

# Media coverage of violent events in war zone

Lucas Gauchoux

Niroshan Vijayarasa  
surname.name@epfl.ch

Remi Coudert

## Abstract

Media coverage impartiality and consistency of armed conflict is often criticized. The press is seldom accused of looking only at profit, resulting in coverage of war events to be poorly correlated with their importance. We show in this study that this criticism might be founded and that media seems to lose interest for some wars over time. We found that interest of medias for Afghanistan in particular seems to have faded these last years even if casualties in this country are still very high.

We show as well a bias towards the U.S press.

## 1 Introduction

Media is the first and often the only source of information in modern society and is supposed to ensure a certain integrity. The information medias provide, determines one global vision of the world. We aim to show here the loss of interest in the media for armed conflicts that tends to last over time and/or are located in developing countries. We believe that this practice suggests to the population that the situation has improved when it might have not. This study is based on violent events in a subset of countries that faced war during the last fifteen years, namely Afghanistan, Iraq, Mexico, Pakistan and Syria. We used the GDELT 2.0 dataset along with the UCDP dataset.

## 2 Presentation of the data

### 2.1 GDELT

The GDELT 2.0 is the result of a the work of the GDELT project. It aims to collect and aggregate information on various events happening around the world. Each event is classified by its type (ranging from minor events to mass shootings) and the

dataset provides various information such as the groups/countries involved, the location, the date or the number of related articles. The original GDELT 2.0 mention only articles published within the next 15 minutes of its realization. We used an alternative version of this database called GDELT full available on Google Big Query that sums all articles talking about an event, leading to a more accurate analysis.

### 2.2 UCDP

The UCDP dataset provided by UPPSALA University, tracks events similarly to GDELT. It is interesting for us in that it provides useful additional information such as the number of deaths per year in a given country. The added information is relevant to our study since the number of deaths is a reliable indicator of the gravity of a set of events in a country, it is indeed very natural to classify aggregated event gravity in term of casualties.

### 2.3 First data extraction

Since the GDELT dataset is huge (1.2TB), we worked on a subset of the data. The extraction was made using SQL and Google Big Query since the GDELT dataset in the cluster was not complete. We aggregated the data by month and summed the number of articles and events. The resulting table thus represent the number of events and articles per month between 2000 and 2016 separated by country and event root code.

Since the data from the UCDP dataset is given by year, most of our analysis has to be done per year, hence one of the first things to do was to aggregate the events by year. It is interesting to note at this point that our study was supposed to be looking at Syria as well but UCDP's data for Syria is unfortunately blank.

Figure 1: Difference of scale over the years

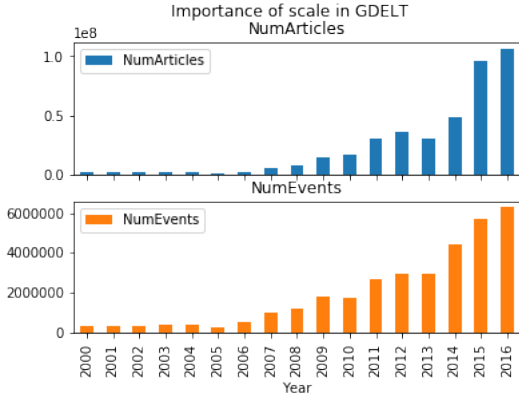
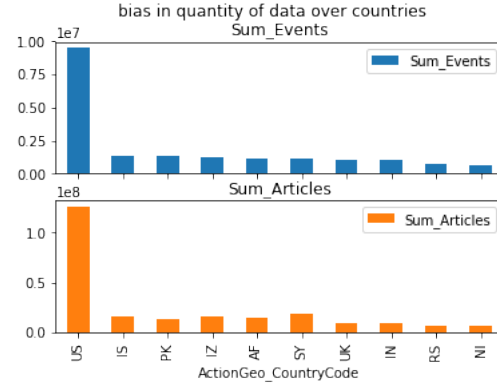


Figure 2: U.S. bias



## 2.4 Data preprocessing and first observations

### 2.4.1 Impact of scale

We first noted that the difference of scale in number of events and articles varies significantly during the years. This is illustrated in Figure 1. Such a difference makes it hard to analyze the data correctly, thus we chose to normalize it by dividing each number of events (resp. articles) by the total number of events (resp. articles) in the world for the current year.

The possible reasons that could explain such a difference in number of events (resp. articles) could be numerous including (and not mutually exclusive) :

- A better data retrieval from GDELT
- An increase in the number of sources
- An increase in the number of violent events over the world.
- An increase in data scale itself, for example thanks to internet.

### 2.4.2 Bias towards the U.S.

By looking at the number of articles by country, we found out that the number of articles for the U.S. was by far the highest and that it was somewhat equivalent between the other countries. Figure 2 showcases this very well.

From this, it lead us to think that the source of articles is most likely U.S. centered.

An important thing to note is that the U.S. was involved in Mexico, Iraq and Afghanistan wars and not in Pakistan and Syria deadly events. Moreover Mexico has the characteristic of being geographically close to the U.S.

## 3 Data processing

Figure 3: Global view of the data

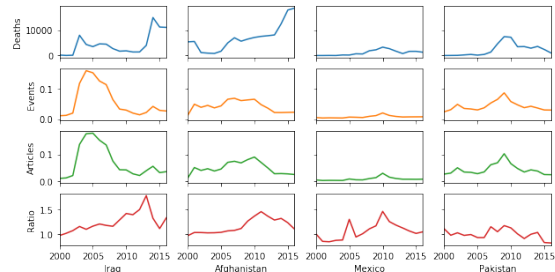


Figure 4: Correlation by country between number of articles and casualties

Iraq	Afghanistan	Pakistan	Mexico
0.155	-0.271	0.752	0.865

After normalization and study of the data, we chose to mostly study the ratio between the number of articles and the casualties by country per year and compare it against the number of deaths by country per year.

If the lost of interest is generalized over all the scrapped sources of the GDELT data, the number of identified event would decrease also and the ratio would be a biased unit of measure. However we have assumed that a violent event is at least mentioned by one of the sources and that ratio is therefore a good measure.

Figure 3 shows a global view of our dataset, difference by country is shown horizontally and difference between parameters vertically. There are several important things to note here. First we can see that the sense of scale between countries is not the same, the data for Mexico for instance is significantly and constantly lower than its counterparts.

Another interesting point to note is the difference in correlation between the number of deaths and the number of articles by country. Figure 4 shows these differences. As we can see, Mexico and Pakistan coverage by the media is highly correlated with the casualties, whereas it is very poorly correlated for Iraq and Afghanistan, even leading to a negative correlation for the latter. Basically when the number of deaths increases, the number of articles decrease.

## 4 Analysis and Results

In this section, we'll discuss our main plots and interesting findings.

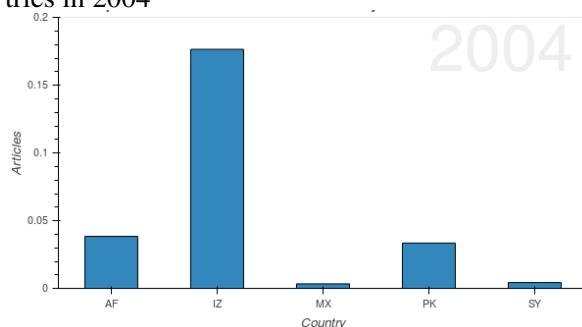
### 4.1 Analysis of proportion of articles

For all of the analysis, it is important to point out the fact that the proportion of the number of articles between the 4 countries studies can vary a lot.

Figure 5 showcases this for the year 2004. We can see that for this particular year, most of the articles were focused on Iraq, among the 5 countries, which is an expected behaviour. Indeed, 2004 marks the second year (and a drastic increase in intensity) of Iraq's war *in which the U.S. are involved*.

Thus when analyzing the ratio between Articles and Events as we do below, we must keep in mind that even if we normalized the data, the difference in proportion of articles can influence the said ratio.

Figure 5: Proportion of articles among the 4 countries in 2004



### 4.2 Ratio Articles/Events compared to number of deaths

The main thing we were interested in was to compare the *shape* (hence a qualitative analysis) that takes the plot of the Ratio (Articles/Events normalized) over the years against the shape of the plot of the number of deaths over the years.

One could expect the two curves to more or less be similar in shape, since they are supposed to be highly correlated.

In the following analyses, we'll refer as **ratio** for our computed value of the (normalized) ratio of number of articles over number of articles.

#### 4.2.1 Iraq analysis

Important dates regarding Iraq :

- 2003 : Beginning of Iraq's war, lead by the U.S.
- 2009 : First announcement of withdrawal of U.S. troops.
- 2010-2011 : Beginning of the Arab Spring
- 2013-2017 : Iraqi Civil War

When looking at Figure 6, we can see that the number of deaths and the ratio have a similar shape and follow a similar tendency from 2003 to around 2008.

From 2008 to 2011, the casualties in Iraq greatly decrease (due to war ending), but we can see that the ratio continues to increase. This can be explained by the withdrawal of U.S. troops, it is indeed expected for the press to talk in length of Iraq at this point.

From 2011 to 2013, the ratio continues to increase even if the number of deaths doesn't. Our hypothesis is that is probably due to the Arab Spring. Having border countries at war, and considering the recent war in Iraq, it seems logical to have a somewhat high coverage.

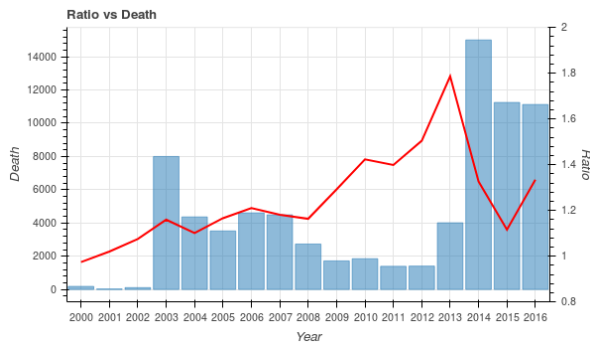
From 2011 however, we can observe a complete reverse in tendencies, the number of deaths explodes whereas the ratio decreases a lot. This can be explained by the high number of war events occurring at this time, in particular in Syria. This seems to show that our hypothesis holds for Iraq, i.e. the press displays a lack of interest for events in Iraq from 2011, even if the casualties in this country **never were this high even during the Iraq War**.

#### 4.2.2 Afghanistan analysis

When looking at Figure 7, we can see a similar tendency between the ratio and the number of deaths in Afghanistan from 2000 to around 2011.

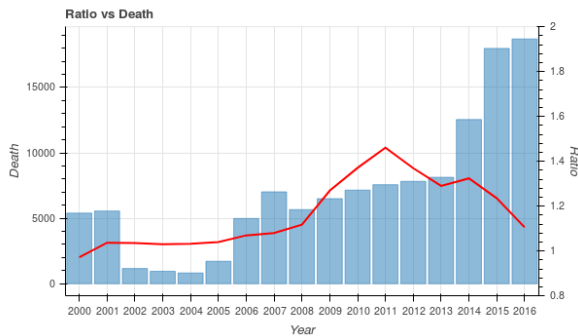
From this year forth, the curve for the ratio drops whereas the number of deaths increases a lot. This looks a lot like for Iraq's case, i.e. the increase of

Figure 6: Ratio vs Deaths for Iraq



events in the region, in Syria in particular, makes the press lose interest for Afghanistan events even if as for Iraq the number of deaths never were this high.

Figure 7: Ratio vs Deaths for Afghanistan



#### 4.2.3 Mexico analysis

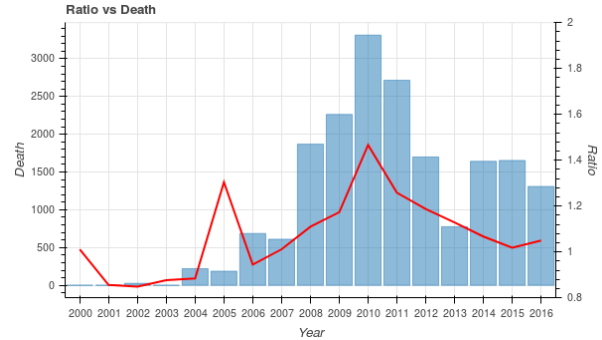
When looking at the plot for Mexico (Figure 8), we can note that the shape of the ratio and the number of deaths over the years is quite similar.

This shows a media coverage that seem logical and expected, as opposed to Iraq and Afghanistan.

Two things are important to recall, leading to two hypotheses :

- Mexico is close to the U.S. and the U.S. is involved in Mexico's events (drug wars). From our hypothesis that most article sources come from the U.S., it looks like the coverage is better when the country is close.
- The scale of data for Mexico is by far the lowest, this might highlight that very violent events are more unusual in Mexico and hence more covered.

Figure 8: Ratio vs Deaths for Mexico



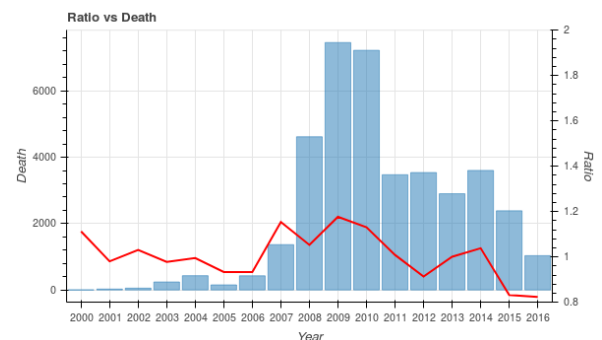
#### 4.2.4 Pakistan analysis

The case of Pakistan, observed in Figure 9, is interesting. Events in this country are often less frequent than in Iraq or Afghanistan, we can observe a logical behaviour for our two parameters up to about 2014, followed by a big decrease in the ratio.

Now this decrease can be explained by the following points :

1. The number of deaths decreases from 2014, as opposed to Iraq and Afghanistan.
2. Pakistan is a border country of Afghanistan and events happening in Pakistan might seem "less important" in term of death scale.

Figure 9: Ratio vs Deaths for Pakistan



## 5 Conclusion

From our study, it seems that our hypothesis is founded and that newspapers lose interest over time for some countries at war even if it is still raging. The wars in Iraq and Afghanistan in particular seem to showcase the most this effect. This may be due to the fact that the U.S. were involved in these wars. The coverage for Mexico and Pakistan looks to be a bit more reasonable.

However we cannot state that our conclusion is definitely true since there are too many parameters to take into account such as location bias, unknown reliability of GDELT and UCDP datasets or scale difference over the years, even if our data processing got rid of part of the biases. We can still reasonably state that our hypothesis is *probably* true.

This loss of interest is really important to note so that it can be worked on. Medias can't just choose the most sensational war and only talk about it. Iraqi prime minister announced the end Iraq Civil War against ISIS less than a month ago. Have you heard of it?

People die every day all around the world but if we don't even hear about it, we might think everything's going fine besides some particular wars.

## References

Le Premier ministre annonce "la fin de la guerre" contre l'EI en

[http://www.lepoint.fr/monde/le-premier-ministre-annonce-la-fin-de-la-guerre-contre-l-ei-en-irak-09-12-2017-2178403\\_24.php](http://www.lepoint.fr/monde/le-premier-ministre-annonce-la-fin-de-la-guerre-contre-l-ei-en-irak-09-12-2017-2178403_24.php)

Iraqi Civil War

[https://en.wikipedia.org/wiki/Iraqi\\_Civil\\_War\\_\(2014%E2%80%93present\)](https://en.wikipedia.org/wiki/Iraqi_Civil_War_(2014%E2%80%93present))

Iraq War

[https://en.wikipedia.org/wiki/Iraq\\_War](https://en.wikipedia.org/wiki/Iraq_War)

Afghanistan War

[https://en.wikipedia.org/wiki/War\\_in\\_Afghanistan\\_\(2001%E2%80%93present\)](https://en.wikipedia.org/wiki/War_in_Afghanistan_(2001%E2%80%93present))

Arab Spring

[https://en.wikipedia.org/wiki/Arab\\_Spring](https://en.wikipedia.org/wiki/Arab_Spring)

UCDP

<http://ucdp.uu.se/>

The full GDELT dataset that we used

<https://bigquery.cloud.google.com/table/gdel-bq:full.events>

Robert West, Jorgen Pfeffer

*Armed Conflicts in Online News: A Multilingual Study*