PeaceRep
Peace and Conflict Resolution Evidence Platform

GLOBAL TRANSITIONS SERIES

Global Realignments: Understanding Peace Process Interventions Through News

Roy Gardner and Christine Bell

The Global Transitions Series looks at fragmentations in the global order and how these impact peace and transition settlements. It explores why and how different third-party actors – state, intergovernmental, and non-governmental – intervene in conflicts, and how they see themselves contributing to reduction of conflict and risks of conflict relapse. The series critically assesses the growth and diversification of global and regional responses to contemporary conflicts. It also asks how local actors are navigating this multiplicity of mediators and peacebuilders and how this is shaping conflict outcomes and post-conflict governance.

This research is an output from PeaceRep - the Peace and Conflict Resolution Evidence Platform and has been funded by the UK Aid from the UK Foreign, Commonwealth & Development Office (FCDO). The information and views set out in this publication are those of the authors. Nothing herein constitutes the views of FCDO. Any use of this work should acknowledge the authors and the Peace and Conflict Resolution Evidence Platform.

Authors: Roy Gardner and Christine Bell
Data Engineering and Research Assistance: Niamh Henry, University of Edinburgh
Series Editor: Mateja Peter, University of St Andrews
Managing Editor: Mia Furlong, University of St Andrews
Design: Smith Design Agency

PeaceRep: The Peace and Conflict Resolution Evidence Platform
School of Law
Old College
The University of Edinburgh
South Bridge
Edinburgh
EH8 9YL

Tel. +44 (0)131 651 4566
Fax. +44 (0)131 650 2005
E-mail: peacerep@ed.ac.uk
PeaceRep.org
Twitter: @Peace_Rep_
Facebook: https://www.facebook.com/PeaceRepResearch
LinkedIn: https://www.linkedin.com/company/peacerep/
Instagram: https://www.instagram.com/peace_rep_/

About the Authors:
Roy Gardner is a consultant specialising in data analysis and natural language processing. Christine Bell is Assistant Principal (Global Justice) and Professor of Constitutional Law, University of Edinburgh, and Director of PeaceRep.

Cover images: All images may be subject to copyright. Getty Images. ©2022

Contents

Key Findings and Recommendations 01

Introduction 04

Methodology 07

Syria Case Study 22

Sudan Case Study 38

UN Case Study 47

Discussion 50

Key findings and recommendations

This report describes the methodology and findings from analysis of in-country and global media perceptions of the interventions of international actors in peace-related activities. The purpose of the study was to explore whether media attention relating to peacemaking or building initiatives, was shifting in terms of the global players involved, and the positive or negative sentiment with which those interventions were viewed.

The research described in this report uses GDELT – a Google-based database of coded global news stories – as a source of data. GDELT has codes for news stories that cover cooperative, non-conflict peacemaking or peacebuilding activities. Data on the number of such stories, together with a measure of the sentiment of the stories, make GDELT potentially valuable in assessing the way that news sources perceive interventions related to peace processes.

The report describes several case studies on media perceptions of global influence in the peacemaking/peacebuilding space and assesses the reliability of GDELT for this type of analysis.

The case studies explore perceptions of (1) the US, Russia, and the UN in Syria since 2015, (2) the US, Russia, the UN, and additional interveners in Sudan after the uprisings, (3) the UN in a range of countries.

Case Studies

Syria

Our results clearly show that in Syria, UN peacemaking and peacebuilding had the most negative sentiment over time. This was true across English and Arabic news sources, and across news sources both inside and outside Syria. Whilst further research is required to determine if this news perception indicates a poor opinion of the UN, it does suggest that mentions of the UN occur in stories that are broadly negative. This would not be surprising since during this period, attempts by the UN to deliver a viable peace process failed.

There was evidence that perceptions of Russian interventions in Syria became less negative over time. This was also true of the US, but perceptions of the US were overall more negative than perceptions of Russia over the same period. This indicates that the Russian move to draw the peace process into a constitution-making process may have been perceived by the media as a success.

Interestingly, when US news sources (of which there are a disproportionate number) were excluded from the analysis of Syria, the difference in sentiment of mentions of Russian and US government actors was reduced. This indicates that US news sources are likely to be more critical of the US than non-US news sources.

The results also showed, unsurprisingly perhaps, that media interest in conflict resolution wanes over time.

Sudan

In Sudan, we examined post-uprising news stories relating to peacemaking and peacebuilding. We selected a range of Inter-Governmental Organisations (IGOs) and a range of Western and non-Western countries for analysis.

The data show that there were no stories related to the interventions by the Gulf Cooperation Council (GCC). The perceptions of the UN were much more negative than those of the African Union, a finding that confirms scholarly assessment that neighbours and regional organisations are more positively viewed than global organisations or non-regional actors (see e.g., Carothers and Samet-Marram, 2015).

When aggregated across all news sources, there were more stories about interventions by non-Western countries than there were about interventions by Western countries. Overall, stories about non-Western countries were more positive.

UN

To further investigate the negative sentiment of stories about the UN in Syria we compared stories about the UN in a range of countries in which the UN operates. There are striking differences in sentiment which probably reflect the type of UN operation in these countries and the perception of these operations. The comparison confirmed the Syrian result, with Syria showing the most negative sentiment followed by Libya and South Sudan. Stories about UN operations in Cyprus had the most positive sentiment.

Reliability of GDELT

Questions remain regarding the quality and possible bias of GDELT data. Like many current machine-learning and AI systems, GDELT is a "black box", and can only be assessed by measuring its behaviour in various benchmarking tasks.

One issue of concern is the number of duplicate mentions which imply over-sampling of news sources by GDELT. We have mitigated this to some extent, but some duplication still exists.

Whilst there are clear biases in GDELT towards US and English-language news sources, our methodology is designed to be able to control for these biases.

Recommendations

- Two principal recommendations are made to build confidence in GDELT. The first is that GDELT is benchmarked by: 1) comparing its coding of selected news stories to the coding of the same stories by human coders, 2) assessing whether trends and differences discovered in our analysis are considered credible by country and subject experts. Other benchmarks are listed in the Discussion.

- The second recommendation is that news sources are assessed for quality and for type, for example, state/non-state, corporate/independent. These types would provide valuable additional comparison dimensions.

Introduction

The Global Database of Events, Language and Tone (GDELT) (Leetaru, K. and Schrodt, P.A., 2013) is a machine-coded database hosted on Google's BigQuery cloud infrastructure. The database contains geo-referenced records derived from the classification of news stories by version 1.1b3 of the Conflict and Mediation Event Observations (CAMEO) (Gerner, D.J. et al., 2002, Schrodt, P.A. et al., 2008) actor and event classification scheme.

A GDELT event is a date-stamped database record containing the classification under the CAMEO scheme, of the content of a news story found by GDELT's search algorithm. Machine-learning and natural language processing algorithms predict actor data, country data, CAMEO event codes etc. from the content of the news story.

A GDELT event is also stored in the database as a GDELT mention. Mentions extend event data to include confidence, sentiment, and information related to the country and language of a news source. Mentions are also created if the original story is covered by other news sources. Because they give us insight into the volume, sentiment, and confidence of coverage, mentions are the unit of analysis in this report.

GDELT does not record mentions before the beginning of 2015, which coincides with the introduction of GDELT version 2. Whilst it is possible to analyse GDELT events before 2015, these events are missing important information about the news source's country and language, and have cruder sentiment measures. Furthermore, without mentions it is impossible to measure the volume of news stories about real-world events on a given day.

Our research focused on whether - and how- GDELT could be a useful source of information for comparing the volume and sentiment of web-based news stories covering peacemaking interventions by non-traditional actors such as Russia and China, with interventions by what we might consider to be normative peacemakers and peacebuilders, such as the United Nations.

The case studies reported here use time series analysis (Shumway, Robert H. and Stoffer, David S., 2017) and visualisation to compare interventions in countries of interest. We have developed a methodology that enables us to address several research questions:

- Is there an association between weight of news stories and interventions, and what do these interventions relate to?
- Are interventions perceived positively or negatively by global, regional, and national media?
- Do these perceptions change depending on language source analysed?
- How do perceptions of emergent power interventions compare with perceptions of interventions by what we might loosely term Western powers (e.g., United States) and IGOs with normative approaches to peace and transition processes (e.g., African Union or the United Nations)?
- Are there trends in engagement and can our methodology be used for long-term trend mapping?

Table 1 provides a definition of some of the terms used in this report.

Term	Definition
Event	This is the classification under the CAMEO scheme of the content of a news story found by GDELT's search algorithm. Machine-learning and natural language processing algorithms predict actor data, country data, event codes etc. from the content of the news story.
Mention	This is the classification of the coverage of an event by other news sources. The original event is also recorded as a mention.
Actor	An actor is a country-affiliated or non-country affiliated (e.g., an IGO) entity that carries out some action in a country. Actors of interest studied in this report include emergent global and regional powers, Western powers, and various IGOs.

Term	Definition
Actor type	Defines the type or role of an actor of interest. For example, the type GOV refers to government country-affiliated actors, and MIL to military country-affiliated actors.
Action	Defines the activity or intervention of an actor in a country. Actions are defined by CAMEO event codes.
Country of interest	Countries of interest are those in which actors of interest are involved in conflict and peace processes. For example, Syria, Sudan, Yemen.
Confidence	Refers to GDELT's confidence in its classification of a mention's news source content. Defined as a percentage. We use confidence to weight mentions, sentiment, and Goldstein Scale scores.
Sentiment	Ranks the emotion expressed by the content classified by a mention.
Goldstein Scale (Goldstein, J.S. 1992)	A system used to rate the severity of a CAMEO event code in terms of harm done by the event. Events leading to the loss of life are highly negative, cooperative activities are positive.
Weight of news stories	Measured as a mentions score - a weighted sum of daily mentions.
Perception	How an actor is perceived by news media in terms of weight and sentiment of news stories.

Table 1: Definitions of terms used in this report

Methodology

We have developed a methodology for the statistical analysis of GDELT data that takes the results of SQL queries run against the GDELT database, and processes these results to generate data in CSV format. Processed data are used to construct multivariate time series. These time series are analysed to detect trends and to compare the weight and sentiment of news stories along dimensions that include actors, news source language, and country of interest.

Correlations and differences in the weight and sentiment of news stories mentioning various actors in countries of interest can therefore be detected in time series. In some cases, there are clear trends in the weight and sentiment of news stories over time. The weight and sentiment of mentions is assumed to reflect the way an actor is perceived by news media.

The methodology provides a basis for comparing emergent powers with regional powers, Western powers, and IGOs in terms of the perception of interventions by these actors in countries of interest. The perception of a single actor in several countries of interest can also be compared. Comparisons can be refined to filter mentions by news source country and news source language. These provide additional dimensions along which the perception of actors can be compared. Data can be de-biased by removing selected news sources, e.g., US news sources.

Of interest too is the absence of correlation in some cases, e.g., significant changes in the number of mentions may not necessarily correlate with significant changes in sentiment and vice versa.

The methodology is implemented in a set of Jupyter Notebooks using the Python programming language. Various outputs are generated in either CSV or JSON format for use by other researchers.

Data Acquisition and Processing

SQL Queries

The GDELT database can be interrogated by formulating a query. Each query must specify the actors, the country or countries of interest, and the CAMEO events of interest over a particular date range.

A query specifies values in the Events database columns listed in Table 2. At a minimum, values must be specified for the SQLDATE and ActionGeo_CountryCode fields.

Column	Description
SQLDATE	Used to define date range. Dates are specified in YYYYMMDD format, e.g., 20161011.
Actor1Code	CAMEO code used to specify actors with no country affiliation, e.g., IGOs.
Actor1CountryCode	A three-letter CAMEO country code that specifies the country affiliation of an actor, e.g., RUS for Russia.
Actor1Type1Code	CAMEO code defining the type of an actor, e.g., GOV (government), MIL (military).
ActionGeo_CountryCode	A two-letter FIPS country code that defines a country of interest.
EventRootCode	CAMEO numeric event code that defines the action of an actor in a country of interest.

Table 2: GDELT Events table columns used in SQL queries

Column	Column	Description
Mentions in the date range 1st Jan 2015 to 31st Dec 2021	SQLDATE	>= 20150101 <=20211231
Mentions referring to UN	Actor1Code	IGOUNO
Mentions referring to Russia and USA	Actor1CountryCode	RUS,USA
Mentions referring to Russian and US government actors	Actor1Type1Code	GOV
Mentions referring to Syria and Libya	ActionGeo_CountryCode	SY,LY
Mentions classified under cooperative, non-military CAMEO codes	EventRootCode	03,04,05,06,07

Table 3: Values used to construct a GDELT database query (see above).

Table 2: GDELT Events table columns used in SQL queries

Data Processing

The functions of the data processing stage are to:

- Remove duplicate mentions which would distort the results by double counting.
- Create fields containing calculated values obtained from lookup dictionaries, e.g., using a news source's domain to determine the source's country.
- Optionally combine the results of several queries into a single dataset to increase the scope of time series analysis.
- Save a processed dataset to a CSV file.

After data processing, each mention is stored as a row in a CSV file. The row fields are defined in Table 4.

Field	Description
Globaleventid	Unique identifier of the mention's event.
SQLDATE	Date of mention in YYYYMMDD format.
Actor1Code	CAMEO actor code. Used for non-country affiliated actors.
Actor1CountryCode	Three-letter CAMEO country code for country affiliated actors.
Actor1Type1Code	CAMEO actor type.
ActionGeo_CountryCode	Two-letter FIPS country code defining a country of interest
EventCode	The fine-grained CAMEO event code of the mention.

Field	Description
SOURCEURL	The URL of the source of the mention's event news story.
MentionSourceName	The domain of a mention's news source. Used to look up the value of the SourceCountry column.
Confidence	Percentage value defining GDELT's confidence in the classification of a mention's news source content. Used to weight metrics in subsequent analysis.
MentionDocTone	Sentiment measure of the emotion of a mention's news source content in the range -100 to 100.
MentionDocTranslationInfo	Provides information about the mention's source language. Used to look up the value of the SourceLanguage column.
RootCode	CAMEO event root code obtained from the value of the EventCode column.
Goldstein	Goldstein Scale score obtained from the value of the EventCode column in the range -10 to 10.
SourceCountry	A two-letter country code defining the country of a news source. Obtained from the value in the MentionSourceName column.
SourceLanguage	News source language obtained from the value in the MentionDocTranslationInfo column. Currently, a proper name, e.g., English.

Table 4: List of data fields of mention records in CSV files generated by data processing.

Time Series

We explore mentions over time using time series analysis. A time series is a series of data points indexed in time order at a given resolution, e.g., day, month, year.

All mentions have a timestamp that defines the day of the mention. It follows that mentions can be arranged in day order to create a time series with a resolution of days within a given date range. Note that there may be days in the range without any mentions. Our analysis also uses aggregation of day data to generate time series with resolutions of months and years. It is also possible to define and analyse a subrange of a time series including a single day.

The three metrics used to construct time series from the processed query data are described below. The default period is a day.

1. The *weighted mentions score* m_t for a period t is the sum of mentions in the period where each mention is weighted by the value of the mention's Confidence field:

$$m_t = \sum_{i=1}^{N} \frac{c_{ti}}{100}$$

where:

N is the number of mentions in period t

c_{ti} is the value of the Confidence field of the i^{th} mention in period t

Weighting reduces the contribution of mentions that are unreliable.

2. The *mean weighted sentiment score* s_t for a period t is: $$s_t = \frac{1}{N}\sum_{i=1}^{N} \sigma_{ti} \frac{c_{ti}}{100}$$

where:

σ_{ti} is the value of the MentionDocTone field of the i^{th} mention in period t

Weighting reduces the contribution of mentions that are unreliable. The mean is calculated to remove the effect of different numbers of mention across days.

3. The *mean weighted Goldstein Scale score* g_t for a period t is given by:

$$g_t = \frac{1}{N}\sum_{i=1}^{N} \gamma_{ti} \frac{c_{ti}}{100}$$

where:

γ_{ti} is the value of the Goldstein field of the i^{th} mention in period t

Weighting reduces the contribution of mentions that are unreliable. The mean is calculated to remove the effect of different numbers of mention across days.

Goldstein Scale scores are included as a control. Firstly, Goldstein Scale scores would be expected to be positive for news stories coded by cooperative, non-conflict CAMEO event codes. Secondly, Goldstein Scale should be negatively correlated with sentiment. Both assumptions are confirmed by case study data.

For each of the metrics above, the following three types of time series are generated.

1. Score series. Score series provide the basic data for measuring the volume, sentiment, and Goldstein scores of news stories over time.

$$\mathrm{M}_{score} = \{m_t\}_{t=1}^{N} \qquad \mathrm{S}_{score} = \{s_t\}_{t=1}^{N}$$

$$\mathrm{G}_{score} = \{g_t\}_{t=1}^{N}$$

Figure 1 shows sample score series for mentions, mean sentiment, and mean Goldstein score. The data come from the Syrian Case Study. The x-axes contain days in a 7-year period from the beginning of 2015. Year boundaries are marked as grey vertical lines. Each day in the mentions series contains the weighted mentions score for that day. Clear peaks in the volume of mentions on certain days can be seen especially in 2015 and 2016 presumably reflecting media interest in real-world events on these days. Each day in the sentiment series contains the mean of the weighted sentiments of the day's mentions. Using the mean controls for the effect of differences in the number of mentions across days and makes it easier to detect significant positive and negative peaks in sentiment. Sentiment scores are generally in the range -10 to +10 and most sentiment scores are below zero reflecting a prevailing negative sentiment in mentions. The Goldstein series contains the mean of the weighted Goldstein Scale scores of the day's mentions.
All Goldstein scores are above zero which is consistent with the use of cooperative CAMEO codes in the Syrian Case Study.

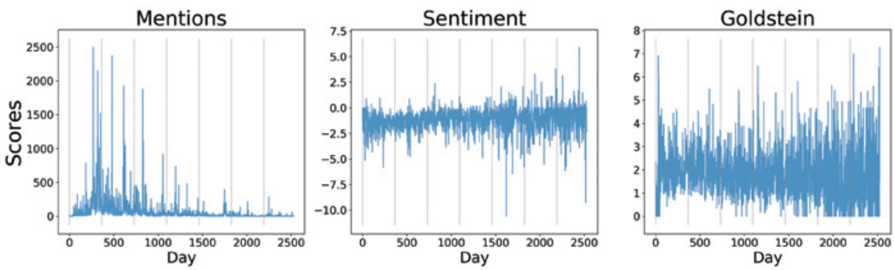


Figure 1: Sample mentions, mean sentiment, and mean Goldstein score series.

2. Moving average series. Moving averages are used to reveal trends that might not otherwise be visible in a score series.

$$\mathrm{M}_{ma} = \left\{ \frac{1}{n} \sum_{j=t}^{t+n-1} m_j \right\}_{t=1}^{N}$$

$$\mathrm{S}_{ma} = \left\{ \frac{1}{n} \sum_{j=t}^{t+n-1} s_j \right\}_{t=1}^{N}$$

$$\mathrm{G}_{ma} = \left\{ \frac{1}{n} \sum_{j=t}^{t+n-1} g_j \right\}_{t=1}^{N}$$

where:

n is the length of the moving average window

Moving average series may provide a better assessment and comparison of changes in volume and sentiment of news stories over time.

Figure 2 shows moving average series derived from the mentions, mean sentiment, and mean Goldstein score series in Figure 1. The moving average window width is 200 days. Trends in all series are clearly visible and there are significant negative correlations between the mentions and sentiment series, and the sentiment and Goldstein series. The trends show that mentions decrease over time from a peak at the end of 2015. In contrast, sentiment becomes less negative over the same period, and Goldstein score show a downward trend indicating that stories towards the end of the period are coded with less positive CAMEO codes.

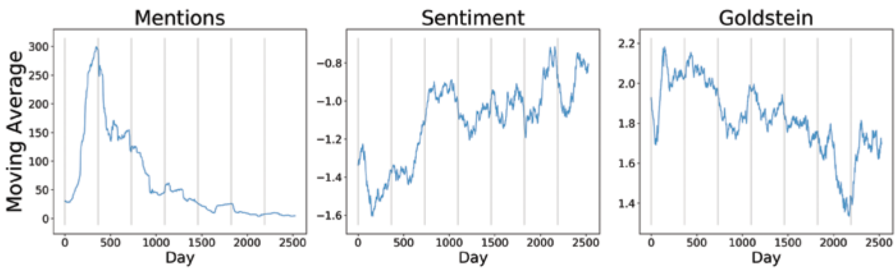


Figure 2: Sample moving average series.

3. Difference series. Differencing is used to remove trends. This may seem counter-intuitive given that moving averages are used to find trends. However, it is important to remove trend-based correlations before looking for score-based correlations. Difference series are also useful for detecting large shifts in score from positive to negative or vice versa.

$$\mathrm{M}_{diff} = \{m_{t+1} - m_t\}_{t=1}^{N}$$

$$\mathrm{S}_{diff} = \{s_{t+1} - s_t\}_{t=1}^{N}$$

$$G_{diff} = \{g_{t+1} - g_t\}_{t=1}^{N}$$

Figure 3 shows difference series derived from the mentions, mean sentiment, and mean Goldstein score series in Figure 1. Difference series are used to enable the measurement of score-based correlations by removing trends. For example, in Figure 2 there is a significant negative correlation between mentions and sentiment moving average time series. However, this correlation disappears once the trends are removed by the difference series. In other words, the correlation in Figure 2 is a result of trends and not a day-to-day correlation between mentions and sentiment scores. This is illustrated by the large shift in sentiment close to day 1700 which does not coincide with a significant feature in the mentions time series.

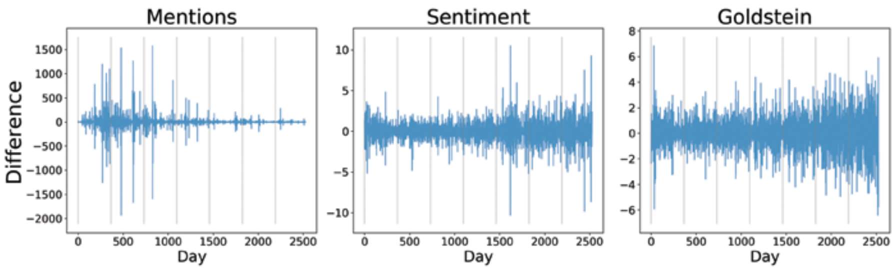


Figure 3: Sample difference series.

Altogether we have nine time series shown in Table 5 below.

	Mentions	Sentiment	Goldstein Scale
Score	M_{score}	S_{score}	G_{score}
Moving average	M_{ma}	S_{ma}	G_{ma}
Difference	M_{diff}	S_{diff}	G_{diff}

Table 5: Types of time series derived from a set of GDELT data.

As will be seen later, this three-by-three grid representation is reproduced in dashboard-type visualisations of time series.

The set of nine time series is represented by a matrix *TS* comprising nine time series row vectors. Matrix columns are time periods, and matrix rows comprise the time series.

Pairs of time series from the matrix can be compared to answer research questions. For example, we can ask whether M_{ma} and S_{ma} are correlated. The statistical comparison of time series is discussed in detail in the next section.

Comparing Time Series

To answer research questions relating to changing perceptions of different country and international organisation involvement in peacebuilding, it is necessary to compare time series along various dimensions. For example, actor country, actor type, country of interest etc. To do this we introduce a variable that contains values of a comparison dimension, and for each value we create a time series matrix.

For example, where we acquired mentions about various actors in Syria and wanted to compare mentions of Russian government actors with mentions of US government actors, we generated two time series matrices; one containing the set of time series generated from mentions of Russia, and the other containing the time series generated from mentions of the USA.

Field	Values
Actor1CountryCode	RUS
Actor1Type1Code	GOV
ActionGeo_CountryCode	SY
EventRootCode	03,04,05,06,07

Table 6: Field values for selecting mentions about Russia government actors in Syria.

To obtain the US time series we selected mentions based on the field values in Table 7.

Field	Values
Actor1CountryCode	USA
Actor1Type1Code	GOV
ActionGeo_CountryCode	SY
EventRootCode	03,04,05,06,07

Table 7: Field values for selecting mentions about US government actors in Syria.

In this example, the field Actor1CountryCode is the comparison dimension, and the comparison variable contains two values – RUS and USA. This provides two matrices: TS_{Rus} and TS_{USA}. This stack of two matrices is referred to as a tensor.

Tensors provide opportunities to compare time series across matrices and therefore answer our research questions. For example, we can measure the correlation and difference of Russian and US S_{ma} time series to explore trends related to the sentiment of news stories that mention Russian and US actors in Syria.

By expanding the number of values in our comparison variable, we increase the number of matrices in the tensor and provide more comparisons. For example, if we obtain mentions of UN actions in the countries in which the UN operates, then these countries are the values of the comparison variable, and we can compare the perception of the UN in different countries. This is the subject of the UN Case Study.

Query Language

To enable time series comparisons, a simple query language was developed with which to specify and label the comparison dimension and define the values of the comparison variable.

Table 8 illustrates three queries for comparing actors of interest in Syria. In this example, the comparison dimension is the actor defined either by the actor country for Russia and the USA, or by the actor code in the case of the UN. Each query is given a unique label used to identify the query's time-series matrix in the tensor which enables us to compare the three actors.

	Query 1	Query 2	Query 3
LABEL	**RUS GOV**	**USA GOV**	**UN**
country	SY	SY	SY
actorcode			IGOUNO
actorcountry	RUS	USA	
actortype	GOV	GOV	
eventroot	03,04,05,06,07	03,04,05,06,07	03,04,05,06,07
source			
language			

Table 8: Illustration of a comparison query designed to compare the non-military actions of Russia, the USA, and the UN in Syria.

The queries specified in Table 8 generate a tensor comprising three time-series matrices, TS_{Russia}, TS_{USA}, and TS_{UN}.

Values can be excluded from a query by prefixing the value with a ! symbol. For example, if the language field in the query above contained the value !English, mentions by English-language news sources would be excluded from the time series analysis.

Statistical Analysis

This section discusses the statistical methods used to compare time series. We use non-parametric tests because the distributions of mention scores, mean sentiment scores, and mean Goldstein Scale scores are non-normal according to the Shapiro-Wilk test.

We use Spearman's rank correlation coefficient to measure the correlation between two time series. For example, we can determine the degree to which the same time series type from two actors are correlated.

The correlation between time series of different types can also be measured. For example, the correlation between mention scores and mean sentiment scores can be measured to determine the extent to which the volume of news stories and the sentiment of those stories are correlated.

To measure the difference between two time series we use the Wilcoxon signed-rank test.

Probability density functions (PDFs) are used to summarise differences between comparison variables in the score time series M_{score}, S_{score}, G_{score}.

Event Marking and Detection

We found it useful to determine the alignment of real-world events with peaks and changes in time series. Event marking is used to display the date of a known real-world event on a time series. An event is marked as a pair of dots at an x-axis position corresponding to the day of the event. Event marking also provides evidence for the reliability and accuracy of GDELT data.

Event detection finds peaks and large changes in time series. Such peaks and changes in mention scores and sentiment may correspond with one or more known significant real-world events. Difference time series help identify dates on which significant changes in sentiment are not correlated with high mentions scores. For example, we can detect real-word events that were not significant in terms of the volume of reportage, but were very significant in altering the sentiment with which an international actor was perceived.

Syria Case Study

The objective of this case study was to explore changes over time of the perceptions of Russian and US government actors, and the UN, in Syria over an almost seven-year period from 01/01/2015 – 30/11/21. We were also interested in differences in perception of the various actors. We selected mentions with CAMEO root codes 03,04,05,06, and 07 which refer to cooperative non-conflict actions all of which are associated with positive Goldstein Scale scores. Various comparisons were made:

- Perceptions of actors by all news sources
- Perceptions of actors with US news sources removed
- Perceptions of actors by Arabic-language news sources
- Perceptions of actors by Syrian news sources

Comparison 1: Russian and US government actors and the UN

The comparison query is defined in Table 8. All news sources are included in the comparison.

Figure 4 shows a dashboard view of the complete set of time series for this comparison as well as the PDFs of the score series. The first column of the grid displays the time series for mentions, the second column the time series for sentiment, and the final column the time series for Goldstein. The first row contains the scores series, the second the moving average series, and the third the difference series. The last row contains the PDFs. The grey vertical lines mark the first day of each year.

Some clear trends are visible:

- The mentions score of all actors decreases over time from peak values in 2015 and 2016. This trend is visible in every time series of the first column.
- After an initial fall through 2015 the sentiment of mentions rises for all actors.
- After an initial increase through 2015, the Goldstein score of mentions decreases for all actors. Goldstein scores are negatively correlated with sentiment.
- The sentiment PDF shows that mentions about Russia are the least negative.

Figure 4: Dashboard graphic showing time series and PDFs of Comparison 1

Some significant findings are:

1. Sentiments of mentions of Russian government actors are less negative than those of US government actors and the UN.
2. Sentiments of mentions of the UN are the most negative compared to Russia and the US.
3. The correlation of sentiments between actors is trend-based not score-based.
4. Mentions and sentiment moving averages are negatively correlated. However, this does not imply a causal relationship between volume of mentions and sentiment since the correlations are absent from the difference series.
5. Mentions of US government actors predominate in the first two years of the time series.
6. All Goldstein scores are positive, consistent with the use of non-conflict CAMEO event root codes in the query.

Event Marking

Figure 5 illustrates event marking in the mention scores time series using markers for two historic events. The first event on 30/09/2015 is the start of Russian military engagement in Syria which coincides with a peak in mention scores of US government actors perhaps reflecting the reaction of the US to Russian military engagement. The second event on 22/11/2017 is the Sochi conference which coincides with a peak in mention scores of Russian government actors.

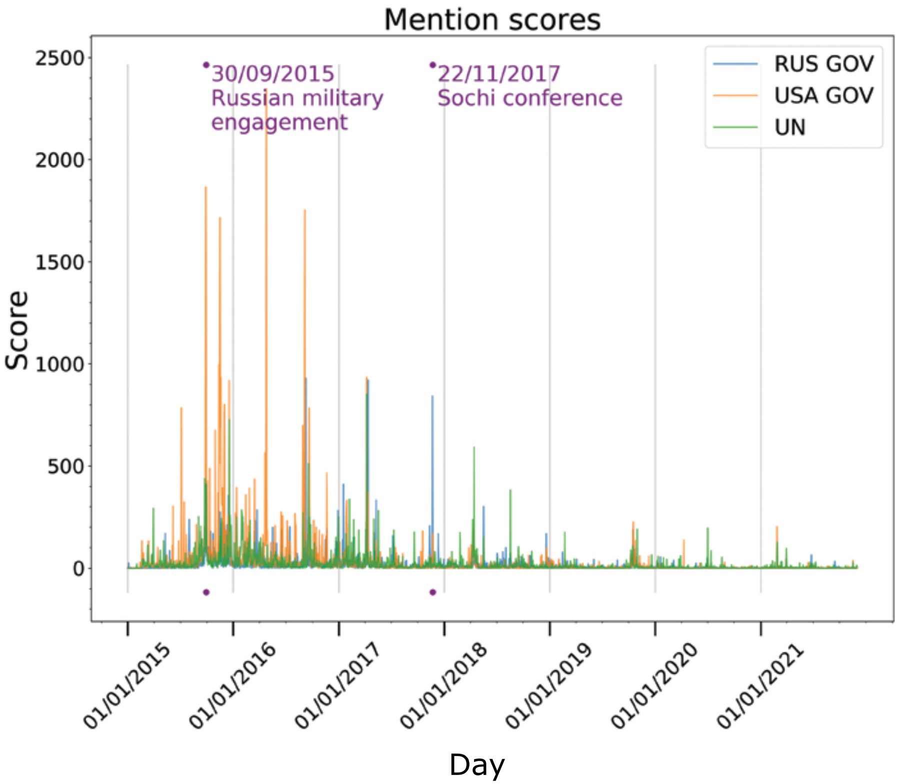


Figure 5: Mentions score time series for the actors of Comparison 1 with markers for two key dates in the Syrian conflict.

Figure 6 below illustrates event detection using the sentiment different series. There was a decrease in the sentiment of mentions about the UN on the 28/12/2019. This is the date of the displacement of thousands of Syrians from Idlib province during an offensive by Syrian and Russian forces.

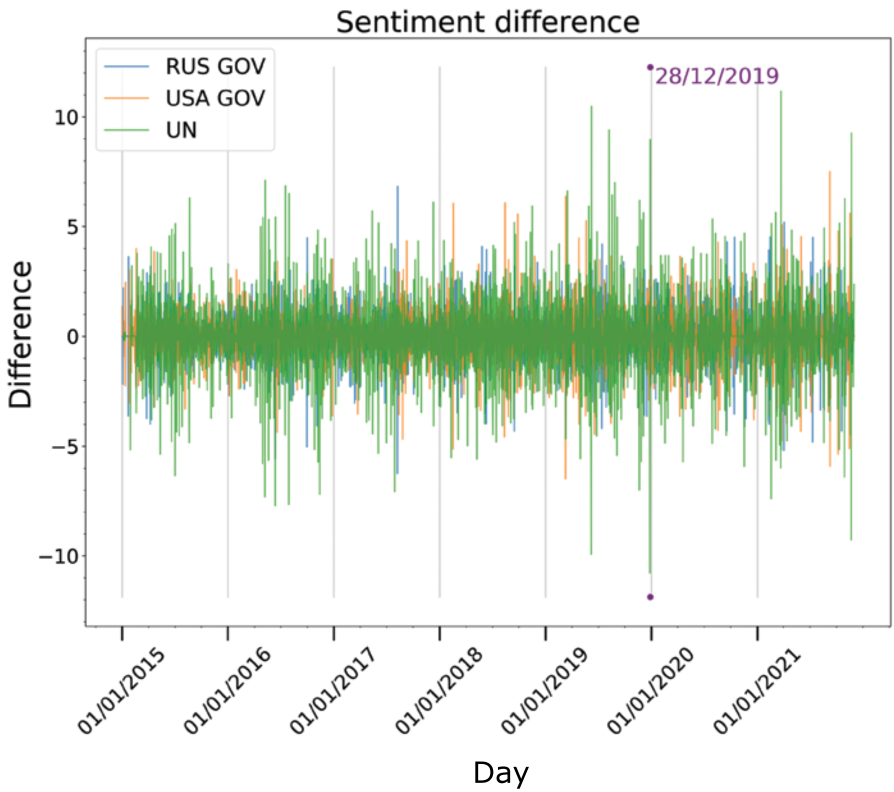


Figure 6: Event detection in the UN sentiment difference time series.

Comparison 2: US news sources removed

This is the same query as Comparison 1 above except for the exclusion of US news sources. The comparison query is defined in Table 9. Because there are many more US news sources than sources from other countries (see Discussion), the objective here is to remove the influence of US news sources from the comparison.

	Query 1	Query 2	Query 3
LABEL	**RUS GOV**	**USA GOV**	**UN**
country	SY	SY	SY
actorcode			IGOUNO
actorcountry	RUS	USA	
actortype	GOV	GOV	
eventroot	03,04,05,06,07	03,04,05,06,07	03,04,05,06,07
source	!US	!US	!US

Table 9: Comparison 2 query field values.

Figure 7: Dashboard graphic showing time series and PDFs of Comparison 2

The basic patterns follow those of Comparison 1 apart from a significant reduction in mentions of US government actors. However, the difference in sentiment of mentions of Russian and US government actors is reduced which implies that US news sources may have seen US involvement in Syria in a more negative light than non-US news sources. This is also illustrated in the sentiment PDF.

Event Detection

Figure 8 shows event detection in the US mentions score time series. On 25/04/2019 the US announced an increased military presence in Syria. The data show that this event registered significantly in non-US news stories.

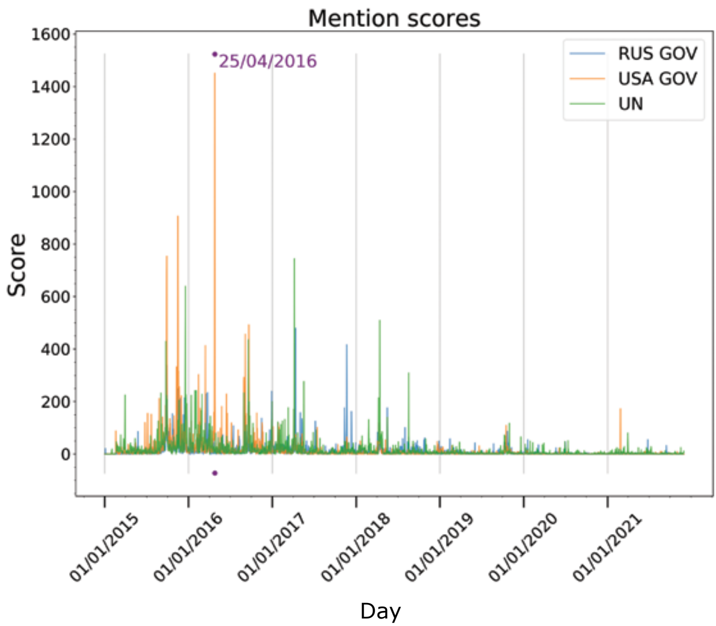


Figure 8: Event detection in the US government mentions score time series.

Comparison 3: Arabic-language news sources only

This is the same query as Comparison 1 and 2 above except only Arabic-language news sources are included. The comparison query is defined in Table 10.

	Query 1	Query 2	Query 3
LABEL	**RUS GOV**	**USA GOV**	**UN**
country	SY	SY	SY
actorcode			IGOUNO
actorcountry	RUS	USA	
actortype	GOV	GOV	
eventroot	03,04,05,06,07	03,04,05,06,07	03,04,05,06,07
language	Arabic	Arabic	Arabic

Table 10: Comparison 3 query field values.

Some significant findings are:

1. The trends and differences of the previous two comparisons are visible in this comparison.
2. The overall mention scores are significantly lower reflecting the predominance of English-language news sources in GDELT (see Discussion).
3. Most mentions are about the UN.

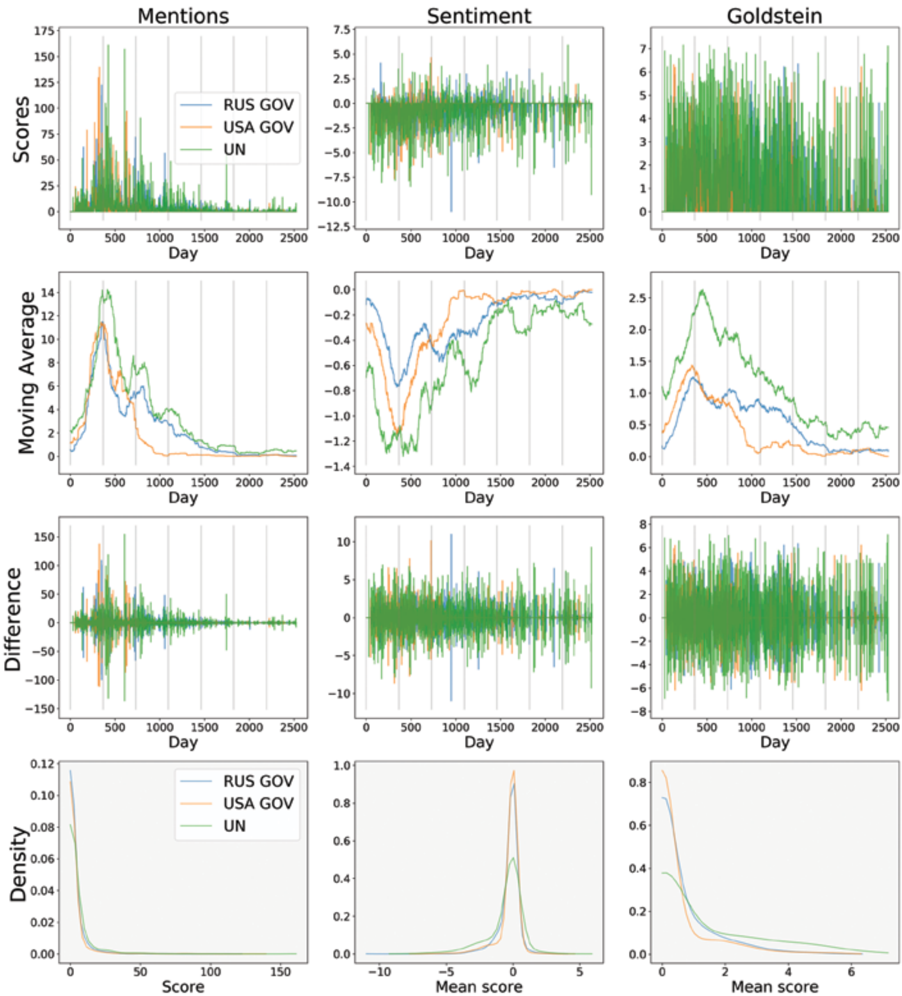


Figure 9: Dashboard graphic showing time series and PDFs of Comparison 3

Event Detection

Figure 10 shows event detection in the UN mentions score time series. On 29/02/2016 the UN announced a funding pledge for families in Syria and for Syrian refugees in Jordan, Lebanon, Iraq, and Egypt. This event is not detected in the previous two comparisons indicating that it is more significant for Arabic-language news sources than non-Arabic language news sources.

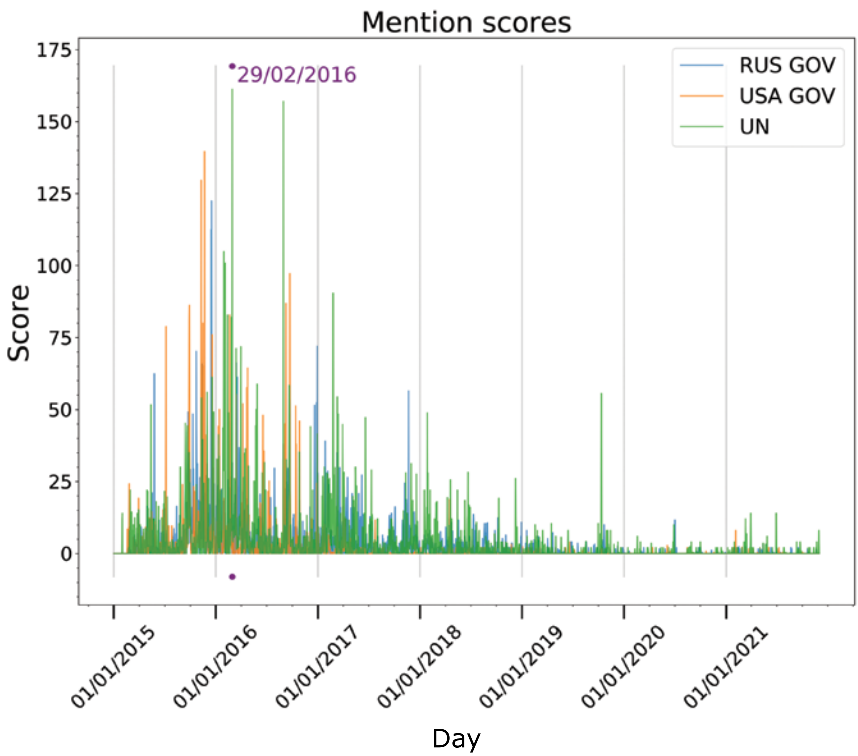


Figure 10: Event detection in the UN mentions score time series.

Comparison 4: Syrian news sources only

Once again this is the same query as Comparison 1 and 2 above except that only Syrian news sources are included. The comparison query is defined in Table 11.

	Query 1	Query 2	Query 3
LABEL	**RUS GOV**	**USA GOV**	**UN**
country	SY	SY	SY
actorcode			IGOUNO
actorcountry	RUS	USA	
actortype	GOV	GOV	
eventroot	03,04,05,06,07	03,04,05,06,07	03,04,05,06,07
source	SY	SY	SY

Table 11: Comparison 4 query field values.

Again, the trends are consistent with the comparisons above. There is a marked reduction in the overall mention scores consistent with GDELT's bias towards US and English-language news sources. There are more mentions of the UN than of Russian and US actors, comparable with the finding for Arabic-language news sources.

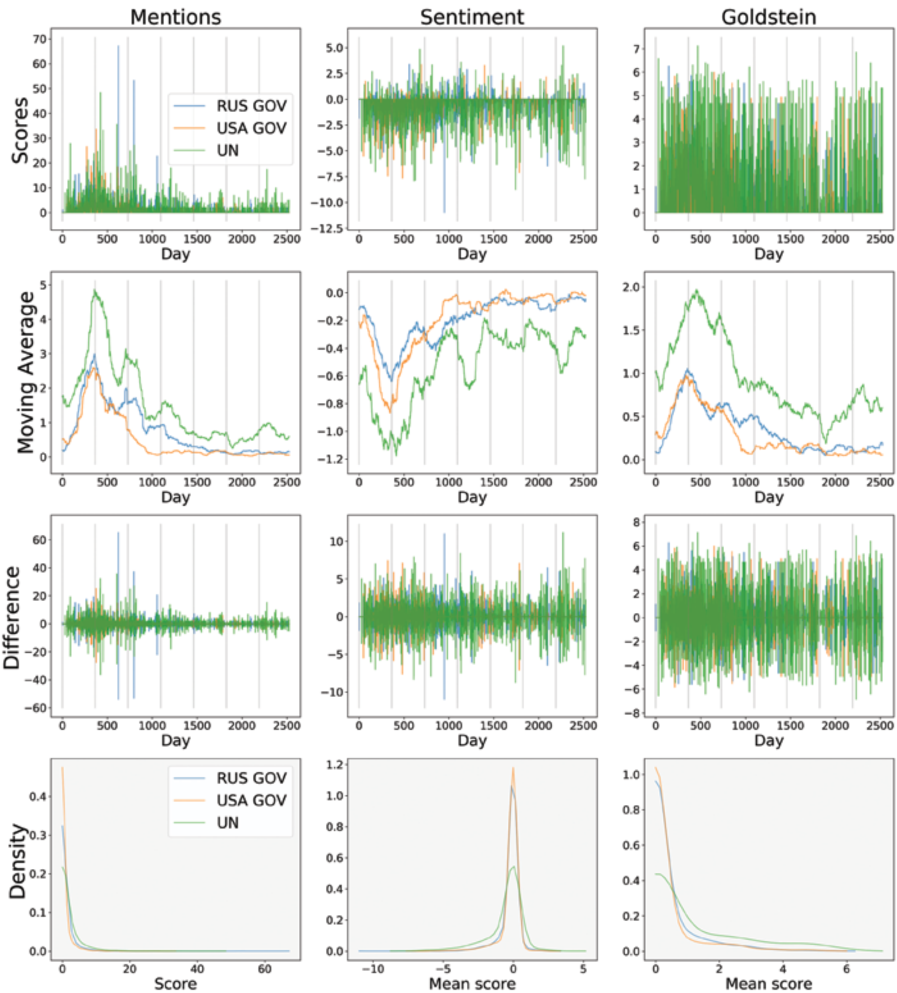


Figure 11: Dashboard graphic showing time series and PDFs of Comparison 4.

Event Detection

Figure 12 shows event detection in the Russian mentions score time series at a prominent peak in mid-September 2016. On 16/09/2016 Russian military chiefs visited Ankara to discuss military cooperation and Russian bombers based in Iran targeted Islamist militants in Syria. This is further support for the assertion that the relevance of an event is country or language dependent.

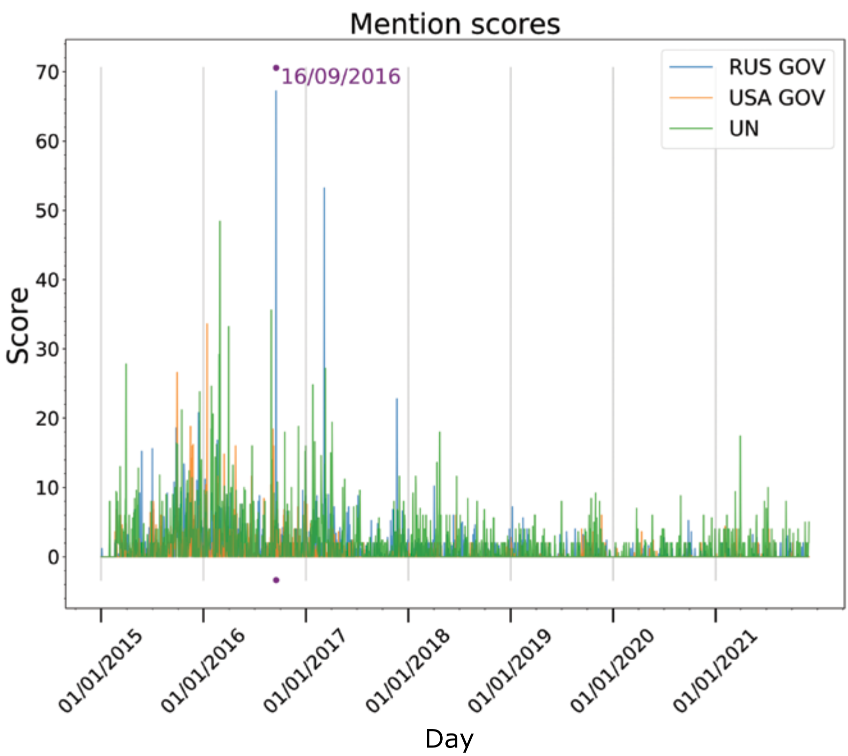


Figure 12: Event detection in the Russian mentions score time series.

Conclusions

Perhaps the most surprising finding of this case study is that the sentiment of mentions of the UN are consistently the most negative even when compared to military actors.

There are two other consistent findings. Firstly, Syria is no longer in the news to the extent it was in 2015 – 2016, and secondly, the sentiment of stories has become less negative as mentions have decreased. As discussed in Comparison 1, once time series data are de-trended, sentiment is not correlated with mention scores. This suggests that the perception of the situation in Syria is truly less negative than it was in the early years of the period studied, and not just a consequence of fewer mentions. This may reflect stabilization of the country contingent on Russia's intervention.

Sudan Case Study

The date range for Sudan comparisons is a three-year period from 01/01/2019 – 19/01/2022.

Comparison 1: Comparing IGOs

This comparison compares time series of IGOs in Sudan for all news sources, actor types, and all CAMEO root codes (non-conflict and conflict). The comparison query is defined in Table 12.

	Query 1	Query 2	Query 3	Query 3
LABEL	**UN**	**African Union**	**Gulf Cooperation Council**	**IGAD**
country	SU	SU	SU	SU
actorcode	IGOUNO	IGOAFRAFU	IGOGCC	IGOIAD

Table 12: Query field values for the Comparison 1 query.

Some significant findings are:

1. There are no mentions of the Gulf Cooperation Council and IGAD, indicating that any intervention did not register at a media level.
2. The mention scores of the UN increases significantly in the second half of 2021. Since the Juba Agreement was signed late October and then developed and implemented, this increase may reflect a period of relative success, which in turn affected perceptions of the UN.
3. The sentiment of mentions of the UN is significantly more negative than the sentiment of mentions of the African Union over the whole period. This bears out scholarly assessment that neighbours and regional organisations are perhaps viewed more positively than global organisations or non-regional actors.
4. There are negative Goldstein Scale scores consistent with the use of all CAMEO root codes. This indicates that stories about the activities of IGOs are not always coded with cooperative CAMEO codes.

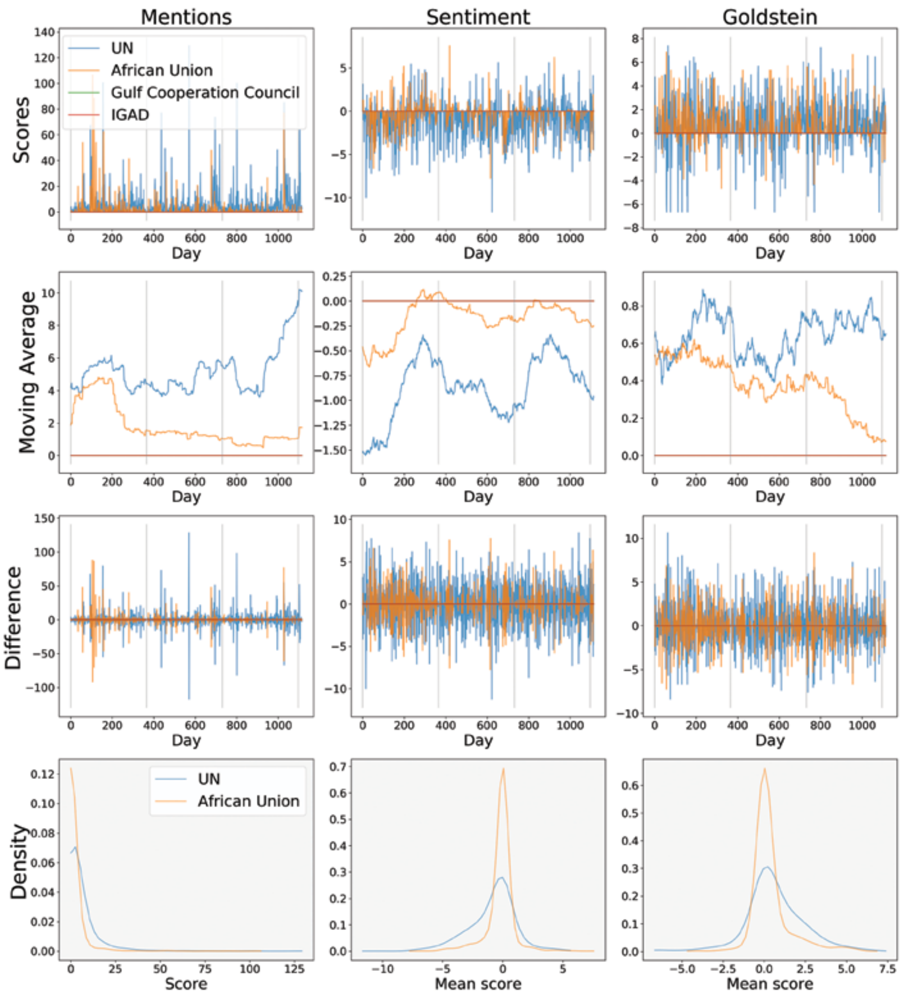


Figure 13: Dashboard graphic showing time series and PDFs of Comparison 1.

Comparison 2: Western and non-Western actors

This comparison compares time series of what we refer to as Western government actors (USA, UK, and Norway) with those we refer to as non-Western government actors (China, Russia, Ethiopia, Qatar, United Arab Emirates, and South Sudan), in relation to Sudan. All news sources are included, and cooperative CAMEO root codes are specified. The comparison query is defined in Table 13.

	Query 1	Query 2
LABEL	**non-Western**	**Western**
country	SU	SU
actorcountry	CHN,RUS,ETH,QAT,ARE,SSD	USA,GBR,NOR
actortype	GOV	GOV
eventroot	03-07	03-07

Table 13: Query field values for Comparison 2.

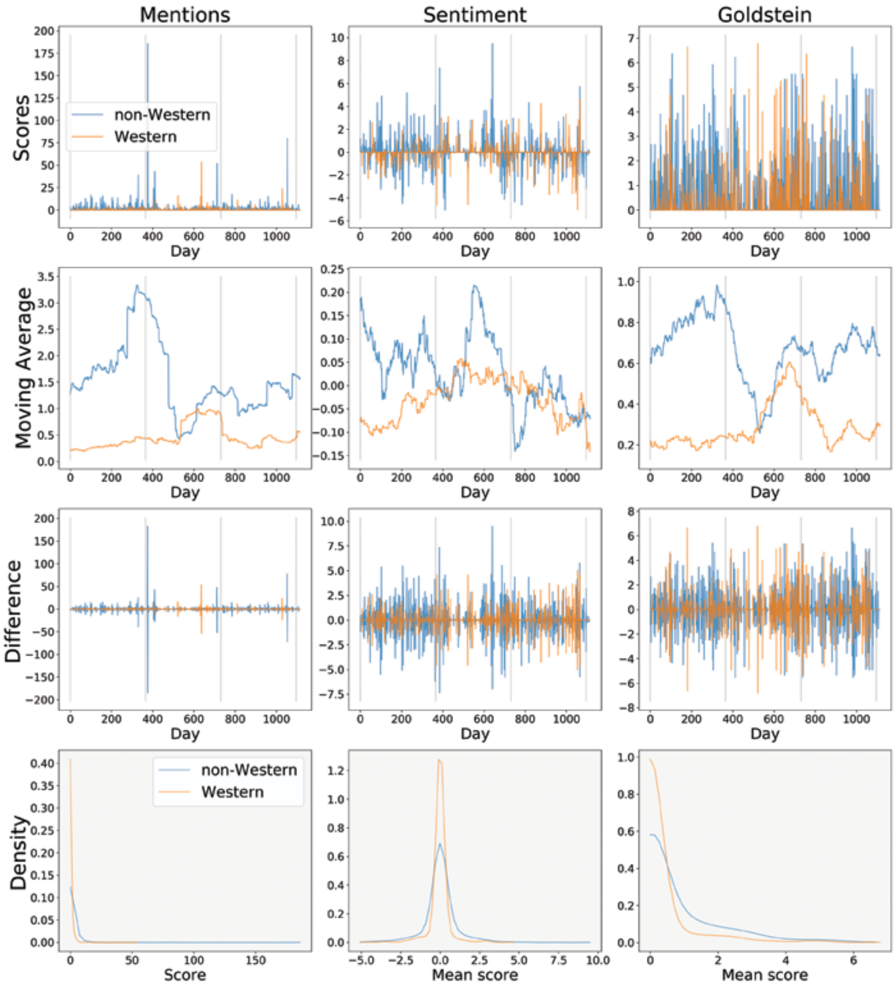


Figure 14: Dashboard graphic showing time series and PDFs of Comparison 2.

Event Marking

Figure 15 marks three key events. Whilst the signing of the Juba Agreement and the coup get peaks for Western actors, there are no significant peaks at the date President al-Bashir was deposed. Figure 16 shows a significant positive sentiment peak for mentions of non-Western actors on the date of the signing of the Juba Agreement.

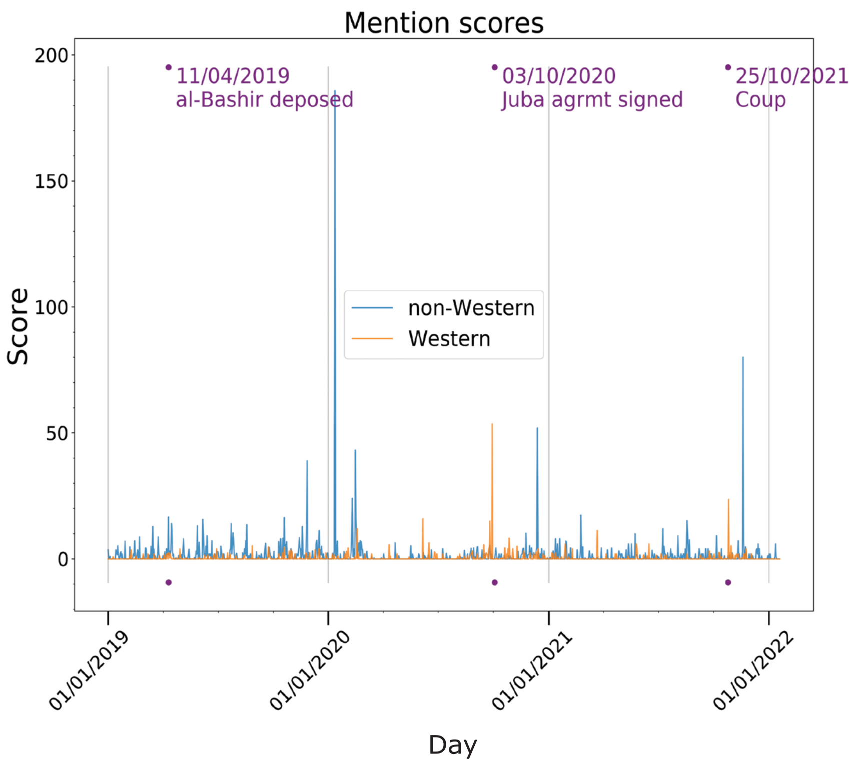


Figure 15: Event marking in mention scores time series.

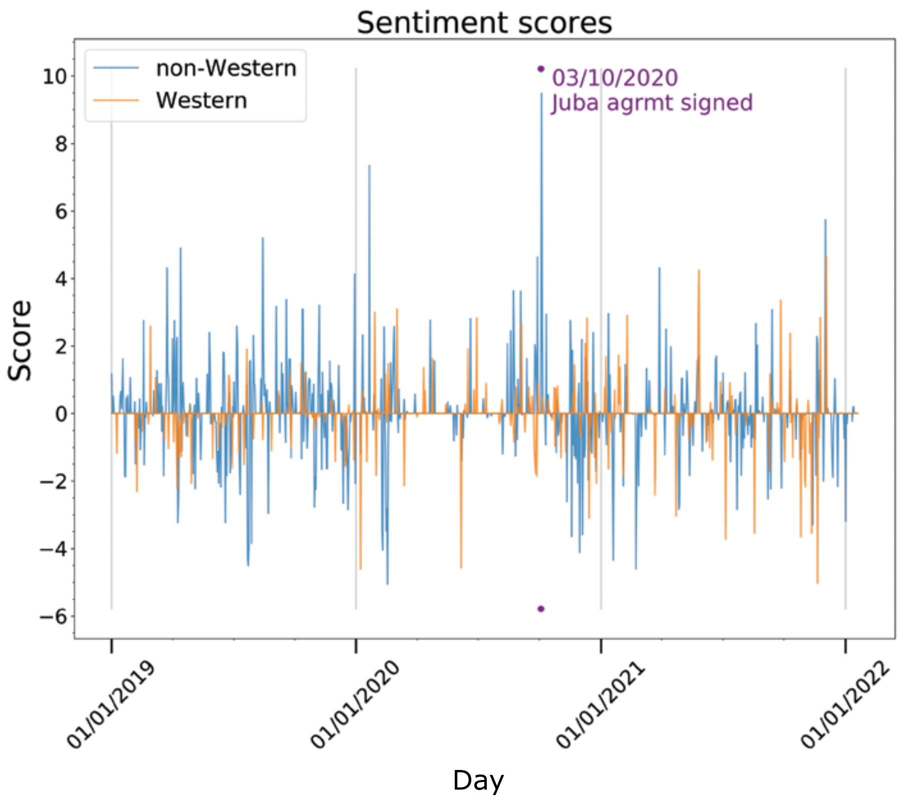


Figure 16: Event marking in sentiment scores time series.

Event Detection

Figure 17 illustrates the detection of two events in the mention difference time series for non-Western actors. The event of 18/11/2021 corresponds to the date of violent post-coup protests. No known events can be determined for the date of 11/01/2020 which is also visible in the Comparison 3 time series.

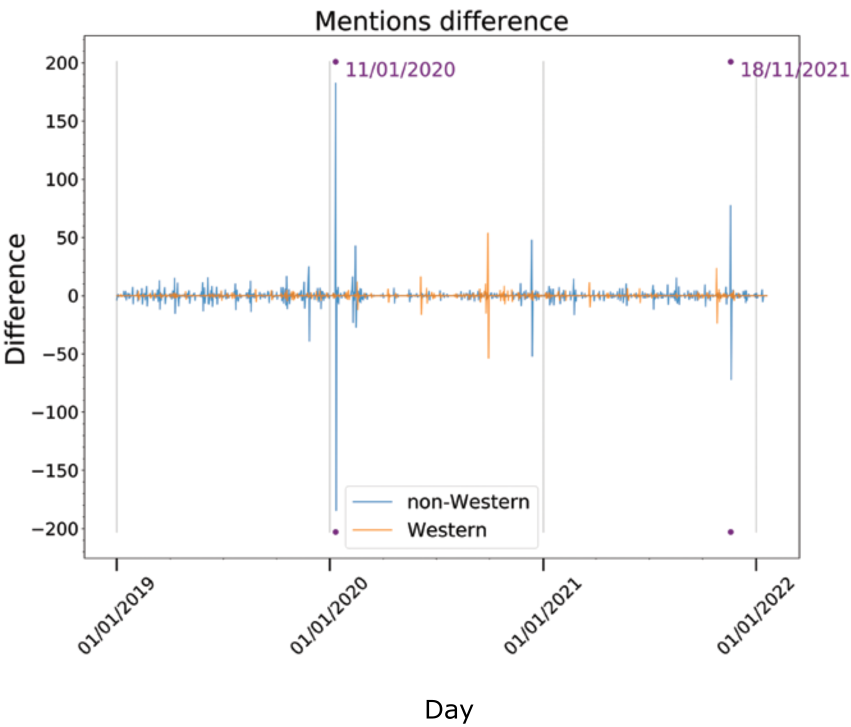


Figure 17: Event detection in the non-Western mention difference time series.

Comparison 3: Arabic-language and English-language news sources

Here we compare Arabic-language and English-language news sources for government actors from all countries. IGOs are excluded from the analysis. CAMEO codes are non-conflict codes in the range 03-07.

	Query 1	Query 2
LABEL	**Arabic**	**English**
country	SU	SU
actorcountry	CHN,RUS,ETH,QAT,ARE,SSD, USA,GBR,NOR	CHN,RUS,ETH,QAT,ARE,SSD, USA,GBR,NOR
actortype	GOV	GOV
eventroot	03-07	03-07
language	Arabic	English

Table 14: Query field values for Comparison 3.

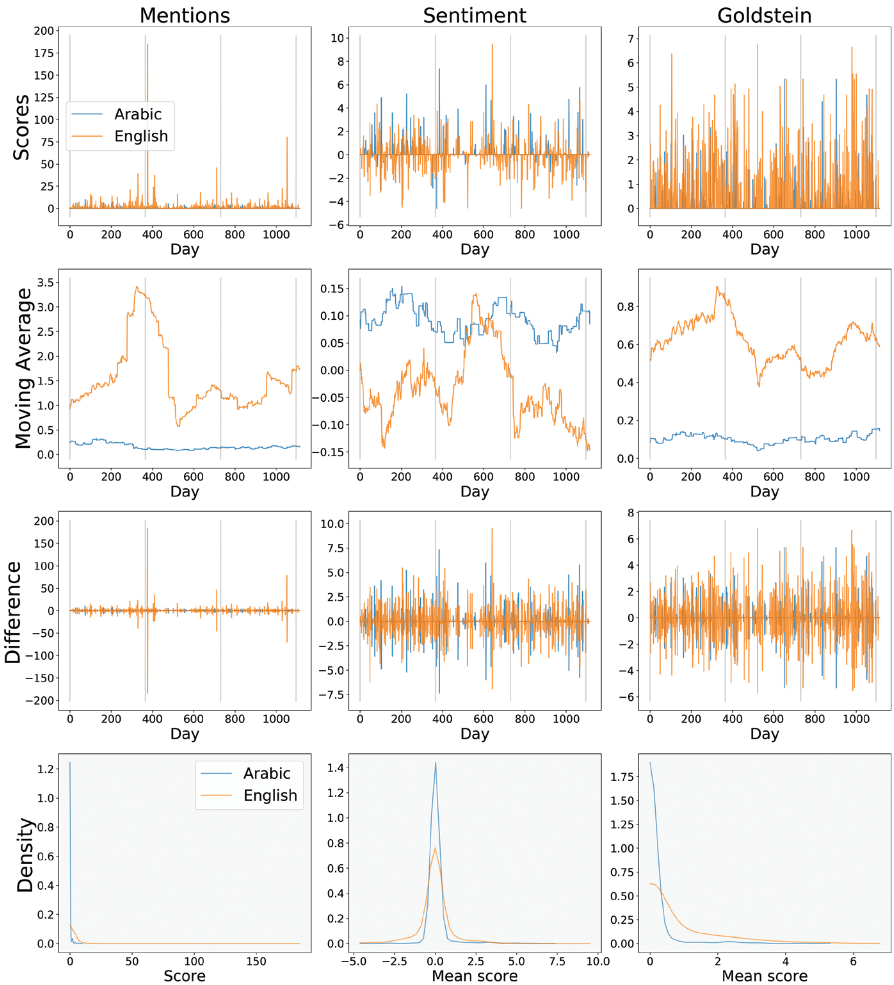


Figure 18: Dashboard graphic showing time series and PDFs of Comparison 3.

English-language news sources dominate in terms of mention scores. The sentiment of English-language mentions is more negative than the sentiment of Arabic-language mentions for all actors. This is clearly reflected in the sentiment PDF.

UN Case Study

Given the negative sentiment scores attached to the UN in Syria, we compared the sentiment of stories about the UN in several countries. All CAMEO codes (conflict and non-conflict) were included. We chose a set of countries where the UN had very different types and scales of engagement.

Comparison: Countries in which the UN operates

The date range for this comparison is a seven-year period from 01/01/2015 – 30/11/21. The time series of countries in which the UN operates are compared. These countries are listed in Table 15.

Query	LABEL	country
1	DRC	CG
2	CAR	CT
3	South Sudan	OD
4	Mali	ML
5	Western Sahara	WI
6	Kosovo	KV
7	Cyprus	CY
8	Lebanon	LE
9	Syria	SY
10	Libya	LY

Table 15: Query field values for the UN case study comparison. The table is transposed compared to other case studies.

These data illustrate that there are significant differences between countries both in terms of mentions scores and sentiment. Figure 20 shows the sentiment moving averages in more detail. Interestingly, the marked negative sentiment of mentions of the UN in the Syria Case Study is confirmed by comparison with other countries as no merely an artefact of GDELT. For example, the sentiment of stories about UN operations in Cyprus are significantly more positive than the sentiment of stories about the UN in Syria.

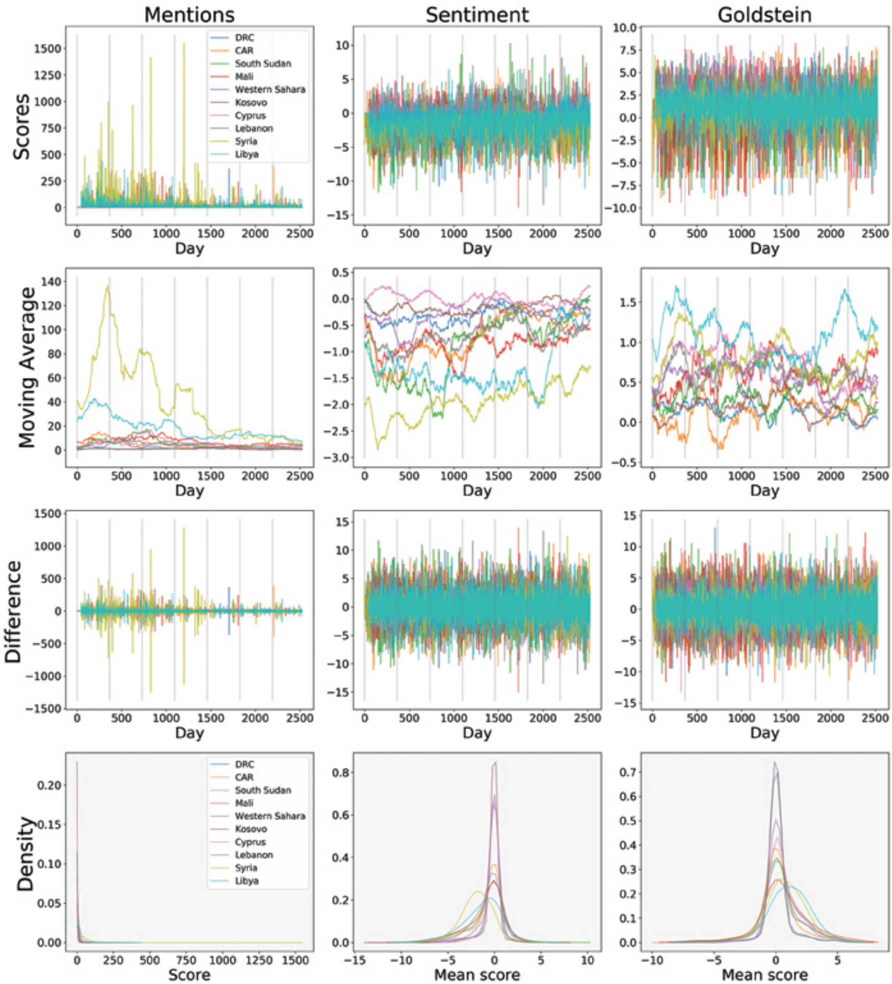


Figure 19: Dashboard graphic showing time series and PDFs of the comparison.

There are negative Goldstein Scale scores consistent with the use of all CAMEO root codes. This indicates that stories about the activities of UN are not always coded with cooperative CAMEO codes.

The trends show a marked increase in sentiment for South Sudan in mid-2017, and a similar increase in Libya from mid-2020.

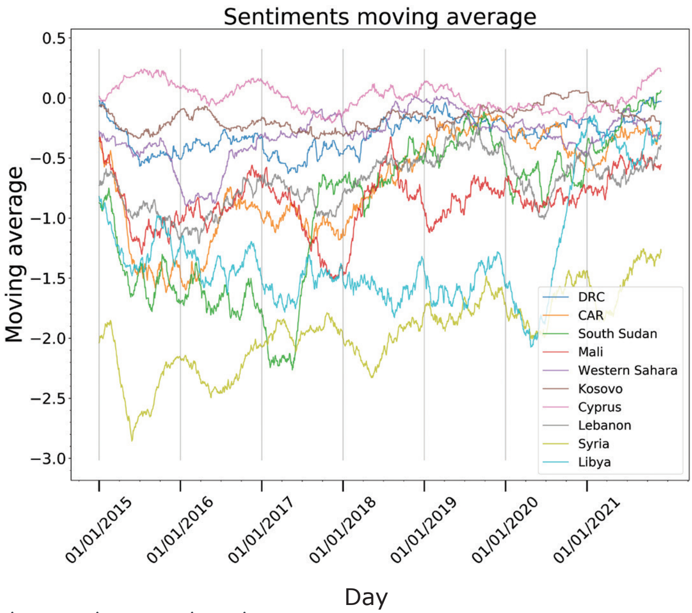


Figure 20: Sentiments moving average time series.

Conclusions

There are clear differences in mentions of the UN between the countries in which it operates. These differences are seen in mention scores, sentiment and Goldstein Scale score and may reflect the type of operations the UN conducts in these countries.

Discussion

Each of the research questions presented in the Introduction is discussed below.

Is there an association between weight of news stories and interventions, and what do these interventions relate to?

The Case Studies provide evidence that significant real-world events related to interventions coincide with peaks in mentions scores.

Are interventions perceived positively or negatively by global and regional media?

Significant changes in mentions score do not necessarily correlate with significant changes in sentiment and vice versa. That is, the amount of media attention an intervention gets may not be related to how positively or negatively that event is perceived. This is illustrated by Figure 17 from the Sudan Case Study. The figure shows a clear sentiment peak for mentions of non-Western actors on the date of the Juba Peace Agreement. No clear peak is visible in the non-Western mentions score time series in Figure 15.

Positive sentiment is a rarity in GDELT data. Figure 21 shows the distribution of sentiment scores for Syrian Case Study data. Whilst the peak at zero appears significant, it only accounts for 3.2% of all mentions, whereas 82% of mentions have a negative sentiment value. In other words, issues with GDELT aside, interpretation of relative sentiment scores must consider that all peacebuilding interventions are perceived negatively by the media.

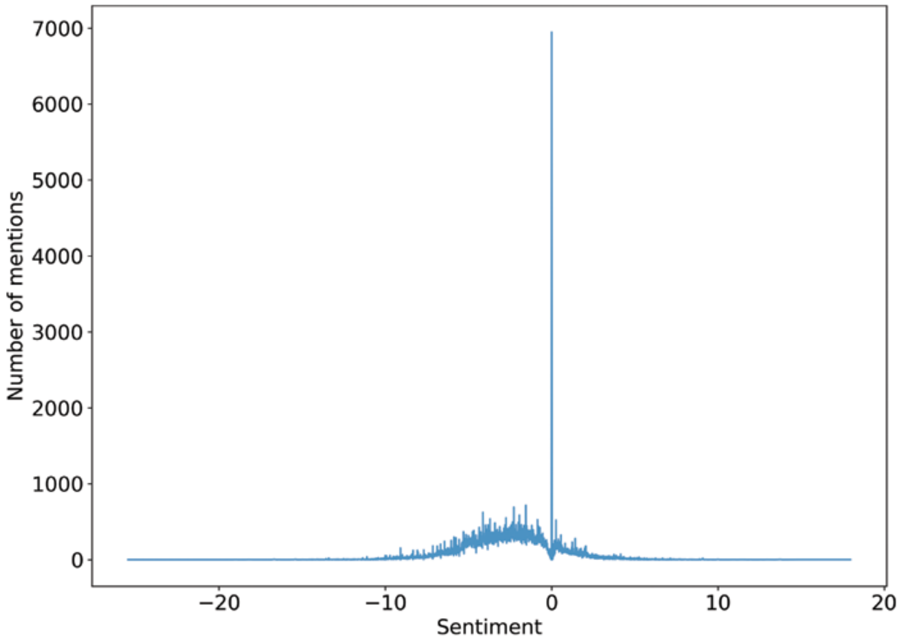


Figure 21: Distribution of the sentiment scores represented as the number of mentions per sentiment score value.

How do perceptions of emergent power interventions compare with interventions by Western powers and IGOs?

Significant differences in the weight and sentiment of news stories about emergent powers, Western powers, and IGOs can be found. For example, in terms of sentiment, the perception of Russian government intervention in Syria is less negative than that of the US and the UN. Perceptions of the UN are the most negative. This might be because the UN has had the largest formal role in peacemaking/peacebuilding, but has had little success.

Are there trends in engagement and can our methodology be used for long-term trend mapping?

Clear trends can be detected in all three types of time series. Historical trend analysis is therefore supported by the methodology. Time series forecasting which extrapolates trends is worth investigating.

Quality of GDELT Data

In this section we present our assessment of the reliability of GDELT for this type of research, and define what must be taken into consideration when analyzing search results.

Duplication of mentions

We have found evidence for duplication of mentions not all of which can be handled at the data processing stage. For example, the same mention - including the mention of the original event - may be duplicated with differences in sentiment and confidence, something which is difficult to deal with algorithmically without introducing an arbitrary judgement.

Duplication implies over-sampling at the mentions level, but the effect of this over-sampling on time series may be small given that over 60% of events only have a single mention. Further investigation is required.

Bias in GDELT data

There is a clear bias towards US English-language news. This is illustrated by data from the Syrian Case Study above which contains 219698 mentions.

Figure 22 shows that 42% of mentions are derived from the content of US news sources. Given this percentage, it is not surprising that 62% of mentions are derived from English-language news sources as shown in Figure 23.

This bias raises questions about the extent to which local and regional news sources are represented in GDELT data. Hammond and Weidmann (Hammond, J. and Weidmann, N.B., 2014) were critical of GDELT's capabilities at the subnational level:

'Our findings indicate that GDELT should be used with caution for geo-spatial analyses at the subnational level: its overall correlation with hand-coded data is mediocre, and at the local level major issues of geographic bias exist in how events are reported'

Figure 22: Number of mentions by news source country.

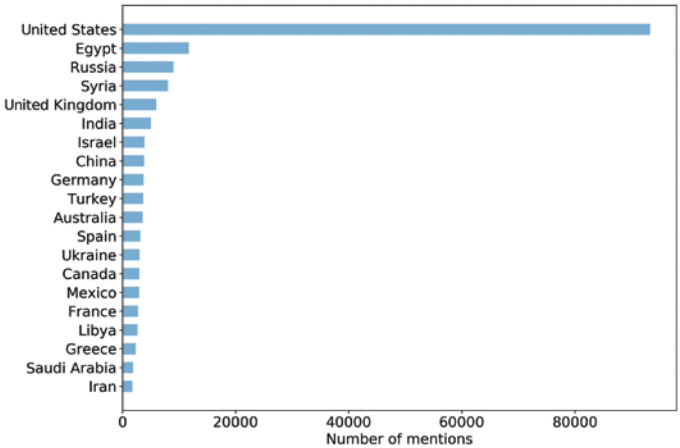
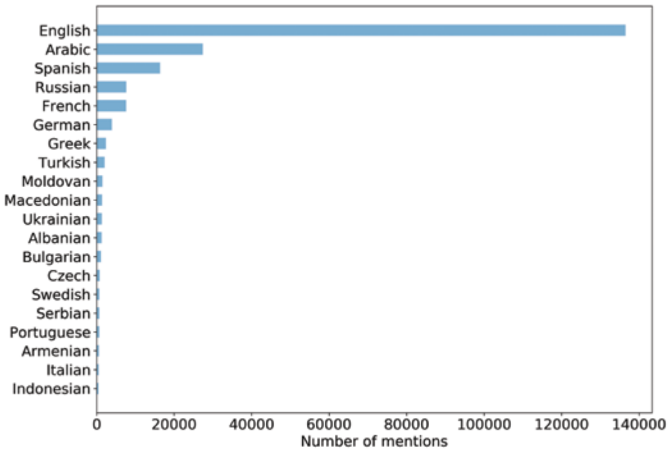


Figure 23: Number of mentions by news source language.



Sentiment

One surprising result of the Syrian Case Study is that the sentiment of stories about the UN is significantly more negative than stories about Russia and the US. This may reflect the true situation, a possibility supported by the more positive perceptions of the UN in other countries in which it has played a key role seen in the UN Case Study. However, given that the sentiments of GDELT mentions are generally negative (see Figure 21), is it possible that a positive actor attracts negative sentiment by being referenced in an overall negative news story? Alternatively, is it possible that the UN tends to be mentioned in stories relating to cooperation which are in themselves negative or critical? Further investigation is required.

GDELT – Next Steps

It is essential to build confidence in GDELT and to understand, and work around, limitations and bias. Two approaches are discussed below.

Benchmarking

- Compare coding of selected news stories by GDELT with coding of the same stories by human coders.
- Are trends and differences discovered by analysis of GDELT data considered credible by country experts?
- Compare GDELT time series with time series derived from similar data gathering systems (see below)
- Compare GDELT time series with time series derived from social media content aggregated at regional, country, and local levels

Similar systems that may be candidates for benchmarking:

- Event Registry
- The Armed Conflict Location & Event Data Project (ACLED)
- UCDP Georeferenced Event Dataset (GED)
- The Integrated Crisis Early Warning System (ICEWS)
- Violent Incident Information from News Articles

News source classification

A better understanding of categories of news source is desirable. This would involve building dictionaries of news sources categorized on various dimensions:

- State/non-state
- Corporate/independent
- Tabloid/non-tabloid
- Global/regional/local

Such a scheme would provide more filters when building time series queries. When considering local news outlets, it would be useful to know their sources of funding.

Methodology – Next Steps

The methodology could be applied to any time-stamped data source including social media. An unexplored application of the methodology is time series forecasting.

References

Carothers, T. and Samet-Marram, O., 2015. The New Global Marketplace of Political Change. Washington: Carnegie Endowment of International Peace.

Gerner, Deborah J., Philip A. Schrodt, Ömür Yilmaz, and Rajaa Abu-Jabr, 2002. The Creation of CAMEO (Conflict and Mediation Event Observations): An Event Data Framework for a Post Cold War World. Presented at the Annual Meeting of the American Political Science Association, San Francisco, 29 August – 1 September 2002.

Goldstein, Joshua S., 1992. A Conflict-Cooperation Scale for WEIS Events Data. *Journal of Conflict Resolution* 36(2), pp.369-385.

Leetaru, K. and Schrodt, P.A., 2013. GDELT: Global Data on Events, Location, and Tone, 1979–2012. In: ISA annual convention, vol.2

Schrodt, Philip A, Yilmaz, Ömür, Gerner, Deborah J, Hermreck, Dennis, 2008. The CAMEO (Conflict and Mediation Event Observations) Actor Coding Framework. Presented at the 49th Annual Meeting of the International Studies Association, San Francisco, 26–29 March 2008.

Shumway, Robert H. and Stoffer, David S., 2017. *Time Series Analysis and Its Applications*, Springer, Switzerland.

Hammond, J. and Weidmann, N.B., 2014. Using machine-coded event data for the micro-level study of political violence. *Research and Politics*, July 2014.

About Us

PeaceRep is a research consortium based at Edinburgh Law School. Our research is rethinking peace and transition processes in the light of changing conflict dynamics, changing demands of inclusion, and changes in patterns of global intervention in conflict and peace/mediation/transition management processes.

PeaceRep.org

PeaceRep: The Peace and Conflict Resolution Evidence Platform | @Peace_Rep_ | peacerep@ed.ac.uk

School of Law, University of Edinburgh, Old College, South Bridge, EH8 9YL

PeaceRep is funded by the Foreign, Commonwealth and Development Office (FCDO), UK.

PeaceRep: The Peace and Conflict Resolution Evidence Platform
peacerep@ed.ac.uk | https://peacerep.org | @Peace_Rep_

University of Edinburgh, School of Law, Old College, South Bridge EH8 9YL

PeaceRep is funded by the Foreign, Commonwealth and Development Office (FCDO), UK