# Title Page

# Uncovering the state of Sudan civil war through remote sensing and google earth engine

Sudan 2023- An empirical study based on satellite FIRMS fire data and ACLED conflict data.

Battle damage assessment of Sudan civil war using satellite imagery or

Battle damage assessment of sudan civil war through the eyes of remote sensing and google earth engine.

Remote sensing of violent conflict: eyes from above

**This dissertation is submitted in part requirement for the MSc in the Centre for Advanced Spatial Analysis, Bartlett Faculty of the Build Environment, UCL.**

Title:

Name : SAMEERA SIDDIQUI

Course : MSc. Urban Spatial Science

Date: 23rd August 2024

Student number: 23140058

# Table of Content

[Title Page 1](#_Toc174387251)

[Uncovering the state of Sudan civil war through remote sensing and google earth engine 1](#_Toc174387252)

[Table of Content 2](#_Toc174387253)

[Acknowledgements 3](#_Toc174387254)

[Abstract 5](#_Toc174387255)

[List of figures 6](#_Toc174387256)

[List of tables 7](#_Toc174387257)

[Abbreviations 8](#_Toc174387258)

[1. Introduction 9](#_Toc174387259)

[2. Literature Review 13](#_Toc174387260)

[3. Study area and data 17](#_Toc174387261)

[4. Methods 21](#_Toc174387262)

[5. Results and discussion 32](#_Toc174387263)

# Acknowledgements

Ha, a sigh of relief I finally manage to finish my dissertation and what I envisioned to do. I put all my learnings into this research, and it finally took a shape. All the hard work of past 4 months has come to an end.

I would like to say thanks to all my teachers from Bartlett CASA. Last year in September when I came here, I could not have imagined myself learning so much and putting this much effort into the course.

A special thanks goes to my supervisor Ollie Ballinger who guided me all through the way with his immense support and constructive feedback. I could not have finished this research without my better half Asif. Whenever I felt stuck or could not think beyond, he helped me to look through and helped me with ideas and motivated me to do better and my 5-year-old boy Azaad though he did not help in writing but he sat there patiently watched me writing.

Lastly, I would like to say thank you to all my friends who believed in me and my vision. I want to dedicate this study to all those people who are suffering in war all over the world. May peace prevails.

In the end I would take responsibility for this research if there were any flaws or errors.

# Abstract

# List of figures

# List of tables

# Abbreviations

AFP- Agence France press

AP- Associated press

CIR- Centre for Information Resilience

EPSG-European Petroleum Survey Group

FIRMS – Fire Information for Resource Management System

MODIS- Moderate Resolution Imaging Spectroradiometer

RSF – Rapid Support Forces

SAF- Sudan Armed Forces

VIIRS- Visible Infrared Imaging Radiometer Suite

# Introduction

In the annals of history, wars and conflicts have cast a somber shadow, leaving behind a trail of sorrow and tragedy. From the ongoing struggles in Gaza, Ukraine, Sudan, Myanmar, Syria, and Ethiopia, the current situation paints a grim picture of uncertainty and unrest.

The internet and smartphones enable the sharing of information through various online platforms and applications. These include Twitter, Facebook, YouTube, Telegram, and discussion boards like Reddit. A crucial aspect of this information sharing is the use of remote sensing data. Remote sensing technology helps track and analyze the links between conflicts and environmental damage, providing valuable insights into the environmental impact of human activities and natural events.

There’s a humanitarian crisis happening since past year in Sudan which has been overshadowed by Gaza and Ukraine and it’s getting worse. (Wim Zwijnenburg and Ballinger, 2023; Kristof, 2024).

1.1 Research Background

The Sudanese civil war centers around a power struggle between two major groups: the Sudan Armed Forces (SAF) led by general Abdel Fattah al-Burhan, and the paramilitary Rapid Support Forces (RSF) commanded by General Mohamed Hamdan Dagalo, also known by ‘Hemedti’. (Ali, T., 2024)

Since, the conflict broke out on 15th April 2023, it has resulted in the loss of more than 18,800 people and over 33,000 injured, according to humanitarian partners. Over 10 million people have fled their homes, and this includes more than 5 million children and over 2 million people who have crossed into neighboring countries. (OCHA, 2024)

The country is facing extreme shortages of food, water, medicine, and fuel and nearly 18 million people are facing acute food insecurity and 5 million of them at emergency levels. Before the current conflict, Sudan had been grappling with violence and displacement since the onset of the Darfur crisis in 2003. (UNHCR ,2024)

‌A large black smoke billowing from a building

Description automatically generated

Figure 1: Smoke rises over buildings after aerial bombardment, during clashes between the paramilitary Rapid Support Forces and the army in Khartoum North, Sudan, May 1, 2023. REUTERS/Mohamed Nureldin Abdallah. (Abdelaziz, K. and Lewis, A., 2023)

According to investigators from the Centre for Information Resilience’s (CIR) Sudan Witness project, fire constitute to devastate villages and settlements in western Sudan. (CIR, 2024) The reports of fighting or airstrikes coinciding with clusters of fires indicates that fire is used as a weapon of war.

1. The Geo-political Aspect

The ongoing civil war in Sudan holds immense geopolitical significance due to the country's strategic location near the Suez Canal in the Horn of Africa, a vital corridor for global trade. Sudan's abundant oil and gold reserves have drawn the interest of various external actors aiming to advance their own strategic interests, further intensifying the conflict. (Vardhan, A., 2023)

The RSF profit from gold trade with Russia and the UAE. The UAE refines most of Sudan’s gold, sourced from RSF leader Hemedti’s military business. Russia, via the Wagner Group, trains the RSF in exchange for gold access. (Doxsee, 2023) Egypt remains neutral due to its economic ties with the UAE. (Rickett, 2024; ADF, 2024; Trad, 2023) This intricate web of alliances and interests positions the Sudanese civil war as a potential proxy battleground, with regional and global powers maneuvering to secure their stakes.

1. The Role of Media – GAP in Coverage

Media coverage of the Sudan conflict has been limited compared to other global crises like the Israel-Gaza conflict and the Ukraine war, which receive more attention due to their direct implications for global security. Restricted access to conflict zones in Sudan, including a ban on foreign media until recently, has also hindered reporting. Major outlets such as The Washington Post and The New York Times offer limited coverage, while agencies like AP, AFP, and Reuters continue to report on the violence and humanitarian issues.

Remote sensing methods come to rescue even if the access was there as the war is happening in entire country. It can monitor areas that are difficult to cover on the ground.

* 1. Research Question and Objective

How can FIRMS fire data and the ACLED conflict data be used to identify and analyze the relationship between conflict induced fires and active warfare in Sudan?

The purpose of this thesis is to explore the relationship between satellite detected fires and fighting in Sudan with the use of ACLED conflict and FIRMS fire data. The approach can be used to develop real-time conflict monitoring system based on satellite imagery for regions which are not covered by media and news.

# Literature Review

This section introduces the background and motivation for the work carried out for this report, which have been introduced above in the form of similar studies done using remote sensing in war impacted regions where some gaps are identified and how with the methods explored in the research can improve the ways of identifying war related fires.

Conflict research can be divided into four main categories followed by the temporal lag for which each effect requires to become visible. For example, conflict-induced structural damage from bomb detonations or fire is generally a very immediate impact, occurring within minutes to hours. Other effects such as environmental damage (hours to days), population movement (days to months), and land-cover/land-use change (months to years) take longer to materialize. Although there is some overlap between different conflict effects, this typology provides a useful way to organize the current body of research. The three-paper discussed here discusses all these categories.

Bromley, L.(2010) paper had covered the conflict of 2003 in Darfur, Sudan using Moderate Resolution Imaging Spectroradiometer (MODIS) fire data to detect fires and burned area. The paper was based on the finding that using remote sensing for violent conflicts produce heat in the form of fires, incendiary bombs, etc., and that satellite-based instruments can detect those heat events. During that time, it was very difficult to know the exact location of the conflict as to gather ground truth data was not possible and the restrictions from government which allowed very minimal contingents of African union to move so the information was sparse that where the attacks took place. It was challenging but one thing which worked was the involvement of social media platform which played an important role in covering this civil war where foreign journalist was banned but organization like CIR which were involved in investigating, actively monitoring and documenting and verify human rights interferences and conflict dynamics on the ground. (Centre for Informati., 2023)

The vast majority of studies on Bromley, L.(2010) paper have been qualitative. The data used only covered major towns where journalist had greater access to refugees leaving the other parts which might have been affected also and the dataset used does not differentiate between the burned settlements from otherwise destroyed settlement and neither indicates the date of destruction. The paper lacks clarity with MODIS fire detection system which was able to sample only larger fires.

Looking at all the limitations the following research aims to develop a methodology which uses Fire Information for Resource Management System (FIRMS) fire data used in google earth engine cloud-based platform and count the number of fires from 2000 to 2023 for the whole country dividing it into grids and validating it by ACLED data to find some correlation between fires and conflicts. The current study main objective is developed real time monitoring system for wars like Sudan where there is little media coverage.

Another work of Hamilton and Drake (2018) explored the efficacy of the MODVOLC algorithm, which utilizes MODIS satellite data to monitor conflict-related fires during the Syrian Civil War. MODVLOC uses infrared satellite data acquired by NASA’s MODIS instrument to monitor earth’s surface for the thermal emission signature of volcanic eruptions, wildfires, and anthropogenic heat sources (e.g., gas flares) (Wright et al., 2002)

By comparing pre-war and post-war fire detections, the study highlighted a significant increase in hotspots within urban areas following the onset of conflict, as opposed to rural regions. Despite identifying a substantial number of conflict-related heat events, the study noted several limitations. MODVOLC often failed to detect fires in areas with documented violence and sometimes recorded hotspots in regions without reported incidents. Additionally, the correlation between MODVOLC detections and news coverage was weak, suggesting that some events detected by the algorithm were either not reported or misidentified as conflict related. These findings underscore the potential and challenges of using satellite-based fire detection in conflict monitoring, indicating the need for integrating satellite data with other reporting methods to improve accuracy and reliability (Hamilton and Drake, 2018)

The main weakness of the study is the failure to address how to detect conflict-related fires, especially in urban areas where tall buildings can obstruct the sensors' view which lead to underreporting of fire events in densely populated areas. Also there is an issue with false positives like the algorithm might record hotspots in areas without any reported conflict incidents, leading to potential misinterpretations of the data. This issue underscores the importance of corroborating satellite data with ground reports which leads to weak correlation between MODVOLC detections and media reports, indicating that some detected events may not be covered by international news. This constraint suggests that relying solely on news reports for validation could result in an incomplete analysis of conflict-related fires. Also, the algorithm does not distinguish between conflict-related fires and other sources of heat such as industrial activities or accidental fires. This impediment necessitates additional contextual analysis to accurately attribute detected fires to conflict activities.

Therefore, the second part of current research aims to identify different types of fires including the anomalous fires which are conflict induced by calculating the z-score value.

Tomchenko et al. (2023) paper conducted an important study on the environmental impact of the Ukraine war, specifically focusing on fires caused by military activities. Utilizing remote sensing technologies, particularly the VIIRS spectroradiometer and data from NASA's FIRMS, the study effectively monitored and assessed fires in conflict zones. The research revealed valuable insights into the spatial distribution and seasonality of fires, demonstrating the efficacy of integrating satellite imagery with traditional ground-based data for environmental monitoring. However, the study also highlighted several limitations, such as potential inaccuracies in data due to missed smaller fires or misinterpreted heat sources, limited ground verification, and a focus on regions near the front line.

Well these findings are relevant to the Sudan conflict suggesting that remote sensing can be a powerful tool for monitoring destruction caused by fires in the form of arsons and by using the dataset from ACLED and GEE cloud platform will make the research more reliable.

Conclusion

The ongoing civil war in Sudan is faced with different challenges and to investigate the damage caused by fires and number of people displaced can be done with the help of ACLED data and FIRMS fire data which uses near real time active fire locations using the standard MODIS MOD14/MYD14 fire and thermal anomalies product. Moreover a spatio-temporal study can help in knowing about changes happened in past 20 years. This research can help all the war-torn regions in world which have very less media attention and struggling with similar kind of civil war. This paper can be used to inform policy and humanitarian response in Sudan.

# Study Area and Data

The study area is country Sudan. The entire country is in war so for the analysis FIRMS fire data and ACLED conflict data is taken for the year 2000 to 2023. The most affected areas in the year 2023-2024 are Khartoum, El-Geneina, El-Fasher in Darfur, Wad Al-Nura in Al Jazirah and Nyala. These are the areas with most population.

A map of the state of south africa

Description automatically generated

Figure2 : Population density Map of Sudan

A blue sky with stars

Description automatically generated

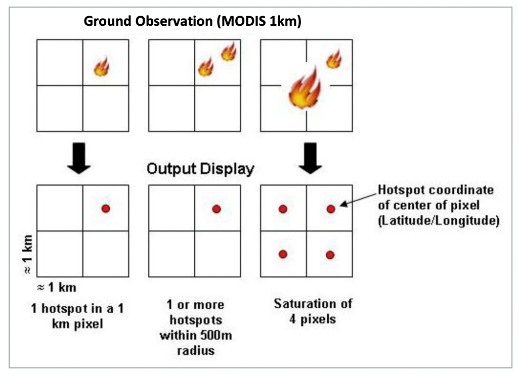
DATA

For this research all the dataset used are open source.

1. FIRMS

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| The study makes use of FIRMS data which can be useful when working to verify statements by actors on the ground, even in in circumstances where little or no imagery or footage is available. The FIRMS data is derived from MODIS. The present study uses the earth engine version of the FIRMS dataset which contains the NASA's Land, Atmosphere Near real-time Capability for Earth observations (LANCE) fire detection product in rasterized form. The near real-time (NRT) active fire locations are processed by LANCE using the standard MODIS MOD14/MYD14 Fire and Thermal Anomalies product. The resolution is 1000 metres. (Earth Science Data Systems, 2021) | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Each active fire location shows the centroid of a 1km pixel which is then captured by the algorithm containing one or more fires within the pixel. The data is rasterized in the form of 1km bounding box for each FIRMS active fire point; pixels in the MODIS sinusoidal projection that intersect the FIRMS BB are recognized; or if multiple FIRMS BBs intersect the same pixel, the one with higher confidence is retained and in case of tie the brighter one is retained. So the satellites take a picture of events as they pass over earth. Each hotspot or active fire detection indicates the center of a pixel identified as containing one or more fires or other thermal anomalies, such as volcanic activity. Therefore, the “location is the center point of pixel and not the co-ordinates of actual fire.( Earth Science Data Systems, N., 2021)



The shapefiles are in the geographic WGS84 projection. Some limitation of active fire pixel location is that it may not always be the most appropriate source of fire related information and the data do not provide any information on cloud cover or missing data.

1. ACLED Data (Raleigh, Kishi and Linke, 2023)

The ACLED is a project focused on the detailed collection, analysis, and crisis mapping of conflict data. It gathers information on the dates, actors, locations, fatalities, and types of all reported political violence and protest events globally. The ACLED team analyzes this data to describe, investigate, and test conflict scenarios, providing both the data and their analyses freely to the public. It is recognized as the highest-quality and most widely utilized near real-time source for political violence and protest data worldwide. The data is updated weekly, covering events up until the most recent Friday.

For the research ACLED data was downloaded for the year 2000 to 2023 as a CSV file for Sudan.

1. The Global Administrative Unit Layers (GAUL) compiles and disseminates the best available information on administrative units for all the countries in the world, providing a contribution to the standardization of the spatial dataset representing administrative units. The GAUL always maintains global layers with a unified coding system at country, first (e.g. departments), and second administrative levels (e.g. districts).

In the paper Sudan boundary has been used taken the ADM0\_Code which is the GAUL country code. (Google for Developers, 2015)

# Methods

A diagram of a company

Description automatically generated

Flowchart

The important point to understand about this war is that the casualties or destruction was mainly due to fires. The attacks and destruction is caused setting fire to a building or a house or burning things. Therefore, the aim of this research is develop some kind of relation between fires and conflict. Bromley, L.(2010) paper also used MODIS fire data which can be considered similar to FIRMS fire data but he did not use ACLED data which is the major difference from the previous research.

This study is using FIRMS data for calculating number of fires and ACLED data for counting number of conflict events in Sudan and then empirically a z-score is calculated which can differentiate between anomalous fires.

## Data Preparation

The objective is to prepare a long data also called panel data. The data is cleaned and validated before calculating the key components. The limitations of the data and calculation methods chosen are reflected upon and outlined in the assumptions.

The methodology in the research has been divided into two parts –

The first half of data preparation is done in Google Earth Engine(GEE) cloud computing platform. It analyzes the data that is stored remotely, each time some code is written it is send to server and speeds up the process as computation is not done locally.

While the study area is Sudan and to get the boundary of Sudan in GEE data catalogue there is Global Administrative Unit Layers (GAUL) which gives the boundary for Sudan and level 0 is picked which gives the country level boundary.

The best way to deal with the entire country is to divide it into grids so that each grid will cover data for all the years from 2000 to 2023. For defining the grid its important to define the scale of grid and the projection. The scale of the grid is taken (50 km x 50 km) in size. The coveringGrid() function in GEE script creates a grid that covers the entire geometry of Sudan. By dividing it into grids a long data is created which will be helpful for analysis.

A map of the state of sudan

Description automatically generated

Figure 3: Screenshot of the grids for Sudan in GEE.

The code sets the projection to (EPSG:4326) which is a common geographical coordinate system. Firms fire data is loaded for the year 2000 to 2023 with 1000m resolution raster image and a threshold of 300 is taken that minimum range based on the value of T21 which is the brightness temperature measured in the thermal infrared band at 21µm. this band is sensitive to heat and is used to detect high temperatures, such as those found in fires.

After the grids are formed, the fire count can be calculated for every grid for each year. Now these fires can be anything like agricultural fire, forest fire, recreational fire and conflict fire. In order to get the anomalous induced fires there is a need for some outside resource to validate and match it with conflicts. In supporting this theory ACLED data is used which accounts for conflicts in the country.

The advantage is that both datasets are available starting from the year 2000, allowing us to examine historical data over the years and compare it to 2023, the conflict year, to see how it differs from past records.

## Data Cleaning

The FIRMS satellite data is used in GEE without any pre-processing but for ACLED data before using it on GEE it has been cleaned and filtered out. The original data consist of 42,156 rows and 32 columns. Due to the volume of data and the need of data manipulation, data cleaning was undertaken using Pandas 2.1.0 and NumPy 1.24.4 package. The steps undertaken were:

* Ensuring datatypes are correct,
* Identification and removal of missing data
* Checking the column names.
* Filtering on the basis of column **sub\_event\_types** and dropping few rows which seems non-violent or not fire related. Also filtering another column geo\_precision whose value is from 1 to 3 indicating the level of certainty of location recorded for event. Thus dropping all rows with geo\_precision = 3

A blue and white chart with text

Description automatically generated

Figure 4: ACLED Events and Sub-event types (ACLED, 2023)

HYPOTHESIS

The hypothesis theory in this study is that all the ACLED events should match with the anomalous fire events and then it can be conclude that there is a correlation between fires and ACLED events.

## Building the Panel Dataset

The current investigation uses these dataset to make a panel dataset where each grid for every year will give one column of fire count and other column of ACLED count. So for year 2000 grid1 will have these inputs then year 2001 for grid 1 will have similar information and so on. Every row is a time instance. Every year has a different layer from where the fire incidents can be seen in each year. Its an iterative process. Table 1 has columns grid\_id, year, fireCount, acled\_count, acled\_event\_ids and respective geometry of fire data and acled data. With this type of dataset conflict induced fires can be calculated by calculating the z-score of anomalous fires and regression can be performed to establish some kind of correlation between fire data and ACLED data.

A screenshot of a computer

Description automatically generated

TABLE 1: PANEL DATAFRAME

## Creation of Geo data frame

The above data frame consist of geographic data stored as JSON strings for ACLED and fire. To perform spatial operations and analyses on the data a Geo Data Frame is required. The above data needs to be converted into Shapely geometry objects.

Table 2: Geodataframe

## Initial Data Exploration

Understanding the data being used is key before undertaking further analysis. Plotting a graph for input data helps in visualization data relationships.

A graph showing the number of companies

Description automatically generated with medium confidence

Figure 5: Trend of the Input Data

The above figure shows the pattern of the input data. In red is the FIRMS data and in blue is ACLED conflict data. The trend can be seen for all the years from 2000 to 2023. The long term average is also taken which gives the mean value of a dataset over a long period. Few peaks can be seen for ACLED in year 2004 and year 2023 which is exactly showing that sudden peak compared to the historical data. This shows that how the conflicts have peaked in these years.

A screenshot of a graph

Description automatically generated

A screenshot of a graph

Description automatically generated

A screenshot of a graph

Description automatically generated

Figure 6: Map of Sudan for Year 2003, 2004 and 2023

The above maps clearly shows the spread of fire data through the grids all over Sudan and ACLED events. The empty cells are the zero fire count so it’s easier to distinguish between actual events and no events.

The first two maps of 2003 and 2004 shows conflicts occurred in Darfur, Sudan which started in latter half of 2003 and continued till early 2004. Here the ACLED map clearly shows a yellow grid which is 100 above conflicts. Most of the conflicts are concentrated in the west Darfur region.

While year 2023 which is the current conflict year, a peak can be seen in red in figure 5 and that can be seen on the map also specially one grid is all yellow which is the capital city Khartoum which has more than 4000 ACLED events. So, all the conflicts are concentrated there. Therefore, a geographical shift can be seen in the history of Sudan conflict from west to all over Sudan especially concentrated in the capital of Sudan. like andaround Khartoum

### Log Transformation

Both ACLED and Fire counts have asymmetric distribution with extreme right skew. Log transformation (using the natural logarithm) was applied to reduce skewness and stabilize variance. Post log transformation, the data is still skewed however with slightly shorter tail. There are signs of heteroskedasticity in the ACLED count (response variable), specifically as variance is not constant for different values of Fire count.

|  |  |
| --- | --- |
| A graph of a graph with red dots  Description automatically generated | A graph of a fire count  Description automatically generated |

## Identification of Anomalous Fires

The second part of the methodology is distinguishing between the anomalous fires. The anomalous fires are those fires which happened due to certain extreme events like in our case is conflict. It will be different from the regular fires.

The data is divided into 2000-2022 that is the pre-conflict years that is the historical data and 2023 is the conflict years. The shape of the data from 2000-2022 has 19550 rows and year 2023 has 850 rows. These are the following steps taken

Step 1: Created sub-dataset for historical period (2000-2022) and current conflict year (2023)

Step 2: Binarize both datasets (each grid and all years) based on ACLED-count (1 ACLED conflict and 0 for no ACLED conflict). Store binary in acled\_binary column.

Step 3: Calculate mean and std. deviation for 2002-2022 period for each panel. Took overall mean and std (acled\_conflict 0 and 1).

Step 4: Calculate z-scores for 2023

By calculating the mean value of fire in each grid cell and standard deviation, the z-score is calculated. Generally z-score above 3 is considered anomalous. By comparing a particular observation to that cell historical data it can be seen that how much higher or lower that fire is.

Applied percentile method to arrive at optimum z-score, which was validated using confusion matrix and calculating accuracy, precision and recall values. Therefore with this calculation of z-score all the values above the threshold value of 1.48 can be considered as conflict induced fires. So the actual fire count known and the z-score known conflict induced fires can be calculated with these equation.

The MODVOC algorithm used in (Hamilton and Drake, 2018) study does not distinguish between fires which are conflict induced and other fires such as industrial activity fires or accidental fires. Therefore calculating the z-score and choosing threshold can solve this problem and we can get a number for conflict fires.

How should we use existing literature with my current work?????

## Panel Regression

The final dataset for 20400 rows. The maximum fire count is 44 and the ACLED count is 4637 which can be seen clearly in the above map of 2023 where Khartoum is showing the maximum conflicts event happening as the population is not evenly distributed. So, for further processing its important to take log of fire\_count and acled\_count.

After taking log, a Pooled OLS regression model is a good starting point and can act as a reference model. Our dependent (endogenous) and explanatory (exogenous) variables are as follows:

Step 1: Transform data into right format for use in regression model including multi index data, with grid\_id (entity) and year (period) index.

Step 2: Apply regression models: PooledOLS, PanelOLS (Fixed-effects) and Random Effect model.

Step 3: Check goodness of fit – normality, heteroskedasticity, correlation etc.

* **Pooled OLS**: This model assumes that there are no individual-specific effects or time-specific effects. It treats the data as if it were a single large cross-sectional dataset, ignoring any panel structure.
* **Fixed Effects (Entity and Time)**: This model controls for unobserved heterogeneity by allowing for individual-specific (entity) and time-specific fixed effects. It captures the effects that vary across entities and over time but are constant within an entity or time period.
* **Random Effects**: This model assumes that the individual-specific effects are random and uncorrelated with the independent variables. It allows for variation across entities but assumes that these variations are randomly distributed.

The pooled OLS regression model equations as follows:

The general equation is like this

A black and white rectangular sign with a letter x

Description automatically generated

**Fixed effect Model**

The Fixed Effects regression model is used to estimate the effect of intrinsic characteristics of individuals in a panel data set. Examples of such intrinsic characteristics are genetics, acumen and cultural factors. Such factors are not directly observable or measurable but one needs to find a way to estimate their effects since leaving them out leads to a sub-optimally trained regression model. The Fixed Effects model is designed to address this problem. (Date, 2022)

The general equation is

A black and white rectangular sign with letters and numbers

Description automatically generated

The above equation suggests an approach for constructing the following two kinds of models — the Fixed Effects model, and the Random Effects model depending on whether or not the Covariance term in the above equation is zero, i.e. whether or not the unobservable effects are correlated with the regression variables.

In this model, it is assumed that the unobservable individual effects are correlated with the regression variables. In effect, it means that the Covariance and in the above equation is non-zero. It is also assumed that the bias introduced due to the omission of the unit-specific factors is group-specific.

To compensate for this bias, a group-specific intercept is introduced called into the model. is assumed to act in a direction that is opposite (in a vector sense) to the effect of the omitted-variable bias.

A black and white sign with a plus and a symbol

Description automatically generated

Here is assumed constant across all time periods.

# ‌Results

Empirical selection of z-scores and validation.

The panel regression is done on panel dataset

Yes, look there is a correlation between fires and fighting. Compare 2003 to 2023.

A screenshot of a computer

Description automatically generated

All models use the same dataset with 20,400 observations. Pooled OLS used a clustered covariance estimator, at the entity level, to account for potential within-entity correlation. The Fixed Effect and Random Effects models used unadjusted covariance estimators.

**R-squared:** The Pooled OLS indicates that approximately 2.15% of variation in Log (ACLED count + 1) is explained by Log (Fire count + 1). For Fixed effect model, the variation explained by Log (Fire count + 1) was only 0.28% (very low), indicating that independent variable explains very little of the variation in the dependent variable when entity and time effects are controlled for. In other words, the unobserved independent variables potentially explain more variation in dependent variables. The Random effect model performed marginally better than Fixed effect model, but still explains very little variance of dependent variable.

**F-statistic and P-value:** All 3 models have significant F-statistic (p-value = 0.000), indicating that the models as whole are statistically significant. The Pooled OLS has strongest fit (F-statistic = 448). The Fixed effect model has a lower F-statistic (12.869) indicating the model significantly weaker fit when controlling for entity and time effect. The Random effect model indicated moderate fit with (F-statistic = 47.236) compared to other models.

**Coefficient estimates and t-statistics:**

const (Constant/Intercept): These are the estimated value of log\_acled\_count when log\_fireCount is zero.

The Pooled OLS has 0.1209 (with a t-statistic of 9.1499). The Fixed Effect has 0.1482 (t-statistic: 35.883) and Random Effects: 0.1400 (t-statistic: 10.018)

Now for log\_fireCount: This coefficient shows the expected change in log\_acled\_count for a one-unit change in log\_fireCount. The Pooled OLS: 0.0969 (t-statistic: 6.3363), the Fixed Effect: 0.0267 (t-statistic: 3.5873) and Random Effects: 0.0477 (t-statistic: 6.8728)

So in fixed effect it accounts for both entity and time effects while random effect takes into both entity and time but assumes they are random.

The positive coefficients across all models strongly suggest that as log\_fireCount increases, log\_acled\_count also increases, though the magnitude of this effect varies by model which proves that there is a relation between both dependent and independent variables and that was the aim of this study to show that a relationship exist.

A screenshot of a computer

Description automatically generated

So, by comparing all these values of different models the Fixed Effect model controls for unobserved heterogeneity by allowing for entity-specific effects, making it suitable as the assumption is that these effects are correlated with the independent variables.

As the research is concerned with the effects of variables within entities over time, the fixed effect model is usually preferred despite the low R-Squared.

The Fixed Effect model, despite its low R-squared, is valuable because it provides insight into the relationship between log\_fireCount and log\_acled\_count while controlling for unobserved, time-invariant factors. The statistical significance of the log\_fireCount coefficient indicates that the model is capturing a meaningful relationship, making it a crucial part of the analysis. Therefore the hypothesis with which we stated is true.

# Discussion

Interpret those results..analyze the spatial pattern of fighting in Sudan.

R4:Can remote sensing data from MODIS and Google Earth Engine effectively monitor and predict fire outbreaks in conflict zones like Sudan?

R5: What are the spatio-temporal patterns of fire incidents in Darfur, Sudan, during the civil war period?

Important to mention about Ukraine War Model by Sondre’s- The Economist and Solstad, Sondre (corresponding author), 2023.

Yes, the results do provide valuable insights about the Fixed Effect model, even though the R-squared value is low. Here's how to interpret the significance and usefulness of the Fixed Effect model in this context:

### 1. \*\*Understanding the Low R-Squared Value:\*\*

- \*\*R-Squared Interpretation\*\*: The R-squared value, particularly the "Within R-squared" for Fixed Effects, measures the proportion of the variance in the dependent variable (within entities over time) that is explained by the independent variables. A low R-squared does not necessarily mean that the model is ineffective; it simply indicates that the independent variables explain only a small portion of the variation in the dependent variable within entities.

- \*\*Nature of the Data\*\*: In panel data, especially with many entities and over time, it’s common for the within R-squared to be lower because individual-level variations may be more influenced by factors not captured by the model, such as unobserved heterogeneity.

### 2. \*\*The Role of the Fixed Effect Model Despite Low R-Squared:\*\*

- \*\*Controlling for Unobserved Heterogeneity\*\*: The Fixed Effect model's primary strength is its ability to control for time-invariant characteristics of the entities (e.g., countries, companies, etc.) that might affect the dependent variable. This means that the Fixed Effect model provides more reliable estimates of the relationship between the independent variable (`log\_fireCount`) and the dependent variable (`log\_acled\_count`) by isolating the effect of `log\_fireCount` from these unobserved, constant factors.

- \*\*Statistical Significance of Coefficients\*\*: Despite the low R-squared, the coefficient for `log\_fireCount` (0.0267) is statistically significant with a p-value far below 0.05 (indicated by the t-statistic of 3.5873). This suggests that even after controlling for unobserved, entity-specific factors, there is a significant relationship between `log\_fireCount` and `log\_acled\_count`.

### 3. \*\*Implications for Research:\*\*

- \*\*Focus on Coefficient Significance\*\*: In Fixed Effect models, the emphasis is often on the significance and magnitude of the coefficients rather than the R-squared value. The significance of the `log\_fireCount` coefficient indicates that changes in this variable are indeed associated with changes in the `log\_acled\_count`, even when accounting for entity-specific effects.

- \*\*Utility of Fixed Effects\*\*: The Fixed Effect model is particularly useful when the goal is to understand how changes in an independent variable within an entity affect the dependent variable over time, rather than explaining variance across entities. The model's ability to control for unobserved, time-invariant factors is its primary advantage.

### 4. \*\*Choosing Between Models:\*\*

- \*\*Trade-off\*\*: While the Fixed Effect model might not explain as much of the overall variance in the dependent variable (as reflected by the lower R-squared), it provides a more accurate estimate of the relationship between the variables of interest by removing the bias from unobserved heterogeneity.

- \*\*Model Suitability\*\*: The choice of model should be guided by the research question. If the research is concerned with the effects of variables within entities over time, the Fixed Effect model is usually preferred despite the low R-squared.

### Conclusion:

The Fixed Effect model, despite its low R-squared, is valuable because it provides insight into the relationship between `log\_fireCount` and `log\_acled\_count` while controlling for unobserved, time-invariant factors. The statistical significance of the `log\_fireCount` coefficient indicates that the model is capturing a meaningful relationship, making it a crucial part of the analysis.

# Conclusion

‌

# References

OCHA (2024). *Sudan*. [online] reports.unocha.org. Available at: <https://reports.unocha.org/en/country/sudan/>.

Wim Zwijnenburg and Ballinger, O. (2023). Leveraging emerging technologies to enable environmental monitoring and accountability in conflict zones. *International Review of the Red Cross*, 105(924), pp.1497–1521. doi:https://doi.org/10.1017/s1816383123000383.

‌

Kristof, N. (2024). Opinion | From the Embers of an Old Genocide, a New One May Be Emerging. *The New York Times*. [online] 15 May. Available at: <https://www.nytimes.com/2024/05/15/opinion/darfur-sudan-genocide.html>.

Abdelaziz, K. and Lewis, A. (2023). Sudan factions agree to extend ceasefire deal amid clashes. *Reuters*. [online] 30 May. Available at: <https://www.reuters.com/world/africa/heavy-clashes-sudans-capital-truce-set-expire-2023-05-29/>.

UNHCR (2024). *Sudan crisis explained*. [online] www.unrefugees.org. Available at: <https://www.unrefugees.org/news/sudan-crisis-explained/>.

Ali, T. (2024). *Sudan in Crisis: Mapping the War’s Toll on People*. [online] ArcGIS StoryMaps. Available at: https://storymaps.arcgis.com/stories/ee9c6513732344448bd6337fdadc11c6 [Accessed 10 Jun. 2024].

Gallopin, J.-B. (2024). ‘The Massalit Will Not Come Home’. *Human Rights Watch*. [online] Available at: <https://www.hrw.org/report/2024/05/09/massalit-will-not-come-home/ethnic-cleansing-and-crimes-against-humanity-el>.

CIR (2024). More than 200 villages and towns damaged or destroyed by fire since the start of the war in Sudan, with April the worst month on record. [online] Centre for Informati. Available at: <https://www.info-res.org/post/more-than-200-villages-and-towns-damaged-or-destroyed-by-fire-since-the-start-of-the-war-in-sudan-w>.

‌ Rickett, O. (2024). *How the UAE kept the Sudan war raging*. [online] Middle East Eye. Available at: <https://www.middleeasteye.net/news/sudan-uae-war-arms-trade-rsf>.

Doxsee, C. (2023). How Does the Conflict in Sudan Affect Russia and the Wagner Group? www.csis.org. [online] Available at: https://www.csis.org/analysis/how-does-conflict-sudan-affect-russia-and-wagner-group.

‌ ADF (2024). *Smuggled Gold Fuels War in Sudan, U.N. Says*. [online] Africa Defense Forum. Available at: <https://adf-magazine.com/2024/02/smuggled-gold-fuels-war-in-sudan-u-n-says/>.

Trad, T.K. (2023). *Sudan’s gold: The precious metal used to fund conflicts*. [online] https://www.newarab.com/. Available at: <https://www.newarab.com/analysis/sudans-gold-precious-metal-used-fund-conflicts>.

REFERENCES

Earth Science Data Systems, N. (2021). *MCD14DL-NRT | Earthdata*. [online] www.earthdata.nasa.gov. Available at: https://www.earthdata.nasa.gov/learn/find-data/near-real-time/firms/mcd14dl-nrt.

‌

Earth Science Data Systems, N. (2021). FIRMS Frequently Asked Questions | Earthdata. [online] www.earthdata.nasa.gov. Available at: https://www.earthdata.nasa.gov/faq/firms-faq#ed-wfm-download [Accessed 31 Jul. 2024].

Firms CSV- firms.modaps.eosdis.nasa.gov. (n.d.). *NASA-FIRMS*. [online] Available at: <https://firms.modaps.eosdis.nasa.gov/country/>.

Google for Developers. (n.d.). FIRMS: Fire Information for Resource Management System | Earth Engine Data Catalog. [online] Available at: https://developers.google.com/earth-engine/datasets/catalog/FIRMS#bands [Accessed 31 Jul. 2024].

Google for Developers. (2015). *FAO GAUL: Global Administrative Unit Layers 2015, Country Boundaries  |  Earth Engine Data Catalog  |  Google for Developers*. [online] Available at: https://developers.google.com/earth-engine/datasets/catalog/FAO\_GAUL\_2015\_level0#terms-of-use [Accessed 9 Aug. 2024].

‌

Raleigh, C., Kishi, R. & Linke, A. Political instability patterns are obscured by conflict dataset scope conditions, sources, and coding choices. Humanit Soc Sci Commun 10, 74 (2023). <https://doi.org/10.1057/s41599-023-01559-4>

The Economist and Solstad, Sondre (corresponding author), 2023. The Economist war-fire model. First published in the article "A hail of destruction", The Economist, February 25th issue, 2023.

Date, S. (2022). *Understanding the Fixed Effects Regression Model*. [online] Medium. Available at: https://towardsdatascience.com/understanding-the-fixed-effects-regression-model-d2fccc2cc27e.

‌

‌

‌

Appendix:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Table 1: BANDS