

Modelling the Reliability of Search and Rescue Operations within the UK Through Bayesian Belief Networks

Ashley Russell, John Quigley and Robert van der Meer
Department of Management Science
University of Strathclyde
Glasgow UK

Abstract

This paper uses a Bayesian Belief Networks (BBN) methodology to assess the reliability of Search And Rescue (SAR) operations within the UK Coastguard (Maritime Rescue) coordination centres. This is an extension of earlier work, which investigated the rationale of the government's decision to close a number of coordination centres. The previous study made use of secondary data sources and employed a binary logistic regression methodology to support the analysis. This study focused on the collection of primary data through a structured elicitation process, which resulted in the construction of a BBN.

The main findings of the study are that approaches such as logistic regression are complementary to BBN's. The former provided a more objective assessment of associations between variables but was restricted in the level of detail that could be explicitly expressed within the model due to lack of available data. The latter method provided a much more detailed model but the validity of the numeric assessments was more questionable. Each method can be used to inform and defend the development of the other.

The paper describes in detail the elicitation process employed to construct the BBN and reflects on the potential for bias.

1. INTRODUCTION

The UK government has recently been concerned with trying to improve the operational performance of the Maritime and Coastguard Agency (MCA). Since 1994, there have been a number of reviews of the effectiveness of Search And Rescue (SAR) operations coordinated by the MCA, culminating in the decision in 1999 to close a number of the Maritime Rescue coordination centres. However, there appears to have been a distinct lack of statistical analysis to justify this decision.

The aim of this research is to conduct a formal statistical analysis of the Coastguard SAR service in order to identify the key factors that influence its effectiveness and to provide a way of measuring this influence. An earlier study had been conducted with this aim, which made use of publicly available secondary data

(in the form of annual Incident Statistics) and developed a binary logistic regression model to support the assessment. The findings of this previous study highlighted the importance of explanatory variables such as the scale of incidents, staff workload and the length of coastline monitored by each coordination centre. Such variables may be interpreted as providing rather crude composite indicators for the multitude of factors that determine the effectiveness of this kind of complex operational system. The contribution of this current study is to support a more detailed explanation of the relationship between the operational effectiveness of the SAR service and identifiable causal factors.

In section 2 we present a background to the problem. In particular, we summarise some key issues that were raised during public inquiry into the closure of the coordination centres. In section 3 we present an outline of the structured elicitation processes we employed for this study and identify various compromises made during the implementation of the process. Section 4 is a summary of the key findings from the study and its shortcomings. In section 5 we reflect on the potential synergies between logistic regression and BBN's.

2 BACKGROUND

UK Coastguard (Maritime Rescue) coordination centres and sub-centres do not themselves undertake rescues. Instead, SAR equipment is strategically positioned around the UK coast to meet specified response times; the personnel using the equipment tend to be local volunteers. There are currently nearly 400 such sites around the UK. However, the staff at the coordination centres play the vital role of communicating with the casualty, assessing the scale of the distress, selecting the appropriate SAR response, and activating that response by alerting the relevant personnel and directing them to the casualty. They continue to fulfil this function until the emergency is over.

In the mid-1990s, there were 21 UK Coastguard districts, each with its own coordination centre. From 1994 onwards, there were a number of reviews of, and reports on, the Coastguard service concerning the organisation of these districts. From 1994 until

1996, the MCA conducted the *'Focus for Change'* review. One of its key recommendations was to introduce a new grade of Coastguard Watch Assistants to assist (and, in certain cases, to substitute for) the Watch Keeping Officers working in the coordination centres. This recommendation was implemented in 1996-97.

In 1997, the MCA set out their view of the future in the *'Five Year Strategy'* consultation document. This document contained two key proposals: (1) to invest in a new 'Integrated Coastguard Communications System' (ICCS), and (2) to reduce the number of co-ordination centres from 21 to 17. One of the main justifications given for the proposed closure programme was to provide greater exposure of Watch Keeping Officers to a wider spectrum of incidents (in order to produce more experienced staff and to provide greater job satisfaction).

In 1998, a report from the (parliamentary) Committee of Public Accounts expressed concern that without remedial action, 'local knowledge' would be lost if closures of coordination centres went ahead. In response to these serious concerns, the UK Government asked Lord Donaldson to conduct an overall *'Review of the Five Year Strategy for HM Coastguard'*¹. His report was published in August 1999. In this review, Lord Donaldson accepted the case for closing three coordination centres but rejected the case for the fourth. These recommendations were based on striking a balance between staff workload and the local knowledge required by staff to deliver an effective service.

Using the probability of a life being saved as the dependent variable, van der Meer et al² developed a binary logistic regression model to explain the difference in performance between the 21 different coordination centres. The average scale of incidents, staff workload and the length of coastline monitored (as a possible proxy for the local knowledge required) were found to be significant explanatory variables. This study found that the model developed was considerably more realistic and complex than any model that may have been used by the UK Government, and also that the coordination centres selected for closure were not necessarily the ones that were least effective in their primary purpose (i.e. to save lives).

Such regression studies, based on the use of aggregate statistics, can only be of limited use as at best they provide rather crude composite indicators for the multitude of factors that are actually involved.

This provided the motivation to conduct interviews with the watchstaff and construct a more detailed model of the factors that influence the effectiveness of operations. A BBN modelling approach was chosen as it would facilitate the integration of information from various sources, such as government reports and interviews with watchstaff personnel.

3 STRUCTURED ELICITATION PROCESS

Three semi-structured interviews were conducted with watchstaff each from different centres about the UK. This information was used to construct the influence diagram. Quantification of the influence diagram was obtained through an elicitation exercise with one person who was watch staff but is presently a union representative and therefore removed from the day-to-day operations of the centres.

This section describes the process used to elicit and validate the data acquired from the interviews. The process for quantifying the BBN was adapted from the general SRI³ approach to structured elicitation for bias management. At each stage we highlight potential biases within the data.

Step One- Identify Variables

Interviews were restricted to those members of the coastguard who responded to a request, which was conducted through the watchstaff union. This brings into question the validity of the interviews, as all respondents were self selected and could not be interpreted as being neutral given the political atmosphere within the organisation due to the closure of centres. It was felt that this bias could be kept to a minimum by careful wording of the questions and ensuring that the interview remained as impartial as possible. Three interviews were conducted from the Scotland and Northern Ireland regions.

The interviewees were asked a number of questions relating to the risk factors associated with the effectiveness of SAR co-ordination such as the effect of local knowledge, which outside factors influence the SAR operation and the staffing levels within the operations room. Each interview was recorded, with the interviewee's permission, to ensure that all details were accurately recorded.

After each interview took place the transcripts were written up within three days. Once all three interviews took place the answers to each section of the interview were

collated in order to find the common elements to each.

This information and other data gained from various government reports, internal communications and MCA reports was then used to identify all the relevant variables to this problem situation.

Step Two- Identify network Structure

The variables identified in the previous stage of the BBN process, along with the linkages identified by the interviewees were then used to build a causal map of the situation in order to gain an understanding of the influences and relationships involved.

There were found to be five nodes which directly influence the outcome "Effective Search And Rescue Operation"- "Scale of Rescue Operation", "Outside Factors", "Technological Failure", "Availability of Auxiliary Coastguards" and "Appropriate Staffing Levels in Operations Room".

Step Three - Model Validation

Once the model had been built it was given to the union representative, himself an experienced watch staff person to identify any shortcomings and errors/omissions within the model. It was intended that it would be circulated about other watchstaff personnel for further validation but due to this request coincided with a particularly busy time for the union as a whole. Therefore, the final influence diagram was only validated by the union representative.

Step Four - Conditional Probabilities

Often experts may find it difficult to express probabilities and find it hard to relate a number to their beliefs⁴. When people talk about probabilities they frequently prefer to use words rather than numbers. Verbal probabilities can be claimed to be more natural to produce and easier to understand than numerical probabilities, however, they can be harder to validate and more vague.

One issue relating to the use of numerical scales is the number of expressions available to describe varying degrees of probability. A scale with a smaller number of expressions is recommended as it is felt that it would be easier for the expert to distinguish between this number of categories, whilst still retaining a large enough number to indicate significant differences in probabilities.

Another issue associated with this form of describing probabilities is the fact that there will be variations in interpretation of each expression. Druzdzel⁵ reports that

previous studies have found significant within-subject consistency (experts are consistently using the same expression to represent the same level of probability). However, between-subject consistency (different experts are consistently using the same expression to represent the same level of probability) has reportedly achieved much lower consistency.

For the purposes of this research it was decided to adopt the scale as developed by Renooij and Witteman⁴, as it was felt that the clear use of the scale, along with relatively simple design facilitated the elicitation process more effectively than the more cluttered, and perhaps confusing scales available. In addition, the representation of the scale upon a line makes it easy for the expert to rank the probabilities in his/her own mind.

Figure 1 is a copy of the scale which was presented to the expert each time a probability was requested. The following five stages relate to the SRI process of eliciting subjective probabilities.

Step 4A Motivating

The purpose of the model was again presented to the union representative to ensure he understood the influence of each of the factors upon the effectiveness of a rescue of a person at peril as represented in the model. The structure of the model was discussed and checked to ensure that it was accurate. This also helped to verify the structure of the model. A number of changes took place as a result of this verification.

Stage 4B- Structuring

The definition of each of the variables was discussed, and the states of them to ensure they made sense.

The next part of this stage involved going over the scale to be used and defining each probability.

Stage 4C- Conditioning

The interviewee was asked if there were any major events that have occurred recently which could influence the probabilities he would be providing. He highlighted that much of the recent Trade Union activity has been involved with the Transport Select committee and evidence from this may have influenced his probabilities. It was discussed that as far as possible he should use only his own experience to base the probabilities on.

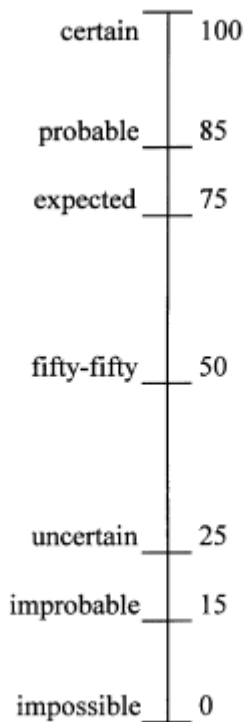


Figure 1 Linguistic/Numerical Scale

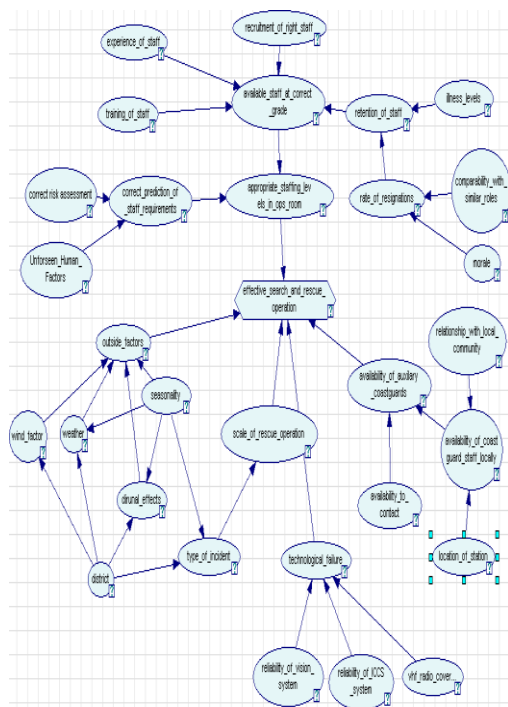


Figure 2: Bayesian Belief Network

Stage 4D- Encoding

The interviewee was invited to use the scale as in Figure 1. The probabilities were often initially assessed as frequencies, such as “once every week of standard watchkeeping routine”. However, it became clear that the interviewee was comfortable stating probabilities which were not on the scale, such as 40% etc. It was

felt that so long as he was producing probabilities which were consistent that this was more acceptable than forcing him to use a scale which quite arbitrarily chosen.

Probabilities for each cluster of nodes were elicited one at a time. At this stage there were also occasions when the structure of the model was changed slightly to reflect the interviewee’s view of the problem situation.

It helped to have the influence diagram in front of both the researcher and the interviewee as it focussed attention on the task at hand and simplified communication.

Stage 4E- Verifying

As stated above the structure of the model had been verified both at the motivating and encoding stages. In addition at frequent points throughout the encoding stage the probabilities were rephrased in order to verify them. The resulting BBN is illustrated in Figure 2.

4 REFLECTION ON THE ELICITATION PROCESS

There are several outcomes and learning points that have been obtained through the elicitation process as described above. These include issues concerning validity and triangulation of results.

There were several changes in the structure of the model, mainly concerned with the linkage between the district concerned, seasonality, weather and type of incidents that occur. In addition nodes relating to wind factor, comparability with similar roles and VHF radio frequency were added to the original model, and the node concerning network coverage was removed. In short, the structure continued to be modified after it had been “verified” when the numerical elicitation was being conducted.

The union representative himself has acknowledged the fact that he has been out of the Operations Room of a SAR facility for a number of years and as such his day to day experience of many of the factors within the BBN model are several years out of date.

Additionally, on several occasions the union representative referred to the fact that much of his information comes from other people within the field, indicating that it is not based upon his own experiences, which is not surprising given he has been operationally inactive as described above. This again calls into question the reliability of the data, as the information obtained for the model is secondary and not primary data.

However, it is felt that with future work these issues can be overcome by carrying out the elicitation process with many more experts in order to gain a consensus, and thus increase both the reliability and validity of the data.

With regards to staffing levels the model built shows that there must be a match between the needs of each station, as identified by the correct prediction of staff requirements and the number of available staff at each grade. If these two do not match there is the serious risk that staff will not be able to cope with the demands of the job, or conversely have so little contact with emergency situations that they will be unable to respond effectively. The importance of valuing staff through training, and recruitment of staff with the appropriate skills and competencies for the roles is also highlighted. This is echoed by the statistical analysis that was conducted with the logistical regression analysis and the findings of the Transport Select Committee.

5 Synergies Between Logistic Regression and BBN's

The findings of the logistic regression investigation into the secondary data identified the number of lives at risk per incident as an important environmental factor influencing the effectiveness of SAR operations. In fact this variable remained the strongest, most influential variable, indicating the size of an incident has the strongest influence upon the likelihood of a desirable outcome.

The BBN analysis concerning environmental factors identified several other factors that the coastguard do not officially measure, and consequently do not publish statistics. However, one of the variables also relates to the scale of the incident, which supports the findings of the statistical analysis.

The process of eliciting the BBN identified differences in performance due to the location of the centres. Differences were detected through the logistic regression between centres, however, due to the inclusion of the explanatory variable of length coastline, much of these differences were attributed to coastline length. Therefore, the outcome of the BBN study will inform the type of data to be collected to improve the logistic regression model.

Both the statistical and BBN analysis highlighted similar issues surrounding the staffing levels within an operations room. Follow up statistical analysis of more recent publicly available data highlighted that there

has been a change in the direction of influence of the workload of officers; simply higher workloads were associated with increasing the chances of a successful rescue but more recent data suggests it reduces the chances of a successful rescue. This was also supported by the findings of several Transport Committee reports, which found that consistent understaffing, deputising above grade and poor quality of new staff at lower levels have all lead to a decrease in the effectiveness of staff within an operations room.

The process of constructing the BBN's through multiple interviews resulted in a deeper understanding of validity (and lack thereof) of the surrogate measures used with the logistic regression. For example, the variable of staff was decomposed to capture the impact of Auxiliary Coastguard.

In summary, BBN's provide more explanatory detail to support a more effective operational improvements but validity of the resulting model is an issue. This was a concern with this study because of the political atmosphere that existed due to the closure of SAR centres. As such, we found that the binary logistic regression study was useful in providing validity to the inclusion of certain variables in the model and could be used to evaluate some of the quantification provided by coastguard watchstaff.

REFERENCES

1. Donaldson Lord (JF) *Review of the Five Year Strategy for HM Coastguard*. Department of the Environment, Transport and the Regions, London 1999.
2. Van der Meer, R.B., Quigley, J. and Storbeck, J.E., *Using Regression Analysis to Model the Performance of UK Coastguard Centres*, **The Journal of the Operational Research Society**, Oxford, June 2005
3. Merkhofer, M.W. *Quantifying Judgemental Uncertainty: Methodology, Experiences, and Insights* **IEEE Transactions on Systems, Man, and Cybernetics**, Vol SMC-17, No 5, September/October 1987
4. Renooij, S., and Witteman, C. *Talking probabilities: communicating probabilistic information with words and numbers*, **International Journal of Approximate Reasoning**, 22; 169-194 1999
5. Druzdzel, M.J. *Qualitative Verbal Explanations in Bayesian Belief Networks*, **Artificial Intelligence and Simulation of Behaviour Quarterly- Special Issue on Bayesian Belief Networks**, 94: 43-54 1996