

Challenges in Obtaining and Using Data in a Post-Emergency Environment

John Créquer
Senior Analyst Statistics
Canterbury Earthquake Recovery Authority
Christchurch
New Zealand
John.Crequer@cera.govt.nz

Abstract

In the immediate post-earthquake environment in Canterbury in 2011 data were shared relatively freely to assist with the emergency response. Once the initial response phase was over the needs changed to planning for the short- and medium-term. Decisions evolved over time from immediate emergency responses with short lead times to less well-defined decisions with greater lead times and longer horizons.

To inform planning and decisions data were needed to establish the current situation and monitor changes. Typically data were required to be highly geographically detailed and frequent which could not so easily be obtained from established and traditional sources. The challenges included: establishing the need for the data, obtaining data from non-traditional sources, obtaining co-operation of suppliers, assessing data sources and benchmarking new data sources, and combining data with differing definitions. The latter is particularly challenging when using these data in models.

The challenges of meeting these data needs in this environment will be discussed. Approaches employed in meeting these challenges and their success will also be traversed.

Key words: Decision analysis, post-emergency, data acquisition.

1 Introduction

A sequence of earthquakes began in the Canterbury region on 4th September 2010. The most significant events in terms of loss of life, injuries and damage were the earthquakes on 22 February 2011. A national state of emergency was declared on 23rd February 2011. The state was terminated on 30 April 2011 when responsibility for earthquake recovery was transferred to the Canterbury Earthquake Recovery Authority (CERA).

This paper is concerned with the data and decision-making challenges following the immediate emergency response. Some elements of that response, and in particular the Civil Defence response, are considered elsewhere (McLean, et al. 2012). The focus here is more on the recovery after the emergency than the response.

1.1 Environment

To understand the situation for decision making and planning, data were required, but the characteristics of the data differed from those normally published and easily

available. Similarly the situation was not a relatively stable steady-state but a quickly changing one with ill-defined competing demands.

Much of New Zealand's data collection infrastructure is designed for the production of frequent national statistics, in some cases with sub-national components. There tends to be a negative correlation between frequency and the fineness of geographic level. For example, monthly publications may have regional estimates but suburb-level data is typically only available after a Census.

The need in Canterbury was for frequent data, being both fine-level for geography and for person characteristics. To be more geographically specific, the three territorial authorities nearest the epicentres were of interest: Waimakariri District, Christchurch City and Selwyn District.

The post-emergency information environment could be characterised as having a need for details to inform policy but also information to refine the policy questions. In more normal circumstances it might be possible to refine a question over time in the light of data gathered for that purpose, given the more pressing circumstances secondary data sources were often used instead. For clarity, primary data are collected and used for a specific purpose while secondary data have been collected for one purpose but also used for another. A commonly used source of secondary data is administrative data, so named because it is collected for the administration of the government. Given the need for prompt decision-making, a number of secondary data sources were investigated.

However, the use of secondary data has a number of challenges, in particular privacy and fitness-for-purpose. The Privacy Act 1993 has a number of provisions relating to addressing privacy concerns but this paper concentrates on the fitness-for-purpose aspects.

1.2 Situation

In the post-emergency situation there was a need to make decisions and take action briskly but perhaps without the previous urgency. The questions of interest move from immediate food, shelter and clothing to medium- and longer-term decisions. Regardless, the need for data and information remain. Although the urgency is lessened, there is still a great impetus for quick decisions and the data to support them.

Given the history of management of events and their aftermath, it seemed reasonable to expect there might be some guidance available on processes. However, despite contacting various organisations, including the Global Risk Forum in Davos, there seemed to be a paucity of information on the post-emergency situation. It would appear others have found the situation similar, for example, Rubin (2009).

One of the difficulties in this situation is a constantly evolving understanding of the question, which further evolves as each datum is obtained, but often lacks an obvious stopping rule. Ideally, the value and cost of additional information can be compared but in this situation both can be unclear. For example, a regularly produced output or even a commissioned work will have a clear cost. The cost of obtaining secondary data and interpreting it is often unclear until it is obtained. The value of the information is clearer with a well-defined problem but can be obscure when the nature of the problem may be revised in its light.

The consequence of this situation is a great deal of revision of the problem definition and the data requirements.

2 Data and Decisions

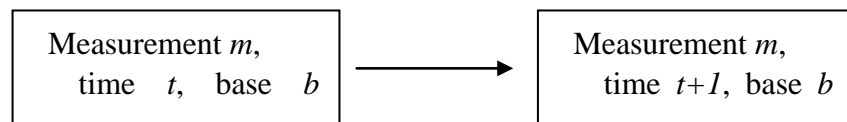
Although there were many decisions which had to be made, two elements which were frequent inputs were population and dwellings.

2.1 Population

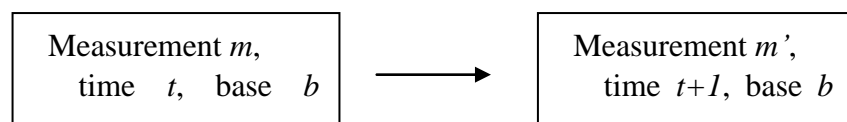
There was a consistent need for good data on population location and characteristics. In a stable situation the quinquennial Census provides a wealth of counts, demographic, location and other data about people at a fine geographic level. These data can be supplemented by annual population estimates. However, in the post-emergency environment population data needs become more frequent, specific and difficult to obtain.

Part of the difficulty in assessing population change is the lack of a base prior to the event. While the Census provides a base at the time it is taken, the ideal would be a base immediately prior to an event. This would provide a starting situation to modify in the light of data following the event. In the absence of such a base, the question of how a given sub-population has changed must be answered by simultaneously obtaining information on the current situation and the situation at some previous time. For example, if the Census represents the base and the event occurs two years later, then not only must the current situation be assessed but somehow the Census also needs to be updated to the event time. However, the measurement at the time of the event needs to be on a comparable basis to the Census measurement to enable a good estimate of the new situation.

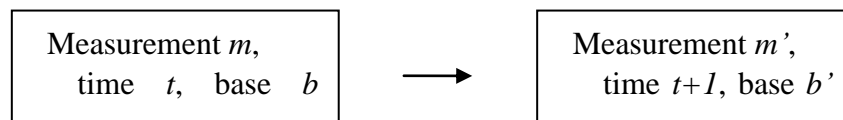
Ideal:



Less than Ideal:



Difficult:



As illustrated above, the ideal situation of a repeated measurement is a measurement of the same thing on the same basis with only time changed. A measurement of something related but using the same base can be helpful. For example, an update of the population estimates, which use Census data as their base. However, by the time we are measuring a different quantity on a different base with time incremented as well, comparisons become difficult. All too quickly alternate measurements and bases at differing frequencies were being compared to attempt to understand population changes.

One fruitful approach was to obtain data on the alternate base or measurement at the same time as the standard measures, enabling benchmarking of the alternative or

secondary source. Such data was not always readily available, whether because it related to a previous financial year, had never been available as a measurement before or was simply not readily retrievable. While goodwill extended to providing current data, finding historical data for comparison was, at times, a bridge too far.

It is of note that there are many organisations which have registers of members which can be used to indicate populations. An evaluation of these was carried out by Dunstan and Ryan (2011). In trying to reconcile some of these to understand population change the question of whether, after future events, some type of overall population register might be useful. While it may not capture the immediate post-event situation it might greatly assist with the recovery phase. A promising source of data for the post-event situation, cellphone data, was described in Nissen (2012).

2.2 Dwellings

Given the damage to dwellings following the earthquake events, and the population changes, estimates of the occupiable dwellings were in demand. The starting stock of dwellings could be estimated from Census data supplemented by building consents. However, the extent of unoccupiability of dwellings was, and remains, a problematic question. Again, there is no one source of information on whether a dwelling is occupiable or whether it is occupied. Whether a dwelling is occupied has been complicated at various times by the imposition of s124 notices under the Building Act 2004, effectively prohibiting occupation of the building.

Accounting for the number of dwellings available faces the usual problems of definition. Distinctions between dwellings, multiple dwellings in a single house, multiple land parcels for one dwelling or vice versa have all contributed to complications in assessing the number of dwellings.

Reconciliation of different data sources is also compounded by the alternative uses for the lists of dwellings or properties etc. For rating purposes councils are interested in the property and improvements, for land ownership the land boundaries, and for services the various connection points to the network. A further difficulty is the lack of an implemented addressing standard, so that the same building may have multiple addresses across different sources.

A register of dwellings would have provided benefits on a number of occasions. Once matched to the register it would have been unambiguous which dwelling was referred to. Most importantly, it would have given all agencies involved a shared reference point for their data.

3 Discussion

The post-emergency period has continued to exhibit the goodwill seen in the emergency period for co-operation in data provision, although the concentration has been moving to the organisations with a longer-term interest in the respective topics. Even as this settling occurs, the matter of obtaining data and reconciling it with other sources remains.

At the start of the post-emergency period, competing sources and data were at times a challenge. While there was the intent of helpfulness in much of what was made available, especially publically, this contributed to confusion at times. It was generally accepted that the Official Statistics were the benchmarks but there was something of a

vacuum between the periodic releases by Statistics New Zealand and others. Given that nature abhors a vacuum, others moved to fill the perceived gap. Where the methodology and data sources were comparable to the Official Statistics this was helpful, but at times the limited range of data available to these providers did not aid reconciliation to the benchmarks. In fact, having additional estimates not only contributed to confusion but reduced the available effort in interpreting higher quality sources due to the diversion to understand and explain how the supplemental estimates might or might not be suitable.

In addition to the competing sources, consistency of interpretation was also problematic. This is not to suggest universal consistency might be desirable but that it might be within the core participants in the post-emergency situation. However, the press of time and the need for one agency to lead tends to be inimical to consensus building.

It would be ideal to have a single repository of data from which data was drawn for the various needs. These data would be available with their metadata. This would aid consistency of interpretation and reduce the problem of using the wrong version of the data. How practical this might be is another question.

At least two further topics arise from the post-emergency Canterbury experience: the role of statisticians and decision analysts, and preparation.

3.1 What is the role of decision analysts and statisticians?

There are two aspects to our role in these situations: what we understand our contribution could be and what those in managerial positions see as our role. To be clear, for this discussion decision analysts includes operations researchers.

I expect that most people at this conference would consider that we could assist with refining the question to be decided and evaluating data for making that decision. Typically we have the skills to understand both the essential elements of the presented question and help refine the problem definition, as well as assist with suitable data for the question.

However, it is perhaps less clear to what degree our possible contribution is understood. While many subjects employ elements of statistics or operations research, the benefit of a specialist, able to work amongst a range of work streams, may not be so well appreciated. How best to communicate what we can offer is a question which has occupied both professions.

3.2 Preparation

The co-operation of people who had access to data which was able to assist was outstanding. Various organisations made available, often at short notice, data which was able to help with decision making. However, as the time from the events increased it was noticeable that the more co-operative groups were those which had direct or indirect responsibility for aspects of the recovery. Understandably, those organisations whose core business was elsewhere naturally concentrated on that once the immediate urgency had passed.

It seems reasonable to assume that New Zealand will again face an emergency situation of a large scale, of a similar magnitude to that faced in Canterbury. In the same way that emergency plans are made perhaps statisticians and decision analysts

should also consider developing processes and procedures for selecting, cataloguing and sharing data in times of post-emergency and recovery.

At an informal level, our own networks and conferences, such as this, allow us to keep up with what our colleagues are working on. However, informal networks can be constrained by managerial, organisational or other prerogatives. Goodwill, while welcome and commendable is not really an adequate substitute for planning. Perhaps a more formal and widely recognised structure would be more appropriate, with support or endorsement from government as required.

A more formally recognised network might assist decision makers in the post-emergency environment and may also supplement current procedures employed by the NZ Police and the Ministry of Civil Defence and Emergency Management. Its importance lies in having the people who are familiar with their data and its analysis able to contact each other and discuss needs and how they may be met.

The idea of a network could be extended to include a neutral custodian where required. For example, some commercially sensitive data may be very useful during recovery but the holder may be reluctant to share it. A neutral custodian, which experience sound experience in confidentiality and security, may be ideally placed to facilitate the receipt and compilation of such data and make it available to the required parties.

Whatever solution was possible, it would be hoped that it would reduce the time taken to consider data sources, approach organisations, establish what data were available, establish protocols for its sharing and use, obtain the data from the custodian's system, provide the data to the requestor's system and interpret the data. Avoiding an ad hoc approach would reduce the time spent on the mechanics of obtaining the data enabling more time to be spent interpreting and using them.

A particular example of an ad hoc request was the use of cellphone data immediately post-event to estimate how many people had left the Christchurch area (Nissen, 2012). If protocols were put in place in advance to obtain these data in specific situations then they may be able to be used more effectively after future events.

4 Conclusion

The post-emergency period required data and information for decision making in the short- to long-term. Non-traditional sources of data were obtained to attempt to address those needs. Much effort was required to deduce how those data were able to be used, due to the wide variation in measurement methods, bases and timing of collection. The necessary meta-data was often even more of a challenge. Benchmarking to established official measures was a useful practice.

While the ad hoc response was commendable there is a clear need for planning and preparation in the data space for future events. Protocols established beforehand would allow a more timely response as well as a better understanding of what data may be available. Informal networks are invaluable and could be supplemented, but should not be supplanted, by more formal arrangements. National standards and possibly registers for dwellings and population would assist future responses, both short- and long-term.

Acknowledgments and Disclaimer

Although the author is an employee of the Canterbury Earthquake Recovery Authority (CERA), the opinions and conclusions reached in this paper are personal to him and do not represent the opinions of CERA. The author acknowledges the time and support from his employer to produce this paper. It is also acknowledged that the paper is of an abstract nature; should readers be interested in further details then they are invited to contact the author.

5 References

- Dunstan, K., and Ryan, M., 2011. "Evaluation of alternative data sources for population estimates." Statistics New Zealand.
- McLean, I., Oughton, D., Ellis, S., Wakelin, B., and Rubin, C. B., 2012. "Review of the Civil Defence Emergency Management Response to the 22 February Christchurch Earthquake". http://www.civildefence.govt.nz/memwebsite.nsf/wpg_URL/For-the-CDEM-Sector-Publications-Review-of-the-Civil-Defence-Emergency-Management-Response-to-the-22-February-Christchurch-Earthquake, last retrieved 12 November 2013.
- Nissen, K., 2012. "Using cellphone data to measure population movements". Statistics New Zealand.
- Rubin, C. B., 2009. "Long Term Recovery from Disasters—The Neglected Component of Emergency Management." *Journal of Homeland Security and Emergency Management*, **6** (1), Article 46.