RiskScape and Wellington Electricity Restoration Uncertainty Analysis

Y Syed
N Horspool
R Buxton

GNS Science Internal Report 2019/08
November 2019

GNS SCIENCE
TE PŪ AO

DISCLAIMER

The Institute of Geological and Nuclear Sciences Limited (GNS Science) and its funders give no warranties of any kind concerning the accuracy, completeness, timeliness or fitness for purpose of the contents of this report. GNS Science accepts no responsibility for any actions taken based on, or reliance placed on the contents of this report and GNS Science and its funders exclude to the full extent permitted by law liability for any loss, damage or expense, direct or indirect, and however caused, whether through negligence or otherwise, resulting from any person's or organisation's use of, or reliance on, the contents of this report.

BIBLIOGRAPHIC REFERENCE

Syed Y; Horspool N; Buxton R. 2019. RiskScape and Wellington Electricity Restoration Uncertainty Analysis. Lower Hutt (NZ): GNS Science. 21 p. Internal Report 2019/08.

Syed, GNS Science,1 Fairway Drive, Lower Hutt, New Zealand
Horspool, GNS Science,1 Fairway Drive, Lower Hutt, New Zealand
Buxton, GNS Science,1 Fairway Drive, Lower Hutt, New Zealand

CONTENTS

ABSTRACT .... **III**

KEYWORDS .... **III**

1.0 INTRODUCTION **1**

2.0 IMPACT MODELLING **2**

2.1 Introduction 2
2.2 Asset Data 2
2.2.1 Building and Population Data 2
2.2.2 Electricity Assets 2
2.3 Fragility Functions 3
2.3.1 Buildings 3
2.3.2 Human Casualties 4
2.3.3 Electricity Network 5
2.4 Hazard Model 6
2.5 Uncertainty Quantification 6
2.6 Limitations 7

3.0 CLUSTERING **8**

3.1 Some context 8
3.2 Introduction 8
3.3 K-Means Algorithm 8
3.4 Hierarchical Algorithm 9
3.5 Approaches for Measuring the Goodness of Clusters 11
3.6 Results 12

4.0 AUTOMATED OUTAGE MAP DEVELOPMENT **15**

4.1 Restoration Assumptions 15
4.2 Integrated approach for Linking Electricity and Road Networks 15

5.0 CONCLUSIONS BASED ON THE CLUSTERING WORK **19**

6.0 ACKNOWLEDGEMENTS **20**

7.0 REFERENCES **20**

FIGURES

Figure 2.1 Electricity assets used in modelling. Red triangles are substations, black squares are transmission towers and blue lines are overhead or underground cables. 3

Figure 2.2 Example of two realisations of ground shaking from modelled Mw7.5 Wellington Fault scenario. 6

Figure 2.3 Example of two different realisations of damage to electricity assets. 7

Figure 3.1 Example dendrogram (Galili, 2019). 10

Figure 3.2 An example Silhouette plot. 11

Figure 3.3 Silhouette plot for the damage data cluster with k = 2. 12
Figure 3.4 Silhouette scores per number of clusters for the recovery time data. 13
Figure 4.1 An integrated methodology for modelling electricity and road networks. 16
Figure 4.2 Electricity network outage maps on different timestamps (a) with and (b) without road network dependency. 17
Figure 4.3 An example output showing comparison of different outage maps which can be seen in different timestamps. 18

TABLES

Table 2.1 Damage state definitions for buildings, and indicative consequences. 4
Table 2.2 Casualty state definitions. 5
Table 2.3 Electricity assets vulnerability models. 5
Table 3.1 The Silhouette scores for different numbers of clusters for the recovery time data. 13
Table 3.2 The membership numbers for each of the 80 clusters in the recovery time data. 14

ABSTRACT

Machine-learning and Artificial Intelligence approaches are becoming increasingly popular in all areas of science. Here we examine clustering techniques as an approach for simplifying the analysis of uncertainties in the domain of hazard and risk modelling, particularly when dealing with large Monte Carlo simulations combined with the RiskScape and MERIT modelling systems.

RiskScape is a multi-hazard risk evaluating tool, jointly developed by NIWA and GNS Science. It provides estimates of the risk to exposed assets (e.g. infrastructure networks such as road) from natural hazard scenarios. MERIT is an economic modelling toolkit designed to take RiskScape outputs as inputs and provides estimates of the wider economic impacts of critical infrastructure outages.

Here we use a scenario relating to a Wellington Fault earthquake and it's impacts on critical infrastructure as an example case study. In particular, the impact to the electricity network is used as an example to refine the approach. Three sets of results were examined, firstly the raw Monte Carlo results, then clustered damage data, lastly clustered recovery time data.

This work has revealed that the damage state data resulted in poorly resolved clustering, but only two clusters and the recovery time data resulted in better resolution for the clusters, but still a reasonably large number of clusters (80) that would theoretically, incur a substantial time overhead for experts to review. It is maintained, however, that conceptually both of the clustered approaches offer substantial advantages over the raw Monte Carlo results where experts cannot be expected to effectively process many different sets of results.

KEYWORDS

Clustering, Artificial Intelligence, Machine Learning, natural hazard modelling, RiskScape, MERIT

This page left intentionally blank.

1.0 INTRODUCTION

Machine-learning and Artificial Intelligence approaches are becoming increasingly popular in all areas of science. Here we examine clustering techniques as an approach for simplifying the analysis of uncertainties in the domain of hazard and risk modelling, particularly when dealing with large Monte Carlo simulations.

RiskScape is a multi-hazard risk evaluating tool, jointly developed by NIWA and GNS Science. It provides estimates of the risk to exposed assets (e.g. infrastructure networks such as road) from natural hazard scenarios. RiskScape is a freely available. It is modular in design and requires a user to specify hazard, asset and fragility models. Hazard models represent the spatial distribution of the hazard and its intensity. Asset data defines the location and attributes of the elements of interest such as buildings, people, infrastructure or other assets. Fragility functions bring together the hazard and asset data and are functions that estimate the impact conditional on the hazard intensity and asset attributes. RiskScape produces estimates of the spatial distribution of damage or loss from specified scenarios (King and Bell, 2009).

MERIT is an economic modelling toolkit designed to take RiskScape outputs as inputs and provides estimates of the wider economic impacts of critical infrastructure outages. MERIT was developed by Market Economics, Resilient Organisations and GNS Science.

Here we use a scenario relating to a Wellington Fault earthquake and it’s impacts on critical infrastructure as an example case study. In particular, the impact to the electricity network is used as an example to refine the approach. When undertaking modelling of this sort it is important to understand the uncertainties arising from the modelling process so that these can be allowed for in any decision making based on modelled results. However, complex models can be difficult to characterise without using approaches such as Monte Carlo simulations, and this is the basis of the work described here. Monte Carlo simulations depend on the modelling of many realisations of a scenario to be able to properly gauge the variance and hence the uncertainty encompassed by the modelling process.

Whilst this is useful, it can be difficult to interpret results particularly when the results are expressed as maps or time stamped maps which is the case with the RiskScape/MERIT process. Here the usefulness of clustering as an approach to arrive at a set of representative patterns is briefly discussed.

This work describes an approach whereby:

1. The RiskScape tool is used to generate 1000 different representations of the Wellington Fault earthquake scenario in terms of the estimated damage sustained by the Wellington Electricity network due to each earthquake realisation.
2. The RiskScape generated information is used as a basis for clustering in an attempt to reduce the number of sets of outage maps from 1000 to some more easily manageable number.
3. The outage information is also used as a basis for clustering as above.
4. All three (Monte Carlo, clustered damage, clustered outage times) sets of information are passed through the MERIT model providing three sets of economic trajectories.

2.0 IMPACT MODELLING

2.1 Introduction

This section provides an estimate of potential impacts to buildings, humans, and electricity assets from a Mw7.5 Wellington Fault earthquake. The impact to buildings (potential physical damage and downtime), people (injuries and fatalities) and electricity assets (potential physical damage and time required to bring back lost service) is estimated using the RiskScape tool with inputs from Wellington Electricity. The impact modelling includes uncertainty in the various scenario components by considering a range of possible earthquake ground motion scenarios, diurnal population variations (day and night time) as well as uncertainty in the estimates of asset damage and casualty states. By including this uncertainty, 1000 different impact scenarios were generated. Results were then used in further modelling for recovery times of the electricity network, relocation of injured people, cordoning of severely damaged buildings, and disruption to businesses in damaged buildings.

2.2 Asset Data

2.2.1 Building and Population Data

Building and population exposure for the impact modelling is based on the national building and population database developed by the RiskScape Project. This database contains building level information for each building in New Zealand, including attributes that are relevant to modelling such as construction type, age, floor area, replacement value and others. The building data is primarily based on information sourced by local councils for rates evaluations and has been updated with local surveys by the RiskScape project. The data is thought to be accurate at a regional level and inaccuracies likely exist for individual buildings due to some of the statistical methods applied to estimate some building attributes (King and Bell, 2009).

The asset database also includes estimates of occupancies for day and night time (as estimated at time of census usually in March 2013 with extrapolations applied for subsequent years). For further information on the RiskScape building and population database see King and Bell (2009) or the RiskScape wiki[1].

2.2.2 Electricity Assets

Data relating to electricity assets were provided by Wellington Electricity for a previous damage modelling project and were available for use in this project. The asset data contains the location and attributes of electricity assets including transmission towers, substations and local distribution networks. The attribute data was mapped to RiskScape fragility classes for electricity assets.

The Wellington Electricity network includes the national grid network and the local distribution networks. The following components (also see Figure 2.1) were considered in the analysis:

1. **Transmission Structure** — no classifications
2. **Substation** — classified by voltage level, construction type, age and condition
3. **Buried Cables** — classified by length, voltage level, and conduction material
4. **Overhead Cables** — no classifications

[1] https://wiki.riskscape.org.nz/

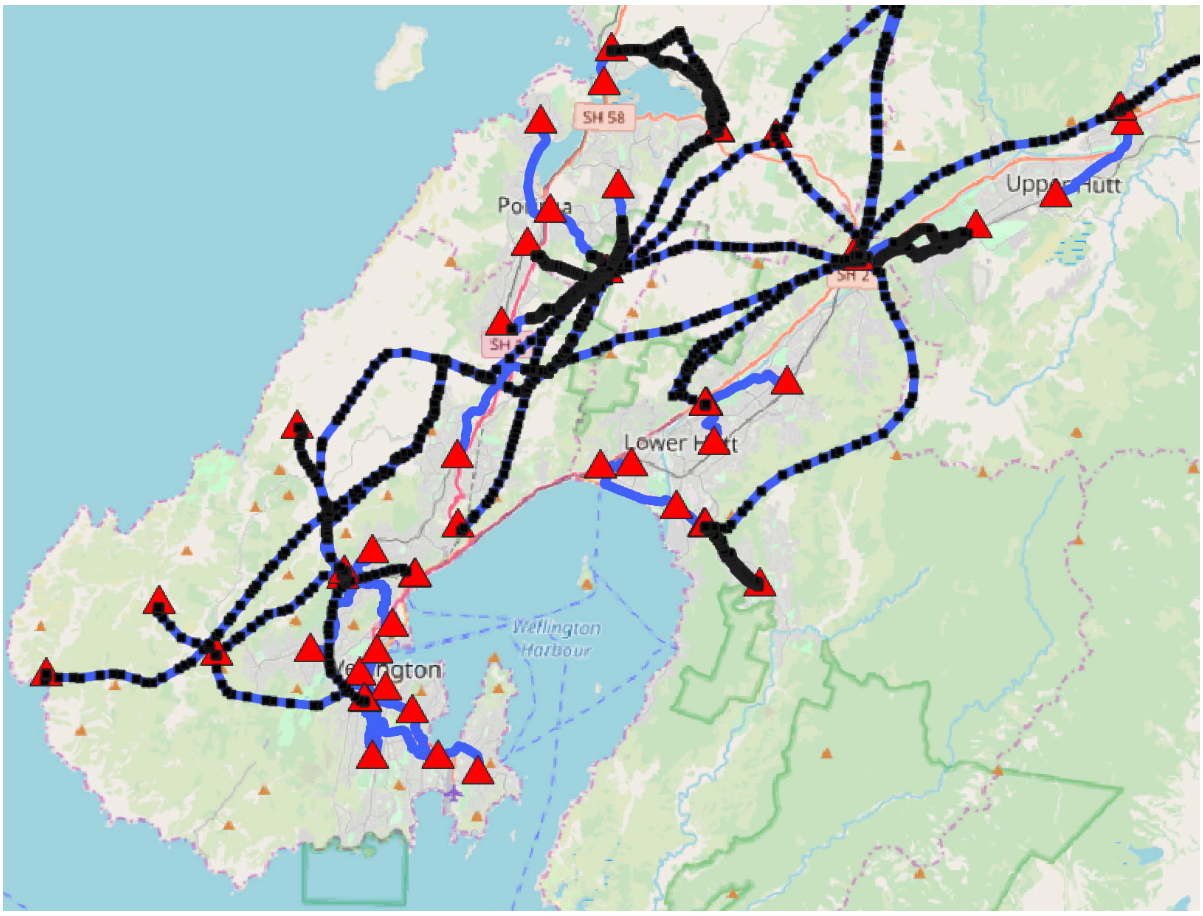

Figure 2.1 Electricity assets used in modelling. Red triangles are substations, black squares are transmission towers and blue lines are overhead or underground cables.

2.3 Fragility Functions

2.3.1 Buildings

Fragility models are used to estimate the probability of being in different damage states for a given shaking intensity. The default RiskScape fragility functions are used for the modelling and use MMI and PGA as the ground shaking intensity measure. Fragility models are assigned to different building types (e.g. timber frame, reinforced concrete, unreinforced masonry etc) based on the characteristics of the buildings. Fragility models are probabilistic in such that for any given level of shaking there is a probability of being in different damage states. In order to capture this uncertainty a total of ten different runs were produced for each scenario. The damage state definitions are shown in Table 2.1.

Table 2.1 Damage state definitions for buildings, and indicative consequences.

Damage State	Downtime	Structural Damage	Injuries	Deaths	Description
DS0: None	-	No	No	No	No damage
DS1: Light	7 days (1–7)	No	Rare	No	Damage to claddings, small cracks in concrete and masonry infill walls, cracks in interior walls and ceilings
DS2: Moderate	30 days (7–180)	Yes	Few	No	Cracks in columns and beams of frames and in structural walls. Cracks in partition and infill walls, fall of brittle cladding and plaster, fall of mortar from joints of wall panels
DS3: Severe	365 days (180–730)	Yes	Some	Rare	Cracks in columns and beam-column joints of frames. Spalling of concrete cover, buckling of reinforcing rods. Large cracks in partition and infill walls, failure of individual infill panels.
DS4: Partial Collapse	730 days (730–1095)	Yes	Many	Some	Large cracks in structural elements with compression failure of concrete and fracture of reinforcing bars. Bond failure of beam reinforcing bars, tilting of columns or buildings. Collapse of a few columns or of a single upper floor. Parapets, gables, and unreinforced walls may fall. Volume Loss < 50%.
DS5: Collapse	730 days (730–1095)	Yes	Many	Many	Collapse of ground floor or parts (e.g. wings) of buildings. Volume loss of 50% or more (e.g. at least 50% of the building has pancaked).

2.3.2 Human Casualties

Casualty models estimate the probability of being in a given casualty state (injuries or deaths) as a function of building damage (Cousins, 2008). The default RiskScape casualty models have been applied and then adjusted with an injury amplification factor to include increased number of injuries seen in recent earthquakes in New Zealand (Johnston et al, 2014) when compared with RiskScape models. Additional casualties are estimated for people outside of buildings and from other causes such as landslides (Cousins, 2008).

Table 2.2 Casualty state definitions.

Casualty State	Description of State
CS1: Uninjured–Light	[First-Aid] Injuries that can be self-treated or treated by a "first aider". Examples are bruising/contusion, minor cuts, sprains.
CS2: Moderate Injury	[Doctor] Injuries that require expert treatment (paraprofessional or doctor), but which are not immediately life threatening if such treatment is not available. Examples are cuts requiring stitches, serious sprains, dislocations, significant burns (first degree, or second degree over small part of body), minor concussion (unconscious < 1 hr).
CS3: Serious Injury	[Hospital] Injuries requiring a greater degree of medical care and use of medical technology such as x-rays or surgery, but not expected to progress to a life-threatening status, full recovery expected with suitable treatment. Examples are: open head or face wounds, concussion (unconscious > 1 hr), fractures (open, displaced or comminuted), dehydration or exposure, serious burns (third degree over small part of body, or second degree over large part of body).
CS4: Critical Injury	[Intensive Care] Injuries that pose an immediate life-threatening condition if not treated adequately and expeditiously, or long-term disability. Examples are brain damage, spinal column injuries, nerve injuries, crush syndrome, internal organ failures due to crushing, organ puncture, other internal injuries, uncontrolled bleeding, traumatic amputations of arms or legs.
CS5: Death	[Undertaker] Well defined state.

2.3.3 Electricity Network

Impacts to the electricity network were modelled in RiskScape with the vulnerability models presented in Table 2.3 applied to each asset. Overhead cables use transmission tower damage as a proxy for impacts, due to a lack of available vulnerability models. Substation buildings are modelled using the same methods as with buildings, with the final impact for each substation being assigned using a logic-based approach to define the critical damage state (i.e. for plant components or building).

The damage ratio curve for buried cables as presented in Lin et al, 2016, was refined following workshops with the network providers, and a new vulnerability model developed to reflect a more robust cable network in the Wellington region compared with that of Christchurch, of which the model is based on. Cables are segmented into approx. 50m lengths for modelling.

Table 2.3 Electricity assets vulnerability models.

Asset Type	Hazards		
	Ground Shaking	Liquefaction	Binary Functions
Substation (zone & GXP)	PGA: Federal Emergency Management Agency, 2015	LSN: (Rosser and Dellow, 2015) and engineering judgement	Within Hazard foot print = DS4, not within hazard footprint = DS0
Buried cable	MMI: Lin, Nayyerloo and Zhang, 2016		
Transmission tower or pole	PGA: Xie *et al.*, 2012		

2.4 Hazard Model

The hazard scenario used for the model is a Wellington Fault Mw7.5 earthquake. The hazard model includes the earthquake source (geometry and magnitude), the site model representing soil properties, and ground motion prediction models that estimate shaking from the source to site of interest (i.e. asset location). OpenQuake was used for generating the hazard layers?

There is no uncertainty in the source geometry or the magnitude applied, as the Wellington Fault is characterized by a single magnitude and source in the New Zealand National Seismic Hazard Model. Further work would be required to develop uncertainty in earthquake source and magnitude for modelling and is therefore out of scope for this project. The site model is a single site model used in RiskScape that is developed from the New Zealand Site Class Map.

Uncertainty in the ground motion estimation is from two primary sources. The first is epistemic uncertainty represented by different choices in published ground motion prediction models (GMPE), and the second is the variability about the median ground motions for a given magnitude, distance to source, and soil class. The latter uncertainty is classified as aleatory uncertainty as it is natural variability in observed data used to develop the models. The GMPE uncertainty is represented by a logic tree based on that of Van Houtte (2018). Five GMPE are used for active crustal sources and each is given a weighting. For each ground motion realisation (N=1000), first the GMPE is randomly selected based on the weights, then variability in ground motion is modelled, by sampling from a normal distribution with a mean of zero and standard deviation that of the GMPE. The result is 1000 ground motion maps which represent variations due to the ground motion uncertainty (Figure 2.2).

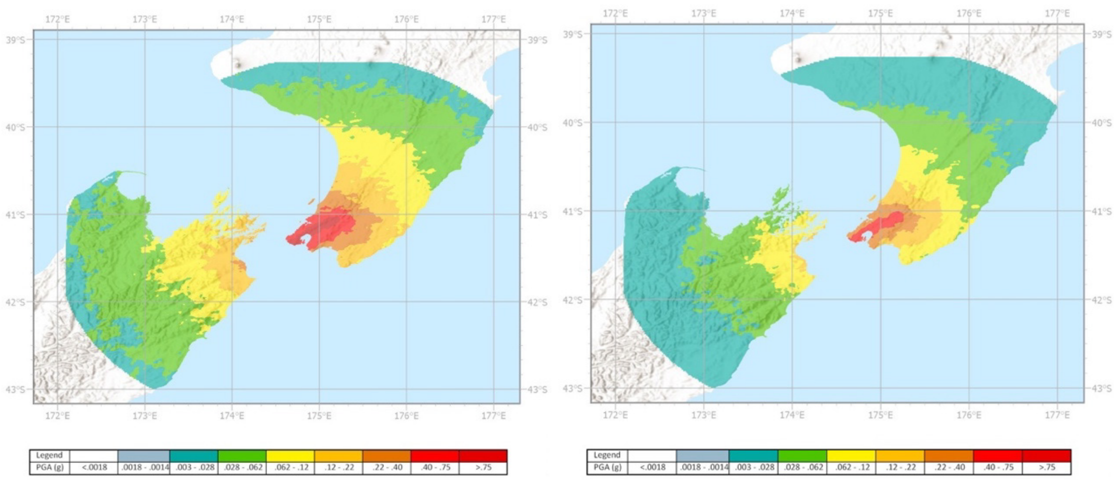

Figure 2.2 Example of two realisations of ground shaking from modelled Mw7.5 Wellington Fault scenario.

2.5 Uncertainty Quantification

Uncertainty is propagated through the various components of the impact modelling. First the 1000 ground motion maps are generated in the hazard model. For each realisation, the assets are spatially joined to the hazard maps. The hazard intensity at each asset is extracted and used as input into the fragility function.

Uncertainty in the fragility function is represented through sampling of the probability of a given damage state conditional on the ground motion intensity. For population data, either a day or night time exposure is randomly selected, and for building and electricity assets, they are fixed for all realisations.

Table 3.1 provides a summary of the various components of uncertainty in the modelling.

The result of the modelling is 1000 realisations of building damage states, human casualties (injuries and deaths), and damage states for electricity assets.

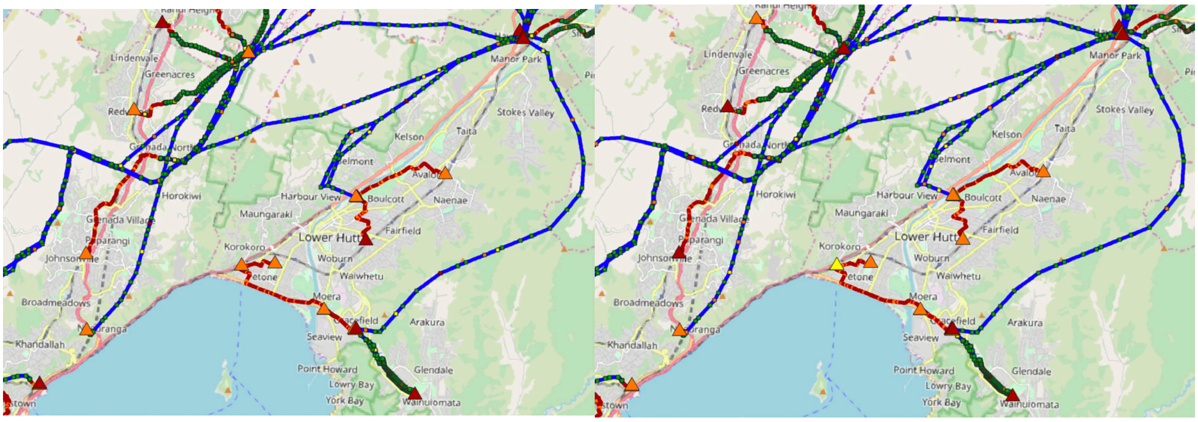

Figure 2.3 Example of two different realisations of damage to electricity assets. Dark red is complete damage, orange is severe damage, yellow is moderate damage, green is minor damage and blue is no damage.

2.6 Limitations

A number of limitations exist in the modelling method used. First, the model does not quantify uncertainty for all model components such as asset attributes or locations, soil properties, earthquake scenario magnitude and geometry, and epistemic uncertainty in what fragility functions are used. However, while this would be an exhaustive quantification of uncertainty it was not practical for this study as it would require development of uncertainty models for those components which are outside the scope of this project. Second, the method used for uncertainty quantification is 'brute-force' Monte Carlo simulation. There are alternative approaches that isolate model components that contribute the most variation to the final output, and these are sampled in a more granular approach. Further work could explore these methods. Finally, the scenario used for the modelling was selected because models already existed for this scenario However, a limitation of this scenario is that ground motions and impact is very severe and for high ground motions the variability in the fragility functions is small as most assets are in high damage states. This results in only small variability in the final results. If a more moderate and 'threshold damage' event was used, or a range of magnitudes, then it would be expected that more variability in final results would be seen.

3.0 CLUSTERING

3.1 Some Context

The previous section of this report outlined the process for allowing for the uncertainties in the hazard modelling process using a Monte Carlo approach. The results of this approach are many different versions of the hazard damage profile, which taken to its logical conclusion would correspond to many different sets of recovery maps. An automated approach for deriving the different sets of recovery maps has been developed and this is discussed in the next section, however, the question we ask here is “can AI approaches reduce the complexity of this dataset so that a smaller set of patterns can be used to represent the 1000 sets” coming from the Monte Carlo simulation. If so, then it is theoretically possible that human experts might be able to validate the automated process.

An unsupervised technique called clustering was chosen as a candidate approach for dimensionality reduction and the background of this approach is now discussed.

3.2 Introduction

Clustering is a technique commonly associated with Artificial Intelligence and Machine Learning that are used to analyse or transform data by grouping similar items close together in information space simultaneously maximizing the distance between dissimilar items.

Being able to group similar data together has many different applications in many different science domains. One of the reasons for this variety is the lack of a universally recognised definition of a cluster. Attempts to undertake clustering operations on datasets predates electronic computers. Initial approaches were proposed by Driver and Kroeber (1932) in the field of anthropology other early applications include work in psychology by Zubin (1938).

More recent work has yielded a bewildering array of models, but k-means and hierarchical approaches still seem to be the most popular and well supported algorithms. These approaches are briefly discussed. For a more extensive treatment of possible approaches to clustering, the reader is directed to Omran et al (2007).

Whereas, often, clustering is seen as a way of classifying like objects, here we propose an extension to the use of clustering in the domain of hazard and risk modelling. We show that, in some circumstances clustering can be used to reduce the dimensionality of hazard/risk outputs in a useful fashion.

3.3 K-Means Algorithm

One of the most well known and most used algorithms for clustering is k-means (first used by MacQueen, 1967; from an idea by Steinhaus, 1956). K-means was originally used in the domain of signal processing but has proven to be versatile in its application. The algorithm partitions observations into “k” clusters, where “k” is a parameter nominated by the user before the commencement of the clustering process.

The objective is to create clusters where the distance between the members and the centroid of the clusters is minimized. Over a number of iterations of the algorithm the for membership of the clusters change to reflect this aspiration.

Example pseudocode of the k-means algorithm below:

1. Generate k cluster centroids at random.
2. Assign each of the observations to a cluster with the least means squared Euclidean distance.
3. Update the centroids based on the new cluster memberships.

The algorithm converges when the observation assignments no longer change.

3.4 Hierarchical Algorithm

The other popular approach to clustering are based on hierarchical algorithms. In hierarchical clustering techniques two high level ideas are used:

1. Agglomerative algorithms that start with each observation as a cluster and then strive to merge the nearest observations to create larger and larger clusters,
2. Divisive algorithms that start with a single cluster and then iteratively splits off the observations that are the furthest from the other observations creating more clusters.

With these techniques there is no requirement for the user to nominate a number of clusters. Results of hierarchical techniques are normally portrayed as a dendrogram (Figure 3.1). The popular, flexible and well supported k-means clustering algorithm was used as a basis for the clustering carried out in this work.

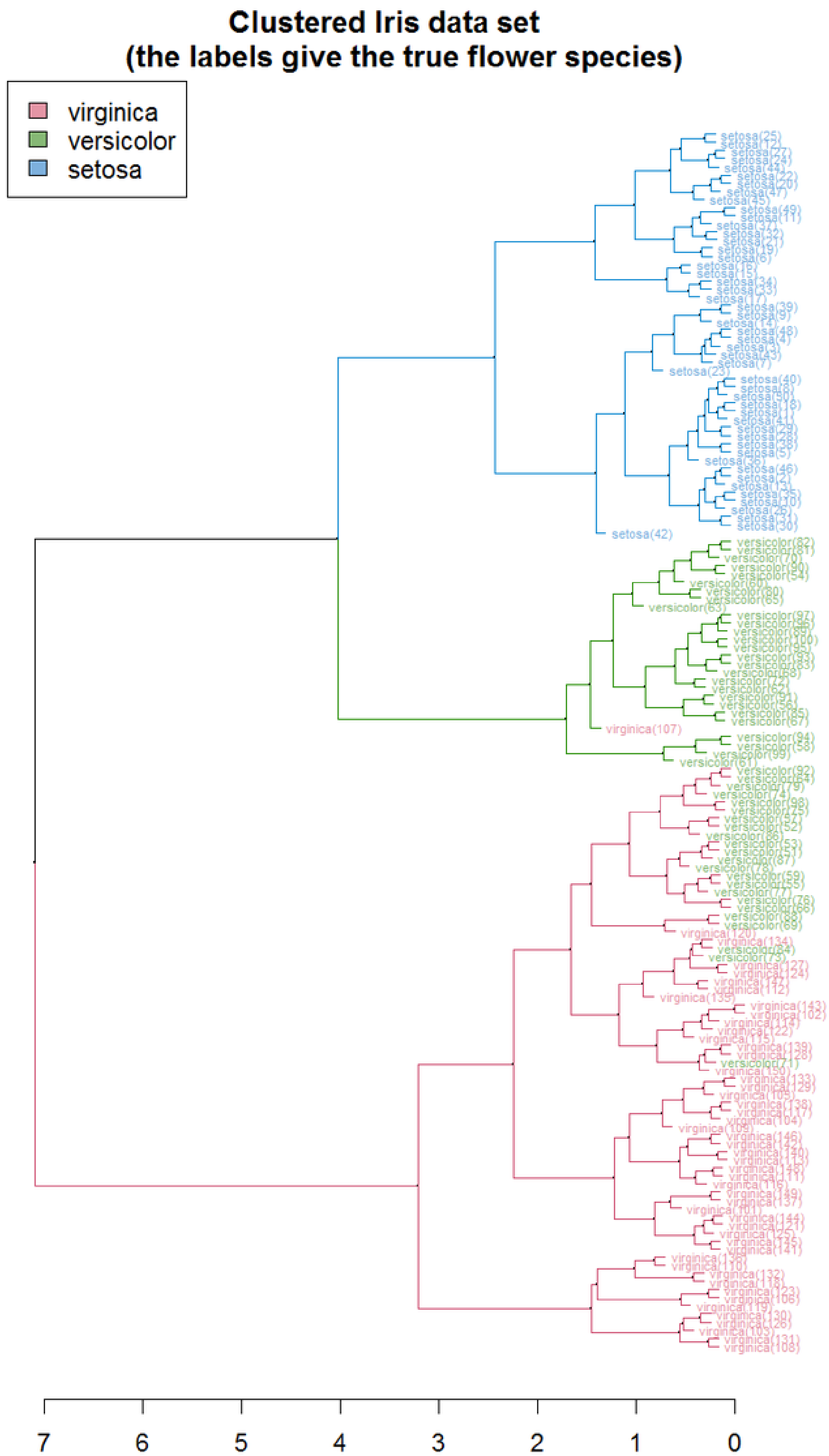

Figure 3.1 Example dendrogram (Galili, 2019).

3.5 Approaches for Measuring the Goodness of Clusters

Many approaches exist for measuring or validating clustering results, but it remains a challenging problem with no single approach gaining universal support. With many approaches there remains some element of subjective manual judgement as to the quality of the clustering results. A thorough discussion of the possible approaches is beyond the scope of this report, but readers are directed to work by Omran et al (2007) for a more extensive discussion of work in this area.

In this work we utilize the Silhouette score as a basis for measuring the quality of the clustering results. The silhouette score is a measure of how similar an observation is to its own cluster compared to other clusters. It is a single score in the range -1 to +1 where a high positive value indicates a "good" clustering configuration and zero or negative values indicate "poor" clustering.

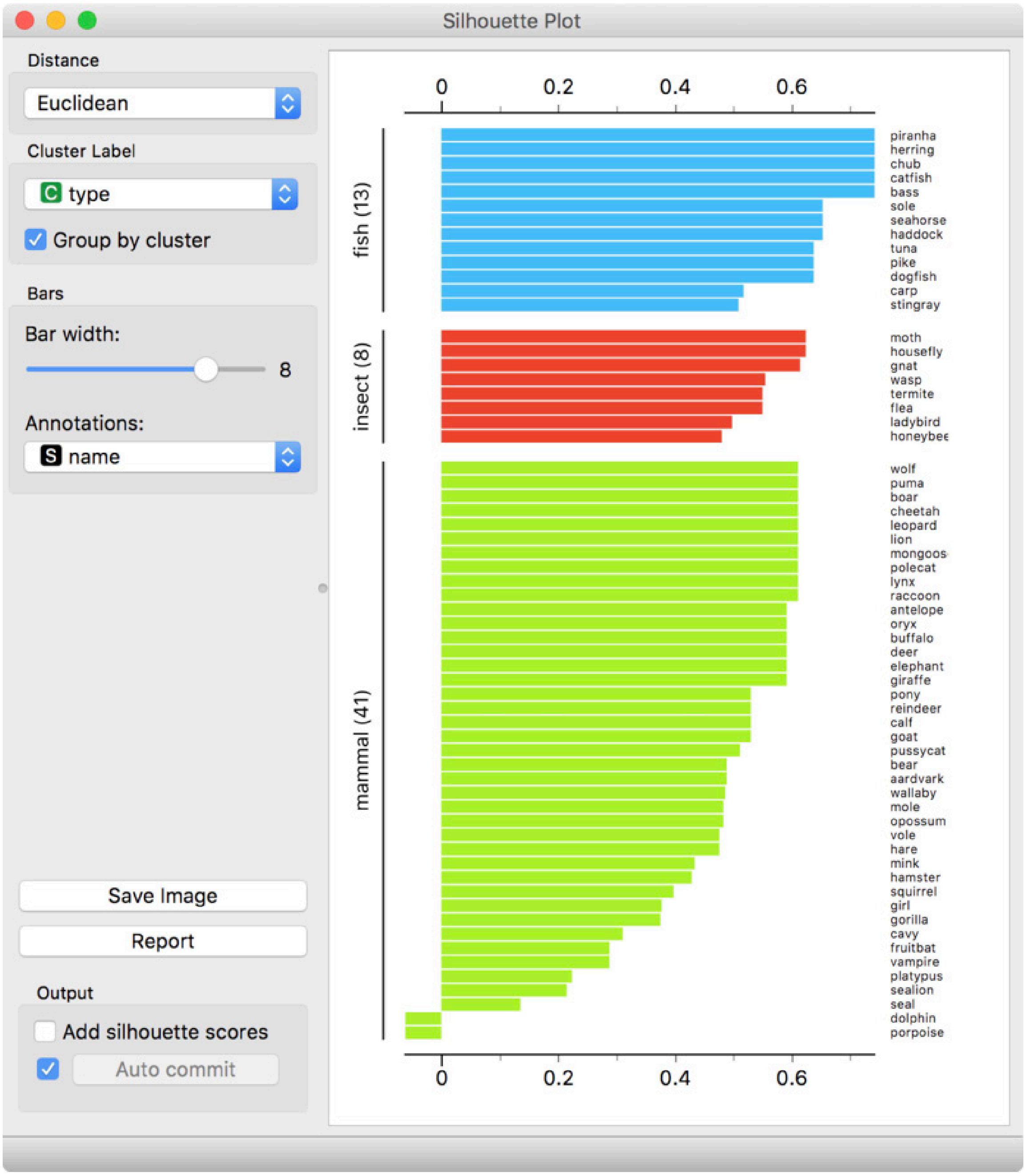

Figure 3.2 An example Silhouette plot.

3.6 Results

The clustering approach was tested on two different types of output data utilizing the k-means algorithm:

1. The direct damage state data output by RiskScape
2. The processed recovery time data

The damage data comprised of 1000 different realisations of a Wellington Fault earthquake scenario. The data is over-parameterised with parameters representing the damage state of over 2000 different components in the electricity network. Many of these parameters are invariant, however, across the different representations and for the purposes of clustering can be effectively removed. The remaining parameters were normalised to values 0–1 before the clustering was carried out.

The recovery time data comprised of 1000 cases but with far fewer parameters, each one representing the recovery time for each of the substations in the electricity network being modelled. A program of repeated silhouette tests resulted in the adoption of 2 clusters for the damage data and 80 clusters for the recovery time data based on the scores from the tests.

Clustering on the damage data outputs from RiskScape resulted in poorly resolved clusters and a silhouette score fractionally above 0.

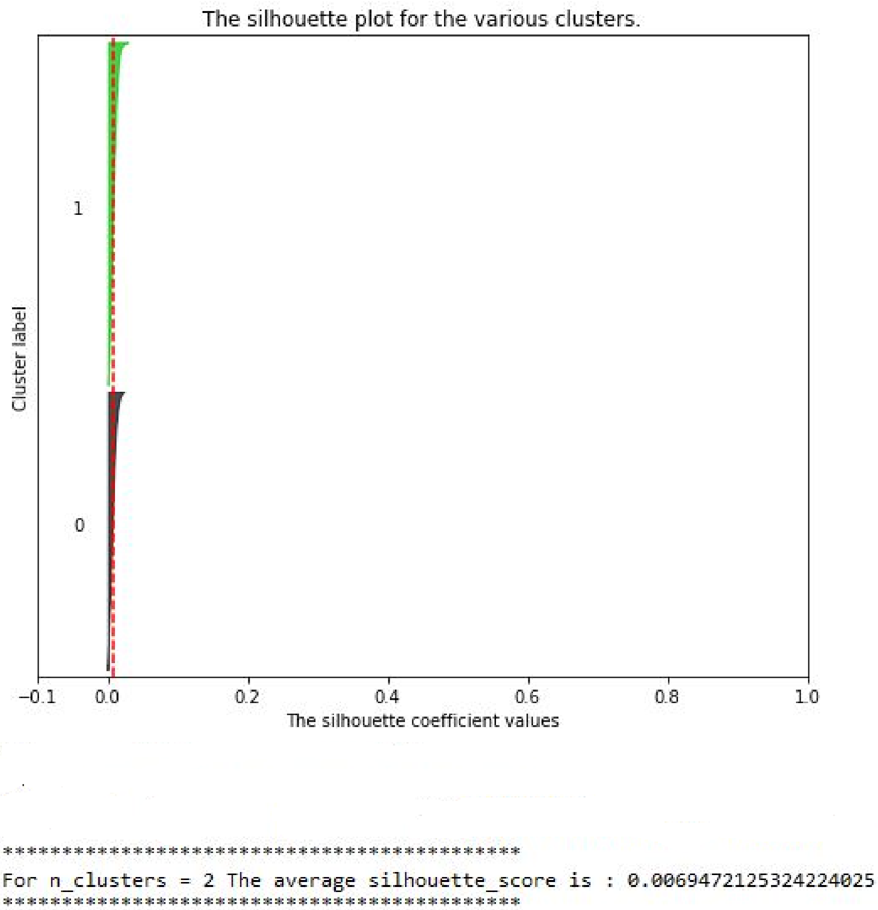

Figure 3.3 Silhouette plot for the damage data cluster with k = 2.

Cases on the fringe of the clusters are not robustly classified so the process was run repeatedly, and cases were apportioned to clusters based on a majority vote. The two vectors

representing the two cluster centres were very similar but there were 25 differences, 4 in the states of the substations and 21 in the other components. Timestamped outage maps based on the two centroid vectors were used as a basis of the information passed to the MERIT model. Cluster 0 has 409 members and cluster 1 has 591 members.

The recovery time data was better suited to clustering and consistently yielded better silhouette scores at all cluster numbers (k in the k-means). A general trend in the results was for the score to increase with increasing numbers of clusters. A plot of the scores indicated a small "elbow" at approximately k=80 with another potential optimal point at k=120. The smaller of these two values were taken as a number of clusters for the purposes of this study. Again, timestamped outage maps based on the 80 centroid vectors were passed to the MERIT model for economic modelling.

Table 3.1 The Silhouette scores for different numbers of clusters for the recovery time data.

Number of clusters	Silhouette Scores
2	0.228
10	0.249
20	0.35
50	0.538
70	0.598
80	0.606
100	0.638
200	0.682
300	0.738
431	0.749

Table 3.1 lists the Silhouette scores for different numbers of clusters. It, plus Figure 3.4, illustrate the diminishing returns in the quality of clustering with increasing numbers of clusters.

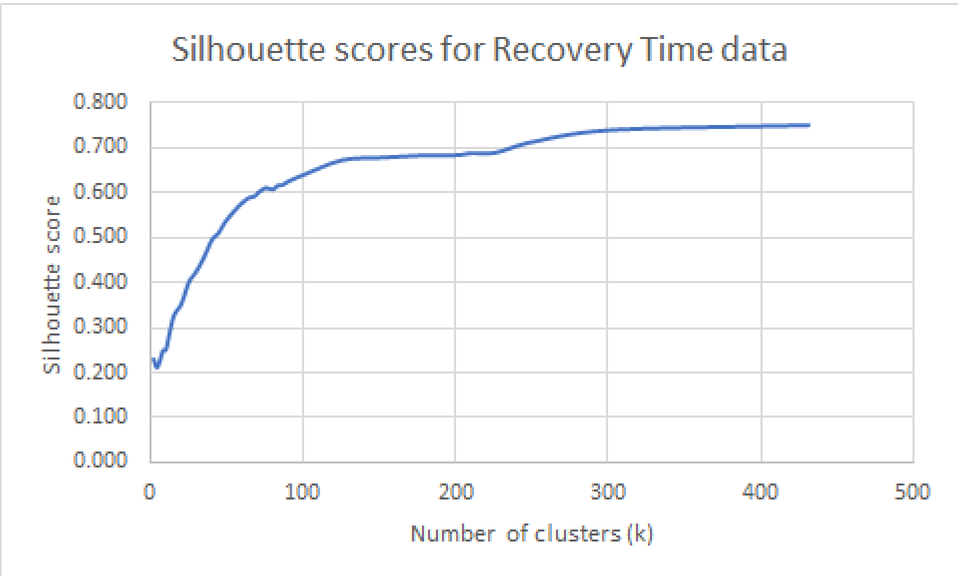

Figure 3.4 Silhouette scores per number of clusters for the recovery time data. Changes in gradient indicate possible elbow points at roughly 80 and 120 clusters.

Table 3.2 The membership numbers for each of the 80 clusters in the recovery time data.

Cluster number	Members	Cluster number	Members	Cluster number	Members	Cluster number	Members
0	17	**20**	3	**40**	11	**60**	1
1	32	**21**	11	**41**	11	**61**	5
2	72	**22**	4	**42**	9	**62**	6
3	22	**23**	12	**43**	15	**63**	9
4	15	**24**	8	**44**	9	**64**	5
5	10	**25**	10	**45**	7	**65**	2
6	71	**26**	12	**46**	2	**66**	1
7	46	**27**	14	**47**	16	**67**	5
8	16	**28**	19	**48**	8	**68**	7
9	8	**29**	18	**49**	4	**69**	4
10	14	**30**	12	**50**	2	**70**	12
11	32	**31**	16	**51**	7	**71**	4
12	27	**32**	4	**52**	8	**72**	2
13	3	**33**	12	**53**	9	**73**	1
14	25	**34**	10	**54**	7	**74**	31
15	14	**35**	4	**55**	8	**75**	4
16	13	**36**	9	**56**	2	**76**	2
17	16	**37**	20	**57**	2	**77**	1
18	24	**38**	17	**58**	5	**78**	2
19	35	**39**	15	**59**	8	**79**	4

Table 3.2 shows that the most popular cluster is Cluster #2 with 72 members, 4 clusters only have 1 member.

4.0 AUTOMATED OUTAGE MAP DEVELOPMENT

4.1 Restoration Assumptions

To successfully model and understand the recovery process with the inclusion of road outages for each electricity substation zone, certain assumptions were made to establish a common understanding of restoration times. Without these base assumptions, the analysis of the restoration times could become overly complicated, to the point where the realistic assessment would become impossible. These assumptions are that:

- The majority of the expected damage would be caused by the initial fault rupture and earthquake. Significant aftershocks would potentially cause further damage and therefore potentially lengthen restoration times.
- The scope of this test case is to model only the transmission network from generation units to the substation level. For the recovery of assets in the distribution network that, a predefined restoration time has been assigned after recommendations from the experts.
- The majority of necessary skilled resource and associated equipment would also be locally available, and there would be enough repair staff for work to proceed in multiple locations at a time.
- The repair times and strategies for various types of cables like Paper Insulated Aluminum Sheathed (PIAS) and cross-linked polyethylene (XLPE) may be different because some could be solid fluid-filled therefore harder to repair (Giovinazzi et al., 2017). If the number of damaged cables exceeds a predefined value, then the repair work on those cables could be abandoned and the existing cables could be replaced with emergency overhead lines for the continuity of electricity services to the customers.
- There are different priority substation zones in which the emergency overhead lines are replaced first. These priority zones are based on the significance of buildings depending on the services they provide like hospitals, fire brigade offices, police stations and other emergency management organisations.
- The road outage times would be computed based on the assumed number of days between different road zones during the response and recovery stages after an event.
- Every component of the electricity network would be mapped to a predefined road zone to understand the road access time to reach the damaged site before starting the repair work.

These assumptions are related specifically to the Wellington region for this scenario, but the automated outage map generation framework has the capability to model any city with new assumptions as well. Based on the assumptions mentioned above, we developed an integrated methodology to link electricity and road networks and then used the approach to create outage maps (key input to the MERIT model).

4.2 Integrated Approach for Linking Electricity and Road Networks

As described in the previous section, the simulation framework in this study is developed to compute the recovery times of electricity services through an understanding of their dependency on the road network. As shown in Figure 4.1, both the models have specific inputs and outputs. The road network model is integrated into the electricity network model as a subroutine so that electricity model uses the outputs of the road access times between various road zones to generate some realistic restoration times.

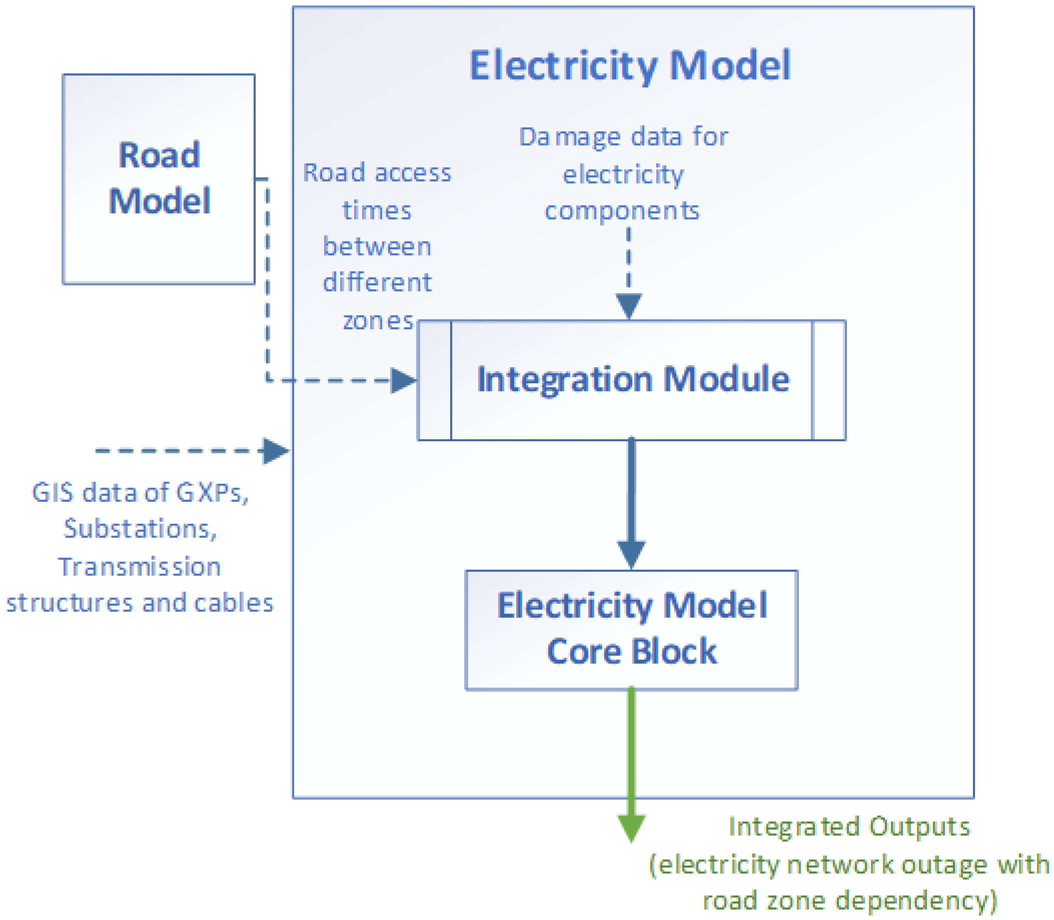

Figure 4.1 An integrated methodology for modelling electricity and road networks.

Further explanation about this integrated methodology is presented through a step by step process. This process has been developed on the restoration assumptions defined earlier in this report. The 4-step integration methodology is described below:

1. During the network setup, the GIS data for modelled electricity components of GXPs, substations, cables and transmission structures is loaded in the simulation framework. Apart from that, a road matrix is also included that shows the number of days required to restore road access from one road zone to another in response and recovery phases following a disaster. The asset data and input road matrix are included only once in the framework.
2. A damage map of 1000 different realizations generated by Riskscape for all the preloaded electricity network components is then loaded in the framework.
3. Within the integration module as shown in Figure 4.1, the simulation framework first combines road zones with substation zones and then uses the road outage times to determine the amount of time needed for damaged electricity components to recover. Different parts of the damaged components e.g. electricity cables can be spread across different road zones. Therefore, it is important to first consider the time needed to reach the site where each of the damaged cable segments are located.
4. The damage to repair calculation times for each component's different damage states is integrated within the electricity model's core block. These calculations consider the restoration assumptions mentioned in the previous section. The calculation setting along with the road access times is then applied to all the 1000 damage realizations of the components to generate 1000 sets of restoration times for each of the substation zone.

After applying the above-mentioned methodology, we developed time stamped outage maps for all the different sets of results, the 2 cluster centroids based on the raw damage data, the 80 cluster results and the results from 1000 Monte Carlo runs. Figure 4.2 shows some examples of these maps with and without the inclusion of road dependency.

The outage maps reveal a substantial increase in electricity outage time due to the road network dependency. The comparison of both the maps show that if there is no dependency on the road network then the electricity zones can recover quickly as compared to when there is road network dependency.

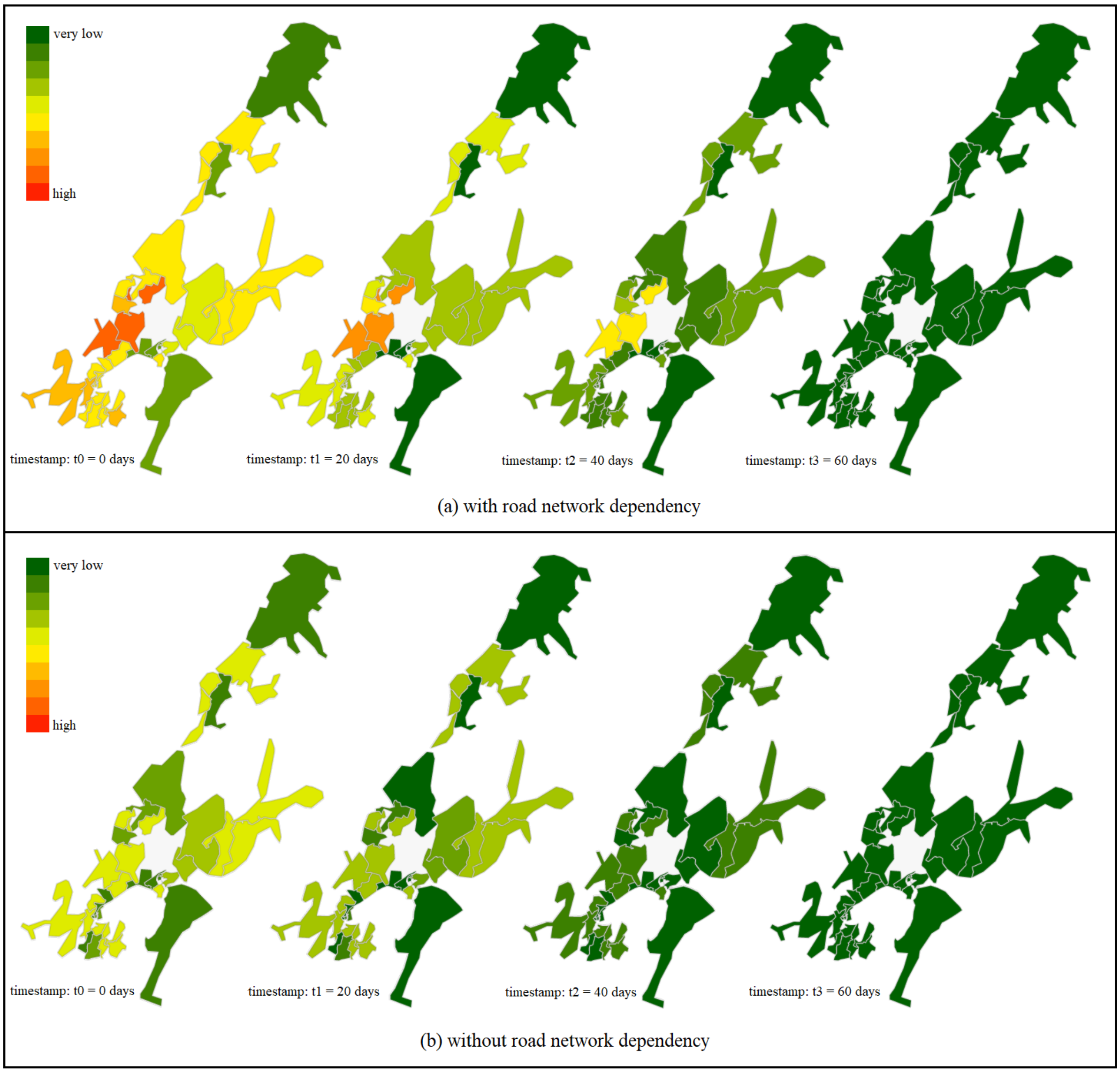

Figure 4.2 Electricity network outage maps on different timestamps (a) with and (b) without road network dependency.

The simulation framework is developed in a way that the electricity and road network of any other city can also be modelled in the same manner if the damage information of the components is provided in the acceptable input format. The outage times can be seen at different timestamps for any day by moving a timeline bar to understand the recovery process and the amount of time needed for the region to recover completely as shown in Figure 4.3. The resultant outage maps not only show the recovery time of a substation zone, but the simulation framework also has the capability to show the detailed damage information of all

the components that provide electricity to a substation from a GXP. This detailed analysis could be useful for identification of vulnerabilities within the electricity network. Moreover, the relevant stakeholders who would use this framework could customize the restoration assumptions and repair strategies to compare different recovery options and their corresponding outage maps.

Figure 4.3 An example output showing comparison of different outage maps which can be seen in different timestamps.

5.0 CONCLUSIONS BASED ON THE CLUSTERING WORK

In theory, clustering can be a useful means of reducing the time/effort overhead in reviewing time stamped outage maps of the sort needed by MERIT to undertake a wider economic analysis. However, in reality, the usefulness of this approach is greatly influenced by the quality of the input data, whether it be damage state data or processed recovery time data, in the cases described here.

This work has revealed both extremes in that the damage state data resulted in poorly resolved clustering but only two clusters and the recovery time data resulted in better resolution for the clusters, but still a reasonably large number of clusters (80) that would theoretically, incur a substantial time overhead for experts to review.

It is maintained, however, that conceptually both of the clustered approaches offer substantial advantages over the raw Monte Carlo results where experts cannot be expected to effectively process a 1000 sets of results.

6.0 ACKNOWLEDGEMENTS

Dr Vinod Sadashiva is thanked for his review and comments.

7.0 REFERENCES

Cousins WJ, Spence R, So E. 2008. Estimated casualties in New Zealand earthquakes. In. *Earthquake engineering in Australia conference 2008: held by the Australian Earthquake Engineering Society, Ballarat, Victoria, Australia, 21–23 November 2008: conference papers*. Ballarat (AU). Ballarat, Vic.: Australian Earthquake Engineering Society. Paper 19 (7 p.).

Driver HE, Kroeber AL. 1932. Quantitative expression of cultural relationships. In: Putnam FW, Kroeber AL, Lowie RH, editors. *University of California publications American archaeology and ethnology*. Berkeley (CA): University Press. Vol. 31, p. 211–256.

Dueñas-Osorio L, Vemuru SM. 2009. Cascading failures in complex infrastructure systems. *Structural Safety*. 31(2):157–167. doi:10.1016/j.strusafe.2008.06.007.

Erdener BC, Pambour KA, Lavin RB, Dengiz B. 2014. An integrated simulation model for analysing electricity and gas systems. *International Journal of Electrical Power & Energy Systems*. 61:410–420. doi:10.1016/j.ijepes.2014.03.052.

Federal Emergency Management Agency. 2015. Hazus–MH 2.1: technical manual. Washington (DC): Federal Emergency Management Agency; [accessed 2019 Nov 6]. https://www.fema.gov/media-library-data/20130726-1820-25045-6286/hzmh2_1_eq_tm.pdf.

Galili T. 2019. Hierarchical cluster analysis on famous data sets - enhanced with the *dendextend* package. [place unknown]: Comprehensive R Archive Network; [accessed 2019 Nov 6]. https://cran.r-project.org/web/packages/dendextend/vignettes/Cluster_Analysis.html#background-3.

Giovinazzi S, Pollino M, Kongar I, Rossetto T, Caiaffa E, Di Pietro A, Porta L, Rosato V, Tofani A. 2017. Towards a decision support tool for assessing, managing and mitigating seismic risk of electric power networks. In: Gervasi O, Murgante B, Misra S, Norruso G, Torre CM, Rocha AMAC, Taniar D, Apduhan BO, Stankova E, Cuzzocrea A, editors. *Computational Science and Its Applications - ICCSA 2017: 17th International Conference, Trieste, Italy, July 3–6, 2017: Proceedings, Part III*. p. 399–414.

Guidotti R, Chmielewski H, Unnikrishnan V, Gardoni P, McAllister T, van de Lindt J. 2016. Modeling the resilience of critical infrastructure: the role of network dependencies. *Sustainable and Resilient Infrastructure*. 1(3–4):153–168. doi:10.1080/23789689.2016.1254999.

Gopalakrishnan K, Peeta S. 2010. Sustainable and resilient critical infrastructure systems: simulation, modeling, and intelligent engineering. Berlin (DE): Springer. 265 p.

King AB, Bell RG, Programme managers. 2009. RiskScape Project: 2004–2008. Wellington (NZ): National Institute of Water and Atmospheric Research. 172 p. Science Report 2009/75.

Lin S-L, Nayyerloo M, Zhang ZH. 2016. Seismic performance of buried electricity cables during the Canterbury earthquake sequence. In: *Australian Earthquake Engineering Society 2016 Conference;* 2016 Nov 25–27; Melbourne, Australia. Melbourne (AU): Australian Earthquake Engineering Society. 10 p.

MacQueen J. 1967. Some methods for classification and analysis of multivariate observations. In. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability: held at the Statistical Laboratory, University of California, June 21–July 18, 1965 and December 27, 1965–January 7, 1966 Vol 1: Statistics*. Berkeley (CA): University of California Press. p. 281–297. [accessed 2019 Nov 6]. https://projecteuclid.org/euclid.bsmsp/1200512992

Mowll R. 2012. Lifeline utilities restoration times for metropolitan Wellington following a Wellington Fault earthquake: Report to the Wellington CDEM Group Joint Committee from the Wellington Lifelines Group. [Wellington (NZ)]: [Ministry of Civil Defence & Emergency Management]. 63 p.

Omram MGH, Engelbrecht AP, Salman A. 2007. An overview of clustering methods. *Intelligent Data Analysis*. 11(6):583–605.

Rinaldi SM, Peerenboom J, Kelly TK. 2002. Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Control Systems*. 21(6):11–25. doi:10.1109/37.969131.

Rosser BJ, Dellow GD. 2017. Assessment of liquefaction risk in the Hawke's Bay. Volume 1: the liquefaction hazard model. Lower Hutt (NZ): GNS Science. 114 p. Consultancy Report 2015/186. Prepared for Hawke's Bay Regional Council.

Steinhaus H. 1957. Sur la division des corps matériels en parties [On the division of material bodies into parts]. *Bulletin de l'Académie Polonaise des Sciences Classe Troisième*. 4(12):801–804. French. https://zbmath.org/?format=complete&q=an:0079.16403

Van Houtte C. 2017. Performance of response spectral models against New Zealand data. *Bulletin of the New Zealand Society for Earthquake Engineering*. 50(1):21–38.

[WE] Wellington Electricity. 2019. Wellington Electricity – 10 year Asset Management Plan, 1 April 2019–31 March 2029. Lower Hutt (NZ): Wellington Electricity Lines Limited; [accessed 2019 Nov 6]. https://www.welectricity.co.nz/disclosures/asset-management-plan/document/177

Xie L, Tang J, Xie Q. 2012. Seismic fragility assessment of transmission towers via performance-based analysis. In. *15th World Conference on Earthquake Engineering*. 2012 Sep 24–28; Lisbon, Portugal. Lisbon (PT): World Conference on Earthquake Engineering. paper 1617 (10 p.).

Zubin J. 1938. A technique for measuring like-mindedness. *Journal of Abnormal and Social Psychology*. 33(4):508–516. doi:10.1037/h0055441.

www.gns.cri.nz

Principal Location

1 Fairway Drive
Avalon
PO Box 30368
Lower Hutt
New Zealand
T +64-4-570 1444
F +64-4-570 4600

Other Locations

Dunedin Research Centre
764 Cumberland Street
Private Bag 1930
Dunedin
New Zealand
T +64-3-477 4050
F +64-3-477 5232

Wairakei Research Centre
114 Karetoto Road
Wairakei
Private Bag 2000, Taupo
New Zealand
T +64-7-374 8211
F +64-7-374 8199

National Isotope Centre
30 Gracefield Road
PO Box 31312
Lower Hutt
New Zealand
T +64-4-570 1444
F +64-4-570 4657