A DISCOURSE-DRIVEN INTERVENTION RECOMMENDATION FRAMEWORK FOR UNITED NATIONS PEACEKEEPING IN POST-COLONIAL AFRICA

Brian Zhou¹ and Alvan Caleb Arulandu²

¹Thomas Jefferson High School for Science and Technology, Alexandria, Virginia, USA
²Harvard University, Cambridge, Massachusetts, USA

ABSTRACT

The United Nations (UN) is the foremost international body helping uphold world peace through peacekeeping missions, ranging from deployments that enforce peace treaties, monitor conflicts, and protect civilians; However, determining when and how to intervene is complex. The updated UN General Debate Corpus (UNGDC), cataloging every speech from the UN's inception in 1946 to 2022, is a treasure trove of national policy, as the UNGD is the only body where every country can speak. We propose a discourse-driven intervention recommendation framework that categorizes ongoing conflicts based on UN precedent and recommends the magnitude of funds and forces that should be committed to addressing a conflict. We employ natural language processing techniques to tokenize, preprocess, and analyze word stem frequencies in the UNGDC, generating a timeseries of the number of UN mentions for any given country. Paired with historical analysis, we show that debate in the UNGDC is a potent indicator to determine UN intervention and response mechanisms for conflicts in Africa; further, by aggregating mention statistics across periods of active conflict, we provide quantitative backing for the correlation of mention dynamics and the presence of an active conflict, for a given country. Finally, we present and test an interpretable, shallow decision tree model that can perform intervention type classification and response magnitude recommendation with 91.7% accuracy. Our results, established by computational experiments and statistical testing, suggest that corpus analysis and broader computational diplomacy methods can drive intervention recommendations to improve the UN's decision-making.

KEYWORDS

United Nations, conflict resolution, natural language processing, decision trees, computational diplomacy, UN peacekeeping, UN General Assembly General Debate, UN General Debate Corpus, political communication

1. INTRODUCTION

United Nations (UN) peacekeepers, run by the UN Peacebuilding Commission, operate around the world to assist in humanitarian interventions, protect civilians and refugees in civil war and conflict, and promote post-war peace and peaceful transitions. Peacekeepers operate as a representation of the international community's will to protect and uphold human rights while promoting and overseeing demobilization and transitions to peace [1].

David C. Wyld et al. (Eds): SIPR, NCWC, BIBC, SOFEA, DSML, ARIA, NLP, CSEN -2023
pp. 349-369, 2023. CS & IT - CSCP 2023
DOI: 10.5121/csit.2023.131927
UN peacekeepers operate around the world with varying levels of success. In Rwanda and Bosnia [2], UN peacekeepers failed to protect civilian lives in devastating massacres and genocides, but UN peacekeepers saw success in other operations such as the UN Interim Administration Mission in Kosovo and the UN Transitional Administration in East Timor (UNTAET) [3]. The efficacy of individual peacekeeping missions, alongside the overall efficacy of the UN peacekeeping body, has been disputed by various scholars and researchers. Scholars and activists have called out varying forms of human rights violations conducted by UN peacekeepers against civilians [1], including sexual exploitation by UN peacekeepers [4]. Scholars also dispute the effectiveness peacekeepers bring, drawing various conclusions on the effects of peacekeeping deployments:

- Prior and in early stages of a conflict, increased budget for UN peacekeeping operations (PKOs) reduces armed conflicts by up to two thirds in a simulation against an alternative without early peacekeeping action [5].
- During conflicts, UN operations are more effective compared to operations conducted by countries to reduce civilian casualties [6], but often hinge on cooperation from actors inside the state [7]. Other scholars dispute the effectiveness of non-UN peacekeepers, finding that non-UN operations are often effective at limiting civilian violence [8] and are similarly effective to mitigate violence [9]. Literature usually agrees that the UN is significantly more effective in creating treaties and agreements for peace [8].
- After a conflict happens, UN peacekeepers increase the likelihood that combatants agree to negotiate and a cease-fire [10]. UN peacekeepers are effective at short-term achievement of peace; longer deployments increase the chance a peace continues [11], and similarly, UN peacekeeping missions are effective at enforcing rule of law in times of peace when reforms are being implemented [12].

In this environment, knowing when, where, and how peacekeeping forces will be deployed is critically important for promoting peace and security in war torn regions. International organizations (IOs) such as the United Nations with peacekeeping capacities are not only more effective at creating and enforcing peace agreements, but create forward-facing effects that the discourse of a potential peacekeeping force creates a credible commitment, a theory in conflict resolution demonstrating that discourse surrounding UN interventions helps bring peace and create more effective conflict resolutions [10]. Thus, predicting a UN Security Council (UNSC) decision to deploy peacekeepers can shine light into evaluating the credible commitment theory and serve as an effective push to promote peace in emerging conflict zones. This knowledge of future operations can assist other humanitarian organizations and allow researchers to compare the effectiveness between an implemented UN PKO scale and a predicted need, allowing future missions to adapt and commit a credible force to promote peace.

Fortunately, linguistics is a flexible tool that is able to turn qualitative data into quantifiable data for data science in the field of international relations and politics, and has the potential to be able to detect and predict discourse surrounding ongoing and future UN peacekeeping deployments. Historically, linguistics and Natural Language Processing (NLP) has been applied to a variety of applications in international relations using a variety of Corpora (compilations of text), including diplomatic documents [13], real-world conflict detection using news [14], and foreign influence from social media [15], to name a few. The growing popularity of data science and NLP approaches in social science allows the use of these Corporuses to answer important questions. One such Corpus is the UN General Debate Corpus (UNGDC) created by A. Baturo [16]. We utilize this corpus to answer questions about how post-independent states in the Global South emerge on the international stage, including their prevalence in UN discourse and distinguishing factors between African regions. The UNGDC is comprised of speeches from the representatives of all 193 UN member states given each year in an address to the UN General Assembly (UNGA) discussing a state’s stance on global issues, outlining the issues they deem most important, and
laying out their national and international agendas. These speeches are important, as the UNGA is often referred to as a "barometer of international opinion" [17]. By being text-based, the UNGDC offers a promising solution to the large gap of quantitative data that existed within the social sciences just decades prior, and which continues to persist today. Notably, the broad nature of the dataset allows for a variety of quantitative studies to emerge.

For example, prior studies with the UNGDC have researched how specific world leaders, countries, or supranational bodies engage or utilize rhetoric at the UNGA, like Pakistan [18] and the speeches given by the former Prime Minister of Pakistan Nawaz Sharif [19] or the European Union [20]. Studies include a wide range of topics, including rhetoric, Islamophobia, organizational management, gender, language policy, and economic theory, respectively:

- Evaluating the impact of cross-cultural differences on the explicitness and persuasiveness of rhetoric between cultures at the UN [21]
- Utilizing Critical Discourse Analysis to study the portrayals of Islam and Muslims in UN speeches [22]
- Studying benign neglect in an organization’s rhetoric [23]
- Examining trends in gender representation in policymakers as well as gender representation and linguistics in topic discourse at the UNGA [24]
- Highlighting language policy and ideology of UN members [25]
- Analyzing how economic scarcity between lower-income and higher-income countries influences narratives of want and wealth in speeches given at the UN [23]

Many studies have also expanded the UNGDC dataset to new parts of the UNGDC, including:

- UNSCdeb8, a corpus containing all verbatim statements of permanent and non-permanent members of the UNSC by [26]
- spaceTexts, a corpus containing all speeches related to the UN Committee on the Peaceful Uses of Outer Space including state and nonstate actors by [27]
- UNSC Afghanistan, a corpus combining quantitative and qualitative approaches to analyze dynamic topics in Afghanistan from UNSC speeches conducted by [28]

We study only the UNGDC, where new novel areas of study have been created by the expansion of the dataset. The most recent dataset created by [29] extends the original UNGDC, which covered all speeches from 1970 to 2017, to the first UNGA session, now covering from 1946 to 2022. With over 10,000 speeches from more than 193 countries, the updated UNGDC is a treasure trove of linguistic data.

Uniquely, the expanded UNGDC dataset allows us to study the complete history of post-colonial states such as Nigeria or Algeria. These post-colonial states often emerged beginning in the 1950s and joined the UN the same year as their independence. All of them joined after the formation and establishment of the UN, meaning that for the entirety of the nation's existence, they have been UN members. This uniquely allows us to examine and catalog the development of these independent states as they develop economically, gain footing on the global stage, and also observe how the international community reacts and resolves turmoil in these new states.

In Materials and Methods, we discuss the composition of the UNGDC dataset, the advantages of using the UNGDC dataset for mapping the shift of a country's national priorities over time, and the specific coverage that the UNGDC dataset offers for representing lower-income countries. We use the UNGDC dataset to first develop a time series of the 'relevance' of a nation in UNGDC speeches, based on the number of references a country receives in a given UNGD. We explain the novel ways that the UNGDC dataset is able to explain trends such as UN involvement in the
peacekeeping and conflict resolution process. Computational Experiments and onward detail how we extract features and tokenize speeches to conduct our experiment, as well as the conditions validated to conduct our statistical tests.

In Results, we conduct a select review of certain UN PKOs in Africa alongside compute differences in discourse between various conflicts, evaluating determining factors to develop an interpretable machine learning model that can predict the type of peacekeeping operation and magnitude of committed funds and forces throughout the length of the operation. We discuss the implications of our analysis and review future work that can develop a multimodal model including other sources of information such as UNSC data and large datasets from traditional press and social media platforms.

2. MATERIALS AND METHODS

2.1. Advantages of the UNGDC Dataset

Every September, the heads of state and other high-level country representatives gather in New York at the start of a new session of the United Nations General Assembly (UNGA) and address the Assembly in the General Debate (GD). The GD provides the governments of the almost 200 UN member states with an opportunity to present their views on international conflict and cooperation, terrorism, development, climate change and other key issues in international politics. As such, the statements made during the GD are an invaluable and largely untapped source of information on governments’ policy preferences and global involvement across a wide range of issues over time. Thus, statements made during the GD can roughly approximate the relevancy of African nations, but conflicts take center stage, dominating discourse rates.

Uniquely, the UNGD provides a platform for all member states to deliver addresses, including smaller nations otherwise unrepresented [16], with the intention of all speeches being to justify their stance on foreign policy and persuade other UN states to adopt positions similarly [29]. This updated corpus contains over 10,000 speeches from more than 193 countries, making it the most comprehensive collection of global political discourse. This corpus can facilitate further research on a range of issues in international relations – from looking at different influences on state preferences to understanding the spread of ideas and norms in international politics.

2.2. Feature Extraction

Before using the UNGD, we first perform feature extraction to generate metrics for further statistical analysis. Firstly, we generate a number of characteristics describing each speech by segmenting the speech text according to stopwords provided by Python’s Natural Language Toolkit [30]. Doing so, we characterize each speech according to its character count, word count, sentence count, average word length, and average sentence length.

After generating these summary statistics, we tokenize each piece of speech text. In the English language, words can be modified in a number of ways while preserving meaning. For example, the words "nation", "nations", and "national" all stem from a common linguistic concept though they have different spellings. In order to perform an accurate linguistic analysis of UNGDC speeches, multiple variations of a single word stem must be consolidated. To execute this, we use WordNet's [31] morphological processing capabilities to isolate the stems of each word, removing affixes and other lexicographic nuances.
3. **Computational Experiments**

After the data has been sufficiently pre-processed in the above manner, we conduct multiple computational experiments to address the discussion of African regions in the UN.

3.1. **Frequency Analysis**

We begin by aggregating the number of mentions for each word temporally segmented by year. Sorting words by total frequency across the temporal span of the dataset, we can compute the most common words used. The results in Figure 1 are quite intuitive, as the most popular words strongly relate to international relations and government structure. Cross-referencing these tokens with a database of countries from Python’s pycountry package, we can compute the number of mentions over time for any given country.

From Figure 2, the most discussed countries are those heavily involved in international conflicts. Using a database of African regions and the nations which consist of them, we aggregate the net number of mentions for each region over time. From Figure 3, we see that nations in Eastern Africa seem to receive the most attention followed by Western, Northern, Southern, and Central Africa. Regions are segmented using the UN Geoscheme for Africa from the UN M49 [32]. Dependencies or overseas territories of European countries are excluded in this data, and the disputed territory of Western Sahara is included.

![Figure 1. Most Mentioned Words in UNGDC Speeches](image_url)
3.2. Summary Statistics

We migrate the time series data from Figure 3 from Python to R for further analysis. We begin by transforming the time series data into a linear regression problem, viewing the quantitative response variable, the total Count of mentions, as a function of the quantitative explanatory variable Year and the categorical explanatory variable Region. Grouping this rearranged data by region, we compute summary statistics.
Table 1. Summary Statistics for African Region Mentions

<table>
<thead>
<tr>
<th>Region</th>
<th>min</th>
<th>Q1</th>
<th>median</th>
<th>Q3</th>
<th>max</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastern</td>
<td>94</td>
<td>311.8</td>
<td>360.0</td>
<td>523</td>
<td>1021</td>
<td>195.3</td>
</tr>
<tr>
<td>Central</td>
<td>46</td>
<td>121.0</td>
<td>165.5</td>
<td>221</td>
<td>539</td>
<td>88.1</td>
</tr>
<tr>
<td>Northern</td>
<td>39</td>
<td>149.0</td>
<td>172.5</td>
<td>211</td>
<td>750</td>
<td>104.5</td>
</tr>
<tr>
<td>Southern</td>
<td>11</td>
<td>76.5</td>
<td>76.5</td>
<td>370</td>
<td>577</td>
<td>194.8</td>
</tr>
<tr>
<td>Western</td>
<td>106</td>
<td>195.2</td>
<td>240.0</td>
<td>297</td>
<td>446</td>
<td>73.9</td>
</tr>
</tbody>
</table>

From Table I, we see that the medians differ between each region, suggesting a potentially significant result.

3.3. Diagnosing Condition Violations

A priori, we propose and fit a multiple linear regression model with \( n - 1 = 4 \) indicator variables to account for the Region. From the Q-Q plot in Figure 4, the majority of the data adheres to normality, but there is concave curvature indicating a short tail on the right-hand side of the distribution. Further, we analyze the deviation in spread according to region.

\[
\text{Count} = \beta_0 + \beta_1 \cdot \text{Year} + \sum_{k=1}^{4} \beta_{k+1} \cdot \text{Count} \approx + - 3315170 \cdot 1.52017 \cdot l - w 1.453 \cdot \\
\text{Year} - 237.652 \cdot l - 227.522 \cdot l - 228.022 \cdot l
\]

From Table I, \( \max \min (\sigma / \mu) = 19573.93 = 2.64 > 2 \).

This violation of normality and equal spread is enough to consider a data transformation.

3.4. Power Transformation

Since our data is grouped by species, we construct a diagnostic plot for computing the ideal exponent for a power transformation using the regression: \( \log(\sigma) = \beta_0 + \beta_1 \cdot \log(\mu) \).
Fitting this model, we find $\beta_1 = 0.538$, so we use the following transformation:

$$x \rightarrow f(x) = x^p = x^{1-\beta_1} = x^{0.462}$$

While this transformation is similar to $\sqrt{x}$, we use this more accurate form in order to maximally adhere to normality.

### 3.5. Validating Conditions

Applying $f(x)$ to both $Count$ and $Year$, we fit the following regression model.

$$f(Count) = \beta_0 + \beta_1 \cdot f(Year) + \sum_{k=1}^{4} \beta_k \cdot I_k$$

$$f(Count) \approx + 188.329 - 5.157 \cdot f(Year) - 5.086 \cdot I_C - 4.810 \cdot I_N - 6.082 \cdot I_S - 3.218 \cdot I_W$$
From the Q-Q plot in Figure 6, we see that while the normality condition is not perfectly satisfied, the data adheres enough to satisfy the condition. From the residual plot in Figure 7, we see that the variance across regions is sufficiently similar for the equal variance condition and note that the residuals have zero mean.

Since the entire population of UNGDC debates is used, the randomized and independent conditions are independently satisfied. We also assume that the data is approximately linear. With these conditions met, we continue to a significance test.

$$p_{c} = 1.2 \times 10^{-12}; p_N = 1.4 \times 10^{-11}; p_s = 2 \times 10^{-16}; p_w = 3.4 \times 10$$

p-values for each indicator are all significant.

### 3.6. Significance Testing

We t-test the indicator weights where $H_0: \beta_i = 0$ and $H_a: \beta_i \neq 0$ for $2 \leq i \leq 5$.

$$p_{c-6}$$

These resulting

$$p_c, p_N, p_s, p_w < \alpha = 0.05$$

These indicator variable t-tests suggest that $\beta_i \neq 0$ for $2 \leq i \leq 5$, rejecting the null hypothesis and suggesting the validity of the alternate hypothesis. This is statistically significant evidence that the mean number of mentions after accounting for time is significantly pairwise different between all five African regions.

### 3.7. Complementary Correlation Analysis
Given that the conditions for multiple linear regression were not satisfied perfectly, we also analyze the Pearson correlation coefficient between each region. From Table II, most regions are weakly associated with one another. The most associated pair is Southern and Central Africa, and yet, even for this pair, $r = 0.466 < 0.5$ which is not large. This result further supports the conclusion from significance testing.

Table 2. Pairwise Pearson Correlation Between Regions

<table>
<thead>
<tr>
<th>Region</th>
<th>Eastern</th>
<th>Western</th>
<th>Northern</th>
<th>Southern</th>
<th>Central</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastern</td>
<td>1.000</td>
<td>0.228</td>
<td>-0.032</td>
<td>0.068</td>
<td>0.004</td>
</tr>
<tr>
<td>Western</td>
<td>0.228</td>
<td>1.000</td>
<td>0.293</td>
<td>-0.344</td>
<td>-0.043</td>
</tr>
<tr>
<td>Northern</td>
<td>-0.032</td>
<td>0.293</td>
<td>1.000</td>
<td>-0.176</td>
<td>-0.138</td>
</tr>
<tr>
<td>Southern</td>
<td>0.068</td>
<td>-0.344</td>
<td>-0.176</td>
<td>1.000</td>
<td>0.466</td>
</tr>
<tr>
<td>Central</td>
<td>0.004</td>
<td>-0.043</td>
<td>-0.138</td>
<td>0.466</td>
<td>1.000</td>
</tr>
</tbody>
</table>

4. RESULTS

Understanding the timing and method by which the UN acts during crises is key to understanding the effectiveness of UN solutions, such as peacekeeper deployments in Africa [33]. The UNGD is key to facilitating these understandings, as it has primarily served as a vehicle to formulate multilateral action and dialogue.

Table III demonstrates trends with UN discourse for all conflicts collected in our data. We select civil wars and conflicts that have lasted 3+ years (a shorter conflict will still be included if a UN peacekeeping action was taken), and compare the average mentions during the conflict with the average of the state(s) throughout its history.

Figure 8 demonstrates a spike in discourse regarding Chad during the Chad-Libyan war. Notably, the selected word matters, as “Chad” was a preferred term by UN delegates to refer to the “Chadian-Libyan War” overall. Although Table III finds discourse trends up during conflict, two general exceptions exist. First, some states have conflicts that dominate UN discussion, such as Libya, which has a high peak due to the Libyan Civil War that overshadows other conflicts involving Libya. This is best demonstrated through Figure 8, which shows that discourse for Libya remained low. Second, some conflicts did not last long enough or were not severe enough to attract discourse, such as Kenya and Uganda’s border conflict.
4.1. UN Peacekeeping Missions

The UN Charter authorizes the UN Security Council to take collective action to maintain international peace and security. This is carried out through UN peacekeepers and UN peacekeeping missions, the primary instrument by which the United Nations decides to intervene to preserve or create conditions for peace in hostile conflicts and environments. Although the UN primarily describes peacekeepers as vehicles to create conditions for lasting peace, various UN peacekeeping missions are dependent on the context of the mission and regional background; missions range from protecting demilitarized or arbitrated ceasefire borders and stabilizing nations post-conflict to preventative missions to ensure peaceful transitions or observe developing conditions on the ground.

In total, more than 1 million personnel including military, police, and civilians from 125 countries have served in 71 peacekeeping missions worldwide [34]. In Africa, the UN has authorized 29 total peacekeeping missions. Of these, 23 missions are complete, and 6 missions are currently ongoing. Of the 29 missions, we did not analyze the United Nations Operation in the Congo (ONUC), as the Congo Crisis (1960-1965) fell outside of the range of the UNGDC dataset.

The UN are good recordkeepers of quantitative counts of past peacekeeping operations. Data on UN PKOs sourced in Table III can be found at [35], but the depth of data differs between each mission; historic missions lack the same level of depth when having exact numbers for military or civilian forces provided by each participating country, for example, while still ongoing missions lack a final expenditure count as budgets are revised and annual expenditures often go below the allocated annual budgets.
<table>
<thead>
<tr>
<th>State</th>
<th>Conflict</th>
<th>Period</th>
<th>Base Rate</th>
<th>Conflict Rate</th>
<th>% Chg. of Rate</th>
<th>UN Action</th>
<th>Budget (USD)</th>
<th>Military Strength</th>
<th>Civilian Strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAM</td>
<td>Namibian War for Independence</td>
<td>1966-1990</td>
<td>169.4</td>
<td>339.5</td>
<td>100.5</td>
<td>UNTAG (1989-1990)</td>
<td>368.6mil</td>
<td>6993/7500</td>
<td>000, 000</td>
</tr>
<tr>
<td>ETH</td>
<td>Ethiopian Civil War</td>
<td>1974-1991</td>
<td>28.2</td>
<td>28.4</td>
<td>0.71</td>
<td>The UN did not intervene with PKO(s) in this conflict.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESH</td>
<td>Western Sahara War</td>
<td>1975-1991</td>
<td>45.0</td>
<td>70.9</td>
<td>57.4</td>
<td>MINURSOU (1991–3)</td>
<td>Ongoing</td>
<td>202</td>
<td>226</td>
</tr>
<tr>
<td>MOZ</td>
<td>Mozambican Civil War</td>
<td>1977-1992</td>
<td>41.1</td>
<td>43.9</td>
<td>6.81</td>
<td>ONUMOZ (1992-1994)</td>
<td>492.6mil</td>
<td>5063</td>
<td>861, ~1561 during election</td>
</tr>
<tr>
<td>SOM</td>
<td>Ethio-Somali War</td>
<td>1977-1978</td>
<td>59.4</td>
<td>44.5</td>
<td>-25.08</td>
<td>The UN did not intervene with PKO(s) in this conflict.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UGA</td>
<td>Ugandan-Tanzania War</td>
<td>1978-1979</td>
<td>22.5</td>
<td>18.5</td>
<td>-17.78</td>
<td>The UN did not intervene with PKO(s) in this conflict.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCD</td>
<td>Chadian-Libyan Conflict</td>
<td>1978-1987</td>
<td>41.8</td>
<td>105.6</td>
<td>152.6</td>
<td>UNASOG (1993-1995)</td>
<td>64,471</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>ERI</td>
<td>Second Eritrean Civil War</td>
<td>1980-1981</td>
<td>14.4</td>
<td>1.5</td>
<td>-89.6</td>
<td>The UN did not intervene with PKO(s) in this conflict.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UGA</td>
<td>Ugandan Bush War</td>
<td>1980-1986</td>
<td>22.5</td>
<td>15.1</td>
<td>-32.88</td>
<td>The UN did not intervene with PKO(s) in this conflict.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDN</td>
<td>2nd Sudanese Civil War</td>
<td>1983-2005</td>
<td>50.5</td>
<td>43.7</td>
<td>-13.47</td>
<td>The UN did not intervene with PKO(s) in this conflict.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LBR</td>
<td>First Liberian Civil War</td>
<td>1989-1997</td>
<td>36.0</td>
<td>86.2</td>
<td>139.4</td>
<td>UNOMIL (1993-1997)</td>
<td>103.7mil</td>
<td>368</td>
<td>284</td>
</tr>
<tr>
<td>NER</td>
<td>Tuareg Rebellion</td>
<td>1990-1995</td>
<td>15.9</td>
<td>24.0</td>
<td>50.94</td>
<td>The UN did not intervene with PKO(s) in this conflict.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 The United Nations Mission for the Referendum in Western Sahara (MINURSO) is the longest ongoing PKO. Figures are from December 2019.
2 The United Nations Aouzou Strip Observer Group (UNASOG) was resolved the dispute between Chad and Libya over the Aouzou Strip after decades of fighting.
<table>
<thead>
<tr>
<th>State</th>
<th>Conflict</th>
<th>Period</th>
<th>Base Rate</th>
<th>Conflict Rate</th>
<th>% Chg. of Rate</th>
<th>UN Action</th>
<th>Budget (USD)</th>
<th>Military Strength</th>
<th>Civilian Strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>DJI</td>
<td>Djiboutian Civil War</td>
<td>1991-1994</td>
<td>11.7</td>
<td>10.2</td>
<td>-0.1282</td>
<td>The UN did not intervene with PKO(s) in this conflict.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DZA</td>
<td>Algerian Civil War</td>
<td>1991-2002</td>
<td>18.7</td>
<td>16.7</td>
<td>-10.7</td>
<td>The UN did not intervene with PKO(s) in this conflict.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDI</td>
<td>Burundian Civil War</td>
<td>1993-2005</td>
<td>35.3</td>
<td>64.2</td>
<td>81.87</td>
<td><strong>ONUB</strong> (2004-2006)</td>
<td>678.3mil</td>
<td>5.665/5650 max</td>
<td>885/1170 max</td>
</tr>
<tr>
<td>NAM</td>
<td>Caprivi Conflict</td>
<td>1994–1999</td>
<td>169.4</td>
<td>42.3</td>
<td>-75.0%</td>
<td>The UN did not intervene with PKO(s) in this conflict.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COD</td>
<td>1st Congo War</td>
<td>1996–1997</td>
<td>38.3</td>
<td>107.2</td>
<td>179.9</td>
<td>The UN did not intervene with PKO(s) in this conflict.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COD</td>
<td>2nd Congo War</td>
<td>1998–2003</td>
<td>38.3</td>
<td>112.0</td>
<td>192.4</td>
<td><strong>MONUC</strong> (1999–2010)</td>
<td>8.73bn</td>
<td>20586/22016 max</td>
<td>4397, no max</td>
</tr>
<tr>
<td>LBR</td>
<td>2nd Liberian Civil War</td>
<td>1999–2003</td>
<td>36.0</td>
<td>58.8</td>
<td>63.33</td>
<td>The UN did not intervene with PKO(s) in this conflict.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUD</td>
<td>War in Darfur</td>
<td>2003–2020</td>
<td>50.5</td>
<td>108.7</td>
<td>115.2</td>
<td>The UN did not intervene with PKO(s) in this conflict.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

³ The United Nations Mission in the Central African Republic and Chad (MINURCAT) operated simultaneously in both regions. ⁴ Figures for military strength are from February 2023, and civilian strength is from May 2018.
<table>
<thead>
<tr>
<th>State</th>
<th>Conflict</th>
<th>Period</th>
<th>Base Rate</th>
<th>Conflict Rate</th>
<th>% Chg. of Rate</th>
<th>UN Action</th>
<th>Budget (USD)</th>
<th>Military Strength</th>
<th>Civilian Strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBY</td>
<td>Post-Civil War Violence</td>
<td>2011-2014</td>
<td>33.0</td>
<td>133.0</td>
<td>303.0</td>
<td>The UN did not intervene with PKO(s) in this conflict.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDN</td>
<td>Sudan-South Sudan Border War</td>
<td>2012-2012</td>
<td>50.5</td>
<td>162.0</td>
<td>220.8</td>
<td>UNISFA(2011–)</td>
<td>Ongoing</td>
<td>3786, 3890 max</td>
<td>232, “as approp.”</td>
</tr>
<tr>
<td>CAF</td>
<td>Central African Republic Civil War</td>
<td>Requires collocations in NLTK, disrupting format</td>
<td>MINUSCA(2014–)</td>
<td>Ongoing</td>
<td>16363</td>
<td>1662</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSD</td>
<td>South Sudanese Civil War</td>
<td>2013-2020</td>
<td>50.5</td>
<td>88.0</td>
<td>74.26</td>
<td>UNMIS(2005-2011)</td>
<td>5.76bn</td>
<td>10352/10519 max</td>
<td>4099/4280 max 2876, “as approp.”</td>
</tr>
<tr>
<td>LBY</td>
<td>2nd Libyan Civil War</td>
<td>2014-2020</td>
<td>33.0</td>
<td>100.0</td>
<td>203.0</td>
<td>The UN did not intervene with PKO(s) in this conflict.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

UN data is historically well kept, especially with recent operations that include tallies of forces each year and even including a breakdown of which countries provide forces and how many. In some cases, we made exceptions based on the conditions of the data. In cases where there was a temporary increase in civilian strength for operations including overseeing elections, we included the temporary civilian personnel and indicated as such under Civilian Strength. Old operations where civilian strength is reported ‘unknown’ on the UN report are left out of regressions. Other operations only had approximate counts accurate to the hundreds, which we entered assuming the numbers were correct.

Notable exceptions to the observed trends are the UNASOG, which had a historically low expenditure of $64,471 in total supporting 9 armed and 6 civilian staffers. Another is Namibia, which up until its independence in 1990, dominated discussion in the United Nations. Namibia’s War for Independence had been going on and off since it became a League of Nations mandate overseen by South Africa following World War I. Discourse over Namibia reached upwards of a

---

\(^2\) Figures are from December 2019

\(^3\) Figures are from February 2023

\(^4\) The United Nations Multidimensional Integrated Stabilization Mission in Mali (MINUSMA) is coming to an end in July 2023.

\(^5\) The UN Mission in the Sudan (UNMIS) took place before the South Sudanese Civil War, and was the peacekeeping force sent to ensure the peaceful creation of South Sudan. This was then followed with the UN Mission in South Sudan (UNMISS), which operated during the South Sudanese Civil War.
rate of 580 at its peak over a two-decade span, making it the second-most mentioned country in the UN just below Israel. This is an example of when one conflict significantly undercuts a later one, in this case, the Caprivi Conflict.

4.2. An Interpretable Machine Learning Model for Intervention Recommendation

Given the clear correlation in Table III between the percent increase in yearly UN corpus mentions of a particular country and the presence of an active military conflict within the country, we create an interpretable framework for UN officials to leverage the presented analysis of their rhetoric to suggest data-driven intervention recommendations. In particular, we present a two-pronged approach that not only categorizes on-going conflicts according to the UN’s existing conflict classification system but also forecasts the required bulk budget, strength of deployed military troops, and strength of civilian personnel to address the conflict in accordance with the UN’s behavior historically.

To optimize the interpretability of our framework, we utilize a decision tree with a fixed maximal depth $d$, there are at most $1 + b + b^2 + \cdots + 2b^d = 4 \cdot b^{d+1} - 1$ nodes. Thus, for a tree with branching factor and decision tree, there are at most $d$ nodes, at most of which are leaf nodes. This significantly bounds the number of nodes in the tree to an extent that any user, such as a government official, can manually interpret the nodes and employ their own decision making.

Our framework uses the discourse of a given country to calculate the Base Rate and Conflict Rate:

1. **Base Rate**: The mean number of annual mentions across the temporal span of the UNGDC database.
2. **Conflict Rate**: The mean number of annual mentions within the target time period of active conflict. We alternatively coin this the *active rate*.

Both rates are helpful, as beyond just the percent change as calculated in Table III, the Base Rate and Conflict Rate provide context into the overall count of how many times a country is mentioned, providing context that helps the decision tree make the right classification.

4.2.1. A Decision Tree Classifier for Intervention Classification

The most important classification to determine is whether the UN decides to intervene based on the knowledge that we have about discourse rates. After we classify a binary yes/no in terms of UN intervention, we then classify the type of UN intervention based on Fortna’s four classifications of Peacekeeping Operations [36]:

1. **Observation Missions**, which are small contingents of usually unarmed military/civilian observers that monitor ceasefire agreements and conditions, observing and reporting on conditions to the UN. UNAVEM II in Angola is one example.
2. **Interpositional Missions**, which are traditional peacekeeping operations where large contingents of armed military forces serve as a buffer after two sides in a conflict. UNAVEM III in Angola is one example.
3. **Multidimensional Missions**, which are large operations carried out by military and police personnel that engage in various tasks outside of the historical understanding of PKOs, including supervising elections, building institutions, developing economies, and overseeing human rights reports. ONUMOZ in Mozambique is one example.
4. Peace Enforcement Missions are substantively different from the three prior missions. Each of these missions invoke Chapter VII of the UN Charter, allowing PKOs to operate in the region without the consent of either side. The prior three classifications, which invoke Chapter VI, depend on the consent of factions on either side to operate, which has led to some to be preemptively canceled when consent is revoked. These missions are authorized to use force beyond self-defense and take an active role in enforcing peace and protecting civilians. UNAMSIL in Sierra Leone is one example.

This is key to understanding UN interventions, as actions differ based on the classification.

Fig. 9. Decision trees for all four indicator variables, in the order: OBS, INT, MUL, ENF/VII.

Fig. 10. Decision tree classification for intervention types, in the same order as Figure 9.
Blue is N/N, Red is N/Y, Orange is Y/N, and Purple is Y/Y for Truth/Predicted Intervention. We introduce four binary indicator variables (OBS, INT, MUL, ENF/VII) corresponding to whether the UN has authorized an observation, interpositional, multidimensional, or enforcement PKO. We proceed to fit our decision tree using Gini impurity [37] as our splitting criterion, resulting in the following models shown in Figure 9 yielding the respective classifications shown in Figure 10. The classification accuracies of each decision tree are: OBS: 92.9% | INT: 92.9% | MUL: 89.3% | ENF: 75%

Given the high classification accuracy from the shallow tree, any UN official can predict the classification type to at least 75% accuracy within two questions regarding the base and conflict mention rate of a particular country.

4.2.2. A Regression for Quantitative Intervention Factors

Once an intervention is classified, we explore if discourse serves as a determining factor to see the overall expenditure, military strength, and civilian strength of a mission. These correlate heavily with the type of intervention; for example, an observer mission will have low-forces and subsequently low expenditures, while an enforcement mission will have expenditures in the billions and a large active military force with a similarly large civilian supplement. We study the following three quantitative metrics of a UN response:

1. The total expenditure of a PKO
2. The maximum military strength of a PKO throughout the mission
3. The maximum civilian strength of a PKO throughout the mission

Numbers like the expenditure and maximum strengths are useful quantifiable metrics of the size and scope of a UN intervention. Helpfully, the UN provides total expenditure and maximum strength of military and civilian forces for almost every prior PKO, but more expansive statistics such as the contribution of forces from each participating country is limited to more recent conflicts with better data and recordkeeping. Additional data, such as the difference between armed police, civilian police, and military officers are recorded for many, but are still not present in most PKOs. Thus, these three metrics serve as useful baselines which can explore the evolution and types of PKOs throughout the past 50 years.

We use mean squared error as our splitting criterion, and do not perform minimal cost-complexity pruning [37] given that the number of analyzed countries is within reason. Fitting our regression model on the historical PKO and UNGDC data detailed in Table III, we obtain the following regression predictions:

Fig. 11. Regression predictions for total mission expenditure (in millions USD).
As with classification, with a shallow tree of depth 2, this approach is able to reasonably recommend the maximum military/civilian strength that the UN should authorize given the base and conflict rate of a given country during a conflict. The United Nations Aouzou Strip Observer Group (UNASOG) is a notable outlier; the regression correctly predicts a much higher estimated force count in the thousands, but there were only 9 military troops and 6 civilian forces deployed. Other cases lie much closer to the regression, with a maximum deviation of approximately 2,000 in extreme cases.

5. Conclusion

While this work uses a representative span of UN discourse via the UNGDC, our framework provides an end-to-end solution for transforming discourse during an on-going conflict into actionable intervention recommendations and classifications using natural language processing and interpretable machine learning. After computational experimentation and statistical testing, we show to significance that the timeseries of mentions for African regions varies region to region based on the time period’s conflict landscape, suggesting that regions with differing mention dynamics may require differing intervention strategies. We identify multiple wars where discourse in the UNGD prerequisites intervention of UN actors and peacekeeping forces, especially in conflicts that lasted more than three years and with significant violence. In general, higher UN discourse in a particular conflict is correlated with the magnitude of the UN's crisis resolution response, and we aim to rigorously explore various response mechanisms in future work.

This correlation allows us to develop a framework which validates that, presented with any given conflict, a decision tree is accurate in predicting the classification of a Chapter VI UN Charter mission (observation, interpositional, and multidimensional) with a ~91.7% accuracy. Our decision tree to predict a Chapter VII UN Charter Mission (enforcement) is less accurate at 75%, indicating the difficulties of predicting when the UN decides to authorize use of force beyond self-defense in PKOs, which is understandable given the qualitative nuances required to authorize force in any given PKO. With these accuracies, we conclude that shallow decision trees are an effective, accurate, and interpretable method to use base and conflict discourse rates to not only determine not only if the UN will intervene, but also the type of a forecasted intervention. This is incredibly important to create qualitative analysis of UN effectiveness and forecast when and what types of aid will be provided by the UN during the conflict.

An important benefit of validating decision trees as an accurate predictor is that while alternative machine learning methods such as deep neural networks may achieve higher accuracies, shallow decision trees do not black box decision making from the policy maker. A fitted depth-two decision tree, such as our proposed classifier, reduces the intervention categorization problem to
an interpretable two-step question-and-answer process that provides accurate conclusions; in practice, computational diplomacy must be actionable by policymakers.

5.1. Future Work

Given the large amount of time series data being used, techniques such as multiple linear regression and Pearson correlation can not detect complex temporal correlations. Even if two random temporal variables are not directly correlated, there may be time-lagged cross-correlation during which changes in one variable correlate with the other after some amount of delay. Similarly, variables may temporally react in a non-Euclidean way detected by more advanced techniques such as dynamic time warping and instantaneous phase synchrony. Measuring autocovariance could also inform the prediction of long-term mention growth trends.

Given the breadth of the UNGDC dataset, we aim to expand our experiments to consider international relations, forecast economic growth, and network relations between countries. Other combinations of statistical analysis, NLP, and ML for speech analysis could be invaluable for computational approaches to diplomacy for other areas the UNGDC covers such as the UN’s Sustainable Development Goals.

To augment our recommendation framework, we aim to include other modes of data, including other text sources such as social media feeds from UN ambassadors on the UNSC and conventional media reporting on conflicts for more accurate and informed recommendations for modern PKOs. We believe that integrating other sources of data can create a novel quantitative algorithm and process to forecast and predict future conflicts alongside effective responses from UN PKOs alongside qualitative analysis.

Beyond the decision trees we created, we hope to use modern time-series classifications techniques that can provide UN intervention recommendations as a foreign conflict progresses across multiple years and perhaps finer-grained temporal resolutions; this is especially helpful to classify between interventions that often happen when a conflict concludes and other intervention types that happen during a conflict. Furthermore, we believe that modern time-series classifications can allow us to segment the interventions for countries such as Angola which had different interventions at different times, helping us gain a more nuanced understanding of each individual conflict for review.

6. Availability of Data and Material

Our work uses Python, R, and associated packages to analyze the open-source UNGDC dataset. Please contact authors for access to the implementation.

Acknowledgements

We would like to thank Professor Catherine Scott at the University of North Carolina Chapel Hill for her guidance.

References


