations with no changes in battle-related deaths. The same procedure is applied to the predictions for the subset.

Our peacekeeping models include seven peacekeeping variables. To reflect the local peacekeeping capacity, the first variable captures the number of peacekeeping troops in the cell-month. Second, we also include this measure at t-1. Third, we include a spatial lag of the number of troops, that captures the extended reach of peacekeepers and spatial effects beyond the cell. Fourth, we include a dummy variable for PKO presence in the country, since peacekeeping data can only improve predictions when there is a peacekeeping operation deployed. Fifth, we include a dummy variable for PKO presence in the cell, to account for any difference between presence and size. Sixth and seventh, we include two “space-time” transformations of the number of peacekeeping troops in the cell-month. Generally, the assumption is made that events occurring at t-1 are equally distant as events occurring in a neighbouring cell. Applying a "space-time" transformation enables us to adjust the weighting of time versus space proximity (Hegre et al. 2019). We set the weights of the space-time features at $t = 0.01$ and at $t = 10$, respectively. Setting the scale to $\mathrm{t} = 0.01$ indicates that peacekeeping troops that are present in the same cell but in a different month are less important for the prediction of our dependent variable of interest than peacekeeping troops that are located in a neighbouring cell. Instead defining $t = 10$ leads to the opposite effect.[2]

The first model that we label the pko_slim model contains another 10 variables—that we also use as benchmark (benchmark_slim). Previous changes in state-based violence is captured by the log change in battle-related deaths over the past one, three and six months, based on best estimates (ged_best_sb_d1, ged_best_sb_d3, ged_best_sb_d6). It should be noted that this is an extension of the simple lags of battle-related deaths as included in the ViEWS benchmark and aims to better capture the specific dynamics of previous changes in the history of violence. The lengths of the history of violence variables are motivated by the small forecasting window of seven months into the future.[3]

We also include space-time transformations of the dichotomous indicator of state-based violence in a cell-month (stdist_k1_t001_ged_dummy_sb, stdist_k1_t10_ged_dummy_sb). As is the case with the "space-time" transformations of the number of peacekeeping troops, these features account for distances to conflict events across space (i.e. longitude and latitude) and time (months) and thus incorporate dynamics that are of special importance for predictions on the pgm level such as the clustering of conflicts over time and space. (Hegre et al. 2019). The spei variables (spei_12, spei_3, spei_1) capture the occurrence of droughts and indicate seasonal changes over 12, 3 and 1 months that may interact with both the deployment of peacekeepers and the dynamics of violence, using data from Vicente-Serrano, Beguera, and Lopez-Moreno (2010). The mean travel time to the nearest city (pgd_ttime_mean) and the logged grid-level population size (pgd_pop_gpw_sum) account for structural local conditions (Tollefsen, Strand, and Buhaug 2012).

Figure 1 contains the complete list of variables included in our pko_slim model, showing their impurity-based feature importance scores. Here we can see that the history of changes in violence are clearly the most important features, while the peacekeeping troops and spatial and temporal transformations are less important. Note, however, that impurity-based feature

[2]Different model specifications that included additional peacekeeping features were also explored, see Supplementary Appendix A.2. for more details.

[3]Additional specifications were tested, such as including all changes between one and seven months, but this did not improve the model performance.

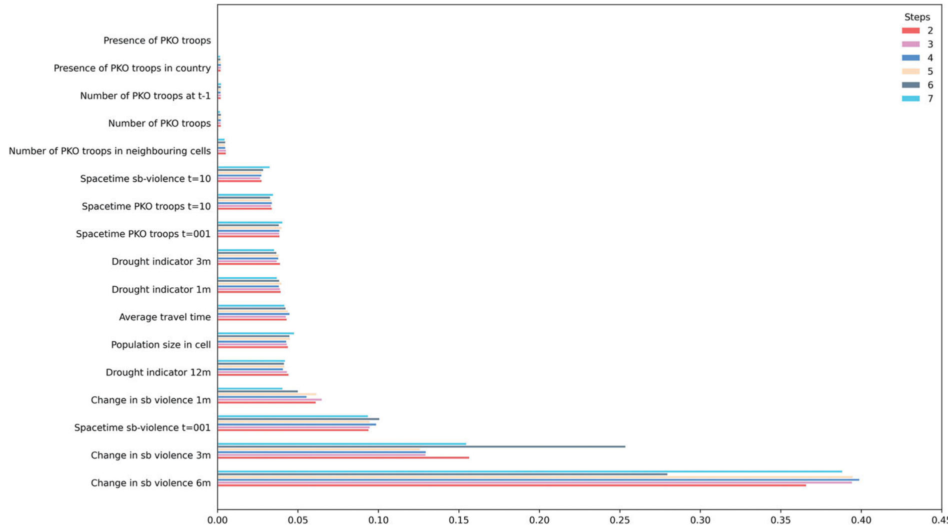

Figure 1. Impurity-based feature importance scores for pko_slim model for all steps.

Table 1. MSE scores.

Model	s2	s3	s4	s5	s6	s7
pko_slim	0.045	0.045	0.047	0.047	0.047	0.047
benchmark_slim	0.048	0.046	0.048	0.049	0.050	0.050
pko_views	0.046	0.046	0.047	0.044	0.047	0.047
benchmark_views	0.045	0.046	0.050	0.048	0.050	0.052

importance scores can be influenced by the correlation of certain features. We think that the combination of these variables, nonetheless, provides a good foundation for the forecasting model. The second model we introduce, referred to as pko_views, contains the same seven peacekeeping variables plus the full ViEWS benchmark model consisting of 40 variables, which is presented in the introductory article of this special issue (Hegre, Vesco, and Colaresi 2022). This second model thus allows us to assess the added value of the peacekeeping variables to the ViEWS framework.

Evaluation: Results 2017–2019

We start evaluating our model performance by comparing MSE scores (Table 1) and TADDA scores (Table 2) for our two peacekeeping models with those of their benchmark models. In general, we can note that the model performances deteriorate over time for most of the models, due to higher uncertainty. Each step refers to one additional month into the future. Looking more closely at the MSE scores, we can see that the pko_slim model is slightly better than the benchmark_slim model; we note a small improvement of the MSE scores over all steps, with a 0.003

Table 2. TADDA scores.

Model	s2	s3	s4	s5	s6	s7
pko_slim	0.117	0.120	0.123	0.124	0.124	0.124
benchmark_slim	0.120	0.120	0.124	0.125	0.126	0.128
pko_views	0.143	0.142	0.145	0.132	0.139	0.140
benchmark_views	0.138	0.140	0.151	0.142	0.151	0.151

difference for three of the steps. The pko_views model performs better than its benchmark for steps 4–7. The biggest improvement of 0.005 can be noted for step 7. These results indicate that including peacekeeping data offers some advantage. The TADDA scores show more variation. The pko_slim model performs better in comparison to the benchmark for five of the six steps. The pko_views model, on the other hand, performs worse than its benchmark for the first two steps, but then improves more significantly for the later steps, with differences up to 0.012. The peacekeeping features thus help us make better predictions for a few months into the future. In general, we can say that the TADDA scores indicate that accounting for peacekeeping increases the models' ability to differentiate between the positive and negative direction in change. We want to note that even though these differences are rather small, the comparison between pko_slim (which is a refined version of the model we submitted to the competition) and benchmark_views shows much larger improvements in TADDA scores. In the Supplementary Appendix, we compare our original competition model to the ViEWS benchmark directly.

Further insights regarding the model performances can be gained by examining the predictions in more detail. We here focus on the pko_slim model. In Figure 2 we show the observed change in violence for January 2019 (step 2). This can be compared to Figure 3 showing the predicted change for the same month. Based on these, we can note that all the locations with observed change in violence have some prediction of change based on the pko_slim model.[4] However, without zooming in, it is difficult to see whether we get the predicted direction right.

Hence, to further illustrate the predictions we zoom in on one region covering CAR, Sudan, South Sudan, and the DRC, where we in addition to the observed and predicted changes, respectively, also plot the troop deployments as of two months before the predictions (Figures 4 and 5). There are a few areas where our peacekeeping model correctly predicts changes in violence. First, the three cells with observed de-escalation in western Sudan shown on Figure 4 are also correctly predicted cells of de-escalation in Figure 5. This is an area where peacekeepers are located in several of the cells and where our peacekeeping model thus contributes to

[4]Note that the light-blue coloring of the grid-cells results from the applied calibration procedure.

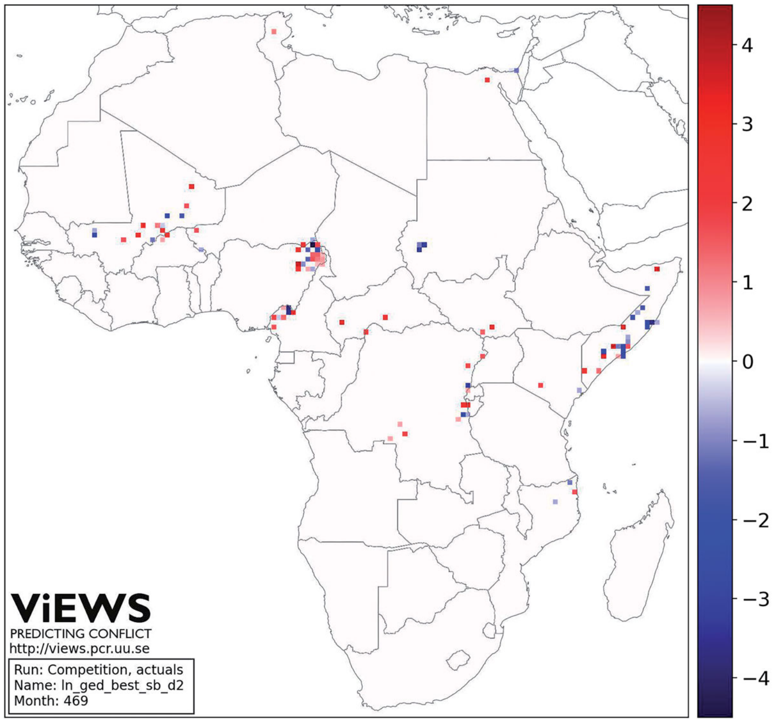

Figure 2. Observed change.

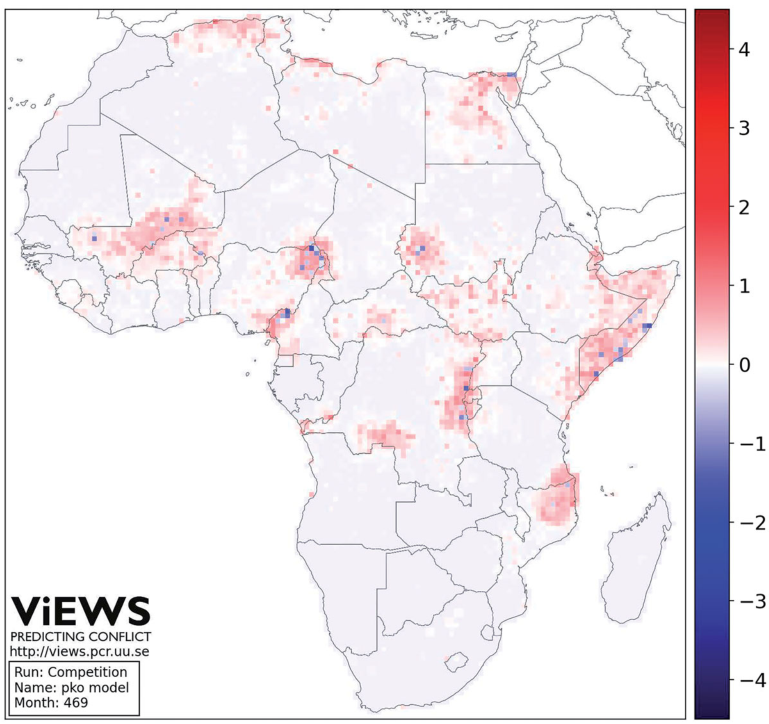

Figure 3. Predicted change pko_slim model.

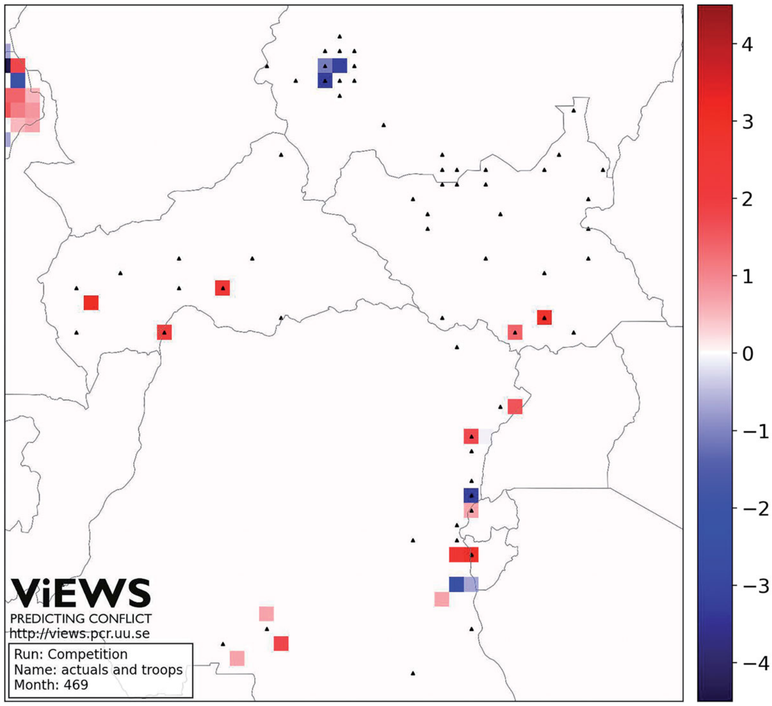

Figure 4. Observed changes & troops: CAR, Sudan, South Sudan and the DRC.

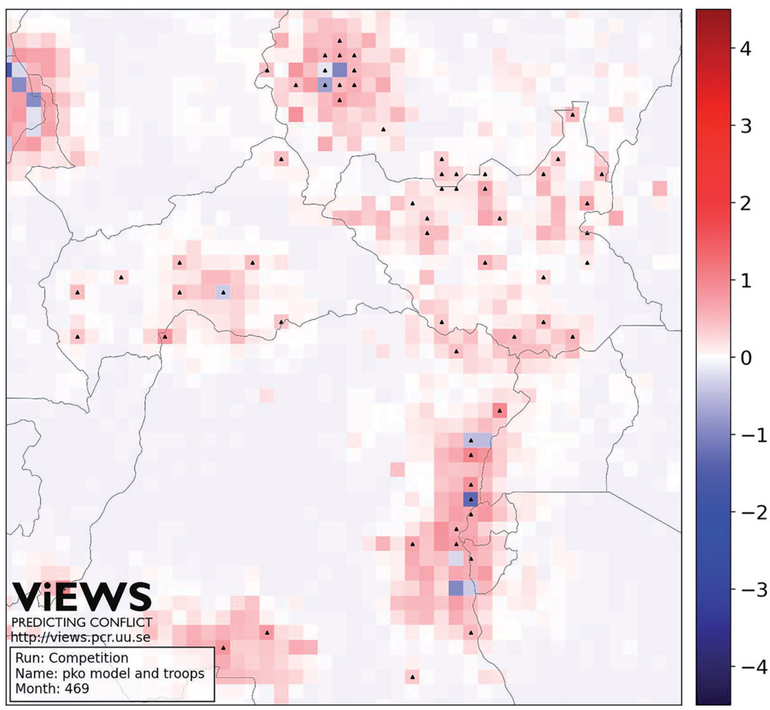

Figure 5. Predicted change pko_slim model & troops: CAR, Sudan, South Sudan, and the DRC.

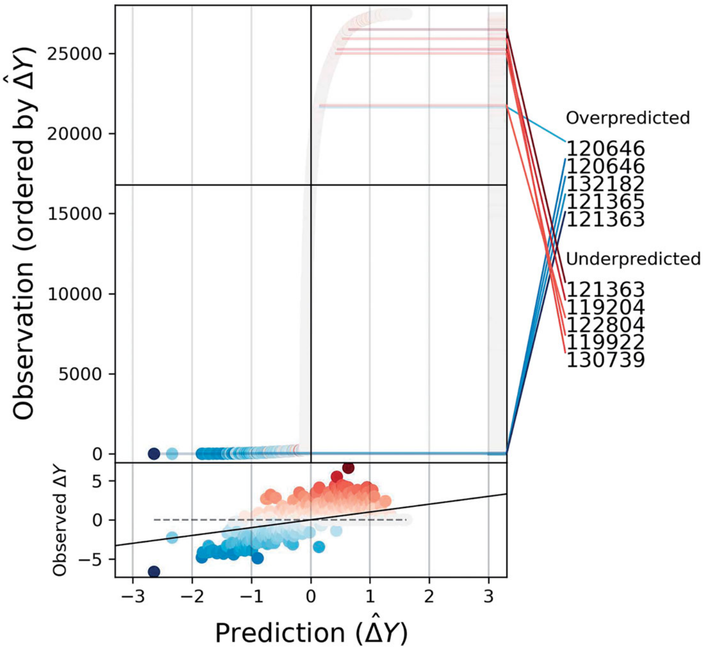

Figure 6. Model criticism plot for pko_slim model, DRC, Step 2.

predicting the changes in violence. Second, in Eastern DRC there are a couple of cells with observed de-escalation that our model identifies as locations of de-escalation. Third, there are a few locations of escalation of violence in South Sudan, in cells where peacekeepers are located, that we also correctly predict. In addition, there are a few observed changes in violence that we get wrong. For example, we miss two out of three cells with observed escalation of violence in CAR. In these specific locations we predict either no change or a de-escalation. However, some of the surrounding cells are predicted to see an escalation of violence. This highlights the need to also take into account the spatial distance between observed and predicted change. Predicting change in a cell neighboring an actual event is often more useful than not predicting anything.[5]

Figure 6 shows the model criticism plot (Colaresi and Mahmood 2017) for the DRC for predictions two months ahead, relying on the pko_slim model. In short, the plot shows the errors between the observed changes in violence and the predicted changes in violence for all observations in the test period for the given step. Dots that are colored in blue indicate

[5]See the article by Vesco et al. (2022) in this special issue for a comparison of all models using the pseudo-Earth Mover Divergence—pEMDiv by Greene et al. (2019), which rewards models that are closer in space even if they miss the exact target.

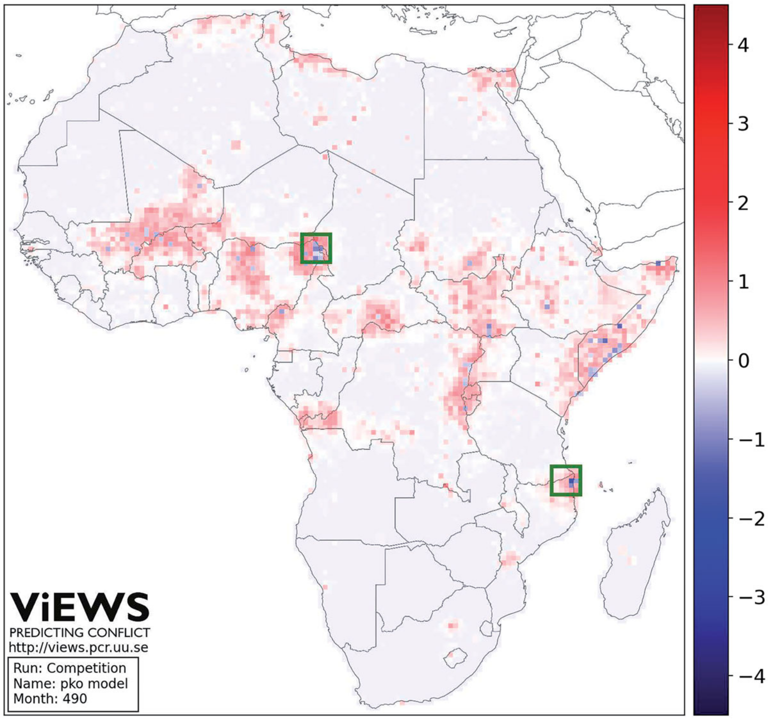

Figure 7. Predicted change in violence pko_slim model (step 2).

observed decreases in the number of battle-related deaths and red dots indicate an observed increase in violence. The darker the color, the greater the changes. We notice that our peacekeeping model (Figure 6) is good at predicting the direction of change (whether positive or negative), as indicated by the smooth separation of the blue and red dots across the diagonal line (bottom end of the figure). The error plot confirms our intuition that including information on the presence of peacekeeping troops, and therefore on factors that contribute to de-escalation, attenuates the problem of overpredicting escalation. However, at the same time it slightly overpredicts de-escalation, i.e. predicting a decrease in violence where we in fact observe an increase. This notion is illustrated by the observation of red dots to the left of 0. Overall, the distribution of the prediction errors of pko_slim confirms that the model is particularly strong in regard to picking up the correct sign of changes in violence, although we tend to predict smaller changes than those observed.

Predictions: October 2020 – March 2021

In this section, we shift our focus from the evaluation of the model performance in the test partition to the description of the true forecasts for October 2020 until March 2021, based on the pko_slim model. Figure 7

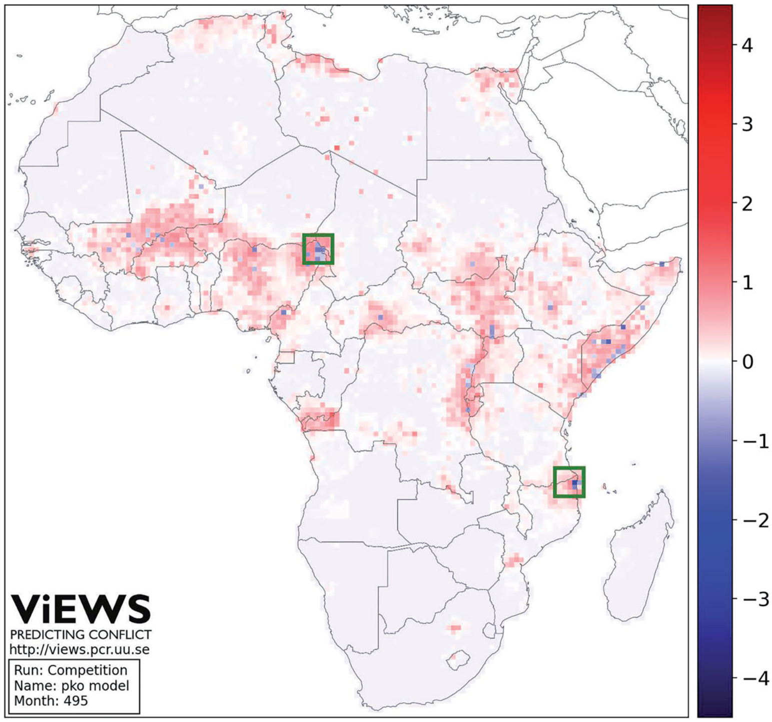

Figure 8. Predicted change in violence pko_slim model (step 7).

shows the predicted changes in violence over two months (step 2) and Figure 8 over seven months (step 7).

We generally expect an increase in battle-related deaths in regions with previous history of violence for October 2020 and March 2021. We can further note that the overall clusters do not change location over time, but the expected intensity of increases and decreases shifts between prio-grid cells. See for example the cells around Juba in South Sudan. It should also be mentioned that the clusters of predicted changes in violence are slightly higher dispersed when predicting for seven months ahead compared to the forecasts for October 2020. This is most likely related to the above mentioned increase in uncertainty over time.

Next, we comment in more detail on two specific regions, marked with green squares on the maps. First, based on Figure 7, we can note that in the border area between Nigeria, Chad, and Cameroon, our model predicts—with an average around 0.2—both increases and decreases in violence. The de-escalation is predicted to take place on the Nigerian side of the border. Since this border region is an area without the deployment of peacekeepers, the predicted changes are thus driven by the other variables in the model, such as the history of violence and seasonal variations.

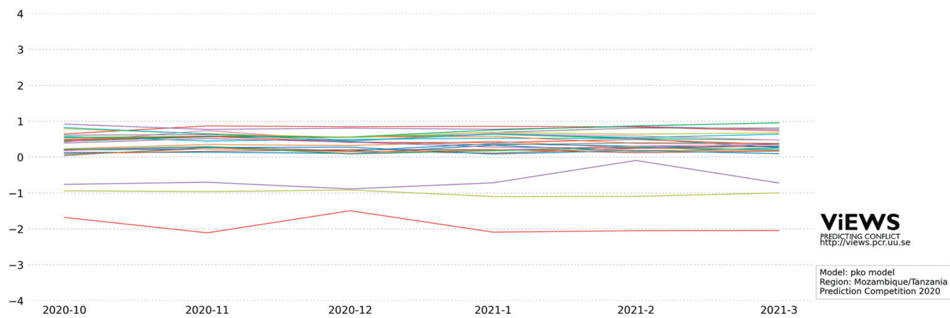

Figure 9. Predicted change in violence over time for border cells between Mozambique and Tanzania (pko_slim model).

Second, the border area between Mozambique and Tanzania show mostly a risk for escalation, with some exceptions. Our model predicts de-escalation in three cells close to the coast in Mozambique. Figure 9 illustrates the predicted changes in violence for the Mozambique-Tanzania border area. The three cells for which we predict a likely de-escalation are captured by the three bottom lines (grey, light brown, and dark brown). This is also a region without peacekeeping presence, and our forecast is thus informed by the other variables in the model. This is highlighted by the cells where we expect a decrease in violence—these have also observed the most extreme fluctuations in violence in the previous months between January and August 2020.

Extensions

Our results are based on the addition of data on UN peacekeeping troops, which means that our model can primarily contribute to improving predictions in countries where peacekeepers are deployed. If one would be interested in making predictions about changes in violence in countries where the UN operates, it would therefore be more efficient to select only those countries for both training the model and making predictions. We evaluate the full potential of our model in this more limited sample. Table 3 shows the MSE scores for the same models we have presented above, but run on a sub-sample of only the countries that at some point in the period 1994–2020 had a peacekeeping operation in place: Angola, Burundi, Central African Republic, Chad, Côte d'Ivoire, Democratic Republic of the Congo, Ethiopia, Eritrea, Liberia, Mali, Morocco (Western Sahara), Mozambique, Rwanda, Sierra Leone, Somalia, South Sudan, Sudan, and Uganda. We include the whole countries for the entire time period, even if they only had peacekeepers for a short period of time, or if they only had deployments on the border.[6]

[6]While one could have considered only focusing on the period when a specific mission was in place, such a setup is not compatible with the ViEWS platform, which we rely on here. Technical limitations regarding how the ViEWS platform implements step-shifting does not allow us to include incomplete time-series data.

Table 3. MSE scores for PKO country subset.

Model	s2	s3	s4	s5	s6	s7
pko_slim	0.059	0.059	0.059	0.061	0.059	0.061
benchmark_slim	0.062	0.059	0.058	0.063	0.062	0.063
pko_views	0.073	0.068	0.074	0.067	0.064	0.064
benchmark_views	0.086	0.087	0.083	0.081	0.080	0.076

Table 4. TADDA scores for PKO country subset.

Model	s2	s3	s4	s5	s6	s7
pko_slim	0.148	0.152	0.151	0.155	0.151	0.153
benchmark_slim	0.149	0.148	0.150	0.156	0.152	0.154
pko_views	0.197	0.188	0.204	0.182	0.175	0.172
benchmark_views	0.219	0.218	0.219	0.215	0.207	0.201

In this extended analysis, we can first note that the MSE scores for the pko_slim model and the benchmark_slim model are similar, but with the peacekeeping model scoring slightly better in four out of six steps. However, the benchmark_slim model is overall a tough comparison, as it includes some key features such as previous changes in the history of violence. Given that we include only countries where peacekeepers have been deployed, we can expect that these countries have also experienced a particularly high level of state-based violence, hence increasing the overall predictive power of the history of violence features. What is noticeable, however, is that when focusing on peacekeeping countries, the pko_views model clearly outperforms the benchmark_views. For example, when comparing them in step 3, the MSE score drops from 0.087 in the benchmark to 0.068 with peacekeeping features included.

Turning to Table 4, we show the TADDA scores for these models using the same subset of PKO countries. Whereas the difference is not so large when comparing the pko_slim model to the benchmark_slim (the TADDA scores are similar with slightly better scores in four out of six steps), we find that the pko_views model clearly outperforms the benchmark_views across all steps. These differences in TADDA scores are much larger than for the full sample. Hence, for a system like ViEWS there is much to be gained by incorporating local peacekeeping data into the forecasting models, especially for those countries where the UN is deployed.

Figure 10 provides an overview of the impurity-based feature importance scores for the pko_slim model for the PKO sample. In comparison to the scores for all countries presented in Figure 1, we can note that the relative importance of the space-time transformation of the number of troops with $t = 0.01$ has increased. The difference in the impurity-based feature importance scores between the space-time transformation with $t = 0.01$ versus $t = 10$ is particularly interesting as it highlights the valuable information we

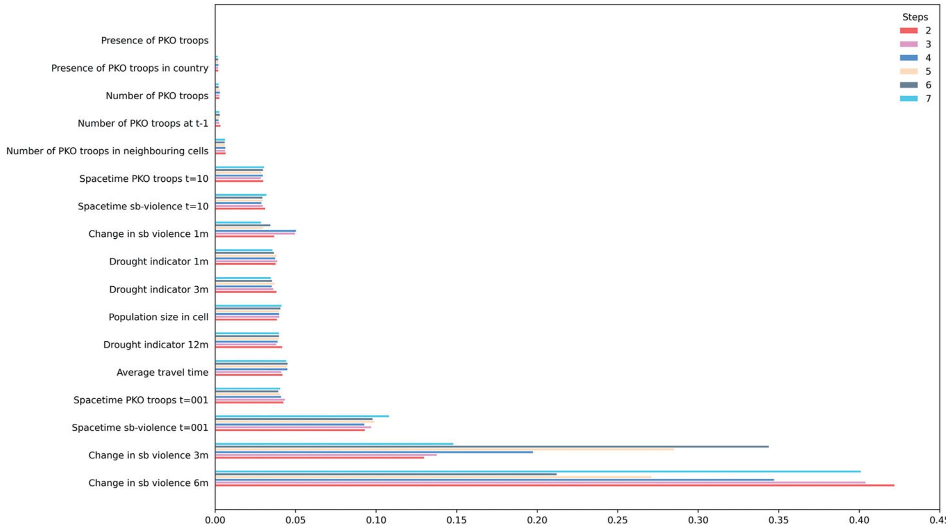

Figure 10. Impurity-based feature importance scores for pko_slim model for all steps for reduced sample.

gain especially from accounting for the spatial variation of peacekeeping troops. This points to the usefulness of including disaggregated data on peacekeeping for predictions in countries where peacekeepers are deployed. In sum, these results confirm our expectation that the peacekeeping model performs best in countries that have hosted peacekeeping operations.

Conclusions

Our paper has evaluated the benefits of adding new and fine-grained data on peacekeeping deployments to a forecasting model of changes in violence. In order to not risk overpredicting violence, we need to consider efforts to mitigate and contain violence. Our contribution points to the usefulness of introducing data on a key policy tool such as peacekeeping.

Adding a set of peacekeeping features to two separate benchmark models, we find that our peacekeeping models generally performs better than each respective benchmark in terms of MSE and TADDA scores, although the differences are rather small. Our strongest results come from focusing only on countries that had a peacekeeping operation during the period under study; our results for adding peacekeeping features to the ViEWS benchmark clearly improves predictions of the direction of change in violence.

Forecasts will always depend on the quality of input data. We draw the conclusion that adding time- and space-varying data that reflect an international policy to counter violence can benefit the forecasts, in particular for the cases that are covered by the data (in our case, countries that have had peacekeeping deployments). While more data in itself is not sufficient to accurately predict the extremely difficult target of changes in violence, it may still be a useful addition to more methodological contributions. For an ambitious forecasting project, like ViEWS, it may thus be highly relevant to incorporate frequently updated data on peacekeeping troops, providing one important piece of the puzzle (especially since maps are often released shortly after the deployment month they refer to). In turn, better forecasts of changes in political violence should be beneficial for third-party actors engaged in efforts to mitigate and contain violence.

Acknowledgments

We would like to thank Håvard Hegre, Paola Vesco, and Michael Colaresi in the ViEWS team, and the participants of the ViEWS competition workshop, for excellent feedback on previous versions. Equal authorship applies. Replication files are available here: http://dvn.iq.harvard.edu/dvn/dv/internationalinteractions

Funding

This work was supported by H2020 European Research Council; (grant number 694640) Knut och Alice Wallenbergs Stiftelse (grant number 2018.0455).

ORCID

Lisa Hultman http://orcid.org/0000-0002-6410-1551
Maxine Leis http://orcid.org/0000-0002-4074-3269
Desirée Nilsson http://orcid.org/0000-0003-0205-2843

References

Altay, Nezih, and Arunachalam Narayanan. 2020. "Forecasting in Humanitarian Operations: Literature Review and Research Needs." *International Journal of Forecasting* Advance online publication. doi:10.1016/j.ijforecast.2020.08.001

Cil, Deniz, Hanne Fjelde, Lisa Hultman, and Desirée Nilsson. 2020. "Mapping Blue Helmets: Introducing the Geocoded Peacekeeping Operations (Geo-PKO) Dataset." *Journal of Peace Research* 57 (2): 360–370. doi:10.1177/0022343319871978.

Colaresi, Michael, and Zuhaib Mahmood. 2017. "Do the Robot: Lessons from Machine Learning to Improve Conflict Forecasting." *Journal of Peace Research* 54 (2): 193–214. doi:10.1177/0022343316682065.

Fjelde, Hanne, Lisa Hultman, and Desirée Nilsson. 2019. "Protection through Presence: UN Peacekeeping and the Costs of Targeting Civilians." *International Organization* 73 (1): 103–131. doi:10.1017/S0020818318000346.

Greene, Kevin, Håvard Hegre, Frederick Hoyles, and Michael Colaresi. 2019. "Move It or Lose It: Introducing Pseudo-Earth Mover Divergence as a Contextsensitive Metric for Evaluating and Improving Forecasting and Prediction Systems." Presented at the 2019 Barcelona School of Economics Summer Forum, Workshop on Forecasting Political and Economic Crisis: Social Science Meets Machine Learning, Barcelona, June 17. https://events.barcelonagse.eu/summerforum/event/7461#date/20220607.

Hegre, Håvard, Lisa Hultman, and Håvard Mokleiv Nygard. 2019. "Evaluating the Conflict-Reducing Effect of UN Peacekeeping Operations." *The Journal of Politics* 81 (1): 215–232. doi:10.1086/700203.

Hegre, Håvard, Marie Allansson, Matthias Basedau, Michael Colaresi, Mihai Croicu, Hanne Fjelde, Frederick Hoyles, et al. 2019. "ViEWS: A Political Violence Early-Warning System." *Journal of Peace Research* 56 (2): 155–174. doi:10.1177/0022343319823860.

Hegre, Håvard, Mihai Croicu, Kristine Eck, and Stina Högbladh. 2020. "Introducing the UCDP Candidate Events Dataset." *Research and Politics*. Advance online publication. doi:10.1177/2053168020935257

Hegre, Håvard, Paola Vesco, and Michael Colaresi. 2022. "Lessons from an Escalation Prediction Competition." *International Interactions* 48 (4).

Hunnicutt, Patrick, and William G. Nomikos. 2020. "Nationality, Gender, and Deployments at the Local Level: Introducing the RADPKO Dataset." *International Peacekeeping* 27 (4): 645–672. doi:10.1080/13533312.2020.1738228.

Jansen, Remco, Håvard Hegre, Michael Colaresi, and Frederick Hoyles. 2020. "Benchmark Models for the ViEWS Prediction Competition." ViEWS working paper, Uppsala University, June 16.

Pedregosa, Fabian, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, et al. 2011. "Scikit-Learn: Machine Learning in Python." *Journal of Machine Learning Research* 12: 2825–2830.

Phayal, Anup, and Brandon C. Prins. 2020. "Deploying to Protect: The Effect of Military Peacekeeping Deployments on Violence against Civilians." *International Peacekeeping* 27 (2): 311–336. doi:10.1080/13533312.2019.1660166.

Ruggeri, Andrea, Han Dorussen, and Theodora Ismene Gizelis. 2017. "Winning the Peace Locally: UN Peacekeeping and Local Conflict." *International Organization* 71 (1): 163–185. doi:10.1017/S0020818316000333.

Ruggeri, Andrea, Han Dorussen, and Theodora Ismene Gizelis. 2018. "On the Frontline Every Day? Subnational Deployment of United Nations Peacekeepers." *British Journal of Political Science* 48 (4): 1005–1025. doi:10.1017/S000712341600017X.

Sundberg, Ralph, and Erik Melander. 2013. "Introducing the UCDP Georeferenced Event Dataset." *Journal of Peace Research* 50 (4): 523–532. doi:10.1177/0022343313484347.

Tollefsen, Andreas Foro, Håvard Strand, and Halvard Buhaug. 2012. "PRIO-GRID: A Unified Spatial Data Structure." *Journal of Peace Research* 49 (2): 363–374. doi:10.1177/0022343311431287.

Vesco, Paola, Håvard Hegre, Michael Colaresi, Remco Bastiaan Jansen, Adeline Lo, Gregor Reisch, and Nils B. Weidmann. 2022. "United They Stand: Findings from an Escalation Prediction Competition." *International Interactions*. doi:10.1080/03050629.2022.2029856

Vicente-Serrano, Sergio M., Santiago Beguera, and Juan I. Lopez-Moreno. 2010. "A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index." *Journal of Climate* 23 (7): 1696–1718. doi:10.1175/2009JCLI2909.1.

Walter, Barbara F., Lise Morje Howard, and V. Page Fortna. 2021. "The Extraordinary Relationship between Peacekeeping and Peace." *British Journal of Political Science* 51 (4): 1705–1722. doi:10.1017/S000712342000023X.