

A Statistical Modelling Approach to Cyclone Origins

Xun Guo Lin
CSIRO Mathematical & Information Sciences
Private Bag 10, Clayton South, VIC 3169

1. Introduction

Cyclones (also called hurricanes in North America or typhoons in Asia) are a major natural hazard impacting Australia as well as many parts of the world. In Australia, over the period of 1967 to 1999 the total cost of tropical cyclones is approximately \$8.8 billion in 1999 prices (aggregated from events each costing \$10 million or more), averaging \$266 million per year. This accounts for 24% of total cost on natural disasters over the same period in Australia (see Bureau of Transport Economics, 2001).

Mathematical modelling is the first step in the assessment of the risk posed by cyclones. These mathematical models must have the ability to simulate the location of a cyclone, its path and the wind that it generates. This paper describes a generic statistical approach for estimating the origin of a cyclone by analysing a historical cyclone dataset from Australian region.

2. Historical Cyclones

The Bureau of Meteorology has made available to public a cyclone dataset of historical cyclones occurred in Australian region. This dataset contains the path and a number of attributes such as the central pressure of the cyclone.

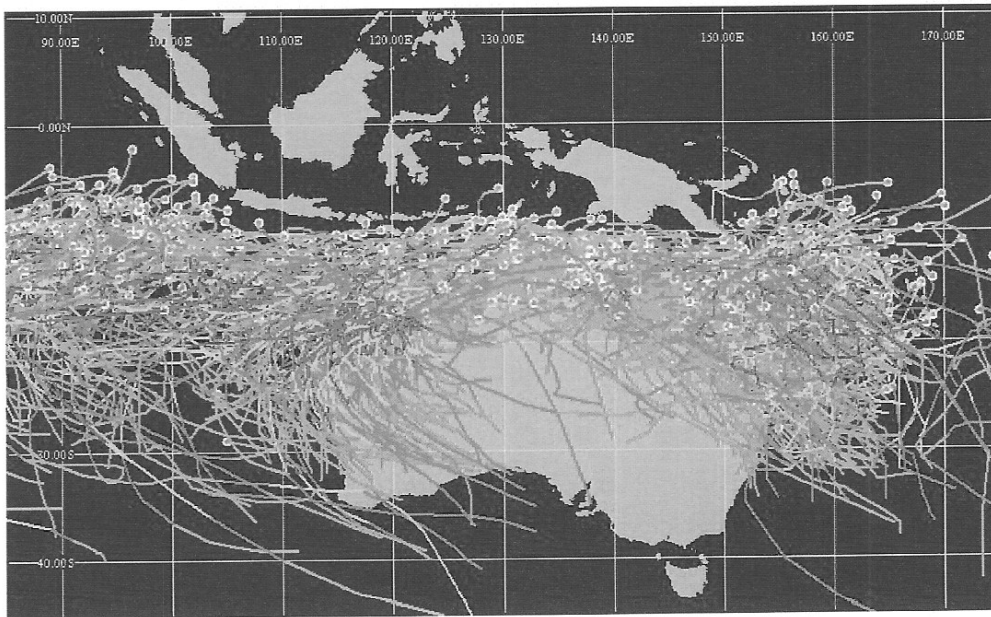


Figure 1: Tracks of more than 900 historical cyclones occurred in Australian region between 1907 and 2001.

Figure 1 shows all observed/recorded cyclone tracks between 1907 and the end of 2000/2001 cyclone season (up to 30 June 2001). The different colours represent different strength (Red is the systems with *hurricane force* winds; Yellow is the *storm force* winds; Green is the *gale force* winds and Blue is below cyclone strength or there was no pressure data in the database).

Note, Holland (1981) provided a comprehensive overview of the quality of the historical cyclone dataset from 1909 to 1979. About half of those 900 historical cyclones in Figure 1 were from the period before the use of satellites (which was from about the mid 60's) and they were rated poor or very poor in terms of their locations, occurrence etc. However, this study aims to introduce a generic statistical modelling approach which uses a historical cyclone dataset as an illustration. The proposed approach is independent from a particular cyclone dataset, though poor quality datasets may mean the statistical models generated from them are less reliable. But throwing away half of the available data will reduce the cyclone data length to its one-third only, which may also introduce the uncertainties.

3. Cyclone Origins

The cyclone origins, that is the location where a cyclone was generated or first observed or upgraded from a tropical depression is indicated by the white circles in Figure 1. The footprints of these origins are shown in Figure 2 and the number counts on a 5 by 5 degree grid of latitude and longitude are displayed in Figure 3. A two-dimensional histogram of the distribution of the cyclones is shown in Figure 4.

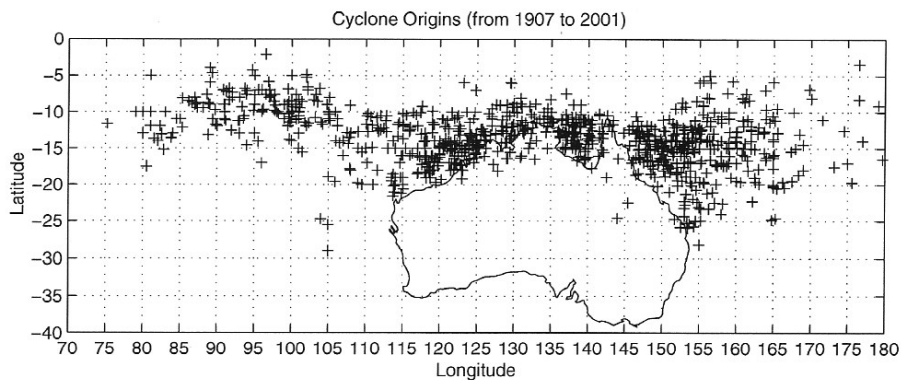


Figure 2: The footprints of cyclone origins.

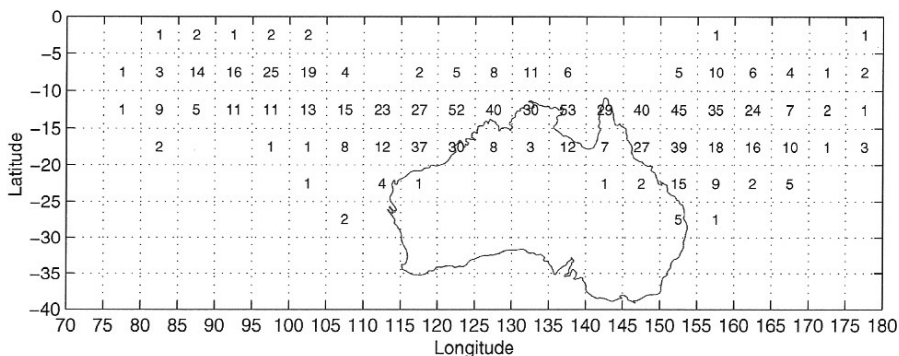


Figure 3: The number counts of cyclone origins on a 5 by 5 degree grid of latitude and longitude.

4. Statistical Distribution of Cyclone Origins

As seen from Figure 4, the two-dimensional distribution of origins has multiple peaks so non-parametric statistical techniques should be used to obtain an estimate of the true spatial distribution of cyclones using the historical data. Particularly, an advanced non-parametric method called kernel density estimation (KDE) (see Silverman, 1986) was applied to the historical data. The probability density function (PDF) of cyclone origins estimated by a Normal kernel with adaptive bandwidth is shown in Figure 5.

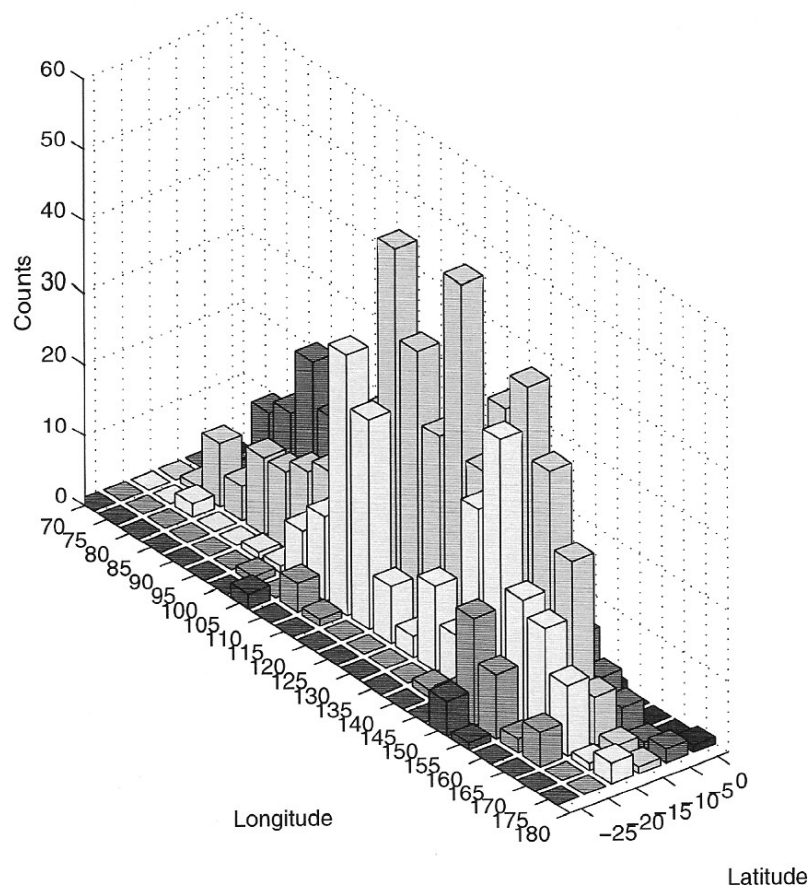


Figure 4: a 3-D view of the cyclone origin counts.

5. Sampling of Cyclone Origins

Based on conditional probability theory, an efficient computational 2-D sampling algorithm has been developed to establish the most likely location a cyclone will develop. The sampling result is shown in Figure 6, where (a) is the contour plot of origin counts on a 5 by 5 degree grid; (b) is the contour plot of the PDF of cyclone origins estimated by KDE; (c) is the contour plot of cyclone origins generated by 1 million random samples drawing from the PDF estimated by KDE. The similarity between (c) and (b) indicates the accuracy of the sampling algorithm developed.

6. Conclusion

A generic statistical modelling technique using advanced kernel density estimation fitted to a historical cyclone dataset from Australian region has been introduced to estimate the distribution of cyclone origins. An effective and efficient computational sampling algorithm has been developed to generate cyclone origins surrounding Australia. This research forms a part of probabilistic cyclone risk assessment currently undertaking by the author.

7. References

1. Bureau of Transport Economics, 2001. *Economic Costs of Natural Disasters in Australia*, Canberra: Bureau of Transport Economics, report 103.
2. Holland, G.J., 1981, *On the Quality of the Australian Tropical Cyclone Data Base*, Australian Meteorological Magazine, V.29, 169-181.
3. Silverman, B.W. 1986, *Density Estimation for Statistics and Data Analysis*, London: Chapman & Hall.

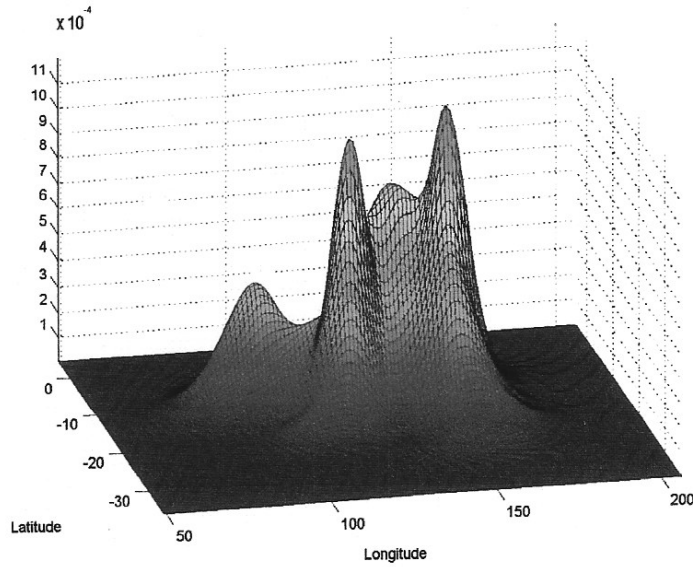


Figure 5: The PDF of cyclone origins estimated by KDE.

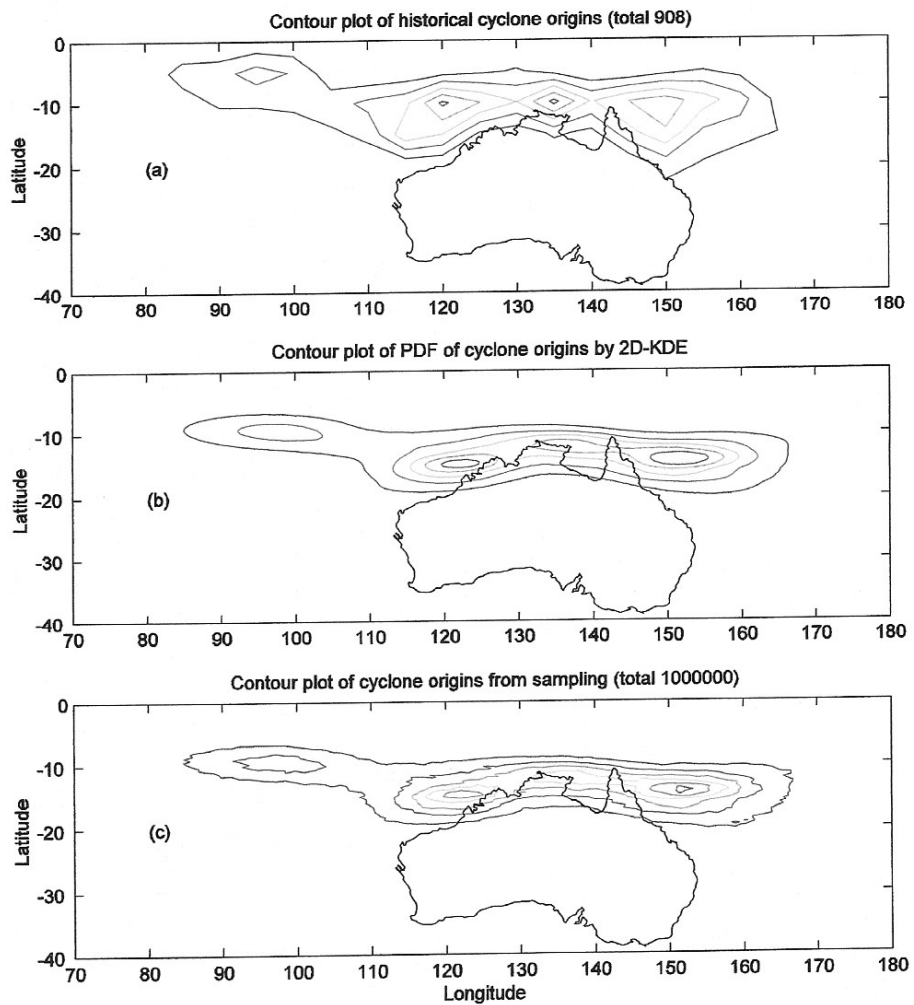


Figure 6: Contour plots of: (a) historical cyclone origin counts; (b) the PDF of cyclone origins estimated by KDE; (c) the PDF of cyclone origins generated by 1 million samples drawing from the PDF estimated by KDE.