Bangor Business School
Working Paper

BBSWP/19/17

A risk-mitigation model driven from the level of forecastability of Black Swans: prepare and respond to major Earthquakes through a dynamic Temporal and Spatial Aggregation forecasting framework

By

Konstantinos Nikolopoulos[a, 1] **, Fotios Petropoulos** [b]**, Vasco Sanchez Rodrigues**[c]**, Stephen Pettit**[c]**, Anthony Beresford**[c]

[a] *for***LAB**, Bangor Business School, Prifysgol Bangor, Bangor, LL57 2DG, Wales, U.K.

[b] *Information, Decision and Operations Division, School of Management,University of Bath*

[c] *Logistics and Operations Management Section, Cardiff Business School, Cardiff University*

September, 2019

Bangor Business School
Bangor University, Hen Goleg, College Road
Bangor, Gwynedd LL57 2DG
United Kingdom
Tel: +44 (0) 1248 382277
E-mail: business@bangor.ac.uk

[1] Corresponding author: kostas@bangor.ac.uk

A risk-mitigation model driven from the level of forecastability of Black Swans: *prepare and respond to major Earthquakes through a dynamic Temporal and Spatial Aggregation forecasting framework*

Abstract

Major earthquakes are black swan, or quasi-random, events capable of disrupting supply chains to an entire country, region or even the whole world as the case of the Fukushima disaster profoundly demonstrated. They are amongst the most unpredictable types of natural disasters, and can have a severe impact on supply chains and distribution networks. This research develops a supply chain risk management model in the anticipation of such a black swan event. The research considers major earthquake data for the period 1985 – 2014, and temporal as well as spatial aggregation is undertaken. The aim is to identify the optimum grid size where forecasting variance is minimized and forecastability is maximized. Building on that a risk-mitigation model is developed. The dynamic model – updated every time a new event is added in the database - includes preparedness, responsiveness and centralization strategies for the different levels of time and geographical aggregation.

Keywords: Risk, Black Swans, Forecastability, Statistical Aggregation, Disaster Relief

A risk-mitigation model driven from the level of forecastability of Black Swans: *prepare and respond to major Earthquakes through a dynamic Temporal and Spatial Aggregation forecasting framework*

1. INTRODUCTION

The prediction *per se*, but most importantly the degree of predictability, of major earthquakes has long been the subject of intensive research. The seminal works of Lane (1966) and Whittow (1980), for example, highlight the semi-predictability of earthquakes, showing that they occur intermittently over long periods of time with a tendency not to cluster into short time periods. Nevertheless, the intensity of particular earthquakes, or of an individual earthquake, is very hard to predict (McKenna, 2011). Nonetheless, the models we have at our disposal offer opportunities to significantly improve our understanding of risk mitigation and to drive research in the best possible direction to anticipate earthquake events. This is an essential topic in Humanitarian Logistics and OR/MS can help by providing models to help take informed decision before and when such situations arise. These models need to be dynamic so as to be fed with data as and when they become available – more data for pre-specified observable variable, but also data for more dimensions than originally envisaged. (He and Zhuang, 2016)

Major humanitarian challenges are usually driven by catastrophic, and, to a large extent, unexpected events. These events are quasi-random and, as such, their occurrence has been likened to the principles of black swan events (Taleb, 2007). Taleb's book "The Black Swan" starts with:

> *'Before the discovery of Australia, people in the Old World were convinced that all swans were white, an unassailable belief as it seemed completely confirmed by empirical evidence' (2007, pp. xxix)*

In forecasting, three types of 'swan' can be identified. First, normal events, usually referred to as 'white swans', represent the mainstream of the data and their occurrence is more or less predictable using well-established statistical models. However, "*the normal is often irrelevant…*" (Taleb, 2007, pp. 36), the greater costs and opportunities are usually associated with the abnormal. So, there are the 'grey' swans, which are rare but expected. "*Somewhat predictable, particularly to those who are prepared for them and have the tools to understand them*" (*ibid*, pp. 37). Last, there are the 'black swans': the unexpected and almost unpredictable "*fewer remote events, but more and more extreme in their impact*" (*ibid*, pp. 38). In trying to predict the appearance of black swans, three questions are of particular importance:

- When is the next black swan going to be observed?
- Where will the next black swan occur?
- What will be the impact of the next black swan event on the probability of subsequent sightings?

But are black swans truly unexpected, or is it all a matter of perspective and perception (Taleb, 2007)? Black swans might be just 'peaks over threshold' (Leadbetter, 1991). However, what is critical is the form and quality of data available for identifying these peaks, in other words maximizing prediction power? The problem therefore is to identify the appropriate lens through which data should be explored and analyzed. The forecasting horizon and width (how far ahead organizations look or plan and over how broad a horizon they scan, and for which geo-region) are usually parameters imposed by the internal or external environment. But is this the horizon and width where the forecasting accuracy is maximized, and thus the uncertainty that prevail in our forecasts and therefore in our decision-making is minimized?

If this argument is followed, the forecasting/foresight horizon and width can be rethought and considered in the model application, in other words, re-thinking the *when* and *where* of strategic planning. One obvious solution would be to match the when and where of strategic decisions to the forecasting horizon and width that minimizes the respective uncertainty and maximizes the forecasting performance. This becomes more important in the presence of black and grey swans, where uncertainty is confronted and even the slightest improvements in forecasting accuracy could prove to be critical (Powell *et al*, 2016).

A particular case of black swans with considerable potential impact is the case of earthquakes. Can earthquakes be predicted or, taking the focus of this research, can a better forecast be made of areas where earthquakes are most likely to occur? Questions such as these have been asked for the last 2,500 years, since Archimedes described the intermittent nature of earthquake occurrences. It is currently contended that the exact timing, location or impact of an earthquake cannot be predicted (McKenna, 2011). Provision of categoric earthquake forecasts to a specific location (city/region) and a narrow time interval (week or even day) is impossible. However, while there are regions that are more seismically active, for example based on plate tectonic movements, it is much more probable that an earthquake of magnitude 5 or greater will occur in Greece compared to UK, even if such areas are focused on, it is still not possible to accurately predict the exact timing or the impact of an earthquake. Therefore, if the exact location and timing of an earthquake cannot be predicted, what actions can be taken? In summary, at which scale are earthquakes the least predictable, and conversely, at which scale are they most predictable?

In recent years there have been several developments in the area of data aggregation. These developments have been in both a temporal context (Nikolopoulos *et al*. 2011; Kourentzes *et al*., 2014; Petropoulos and Kourentzes, 2014; Petropoulos and Kourentzes, 2015; Athanasopoulos *et al*., 2015), and a spatial context using hierarchies of data, and respective forecasting approaches (Hyndman *et al*., 2011). Thus models now exist that could help to identify the most predictable temporal and spatial windows that could then drive our respective strategic decisions. The most useful property of this family of models is that they are dynamic in nature as they are driven by responsive, continuous updates of the empirical databases. Thus the impact of such events can be addressed by a vector of variables that can suggest more options for what drives the decisions.

With reference to impact, one approach is to try and improve existing response systems so that communities are better prepared should an earthquake occur. In order to improve such systems, it might therefore be possible to use aggregation in terms of both time and geographical regions in order to establish the optimal levels of positioning and stock volumes that are used for strategic planning. In this research, various temporal and spatial aggregation scales are evaluated with specific reference to earthquake event prediction. Section 2 provides a literature review focused on earthquake preparedness and disaster response. In Section 3 the experimental design and empirical results of a forecasting exercise employing real-life data and various aggregation strategies are presented. In Section 4 the proposed model is discussed while in Section 5 the implications for theory and practices of RM/MS and humanitarian logistics are demonstrated Lastly, in Section 6, broader inferences concerning emergency decision-making and strategic planning are made, conclusions are drawn and insights into possible areas of future research are provided. In this sense, the research is intended to inform the relevant organizations, in broad terms, when and where to pre-plan, and in which situations to invoke a response strategy, the key question being: which strategy is likely to be more effective overall?

2. LITERATURE REVIEW

In recent years, academic reviews of humanitarian aid and emergency relief logistics have been elevated from essentially descriptive and observational (Pettit and Beresford, 2009; Kunz and Reiner, 2012; Kovacs and Spens, 2011) to methodological and analytical (Naji-Azimi et al, 2012; Paul and MacDonald, 2016; Powell et al, 2016). The rapid growth in academic interest in the applied field of humanitarian aid and emergency relief logistics, as well as adding energy to the debate, has increased its scale and scope.

The frequency of occurrence of natural disasters in recent decades has led to a growing awareness of their impact on communities and society in general. This, in turn, has triggered increased interest in modelling the predictability of the events themselves, and assessing the degree to which impact can be mitigated by improved levels of preparedness or better responsiveness. Galindo and Batta (2013) and Gutjahr *et al* (2016), for example, have reviewed the growing body of literature in the operational research field which has focused on humanitarian aid distribution or emergency relief provision. It is suggested that, although modelling has become more sophisticated and increasingly granular, the underlying pattern of research has not significantly changed. Management of disasters in general terms has persisted as one of the main research threads (see, for instance Edrissi *et al*, 2013) and a second thread has followed a case approach looking at, for example, Brazil (Alem *et al*, 2016), Iran (Tofighi *et al*, 2016) or Turkey (Kilci *et al*, 2015). A third branch of research embraces cross-cutting studies such as that by Ozdamar and Ertem (2015). These embrace several dimensions which include organizational as well as operational parameters. They typically focus on the importance of taking an integrated approach in order to fully understand uncertainty. The papers referred to above, endeavour to make sense of, and parameterize, a range of challenges which are either implicitly, or explicitly, an integral part of the humanitarian logistics problem in different circumstances.

2.1. Earthquake Preparedness

The goal of emergency response is to provide shelter and assistance to the victims of disasters as soon as possible after an emergency occurs. Pre-positioning of key supplies at strategic locations is essential in ensuring their availability both when required and for faster response (e.g. Rawls and Turnquist 2010; Balcik *et al*. 2010). It has been suggested that in the long run such an approach aids in the reduction of the cost of deliveries to those locations due to regular replenishment (Gatignon *et al*. 2010). Many studies have addressed the importance of the preparedness phase and the need for pre-positioned warehouses in humanitarian relief logistics, whereas only a small number of papers are related to the specifics location decision (e.g. Rawls and Turnquist 2010; Campbell and Jones 2011; Roh *et al*, 2015). Gatignon *et al.* (2010) illustrate the implementation of a decentralized model at the International Federation of the Red Cross using the pre-positioned warehouse concept. Campbell and Jones (2011) use a cost model to examine the prepositioning of supplies and the volume of goods in preparation for a disaster. Nevertheless, where the above studies discuss the optimal location based on a single criterion (e.g. minimum total costs), the assessment process for strategic decision-making often involves several attributes, and it is usually necessary to make compromises among possibly conflicting tangible and intangible factors (Onut and Soner 2007).

The multi-criteria decision-making (MCDM) approach has been widely adopted as a tool for optimizing the location of stocking points for emergency relief goods (see, for example Roh *et al*, 2015). However, where and when an emergency event might occur has been considered less frequently, yet is a very important part of effective emergency response. Prediction of major events in terms of their timing, location and intensity form the focus of the research in this paper. In specific terms, therefore, the research gap is addressed by specifically considering the overall pattern of humanitarian relief organizations' strategic stock locations in both international (macro) and local (micro) contexts in relation to the historic pattern of earthquake occurrence at a global scale.

Distribution/logistics centre attributes have been discussed by, for example, Li *et al*. (2011) who highlight the importance of parameters such as accessibility, security, connectivity, costs, and proximity to customers and suppliers as key to successful logistics. Although this research was in the context of commercial operations, in principle all of these measures are transferrable to the humanitarian sector. If these measures are superimposed on robust event forecasts, their value is maximized.

Locating a pre-positioned warehouse near to the beneficiaries and potential disaster location potentially reduces delivery time and cost, but more importantly, it has the potential to save lives. The geographical location of the warehouse does not have to be near the disaster prone area, but rather it could be in the headquarter country or next to a regional office for strategic reasons. Proximity to beneficiaries for a potential warehouse is thus one of the important considerations and can be viewed in a similar way with proximity to disaster prone areas. Critical to the question of locating emergency response depots, and hence materials, is having the best possible understanding of the probability of earthquake occurrence as measured by its location, timing and intensity. This can be viewed as a three-dimensional construct involving X, Y and Z variables which can be assembled into a three-dimensional model.

There is substantial literature on probability forecasting which, though mostly outside earthquake prediction, is useful for improving understanding of such three-dimensional models. In the context of weather forecasting, for instance, three-dimensional models are common and outcomes are in the form of probability forecasts (Palmer, 1999). Central to the application of probability is the level of aggregation of data on both temporal and spatial scales. An example of this is the UK Meteorological Office which has developed techniques to understand such uncertainties, called ensemble forecasts. In this forecasting procedure, simulations are run many times rather than just once, with very slight differences in the inputs in order to slightly vary the starting conditions. The range of outcomes thus generates a measure of confidence or certainty in the overall forecast (Met Office, 2016). While using ensembles gives an indication of certainty / uncertainty it also creates a problem in communicating the results. The main issue being: how high is the confidence about certain (likely) outcomes in relation to the low confidence in (unlikely) outcomes of low probability?

The key measures in the case of earthquakes, and therefore the parameters of concern for forecasting them, are: location of occurrence (epicentre), magnitude (or power), duration, depth of the disturbance and proximity to areas of population; this last parameter largely determines the impact of the event expressed in terms of material damage or loss of life. The United States Geological Survey (USGS) National Earthquake Information Center (NEIC) estimates that over a million earthquakes occur in the world each year (NEIC, 2016). Many have no impact because they occur in remote areas which are virtually uninhabited and beyond the reach of detecting mechanisms. Table 1 details the estimated frequency of earthquakes worldwide by annual average, according to magnitude, while Table 2 details the number of earthquakes recorded annually from 2000 to 2012 according to magnitude.

Table 1. Annual estimate of earthquake occurrence by magnitude

Descriptor	**Magnitude**	**Annual average**
Great	8 or higher	1
Major	7–7.9	17
Strong	6–6.9	134
Moderate	5–5.9	1,319
Light	4–4.9	c. 13,000
Minor	3–3.9	c. 130,000
Very minor	2–2.9	c. 1,300,000

Source: NEIC (2016)

Clearly, as the scale of earthquake analysis reduces, the more challenging the forecast of ‘when, where and how strong’ becomes. That is to say, the more precisely the location of a potential earthquake is stipulated, the less likely the event forecast is likely to be correct. At a global scale, the total number of earthquakes is reasonably constant, but the predictability of the major earthquakes, especially at a granular level where approximate locations are specified, is low. Although earthquakes of magnitude and 6 and above are relatively predictable by annual average frequency, earthquakes of magnitudes from 2 to 5.9 are much more variable in terms of frequency per annum. Earthquakes of below 2 magnitude are so small that they are often not detected; these can be neglected and omitted from any analysis as their impact is negligible.

Table 2. Number of Earthquakes Worldwide for 2000 - 2012 Located by the US Geological Survey

Magnitude	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
8.0 to 9.9	1	1	0	1	2	1	2	4	0	1	1	1	2
7.0 to 7.9	14	15	13	14	14	10	9	14	12	16	23	19	11
6.0 to 6.9	146	121	127	140	141	140	142	178	168	144	150	185	96
5.0 to 5.9	1344	1224	**1201**	1203	1515	1693	1712	2074	1768	1896	2209	2276	1295
4.0 to 4.9	8008	7991	8541	8462	10888	13917	12838	12078	12291	6805	10164	13315	8710
3.0 to 3.9	4827	6266	7068	7624	7932	9191	9990	9889	11735	2905	4341	2791	2174
2.0 to 2.9	3765	4164	6419	7727	6316	4636	4027	3597	3860	3014	4626	3643	2721
1.0 to 1.9	1026	944	1137	2506	1344	26	18	42	21	26	39	47	34
0.1 to 0.9	5	1	10	134	103	0	2	2	0	1	0	1	0
No Magnitude	3120	2807	2938	3608	2939	864	828	1807	1922	17	24	11	6
Total	22256	23534	27454	31419	31194	30478	29568	29685	31777	14825	21577	22289	15049

Source: NEIC (2016)

In order for aid agencies to be as well prepared as possible to implement relief operations, it is clear that any improvement in the understanding of where and when events are likely to occur would improve both locations of pre-positioned warehouses, and, from that, the speed of response. Agencies such as the United Nations High Commission for Refugees (UNHCR) already have pre-positioned warehouses which respond to all forms of crisis (UN, 2015). While this paper only considers the most important locations and only in relation to earthquakes, it is recognized that further development of the research to include other disaster types will improve the locational precision of the work and widen the value of the research.

3. EMPIRICAL EVALUATION OF AGGREGATION STRATEGIES

3.1. Experimental design

In order to identify the optimal aggregation levels for predicting earthquakes, the Significant Earthquake Database[2] is used. This database contains information on destructive earthquakes which meet at least one of the following criteria:

- Moderate damage (approximately $1 million or more)
- 10 or more deaths
- Magnitude 7.5 or greater
- Modified Mercalli Intensity X or greater
- The earthquake generated a tsunami

This research focuses on earthquake events of the 30 year period, 1985-2014. The focus is on three primary variables to evaluate the forecast accuracy of different aggregation levels: the occurrence of an earthquake, and the number of deaths and injuries caused. At the time the data were extracted from the database it was only partially complete in respect of data about the exact epicenter (longitude and latitude) of an earthquake, and the number of deaths and injuries. Thus a manual search for the missing data was conducted in order to populate the missing fields in the database. On completion of this exercise the dataset now offers the best opportunity for analysis as the information is consistent, containing exact latitude and longitude of earthquake epicenters and precise dates of occurrence. At this stage, account is not being taken of the respective earthquakes' magnitudes other than by setting the threshold as detailed above; the earthquake magnitude therefore acts as a 'qualifier' for the dataset. Similarly, account is not being taken of earthquake duration or depth below surface, although these have long been acknowledged as major determinants of the severity of earthquakes and of their impacts, especially as judged by the number of casualties (see, for example, Whittow, 1980; Carter,

[2] Accessed through: https://www.ngdc.noaa.gov/nndc/struts/form?t=101650&s=1&d=1

1999, UN, 2006). More recent work, e.g. the research of Paul and MacDonald (2016), also highlights the importance of earthquake magnitude, and number of foreshocks and aftershocks, as key determinants of material damage and the number of casualties.

Five temporal and five geographical levels of aggregation are considered. These are depicted in Table 3. The data are then re-aggregated in order to take into account all possible combinations for temporal and geographical aggregation levels. In total, therefore, there 25 different aggregation levels. The data (earthquake events) are visually depicted in Figures 1 to 5 for different levels of geographical aggregation. Regions shaded in grey indicate that at least one earthquake has occurred over the thirty-year period. Degrees of latitude and longitude are used to fix the size of the geographical grids.

Table 3. Earthquake event aggregation levels

Temporal Aggregation	Geographical Aggregation degrees (^{o})
Monthly	10^o x 10^o
Quarterly	$30^o \times 30^o$
Yearly	$60^o \times 60^o$
Three-yearly	$90^o \times 90^o$
Five-yearly	$180^o \times 180^o$

To evaluate the suitability of these aggregation strategies, a small-scale forecasting exercise is performed. The first 25 years of data (corresponding to 5 upto 300 data points, depending the level of temporal aggregation) are used to produce forecasts for the next 5 years (60 months). Forecasts are produced using the Simple Exponential Smoothing method where the parameters (alpha smoothing parameter and initial level) are optimized. The occurrence of an earthquake, the number of deaths and the number of injuries for a region (as specified by the geographical aggregation windows) are forecast. Forecasts are generated at the respective aggregation level; e.g. we use the quarterly – 60 x 60 degree ($60^o \times 60^{o)}$ data to produce 20 (5 years × 4 quarters) point forecasts referring to predictions of earthquake events over geographical areas of $60^o \times 60^o$. Subsequently, all predictions are then disaggregated to the most granular level considered in this study (monthly frequency - 10^o x 10^o regions) so as to be able to evaluate all strategies equally.

Temporal disaggregation takes place assuming equal weights. For example, the yearly forecast is equally distributed in 12 monthly forecasts. This assumption makes sense, as it would not be expected that earthquakes event occurrences have seasonal and/or trend patterns. Geographical disaggregation is employed using the top-down hierarchical approach (e.g. Gross and Sohl, 1990; Fliedner, 1999).

Disaggregation weights that are directly calculated from the historical averages of the bottom-level series are selected.

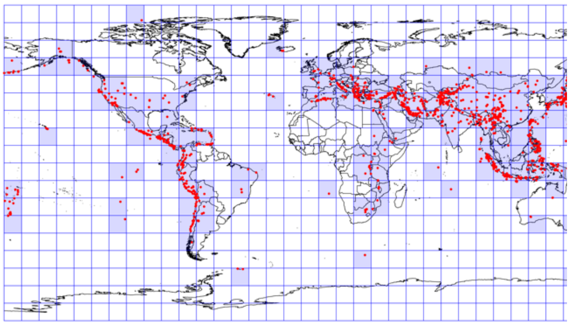

Figure 1. Empirical data aggregated in 10º × 10º geographical regions

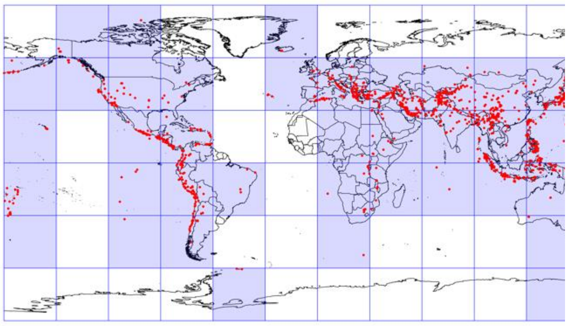

Figure 2. Empirical data aggregated in 30º × 30º geographical regions

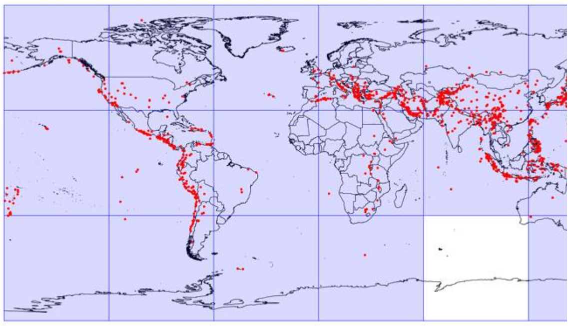

Figure 3. Empirical data aggregated in 60º × 60º geographical regions

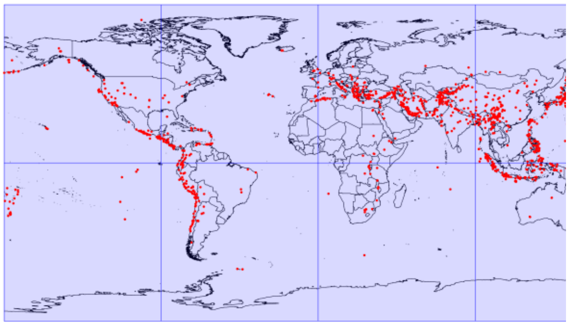

Figure 4. Empirical data aggregated in 90º × 90º geographical regions

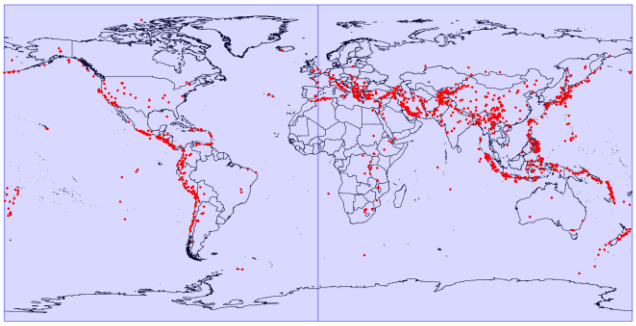

Figure 5. Empirical data aggregated in 180º × 180º geographical regions

The forecasts are then contrasted with the withheld actual events of the remaining 5 years of data. The comparison of the different strategies is based on the scaled Mean Absolute Error (sMAE), which is appropriate for measuring accuracy. This measure is based on the absolute scaled error, which is the absolute error scaled by the arithmetic mean of the in-sample data:

$$absolute\ scaled\ error = \frac{|Y_{t+h} - F_{t+h}|}{\frac{1}{n}\sum_{i=1}^{n} Y_i}$$

where Y_{t+h} is the actual h-steps-ahead from the forecast origin and F_{t+h} is the respective point forecast. sMAE is derived as the simple average (arithmetic mean) over horizons (months, 1 up to 60) and series (regions of area 10^o x 10^o).

3.2. Empirical results

Figures 6, 7 and 8 present the empirical results of the forecast accuracy for earthquake events, the number of deaths and the number of injuries respectively. For each figure, the accuracy is presented in 25 different panels representing the 25 combinations of temporal and geographical aggregation levels that have been considered. The results are presented visually using a grey scale gradation as follows: black - least accurate, dark grey – less accurate, light grey – more accurate and lightest grey – most accurate. Areas of white on the maps signify no data i.e. where no earthquakes have been recorded, or where the data are so scarce or fuzzy that the level of confidence in the data approximates to zero[3].

Each row presents the accuracy for the forecasts produced at different geographical levels (10^o x 10^o, $30^o \times 30^o$, $60^o \times 60^o$, $90^o \times 90^o$, $180^o \times 180^o$) while each column presents the accuracy for the forecasts produced at different temporal aggregation levels (monthly, quarterly, yearly, three-yearly and five-yearly). As all the forecasts are disaggregated to the most granular level, the accuracy of all the forecasts are represented in 10^o x 10^o geographical areas. Thus, it is possible to directly compare and identify where different strategies perform better or worse than others.

Based on the grey scale gradation described above, an initial scan of the maps suggest that the following key observations can be made:

- There is no single aggregation level that is better than all other aggregation levels across all geographical regions. Different temporal and geographical levels can bring superior forecasting performance in the various continents.
- For most of the regions and for all variables of interest, temporally aggregating the results in three-yearly or five-yearly bucket windows does not seem to offer advantages over higher frequencies (monthly, quarterly and yearly).
- There is a significant deviation of the results in terms of the variable of interest. If a strategy performs best in certain regions when measuring the accuracy of, for example, forecasting earthquake occurrences, it cannot be assumed that the same strategy will result in the best performance when forecasting either the number of deaths or injuries.
- There are a few regions where there is little to distinguish between the performance of the different strategies.

[3] Siberia, Arctic Canada, Greenland and the Antarctic are notable examples of 'white' areas, as are most open ocean areas.

It should be noted that as this set of results is based on a single evaluation window of 5 years of data, it is quite possible that results for different geographical regions might change when various aggregation strategies on new data are evaluated. But this is part of the beauty of this approach as the methodology and respective models is fully dynamic: every time a new data point is added – for every new earthquake – or any natural disaster in principle – the maps would be updated respectively. Also, new shorter or larger evaluation windows could be prescribed and that would give new maps while also, obviously, a different database would give a new model as well. Finally, if the decision maker decides to add a new variable that captures the desired decision outcomes, then yet again a new database is formed and that would create in, itself, new maps.

We do in fact cherish this dynamic aspect of the whole endeavor. We do not expect however to change very often as only very major events could influence immediately the predictability of the underlying forecasting model within the aggregation framework – that this could be changed as well at any time. So, in practice these are expected to be observed and updated frequently but not to an extent that plans are redrawn all the time rather than only when a major event has happened or when decision makers decide ot shift the decision variable.

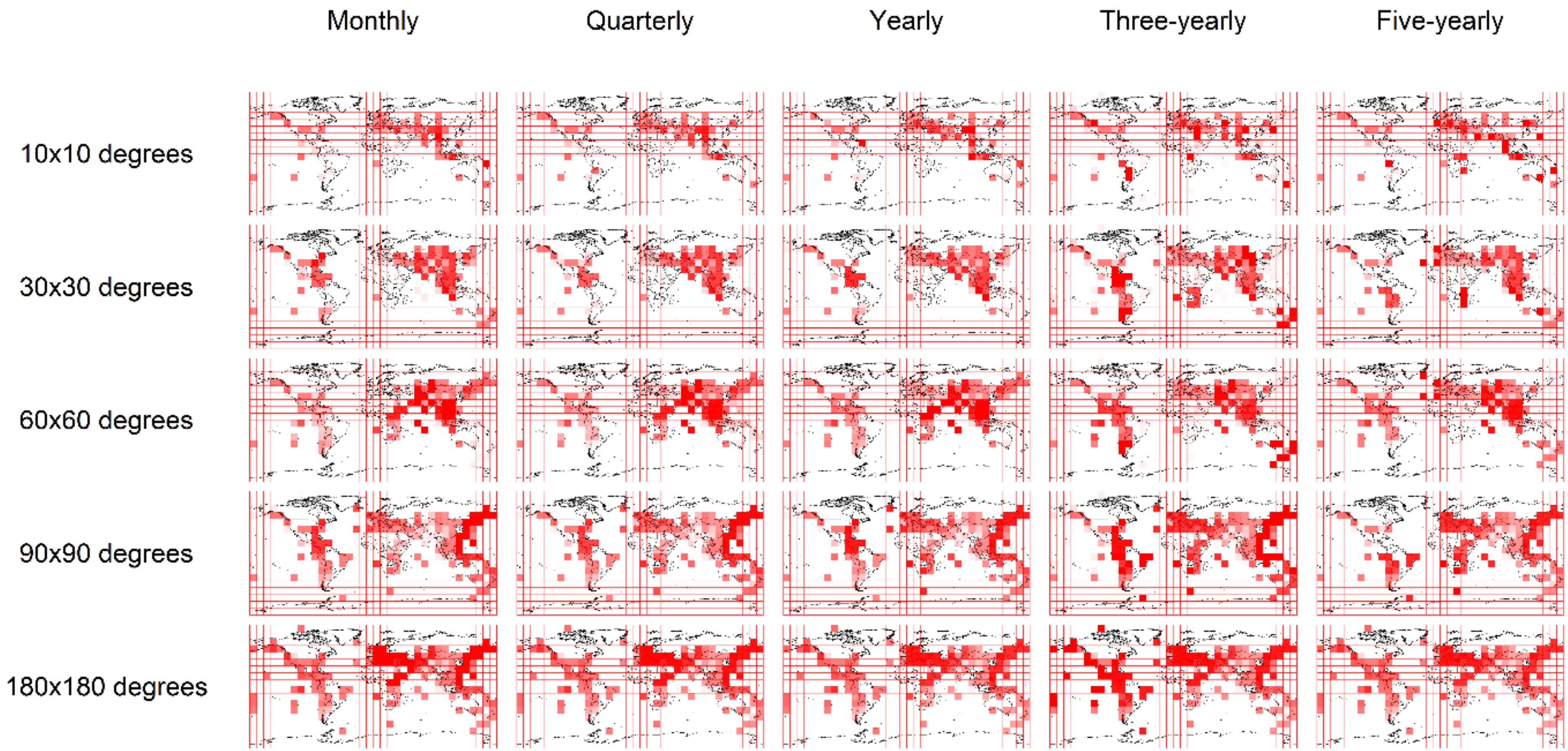

Figure 6. Forecast accuracy results on earthquake events for the various aggregation levels.

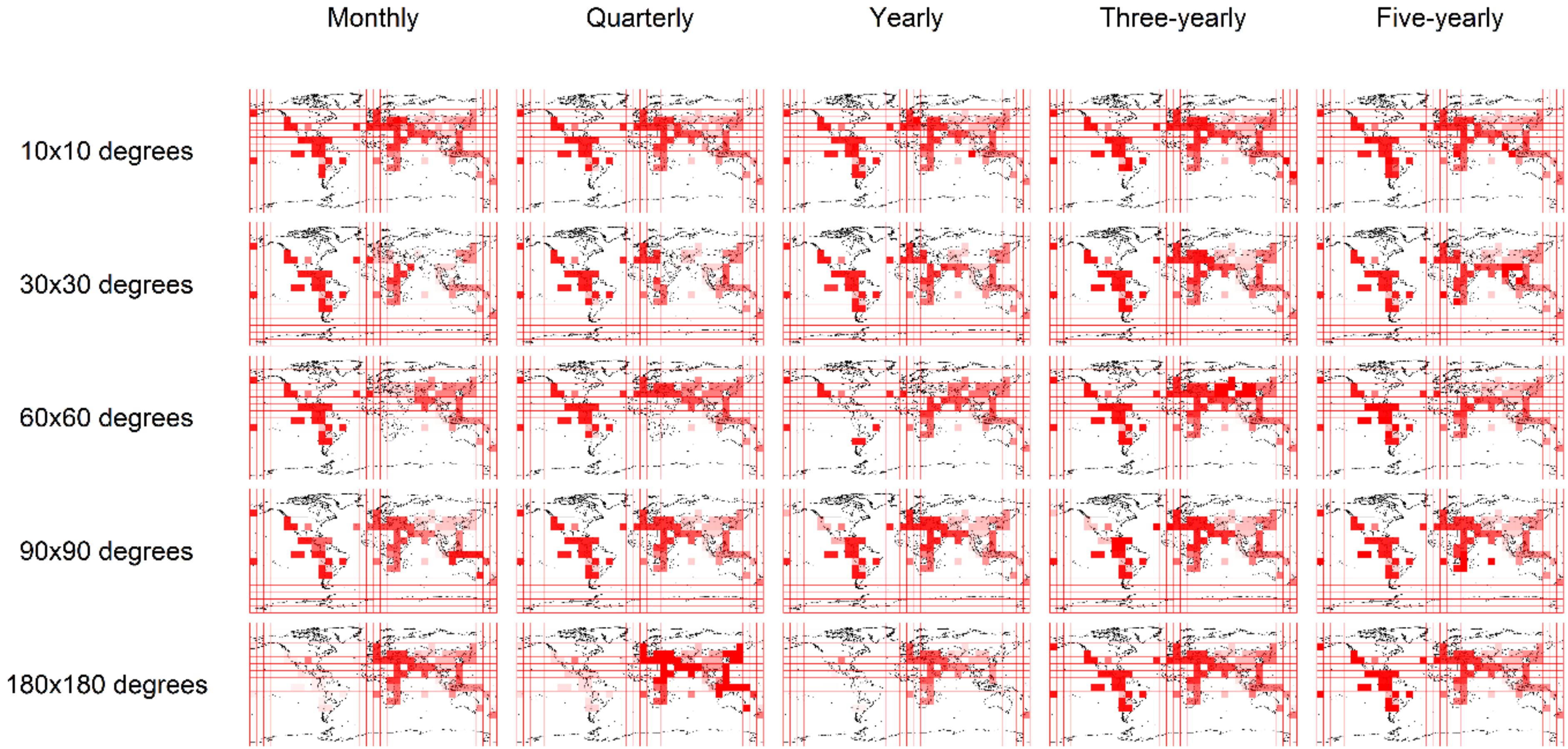

Figure 7. Forecast accuracy results on deaths for the various aggregation levels.

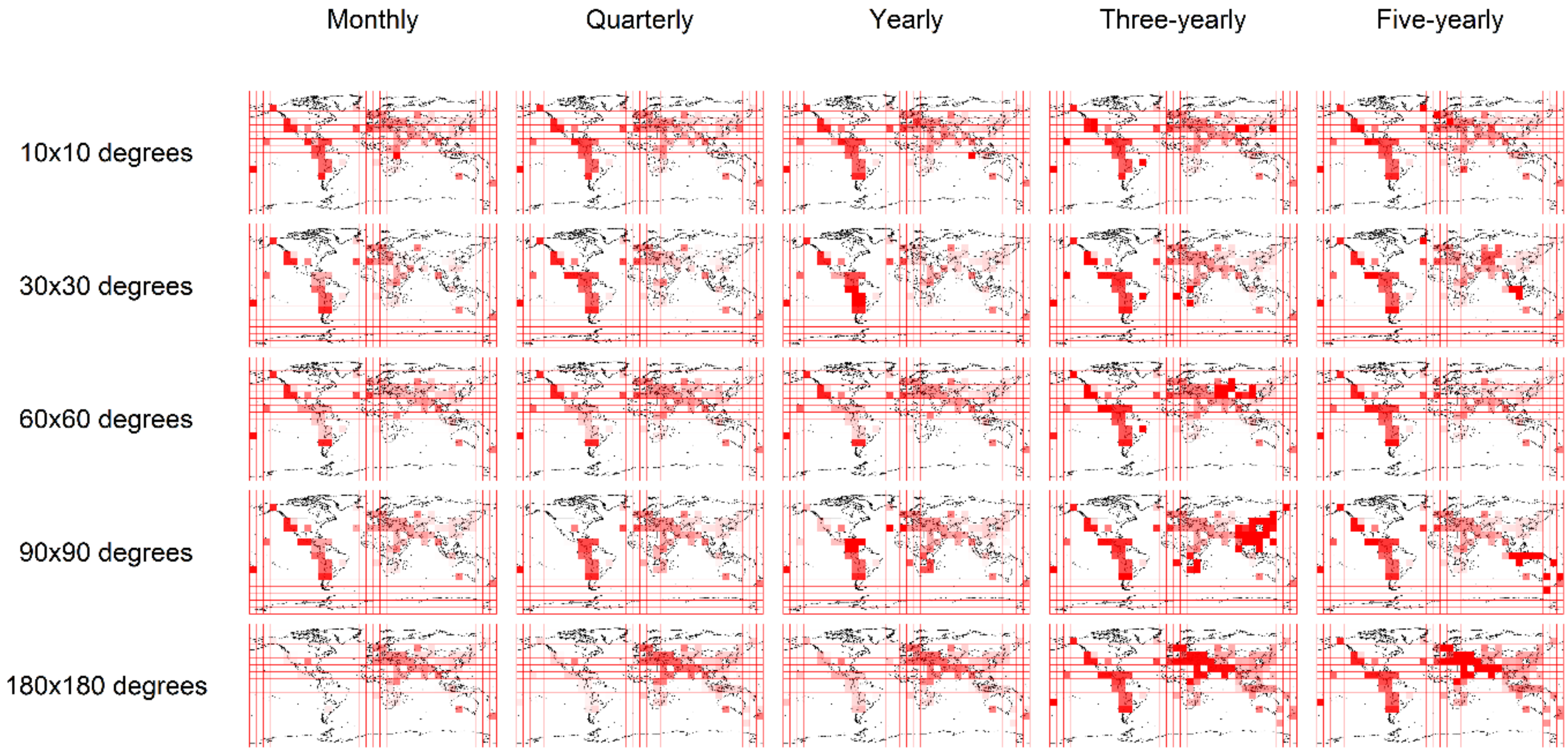

Figure 8. Forecast accuracy results on injuries for the various aggregation levels.

3.3. Interpretation of results

In order to consider the data systematically, and in greater detail, the prediction matrix is interpreted under the three headings suggested above: events, deaths and injuries. While these are not necessarily the only aspects of disaster events for which some form of prediction could be made, they are the primary aspects which can be measured for any earthquake and which therefore need to be considered. Thus, the results presented in grid form in Figures 6, 7 and 8, for events, deaths and injuries respectively can be usefully discussed in more detail as below.

3.3.1. Events prediction

For overall patterns across all spatial scales, some broad observations can be made. Using the finest grid (10º × 10º), there is a general tendency for less predictability as the time period is extended, implying that earthquake event forecasts are rather more accurate over monthly and quarterly timescales than they are over periods of 1,3 or 5 years. This reflects the robustness of the record with lower levels of time-aggregation which allows repeat patterns or trends to be evidenced more clearly. With longer time periods, repeat patterns cannot emerge as the overall dataset itself, in this instance, spans only thirty years. For such patterns to be identified at annual, 3 yearly or 5 yearly levels, a much longer data period would be required.

Using a 30º × 30º grid, there is a consistent low to medium level of accuracy of predictability in all regions. However, in Europe and the Mediterranean, the data appear to yield the best forecasts at this level of spatial aggregation. At the 60º x 60º scale, monthly and quarterly time periods appear to yield the best predictions. Annual, 3 year and 5 year aggregation appear to weaken the forecast accuracy. At the 90º x 90º scale, there is low forecast accuracy so the forecasts at this level of spatial aggregation are of limited value except for earthquake forecasts in the Andean chain which are noticeably more accurate than elsewhere on a 60º x 60º grid. The 180º x 180º scale of aggregation is of least value as the data are simply too spatially aggregated to make any meaningful predictions.

Elsewhere, forecasts in North America appear to be relatively accurate in terms of general predictability at the 180º x 180º scale, but as location precision increases, prediction accuracy decreases. For Africa, the main earthquake events most commonly occur in eastern and southern Africa, as well as in the Maghreb and Mediterranean belt. The highest levels of predictability are observed at 30º × 30º for monthly and quarterly timescales. However, overall, no matter what the grid scale, and no matter what the time-period, predictability of earthquakes for eastern and southern Africa is still low. Forecasts for southern Africa perform better than for eastern Africa at 180º x 180º for all time periods while forecasts for eastern Africa perform better than southern Africa at 90º x 90º for monthly and quarterly time periods. For Central America and the Caribbean, 60º x 60º aggregation at monthly and quarterly time periods are the best performing prediction scales. Lastly, for the south

Pacific and New Zealand, the lowest levels of prediction accuracy are at 3 yearly, 90º x 90º, while all other spatial scales and timescales perform better.

3.3.2. Deaths prediction

The best predictions are achieved at monthly and quarterly time-scales for an 180º x 180º grid, all other time and spatial scales yield much weaker forecasts.

For Africa, there are reasonable levels of prediction at monthly and quarterly timescales. Less predictable patterns of deaths from earthquake events are at 10º × 10º, 30º × 30º, 90º x 90º and 180º x 180º. More predictable patterns are at a spatial aggregation scale of 60º x 60º , implying that the pattern is U shaped with prediction accuracy decreasing both above and below the 60º scale. All other timescales and grids for Africa yield poor results i.e. prediction is inaccurate. For the Far East and Central Asia, the best forecasts appear to be at the spatial scales of 90º x 90º, while for South America the best prediction scale for deaths arising from earthquakes is at the 180º x 180º scale. While these appear to give relatively accurate forecasts, again they are of limited value as they are spatially too vague to be useful.

For west and central Europe, a grid of 60º x 60º at a monthly timescale produces the best forecasts. This is similar for Africa where the results are U shaped i.e. the accuracy of predictions of earthquake deaths is highest using a 60º x 60º grid, but significantly worse using both finer grids (10º × 10º and 30º × 30º) and larger grids (90º x 90º and 180º x 180º).

3.3.3. Injuries prediction

The overall patterns for injuries are rather different *vis-a-vis* the patterns for earthquake events and deaths. At the 5 yearly time-scale, while more data are used in the computation, the forecasts become less accurate, particularly moving from the 60º x 60º through to 180º x 180º grids.

The best prediction scales for South America are at 60º x 60º and 180º x 180º with the value of the forecasts over longer time periods again reducing. However, the longer time periods are essentially too vague to be of any meaningful use due to the reduced opportunity for patterns to repeat. Furthermore, all forecasts are inaccurate at all levels of spatial aggregation for both 3 and 5 year periods. For Asia, the best forecasts are yearly at all levels of spatial aggregation, the worst are the 3 yearly on a 90º x 90º grid.

4. A SUPPLY CHAIN PLANNING AND RISK-MITIGATION MODEL FOR HUMANITARIAN LOGISTICS

The logic of the development of the model is visualized in Figure 9. The model is dynamic as it is based on live data which is updated each time a new earthquake data point is inserted into the database and the model can then be rerun to provide decision-making graphs as presented at the end of this section. The model is also dynamic in terms of the variable space on which decisions are taken: in the model as presented here, the frequency of earthquakes number of casualties and number of injuries are considered and recorded in the respective dynamic database; but this can be expanded at any time so that policy makers can decide on their strategies based on the variable they are interested in most. For example, chief consideration could be given to: resources allocated, economic impact, time to return to normal, and so on; furthermore, a multiple criteria decision analysis (MCDA) approach could also be followed, as well as considering all available variables with different (or equal) weights.

The model is also empirical and evidence-based as it is not driven from deductive theory, rather by actual data that could be live-fed into the database enabling the model to be rerun if needed, even as each new data point becomes available. This can be therefore be done on a rolling basis (by keeping the width of the rolling window fixed) or by expanding the dataset by including all available historical data.

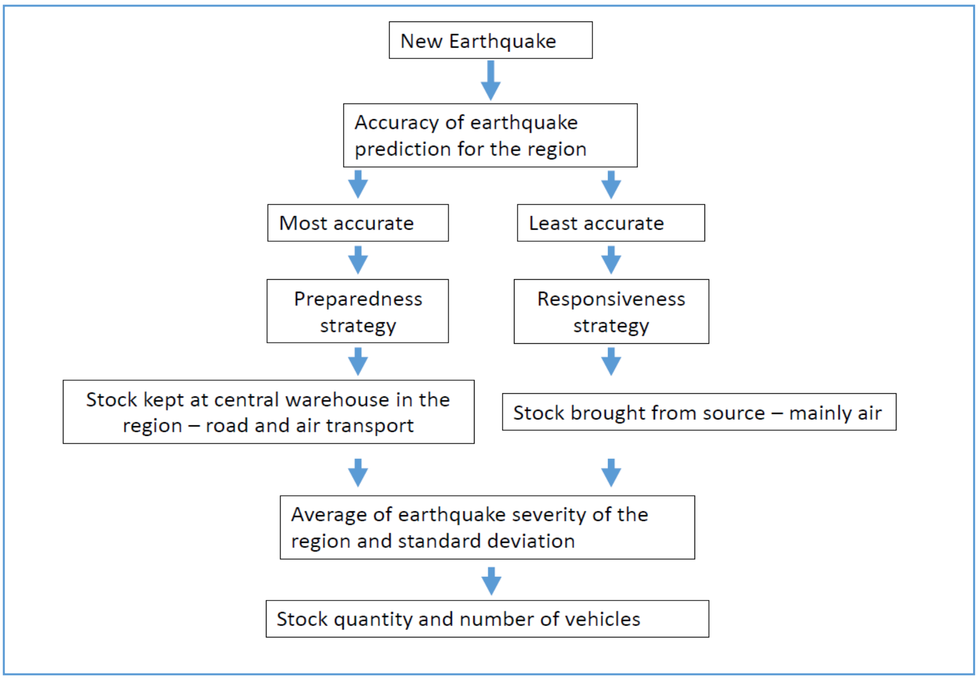

Figure 9. Model development

The logic of the model is quite simple, but at the same time it is intuitively appealing:

- When earthquake predictability is relatively high, as derived from the models and respective graphs in Section 3, a preparedness strategy is encouraged, and thus *preparedness* is promoted from the model. On the other hand, when earthquake predictability is relatively low, then *responsiveness* is suggested as the better impact mitigation strategy. This is highlighted by the grey (preparedness) and dark grey (responsiveness) in the respective decision maps in Figure 10. This essentially suggests that taking a resource optimization approach could be fruitful in order to evaluate the relative effectiveness of expenditure on pre-event mitigation measures versus spend on post-event disaster relief (He and Zhuang, 2016).
- The degree of strategic centralization is determined by the identification of the maximum number of neighboring areas with common characteristics – these are the same shade of grey in the decision maps in Figure 10. This gives a sense of 'which area should be working with which' or, more specifically, 'who should we be working with whom' in responding to major humanitarian situations. The key areas are not necessarily the areas where frequency of earthquakes is greatest, as neighbouring countries (e.g. Mexico and USA) have different response infrastructures and thus different levels of exposure to the aftermath of earthquakes in the decision space (casualties, injuries, etc.). The level of temporal and geographical aggregation which provide the most accurate forecast of a cluster of earthquakes can determine the degree of centralisation or localisation of the distribution operation responding to that cluster of earthquakes.
- The temporal dimension can either be a strategic decision (based on financial budgetary periods, for example) or selected so as to change the predictability levels and shift strategies from a responsiveness model to a preparedness model, or vice versa.
- In most cases, i.e. over most time-scales and in most regions at most scales, a hybrid strategy which combines a fair degree of preparedness with a reasonable degree of responsiveness can be the most educated approach.

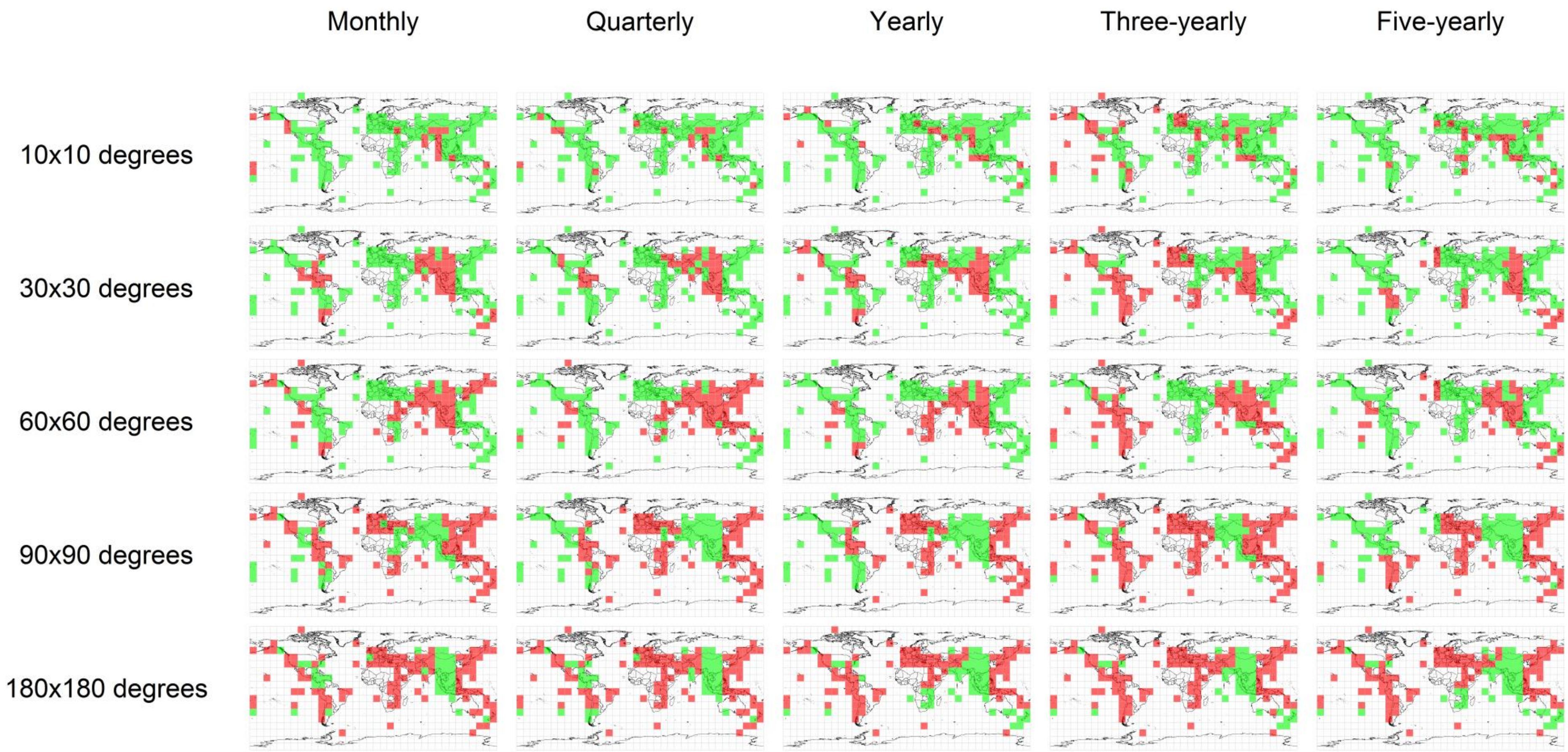

Figure 10. Decision maps for various aggregation levels for earthquake occurences.

5. IMPLICATIONS FOR THEORY AND PRACTICE

The general principle underlying this analysis is that the greater the level of aggregation of earthquake events, the more likely the forecast is to be at least broadly correct; but at the same time the more spatially aggregated the forecast, the lower its value because the higher the degree of spatial aggregation, the more vague is the location of any potential earthquake event. Essentially, earthquake forecasting hinges on a trade-off between accuracy and value (McKenna, 2011). In some cases, however, the result patterns follow a U-shape such that, for example, below a certain level of aggregation, forecasts worsen and above that level of aggregation the forecasts also worsen. This hints at an optimum scale for forecasting which varies depending on region, and this is irrespective of whether consideration is being given to events, deaths or injuries. This is an important finding of this research.

With the initial forecasts consistently over-estimating earthquake frequency, this suggests that the prediction technique used in this research could be refined further to narrow the gap between forecasts of earthquake frequencies and actual earthquake events. As the 3-year and 5-year forecast aggregations are generally the weakest, the more focused monthly, quarterly and yearly forecast aggregations are more appropriate for determining the best locations for the prepositioning of aid. Agencies will need to preposition aid in areas where the highest accuracy forecast is, which is generally the case for 90º x 90º and 60º x 60º forecast aggregations. Also, it is appropriate that yearly levels of aggregation outperform more granular levels of aggregation as it is more likely to be the case that one major event will occur in a yearly time period than that one event would occur within a specific month. In order to show the practical relevance of the forecasting procedure to the question of aid pre-positioning, a model is devised (Figure 11) combining levels of responsiveness, degree of stock centralisation, level of stock holding and the preparedness related to each forecast.

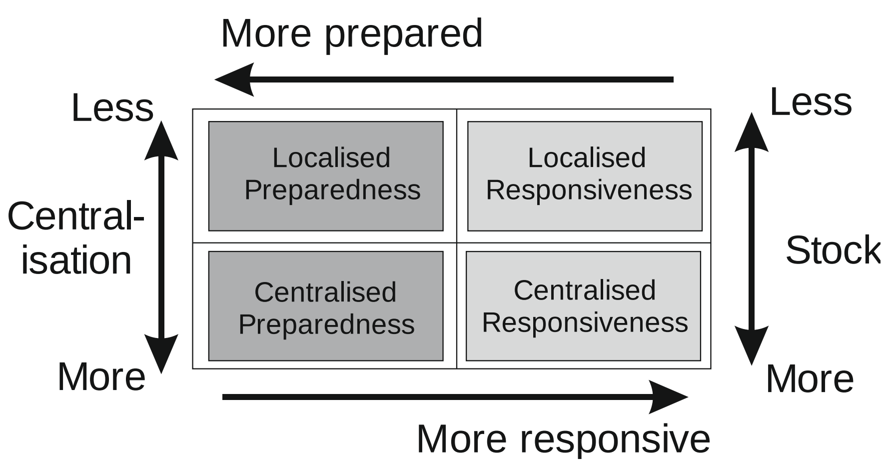

Figure 11. Decision maps for various aggregation levels.

Each level of temporal and geographical aggregation can be placed into the model to show which disaster relief strategy is the most relevant to respond to a particular disaster depending on the accuracy of temporal and geographical aggregations of the region where the earthquake occurs. The most appropriate geographical scale is regional (generally, the 90º x 90º and 60º x 60º forecast aggregations) and the best temporal scale is yearly. Thus, if the model is populated with a particular level of temporal and geographical aggregation, generally the most robust levels of aggregation are 90º x 90º and 60º x 60º, with the 10º × 10º and 30º × 30º forecast aggregations having the lowest accuracy. This suggests that the appropriate strategy to be adopted is to be prepared at regional level, but to be responsive at country and local levels.

This has implications for disaster relief supply chain strategies adopted by decision makers. The research found that for the 10º × 10º scale at all temporal aggregation levels, the earthquake data used in the paper required more responsive and localized strategies; however, for the forecasts run at the 60º x 60º scale with the monthly, quarterly and yearly aggregations are the most accurate forecast of all the combined temporal and geographical aggregations used in the paper, so it could be argued that a region-based preparedness strategy is required at this level of aggregations with a fair degree of stock centralization. The 90º x 90º aggregation levels have generally more accurate forecasts; this also seems to be an indication of the need for a more prepared and centralized stockholding strategy. The 30º × 30º aggregation levels seem to have lower forecast accuracy for monthly and quarterly and yearly temporal aggregations and medium earthquake forecast accuracy for yearly and 3-year temporal aggregations, which could mean that the most appropriate strategy to be taken is to be responsive at local and country levels, but with increasing stock quantity held at regional level when temporal aggregation is extended. This suggests that a hybrid preparedness-responsiveness strategy is required, and the balance between preparedness and responsiveness should vary somewhat from region to region. Just as important is the suggestion that the optimum scale for mitigation planning also varies from region to region

6. CONCLUSIONS AND FUTURE RESEARCH

There is a clear argument for aid pre-positioning and such strategies are already followed by a range of organisations, for example the UN and the IFRC. There are, however, a number of factors which need to be considered in the overall picture when making decisions on warehouse location, for example facility operations, fixed overheads, staffing and stock levels which will all add costs. From a supply chain perspective, there is also the need to balance the number of facilities against the increase in inventory holding costs associated with more facilities. A parallel debate therefore relates to how many facilities and which are the most effective locations for them. Earlier modelling based simply on macro patterns of population distribution suggested that six facilities in Southern Europe, South Central Asia, East Asia, South America, Eastern Africa, and South eastern Asia would be most appropriate (Akkihal, 2006). Western USA, Central America and the southwest Pacific are examples of regions which are conspicuously absent from this list. However, this paper provides new light which could be used for decision making on network redesign of regional disaster relief operations, if other kind of disasters are included in the database and the model is re-run with those disasters.

The model developed in this paper can guide policymakers and the third (non-governmental) relief sector in terms of the range of supply chain risk mitigation strategies which can be adopted in the context of disaster relief distribution. The paper argues that an improvement in the prediction of earthquake events through temporal and geographical data aggregation could influence the location and size of disaster relief distribution facilities positioned in different world regions, the stock policy adopted to supply areas affected by disasters, and how disaster relief supply chains respond to such black swan events. The research can be improved further by adding data for additional natural disaster types such as tsunami, flooding, storm and drought. Aggregating across all disaster types would produce a more robust, although not necessarily different, network configuration. A combination of hazard type, magnitude, and regional characteristics such as population and infrastructure, could improve the disaster "footprint" and assist in predicting inventory locations, ultimately improving the relief system (Akkihal, 2006).

This points towards building a composite natural disaster ‘heat map’ or three-dimensional model as being a natural ‘next-step’ for this research.

One thing must be made crystal clear. This research is not for the method to be used per se used when forecasting earthquakes; it is more for the strategy that follows once we realize our predictability limits, so it is about the risk mitigation model that comes after. One could argue that maybe we could forecast better if we use extreme value theory or even maybe other computational intensive methods – but this is not what we are trying to do here. We have seen evidence in the respective literature that temporal aggregation works well in an intermittent demand context, and we use it without having an empirical forecasting competition in mind to set – we leave that as future research. What we strongly argue however is that temporal and spatial aggregation can give the geographic areas and timeframe within which centralization of resources should take place; and that you cannot achieve through the other alternative forecasting methods that do not consider aggregation. This latter contribution plus the responsiveness/preparedness risk mitigation model built on that, we consider to be the fundamental contribution of this research; and we are sure it will create the necessary discourse and discussion on the development of similar models, and we do cherish and anticipate such activity.

Several questions arise when considering the outcomes of this data-driven modelling exercise compared to existing strategies. The UN network, for example, is based on all types of disaster not just earthquakes. However, would alternative locations for response strategies to different disaster types perform better, or does one network covering all disaster types provide a sufficient level of coverage to ensure an effective response at all times? In respect of future research, a systematic evaluation of this forecast method against alternative forecasting tools and against current in practice in disaster relief would be fruitful.

REFERENCES

Akkihal AR (2006) *Inventory pre-positioning for humanitarian operations*, Unpublished Thesis, Massachusetts Institute of Technology, 109 pp.

Alem D, Clark A, Moreno A (2016) Stochastic network models for logistics planning in disaster relief, *European Journal of Operational Research*, 255, Issue 1, pp. 187-206

Athanasopoulos G, Hyndman RJ, Kourentzes N, Petropoulos F. (2015) Forecasting with temporal hierarchies, *Lancaster University Management School, Working Paper Series*, 2015:3.

Balcik B, Beamon B, Krejci CC, Muramatsu KM, Ramirez M. (2010) Coordination in humanitarian relief chains: Practices, challenges and opportunities. *International Journal of Production Economics*, 126, 22 – 34.

Campbell AM, Jones PC (2011) Prepositioning supplies in preparation for disasters. *European Journal of Operational Research*, 209, 156 – 165.

Carter WN (1999) Disaster Management: A Disaster Management Handbook, Asian Development Bank, Manila, 215 pp.

Fliedner G. (1999) An investigation of aggregate variable time series forecast strategies with specific subaggregate time series statistical correlation. *Computers and Operations Research*, 26, 1133–1149.

Galindo G and Batta R (2013) Review of recent developments in OR/MS research in disaster operations management, *European Journal of Operational Research*, 230, Issue 2, 201-211

Gatignon, A, Van Wassenhove LN, Charles A, (2010) The Yogyakarta earthquake: Humanitarian relief through IFRC's decentralized supply chain. *International Journal of Production Economics*, 126, 102 – 110.

Gross C, Sohl J. (1990) Disaggregation methods to expedite product line forecasting. *Journal of Forecasting*, 9, 233–254.

Gutjahr WJ and Nolz PC (2016) Multicriteria optimization in humanitarian aid, *European Journal of Operational Research*, Volume 252, Issue 2, 351-366

He F, Zhuang J (2016) Balancing pre-disaster preparedness and post-disaster relief, *European Journal of Operational Research* 252, 246 - 256

Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., Shang, H. L. (2011) Optimal combination forecasts for hierarchical time series, *Computational Statistics and Data Analysis* 55 (9), 2579–2589.

Kilci F, Kara BY, Bozkaya B (2015) Locating temporary shelter areas after an earthquake: a case for Turkey, *European Journal of Operational Research* 243, 323 - 332

Kourentzes N., Petropoulos F., Trapero Arenas J.R. (2014) Improving forecasting by estimating time series structural components across multiple frequencies, *International Journal of Forecasting*, Vol. 30, No. 2, 291-302.

Kovacs G and Spens KM (2011) Trends and developments in humanitarian logistics – a gap analysis, *International Journal of Physical Distribution and Logistics Management* 41 (1) 32 – 45

Kunz N and Reiner G (2012) A meta-analysis of Humanitarian logistics Research, *Journal of Humanitarian Logistics and Supply Chain Management* 2(2) 116 - 147

Lane FW (1966) *The elements rage*, David and Charles, Newton Abbott, 280 pp.

Leadbetter, MR. (1991) On a basis for 'Peaks over Threshold' modeling, *Statistics and Probability Letters* 12 (4) 357–362.

Li, Y, Liu, X, Chen, Y (2011) Selection of logistics centre location using Axiomatic fuzzy set and TOPSIS methodology in logistics management, *Expert Systems with Applications*, 38, 7901-7908

McKenna P (2011) Why earthquakes are hard to predict, *New Scientist*, available at: http://www.newscientist.com/article/dn20243-why-earthquakes-are-hard-to-predict/html>

Meteorological Office (2016) Using probabilities in weather forecasting. Available at <http://blog.metoffice.gov.uk/2012/10/04/using-probabilities-in-weather-forecasting/>

Naji-Azimi Z, Renaud J, Ruiz A, Salari M (2012) A covering tour approach to the location of satellite distribution centres to supply humanitarian aid, *European Journal of Operational Research* 222, 596 - 605

National Earthquake Information Center (NEIC) (2016) United States Geological Survey: Earthquake Facts and Statistics, Available at <http://earthquake.usgs.gov/earthquakes /eqarchives/year/eqstats.php>

Nikolopoulos, K., Syntetos, A., Boylan, J., Petropoulos, F. and Assimakopoulos, V. (2011) ADIDA: An aggregate/disaggregate approach for intermittent demand forecasting, *Journal of the Operational Research Society*, 62: 544-554.

Onut, S, Soner, S (2007) Transhipment site selection using the AHP and TOPSIS approaches under fuzzy environment. *Waste Management*, 28 (9), 1552 – 1559.

Ozdamar L and Etem MA (2015) Models, solutions and enabling technologies in humanitarian logistics, *European Journal of Operational Research*, 244, Issue 1, 55-65

Palmer, T (1999) Predicting uncertainty in forecasts of weather and climate, *Rep. Prog. Phys*. 63 (2000) 71–116

Paul JA, MacDonald L (2016) Location and capacity allocations decisions to mitigate the impacts of unexpected disasters. *European Journal of Operational Research*, 251, 252 – 263.

Petropoulos F., Kourentzes N. (2015) Forecasts combinations for intermittent demand, *Journal of the Operational Research Society,* Vol. 66, No. 6, 914-924.

Petropoulos F., Kourentzes N. (2014) Improving forecasting via multiple temporal aggregation, *Foresight: The International Journal of Applied Forecasting*, Issue 34 (Summer 2014), 12-17.

Pettit SJ and Beresford AKC (2009) Critical Success Factors in Humanitarian Aid Logistics, *International Journal of Physical Distribution and Logistics Management*, 39 (6) 450 468

Powell JH, Mustafee N, Chen AS, Hammond M (2016) System-focused risk identification and assessment for disaster preparedness: Dynamic threat analysis, *European Journal of Operational Research* 254, 550 - 564

Rawls, CG, Turnquist, MA (2010) Pre-positioning of emergency supplies for disaster response. *Transportation Research Part B*, 44, 521 – 534.

Roh S, Beresford AKC, Pettit SA, Harris I (2015) The Pre-Positioning of Warehouses at Regional and Local Levels for a Humanitarian Relief Organization, *International Journal of Production Economics*, Volume 170, Part B, December 2015, 616-628

Taleb, NN (2007) *The Black Swan: The Impact of the Highly Improbable*. New York: Random House and Penguin.

Tatham P and Pettit SJ (2010) Transforming Humanitarian Logistics – The Journey to Supply Network Management, *International Journal of Physical Distribution and Logistics Management*, 40 (8/9), 609-622

Tofighi S, Torabi SA, Mansouri SA (2016) Humanitarian logistics network design under mixed uncertainty, *European Journal of Operational Research*, 250, 239 - 250

United Nations (2016) UN Humanitarian Response Depot, Available at < http://www.unhrd.org/>

Whittow J (1980) *Disasters – the anatomy of environmental hazards*, Allen Lane, London, 411 pp.