Proceedings of the 2020 Annual General Donald R. Keith Memorial Capstone Conference
West Point, New York, USA
April 30, 2020
A Regional Conference of the Society for Industrial and Systems Engineering

Causal Threat Modeling Applied to the Horn of Africa

Steven Cromer[1], Connor McDonald[1], Thomas Monahan[1], Seth Shields[2], and Patrick DuBois[1]

[1]Department of Systems Engineering
[2]Department of Mathematical Sciences
United States Military Academy, West Point, NY

Corresponding Author: Steven.Cromer@westpoint.edu

Author Note: Cadets Cromer, McDonald, Monahan, and Shields are seniors at the United States Military Academy (USMA) participating in a year-long capstone design course under the direction of Major Patrick DuBois. In May 2020, they will commission into the United States Army as Second Lieutenants. The client for this project is the Defense Threat Reduction Agency-Joint Improvised Defeat Organization (DTRA/JD), headquartered in Washington, DC. The team is thankful for added guidance from the United States Africa Command (AFRICOM) and United States Agency for International Development (USAID).

Abstract: Initially developed to defeat the increasing threat of improvised explosive devices (IED) during the height of the Iraq War in 2003, DTRA/JD quickly evolved into the Department of Defense's (DoD) main effort in countering and reducing the effect of improvised threats. Following a suggestion from DTRA/JD about project leads, our team reached out to AFRICOM and began working on a problem narrowly tailored toward their mission. AFRICOM's strategic focus in East Africa and the complex situation involving refugees and internally displaced persons in the region require a systematic method to identify the most prevalent threats and their relationship with one another. This paper describes a method to leverage publicly available information (PAI) and K-Means Clustering to identify threats and model their interdependence using a Systems Dynamic model. The output will show the greatest threat to a region enabling a decision maker within AFRICOM to enact policy to reduce the overall threat level.

Keywords: DTRA/JD, improvised threats, East Africa, K-Means Clustering, Systems Dynamic model, PAI

1. Introduction

1.1 Background

Africa is a dynamic environment consisting of mostly developing countries. Economic, political, and environmental factors present consistent and damaging threats to the future of the continent. Due to the tumultuous nature of Africa there is severe instability, leaving a large opportunity for developed countries to pursue their political interests, working to guarantee increased development (Signé and Gurib-Fakim, 2019). Djibouti, Eritrea, Ethiopia, Kenya, Somalia, South Sudan, and Sudan make up the Horn of Africa. The United States' interests in the region consist of reducing terrorist activity, protecting exports such as oil and precious metals, and mitigating influence from near peers such as China (CIA,1978). An updated take on American policy in the region confirms the need to protect the economic interests of Sub Saharan African countries by countering the predatory pursuits of China (Brookings, 2018). Several of the countries in the Horn of Africa exist as breeding grounds for Boko Haram, the Islamic State, and Al-Qaeda. Attacks from these terrorist actors, as well as civil war within countries, often displace a significant percent of the population, whether intra or interstate (Campbell, 2019). Lack of development and infrastructure prevent mitigation of crime and increase the prevalence of water shortages, and lack of proper medical care contributes to the spread of disease within the populace. There are a significant number of threats outside of human control potentially affecting any populace. Any of these vulnerabilities, natural disaster for example, may present potential for grave harm to a region. These, however, are outside of the scope of mitigating policies. For this reason, this study on threat mitigation focuses on a limited number of threats within the locus of control for an intervening agent. This study limits its focus to disease, murder, access to clean water, education rates, economic growth, and overpopulation.

1.2 Research Question

Considering the interests of the United States, which vulnerability factors present the most harm to a populace? How can mitigation of these factors contribute to holistic threat reduction and support to American interests within a region?

1.3 Case Study Problem Statement

At present, the growth of, and access to publicly available information (PAI) is unprecedented. This presents the opportunity to use rapidly developing information to solve problems in ways previously unexplored by AFRICOM. The purpose of this study is to develop a method that leverages PAI, systems engineering, and advanced analytics to systemically identify, characterize, and assess threats to the United States' interests in a region. This paper will apply that methodology to identify prevalent threats within the Horn of Africa, determine intra-threat relationships, and analyze threat mitigation strategies. The Horn of Africa is a highly vulnerable region, making it a prime location for a case study of the application of this methodology.

2. Methodology

2.1 Design Thinking and Lean Startup Collaboration

Our research group created a hybrid project methodology, combining principles of Design Thinking with the Lean Startup (Ries, 2019). This hybrid approach enables innovative idea generation and iterative learning and development. Design Thinking, a human-centered problem-solving methodology, was used in place of the Lean Start-up's 'learn' and 'ideate' phases to better understand stakeholder needs and iteratively define the problem, generate innovative solutions, and prototype and test them for increased learning. Prototypes ranged from verbally explained ideas to coded solutions and tests were conducted with stakeholders regularly to assess the value of the ideas. Armed with a significantly better understanding of the problem and stakeholder requirements, the process shifted to the Lean Start-up cycles of build, code, measure, and data (Ries, 2019). During this part of the iterative process, the team reached out to clients (DTRA and AFRICOM) to gather feedback on the products we created. This feedback completed one iteration. We continue to follow the same process, increasing the complexity and accuracy of our analysis of regional instability. Each completed cycle takes time but improves the quality of the product. Initial iterations of the model occur often focused on quantity over quality, but through time and client feedback, this method leads to a more refined model that iteratively generates more value for the client (Figure 1).

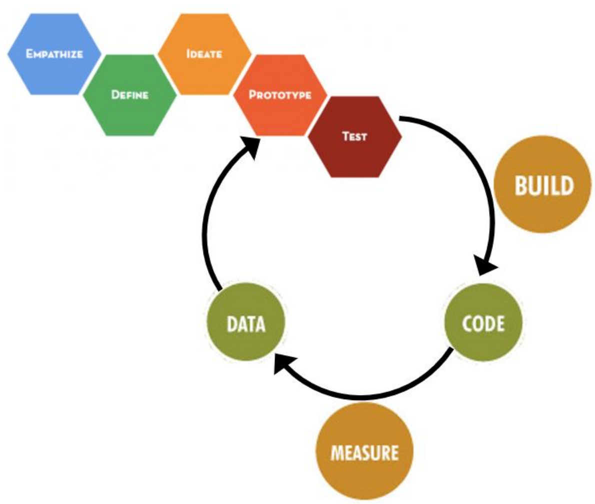

Figure 1. Design Thinking and Lean Startup Hybrid

ISBN: 97819384961-8-9

2.2 System Dynamics Modeling

System Dynamics modeling depicts the change of a system with variation in policies over time. One traditional system dynamics model is the stock and flow diagram. The stock and flow diagram consist of two primary components. The stock is considered a measurement of a variable at a certain point in time. The flow describes the inflow and outflow rates that vary the level of the stock. Combined with other variables, creating a stock and flow diagram for a system is an excellent way to depict the change of the system over time by varying conditions (Sterman, 2000). Figure 2 includes a representative portion of this project's stock and flow diagram.

2.3 Value Functions

Regularly used in multi-criteria value modeling, value functions enable an analyst to aggregate metrics with differing units and scales for comparison of competing alternatives. In value modeling specifically, value functions a mathematical way for converting all metrics of assessment to a singular unit of measure known as value (Parnell et al, 2011). Research on the Horn of African and stakeholder analysis allowed the derivation of each respective value function. The benefit of value functions is that after their initial formulation, entering raw data automatically creates a value standardization on a scale with unrelated data. In the context of this case study threats to refugee camps in the Horn of Africa were measured using a variety of metrics on varying scales. The use of value functions allowed the aggregation of these metrics into a single threat score.

3. Application

3.1 System Dynamics Modeling for the Horn of Africa

The model uses six different threats captured in stocks, showing their relationships with other stocks and certain variables present in the environment. Using the value functions, the level of the stock transforms into an aggregate threat score, which then combines with an overall threat weight to form the cumulative threat level for a specific country. These threat weights vary depending on their impact on the country overall. Research on environmental factors allows weighting that depicts the relationship between the threats on the country with respect to each other. Higher weights correspond to a larger impact on the country from a utilitarian viewpoint while lower weights correspond to less of an impact. We used a rank-weighting method to prioritize each threat based on relative importance and then assigned it a weight (Parnell et al, 2011). Table 1 is an example of weights used in this process. The outcome of the system dynamics model is aggregate threat level using the value functions of each factor.

Table 1. Global Weights

Threat	Weight
Disease	.1940
Population	.1758
Violence	.1211
Economy	.1576
Water Access	.2122
Education	.1393

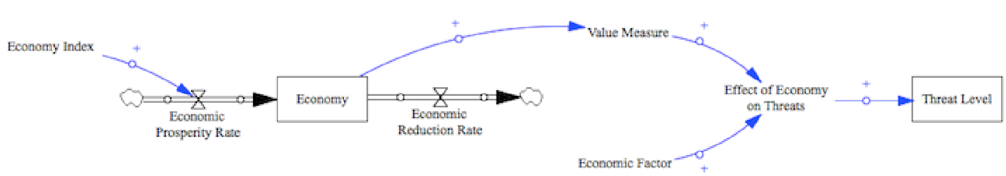

Figure 2. Stock and Flow Diagram for Economic Threat

3.2 Data Collection and Cleaning

Data sets used during this analysis came from open sources only. This simplified the collection process in some ways, but in others it increased the difficulty of obtaining enough data for each country and threat. Most of the open source data came from the United Nations High Commissioner for Refugees (UNHCR), the World Bank, and Comstat (COMSTAT Data Hub). Data collection from these sources consisted of GDP growth rate, percentage of population living with HIV incidence of malaria, secondary education enrolment, access to clean water, and population density. The organizations produced data through a combination of surveys and predictive analytics using small data sets. Due to lack of infrastructure in Africa there is a significant lack of data about some countries. This same challenge extends to comparable countries, for example a developing country within the Middle East, but would not be a challenge when examining a developed country. Statistical distributions enable prediction of future or missing values if a data set is sufficiently large.

3.3 K-Means Clustering Methodology

Following data amalgamation, K-Means Clustering determined which countries were the most affected by each threat. K-Means Clustering is a form of machine learning that uses the mean across the rows and columns of a data set to place subjects into k number of bins based on their distance from the mean. The algorithm typically starts by choosing k observations at random from the input data and assigns every other data point to k bins by minimizing the least square distance to the mean (Garbade, 2018). At each iteration it calculates the new mean of each of the k bins and reiterates through the algorithm, reassigning each of the points to one of k bins. When determining the value of k, it is important to consider the number of inputs being clustered. Our focus is the Horn of Africa which consists of seven countries. A value of $k = 1$ would produce the overall mean observation, and a value of $k = 7$ would place each country in its own bin, leading to meaningless results. For the sake of classifying our chosen threats we used $k = 3$ bins to classify them as high, mild, or low level for each country. Since the bin number only serves as a label, rather than a ranking, the average of each bin determined which bin corresponded to each group.

3.4 K-Means Assignment Analysis

Output from the K-Means Clustering provides an easily interpretable summary of the data. The algorithm clusters each threat into a high, mild, or low bin in comparison to the other threats within each country. Table 2 below shows this comparison across rows.

Table 2. K-Means Output of Threat Level Classifications

		Threat						
Countries	**FSI Rank**	**Economic**	**Disease**		**Education**	**Violence**	**Water Scarcity**	**Overpopulation**
Djibouti	43rd	High	Mild	Low	-	High	High	Mild
Eritrea	17th	Low	Mild	Low	High	High	Mild	Mild
Ethiopia	23rd	High	Mild	Mild	Low	High	Low	High
Kenya	25th	High	High	High	High	High	Mild	High
Somalia	2nd	-	-	-	-	-	-	Low
South Sudan	3rd	-	Mild	-	Low	-	-	Low
Sudan	8th	Mild	Low	Mild	Mild	Mild	Mild	Low

The level of threat within each country shows an initial direction for the weights used in the value function. Higher threat levels within each country correspond to a higher weight in the model. Datasets lacked sufficient data to cluster the threats for Somalia. The Fragility State Index (FSI) from the Fund For Peace (FFP) assesses the risk countries are at using multiple attributes. Countries at greater risk of state failure are closer to ranking first. The output derived from the K-means clustering differs from the conclusions drawn by the FSI. All seven countries are in the Top 50 Most Fragile States. Kenya ranked 23rd on the FSI and Sudan ranked 8th. The output of our model shows that Kenya's threat level is high for five of the six factors and Sudan was did not receive a high score in any. These differences come from incongruencies in datasets and the age of the data in our model. Improvements can be made to the model with more current and accurate data.

3.5 Data Visualization

Part of the solution design involved conveying the results in a meaningful and understandable fashion. To begin with, our team used the open source data visualization tool RStudio with additional packages (such as Leaflet, a library that allows users to convey data onto a variety of visuals, such as geographic maps or scatterplots) to take the data used in the K-Means Clustering algorithm and classified the threat level: high, mild, or low, for each country in the Horn of Africa. Much of the data were specific to just the country level, with violent events being the only threat to have a finer geographic precision of latitude and longitude, as shown in Figure 3 (AFDB Socio Economic Database, 1960). This lack of precision limited the level of detail that the graphic could show, with the only variation being between countries. However, it is still able to show visual insight into the nature of the problem in which we are concerned. Just like the K-Means Clustering model, the visual model can display time dependence as well. Using the Shiny package in RStudio, a slider gives the user control to display data by year. This feature added a layer of depth to the understanding of the issues. It is a way to not only compare the quantitative threat values between countries, but also parametrically over time.

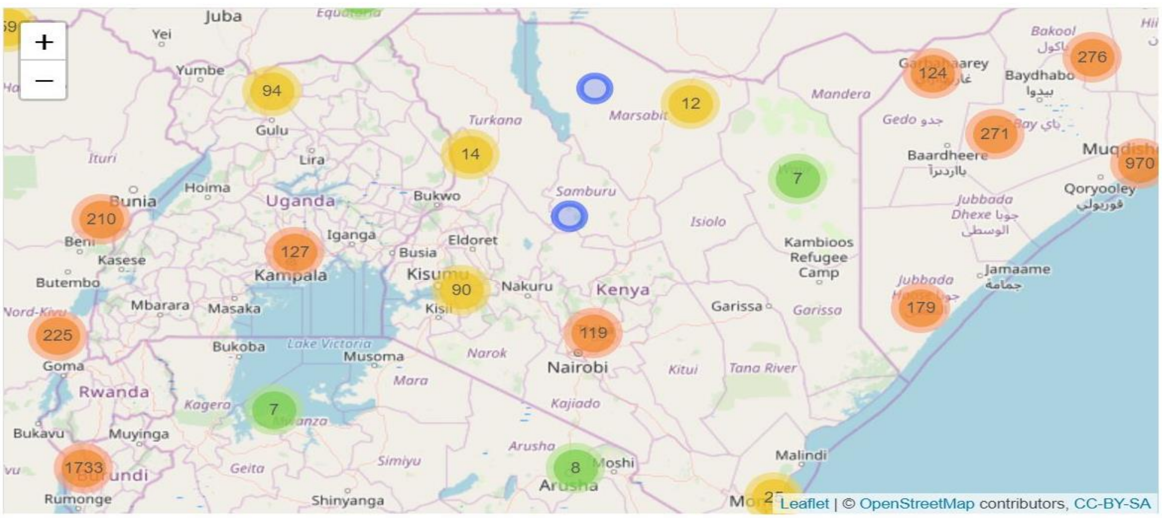

Figure 3. RStudio Shiny App displaying Violent Event Groupings in Uganda, Kenya, and part of Somalia for 2015

4. Recommendations

The end goal for each of the models is that they serve some organization as an informative tool for prescribing policy. Ideally, we would like to put an end to the threats and dangers that people in the Horn of Africa face. Finite resources limit the overall affect external actors can have on the region. Knowing how and where to allocate these resources so that the greatest impact may be made is difficult, but the process may be alleviated through research-backed courses of action. The output from the K-Means Clustering algorithm provides the user with groupings of countries based on their relative threat imminence. The systems dynamics model compares the relationships between threats within each aggregate threat score. However, the value of the model comes from the different country profiles by examining and comparing the threat impacts between different factors. Stakeholder beliefs develop specific weighting criterion for comparing threats against one another. We recommend using complete data sets to complement our system dynamics model to rank the overall threat level of countries.

5. Conclusion and Future Work

Examination of threats begins with stakeholder analysis. A stakeholder, such as the United States or any one of its agencies, must determine its interests within a region, and since relationships between threats can vary by region, the model must change to reflect the attributes of the given region. After finding these interests, the next step is an examination of

ISBN: 97819384961-8-9

vulnerabilities faced by the region of choice. Open source data served as a limitation to the scope of our study, but a stakeholder with greater privilege could, and should, consider collection through other sources.

Clustering the threats within countries gives an initial assessment on the prevalence of each threat. This translates to the eventual weight of each threat. The weakness of this analysis is that the K-means algorithm is incapable of discerning between the differing scales of each vulnerability. The value function closes the gap in this capability, and the systems dynamic model becomes a framework for examination of the overall effect of reducing one or more threats provides a method for assessing competing policies over time. Other regions of interest to the United States can be examined using this framework. Our research group will draft potential policy recommendations and analyze the policy's effect on overall threat level. Given more time and access to more privileged information it would be interesting to develop a regression to produce more data for each country.

The largest improvement in this project would come from filling in the voids of data for countries that have little to no information. Compiling complete data is a great endeavor and it could take years to derive a dataset that was complete and large enough for use with a project of similar scope. The methodology could also be derived from regions of the world with similar issues but more complete sets of data. An interesting and useful follow on study could compare the relationships between threats for two separate regions, which would serve to validate both models.

A future application of this process may examine threats in the Middle East. As the United States pulls out of Afghanistan militarily, it may use this process to provide threat mitigating policy recommendations to the Afghan government. Allowing the local government to reduce the impacts felt by its populace due to various vulnerabilities may increase regional stability and the sustainability of peace.

6. References

AFDB Socio Economic Database, 1960-2019 - COMSTAT Data Hub. (n.d.). Retrieved from http://comstat.comesa.int/wiqcbkg/afdb-socio-economic-database-1960-2019?tsId=1880770.

Bair, J. (2018). *IMPROVISED THREATS: Warfighter Support Maintained, but Clearer Responsibilities and Improved Information Sharing Needed* (GAO-18-509). Government Accountability Office.

Campbell, J. (n.d.). U.S. Interests in Africa. Retrieved from https://www.cfr.org/project/us-interests-africa

COMESA Brief Overview - COMSTAT Data Hub. (n.d.). Retrieved from http://comstat.comesa.int/tcagt/comesa-brief-overview

The Fund For Peace (2019). Fragile States Index 2019. Retrieved from https://fragilestatesindex.org/

Garbade, M. J. (2018, September 12). Understanding K-means Clustering in Machine Learning. Retrieved from https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1.

Parnell, G.S., Driscoll, P.J., & Henderson D.L. (2011). *Decision Making in Systems Engineering and Management* (2nd ed.). John Wiley & Sons.

Ries, E. (2019). *The lean startup: how constant innovation creates radically successful businesses*. London: Penguin Business.

Signé, L., & Allen, N. D. F. (2018, March 27). Trump's Africa policy takes form with focus on security (and China). Retrieved from https://www.brookings.edu/opinions/trumps-africa-policy-takes-form-with-focus-on-security-and-china/

Signé, L., & Gurib-Fakim, A. (2019, January 25). Africa is an opportunity for the world: Overlooked progress in governance and human development. Retrieved from https://www.brookings.edu/blog/africa-in-focus/2019/01/25/africa-is-an-opportunity-for-the-world-overlooked-progress-in-governance-and-human-development/

Sterman, J. (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World*. McGraw-Hill Education.

U.S. Economic Interests in Africa. (1973). Intelligence Handbook. Retrieved from https://www.cia.gov/library/readingroom/docs/CIA-RDP79S01091A000300050001-3.pdf