Many groups face the challenge of trying to make evidence-based decisions about threats such as terrorism. Resource allocation by countries for security and counterterrorism activity is a well-known challenge. While many countries keep this information extremely secret, the USA has had its own methods reviewed publicly by a number of respected bodies, such as the National Academy of Sciences. As recently as 2001, the US National Commission on Terrorism (National Research Council, 2011). The UK and the Netherlands have also had their own national risk register processes reviewed in the open literature. Commensurate with some of the major countries, the UK Government has also had to publish its own academic and commercial forecasts of terrorism. While some catastrophic terrorism models exist in the market it has been regularly asserted that government backtracking is required because of the number of disastrous failures in traditional forecasting models.

AIC

The aim of this paper is that existing quantitative data can better forecast terrorist activity and inform resource allocation if structured correctly. Over reliance on elicitation techniques and probability trees opens up assessments to a variety of heuristic errors and biases. A stronger quantitative model would anchor assessments reducing these errors. This in turn would allow greater effectiveness in fighting terrorism and in the approach to assessing the relative merits of countermeasures. Before this improved approach is viable though some initial hypotheses must be tested.

Hypothesis 1 (H1): Terrorism data either does not exist, or is of poor quality for making forecasts.

Hypothesis 2 (H2): It is not feasible to make forecasts on the basis of this data, with confidence intervals and understanding of error margins.

Method

In order to fulfill this requirement the main open source databases of historical terrorist incidents were examined and compared. Particular focus was given to the presence or absence of codebook, lineage, time period covered and collection/methodology. Based on this, a period of time was defined for which the most databases had coverage in order to compare their records of terrorism. This cross comparison was focused on UK data due to the UK’s clear, public counter terrorism strategy & national ambition. In further studies it may be useful to compare parts of the US with the UK systems because they also have available public counter terrorism databases (HGS; 2007; HMG; 2016; HMG; 2006).

The databases used for this initial pilot were:

RAND Terrorism database (RTD). A think tank funded and administered database of terrorism incidents.

Globelaw Terrorism database (GTD). An amalgamation of the records of the Pakistan agency terrorism database which was purchased by the University of Maryland, and 21 other separate projects. DHS funding has paid for its creation by merging all the sources and its purpose was to create a better database from which to start empirical analysis of terrorism.


Movtency database of MI6 incident. An academic private access resource that tracks current and over 100 years history of CBIN incidents, plots and hoaxes.

EuropeanSituation and trend report. As the name suggest and annual report produced on the trends of terrorism in the EU alongside figures for counter terrorism activity in order to show the efficacy of certain measures.

These were selected as being the largest part processed (i.e. not raw press or reporting data) databases. There are other sources available specialized to types of attack and region or groups. These will be reviewed in the creation of an improved model in later phases of the study.

Terrorism

The best data source available was identified using the analysis of the data sources. This was then imported into a standalone database to allow manipulation and statistical analysis.

Step 1: A new database was created for each geographical region being studied

Step 2: A query was run to select only the data for the selected time period.

Step 3: A further query generated 3 more data sets for each region:

† All incidents

‡ Incidents that killed at least 1 person

§ Incidents that killed more than one person

Step 4: An analysis was run on each data set to assess its suitability for forecasting. This included autoregressive models and moving averages and stationary time series analysis. Some models were optimised using the data set, others such as the moving averages were limited to only being able to predict after a number of years. Only the optimum models are included in this paper for brevity – these being the 3 and 6 year moving averages.

Comparison of terrorism databases

The fig 1 below of the number of recorded actual attacks by the five-geographies terrorism databases shows the variation in data across the same time period in the databases for the UK. Fig 2 shows a comparison of the databases by attributes. The GTD and RAND raw counts are based on the original 1975 recording that has persisted since 2007 when they were at their peak funding point, which has recently reviewed for the GTD. Variation is on the whole down to the definition of terrorism used (or not) as the high level of media coverage for the GTD. Variation is on the whole down to the definition of terrorism used (or not) as the high level of media coverage for the GTD.

The databases are heterogeneous it was not expected to yield good results. However when looks at attacks that killed more than 1 person demonstrate some interesting results for the accuracy of a single year prediction. While any level of inaccuracy may feel unacceptable for security policy decision-makers it appears that there may be further for exploration for policy and insurance forecasting purposes with acceptable levels of [precision].

Further research

Defining terrorism frequency

The most immediate problem of any database that attempts to capture and record incidents of terrorist activity is to define what a terrorist incident or terrorist attack actually means. There exists no internationally recognised term, but there have been many attempts over time to define these terms. APTD in the database review it is suggested to use the UK legal tests from the 2000 Terrorism Act. This fulfills the initial objective of creating a database which is appropriate for UK terrorism modelling and forecasting to allow the international community to build on.

Any multi-geographical database will automatically run in to the problem of whether to define the inclusion of an incident by its own countries definition or by the attacked countries definition. A prime example of this is the Japanese did not consider Aum Shinrikyo to be a terrorist group but an extremist religious group. The US, and others, refer to the group attacks as terrorism though, due to their own definition. Sometimes these definitions are by exclusion. i.e. Aum’s attacks weren’t war; nor were they simply crime, therefore they were terrorism. So in any analysis of historical data it make more sense to begin the analysis on a country by country basis in order not to confuse definitions.

The next problem comes in the definition of an attack or incident itself. Some databases include only ‘successful’ attacks, i.e. on attack which is delivered to the target although it may not function as well as hoped. Others may include plots, threats and hoaxes but in the first two instances these are often the hind sight reporting of evidence coming out in court and can be years old. The Rajnavillaw use of patroes to contumaciously assassinate in the US, and other plots, were only discovered years after they came out as part of another court case. All this information can be indicative of group intentions to carry out attacks, and which type of attacks and targets they might choose so recording it in a useful manner which can be compared with data in other countries is vital.

Defining Attack types

The study showed large variance in the number of attack types within this databases. This is further complicated with the inclusion of more specific details about the weapon used within an attack. A newly constructed database will need to ensure appropriate selection of attack types. A good starting point could be basic attack types on previous DHS analysis (DHS National Planning Scenarios, 2006) and FEMA guidance manuals (FEMA, 2010), combined with analysis of the history of incidents and plots (US Office of Cabinet Affairs, 2006; The DHS planning scenarios. It particularly describes potential terrorist attack types based on a combination of open source information and counter intelligence. Worth noting is the ‘inception’ of the International Roadmap for Security by the Department of Homeland Security (DHS) and the National Counterterrorism Center (NCTC). The reference manual (2011) and John Tucker of MIT (Tucker, 2005). Combining these sources of information on attack types with the attack types included within the databases can generate a better structure related to the effective impact of the attack.

Defining Target Types

There is an established government principle in generating lists of threatened targets which are often grouped by type. This is in use by the US StateDef. The National Asset Database, 2007, and the UK (UK Cabinet Office, 2010; HMG, 2006). This is it only meaningful to separate categories for targets if there is a shared characteristics within that group that is distinct from another group.

Conclusion

Hypotheses 1 and 2 were disproved although only partially. The quality of terrorism databases could and should be improved and necessary enrichments and transformation.