Chapter 2

Methodology

This chapter describes the methods used in the study. The methodology comprises the following components.

(a) Data management.

(b) Methods used for quantitative analysis.

(c) Methods used for create the geographical maps.

2.1 Data management

The following computer programs were used for data analysis and thesis preparation.

Microsoft Excel was used to manage the data for this research. It has many functions that are useful for manipulating data. Microsoft Word was used to create some of the figures and to format the thesis.

WebStat, a suit of web-database software engineering tools written in HTML and ASP (Microsoft's Active Server Pages) is used to create graphs for statistical analysis, when the data are stored in an SQL Sever database. It runs on a local area intranet, and is available in the postgraduate computer laboratory in the Department of Mathematics and Computer Science at Prince of Songkla University in Pattani.

The open source statistical package R is used to fit the negative binomial distributions to the risk rate.

2.2 Methods used for the Quantitative Analysis

Graphical Methods

The following graphical display was created using Microsoft Excel.

Histograms of data for The number of events representing the distribution including mean, minimum and maximum of number event by week.

Statistical Methods

In our study the outcome is defined as the occurrence of a terrorist event at any location in the region during the period from the beginning of January 2004 to the end of December 2005. The severity of the outcome was coded as an integer from 1 to 9 as given in Table 1.1.

The region comprises all districts in the provinces of Narathiwat, Pattani and Yala, together with the four westernmost districts of Songkla Province. The hour of the day, date, and subdistrict where the event took place were recorded. There are 36 districts and 290 subdistricts in this region.

We computed the risk rate of events for each subdistrict in a specified period of time by dividing the total number of events recorded in the subdistrict over the period by the corresponding population resident in the subdistrict in 1000s according to the 2000 Population Census of Thailand. The average population of the 290 subdistricts was 6,694, but these populations varied substantially. Six subdistricts had populations greater than 20,000 (Sateng, SatengNok and Betong in Yala Province with 73,077, 24,745 and 23,531 residents, respectively, BangNak and SungaiKolok in Narathiwat Province with 42,010 and 37,671 residents, respectively, and Sabarang in Pattani

Province with 23,702 residents). At the other extreme, nine subdistricts had fewer than 2,000 residents, the smallest being Tachi in Yala Province with 1,288 residents.

Since the research question involves studying how the risk of a terrorist event depends on various factors including the time of day, the day of the week, the month of the year and the district or subdistrict in the region, we focused on the rate of occurrence of such events with respect to an appropriate denominator reflecting the corresponding population at risk.

In preliminary analysis we used MapInfo Version 6.5 (MapInfo Corporation 2001) to create thematic range maps based on event rates per year in subdistricts. The maps were coloured using five bands with graduated 0-255 scaled (red, green, blue) amounts, namely, (1) less than 0.5 events per 1000 per year (255 red, 208 green, 232 blue), (2) 0.5 to 0.5 (255, 144, 200), (3) 0.5 to 1 (255, 64, 160), (4) 1 to 1.5 (255, 0, 0) and (5) 1.5 or more (80, 0, 0). Since such thematic maps show fluctuations due to unpredictable random variations, we then created grid thematic maps based on smoothing the range maps with IDW (inverse density weighted) interpolator (cell size, exponent, search radius) settings (0.25 miles, 4, 20 miles), creating contours using the same colour bands centred at (0.25, 0.65, 1, 1.5 and 2) events per 1000 per year. We also considered event rate increases from 2004 to 2005, starting with a range map based on four specified bands, namely (1) less than -0.2 events per 1000 per year (255 red, 255 green, 208 blue), (2) -0.2 to 0.2 (0, 255, 0), (3) 0.2 to 0.75 (255, 0, 0) and (4) 0.75 or more (80, 0, 0), and finishing with a grid map using the same IDW interpolator settings centred at increases of (-0.5, 0, 0.5 and 1) events per year.

The statistical methods comprise comparison of proportions including Pearson's chi-squared test, and generalized linear models for analysing risks including Poisson and negative binomial regression (Venables and Ripley 1999). If we assume that the event occurrences are described by a Poisson or negative binomial process with mean λ_{it}, where i denotes place (district, say) and t denotes period, the model takes the form

$$\ln(\lambda_{it}) = \ln(P_i/1000) + \mu + \alpha_i + \beta_t, \qquad (2.1)$$

where P_i is the population of place i. For the Poisson distribution the variance is the same as the mean, whereas the negative binomial distribution is relatively over dispersed, having variance $\lambda_{it}(1+\lambda_{it}/\theta)$. The negative binomial model is thus the special case of the Poisson model arising in the limit as θ tends to infinity.

The overall goodness-of-fit of the model may be assessed by using Pearson's chi-squared test to compare the residual deviance (defined as –2 times the difference in the log-likelihood between the fitted model and a saturated model that fits the data perfectly) with the number of degrees of freedom (defined as the sample size minus the number of fitted parameters). A more detailed analysis of the goodness-of-fit that highlights individual anomalies involves graphing 'deviance residuals' against corresponding quantiles from a standardized normal distribution. We used the R statistical system (Venables and Smith 2004) for statistical model fitting, plotting confidence intervals and assessing goodness-of-fit.

2.3 The methods used for creating the geographical maps

Software used: MapInfo 6.5

Thematic maps

A thematic map is a type of map that uses a variety of graphic styles (usually colours of fill patterns) to graphically display information about a map's underlying data. Thus a thematic map using data in region might show one region in dark red (to indicate that the region has low values).

A range map is a type of thematic map that displays data in a region according to the range set by the users. The ranges are shaded using colors or patterns.

A grid map is a smoothed range map, using a method such as that described in Section 2.2.

Contour map

A contour map is a way of graphing a variable that takes values in a two-dimensional region. It is not the only way of graphing such data you could also use a 3D wire-frame map or a prism map. A contour map may be obtained from a grid map by using curves to separate the areas having the same shadings or colours.