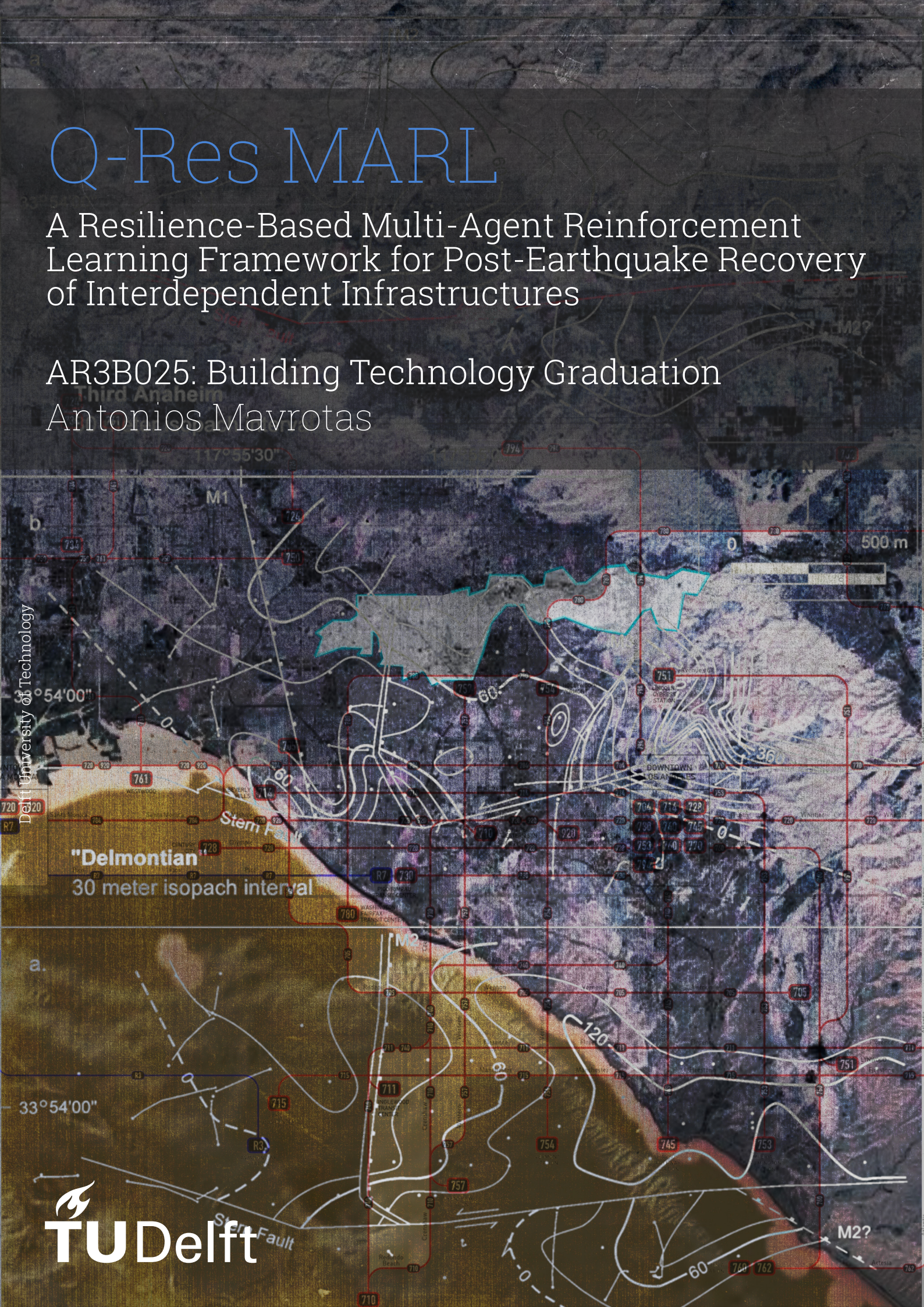
Q-Res MARL

A Resilience-Based Multi-Agent Reinforcement Learning Framework for Post-Earthquake Recovery of Interdependent Infrastructures

AR3B025: Building Technology Graduation

Antonios Mavrotas



Delft University of Technology

TU Delft

Q-Res MARL

A Resilience-Based Multi-Agent Reinforcement Learning Framework for Post-Earthquake Recovery of Interdependent Infrastructures

by

Antonios Mavrotas

Instructors:	Charalampos Andriotis and Simona Bianchi
Project Duration:	Nov, 2024 - June, 2025
Faculty:	Faculty of Architecture and the Built Environment, Delft

TU Delft

Acknowledgements

This work was difficult to compile and produce and I would like to dedicate this section to several people that helped me push through. Firstly, I would like to thank my supportive family and my wonderful girlfriend, Jemima who spent several holidays looking at me look at my computer screen for hours. I would also like to extend my gratitude to my two mentors: Charalampos Andriotis and Simona Bianchi, who shared their incredible expertise with me and helped me keep my research focus on a topic that can easily venture into many directions. Furthermore, I want to thank Prateek Bhustali for meeting with me regularly and developing the *DRL* framework for his PhD, which is used in the thesis.

Abstract

This thesis focuses on using MARL as a decision tool for post-earthquake repair scheduling of interdependent infrastructure. MARL is a multi-agent ML paradigm which combines traditional ML research and game-theoretical approaches. Given the relative increase in natural disaster frequency and the lack of available post-disaster tools such tools are crucial in increasing the climate resilience of cities. Given the stochastic nature of earthquake events and subsequent losses, MARL can be helpful in navigating this uncertainty and finding preferable joint policies.The methodology involves multi-scenario-based seismic hazard assessment, stochastic fragility modelling and prediction of several direct and indirect losses to aggregate them into a holistic community resilience metric. This is then used to compute the instantaneous and cumulative recovery resilience loss values. The tested approach uses two custom built test-beds of 4 and 30 components, and MARL is compared against baseline solvers, including random and importance-based policies. Value Decomposition Network with Parameter Sharing ($VDN-PS$), Q-Learning with Mixer Network and Parameter Sharing ($QMIX-PS$), Deep Centralised Multi-Agent Actor Critic ($DCMAC$) are the algorithms tested. VDN and $QMIX$ are shown to perform similarly to each other and sub-optimally relative to $DCMAC$. $DCMAC$ is shown to match importance-based policies when considering full recovery, but convincingly outperforms all other DRL methods and importance-based policies when considering partial recovery. This shows that $DCMAC$ and DRL more generally is effective at swift early recovery by prioritising components that contribute most to community functionality.

Contents

Nomenclature

Abbreviations

$GMPE$	Ground Motion Prediction Equation
PGA	Peak Ground Acceleration [m/s^2]
PGD	Peak Ground Displacement [m, cm, inches]
PGV	Peak Ground Velocity [m/s]
SA	Spectral Acceleration [m/z^2]
$FEMA$	Federal Emergency Management Agency
FHA	Federal Highway Administration
NSI	National Structures Inventory
NBI	National Bridge Inventory
RL	Reinforcement Learning
$MARL$	Multi-Agent Reinforcement Learning
$QMIX - PS$	Q-Mixer with Parameter Sharing
$VDN - PS$	Value Decomposition Network with Parameter Sharing
$DCMAC$	Deep Centralised Multi-Agent Actor Critic
$IMPB$	Importance Based Scheduling

Symbols

Symbol	Definition
$\mathcal{M}$	The set of RL Agents
$\mathcal{S}$	The joint state space of RL agents
$\mathcal{A}$	The joint action space of RL agents
$\mathcal{T}$	The transition model an RL environment
$\mathcal{R}$	The common reward of RL agents
$\mathcal{O}$	The joint observation space of RL agents
t_H	The finite time horizon of an RL environment
Q	Functionality of an infrastructure network
Res	Resilience
$Loss$	Resilience Loss
T_{rec}	Time to full recovery after a disruption
t_d	Time of disruption
t_m	Time of mobilisation
t_r	Time when repair begins
$T_{N\%}$	Time to recovery $N\%$ of impact losses

1

Introduction

1.1. Natural Disaster Preparedness

Natural disasters such as floods, earthquakes and tornadoes cause immediate injury, loss of life, stress on infrastructure, all of which have cascading effects, which might persist long after the disastrous event. Cities can be thought to act as complex systems of different demographic groups and infrastructure networks, where initially observed structural damage and injury is subsequently seen to extend and impact various aspects of a community's functionality. The manifestation of such cascading effects is described as the *interdependency* between different infrastructure networks, but also people, [44]. For instance, the traffic patterns of a community after a disaster can be interdependent to the amount of injury caused as rescue missions increase in frequency. Such cascading effects and relationships are often clear to see retrospectively but are seldom strategised for pre-emptively. This is because the analysis of such relationships does not take priority when time is of the essence and rescue and recovery are paramount. Therefore, the nature of interdependencies calls for a heightened level of preparedness to natural disasters, as decision makers are faced with the non-trivial and complex nature of cascading damaging effects.

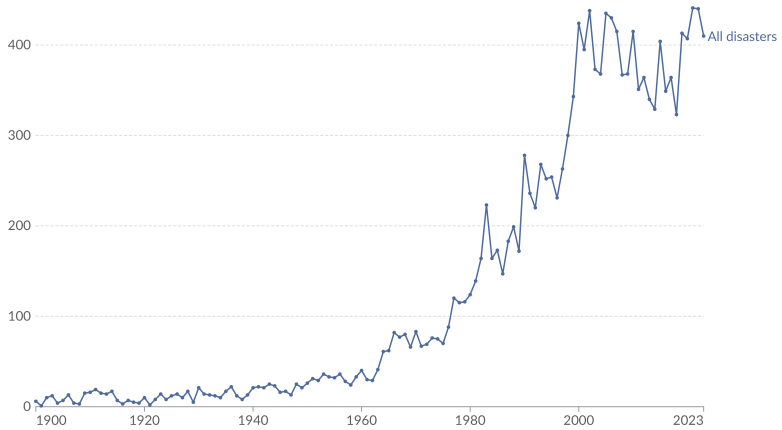


Figure 1.1: Global trend showing the number of recorded natural disaster events over recent years, Our World in Data, 2024, Ritchie, Rosado, and Roser [73]

Additionally, the apparent frequency and severity of certain natural disaster types, or hazards, is seen to increase in many places around the world, [91, 85, 19]. While this is not true for all hazards, this increase in frequency can be at least partially attributed to climate change. The disaster-specific mechanisms that cause this are either difficult or sometimes impossible to model. In effect, this increases the exposure of communities to natural disasters and thus makes the need for recovery more frequent. The combination of increasing disaster frequency and complex nature of interdependent infrastructure increases the hazard risk of many communities around the world. There is also a correlation with many underdeveloped areas and areas of extreme climate events which increases this pressure, [5, 29, 76].

(a) Landslide disaster in the town of Blatten, Switzerland. Landslides such as this can exhibit the combined effects of exceeding amounts of fast-moving debris and flooding. Foulkes [38]

(b) Floods causing exceeding transportation, building and economic losses in Spain, showing the results of exteme flow carrying vehicles to one concentrated road segment. Sofia Ferreira Santos [88]

(c) First-hand photograph taken in the El Haouz region of Morocco, approximately 5 months after the 2023 Marrakesh-Safi 6.9M earthquake. Uniquely, residences were interdependent to each other as the lack of centralised planning policies allowed different families to expand on existing residences organically, making the financing of repairs challenging.

Figure 1.2: Natural Disasters can come in many different forms and their impact can vary between different communities.

Damage after natural disasters is often formulated in terms of losses. Losses can be seen as either direct, indirect or human losses, [36, 28, 16]. Direct losses are related to the induced structural stress and response of infrastructure, and are manifested as repair times, costs, and reduced access to infrastructure. Indirect losses are related to the indirect damage caused by a disaster and can be manifested as income losses, relocation costs etc. Human losses are often seen during or directly after a disastrous event and are the injury or death of people. More generally, the expected physical damage to a community's infrastructure after a disaster is described as the fragility of that community's infrastructure. Likewise,

the expected indirect losses of a community are described as the vulnerability of a community to a disaster. Human losses are different from direct and indirect losses as they can depend a lot more on first-response efforts and general behaviour of healthcare professionals, more so than the infrastructure itself. It is important to note that losses are usually computed probabilistically; thus, they are usually referred to as expected. This is done because the scale and complex nature of cities as described above make the deterministic computation of losses exceedingly difficult. Specifics on vulnerability and fragility are discussed in the literature review. Considering the above, the risk of a community to hazards can be thought of as a combination of exposure, vulnerability, fragility and ability to recover.

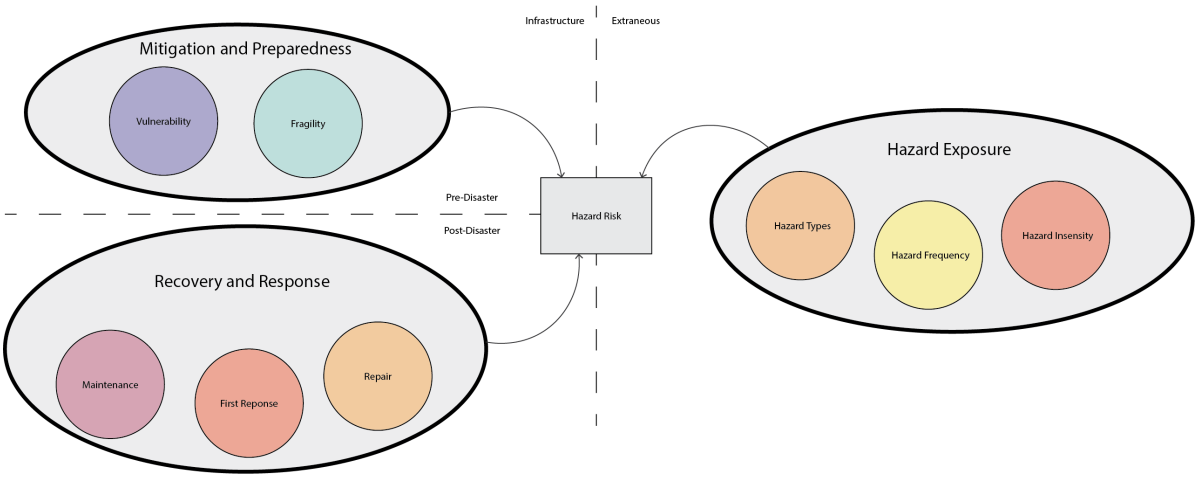


Figure 1.3: A general framework to conceptualise hazard risk as a function of hazard exposure, recovery and response and mitigation and preparedness

The significance of the problem at hand is the heightened risk posed to communities around the world due to natural disasters. Risk can be mitigated by reducing exposure, fragility or vulnerability to disasters, or improving recovery ability by means of strategising for optimal repair strategies. Exposure is evidently very difficult to reduce as it is inherently dependent on the location of a specific city. Fragility of infrastructure can be reduced by increasing its capacity to loads caused by disasters. Vulnerability can be reduced by developing a community such that its socioeconomic indicators increase and expected indirect or human losses are reduced. This might be in the form of emergency shelter construction, healthcare system improvement, unemployment reduction etc. Risk mitigation via fragility and vulnerability reduction focuses on preventative measures and is studied more often than recovery ability improvement, [95]. Recovery is purely concerned with the time during and after a disaster. Thus, improving recovery ability is principally concerned with the decision-making process of post-disaster repair. The specifics of this thesis' objectives are outlined in the following section; however, it is important to note that risk mitigation should always be a continuing and holistic effort to reduce fragility, vulnerability and improve recovery ability. This assumes a continuing collaboration between international, national and municipal decision makers, engineers and designers, contractors and healthcare and rescue professionals. In doing so, the aim should be to mitigate risk via the most cost- and time-effective manner available, aiming to maintain a community's functionality as close to pre-disaster levels as possible.

(a) People requiring assistance after disasters.

(b) Local disaster strategies vs. national alignment.

Figure 1.4: Disaster impact and preparedness strategy indicators, Ritchie, Rosado, and Roser [73]

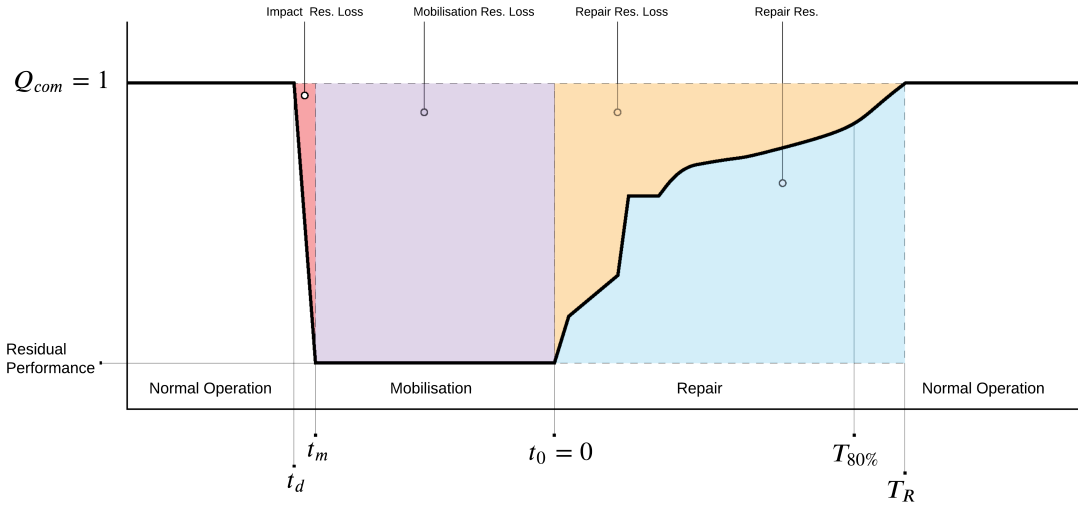


Figure 1.5: A community functionality-time curve for a community after a severe disruption, such as an earthquake

1.2. Engineering Reliability and Resilience

Before introducing the specific aims of the thesis, it is important to define resilience, which is the principal concept governing the measurement of successful natural disaster risk mitigation approaches. While specifics on mathematical definitions of resilience are given in the literature review, the core idea of resilience is centred around minimising the effect of expected disruptions on infrastructure networks and communities, thus increasing the reliability of infrastructure, [71, 14, 33, 41], . More specifically, resilience and resilience loss are usually computed as areas under or over a functionality-time curve of a specific system or community. That is, functionality is a dimension-less quantity which is an aggregate of sub-system functionalities and usually ranges between 0 (*completely dysfunctional*) to 1 (*completely functional*). For instance, in the event of a sever disruption, Fig 1.5 shows such a curve. A disruption happens at time t_d which causes a decrease in functionality, funds and expertise are mobilised at time $t_m \longrightarrow t_0$ while functionality is assumed to stay constant and an appropriate repair strategy is carried out to return to a desired fraction of pre-disruption functionality. The impact time interval $t_m - t_d$ is often not explicitly considered when modelling disasters; the decision to explicitly model it is predominantly dependant on whether decision-making during its duration can affect the resultant losses. For instance, given the short duration of earthquake action, the decision-making process during an earthquake is largely placed on the individual, and there is little centralised decision makers can do to affect losses in the minutes or hours during which ground motion occurs. Conversely, a wildfire can last weeks or months, leaving crucial time during which decision makers can gather and centrally strategise and allocate tasks accordingly, before the disaster event terminates. Of course, one can define the behaviour

of functionality before, during and after a disruption in non-constant terms. For instance, within agriculture, the economic performance of a wheat seed index before economic disruptions is likely to cycle in a manner that follows growing and harvesting seasons, or other extraneous factors. The principal approaches for measuring resilience are similar.

The concept of resilience is thus not unique to engineering structures and can be used for gauging the performance after disruptions in economic markets, food security, water scarcity etc. However, in the context of this thesis, resilience is applied as it relates to civil engineering systems and infrastructure. In doing so, the aim is to use the functionality of infrastructure after disruptions to predict direct and indirect losses and thus predict resilience and resilience loss. It is important to note that this concept is not only applicable to natural disasters, but rather any disruption that causes a decrease in functionality. This thesis specifically considers disasters as the disruption type, however related work in infrastructure management under deterioration disruptions is also of interest, [14]. An example of a non-severe disruption of infrastructure is the effect of fatigue, which is the deterioration of structural systems due to continuous cyclical loading, [72]. More generally, the study of measuring the resilience of infrastructure systems under deterioration and strategising inspection and repair actions is usually referred to as *maintenance planning*, while the study of measuring resilience and strategising interventions after severe disruptions or disasters is referred to as *repair scheduling*. The core difference is that disruptions under maintenance planning are assumed to be ongoing, while disruptions under repair scheduling are unique, severe and discrete events with high probabilities of complete or partial failure. Both types of disruptions can coexist and affect infrastructure concurrently.

Conclusively, the aim of both approaches is to increase the reliability of engineering systems by considering minimising of any of the associated resilience loss metrics, e.g impact resilience losses or mobilisation resilience losses. For instance, one could aim to increase the residual performance of a system under disruptions by increasing its capacity to the expected loads. Conversely, one could strategise for desired post-disruption policies that aim to rebound the functionality of the system as quickly and as effectively as possible. Thus, reliability is not only related to the performance of the engineering system of question, but also to the decision-making process of designers to prioritise intervention and retrofit actions that might increase expected residual performance or inspection and repair actions that aim to quickly rebound performance after a disruption. Decision making science is then seen as being central to the practice of reliability engineering as engineers are often faced with making long-term decisions using fuzzy data, and having to advise multiple stakeholders on decisions with critical consequences, [21].

1.3. Decision Making under Uncertainty

The above sections lay out a basic introduction on the growing concern of infrastructure reliability as it relates to an increased natural disaster hazard risk. This thesis focuses on civil infrastructure and the associated decision-making process of increasing their reliability in face of increasing natural disaster frequency by looking at their resilience. Regardless of the severity of the disruption or the type of infrastructure, the decision-making process often involves many stakeholders and is accompanied with considerable uncertainty. A city's infrastructure usually depends on the decisions of municipal- and national-level officials, but also on lawyers, insurance companies, engineers, designers and trade professionals. They each have to take decisions individually, but also work collectively to find ways of maintaining their city's infrastructure access or improving its performance. Furthermore, decisions are often made using incomplete information on infrastructure performance but also on the expected behaviour of infrastructure under a given disruption. This is because the current state of different infrastructure components is often known to a partial degree and decisions are made based on estimates and expert knowledge. For instance, a city's catalogue of street lights might include information on their construction date, materials, required voltage and links to adjacent power links. However, even if this information is available, which it often is not, it is difficult to predict the current performance of all street lights with high certainty, and even more so to predict the performance of any one street light when subject to disruptions from high winds, floods or earthquakes. For example Fig 1.6 shows a collection of infrastructure components and an example distribution of how their expected repair times might vary probabilistically given a certain loading condition. For this reason, any decision-making framework

aimed at increasing infrastructure reliability should take into account the relative uncertainty associated with a city's infrastructure portfolio. This is especially true as the number of infrastructure components and the number of decision makers considered grows.

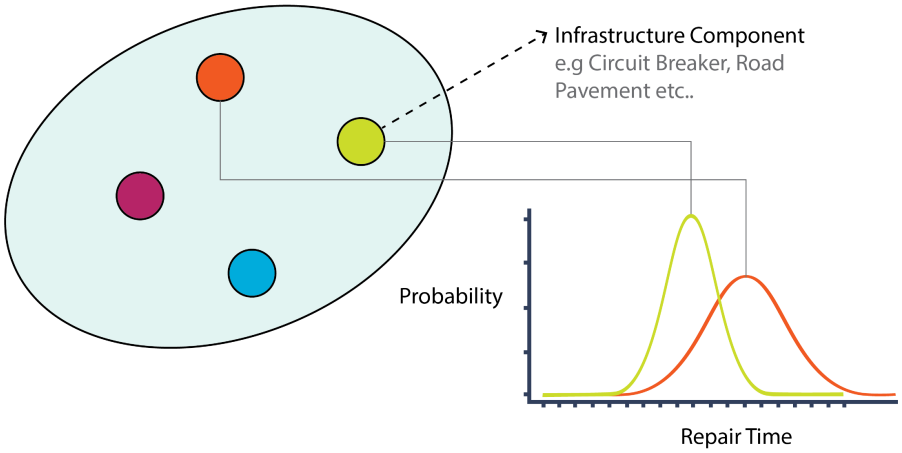


Figure 1.6: The deterioration properties of an infrastructure portfolio can often be modelled stochastically as their exact state is often hard to deduce

While highly bespoke and location-specific studies might include infrastructure portfolios with well-described and certain behaviour, the aim of this thesis is to develop decision-making tools that can generalise well across different cities and infrastructure networks. For this reason, this thesis focuses on decision-making tools with methods that deal well with relatively high levels of uncertainty and with many decision makers. Specific state-of-the-art approaches in decision making for infrastructure management under disruptions is presented in later sections, however this thesis principally considers the use of Multi-Agent Reinforcement Learning ($MARL$) as the test method for infrastructure management under uncertainty, [4, 11, 7]. $MARL$ is the multi-agent version of Single-Agent Reinforcement Learning ($SARL$) and draws upon research in Game Theory (GT) and Machine Learning (ML) to effectively model and solve multi-agent interactions in complex and uncertain decision-making scenarios, [18]. It is chosen as it has produced limited but promising results in the field of infrastructure management, thus demonstrating that its use can be favourable, but requires further research efforts to standardise its use.

GT is a field of knowledge which developed separately from ML and attempts to describe the cooperative or competitive interactions between many decision makers with shared or individual goals, [68, 63]. Traditionally, game theory research is focused on competitive economic and political games, with cooperation often being an unexpected, but favourable behaviour in many competitive games. Thus, the word game is used to describe the multi-agent interactions of decision making scenarios, specifically in infrastructure management under disruptions; it is not intended to take away from the severity of natural disasters or other disruptions. This thesis specifically considers cooperative games, as the various decision makers such as engineers and municipal officials are assumed to be willing to cooperate to increase infrastructure reliability and reduce risk. Machine Learning is a field of study in computer science, mathematics and biology which aims to approximate the functions of neurons in brains to process and learn from large amounts of data, with little explicit rule-based instructions. Principally, ML aims to make predictions on unseen data that are otherwise extremely difficult or sometimes impossible for humans to dissect with traditional data analysis. $MARL$ uses GT to effectively describe and abstract complex decision-making scenarios and ML to learn from the data associated with specific decisions; it then aims to choose optimal actions, given the associated observations after taking an action. Figure 1.7 shows a simplistic version of an $MARL$ framework for infrastructure management. A set of decision-making agents act on a set of infrastructure components, the state and outcome of their

decisions is collected and passed into an ML learning function, which aims to learn an optimal action to maximise the outcome, or reward, of all agents. Of course, this framework can vary in architecture and complexity, agents can act on one or multiple infrastructure components, they can each have their own learning function, their rewards can be individual or shared, and the learning functions can directly learn an optimal action or try to predict the quality of a given action.

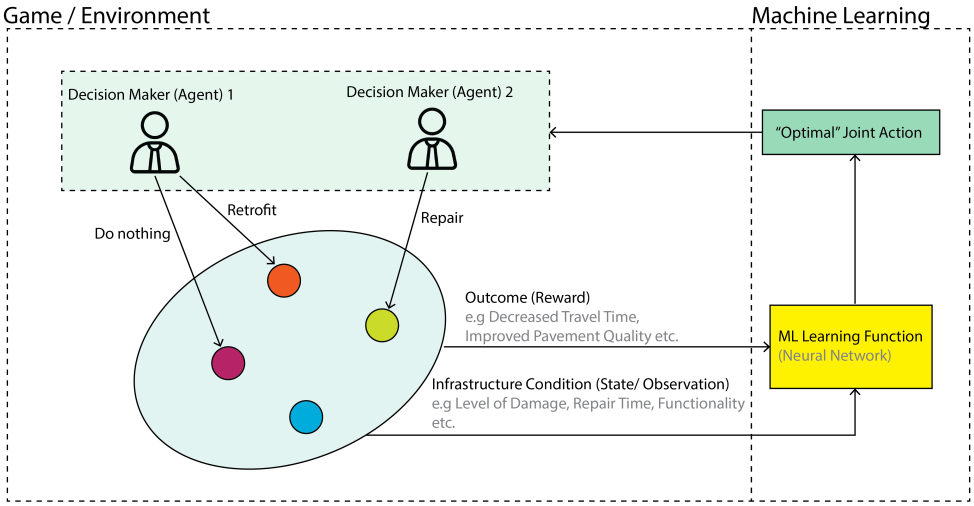


Figure 1.7: Multi-Agent interaction in an infrastructure management decision making scenario

Decision making frameworks for infrastructure management traditionally include importance-based methods, i.e repairing all hospitals before repairing residential buildings. This is described as heuristic, or rule-based decision making. Rule-based decision making can be highly complex, include many conditional rules and yield very good results. However, the use of MARL as a competitive strategy to rule-based decision making is not yet fully explored in the field of infrastructure management. This thesis tests the hypothesis that ML, and specifically $MARL$ is favourable to importance-based decision making and can thus help decision-makers make smarter decisions based on data derived from the computational simulation of a plethora of decision-making events. The ML learning function is an Artificial Neural Network (ANN/NN) which is the principal way most ML tools and frameworks operate, [74, 94, 77, 2]. $ANNs$ are interconnected networks of neurons that receive a set of input features and predict a single or a set of output labels on that data. The layers of neurons between input and output layers are called hidden layers. The way that hidden layers are connected to each other, the way that neurons *activate* and the type of neurons can all vary between neural networks to allow them to make predictions on specific types of data such as text-based prompts, graph-based network data, financial data, etc. Figure 1.8 shows a highly simplistic version of an NN with one hidden layer, this is described as a *fully connected* NN as all neurons of one layer are connected to all neurons of the following layer.

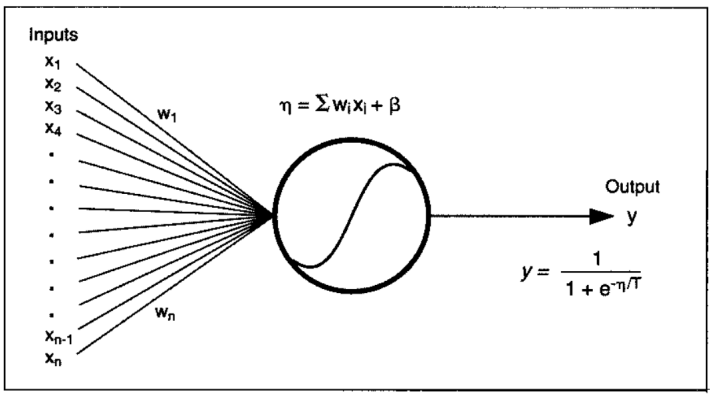


FIGURE 1. A simple sigmoidal output function

FIGURE 2. A simple three-layer network. The key to the effectiveness of the multilayer network is that the hidden units learn to rerepresent the input variables in a task-dependent way.

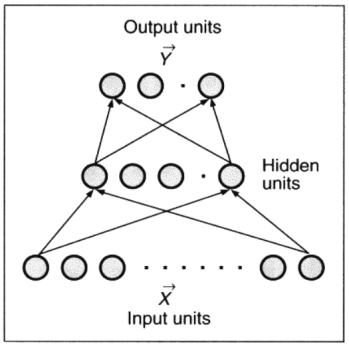


Figure 1.8: The two underlying ideas in $ANNs$ from one of the foundational papers by Rumelhart, Widrow, and Lehr [78]. *Figure 1* shows the behaviour a single neuron as having three key aspects: (a) a weighted sum of input vector $[x_1, x_2 \dots x_n]$ (b) a bias β which shifts the activation function in x and (c) a, usually, non-linear activation function such as the sigmoid function. Learning is the act of tuning all weights $\mathbf{w} = [w_1, w_2, \dots, w_n]^\top$ and biases $\mathbf{b} = [b_1, b_2, \dots, b_m]^\top$ in the network, by minimizing a loss function $\mathcal{L}(\hat{\mathbf{y}}, \mathbf{y})$ with respect to the model's predictions $\hat{\mathbf{y}}$ and the true labels $\mathbf{y}$, typically using gradient-based optimization. *Figure 2* shows the architecture of such a network, such that the connectivity, amount, layering and other *hyper-parameters* affect the overall prediction performance of the network.

In the case of this thesis, the NN receives some numerical description of the state of infrastructure components and makes a direct prediction on the optimal intervention action given that state, or on the quality of an action given that state. Learning is the act of tuning individual neuron weights such that the output intervention action is optimal. Optimality of decisions is measured via the reward function, which in the case of this thesis is done in terms of resilience, such that an action is most optimal if it yields the lowest resilience loss. NNs are not fully explained in this section, however some notable examples of highly complex NNs that make impressive predictions are tools such as OpenAI's ChatGPT or Google's Gemini, which effectively auto-correct user prompts. NNs such as those include billions or trillions of neurons, complex architectures and are able to make context-dependent predictions. In the context of this thesis, the number of neurons, or parameters used, is in the order of tens or hundreds.

In complex infrastructure management scenarios modelled as games, finding a mathematically optimal, or best solution is often infeasible due to the high degree of uncertainty and the vastness of the action and observation spaces. As a result, the performance of a given strategy cannot be measured against a known optimal solution, but rather must be evaluated in relation to other competitive strategies. This benchmarking approach makes objective performance measurement inherently comparative. Consequently, the goal shifts from identifying an optimal strategy to developing one that consistently outperforms a given baseline. Conversely, in scenarios where an optimal strategy is known, performance evaluation can be conducted through optimality-based assessment, where a strategy's effectiveness is measured directly against the best-known solution. This allows for an absolute, rather than relative, evaluation of performance. Considering the complexity of interdependent infrastructure networks and cities as a whole, the use of MARL can be intuitive as it is easy to see that human decision makers and rule-based frameworks can have limitations in concurrently processing and analysing the complex interactions of interdependencies, but also the uncertainty associated with infrastructure performance. However, it must be noted that the aim of the thesis is not to suggest a MARL-only methodology, but

rather use MARL as a decision-making tool to assist engineers and other officials. This is because many ML methods, including MARL, are often eventually understood as black-box functions where the output is impossible to deduce logically from a set of inputs. Given the critical nature of infrastructure management, decision-makers should always be aware of the consequences of their decisions and should not solely rely on such tools to make decisions, but rather use them to benchmark their current approaches. In this way, MARL can allow decision makers to think critically about what interventions to make and potentially offer valuable insights in situations where rule-based decision-making falls short.

1.4. Objectives and Research Questions

This thesis addresses the critical challenge of infrastructure management and reliability in the face of extreme weather events. The decision-making processes involved are characterized by high levels of uncertainty, creating an opportunity for innovative research approaches. This thesis proposes and evaluates the use of Multi-Agent Reinforcement Learning (MARL), benchmarking it against importance-based decision-making approaches to provide a comparative analysis.

While risk mitigation can address various factors, this study concentrates on the post-disaster recovery phase. Specifically, the thesis investigates the problem of post-earthquake repair scheduling of interdependent infrastructure networks. The focus on the recovery phase is motivated by a recognized gap in the existing literature, which has more extensively covered mitigation and preparedness. Earthquakes have been selected as the specific hazard due to the existence of mature and validated tools for modelling their geological mechanisms and the resulting structural response of various infrastructure categories. In contrast, modelling other hazards, such as floods or tsunamis, often involves more prolonged simulation times and bespoke data requirements. The computational efficiency of both disaster and recovery simulations is a governing factor in the success of the proposed methodology, making the choice of earthquakes particularly suitable.

The specific interdependency considered in this thesis is the impact of building debris on adjacent traffic links and, by extension, overall transportation network performance. Consequently, the analysis focuses on two networks: the Building Portfolio (BP) and the Transportation Network (TN). Community resilience losses are measured through repair times, repair costs, relocation costs, income losses, and traffic delay costs, as these metrics align well with the interdependency between roads and buildings. The table below positions this thesis within the broader field of infrastructure management by outlining the key study areas of focus.

Category	Model Component	Included
Hazard Model	Earthquake	Yes
	Flood	No
	Tsunami	No
	Hurricane	No
	Drought	No
	Multi-Hazard	No
Infrastructure Model	Transportation Network (TN)	Yes
	Building Portfolio (BP)	Yes
	Gas Network (GN)	No
	Electrical Power Network (EPN)	No
	Interdependencies	Yes
Phase	Mitigation	No
	Recovery	Yes
	Maintenance (Deterioration)	No
Stochasticity	Stochastic Disaster Model	Yes
	Stochastic Loss Model	Yes
	Stochastic Structural Response Model	Yes

Table 1.1: Model Components and Their Inclusion Status

The problem statement of this thesis is:

- Natural Disasters are increasing in frequency and are causing exceeding economic and human losses to communities around the world. Decision-makers have little access to tools that can help them reduce their communities' disaster risk before, during and after disastrous events.

The research questions and objectives of this thesis are outlined as:

1. **Principal Research Question:**
   (a) How effective is Reinforcement Learning when used as a decision-making tool for post-earthquake repair scheduling of interdependent infrastructures and when compared to baseline methods?
2. **Sub-Questions:**
   (a) How accurate is the computational modelling of earthquakes for different locations?
   (b) What are the factors contributing to community functionality before and after an earthquake disaster?
   (c) How can different community functionality metrics be distilled into an aggregate community functionality metric?
   (d) How can an aggregate community functionality metric be used to describe the resilience of a community in terms of its ability to rebound after an earthquake?

The main contribution of this thesis while answering the above questions is the modelling of a community before, during and after a disaster as a holistic simulation environment. This is implemented in the Python programming language. The key objectives are then to:

- Model building and road infrastructure, their interdependencies, and their response to an earthquake in an integrated simulation environment.
- Formulate a description of community functionality in an aggregate metric that uses existing knowledge on community performance indicators.
- Provide means of modelling the environment such that it can be applicable in a wide range of contexts.
- Use Deep Reinforcement Learning as a decision-making tool for favourable repair scheduling policies
- Compare Deep Reinforcement Learning to heuristic repair scheduling strategies.

1.5. Related Work

Having established the problem statement and objectives of this thesis in the above sections, this section serves as a brief overview of the general approaches for providing decision-making support frameworks for infrastructure management. This section does not provide details on the methodology of each research but rather points out the efforts that have been made. The respective methodologies are explained in detail in the literature review. The two overarching objectives of this section are to then provide an insight into existing work on Deep Reinforcement Learning (DRL) for infrastructure management, but also infrastructure resilience modelling. In the context of this thesis, DRL for infrastructure management can be seen as the *solution concept* of a *game*, where infrastructure resilience modelling is the *game* or *environment*. The two study aspects work together to form a decision support framework.

While the application of Deep Reinforcement Learning (DRL) to infrastructure management has predominantly focused on deterioration rather than post-disaster repair scheduling, the underlying optimization problem remains the same: scheduling inspection and repair actions to maximize an objective function. The key modelling dimensions of an infrastructure management decision making environment are:

- Infrastructure component type-agnostic or type-specific formulations.

- The observability of component states, i.e fully-observable or partially observable.
- The stationarity of the system's dynamics, i.e the change in the confidence of observations across, usually, time.

Firstly, several research efforts have been conducted on abstract and component-specific environments that are aimed to generalise well and formalise the overarching methods concerning DRL for infrastructure management, [6, 7, 11, 12, 57, 62]. While these works often show crossover between component-specific and component-agnostic environments, they aim to define the mathematical formulation of the environment and the objective function rigorously. The infrastructure environments are often *parallel*, *in-series* or are described by *k-ou-of-n* failure modes. *Parallel* and *in-series* environments describe the way components might be dependent to each other and are useful for systems that rely on inter-component connectivity, [14]. *K-out-of-n* environments are ones where failure modes are defined when k components out of a total of n components are in a severe damage state. The authors use DRL and conduct comparative analysis of its performance against classical or custom inspection and maintenance planning strategies. These approaches are also applied to transportation networks, [79, 47, 55]. The key takeaways from this area of research principally relate to the formulation of the environment (i.e type of infrastructure, type of actions etc.) and by extension to the DRL solver used to maximise the objective function. There are summarised as:

- Maintenance planning DRL environments are most often modelled with agents that have partial observability over their attributed components. Inspection actions increase an agent's confidence in the observation they already hold, or change it altogether.
- Repair actions can either major or minor and relate to the repair effort and the cost associated with carrying out that repair effort.
- Without considering the objective function or the DRL algorithms used, the behaviour of inspection actions the belief updates of agents is the crux of modelling such environments. Inspection behaviour can be non-stationary across a realisation and it can also affect the belief over states of neighbouring components, such as *correlated k-out-of-n* environments, [57].
- In practical terms, DRL is extremely resource-intensive and requires state-of-the-art hardware to *train* within reasonable timescales.
- A researcher's choice of DRL algorithms is usually governed by the training and execution being either centralised or decentralised. Centralised training defines that all agents have access to all other agents' beliefs / states. This is to say that the input feature vector to the $ANNs$ grows in proportion to the number of agents. Centralised execution defines that the controller network has access to all agents' beliefs. Generally, centralised methods perform better for smaller environments but don't scale well to larger ones due to the curse of dimensionality, i.e all things considered, many smaller networks are easier to train than one big network, [48].

When looking closer at work focusing on post-earthquake recovery and DRL, there are few efforts that explicitly address this problem. Two efforts exists that address this problem directly, [31, 96]. The first looks are a multi-hazard approach for road network recovery using Singe-Agent DRL and the second focuses on Multi-Agent DRL for post-earthquake recovery of interdependent infrastructure networks. They both used graph-based state definitions to describe the state of the networks as they can be useful in describing transportation networks. Secondly, both research efforts make use of *community resilience* as a concept around which the objective function is defined. This contrasts the other works mentioned that usually employ monetary costs as negative rewards. Some key takeaways are:

- Graph Neural Networks are intensive to train as the graph embedding needs to be learnt concurrently with the action prediction. This adds a significant layer of complexity to the overall approach. However, $GCNs$ should generally be expected to perform better in this kind of task, even though there doesn't exist work on post-earthquake recovery that does not use $GCNs$.
- Reward is formulated as positive when it is a function of resilience or negative when it is a function of resilience loss.
- Single- and Multi-Agent DRL should be expected to perform better than rule-based scheduling algorithms, given training convergence.

To contextualise these research efforts, approaches that do not use DRL might include networks in the order of $10^2 - 10^3$ components, whereas Single-, and especially Multi-Agent DRL has not seen any use for environments above 10^2 components. Overall, the benefits of DRL are clear but still require further research to be applied on a large scale. While approaches on the side of reliability research and structural risk are certainly advancing in being able to use larger environments, such breakthroughs are also largely expected due to advances in computer science, where the theory and algorithms for DRL are developed before they see use in reliability research. In general, the research falls into two main categories: one that uses deteriorating environments and one that uses post-hazard environments. The former approaches usually employ monetary cost-based rewards and partial observability over damage states, while the latter uses resilience- or resilience-loss based reward or costs and employs full observability over damage states.

1.6. Contributions

The contributions made in this thesis are principally concerned with the environment formulation and community functionality / resilience definitions. The thesis uses adapted but existing DRL frameworks for infrastructure management. The proposed contributions are principally seen the integration of such frameworks in a post-disaster scenario, but also in the definition of community functionality as a holistic measure that aims to incorporate various indirect losses related to a community's well-being. An overview of the methodology is provided in figure 1.9. First, data is collected to produce an test bed environment, then the various losses are predicted. The interdependency of roads to building debris and the computed losses are all used to compute community functionality and thus resilience losses. This is then used as the cost, or negative reward of an MARL agent. The results of which are benchmarked against random and importance-based policies.

Generalisable Study Environments: A methodology is introduced for automatically generating generalisable test-beds for any metropolitan area within the United States. This framework leverages open-source data from OpenStreetMap (OSM) and the National Structure Inventory (NSI), integrating established vulnerability and fragility functions from HAZUS and the Federal Highway Administration (FHWA). This approach ensures a unified data schema across generated environments, a significant improvement over previous single-context studies. While highly scalable within data-rich environments like the U.S., global generalisation remains a challenge due to data availability. While these environments are not specifically tested using DRL given their large scale, the functions for doing so in the future are provided.

Holistic Loss and Interdependency Modelling: This research incorporates indirect losses and an interdependency which are not seen in the studies looking at DRL for infrastructure management. Beyond direct building and road repair costs and times, the thesis accounts for traffic delay costs, income losses, and relocation costs. Crucially, the interdependency between building debris and traffic link capacity reduction is modelled, impacting community functionality through cumulative traffic delay costs. Additionally, the normalised performance indicators of hospitals and other lifeline infrastructure components are considered and are aggregated with economic functionality to compute a holistic measure of community functionality. This expands upon prior work, such as that by Yang et al. [96], by integrating economic functionality (income losses and relocation costs) alongside infrastructure performance, thereby offering a more complete assessment of post-disaster community resilience.

Probabilistic Seismic Hazard: Unlike other approaches that use a deterministic earthquake scenario. This thesis considers the simulation of a multitude of earthquake ground motion before training and samples a given ground motion profile in each training realisation. Ground motion prediction is done multiple times per earthquake magnitude, thus providing a diverse set of data that can be used to simulated losses. Fig. 1.12 shows the expected initial functionality robustness for each earthquake that is simulated. Robustness in this case is the residual community functionality just after the earthquake.

While not expanding on the specifics of the results themselves, the comparison of performance between heuristic and DRL approaches is significant in the prioritisation of recovery and the $N - percent$ resilience. In essence, the tested heuristic policy is seen to perform well in effectively recovering 100% of community functionality, whereas the best performing DRL policies perform better when considering

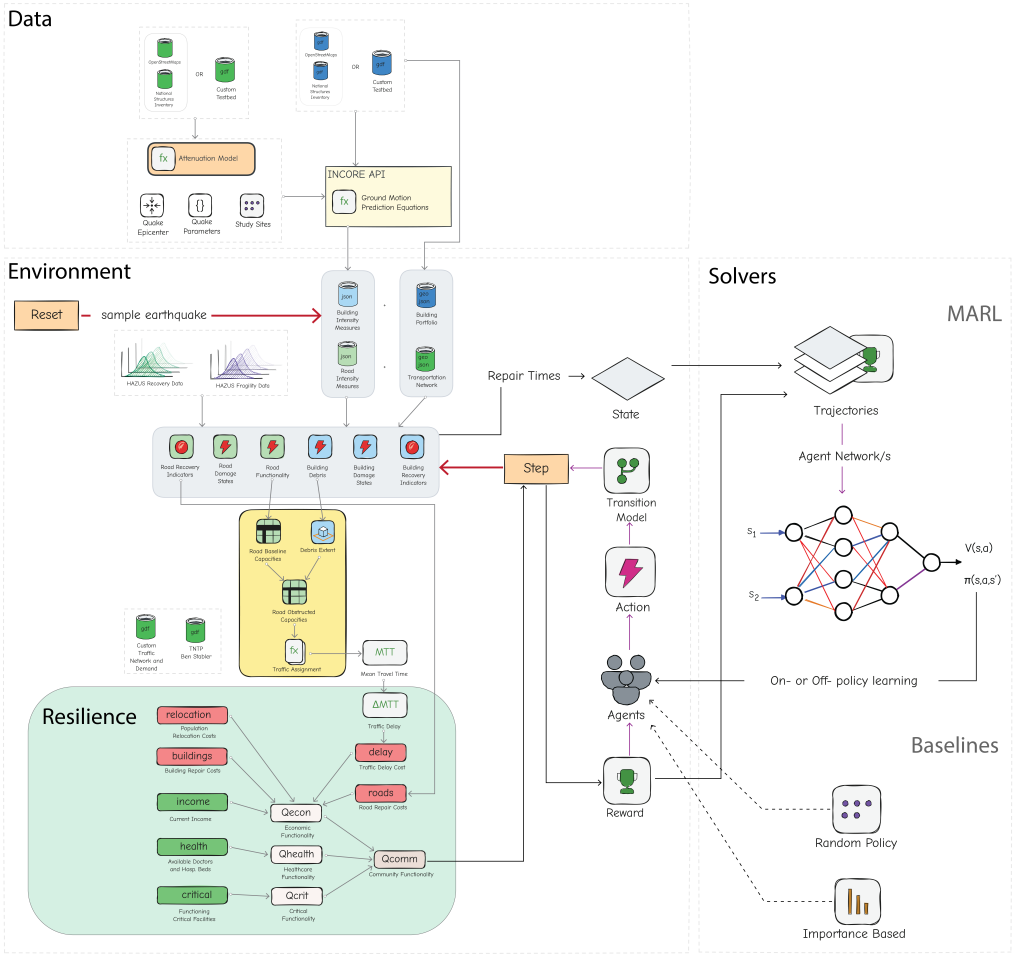


Figure 1.9: Methodology Flowchart

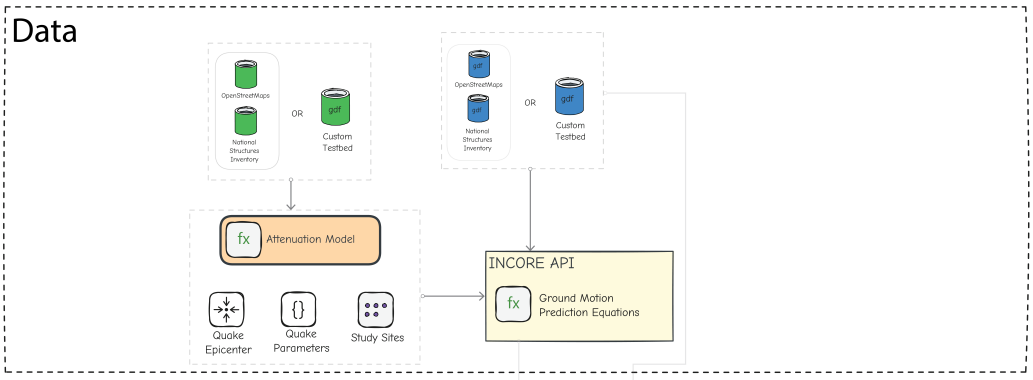


Figure 1.10: The methodology includes functions for incorporating custom test beds or download existing datasets from US-based metropolitan areas

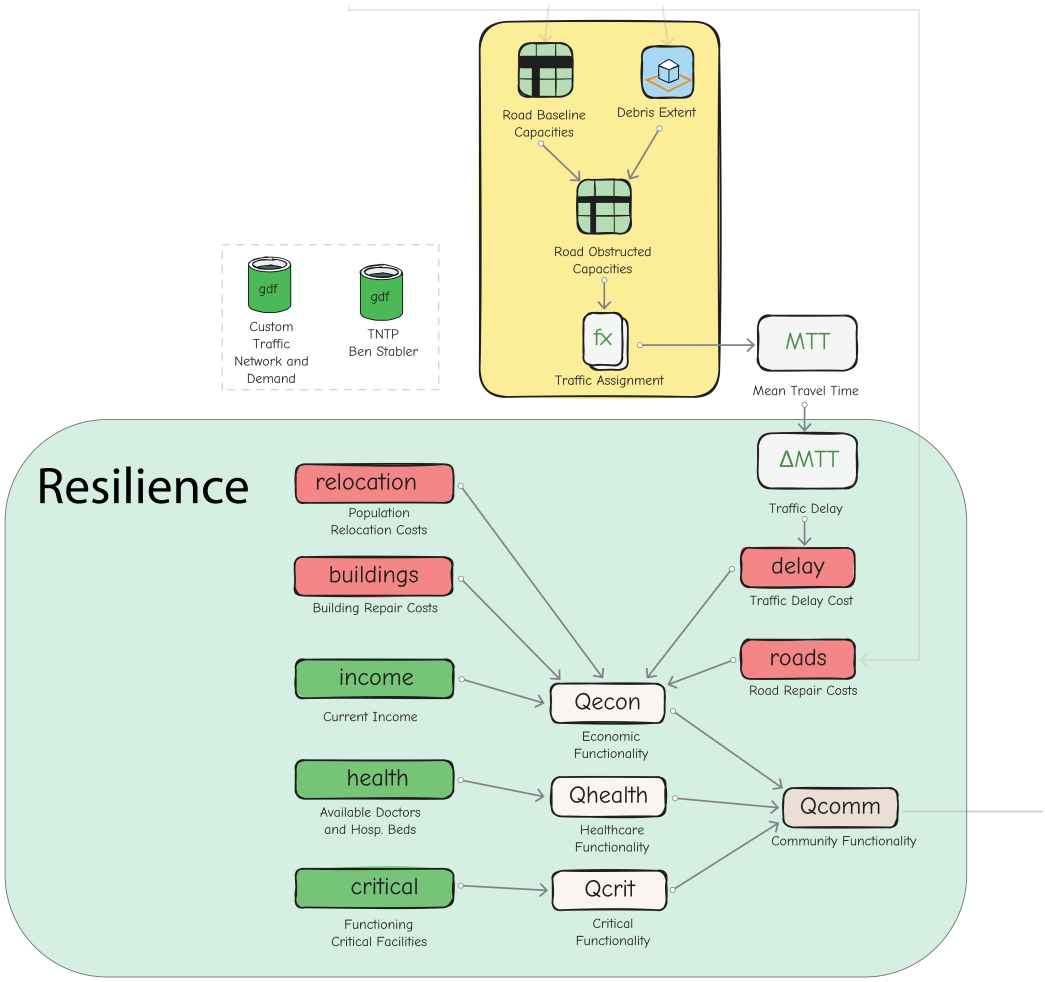


Figure 1.11: The proposed resilience formulation uses critical, economic, transport and healthcare functionalities to compute an aggregate metric.

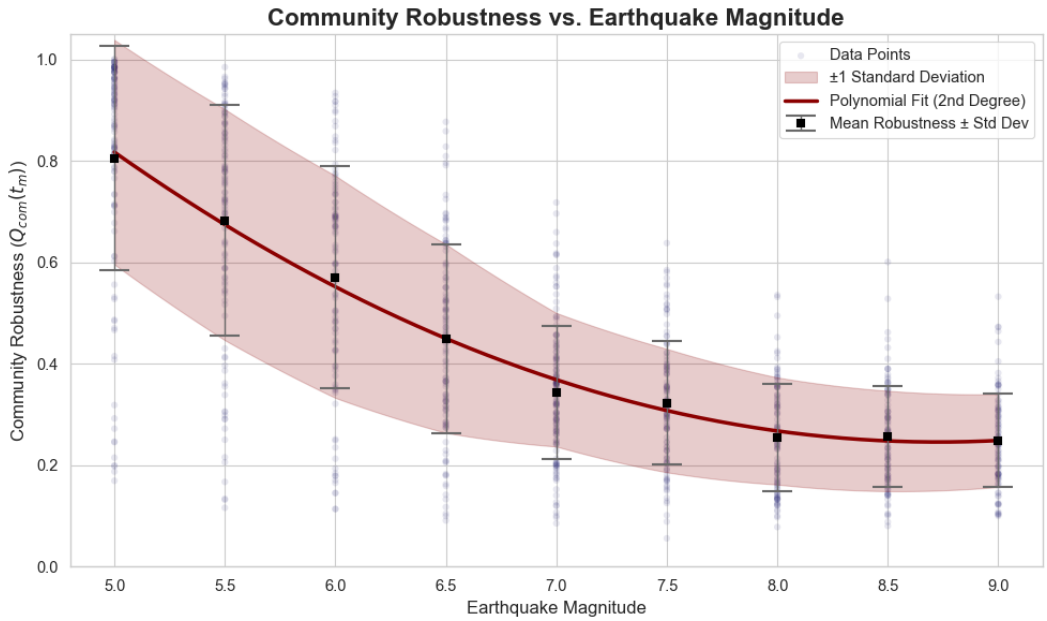


Figure 1.12: A scenario-based seismic hazard assessment is used. Synthetic earthquake IMs are attenuated with magnitudes ranging from 6.0 to 8.0 M with 0.5 increments. 100 realisations are generated per magnitude. The residual functionality (y-axis) is the functionality that is *left-over* just after the earthquake.

the partial recovery of the network. In doing so, DRL shows poor performance in fully recovering the community, but with steeper early recovery curves, achieving fractional recovery quicker than the heuristic policy. This shows that DRL performs favourably early on in the recovery process, but is unable to achieve full recovery as the rewards for doing become negligible later on in the recovery process. Figures X and Y show a snapshot of the results; cumulative losses when using DRL policies generated from three different algorithms: $DCMAC$, $QMIX-PS$ and $VDN-PS$ on average only match or fall short of $IMPB$ (Importance-Based) repair scheduling and are even poorer-performing than a random policy in the case of $VDN-PS$ and $QMIX-PS$. Conversely, when considering recovery of 70% of the initial losses, the three algorithms perform better than $IMPB$, yielding lower losses for partial recovery.

The reasons for the behaviour exhibited by the DRL algorithms are discussed at length in the results section; however, an important aspect of this recovery behaviour is the limitation posed by computational complexity. DRL training is seen to take in excess of 30 hours, given the available hardware, and thus full training convergence is not achieved. It is difficult to say that this behaviour can be completely eradicated just by training for more time-steps, however the lack of convergence points in that direction. Additionally, other limitations relating to the production of environments and how closely they can model a realistic infrastructure network are also of interest and are discussed.

Conclusively, the environment modelling contributions lie in the modelling of an interdependent infrastructure network such that it reflects holistic community resilience. Additionally, probabilistic seismic hazard assessment allows for the evaluation of stochastic disruptions to the network. DRL is shown to perform very well early on in the recovery process by *rebounding* the initial loss of resilience swiftly; however, it falls short in fully recovering all losses induces by seismic impact. Nevertheless, the results show that the advantage of using DRL lies in the early recovery phase; doing so by effective repair prioritisation in a stochastic infrastructure management environment.

1.7. Summary

This thesis addresses the challenge of infrastructure reliability in the face of extreme weather events, specifically focusing on post-earthquake repair scheduling of interdependent infrastructure networks. It proposes and evaluates Multi-Agent Reinforcement Learning ($MARL$) as a decision-making tool,

benchmarking its effectiveness against importance-based approaches under uncertainty.

The research concentrates on the post-disaster recovery phase, identifying a gap in existing literature that predominantly focuses on mitigation and preparedness. Earthquakes are chosen as the specific hazard due to the maturity of available simulation tools, but also their relatively swift simulation time. The study investigates the interdependency between building debris and adjacent traffic links, impacting mean travel time. The analysis centres on the Building Portfolio (BP) and Transportation Network (TN), with economic resilience losses quantified by repair times and costs, relocation costs, income losses, and traffic delay costs. Critical and healthcare resilience losses are integrated with economic ones to derive an aggregate measure of holistic community losses.

Key contributions include a methodology for generating generalizable test-beds for U.S-based metropolitan areas using open-source data and established vulnerability and fragility functions. Furthermore, the thesis introduces holistic loss modelling by using existing research, encompassing direct and indirect losses and the impact of building debris on traffic capacity. Unlike deterministic approaches, this work incorporates a probabilistic seismic hazard by simulating a multitude of earthquake ground motions. The research leverages an adapted DRL framework used for deteriorating infrastructure environments, with the developed code openly published on GitHub to foster transparency and future research.

2

Literature Review

This research investigates strategies to enhance the recovery resilience of communities following earthquake disasters. As outlined in previous sections, the general approach models the post-earthquake repair scheduling problem as a *game*, and explores the application of *Multi-Agent Reinforcement Learning (MARL)* as a *solution concept*. The performance of MARL is benchmarked against baseline methods, such as importance-based repair scheduling. This literature review is structured as follows:

1. **Post-Earthquake Repair Scheduling Environments**
   (a) Studies on computational modelling of earthquakes, Ground Motion Prediction Equations ($GMPEs$) and Probabilistic Seismic Hazard Assessment ($PSHA$),
   (b) Research on modelling interdepent infrastructure networks,
   (c) Literature on community resilience modelling,
2. **State of the Art of Infrastructure Management Literature**
   (a) Research on using DRL for deteriorating infrastructure environments,
   (b) Research on using DRL for post-earthquake repair scheduling environments,
   (c) Research on heuristic or rule-based approaches for infrastructure management.

2.1. Seismic Hazard Assessment

Predicting ground motion during an earthquake event is usually done with attenuation models, which use $GMPEs$ (Ground Motion Prediction Equations). While approaches differ, the conventional method to model earthquake ground motion in seismically active regions begins in the collection of strong motion data on the field using accelerograms and is followed by regression analysis [56]. Strong motion refers to the motion caused by earthquake shaking within a certain distance away from the quake epicentre, usually around 50 km but varying based on the specifics of the collected data. Regression analysis results in $GMPEs$, which are algebraic expressions that predict PGA (Peak Ground Acceleration) and SA (Response Spectral Acceleration). The variables contributing to $GMPEs$ are the earthquake magnitude (M), site-source distance (R) and the qualitative site-class, which is a categorical variable that usually considers an aggregate of metrics related to soil-type and earthquake fault-type [56]. Other intensity measures (IMs) can be derived from the three mentioned above using established theorems in seismology, and more generally in physics.

Intensity Measures are related to the disaster event itself and their prediction is an area of research in seismology and geology. Many different IMs can be produced by an earthquake, however three main ones are generally used in fragility assessment:

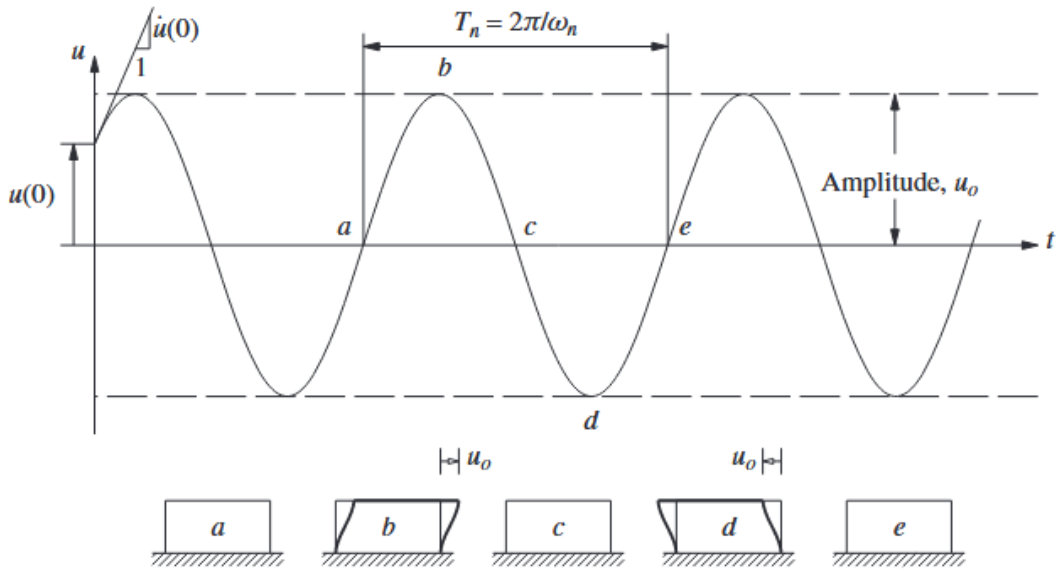


Figure 2.1: Illustration of a structure oscilating at its antural frequency, Chopra [23]

- **Peak Ground Acceleration**, PGA (g) is the highest rate of change of ground velocity at a particular site. PGA is the IM that is used most commonly.
- **Permanent Ground Displacement**, PGD $(inches, cm, m...)$ is the ground displacement that remains after an earthquake at a particular site. It is often caused by lateral spreading and soil liquefaction. This IM is not often directly attenuated using $GMPEs$, it often derived using PGA and certain soil classification variables.
- **Spectral Acceleration**, SA (g) is earthquake-agnostic and structure-specific. It is used to describe the resulting acceleration of a structure when subject to oscillation with a particular period of vibration.

PGA and PGD are more intuitive to understand than SA as it is dependent on the natural period of vibration of a structure. The natural period of a system is the period with which a system oscillates such as to produce the largest displacement from its equilibrium position, and is the frequency it will oscillate with when displaced and allowed to move freely.

Figure 2.1 shows the free vibration of a structure oscillating between states a, b, c, d, e. *Free* is taken to mean without the effect of external forces. The natural frequency is then defined as the frequency T_n at which the maximum displacement u_0 is achieved. Considering an earthquake with a period of vibration T_n then this structure would vibrate with a maximum displacement u_0. This displacement, in the context of buildings, is often said to be the inter-story drift ratio, i.e. the relative horizontal movement of one floor to the next. Therefore, the spectral acceleration of a structure is the maximum acceleration it will experience given a specific natural period of vibration.

Considering the overarching theme of fragility, specific spectral accelerations are used based on the expected natural period of vibrations of different structures. For instance $FEMA$ prescribes the use of $SA_{0.3s}$ and $SA_{1.0s}$ for the fragility analysis of bridges as through empirical research it was deduced that most bridges tend to have a natural period of vibration of 0.3 to 1.0 second, $FEMA$ [35]. Acceleration is not the only IM that can be analysed spectrally; however, spectral velocity and displacement are rarely mentioned in fragility analysis research and are thus not explored in this thesis.

Probabilistic Seismic Hazard Assessment ($PSHA$) is the underlying objective of using $GMPEs$ [56]. $PSHA$ is concerned with producing seismic hazard maps and using those to model earthquake risk in seismically active regions. The crux of this area of research comes in identifying the distribution of active faults in regions with low to moderate seismic activity and allowing different models to generalise over a broad range of site-source distances and site-conditions. Civil and other engineers don't usually

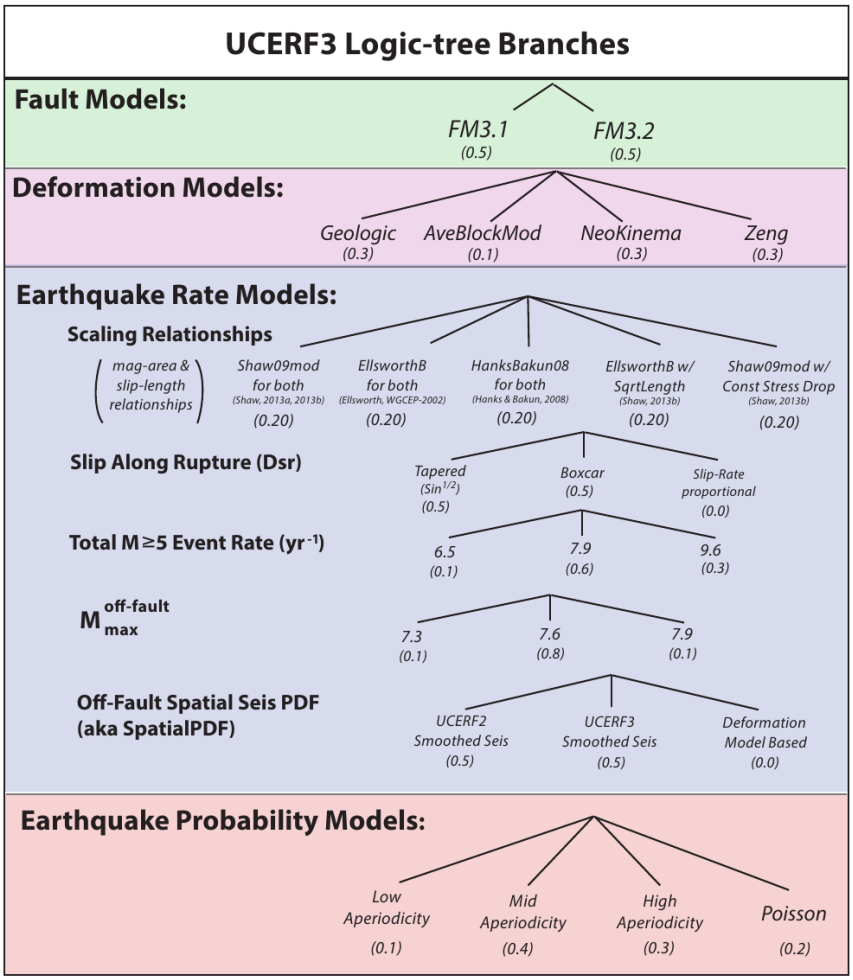


Figure 2.2: Earthquake forecast logic tree for the $UCERF3$ Fault database. Field et al. [37]. Each child node is connected to the following parent node of the next model, thus the illustration shown here is condensed for visual purposes

concern themselves with $GMPEs$ specifically; rather, international, national and regional building codes specify design response spectra which agree (within reason) to predictions made by $PSHA$ models, like in $Eurocode8$, [24]. It is important to note that the crux and benefit of employing $PSHA$ is in *logic-trees* which are decision trees that aim to hedge the epistemic uncertainty associated with different fault, attenuation, magnitude scaling and other seismic component models [8]. Epistemic uncertainty is a *lack-of-knowledge* uncertainty, where we say that we are not sure how exactly an earthquake should be modelled, but given a set of of modelling techniques, we can sample a sequence of models and given a large enough set of samples, $PSHA$ should approximate ground motion with an acceptable degree of error [8]. The type of models used in each phase of the logic tree and the weights of each model are not a trivial task. $PSHA$ at its core remains an open question, with several available approaches that usually site-specific. For example a widely accepted $PSHA$ logic tree for California earthquake forecasting is shown in Figure figure 2.2, following the Third California Earthquake Rupture Forecast ($UCERF3$) [37]. Note that this model does not include attenuation and is purely concerned with the occurrence model of earthquakes in California. It can then be easy to see that holistic models that include rupture and occurrence modelling as well as attenuation modelling are exceedingly difficult to formulate.

The potential disconnect between design response spectra and $PSHA$ becomes particularly important for site conditions that are *unique* or severely out of the scope of a specific $PSHA$ site-class and site-source distance domain. Other practical challenges in this area of research are the inherent difficulty of

collecting data given the relative low frequency of earthquakes, and that the area for which data should be collected is exceedingly large.

Foundational early research and ongoing large-scale efforts on the field are seen to usually originate from two countries: USA and Japan. In the US, early research on the field was made by James Brune in the 70s and David M. Boore in the 90s, with more recent efforts seen in the development of Next Generation Attenuation Models (NGA), which is an effort by the Pacific Earthquake Engineering Centre (PEER), based at the University of California, Berkley [61] [22], [20], [1]. NGA includes two sub-datasets, $NGA-West2$ and $NGA-East$. Both datasets solely consider shallow-crustal earthquakes; the former is broader and applies to quakes worldwide, in active tectonic regions; the latter is specific to central and eastern US regions. NGA includes 5 attenuation models and a dataset of 3552 recorded earthquake events. The 5 models are:

Reference	$V_{S,30}$	Demand Types	M Range	Period Range	R_{JB}
Atkinson and Boore, 2008 (AB2008)[61]	180 to 1300 m/s	PGA, PGV, 5% PSA	5.0 to 8.0	0.01s to 10s	0 to 200 km
Campbell and Bozorgnia, 2008 (CB2008)[20]	< 180 m/s	PGA, PGV, 5% PSA	4.0 to 8.0	0.01s to 10s	0 to 200 km
Chiou and Youngs, 2008 (CY2008)[22]	300 to 400 m/s	PGA, PGV, 5% PSA	4.0 to 8.0	0.01s to 10s	0 to 200 km
Idriss, 2007 (I2007) [50]	450 to 900 m/s	PGA, PGV, 5% PSA	5.0 to 8.0	0.01s to 10s	0 to 200 km
Abrahamson and Silva, 2010 (AS2010)[1]	1000 m/s	PGA, PGV, 5% PSA	5.0 to 8.5	0.01s to 10s	0 to 200km

Table 2.1: Summary of Ground Motion Prediction Equations (GMPEs) used in NGA

All the reviewed models include simillar boundary conditions and are less accurate for very extreme earthquakes, but also very small earthquakes. Nonetheless, they have all gone under review by the USGS (United States Geological Society) and are in use in many derivate resilience-based hazard-analysis tools like INCORE [90]. Allthough less broad, research in Japan is highly rigorous and consistent and focuses on analysis of soil-conditions specific to Japan (crustal earthquakes), but also the development of advanced GMPEs specifically focusing on energy infrastructure risk reduction and accordance to Senior Seismic Hazard Analysis Committee (SSHAC) Level 3 recommendations [39]. Specifically, given the high density of nuclear power stations in Japan and the combined effect of the country's high seismicity, the research is crucial given the severe implications of nuclear power plant disasters. However, this thesis principally focuses on research efforts originating from the US as they align well with ongoing efforts for structural fragility assessemenet of US-specific infrastructure archetypes.

Computational $GMPE$ tools can be sparse as their integration is often secondary to an different overarching project goal; thus, they are not often produced as stand-alone tools. $INCORE$ [90] and $OpenQuake$ [86] are some of the most accessible tools; while INCORE focuses on US-based integrated community resilience, $OpenQuake$ is centered around world-wide, location-specific attenuation modelling. $INCORE$ is a project from the Center of Excellence for Risk-Based Community Resilience Planing (CoE), which is funded by the National Institute for Standards and Technology ($NIST$). It specifically focuses on post-disaster resilience loss and recovery of interdependent infrastructure systems and communities. It is built as a Python client which connects to distributed servers running on Docker and $Kubernetes$, which are tools to effectively deploy large-scale applications in a wide range of machines. $OpenQuake$ is part of the Global Earthquake Model (GEM) and is a global effort to bring together national and international organisations for the development of both uniform standards and

computational *GMPE* tools [86].

This thesis uses *INCORE* as it is native to a Python environment and does not require re-building it from binaries. *OpenQuake* is more easily used in a web-based interface or integrated desktop client. Additionally, *INCORE* allows for the future integration of more infrastructure networks when appropriate data schemas are used. One can perform many different analyses through *INCORE* that relate to direct, indirect, and human losses. A simplified illustration of *INCORE*'s tech stack is shown in figure 2.3 which is built as a containerised application and includes several web-based tools and the python package itself.

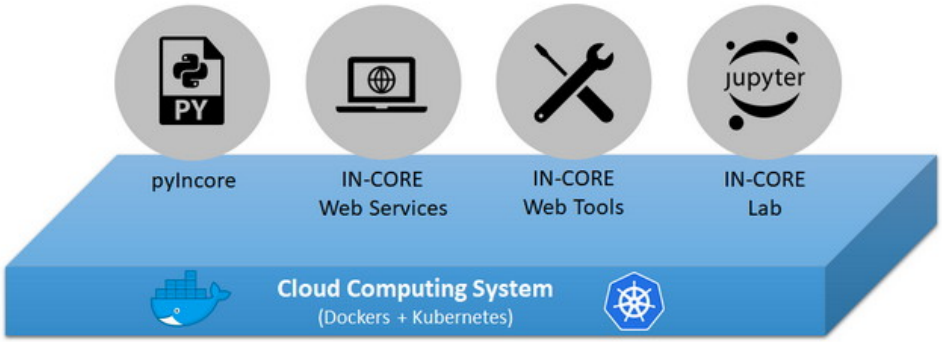


Figure 2.3: INCORE Tech Stack is built as a Docker container and distributed using Kubernetes on AWS (Amazon) servers. This is a common approach for many cloud-based software solutions Standards et al. [90].

2.2. Structural Fragility Assessment

Attenuation models discussed principally focus on shalow-crust earthquake strong motion. This is noted as other types of earthquakes or human-induced earthquakes are not considered of interest. An example of research conducted on human-induced earthquakes and fragility is in the Groningen gas field, where chronic natural gas extraction has led to ground failure and thus human-induced earthquakes [26]. While this research is crucial for regional fragility assessment, it does not generalise for intraplate shallow-crust earthquakes. Large-scale efforts in national- and international- fragility assessment are a topic of research which has only seen significant results in the past decade and is still ongoing.

This section focuses on laying out the foundation for predicting the structural damage caused by earthquakes on infrastructure components. Structural response and thus damage is often analysed at the scale of one or a handful of structures with an existing detailed description of their structural makeup, materials and behaviour. This is done using accurate but computationally expensive methods like Finite Element Analysis (FEA). Conversely, such a task on the scale of a town or a city is difficult as detailed structural descriptions of large building portfolios are not commonplace and the cost of performing FEA at such a scale becomes a limiting factor. Thus, the usual approach is to classify structures and components and use empirical data to fit statistical models for predicting structural response given a loading scenario. This can generally be described as the concept of fragility. Fragility is defined as:

$$Fragility_i = P(DS_n | IM = y) \tag{2.1}$$

where:

- $Fragility_i$ is the probability that a component or system c is at or above damage state DS_n given a realised value y of intensity measure IM.

The *damage state*, DS is a categorical variable used to classify the severity of damage a component or structure has incurred, based on observed indicators. For instance, a building in a *slight* damage state

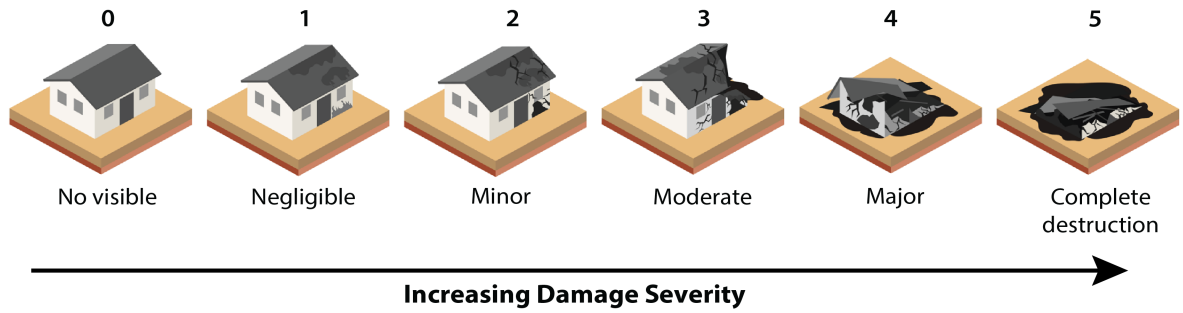


Figure 2.4: A set of example damage that illustrate building damage due to lava flow, *USGS* [87]

might show minor wall cracks and fallen interior panelling. These states, whether defined qualitatively by specific performance indicators or quantitatively, must align with empirical expectations of a structure's usability and collapse risk. Essentially, predicting damage states involves categorizing structural damage into discrete levels, with each level correlating to a specific fraction of baseline functionality or an increased risk of collapse or injury. For instance, a building with *no* or *negligible* damage could be fully functional, while one with *minor* damage might retain half its original functionality, and damage exceeding a moderate level could render it completely dysfunctional. figure 2.4 shows an example set of damage states for buildings affected by lava flows. This thesis employs five damage states, including *None* or *No Visible*, but any number can be used, provided there's enough data to establish the necessary probability distributions.

The task of inventory data collection and structural classification is not trivial and very few counties have made extensive efforts in doing this. Furthermore, a fragility model is only useful for structures whose behaviour is close to that of the data collected. For instance, a three-floor masonry building with a footprint of 1000 m^2 in America is likely to have high similarity in its structural response to an earthquake as another building of the same type within the US. However, the structural response may differ significantly if the same building were located in Europe. This point undermines an overarching theme of this thesis which is that any work in aiming to improve recovery resilience should start with thorough inventory and data collection.

Specific intensity measures are discussed later; however, a brief description is that an intensity measure, in the case of an earthquake, is a type of ground strong-motion which induces a structural demand on a structure. This movement can come in the form of velocity, acceleration, displacement, or any measure which can affect structural damage. While different statistical formulations exist for the definition of fragility, the most normal includes a log-normal cumulative probability distribution function:

$$P(DS_n|IM = x) = \mathbf{\Phi}\left[\frac{1}{\beta_{DS_n}} * \mathbf{ln}\left(\frac{x}{\overline{IM}_{DS_n}}\right)\right] \tag{2.2}$$

where:

- $\mathbf{\Phi}(\cdot)$ is the standard cumulative distribution function (CDF),
- $\beta_{DS_n} = \sigma_{\ln(\mathrm{IM})|\mathrm{DS}_n}$ is the standard deviation of the natural logarithm of the intensity measure at which damage state DS_n occurs, (i.e., the dispersion in naural log space),
- $\overline{IM}_{DS_n}$ is the mean intensity measure at which damage state DS_n occurs,
- x is the realised intensity measure value of intensity measure type IM.

or using the median $\hat{\theta}$ instead:

$$P(C|IM = x) = \mathbf{\Phi}\left[\frac{\ln(x/\theta)}{\beta}\right] \tag{2.3}$$

where $P(C|IM = x)$ is the probability of collapse, given that intensity measure IM is equal to x. Figure 2.5shows two sets of fragility curves for two different building archetypes, as defined in HAZUS's 2024 Earthquake Model Technical Manual, *Equivalent-PGA Structural Fragility - Low-Code Seismic Design Level*, *FEMA* [35]. The specifics of the manual are explored later in the literature review; however, the two sets of curves are shown to highlight the difference between two different types of buildings.

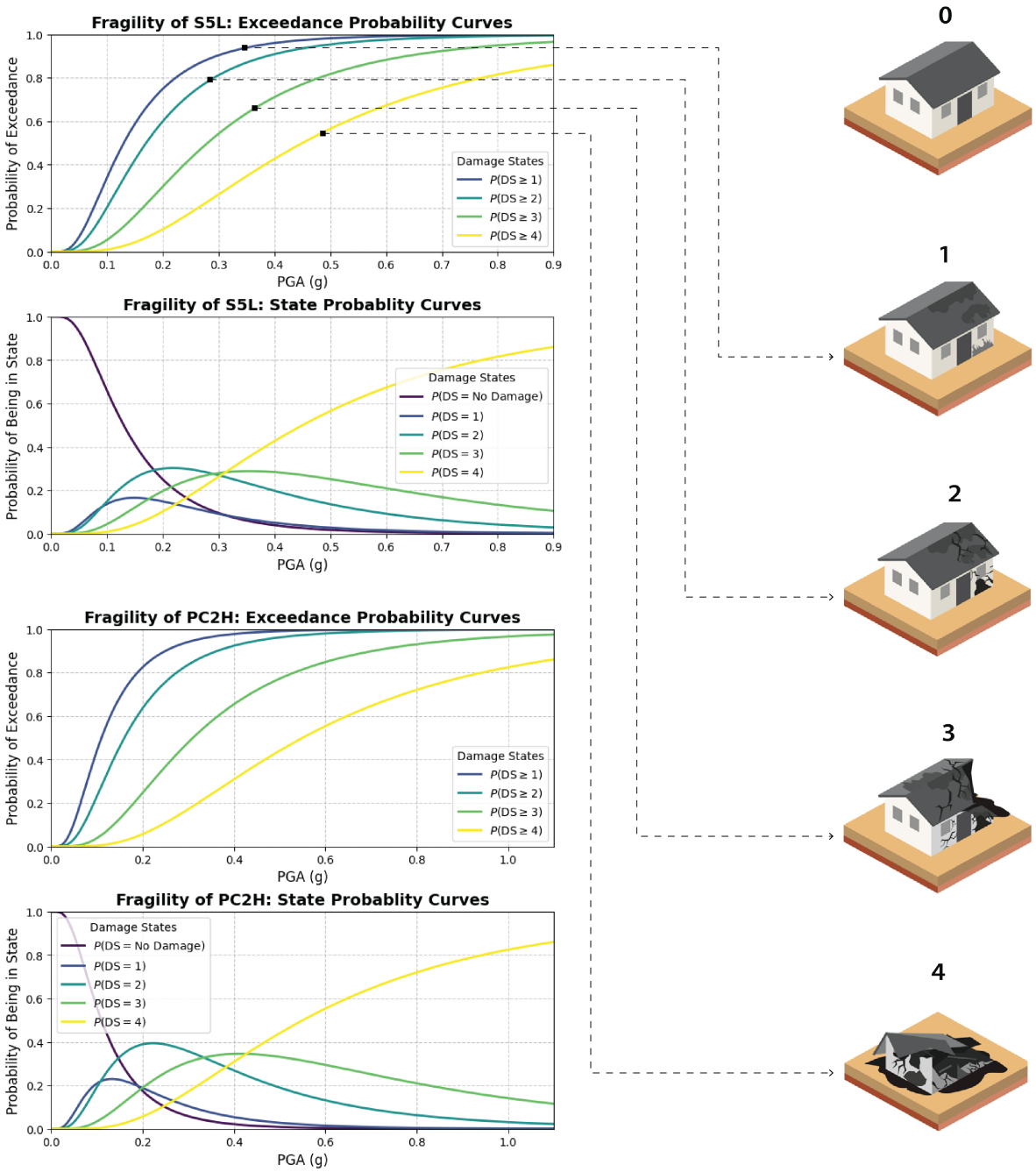


Figure 2.5: Example fragility curves for damage state-exceedance and damage state probabilities from FEMA [35]. S5L = Steel building of 5 floors with a footprint less than 1500m. PC2H = Precast concrete building of 2 floors with a footprint more than 3000m.

One can sample the probability distribution for each damage state, ensuring that the total probability sums to 1. The cumulative probabilities for each damage state can be represented as follows:

- Sample the probability for the most severe state $P(\text{DS4})$:

$$P(\text{DS4}) = p_{\text{sampled, DS4}}$$

- Sample the probability for DS3, and subtract $P(\text{DS4})$ to get the probability of being in DS3 but not in DS4:

$$P(\text{DS3}) = p_{\text{sampled, DS3}} - P(\text{DS4})$$

- Sample the probability for DS2, and subtract $P(\text{DS3})$ and $P(\text{DS4})$ to get the probability of being in DS2 but not in DS3 or DS4:

$$P(\text{DS2}) = p_{\text{sampled, DS2}} - P(\text{DS3}) - P(\text{DS4})$$

- Sample the probability for DS1, and subtract $P(\text{DS2})$, $P(\text{DS3})$, and $P(\text{DS4})$ to get the probability of being in DS1 but not in higher states:

$$P(\text{DS1}) = p_{\text{sampled, DS1}} - P(\text{DS2}) - P(\text{DS3}) - P(\text{DS4})$$

- Finally, calculate the probability for "None" (no damage):

$$P(\text{None}) = 1 - P(\text{DS1}) - P(\text{DS2}) - P(\text{DS3}) - P(\text{DS4})$$

Estimating the parameters of analytical fragility functions such β and $\overline{IM}$ (or μ) can be done in a variety of different ways including incremental dynamic analysis, truncated incremental dynamic analysis, multiple stripes analysis. These methods assume the availability of observations or structural analysis results [9]. Incremental dynamic analysis involves scaling a given ground motion parameter until it causes the simulated or observed collapse of a structure; the estimation of the probability of collapse at a given IM intensity level, x is then the fraction of records that collapse below x, [9]. The dispersion β and median θ parameters can then be estimated by taking the natural log of x associated with the onset of collapse. These can be estimated using:

$$\ln(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n}\ln(\text{IM}_i) \tag{2.4}$$

$$\hat{\beta} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\ln\left(\frac{\text{IM}_i}{\hat{\theta}}\right)\right)^2} \tag{2.5}$$

where:

- $\hat{\theta}$ is the IM level with a 50% probabiltiy of collapse (the median)
- β is the standard deviation in natural log space
- n is the number of samples

These are used to fit the fragility function following equation (2.3). An example of this fitting is seen below:

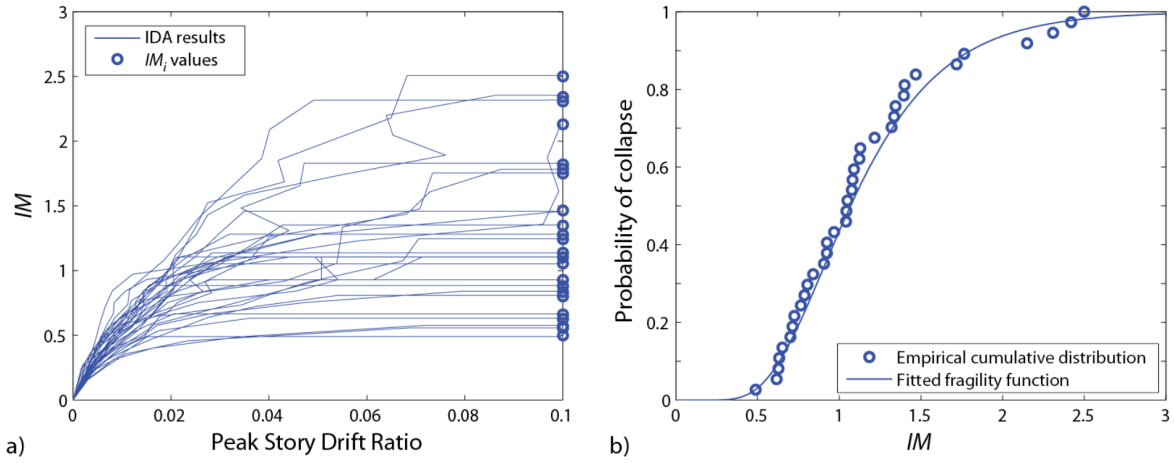
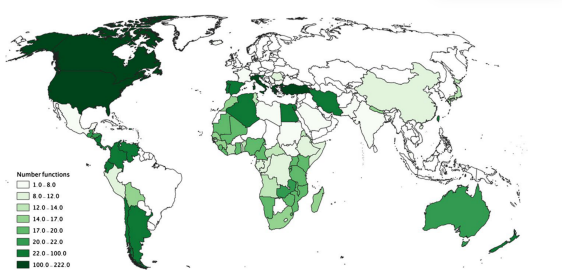


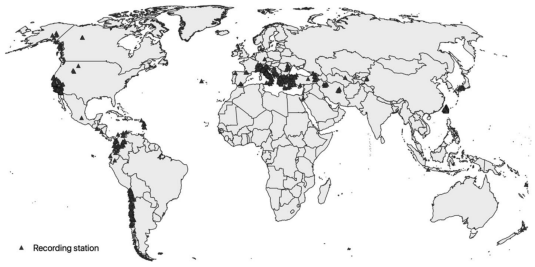
Figure 2.6: a) Example incremental dynamic analyses results, used to identify IM values associated with collapse for each ground motion. b) Observed fractions of collapse as a function of IM, and a fragility function estimated using equations (2.4) and (2.5). Extracted from Baker [9]

Truncated incremental dynamic analysis truncated the samples used for fitting below a certain threshold as analysis of exceedingly high IM values is computationally expensive and normally lower IMs are of higher interest. Multiple stripe analysis involves the discretisation of IM levels and different ground motions for each IM level [9].

$OpenQuake$ has made a significant effort in aggregating fragility assessment models around the world and has compiled them into open computational tools[86].



(a) Available regional fragility and seismic hazard assessment tools using $OpenQuake$ [86]



(b) Location of ground motion recording stations around the world (accelerometers) [86]

Figure 2.7: Available GMPE / Fragiltiy modelling tooling and locations of ground recording stations[86].

As can be seen by figure 2.7a and figure 2.7, there is a clear correlation between the number of recording stations (accelerometers) and the extent of fragility assessment modelling. Four countries that show more extensive efforts to model regional infrastructure fragility are the US, Turkey, Italy, and Greece. Fragility assessment is normally done by predicting the occurrence of complete or partial structural collapse, given a ground motion and a structural analysis result, [9]. These functions can be derived in a number of different ways including field observations and qualitative judgment; however, the focus in this thesis is to use analytical fragility functions which can accept varied IM values, given enough analyses for those IMs.

Fragility assessment for infrastructure in Greece includes assessment of timber-frame masonry buildings [53], fragility assessment of retaining walls [54] post-earthquake recovery phase monitoring using digitally sourced UAS video footage [89], and rapid damage assessment, recovery, education, and resilience in post-earthquake scenarios [70]. Efforts in Italy are seen in scenario-based fragility assessment of non-ductile RC buildings [40], empirical derivation of fragility curves from post-disaster survey data [76], empirical derivation of fragility curves for residential RC buildings [75], and URM buildings [27], and urban-based fragility curves for URM buildings in small- or medium-sized communities in Italy [80].

Existing work in Turkey includes studies like the assessment of one-story precast RC buildings [69] and [83], seismic fragility assessment of typical high-way bridges, fragility assessmenent of public buildings [13] and fragility assessement of mid-rise RC buildings [30]. All though these studies are highly extensive, the lack of large-scale GIS data on transportation and building portfolios makes the use of these analysis methods difficult as crucial large-scale data on the structural makeup of buildings is hard to come by in southern Europe.

Conversely, the US provides the most in-depth and publicly-available fragility and vulnerability assessment frameworks that exist to date. These efforts are seen principally in organisations like FEMA, who have developed HAZUS. HAZUS is an integrated software package for earthquake risk assessment and also includes extensive documentation, results and relevant datasets available for public use. Furthermore, FEMA P-58 methodology Volumes 1 to 7 provide building-specific tools for fragility and vulnerability assessment [36]. FEMA P-58 spans from basic fragility assessment to environmental, human, direct and indirect losses and implementation guides. It is important to note that the US is in the favourable position of having relatively newly- and thus more uniformly-constructed buildings as compared to southern Europe. This makes the act of fragility and vulnerability assessment less uncertain and allows for a higher level of standardisation; this is because infrastructure can be categorised into archetypes / classes more easily. Furthermore, one has to keep in mind that this assessment is specific to earthquake disasters. For instance, all four countries are faced with high and increasing wildfire risk and little available research and tools to reduce or combat it.

FEMA's HAZUS Earthquake Model Technical Manual [35] and Inventory Technical Manual [34] are comprehensive guides on the topic and were released in 2024 and 2022 respectively. The former includes FEMA's HAZUS Loss Estimation Methodology. Within it, one can find data and methods on US site-classification and ground motion, physical (structural and non-structural) damage through fragility assessment on buildings and infrastructure (EPN, WN, GS, TN), vulnerability assessment in terms of human, direct and indirect losses and induced damage functions such as urban fires following earthquakes. The inventory technical manual includes sources and specific data on building and infrastructure inventory data collection, demographic data such as income, rent and healthcare facilities, and various factors used for different loss estimations.

2.3. Infrastructure Interdependencies

This section of the literature review focuses on existing formulations of modelling interdependencies in infrastructure systems. Interdependencies generally refer to the cascading effects of one type of infrastructure to one or more other infrastructures. For instance, one can think of the inability of water wells to functions without functioning EPN substations. Interdependencies of different infrastructures usually rely on associative relationships between different infrastructures to couple them together in such a way that performance reduction in one type of infrastrustructure affects another type of infrastructure. Guidotti et al. describe and formalize methods for modelling critical infrastructures while considering these dependencies under a post-disaster scenario [45]. The authors' goals are to develop a methodology that models network dependencies, integrate it into a probabilistic framework which considers the stochastic aspects of infrastructure modelling and apply the framework to a case study involving Water Networks (WN) and Electrical Power Networks (EPN), both having either demand or supply nodes. More generally, the authors define four types of interdependencies (Guidotti et al., 2016):

- **Physical** : the functionality state of one infrastructure affects the material outputs of another infrastructure (i.e non functional EPN substations affect co-located water pumps).
- **Cyber**: the transmission of information from one infrastructure affects the state of another infrastructure (i.e lack of control in a telecommunications network result in EPN functionality reduction).
- **Geographic**: the functionality of an IC is dependent on a proximal or co-located IC of a different infrastructure network (i.e building debris might affect traffic performance in the Transport Network (TN))
- **Logical**: the functionality reduction of one type of infrastructure has wider societal effects that

can likely influence other networks (i.e a drop in performance of the EPN might induce a change in the price of fuel).

The categorical classification of different interdependencies is useful as it allows clarity in what kind of interdependencies are modelled and how exactly cascading effects are predicted. For example a geographic interdependency can be that of building debris having an effect on the transportation network; a logical interdependency can be when the transportation network has an effect on the income of a community if traffic delays increase. Thus, interdependencies are thought of as cascading, such that an effect on one infrastructure affects another, which in itself might affect more systems down the line. The adjacency matrix of a network and the formulation that the authors use is as follows:

Let $A = [a_{i,j}]$ be the adjacency matrix of a graph with n nodes, defined as:

$$a_{i,j} = \begin{cases} 1, & \text{if there is an edge between node } i \text{ and node } j \\ 0, & \text{otherwise} \end{cases}$$

The matrix A can be represented as:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix}$$

The authors then describe the augmented adjacency table $A^{s,t}$ for two networks s and t, with nodes n and k, as a matrix in which each entry represents the dependency of node i in network s on node j in network t. This formulation allows for either a symmetrical or an asymmetrical matrix. For example, debris from building i might directly influence the accessibility or damage of road j, but damage to road j might not necessarily affect the structural integrity of building i.

The augmented adjacency matrix $A^{s,t}$ is defined as:

$$A^{s,t} = [a^*_{i,j}], \quad \text{where} \quad a^*_{i,j} = \begin{cases} 1, & \text{if node } i \in s \text{ depends on node } j \in t \\ 0, & \text{otherwise} \end{cases}$$

Its general form is:

$$A^{s,t} = \begin{bmatrix} a^*_{s_1,t_1} & a^*_{s_1,t_2} & \cdots & a^*_{s_1,t_k} \\ a^*_{s_2,t_1} & a^*_{s_2,t_2} & \cdots & a^*_{s_2,t_k} \\ \vdots & \vdots & \ddots & \vdots \\ a^*_{s_n,t_1} & a^*_{s_n,t_2} & \cdots & a^*_{s_n,t_k} \end{bmatrix}$$

Here, n is the number of nodes in network s, and k is the number of nodes in network t. The structure of A^* enables modeling of complex interdependencies across different network layers, including unidirectional or asymmetric influences.

Furthermore, the authors describe the use of a *likelihood table* L of size $s \times t$, where s and t are the number of nodes in networks s and t, respectively. The likelihood matrix L introduces probabilistic relationships between nodes in different networks. Specifically, if a water well i in network s is powered equally by two electrical power network (EPN) substations j and k in network t, and both substations contribute equally to the operation of the well, then the likelihood values of j and k influencing i are each 0.5.

The likelihood matrix L is defined as:

$$L = [\ell_{i,j}], \quad \text{where} \quad \ell_{i,j} \in [0, 1]$$

To capture both the dependency (from the augmented adjacency matrix A^*) and its weight (in L), the authors define a *dependency matrix* P as the element-wise product of A^* and L:

$$P = [p_{i,j}], \quad \text{where} \quad p_{i,j} = a^*_{i,j} \cdot \ell_{i,j}$$

This matrix P captures both a binary relationship between nodes in networks s and t, as well as the likelihood or weight of that relationship, expressed as a continuous value between 0 and 1. This allows for more nuanced interdependency modelling between infrastructure networks.

Different literature on infrastructure interdependencies and resilience employ similar formalisms to describing the interdependencies of infrastructures. Gonzalez et. al. model EPN, WN and GN networks with a similar approach [44]. The authors use 4 types of interdependencies structured as matrices and related to Guidotti et al, 2016 like so:

- **Geographical**: interdependencies of one node directly affecting another node due to co-location
- **Logical**: a node in a network can only be functional is there is at least one node in its dependent network that is functional.
- **Physical**: a node in a network can only be functional if there is a subset of functional nodes in another network.
- **Physical**: a node in a network can be partially functional if there is a subset of functional nodes in another network with the origin node having partial dependencies (likelihoods) to each node in the subset.

Additionally, the authors also consider that these interdependencies not only affect the performance of interdependent nodes, but also their repair. Specifically, the repair of a node is broken down into a repair preparation and a repair action, i.e preparing the site and conducting the repair. They consider that given two interdependent nodes i and j, repairing node i reduces node j's repair time by a repair preparation time (Gonzalez et al., 2016). The authors then use these interdependencies to solve a traffic assignment problem with custom cost and flow functions based on the interdependencies affecting flow to and from nodes in different networks but also their repair times. A similar formulation is seen in Yang et al. where the authors model EPN, TN and WN as interdependent and attempt to use RL in predicting an optimal sequence of post-earthquake repairs. The authors model the EPN and WN as graphs and a hyper-graph of the geographical interdependencies of the WN to the EPN due to co-location (Yang et. al, 2024). In this case the WN and EPN graphs are two adjacency tables but with added attributes such as repair time, size etc. The hyper-graph connecting them is structured like the augmented adjacency table seen in Guidotti et al., 2016.

Similarly, Sedieket al. model a transport network and healthcare network and their interdependencies to building debris due to building damage under an earthquake [82]. The authors use geographical interdependencies to describe the effect of building debris reducing the capacity of nearby roads and the effect of capacity reduction to traffic performance and healthcare performance, i.e ambulances taking longer to reach the hospital. Furthermore, the authors use the logical interdependency of hospital capacity to building damage, i.e the amount of injured victims affects the capacity of hospitals in terms of the available hospital beds. Ghorbani-Renani et al. model WN, EPN and GN networks with binary relationships (not specifically weights) [42]. Conversely to the other research efforts, the paper considers resilience related to an 'attack' like an earthquake in a three-level approach. This approach includes defending from the attack, intruding the attack and recovering from the attack. They associate a cost for defending, intruding and recovering from an attack to each node. These papers show a varying level of analysis of nodal interdependency in infrastructures; some authors choose to include weights for differently interdependent nodes and different kinds of interdependencies like in Gonzalez et al. ([44]. On the other hand Ghorbani-Renani et al. [42] employ simpler methods of binary interdependency (non-weighted). This shows that both methods can be applicable and eventually relate to the definition of resilience. Sharma et. al. employ a model with binary interdependencies between networks to predict optimal repair scheduling for interdependent infrastructures [84]. What is interesting in their research is that they organize recovery actions of EPN substations into recovery zones which might have an associated set of sub-actions, like repairing circuit breakers or transformers. This allows for the model

to scale up to larger city environments and potentially consider more large-scale interdependencies. However, it is difficult to maintain accuracy to realistic recovery situations as the scale of the environment grows. It is therefore imperative to relate the interdependencies of a model, its scale and how they affect the measurement of resilience.

2.4. Community Functionality and Resilience

This sections provides an overview and review of different approaches to defining and measuring recovery resilience. Focusing on infrastructure systems, it highlights that the combined aim of resilience formulations are to well describe community functionality but also be formulated in such a way as to allow for better post-disaster *bounce-backs* in performance are represented by higher resilience values. Different researchers measure resilience using various sub-system types and behaviours. Gerges et al. provide a general formulation for resilience which is the basis for resilience formulations in this thesis. They define sub-system resilience attributes as weighted sums of sub-system attribute sub-functionalities [41]. For example, the transportation resilience might dependent on both healthcare trips and business trips. They define a given sub-system resilience Res_{sub} as:

$$Res_{sub}(t) = \sum_{i}^{sub} w_{sub}(i) * R_{Q,sub}(t) \tag{2.6}$$

Furthermore, they define a variable, $P \in [0, 1]$ which conditional on the essential facility functionality and discounts total resilience if an *n-out-of-k* essential functionalities are below a threshold minimum value. Sub-system resilience attributes are then defined as:

$$R_{Q_i} = \frac{\int_{t0}^{t} Q_i(t)dt}{T - t0} \tag{2.7}$$

where s is timestep. This general formulation for resilience is commonplace for most literature where resilience and resilience loss are associated with an area under a functionality-time graph.Guidotti et al. focus on Water Network (WN) performance and its interdependency with the Energy Power Network (EPN) [45]. They use two metrics: the fraction of demand nodes in a network that meet a specified baseline demand, where q_n^i is the flow delivered to node n and Q_n^i is the baseline demand for node n at time t (a node meets demand if $q_n^i \geq Q_n^i$); and whether a node meets a pressure threshold of 180 kPa. This approach directly analyzes one network's performance as the measure of resilience, with the EPN interdependency being implicit. Relating this to Firas Gerges et al. [41], this work considers only the Energy attribute of resilience in their measurement.

Gonzalez et al. (2016) focus on interdependencies between EPN, GN and WN. In their work, they consider interdependencies of components in terms of costs of site preparation due to adjacent damage. They define their experiment as an optimisation problem. As well as site preparation costs, five other costs are considered: guaranteeing flow balance, ensuring arc flows are within capacity, linking flow to the functionality of head and tail nodes, and limiting resource use in construction. This is highlighted as it models interdependencies of traffic networks very effectively and in a high level of detail by considering many cascading effects due to interdependencies. If related to Gerges et al. [41], this research considers Energy and Transportation resilience attributes.

Yang et al. (2024) model three networks (TN, EPN, WN) and use Reinforcement Learning (RL) to predict optimal post-earthquake repair sequences [96]. Unlike the previous two, they measure resilience based on the *buildings* that access these networks, not the networks themselves. The loss of resilience after an event ($t = 0$) at time t is defined as:

$$L_R(t) = \int_0^t [Q_0 - Q(t)]dt \tag{2.8}$$

where $Q(t)$ is the community functionality at time t after the earthquake, Q_0 is the pre-disaster community functionality. Community functionality $Q(t)$ is defined as:

$$Q(t) = \frac{1}{C} \sum_i \alpha_i Q_{phy_i}(t) \tag{2.9}$$

where:

- α_i is the importance weight of each building i,
- $Q_{phy_i}(t)$ is the physical functionality of building i,
- C is a normalization constant.

The normalization constant C is defined as:

$$C = \sum_i \alpha_i S_i \tag{2.10}$$

where is the total surface area of each building i. The physical functionality of building i, Q_{phy_i}, is then defined as:

$$Q_{phy_i}(t) = \sum_j \beta_{ij} I_{ij}(t) \tag{2.11}$$

where:

- $I_{ij}(t)$ indicates whether building i has access to use infrastructure j at time t
- β_{ij} represents the importance of infrastructure j to building i

Considering Firas Gerges et al. (2023) [41], Yang et al. employ Transport, Socio-Economic, and Energy attributes to define resilience [96]. This approach focuses on buildings as the source point of losses and resilience is more broadly taken as the approach in the methodology. Furthermore, this conception of resilience is intuitive as most community users more directly interact with buildings than any other infrastructure.

Sediek et al. focus on post-earthquake performance of the healthcare and transportation networks by considering the interdependency between building debris and road capacity [82]. They model the expected human losses in terms of injury or death and measure traffic performance both using general network performance and healthcare-specific post-earthquake trip performance. They predict building debris using available empirical data and NNs. Their formulation for debris prediction and road capacity reduction is used in this thesis methodology. As Fig X shows, they define modes of building collapse for (a) RC Frames, (b) Masonry and (c) other types of buildings. For masonry and other types of buildings, collapse is deterministic through relationships found in associated literature, [81]. For masonry buildings they define four modes of collapse which they predict using a trained NN and median and dispersion values for a,b, c and c; these are properties of a lognormal cumulative distribution function.

For transportation resilience, they propose the Network Resilience Index (NRI), defined as:

$$NRI = \int_0^{T_{NF}} Q_{TN}(t) dt \tag{2.12}$$

where T_{NF} is the time required for the functionality of the transportation network to return to pre-disaster levels, and $Q_{TN}(t)$ is the weighted functionality of the transportation network. $Q_{TN}(t)$ itself is weighted based on link capacity and defined as:

$$Q_{TN}(t) = \frac{\sum_{i=1}^{n} C_i q_i(t)}{\sum_{i=1}^{n} C_i} \tag{2.13}$$

where n is the number of links, C_i is the traffic flow capacity of link i, and $q_i(t)$ is the functionality of link i at time t. The functionality $q_i(t)$ is defined as:

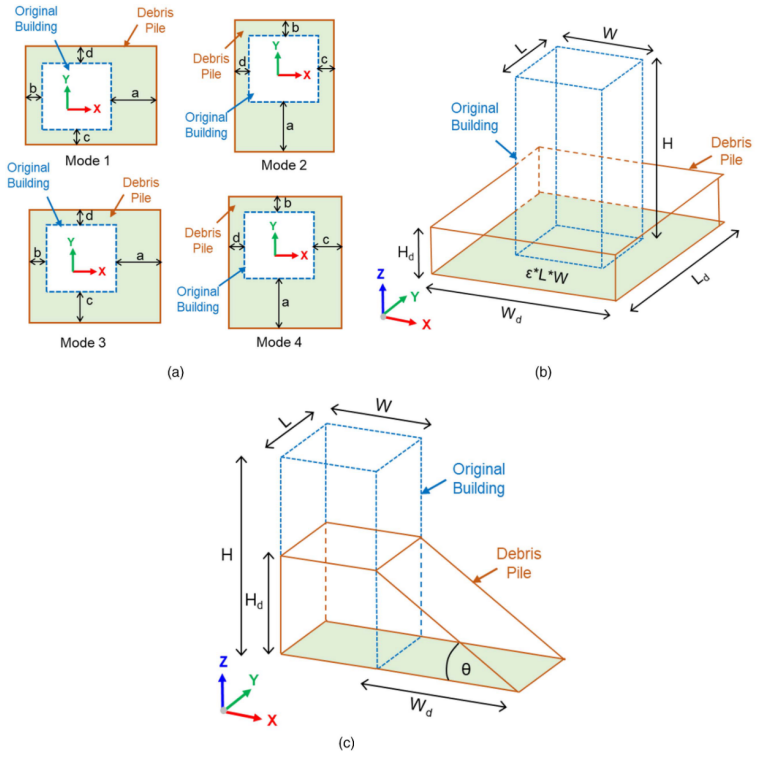


Figure 2.8: Debris collapse per building type, Sediek, El-Tawil, and McCormick [81]

$$q_i(t) = 1 - \%B_i(t) \tag{2.14}$$

where $\%B_i(t)$ is the percentage of blockage of link i at time t, or in other words the capacity reduction of link i at time t. The removal, storage and recycling of debris are all defined during simulation. Given a number of available trucks and the time taken to transport the debris from a building to the Temporary Debris Management Site (TDM), the flow chart of the debris DES simulation is shown in figure 2.9.

the amount of debris that can be collected is given as:

$$Debris/Truck = \frac{\sum T_{debris} + t_{travel} + t_{unload}}{t_{work}} \tag{2.15}$$

where:

- $\sum T_{debris}$ is the sum of debris clearance time components, considering there can be structural, non-structural and other debris.
- $t_{travel} + t_{unload} + t_{work}$ are the travelling, unloading and working hours. Working hours are computed per decision step.

Considering traffic performance, The NRI thus captures the interdependency of traffic performance with building debris and bridge damage. Sediek et al. also propose the Network Performance Index (NPI):

$$NPI = \int_0^{T_{NP}} (1 - MTTR(t))dt \tag{2.16}$$

It represents the area under the Mean Travel Time Ratio ($MTTR$) graph, where T_{NP} is the time required to return to pre-disaster levels, and $MTTR(t)$ is the mean travel time ratio at time t. NPI is noted as similar to Yang et al.'s Resilience Loss metric in terms of the general mathematical approach. Both

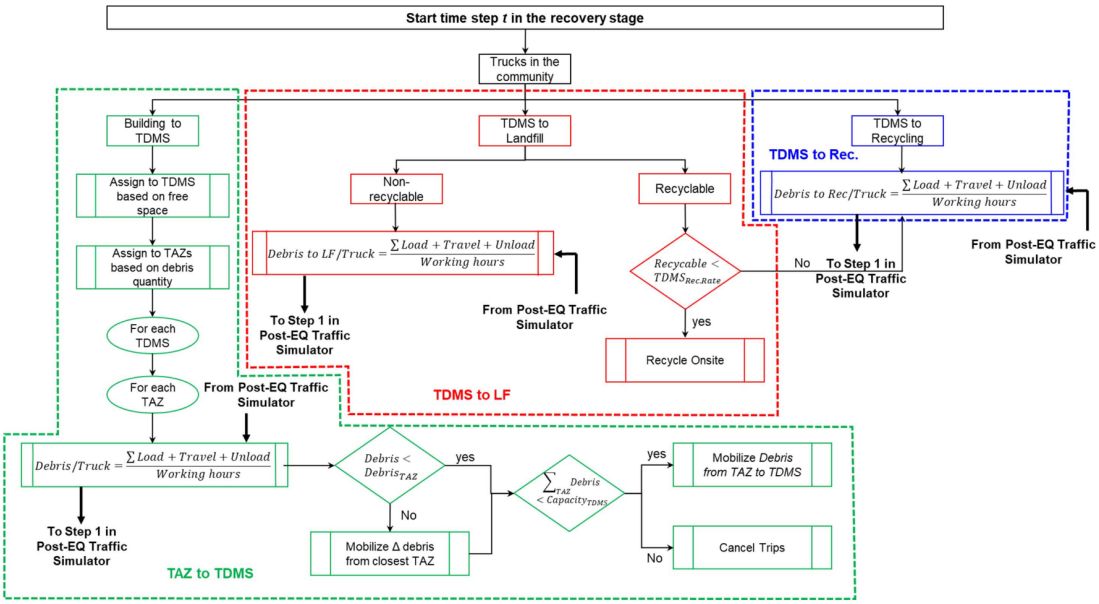


Figure 2.9: Debris DES Simulation from Sediek, El-Tawil, and McCormick [82]. Given one of three possible sites for the debris to be taken (TDMS, Landfill, Recycling), the simulatore defines several rules for the amount of debris that is transported, how much of the debris is recyclable and the effect on traffic.

approaches measure the index of performance as the finite-time interval of a system's functionality, $Q(t)$. The only difference is that Sediek et al. use a cost-based variable (MTT) rather than a reward-based $Q(t)$ variable like seen in Yang et al, [96].

Regarding healthcare functionality, Sediek et al. [82] define three indices: Hospital Use Index (HUI), In-community Mobilization Index (IMI), and Waiting for Admission Index (WAI), all being areas under healthcare performance-loss to time curves. The three metrics use absolute functionality values, like the number of injuries still needing to be treated. For instance, the HUI is given as:

$$\%HUI = \frac{\int_0^{T_H} \hat{I}_{TRE}(t)}{T_H} \times 100 \tag{2.17}$$

where $\hat{I}_{TRE}(t)$ is the normalised value of the number of injuries currently being treated in hospitals against the capacity of the hospitals. Likewise, WAI uses the number of injuries waiting for admission against the hospital capacity. The computation of the healthcare performance indices is made using route-specific performance of trips to and back from the hospital as they are being affected by the load posed on the hospitals. Notably, this is possible at high fidelity since the authors use a dynamic time simulation with shorter early recovery durations. Both healthcare and traffic Discrete Event Simulations (DES) run at a time-step of seconds during the earthquake, $2h$ in the first 30 days after the earthquake, and only for the traffic network, 10 days after the first 30 days. The authors employ the use of traffic assignment zones within the network in order to generate realistic $O-D$ matrices.

Sharma et al. present a resilience framework that can work across scales and takes into account spatial and temporal variation in the recovery phase [84]. They define rigorous mathematical formulations for spatial and temportal resilience and apply it in the case study of component-level repairs of EPN sub-stations. They define the effective area of a repair action as recovery zones which are computed from tributary areas of sub-stations. They define two resilience metrics, both of which defined as *Temporal* versions as well:

- Center of Resilience, (ρ_Q),
- Resilience Bandwidth, (χ_Q).

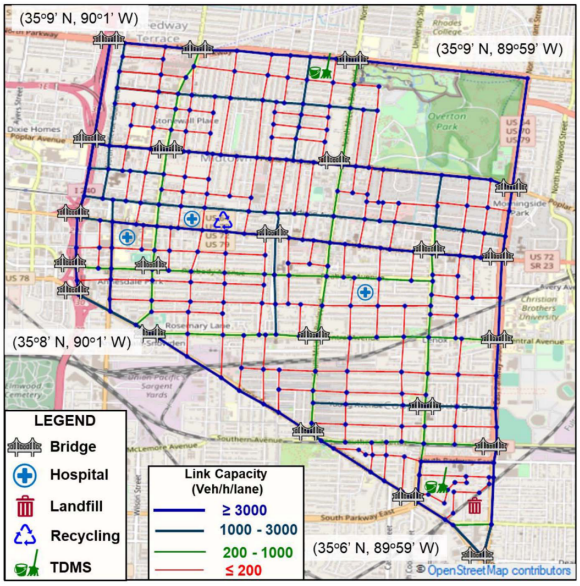


Figure 2.10: The environment used by Seidk et al. [82]. It includes 3 hospitals, 1 debris management site, 1 landfill and 1 recycling facility. Their approach is simulated before, during and after and earthquake; they model traffic using dynamic post-earthquake $O - D$ matrices considering the increase in hospital trips

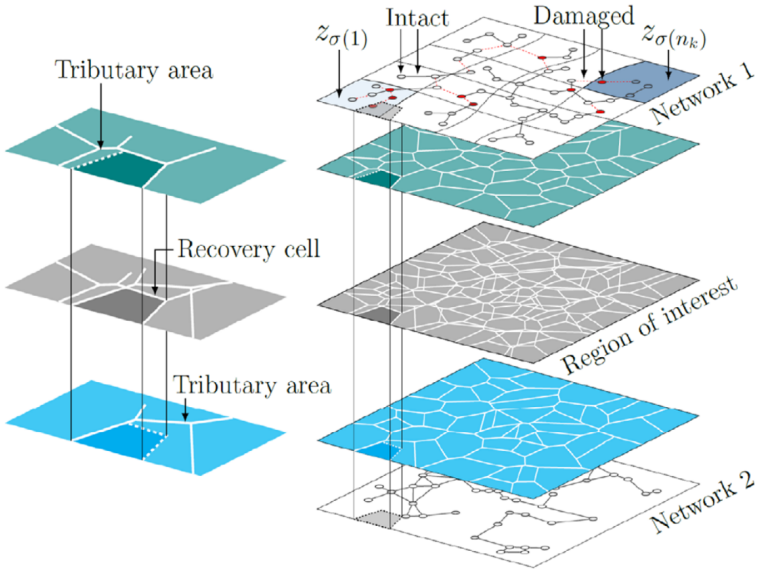
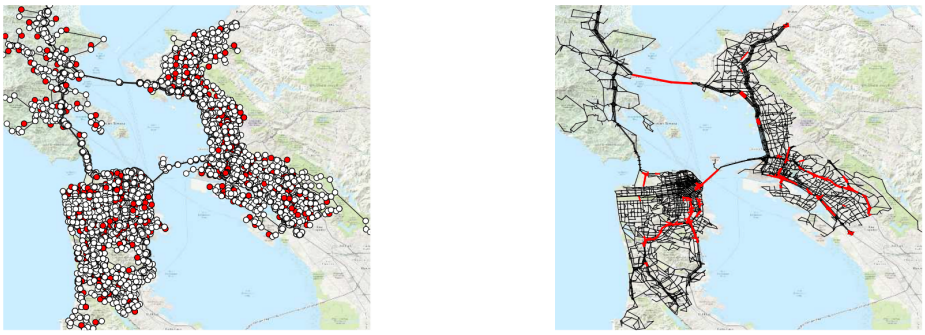


Figure 2.11: Region of Interest partitioning of interdependent infrastructures into recovery zones, Sharma et al. [84]

As defined in their study, temporal metrics are associated with the instantaneous resilience at each timestep. The exact definitions of the two metrics are not given. However, more generally, the center of resilience considers the residual post-disaster performance combined with the total recovery duration [84]. Resilience bandwidth is a measure of the spread of recovery in time, i.e the concentration of effective recovery in time. The researchers then solve post-disaster repair scheduling as a multi-objective optimisation problem, which is conducted using linear optimised paper flow solver using the Python package $PyPSA$. $PyPSA$ is a tool specifically developed for power system analysis [17].

Gomez and Baker employ an optimisation-based approach for coupled pre- and post-earthquake decision making, [43]. They look at the time frame before, but also after an earthquake concerning a large-scale transportation bridge and road network in San Francisco. The decision making problem is whether to pre-emptively retrofit structures before an earthquake or strategise for optimal post-earthquake repair. They conduct spatially correlated scenario-based probabilistic seismic hazard assessment for the San Francisco Bay area. The network is shown in figure 2.12.



(a) (b)

Figure 2.12: an Francisco/Oakland network: (a) nodes in origin-destination pairs (in red); and (b) bridges considered in the optimization problem (in red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article. Extracted from Gomez and Baker [43])

The decision problem is to minimise the investments needed, given the expected consequences generated from seismic hazard assessment:

$$min(investments + E(consequence)) \tag{2.18}$$

The consequences the authors measure are related to retrofit and repair actions and include indirect consequences under a given scenario, monetary costs of retrofit and repair actions and travel time [43]. The master objective function is then defined as:

$$min\left[\sum_{a\in G} c_a x_a + \sum_{\xi\in\Xi} r_\xi \theta_\xi\right] \tag{2.19}$$

where:

- where ξ is a damage scenario in the scenario set Ξ,
- a is a traffic link or bridge in transportation network G,
- c_a is the cost of a retrofit action on a traffic link or bridge,
- x_a is a binary decision variable, where $x_a = 1$ indicates a retrofit action on arc a,
- θ_ξ is an artificial variable used to estimate the value of the function iteratively by means of constraints.

They use the *Gurobi* solver in *Python* 2.7 and show results of optimal decision-making policies for several intervention scenarios of $N - \%$ of the bridges, random policy, and *fix-all* policy. Their results are shown in figure 2.13

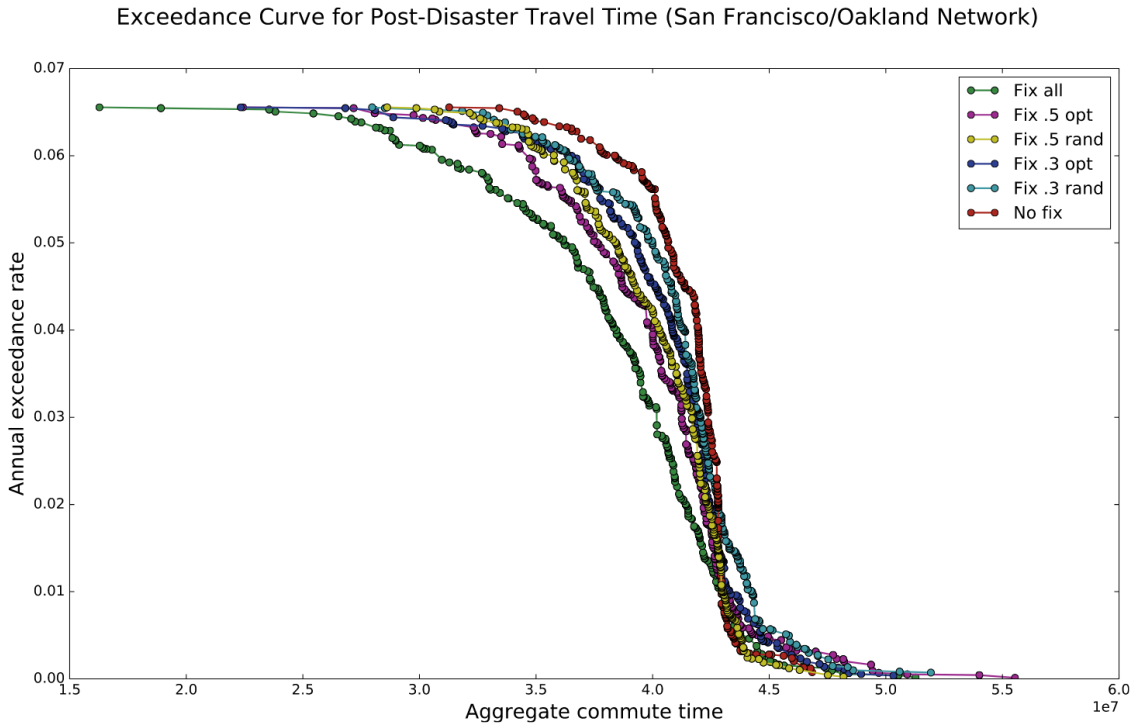


Figure 2.13: Exceedance curves for aggregate users' travel time for the network: without intervention (red), with full intervention (green), with optimal intervention of [jT of bridges (blue), with optimal intervention of]jT of bridges (purple), with greedy randomized intervention of [jT of bridges (cyan), and with greedy randomized intervention of]jT of bridges (yellow). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.). Extracted from Gomez and Baker [43]

Kammouh et al. present a multi-system optimisation approach for interdependent infrastructure management by introducng the 3C concept, integrating multiple infrastructure networks and stakeholders, [52]. The 3C concept is a complex system-based methodology, focusing on the effect of individual and clustered actions not on individual components but on the totality of the system itself [52]. The concept is divided into three phases, the first two are Centralisation and Clustering:

- **Centralisation**: Intervention actions are either *central* or *non-central*. Central actions are pre-emptively planned intervention actions, whose time of occurrence cannot be delayed or advanced. They are implemented following a maintenance-planning approach where an intervention action on the same component is time-dependent. Non-central interventions occur during the downtime due to central interventions and are component-condition-dependant [52]
- **Clustering**: Non-central intervention actions are clustered with central intervention actions such that they respect some overlying constraints, such as the planned interval between intervention actions on the same component type, [52].

The third phase is to calculate the intervention program such that it meets the optimisation objective, which in the authors' case is minimising global cost. The novely of this approach is that it considers an environment that can effectively simulate the decision-making process across a tree of operator - network - asset - object - intervention dependencies. This tree is shown in figure 2.15.

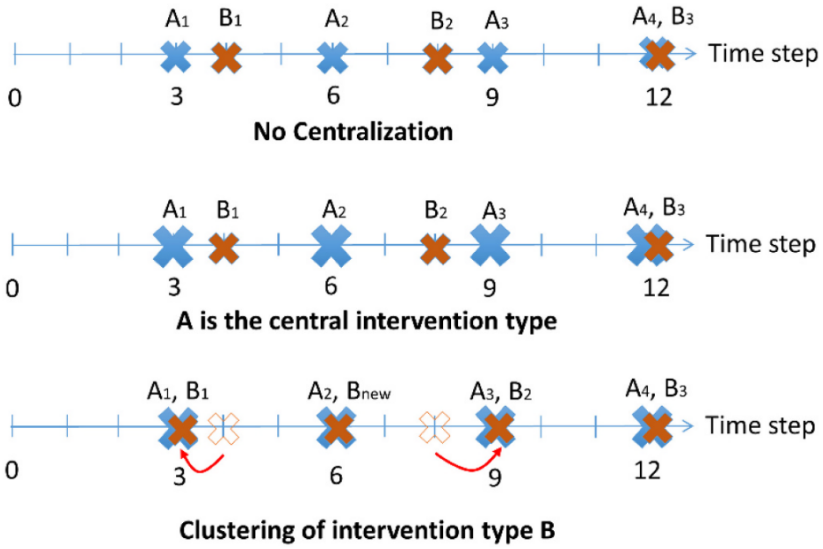


Figure 2.14: Clustered Non-Central and Central actions involve the temporal offset of non-centralised actions to a centralised planned action. Extracted from Kammouh et al. [52]

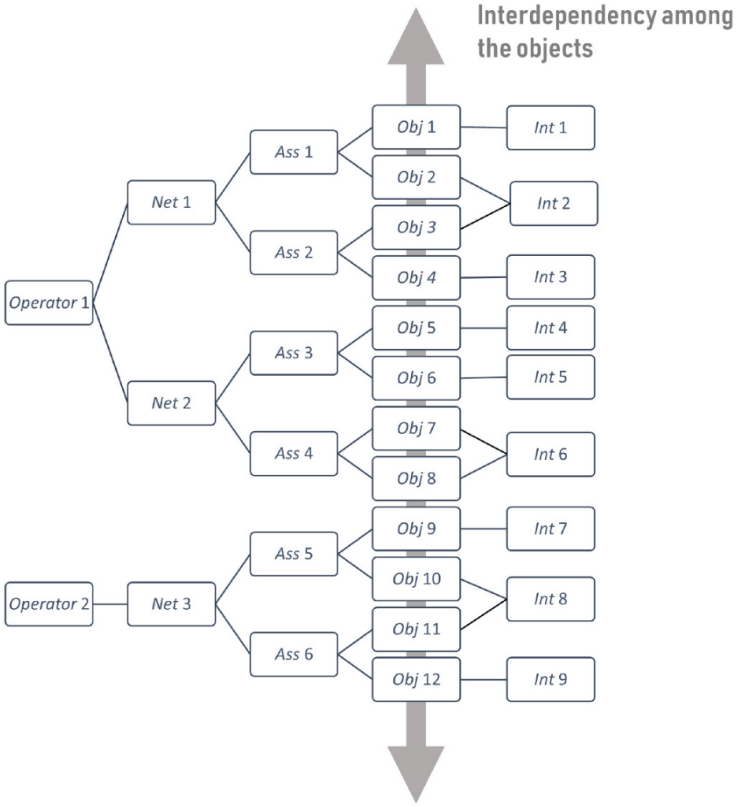


Figure 2.15: Relationships between operators, infrastructure networks (Net), Assets (Ass), objects (Obj), and intervention types (Int). Extracted from Kammouh et al. [52]

They define their objective function as:

$$\min(f_1 + f_2) \tag{4}$$

where f_1 is the total cost of interventions, given by:

$$f_1 = \sum_{t=1}^{T} c_k M = \sum_{t=1}^{T} c_k [m_{k,t}] \tag{2.20}$$

where $c_k \in \mathbb{R}^+$ is the cost of performing an intervention of type k with $k = 1, 2, \ldots, K$, $K \in \mathbb{N}^+$ being the number of intervention types, $T \in \mathbb{N}^+$ is the number of time steps considered in the analysis, and $m_{k,t} \in \{0, 1\}$ are the components of M indicating at which time steps each intervention type is conducted over the total time of analysis. It is assumed that each intervention is entirely performed within a time interval.

And f_2 is the total service unavailability cost caused by the interventions:

$$f_2 = \sum_{t=1}^{T} c_{li}\, \delta(I_t \times R \times M) = \sum_{t=1}^{T} c_{li}\, \delta\left([I_{ij}]_t [r_{i,k}][m_{k,t}]\right) \tag{2.21}$$

The optimisation problem is subject to two constraints considering a discrete time-step g modelled as a positive natural number. The first is that that any two successive intervention actions k are to have a time-interval of at least $G_{min,k} \in \mathbb{N}^+$. The second constraint restrict any two successive intervention actions of type k to have a time-interval larger than $G_{max,k} \in \mathbb{N}^+$. The authors test their method on a network with 12 components as shown in figure 2.16.

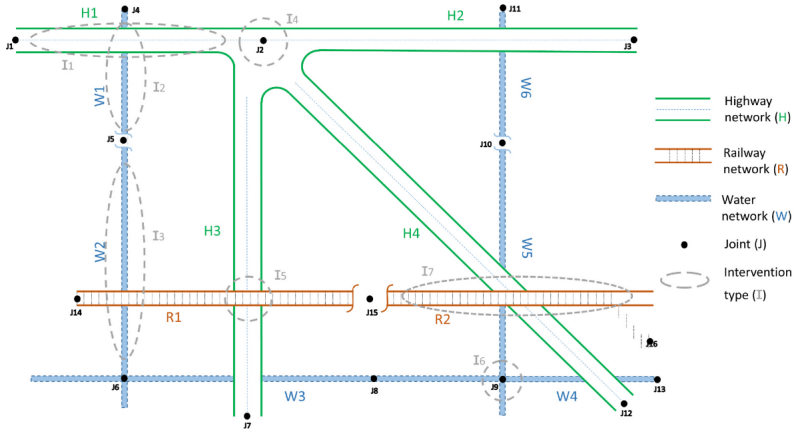


Figure 2.16: Infrastructure networks with preventive interventions to be planned. Extracted from Kammouh et al. [52]

The authors solve the optimisation problem using *Matlab* and achieve approximately 15% lower costs than the baseline policy of each operator individually planning their own intervention. The results are shown in figure 2.17.

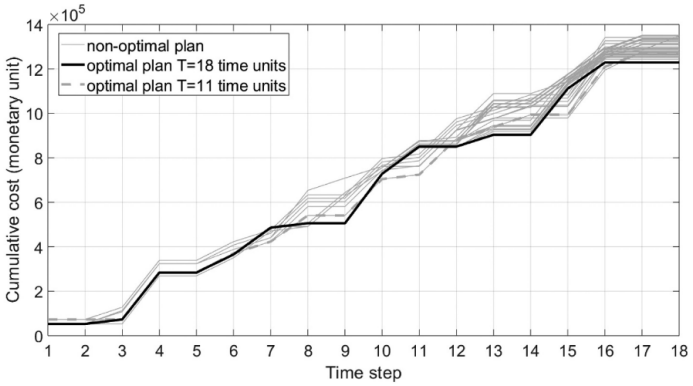


Figure 2.17: Cost comparison between optimal and sub-optimal intervention programs. Extracted from Kammouh et al. [52]

2.5. State of the Art

This section presents state-of-the-art approaches for infrastructure decision making. Specifically, it looks at three key areas of research:

- DRL for deteriorating infrastructure decision making.
- DRL for post-earthquake recovery infrastructure decision making.

The aim in presenting research on the three topics is to provide an overview of how DRL has been used both in deteriorating but also post-disaster environments. Additionally, rule-based approaches that focus solely on post-earthquake recovery are presented as they provide highly complex and effective algorithms for baseline approaches against which DRL can be compared. Rule-based approaches to deteriorating infrastructure decision making that don't focus on post-earthquake recovery are of less interest as the specific rules don't generalise well to the research proposed in this thesis.

The use of DRL for infrastructure decision making is seen to principally focus on modelling infrastructure environments stochastically, such that the observed state of components is uncertain at initialisation as well as during simulation. Considering that the principal damage model is deterioration, the incurred damage on a components happens gradually and is associated with high uncertainty; therefore, the modelling of component states as stochastic is crucial as the effects are long-term and rely heavily on field or remote inspection. Conversely, this thesis as well as other work that focuses solely on post-earthquake recovery that does not include deterioration effects usually employs fully-observable component damage or repair states. Even though a stochastic model is almost always more realistic, the action of inspection is less crucial in post-earthquake scenarios as the damage is usually more apparent and more severe. This is the key difference between work that focuses on deteriorating effects when compared to work that looks at severe disruptions such as earthquakes. In essence, the damage model informs the dynamics of the decision making model.

2.6. Supporting Materials

In modelling decision-making environments, particularly in multi-agent settings under uncertainty, it is important to consider how information is represented and shared among agents. A decision making scenario involving two or more agents is described as a *game*, [66]. Frameworks for mathematically describing and modelling games come in many forms, but usually Markov Decision Processes ($MDPs$) are used. $MDPS$ can be modelled stochastically such that the perceived state, or observation of each agent is stochastic and might also be non-stationary across a trajectory. These are described as partially observable $MDPs$, or $POMDPs$. They are a case of Partially Observable Stochastic Games ($POSGs$), [4].

2.6.1. Markov Decision Processes (MDPs)

An MDP is a mathematical framework used to model decision-making in environments that are fully observable. An MDP is defined by the tuple $\langle \mathcal{S}, \mathcal{A}, P, R, \gamma \rangle$, where:

- $\mathcal{S}$ is a finite set of states,
- $\mathcal{A}$ is a finite set of actions,
- $P : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow [0, 1]$ is the transition probability function
- $R : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ is the reward function,
- $\gamma \in [0, 1]$ is the discount factor.

In the context of infrastructure management the state is usually some deterioration condition of a combination of many deterioration conditions. For instance, the state of a component could be the tuple of its repair time and damage state, or it could just be its damage state. A realisation of acting through an MDP happens by beginning at some starting state s_0, acting on the MDP through transition states s and reaching an absorbing state $\overline{\mathcal{S}}$, or truncating to a state $\mathcal{S}_{\mathcal{H}}$, given a time horizon t_H. A realisation τ with a truncation condition is described as the sequence of state-action-reward-transition state items:

$$\tau = (s_0, a_0, r_0, s_1, a_1, r_1 \dots s_{H-1}, a_{H-1}, r_{H-1}, s_H) \tag{2.22}$$

and for an absorbing state:

$$\tau = (s_0, a_0, r_0, s_1, a_1, r_1 \dots s_{T-1}, a_{T-1}, r_{T-1}, \overline{S}, r_T) \tag{2.23}$$

The cumulative discounted reward of a trajectory is described as the *returns* of a trajectory, $G(\tau)$. In the finite horizon case, this is described as:

$$G(\tau) = \sum_{t=0}^{t_H-1} \gamma^t r_t \tag{2.24}$$

The discount factor has two main purposes: (a) bound the returns between negative and positive infinity (assuming that $R \in \mathbb{R}$), (b) provide the agents a sense of *sight*, such that they value current rewards more than future rewards. The *strategy* of a given solver is the policy, π, which is the stochastic or deterministic mapping between states and actions. Stochastic policies are more often used and describe the probability of ttaking action a when in state s:

$$\pi(a \mid s) = P[a_t = a \mid s_t = s] \tag{2.25}$$

The two metrics that are most commonly used during DRL training are the state-value function $V^\pi(s)$ and the state-action-value function, $Q^\pi(s)$. The state-value function describes the *goodness* of being in a state and is action-agnostic, it is defined as the sum of expected future rewards when being in state s and following policy π:

$$V^\pi(s) = \mathbb{E}\left[\sum_{t=0}^{t_H-1} \gamma^t r_t | s_0 = s\right] \tag{2.26}$$

The state-action-value function describes the *goodness* of taking an action a when being in state s and following policy π thereafter:

$$Q^\pi(s_t, a_t) = r(s_t, a_t) + \gamma \mathbb{E}[Q^\pi(s_{t+1}, a_{t+1})] \tag{2.27}$$

2.6.2. Partially Observable Markov Decision Processes (POMDPs)

$POMDPs$ extend $MDPs$ to handle uncertainty in state information. In a $POMDP$, the agent does not have direct access to the true environment state but instead receives observations that provide partial information about the state. A $POMDP$ is defined by the tuple $\langle \mathcal{S}, \mathcal{A}, P, R, \mathcal{O}, O, \gamma \rangle$, where:

- $\mathcal{O}$ is the set of possible observations,
- $O : \mathcal{S} \times \mathcal{A} \times \mathcal{O} \to [0, 1]$ is the observation function, where $O(o|s', a)$ gives the probability of observing o after taking action a and transitioning to state s'.

The key challenge in solving a $POMDP$ lies in maintaining a belief over the possible states, typically represented as a probability distribution $b(s)$ over all $s \in \mathcal{S}$. Agents then choose actions based on their belief states rather than true states. The value function in a $POMDP$ then becomes:

$$V(b_t) = \max_{a \in \mathcal{A}} \left\{ \sum_{s \in \mathcal{S}} b_t(s) r(s, a) + \gamma \sum_{o \in \mathcal{O}} p(o \mid b_t, a) V(b_{t+1}) \right\} \tag{2.28}$$

where:

- $b_t(s)$ is the belief (probability) that the system is in state s at time t.
- $r(s, a)$ is the immediate reward received for taking action a in state s.
- γ is the discount factor.
- $p(o \mid b_t, a)$ is the probability of observing o given belief b_t and action a.

2.7. DRL for Infrastructure Decision Making

Andriotis and Papakonstantinou present a framework for managing both abstract and applied infrastructure environments with large action and observation spaces [7]. They address the issue that arises when using numerical optimisers such as linear and mixed integer programming for infrastructure environments with many components, stochastic state dynamics and large action spaces. These simulators are seen in the work presented above where an objective function is directly optimised, without the use DRL. The authors navigate this by using DRL, specifically Deep Centralised Multi-Agent Actor Critic ($DCMAC$) solvers. DCMAC is used as the solver of the three decision making environments presented; they model the environments as Partially Observable Markov Decision Processes ($POMDPs$) which allow for the effective simulation of decision-making trajectories.

The authors test three different environments: (a) An abstract parallel-series system with 5 components modelled as a $POMDP$ with stationary state-transition probability matrices, (b) A *k-out-of-n* system with 10 components, 2 failure modes and non-stationary state-transition probability matrices, (c) A steel-truss bridge system with 25 components and non-stationary state-transition probability matrices, which is modelled as a numerical FEA model. These are shown in

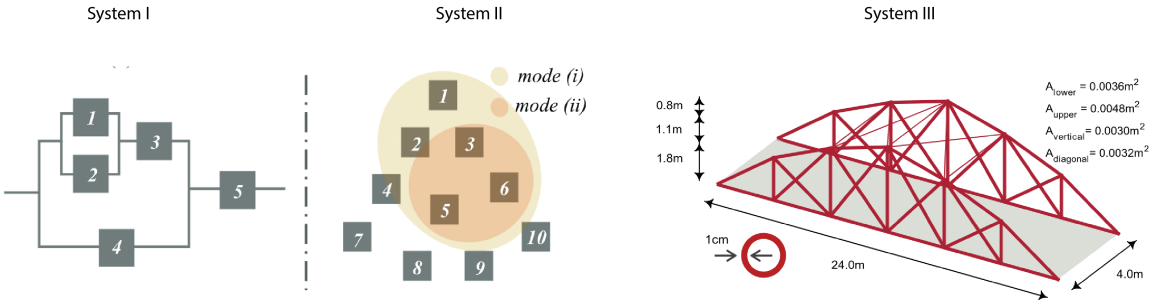


Figure 2.18: System I: 5 components, stationary dynamics, parallel-series. System II: 10 Components, non-stationary dynamics, k-out-n with 2 failure modes. System II: Steel truss bridge, non-stationary dynamics, 25 components. Extracted from Andriotis and Papakonstantinou [7]

The authors principally test $DCMAC$ (Deep Centralised Multi-Agent Actor Critic), which is an on-policy actor-critic DRL method for multi-agent environments, [59]. $DCMAC$ is an example of a Centralised Training with Centralised Execution ($CTCE$) method, $DCMAC$ and other $CTCE$ methods are shown to outpeform other approaches but face challenges in large environments. Their work is seminal in the field as it is the first use of $DCMAC$ for multi-component, partially observable abstract and applied engineering infrastructure systems. $DCMAC$ employs two learning networks with centralised learning, i.e the number of networks does not grow with the number of agents, but the input features to the networks is the joint trajectories of all agents. The actor network directly learns the stochastic policy $\pi(a_t|s_t)$. This is evaluated by the critic network through the advantage function $A^\pi(a_t|s_t)$, and through back propagation the networks' parameters are tuned. This architecture is illustrated in figure 2.19.

The advantage function $A^\pi(a_t|s_t)$ when following policy π is the difference between the state-action-value function $Q^\pi(s_t|a_t)$ and the state-value function $V^\pi(s_t)$:

$$A^\pi(s_t|a_t) = Q^\pi(s_t|a_t) - V^\pi(s_t) \tag{2.29}$$

The authors test the use of $DCMAC$ against several iterations of 4 different baseline policies:

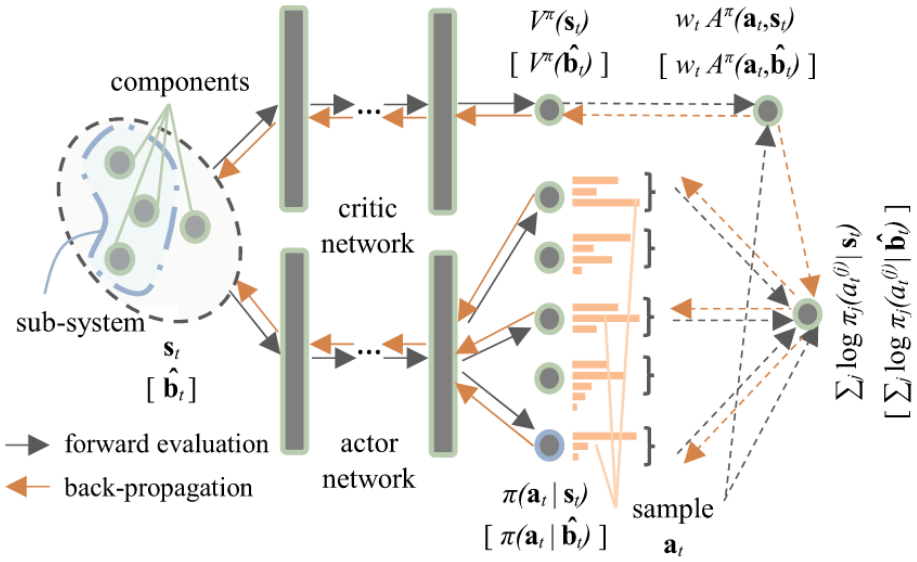


Figure 2.19: Deep Centralized Multi-agent Actor Critic (DCMAC) architecture. Forward pass for function evaluations (black links) and weighted advantage back-propagation for training (red links). Dashed lines represent operations and dependencies that do not involve deep network parameters. Extracted from Andriotis and Papakonstantinou [7]

.

- **Condition-Based Maintenance I** (CBM-I): Components reaching or exceeding severe damage state (3) are replaced.
- **Condition-Based Maintenance I** (CBM-II): Components with minor or no damage state are assigned *do-nothing* actions. Minor repair is assigned to severe damage state components and replace failure is assigned for all probabilities $p = [0.9, 0.8, .0.7]$, making three different CBM-II versions.
- **Time-Condition-Based Maintenance I** (TCBM-I): Similar to CBM-I but with the inclusion of *major-repair* and *replacement* actions.
- **Time-Conditioned-Based Maintenance II**: (TCBM-II): Simillar to CBM-II but with the inclusion of *major-repair* and *replacement* actions.

The authors model the environments with a discrete decision time-step of 1 year. In System I, the authors tested $DCMAC$ and DQN which all converged to the same solution, as can be seen in figure 2.20.

Regarding System-II, $DCMAC$ with a probability of replacement, $p = 1.0$ outperformed all other policies as can be seen in figure 2.21.

Leroy et al. look into inspection-maintenance planning using MARL and produced a suite of environments and agents, [57]. They model four different environments:

- **k-out-of-n**: A system of n components which fails if $(n - k + 1)$ components fail.
- **Offshore wind farm**
- **Correlated k-out-of-n**: an environment that is partially observable and inspections on one component might affect the observed state of another component.
- **Campaign Cost Environment**:

The researchers develop Multi-Agent systems which are similar to single agent RL but can include behaviour such as competition, cooperation or critique. The MARL experiment setup of the environments can include 2 to 100 agents like seen below:

The actions are defined as *do-nothing*, *inspect* and *repair*. Inspect and Repair actions have significant costs associated with them which are included in the $Dec - POMDP$ (Deconstructed Partially Observable Markov Decision Process). A $Dec - POMDP$ is simillar to a $POMDP$, but includes the joint observation

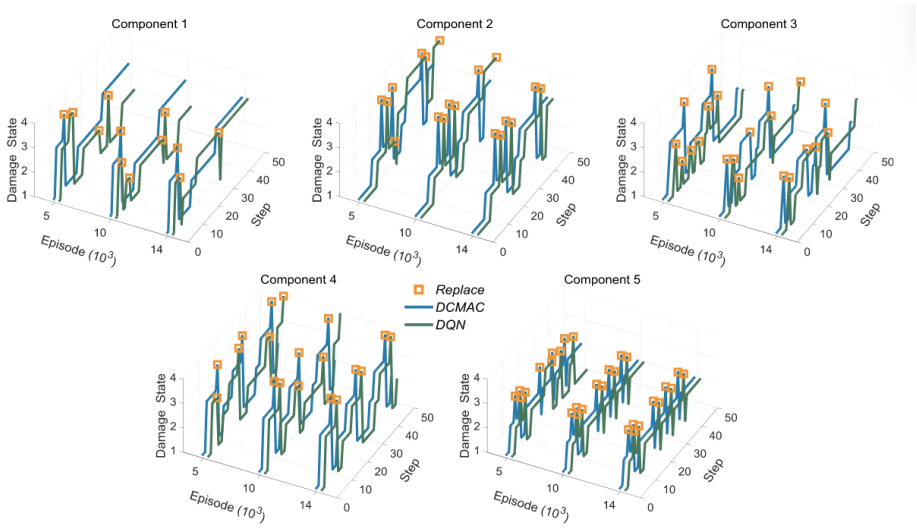


Figure 2.20: Policy realizations for all System I components at three different training episodes, based on DCMAC and DQN solutions. All component policies converge to the exact solution for both methods after 10 thousand episodes, expect for components 1 and 3 for the DQN solution. Extracted from [7]

.

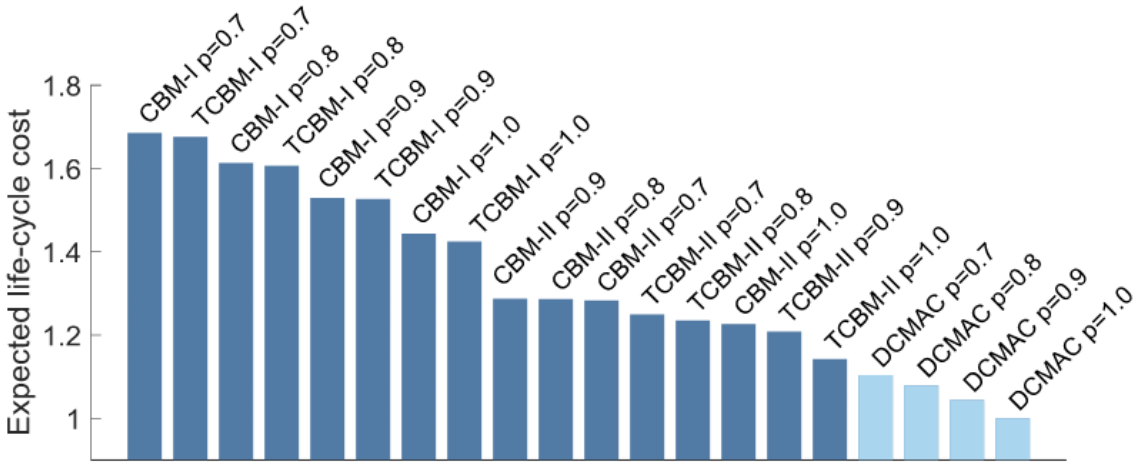
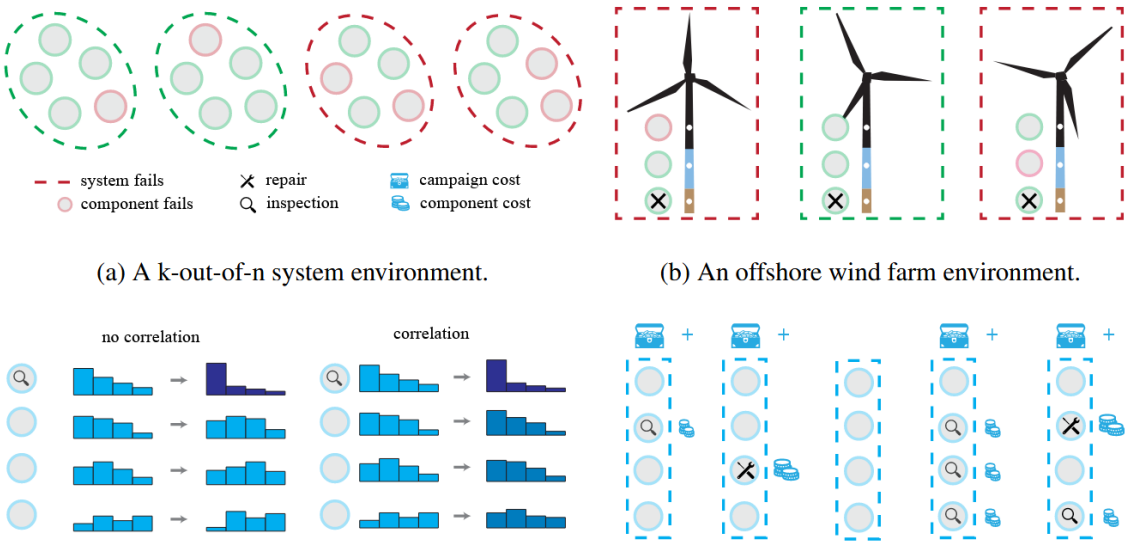


Figure 2.21: Expected life-cycle cost estimates of DCMAC solutions and baseline policies, for System II, for different observability accuracies. DCMAC outperforms all optimized baselines even when these operate under better observability. Extracted from Andriotis and Papakonstantinou [7]

.

Table 2.2: Number of agents specified in IMP-MARL, [57]

IMP environments	Number of agents
k-out-of-n system	3, 5, 10, 50, 100
Correlated k-out-of-n system	3, 5, 10, 50, 100
Offshore wind farm	2, 4, 10, 50, 100



(a) A k-out-of-n system environment.

(b) An offshore wind farm environment.

Figure 2.22: Visual representation of available IMP-MARL environment sets and options. In 2a, a 4-out-of-5 system fails if 2 or more components fail. In 2b, a wind turbine fails if any constituent component fails. In 2c, when the environment is under deterioration correlation, the information collected by inspecting one component also influences uninspected components. In 2d campaign cost environments, a global cost is incurred if any component is inspected and/or repaired plus a surplus per inspected/repaired component. Extracted from Leroy et al. [57]

space O and observation probability function $\mathcal{Z}$, [4]. Reward is given as the negative cost of following an action R_{ins}, R_{rep}, R_{no} in addition to a system failure risk R_f. The authors use two reward models, first a campaign reward which is incurred if at least one component is inspected or repaired, which is added together with the R_{ins} and R_{rep} per selected / repaired component. Second, they use no campaign reward. The experiment is tested over the finite-horizon case and agents act to maximise their expected sum of discounted rewards. The researchers test 7 MARL approaches, one fully centralised, one fully decentralised and 5 $CTDE$ (Centralised Training for Decentralised Execution) methods, [57]. DQN (Deep Q Learning) is used for the fully centralised approach, where all agents receive the same observations about the state of the environment and are controller by one controller. IQL (Implicit Q-Learning) where each agent is independently trained.

The five $CTDE$ methods are not examples of off-policy learning. Specifically, they use $QMIX$, $QVMix$, $QPLEX$, where the value function is factorised to each agent during training; this way agents can choose different actions that follow the same value functions, [57]. The 4 techniques for $CTDE$ are not explored fully, but more generally they rely on Q-learning, specifically the Q Action-Value function. In essence, Q-learning is concerned with the value associated with taking any action from a particular state and only then following the optimal policy. The researchers explain that the results of using MARL were superior to heuristic methods as they yield higher expected sum of total rewards. However, performance varies across RL methods; for example, when dealing with a high number of agents, the campaign cost reward model yielded better results. This is because of the explicit incentive of agent cooperation in the campaign reward. The results are summarised in figure 2.23

The authors show that $CTDE$ methods outperform IQL. Fully centralised approaches such as DQN perform better than heuristic policies but are outperformed by $CTDE$ methods. They point out that

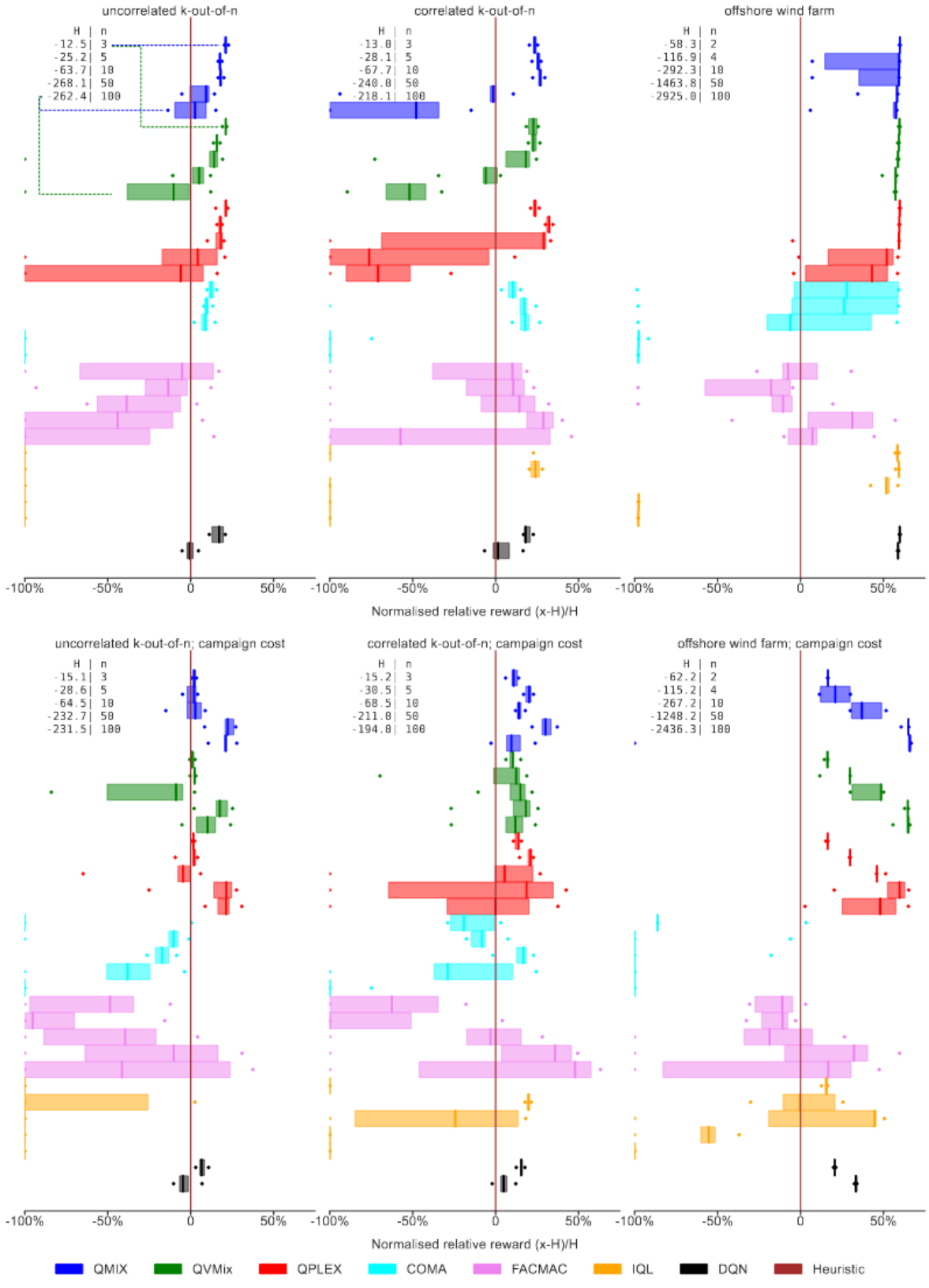


Figure 2.23: Performance reached by MARL methods in terms of normalised discounted rewards with respect to expert-based heuristic policies in all IMP environments, H referring to the heuristics result. Every boxplot gathers the best policies from each of 10 executed training realisations, indicating the 25th-75th percentile range, median, minimum, and maximum obtained results. The coloured boxplots are grouped per method, vertically arranging environments with an increasing number of n agents, as indicated in the top-left legend boxes. Note that the results are clipped at −100%. Extracted from Leroy et al. [57]

further work is needed to make training DRL agents more stable and increase the complexity of the engineering environments tested.

Bhustali and Andriotis test 7 DRL methods for inspection maintenance planning of deteriorating infrastructure environments, [12]. They test a *k-out-of-n* system with 5 components, 4 damage states and 3 actions (*do-nothing*, *repair*, *inspect*). They test 4 *k-out-of-n* failure modes for $k = [1, 2, 3, 4]$. They test the performance of DRL against a baseline policy of time-periodic inspections with condition-based maintenance ($TPI - CBM$). This policy involves inspecting components at fixed time-intervals Δt_{insp}, and based on the observations o_{insp}, the policy aims to minimise the expected discount sum of future costs (negative rewards). The 7 frameworks they test are as follows:

Paradigm	*Mathematical Framework*	*Algorithm*	*Observation*	*Action*	*Critic*	*Actor*
CTCE	POMDP	**JAC** (SADRL)	Joint	Joint	Centralized	Shared
	MPOMDP	**DCMAC**		Factored		Shared
		DDMAC		Factored		Separate
CTDE	Dec-POMDP	**IACC**	Independent	Independent	Centralized	Separate
		IACC-PS		Independent		Shared
DTDE		**IAC**	Independent	Independent	Decentralized	Separate
		IAC-PS		Independent		Shared

Figure 2.24: The 7 $MARL$ methods tested in Bhustali and Andriotis [12]

The authors show that the results show that most DRL approaches can outperform heuristic methods as seen in figure 2.25. They show that $CTCE$ methods likeJAC and $DCMAC$ outperform all other $MARL$ methods but are also susceptible to sub-optimal policies arising from the issue of exploration induced by the joint action spaces, [12]. They show that $DTDE$ and $DTDE$ approaches scale better than $CTCE$ approaches but face challenges in training due to the decentralisation of observation and action spaces.

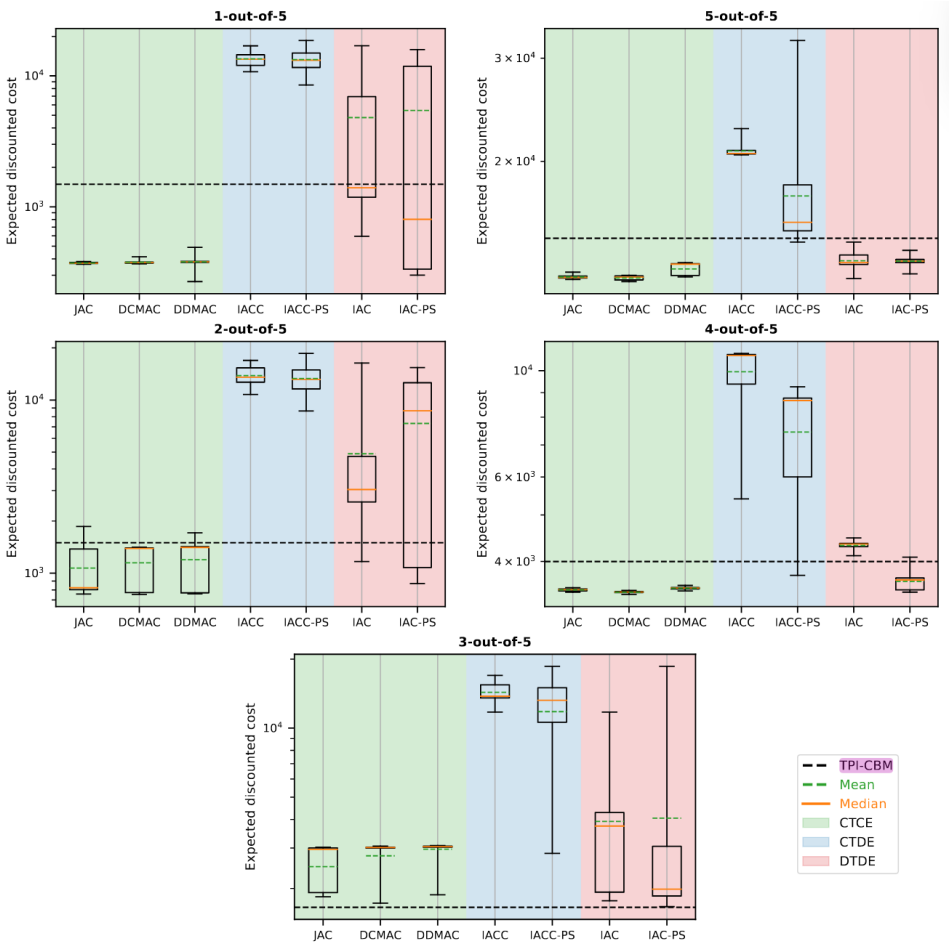


Figure 2.25: Box plots summarizing the performance of the best policies across fifteen training instances for all k-out-of-n settings. The dotted line indicates the TPI-CBM heuristic and the whiskers denote the minimum and maximum values observed. Extracted from Bhustali and Andriotis [12]

2.8. DRL for Post-Earthquake Infrastructure

In this line of research, two key research effort are identified that most well aligns with this research focus. The approaches use Single- or Multi-Agent DRL for post-earthquake recovery of transportation, lifeline or energy infrastructure systems, [96, 81]. More generally, the section begins by indicating some approaches that can be used as baselines, to which DRL can be compared.

- **Rule-Based Methods**: Methods that assign some online or offline ranking on components, and use that to prioritise intervention actions, [65, 58, 67, 49]. Offline ranking does not have access to dynamic information about components such as damage states repair times etc. Most online approaches make use of importance and performance indices in the form of:

$$\text{Importance Index} = \frac{dem(t)}{cap(t)} \tag{2.30}$$

$$\text{Performance Index} = \frac{q(t)}{q^*(t)} \tag{2.31}$$

  where $dem(t)$ and $cap(t)$ are the demand and capacity of a component; $q(t)$ and $q^*(t)$ are the current and nominal functionalities of a component. Demand and capacity can be related to the criticality of infrastructure or some absolute value of performance such as the number traffic trips, number of injuries etc. Methods for computing these can be complex and include rules for bounding their value when critical, or other indicators, are not met. These often do not account for dynamic effects such as interdependencies explicitly. Interdependencies can be easier to describe than use in rank-based decision-making directly. This is because the significance of an interdependency to a network is potentially cascading, making its behaviour difficult to predict using rule-based approaches.
- **Genetic Algorithm**: This is a search algorithm that can be competitive to DRL and was originally developed before major ML breakthroughs. When using the GA, the action component of a trajectory is considered as the *genome*, with step-wise joint actions being *genes*; genomes undergo processes like *mutation* and *crossover* to develop action trajectories, or genome configurations that maximise the objective function, [46]. GA performance is often cursed with finding local optima.

Yang et al. use Deep $MARL$ for post-earthquake repair scheduling of interdependent infrastructure networks, [96]. This is contrasted with Sediek et al who use a heuristic method that integrates traffic modelling, the healthcare network and the interdependency of debris to traffic links, [82].

Yang et al employ the use of one $CTCE$ method, $DCMAC$, and model a fully observable MDP of components of a Water Network (WN) and Electrical Power Network (EPN). They use two actions *repair* or *do-nothing* and measure community functionality and resilience by measuring the derivative effect of WN wells and EPN substation downtime on the community's buildings, which themselves are not included in the MDP. Their work is novel as it uses a graph-based state description of the environment and uses that as the input features to the actor and critic networks, [96]. The graph-based nature of their envirnment formulation is illustrated in figure 2.26.

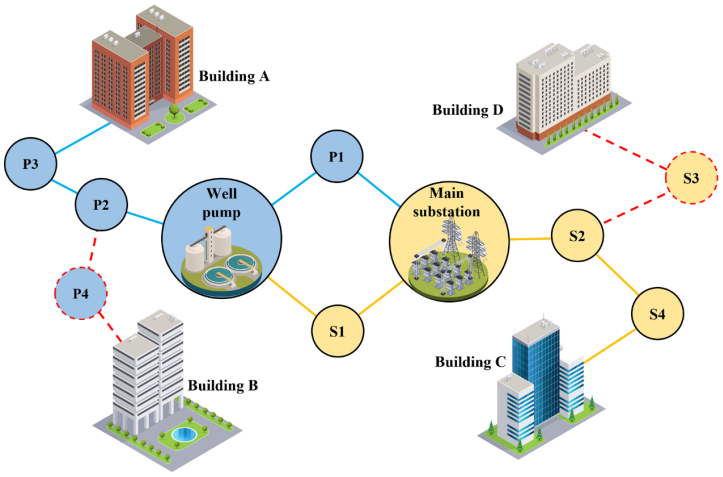


Figure 2.26: The *MDP* components of Yang et al.'s approach are the water wells and EPN substations, but agent reward is given form the cascading effect of WN wells and EPN substations to accessing buildings. Extracted from Yang et al. [96]

The reward function is defined as:

$$R(t) = RL(t) = \int_{t-1}^{t} [1 - Q_{com}(t)]dt \tag{2.32}$$

where:

- $R(t)$ is the common reward of all agents at time t,
- $RL(t)$ is the loss of resilience at time t,
- $Q_{com}(t)$ is the community functionality as time t

Community functionality is defined as:

$$Q_{com}(t) = \frac{\sum_{b \in B} Q_b^{phys}}{\sum_{b \in B} S^b} \tag{2.33}$$

where S^b is the total area of building b and Q_b^{phys} is the physical functionality of building b:

$$Q_b^{phys} = S_b \times (I_b^w \beta_b^w + I_b^p \beta_b^p + I_i^t \beta_b^t) \tag{2.34}$$

where the β's are the importance factors of the water wells, water pipelines and power transmission networks respectively. The I's are the binary performance indicator factors of each building's access to the respective infrastructure. This formulation of community functionality and resilience is highly appropriate for post-earthquake infrastructure environments and is the main anchor point for this thesis's methodology. The authors test their approach against an importance-based algorithm, random policy and simulated annealing and show that *DRL* outperforms all of them especially early on in the recovery process. The key limitations of this work are in that the authors use a deterministic seismic hazard and fragility scenario and the only source of uncertainty is the repair times which are sampled form a distribution. Their results are summarised in figure 2.27

Fan et al. propose a GCN-based Single-Agent *DRL* framework for the recovery of road networks, [31]. The use a resilience-based reward and define resilience as it relates to the distance of points in the network to emergency facilities. Their methodology overview is seen in

Damage to road links results in a lower *reliability* of road links as the authors put it, and by assigning a weight to each intersection in the network, the authors derive an aggregate network performance metric.

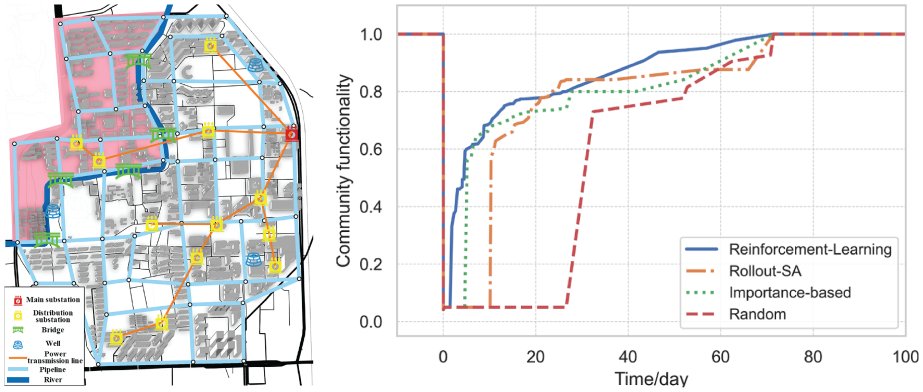


Figure 2.27: Environment and results from Yang et al. MARL is seen to perform better than Importance-Based scheduling, especially in the early recovery phase.

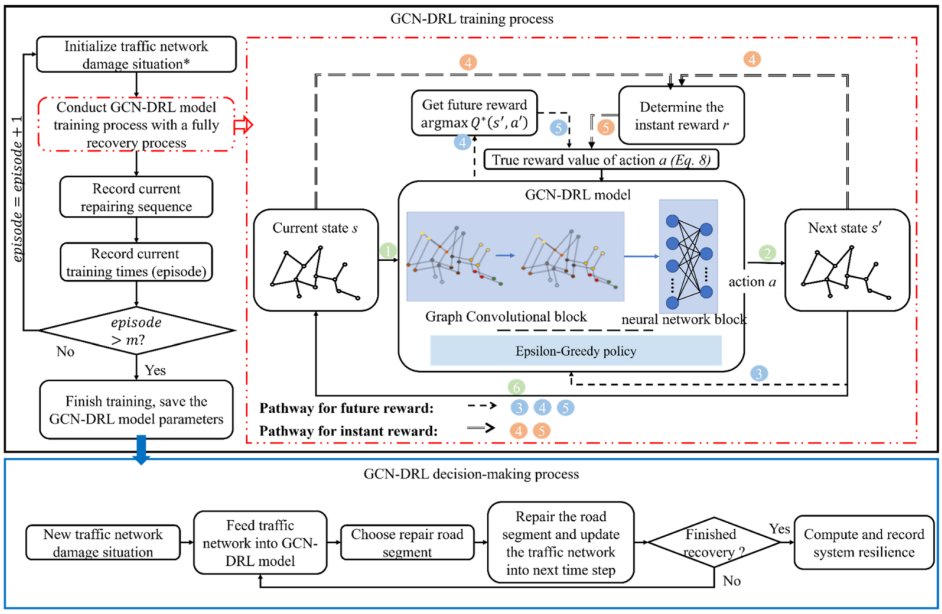


Figure 2.28: The single-agent *GCN-DRL* methodology employed in Fan et al. [31]

It is important to note the authors do not conduct traffic analysis, but rather use the metric explained below as an indicator for network performance. The weight w_i of each intersection i is defined as:

$$w_i = \frac{\Omega_i}{\sum_{j=1}^{n} \Omega_j} \tag{2.35}$$

where:

$$\omega_i = \begin{cases} \frac{1}{\min(D_i)} & \text{if } \min(D_i) > 1\,\text{km} \\ 1 & \text{otherwise} \end{cases} \tag{2.36}$$

where D_i is the set of distances of intersection i to a set of pre-defined emergency response sites. The authors then calculate the average number of reliable independent pathways of each intersection to the set of emergency response sites as:

$$r_i = \frac{1}{n-1} \sum_{j=1, j\neq i}^{n} \sum_{k=1}^{K_{(i),j}} v_k(i,j) R_k(i,j) \tag{2.37}$$

where the reliability $R_k(i,j)$ of path k between nodes i and j is defined as:

$$R_k(i,j) = \prod_{\forall l \in P_k(i,j)} R_l \tag{2.38}$$

and wight of k^{th} independent path through an intersection is defined as:

$$v_k(i,j) = \frac{L_{max(i,j)}}{L_{P_{k(i,j)}} \cdot \sum_{k=1}^{K(i,j)} \left(\frac{L_{max(i,j)}}{L_{P_{k\{(i,j)}}} \right)} \times K_{(i,j)} \tag{2.39}$$

where $K(i,j)$ is the number of all independent paths between nodes i and j. $L_{max(i,j)}$ is the maximum length of all paths and $L_{P_{k(i,j)}}$ is the k^{th} path's length, [31]. The performance of the network is then given as $p(t)$:

$$p(t) = \left[\sum_{i=1}^{n} w_i r_i \right] \times 100 \tag{2.40}$$

Essentially, this approach uses a weighted sum of sub-system functionalities to derive an aggregate system functionality like seen in other resilience approaches, [41, 71]. In this case, the definition of system functionality does not necessarily extend to community functionality but instead narrows down on transportation network functionality. The authors then use the following reward defintion:

$$R_t = \frac{p(t+1) - p(t)}{T_t} \tag{2.41}$$

where T_t is the sum of all road repair times. The authors calculate the future performance using bootstrapping. The authors use four deterministic damage scenarios and four distinct training experiments and predict damage using HAZUS fragility function [31, 35]. The authors train the $SARL$ model for 500 training time-steps and benchmark it against two baseline strategies: (a) Genetic Algorithm (GA), (b) Betweenness-centrality-based scheduling. The GA is not explained in depth; however, it is an optimisation-based approach that aims to simulate biological process of mutation and crossover of genes to choose optimal actions. A common limitation of the GA is that it is usually susceptible to

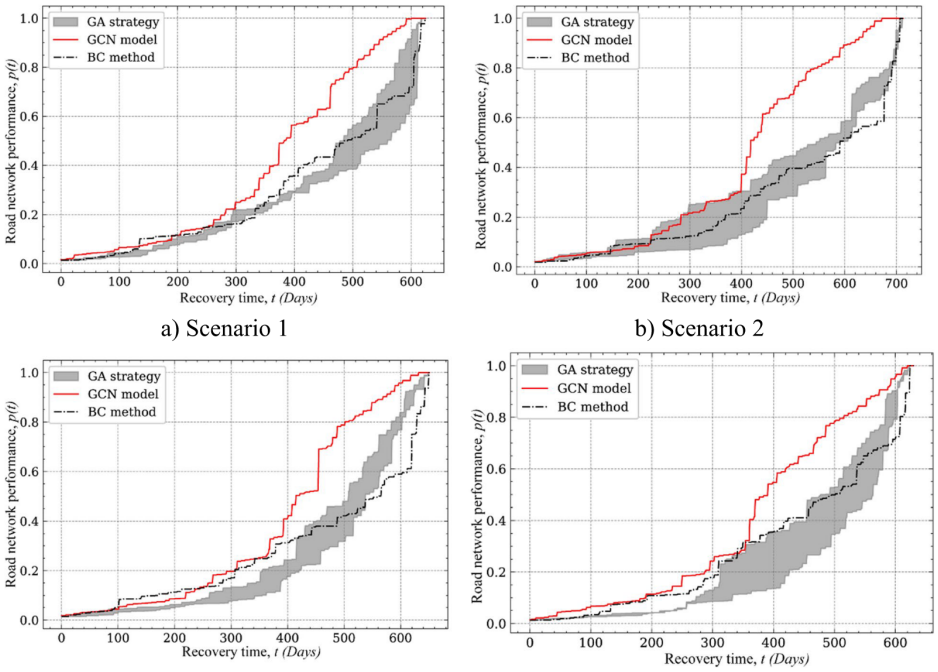


Figure 2.29: Results showing DRL performs better than the two chosen baselines by Fan et al. [31]. It is interesting to note that DRL and the other baselines basicallt perform the same for the first 300 days of repair.

following locally optimal solutions and its overall methodology does not allow for much exploration to be explicitly targeted for. The second baseline uses the graph attribute of betweenness-centrality to rank roads. Betweenness-centrality of a node i is the ratio between the number of of shortest paths that include node i over the number of all shortest paths in the graph. This metrics essentially ranks the *connectivity* of each node. The authors then train a slightly bigger network over 4000 training time steps with a flood hazard. The results for post-earthquake recovery are shown in figure 2.29 show that DRL outperforms the other two baselines across the 4 scenarios.

What is interesting to note here is that DRL outperforms the other policies mostly after 300 days of repair. This could be due to their approach being single-agent. Considering that all actions in a single-agent realisation happen only consequently and never concurrently then at the beginning of a repairing sequence when most roads are highly damaged, the optimal repair action might be easy to find the the GA and the BC baseline. Conversely, as most roads begin to get repaired and their damage is lower, it is harder to predict which one road is of most interest. Aside from the results themselves, it is easy to see the stark difference in performance of $SARL$ and $MARL$. The authors achieve very good performance within 500 timesteps for a network with 136 road segments. This kind of performance for such a large network is not yet possible using $MARL$. However, $SARL$ as a method avoids the concept of the environment being a *game*, as there is only one decision-maker involved. There is no room for cooperation between multiple agents, which can better describe the recovery of a network or community.

3

Methodology

This section presents the various steps and techniques used to answer the principal research question as well as sub-questions. The methods presented are integrated into a simulation environment using the programming language $Python$. The methods for generating the environments and the various losses as well as interdependency modelling are novel and were developed solely by the author. Traffic modelling is conducted using a wrapper of the repository by Matteo Bettini, [10]. The $MARL$ framework used is a very slightly adapted version of Prateek Bhustali's framework for deteriorating infrastructure environments, [11].

The sub-sections that follow begin by describing the general methodology of conducting this experiment, followed by details on data collection and fragility/vulnerability functions and resilience formulation. Consequently, the MDP tuple formulation which describes the environment is described along with the reward function. The algorithms used for MARL are shortly explained. The goal of this experiment is to provide means both for creating a repair-scheduling environment, as well as for solving that environment using $MARL$ and baseline solvers.

3.1. Overview

As outlined above, the proposed methodology aims at being applicable on custom test beds as well existing communities for the general metropolitan United States. The objective is to test the hypothesis that DRL is favourable for post-earthquake repair scheduling policies. This hypothesis is tested using two custom test beds. This section describes the key steps of the methodology; specifically, the sources and nature of the collected data that is used to construct the environment, the environment formulation and MDP and the solvers used to predict post-earthquake repair scheduling policies.

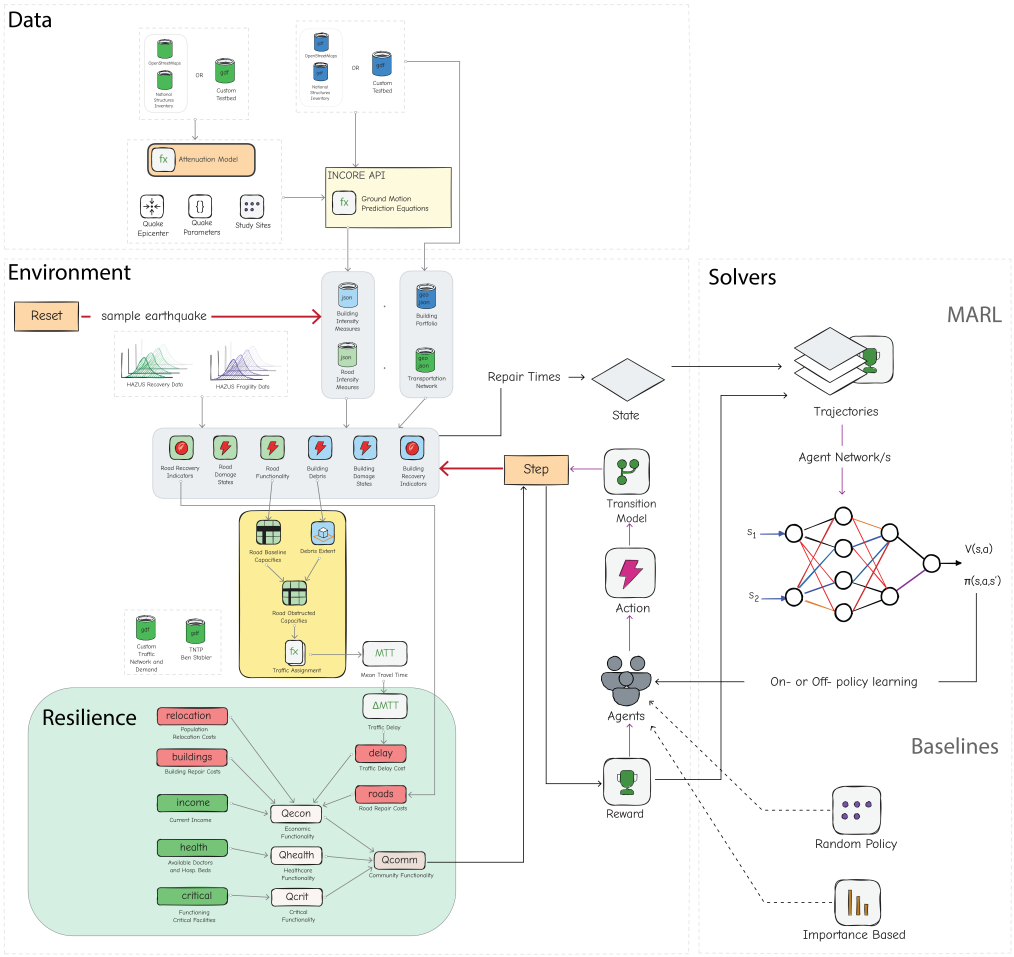
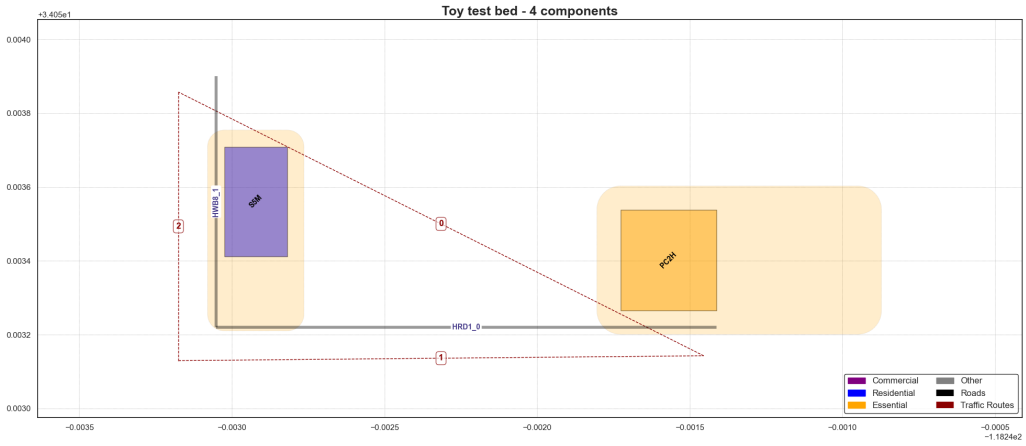


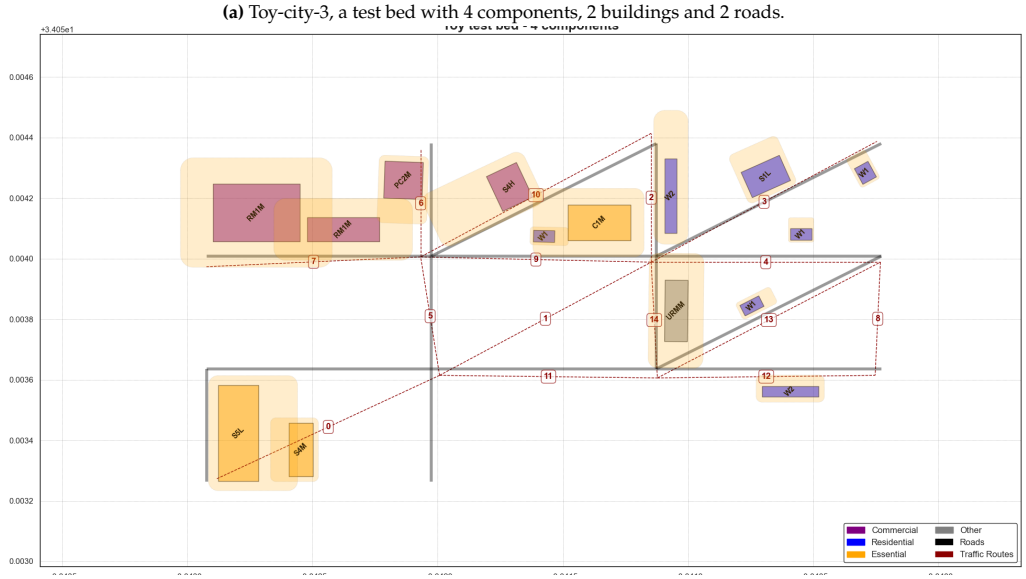
Figure 3.1: Overview of the methodology. Data for constructing an environment is either collected from OSM, NSI and NBI or constructed locally using the appropriate data schemas. An environment is constructed from that data. Two baseline solvers are used to benchmark the performance of $MARL$ algorithms in providing stochastic and performative post-earthquake repair scheduling policies

The experiment conducted in this research is visually outlined in Fig 3.1. Data is produced locally in the form of a building portfolio and transportation network. This data is in the form of geo-referenced tabular data. $IN-CORE$ is used to retrieve Intensity Measure (IM) values at building and road centre points [64]. $IN-CORE$ is a suite of tools for multi-hazard analysis and recovery which is accessible as a python package and is explained in detail in the literature review. $HAZUS's$ earthquake model technical manual [35] provides fragility and recovery functions for building, road and bridge classes. Specifically building structural damage states are predicted using PGA, road damage states are predicted using PGD and bridge damage states are predicted using PGD and $SA_{0.3}$ and $SA_{1.0}$. Non-structural damage to buildings is not considered.

Traffic is modelled using the interdependency of building debris affecting adjacent road capacities and is then used along with other costs and functionalities to compute an aggregate metric for the resilience of the community. This is then used to give a scalar shared reward to all agents in the environment. MARL is benchmarked against two baseline solvers: random policy and importance-based policy. The two test beds considered in the case study are shown in figure 3.2. The smaller one is composed of 4 components, a highway road, a highway bridge, a 6-story residential building and an 8-story hospital. The second environment is composed of 30 components, including two bridges, a hospital and a fire-station. The methods and data used for each of the methods described in this section are explained in detail in the following sections of the methodology.



(a) Toy-city-3, a test bed with 4 components, 2 buildings and 2 roads.



(b) Toy-city-30, a test bed with 30 components, 15 buildings and 15 roads, two of which are bridges.

Figure 3.2: The two test beds used for conducting the underlying *MARL* experiments. Semi-transparent yellow areas surrounding buildings indicate debris, whose direction is sampled stochastically at every realisation following [82]

3.2. Network Interdependency

The set of infrastructure components is denoted by C, such that: $C := C_{\mathcal{B}} \cup C_{\mathcal{R}}$, where $C_{\mathcal{B}}$ is the set of buildings and $C_{\mathcal{R}}$ is the set of roads, with $b \in C_{\mathcal{B}}$ and $r \in C_{\mathcal{R}}$. While not directly considered as repair components, the environment makes use of a traffic network. This network is different from the road network as transportation networks are often modelled with a lower level of detail than a full road network and don't directly match spatially. Roads form a network $\mathcal{G}_R = (\mathcal{V}_R, \mathcal{E}_R)$, where $\mathcal{V}_R$ represents the set of road nodes and $\mathcal{E}_R \subseteq C_R$ represents the road segments connecting them. Mean Travel Time (MTT) is the mean travel times of all trips in the networks and is the basis with which traffic performance is calculated using the traffic network $\mathcal{G}_T = (\mathcal{V}_T, \mathcal{L}_T)$, where $\mathcal{V}_T$ is the set of traffic nodes and $\mathcal{L}_T$ is the set of traffic links. All, or a subset of these nodes are defined as either origin- or destination-nodes and are used for traffic assignment; this is explained in later sections. Each traffic link $\ell \in \mathcal{L}_T$ is mapped to a unique shortest path on $\mathcal{G}_R$, defined by the path $Path(\ell)$ between the closest pair of start and end nodes in $\mathcal{V}_R$ corresponding to the origin and destination of ℓ.

for each $\ell = (o, d) \in \mathcal{L}_T$:

$$Path(\ell) = \arg \min_{p \in \mathcal{P}(v_o, v_d)} \sum_{e \in p} cost_e, \tag{3.1}$$

where:

- $v_o = \arg\min_{v \in \mathcal{V}_R} \text{dist}(o, v)$, closest road node to traffic origin-node
- $v_d = \arg\min_{v \in \mathcal{V}_R} \text{dist}(d, v)$, closest road node to traffic destination node
- $\mathcal{P}(v_o, v_d)$ is the set of all paths from v_o to v_d in $\mathcal{G}_R$, and $cost_e$ is distance.

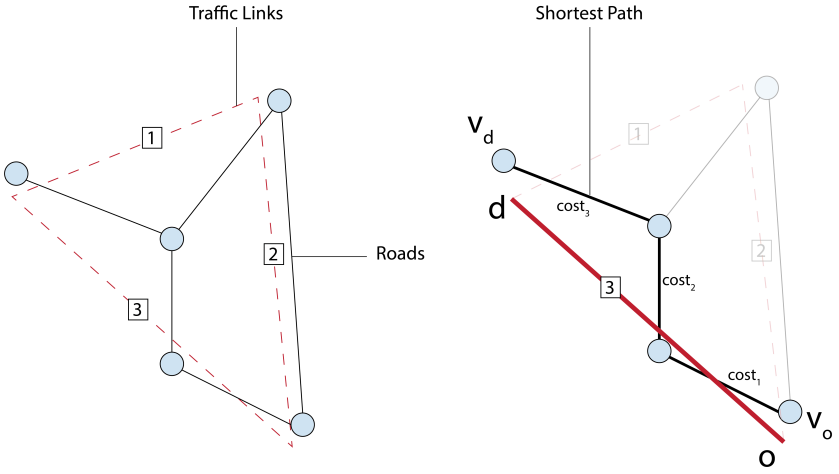


Figure 3.3: Path mapping between the high-fidelity road network and the traffic network. This is implemented as a practical consideration as most traffic O-D matrices do not included for all nodes, thus the attributes of a traffic link can be assumed to depend on a route in the road network rather than a single road.

This can be seen in Figure X. Distance is the minimisation objective because this mapping considers ideal conditions, i.e the undisturbed traffic network before the earthquake. This is minimised using any established weighted shortest path algorithm, in this case A* search is used to find the shortest path. Each building is mapped to a road in $\mathcal{G}_R$, this is the closest road that the building accesses and is represented as a matrix $A_{b,r} \in \{0, 1\}^{|\mathcal{C}_B| \times |\mathcal{E}_R|}$.

$$A_{b,r} = \begin{cases} 1, & \text{if building } b \in \mathcal{C}_B \text{ is assigned to road segment } r \in \mathcal{E}_R, \\ 0, & \text{otherwise.} \end{cases} \tag{3.2}$$

Consider the network shown in figure 3.4 , each building is accessed by one road and the dependency matrix $A_{b,r}$ is visualised in Table 3.1.

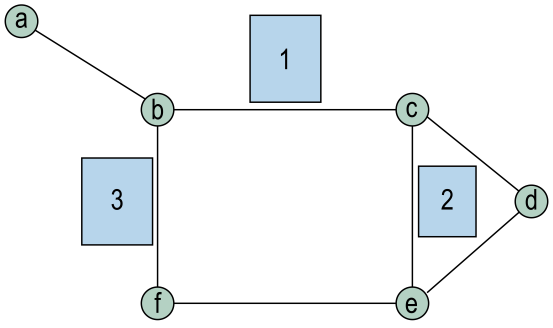


Figure 3.4: Example network of building and roads, not used in the case study of this thesis.

Edge	1	2	3
(a, b)			
(b, c)	■		
(b, f)			■
(f, e)			
(e, c)		■	
(c, d)			

Table 3.1: Road-Building Interdependency matrix, where black = 1, white = 0

The network is modelled as interdependent in its modelling of traffic, using a co-location interdependency of roads to buildings, using $A_{b,r}$, this interdependency is chosen because it aligns well with the available infrastructure components chosen: buildings and roads. Other infrastructure networks such as EPN and GN can also have complex and cascading effects on the functionality of critical facilities, residential buildings, rescue efforts etc. However, in the interest of facilitating a feasible runtime for the environment, the number of infrastructure categories was limited to two as computation time for the various methods discussed can become the limiting factor fairly easily. Furthermore, a major assumption made here is that the debris of one building can only affect one road, this can be far fro reality as one might expect that building 2 for example produces debris which affects road segments (e, c) and (c, d). This assumption is made as for The specific debris-road interdependency outlined is formalised in Guidotti et l, 2016 [45] and Gonzalez et al., 2016[44]. Sediek et al. The debris area prediction differs from Sediek et al. in that they use a trained NN to predict the direction of debris fall in reinforced concrete structures; this thesis considers a random direction. Debris from damaged buildings affects the associated roads' capacities, and thus the post-earthquake MTT of the whole network. The debris area prediction differs from Sediek et al. in that they use a trained NN to predict the direction of debris fall in reinforced conrete structures, this thesis considers a randomly sampled direction. Debris from damaged buildings affects the associated roads' capacities, and thus the post-earthquake MTT of the whole network. For each building $b \in C_B$, let $e_b \in \mathcal{E}_R$ denote the single road edge affected by debris from b. The capacity μ_e of each affected edge $e \in \mathcal{E}_b$ is reduced according to a debris impact function $f_{\text{debris}}(b)$, such that:

$$u'_e = u_e \cdot \left[1 - \max_{\substack{b \in C_B \\ A_{b,e}=1}} f_{\text{debris}}(b)\right], \quad \forall e \in \mathcal{E}_R, \tag{3.3}$$

where u'_e is the effective (post-debris) capacity of edge e. If no building affects edge e, the maximum is taken to be zero. f_{debris} is defined as the maximum lane width reduction that debris imposes on the road. Like seen in Fig X. This is calculated using a predicted lane width using HAZUS's earthquake model technical manual. As centerlines are normally the available geometry for roads, they are used for the distance calculation.

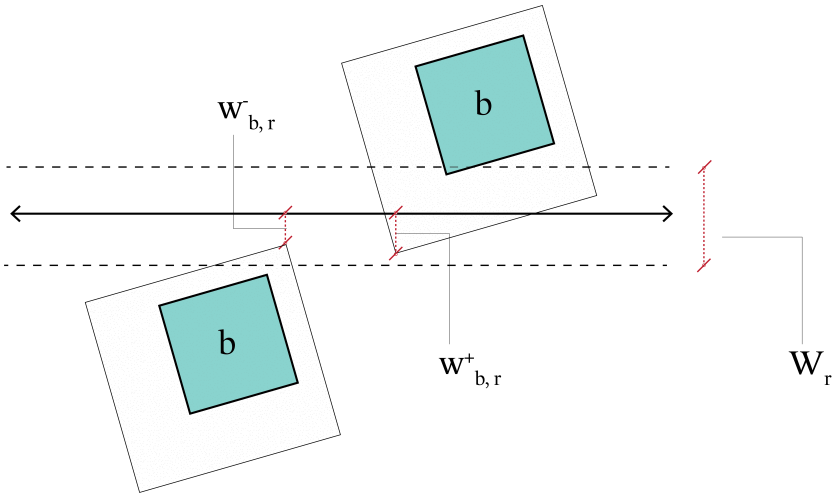


Figure 3.5: Width reduction due to debris is calculated using the road centreline

The capacity reduction f_{debris} of road r due to building b is as follows:

$$f_{\text{debris}}(b) = \begin{cases} \dfrac{w^+_{b,r} + \frac{1}{2}w_r}{w_r}, & \text{if debris overlaps road centreline,} \\ \dfrac{w^-_{b,r}}{w_r}, & \text{otherwise} \end{cases} \tag{3.4}$$

where $w^{+}_{b,r}$ is the minimum distance from any point on building b's debris rectangle perimeter to the centre line of road r. This interdependency only affects the resulting MTT when conducting post-earthquake traffic assignment. When acting in time-step t after an earthquake, $MTT(t)$ is the Mean Travel Time of the traffic network and is used to calculate a cost using pre-earthquake MTT, $MTT(t_d)$.

The time taken to remove debris is principally a function of weight, which is calculated deterministically using HAZUS's Earthquake Model Technical Manual, [35]. The overarching equation used is simillar to Sediek, El-Tawil, and McCormick [82]. Debris clean-up time is defined using the following variables and method:

- D : Total debris weight (tons)
- T : Number of trucks available per day, (default = 0.1)
- L : Loading time per truck (hours), (default =1), [82]
- C : Truck capacity (tons) (default=5), [82]
- τ : One-way travel time to depot (hours), (default=2)
- H : Working hours per day (hours), (default=8)

Trips, N Required:

$$N = \frac{D}{C} \tag{3.5}$$

Total Travel Time, T_{travel}:

$$T_{\text{travel}} = 2\tau \cdot N - \tau = (2N - 1)\tau \tag{3.6}$$

This computes travel time to the disposal site twice per trip (there and back) for a total of N trips, and subtracts the last trip from the the disposal to the building, which is assumed to not be taken as all debris is cleared. Total Loading Time is computed as:

$$T_{\text{load}} = \frac{N \cdot L}{T} \tag{3.7}$$

Total Working Time:

$$T_{\text{total}} = T_{\text{travel}} + T_{\text{load}} = (2N - 1)\tau + \frac{N \cdot L}{T} \tag{3.8}$$

Number of Working Days are computed using a conservative upper estimate:

$$\text{Days} = \left\lceil \frac{T_{\text{total}}}{H} \right\rceil \tag{3.9}$$

This method is rudimentary compared to the supporting work of Sediek, El-Tawil, and McCormick [82]. The assumptions on debris weight prediction and truck loading time are deemed to be appropriate as they can have less variability. However, the time for trips to disposal sites is assumed at a static value. This does not consider any extraneous traffic changes that occur due to capacity reduction and could become an issue for larger environments where trucks have to travel *deep* into the network to recover debris. Furthermore, considering that most debris is usually clear somewhat concurrently, as the buildings and roads can't be repaired if debris is around, the effect of a multitude of large debris

clearing trucks on the traffic network is almost certainly going to cause bottlenecks. In this thesis, it is assumed that even if all the buildings are cleared of debris concurrently, there is no effect on the traffic network.

3.3. Vehicle Traffic Simulation

Traffic simulation is used to compute the MTT of the traffic network. MTT is the sum of the network's trips at time t after the earthquake, $MTT(t)$. Formally, traffic simulation occurs for a type of commodity flowing through a network and can be modelled for different kinds of flowing behaviours, commodities and networks settings [15]. In the case of this thesis, the commodity of the network is driver vehicle, commercial personal; the network is the *arterial* traffic network $\mathcal{G}_{\mathcal{T}}$, which has pre-specified O-D (Origin-Destination) pairs. These are trips that are expected to be made over the course of a day or another time window. An O-D pair is a pair of nodes, trip = (o, d), in the traffic network $\mathcal{G}_{\mathcal{T}}$, such that $o, d \in \mathcal{V}_{\mathcal{T}}$, s.t. $o \neq d$. These trips are facilitated through links $\mathcal{L}_{\mathcal{T}}$ and can go through a single link or be a path along many links. Considering the interdependency of roads to building debris, the properties of each link $l \in \mathcal{L}_{\mathcal{T}}$ are inherited from the set of road links $Path(l) \subseteq \mathcal{E}_{\mathcal{R}}$. Specifically, the effective post-earthquake capacity of link l at time t is $u_l(t)$, which is the minimum capacity $u_r \forall r \in Path(l)$, and is measured in vehicles / hour. Along with capacity, these are the input variables used in traffic assignment.

- Flow x_l, measured in total number of vehicles wanting to use link l during simulation [15]
- free-flow travel time t_l^0 which is the travel time with no congestion effects.

The goal of traffic assignment is to then assign paths to all routes from origins O to destinations D, $paths_{\mathcal{G}_{\mathcal{T}}}$, and minimise the cost of the assignment, considering that congestion effects do not affect route choice. Minimisation is done on the basis that each driver aims to minimise their individual cost (User-Equilibrium, UE). This is in contrast to System-Optimal (SO) traffic assignment where all drivers aim to choose trips such that the total sum of all costs is minimised. This is done as drivers are assumed to be selfish rather than cooperative [15]. The cost of each link is the travel time of that link and the BPR cost function is used. It was developed by the Beaurau of Public Roads and is defined as:

$$t_l(x_l) = t_l^0 \left[1 + \alpha \left(\frac{x_l}{u_l} \right)^{\beta} \right] \tag{3.10}$$

where:

- $t_l(x_l)$ is the travel time of link l under flow x_l
- t_l^0 is the free-flow travel time of link l
- α and β are shape parameters which can be calibrated to different traffic data, 0.15 and 4 are used respectively which are commonly used in literature.

The exact workings of the algorithms used for traffic assignment are not explained. However, as a brief overview the path assignment used it all-or-nothing, where assuming that current travel times are fixed (i.e not changing with congestion) drivers will choose the shortest path between an $O - D$ pair, and all of the demand for that pair will be mapped to that shortest path. The optimisation algorithm used to minimise link travel time is the Franke-Wolfe algorithm which is one of the most competitive algorithms in terms of accuracy, while not being slow. The only change in the traffic network at timestep t after the earthquake is the capacities of traffic links $u_l \forall l \in \mathcal{L}_{\mathcal{T}}$.

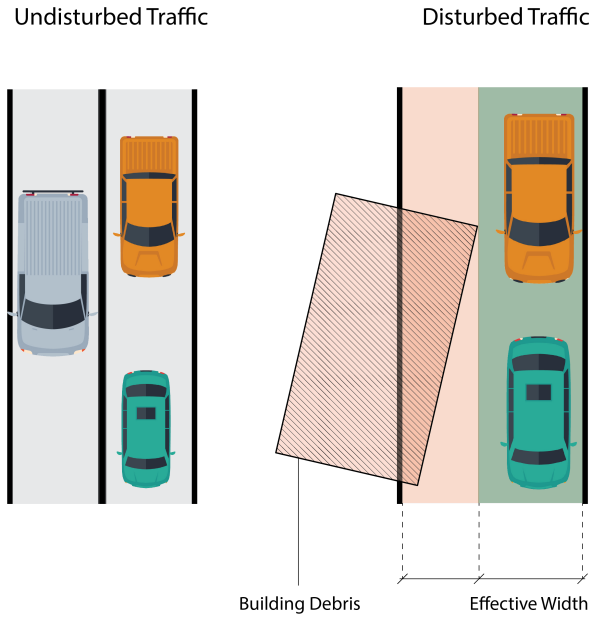


Figure 3.6: Effect of building debris to road capacity. The capacity reduction that any one building causes to its adjacent road is a function of the effective width available for drivers to use.

There are several assumptions made in this method of traffic assignment. The assumed objective that drivers have is to minimise their travel time, which is not always the case. Since assignment is static, it is assumed that drivers have perfect knowledge of the links' travel times before beginning to drive. More specifically, in the context of this thesis traffic patterns are assumed to not change after an earthquake, i.e the $O-D$ pairs remain static across timesteps. This is done so that a cost can be computed fairly. If $O-D$ pairs change between timesteps then the MTT of the network between two timesteps cannot be directly compared. However, it does not reflect realistic post-disaster traffic patterns, where drivers are likely to commute much less for leisure, and much more for emergencies and healthcare trips.

Capacity is measured in vehicles / hour and is directly related to the effective width of the road, see Fig. X. Travel time on link l depends on the flow of x_l. Flow is measured in the total number of vehicles wanting to use link l during simulation [15]. The cost of travelling on link

3.4. Fragility and Seismic Hazard Assessment

Damage and vulnerability functions for buildings and roads are derived using established tools and datasets such as $IN-CORE$ and $HAZUS$. The simulation environment can be constructed either by directly downloading data from sources compatible with $IN-CORE$ and $HAZUS$, or by generating datasets that adhere to the required schema formats of these tools. While the framework developed in this thesis supports integration with real-world datasets—such as the National Structures Inventory (NSI) [25] for buildings, OpenStreetMap (OSM) for road networks, and the National Bridge Inventory (NBI) for bridge data, the results presented later are generated using two custom-designed test-bed environments. These environments and their associated data is consistent with the ones one can download from the sources listed above, but involve less components than an actual environment from a metropolitan area in the US. The initial motivation for using real-world data stemmed from attempts to solve an instance based on Anaheim, California, which includes thousands of buildings and roads. However, due to the increasing computational complexity of such large-scale problems, custom environments were developed to facilitate controlled and tractable experimentation while maintaining schema compatibility with $IN-CORE$ and $HAZUS$. This allows for future extension of the methodology to a larger environment, assuming either an improvement on computational

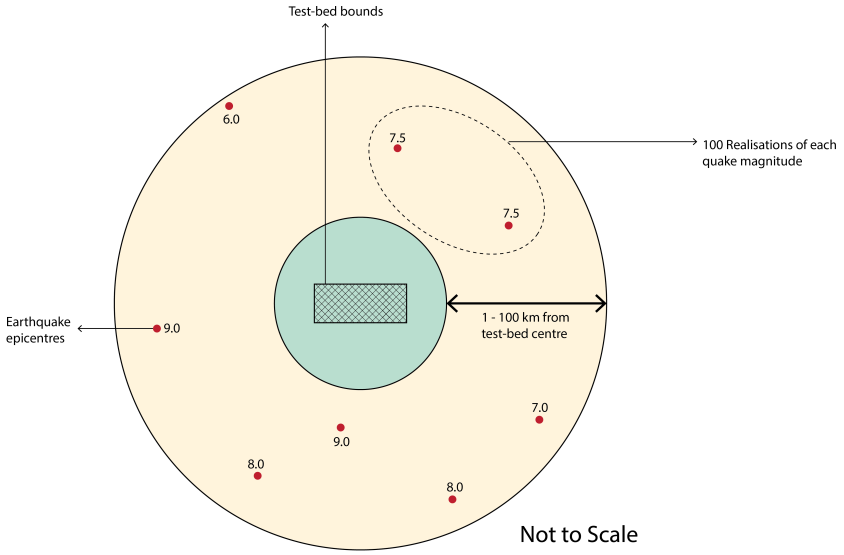


Figure 3.7: *Scenario-based earthquake generation, Author's Own Work.*

complexity or simulation using faster hardware.

Given a set of buildings and roads, $INCORE$ is used to generate earthquakes ranging from 6.0 to 9.0 M with 0.5 M increments. IMs are attenuated at the centres of all building and road geometries. For buildings, PGA is calculated and for Roads and Bridges PGD, $SA_{0.3s}$ and $SA_{1.0s}$ are used; these are the specified IMs in $HAZUS$, [35]. For each earthquake magnitude, 100 realisations of earthquake scenarios are generated and stored in a $Json$ file. Thus, at each RL realisation the *seed* is the earthquake magnitude, from which one of a hundred earthquake realisations is chosen at random. The source point of succeeding earthquakes is chosen as a random point between 1 and 100 km from the average position of all components. This approach to seismic hazard assessment assumes that earthquakes can be generated from any 2D coordinate at a distance away from the test-bed and does not consider the specific fault locations or Magnitude-Frequency (MFD) distributions of each earthquake source (fault). However, it is chosen as it allows for relatively swift simulation times of under 0.2 seconds considering that IMs are saved locally and are only read once training has begun. Conversely, a $PSHA$ methodology was tested using California's $UCERF3$ fault $MFDs$ by modelling the earthquake hazard at each DRL realisation as a $Poisson$ process of all the possible earthquake scenarios, considering each fault can have a stochastic set of rupture scenarios, [8].

The specifics of this methodology are not discussed as it was not used for DRL training. On average $PSHA$ took around 4 to 5 minutes to compute hazard values for a given return period. The amount of simulation time proved to be a bottleneck in using $PSHA$ for $MARL$, thus the simpler methodology is presented. Figure 3.7 illustrates the seismic hazard assessment methodology used.

In order to predict damage and subsequent recovery of buildings, HAZUS specifies the use of certain building attributes which are used along with the IMs to derive a set of damage state distribution for each building. Specifically, the structure type is used for fragility assessment of buildings. Structure type naming follows the general format of material-area-height, where material is the principal mode of construction, area is the size of the building's footprint and height is the category of stories. Some of the structure types specified in Hazus are [35]:

Each building type has a certain fragility curve associated with it, which is defined by a log-normal mean and standard deviation, which are used to predict damage states given IMs. HAZUS specifies three code design levels: Low- Moderate- and High-Code Seismic Design Levels [35], this thesis uses low-code design level as it has the most filled data. This is described in Table 3.3. Damage States are predicted using (2.2), followed by the random sampling of a damage state. Both the sampled damage state and the distribution are used for recovery functions.

Roads follow an almost identical procedure to predicting damage, PGD is used instead of PGA and

Name	Description	Stories
W1	Wood, Light Frame (≤ 5,000 sq. ft.)	1 - 2
W2	Wood, Commercial & Industrial (> 5,000 sq. ft.)	2+
S1L	Steel Moment Frame Low-Rise	1 - 3
S1M	Steel Moment Frame Mid-Rise	4 - 7
S1H	Steel Moment Frame High-Rise	8+
S2L	Steel Braced Frame Low-Rise	1 - 3
S2M	Steel Braced Frame Mid-Rise	4 - 7
S2H	Steel Braced Frame High-Rise	8+
S3	Steel Light Frame	All
C1L	Concrete Moment Frame Low-Rise	1 - 3
1M	Concrete Moment Frame Mid-Rise	4 - 7
C1H	Concrete Moment Frame High-Rise	8+
C2L	Concrete Shear Walls Low-Rise	1 - 3
PC2L	Precast Concrete Frames with Concrete Shear Walls Low-Rise	1 - 3
PC2M	Precast Concrete Frames with Concrete Shear Walls Mid-Rise	4 - 7

Table 3.2: Part of Building Structural Designation types from *Table 5-1 Specific Building Types* in FEMA [35].

Building	Slight		Moderate		Extensive		Complete	
Type	Median	Disp.	Median	Disp.	Median	Disp.	Median	Disp.
W1	0.20	0.64	0.34	0.64	0.61	0.64	0.95	0.64
W2	0.14	0.64	0.23	0.64	0.48	0.64	0.75	0.64
S1L	0.12	0.64	0.17	0.64	0.30	0.64	0.48	0.64
S1M	0.12	0.64	0.18	0.64	0.29	0.64	0.49	0.64
S1H	0.10	0.64	0.15	0.64	0.28	0.64	0.48	0.64
S2L	0.13	0.64	0.17	0.64	0.30	0.64	0.50	0.64
S2M	0.12	0.64	0.18	0.64	0.35	0.64	0.58	0.64
S2H	0.11	0.64	0.17	0.64	0.36	0.64	0.63	0.64
S3	0.10	0.64	0.13	0.64	0.20	0.64	0.38	0.64
S4L	0.13	0.64	0.16	0.64	0.26	0.64	0.46	0.64

Table 3.3: Part of HAZUS Earthquake Model Technical Manual Table 5-39: Equivalent-PGA Structural Fragility - Low-Code Seismic Design Level.

roads fall within two classes in HAZUS: "HRD1" and HRD2" which are two designations of highway roads, the former being arterial, inter-state roads and the latter being intra-state narrower highway roads. Bridges, on the other hand, have a more lengthy procedure for predicting damage. HAZUS specifies a table of median $SA_{1.0s}$ and PGD values for each bridge type (HWB1-HWB28). Depending on the bridge's number of spans, skew angle, span width, bridge length and maximum span length, several modification factors are applied to the appropriate median values to retrieve the ground shaking- and ground failure-related damage state probabilities. The most severe of the two is taken as the actual distribution as it governs design. This thesis only uses ground-shaking related damage state probabilities for bridges as it was found that ground-failure was rarely producing more severe results than ground-shaking related failure (SA). This was done as the computation for modifying the PGD means (ground-failure) is a lot lengthier and made overall computation time unnecessarily long. This calculation procedure i found in equations 7-1 to 7-7 and tables 7-6 to 7-8 in FEMA's HAZUS Earthquake Model Technical Manual [35].

Overall each infrastructure component has 6 key attributes after this analysis. These are the limit state probabilities of each damage state (None, Slight, Moderate, Extensive and Complete) and the sampled damage state as an integer value ranging from 0 to 4. These are used to predict losses and repair times. During repair the damage state of each component is considered fully observable, thus the following procedure is followed to change the damage state of components:

1. **Sample Initial Damage State:**
   $P(DS_n|b,r) = P(DS_1), P(DS_2), P(DS_3), P(DS_4), P(DS_5)\} \forall b \in C_{\mathcal{B}}, \forall r \in C_{\mathcal{R}}$ is the probability distribution over damage states for each component. It is sampled from a probability distribution of five

limit state probabilities, which are themselves sampled from a log-normal cumulative distribution function, given a realised IM, log-normal mean σ and dispersion β. The damage state distribution must always sum to 1 for all buildings and roads.

2. **Initialize Damage State Distribution:**
Define the initial damage state distribution as a one-hot vector $\boldsymbol{d}_0(c)$ for component c. This is done so that HAZUS recovery functions usually depend ona distribution over damage states. Furthermore, it allows for the future formulation of states not as one-hot vectors but as full left-skewing distributions over timesteps. The one-hot damage state distribution is defined as:

$$d_{0,ds}(c) = \begin{cases} 1 & \text{if } ds = ds_0 \\ 0 & \text{otherwise} \end{cases} \tag{3.11}$$

3. **Repair Progress and Damage State Reduction:**
At each repair update step t, if the equivalent amount of repair time passed corresponds to a reduction by one damage state level, update the distribution:

$$\boldsymbol{d}_t = \text{ShiftLeft}(\boldsymbol{d}_{t-1}) \tag{3.12}$$

where ShiftLeft moves the 1 in the vector one position to the left, modeling a stepped, linear decrease in damage state. For example:

$$\boldsymbol{d}_{t-1} = [0,0,0,0,1] \quad \Rightarrow \quad \boldsymbol{d}_t = [0,0,0,1,0] \tag{3.13}$$

4. **Stopping Condition:**
The process continues until the damage state reaches the lowest possible level ($ds = 1$), i.e., $\boldsymbol{d}_t = [1,0,\ldots,0]$.

3.5. Markov Decision Process

As described in previous sections, an MDP is used in this thesis to describe the decision-making environment. This MDP is defined as fully-observable, with nevertheless stochastic seed conditions. This means that given an action a in state s, then we know the next state s' with a probability of 1; however, the initial condition, or damage state of the community is stochastic. This is to say that the income loss, repair times and other losses are stochastically sampled each time the environment is reset to simulate a new earthquake.

- $\mathcal{M}$ is the set of agents that take actions on the components.
- $\mathcal{S} := \times_{m \in \mathcal{M}} \mathcal{TR}_m$, where each $\mathcal{TR}_m \in [0, 500]$, is the joint state space representing the repair times of all components, and thus for each agent $m \in \mathcal{M}$. The upper bound of repair time is taken as 500 as most HAZUS recovery functions for buildings, roads and bridges do not go above 500 days of repair.
- $\mathcal{A} := \mathcal{A}^1 \times \mathcal{A}^2 \times \cdots \times \mathcal{A}^{\mathcal{M}}$ is the joint action space, where each $\mathcal{A}^m$ represents the individual action space for agent m. There are two actions: do nothing and repair.
- $\mathcal{T}(s, a, s')$ is the dynamics model of the environment, for any state s and action a, the next state s' is always known with a probability of 1
- $\mathcal{R}(s, a)$ is the reward model, which includes the costs considered in the recovery process after an earthquake and which affect resilience loss.
- O is the observation space and is identical to the state space $\mathcal{S}$, which means that the agent has full access to the state of the environment at each time step: $O = \mathcal{S}$. The environment is fully observable.
- $t_H \in \mathbb{N}_0$ is the finite time horizon at which the environment truncates.

The main assumptions made in this MDP is that it has a stochastic dynamics model and thus, it is fully observable.

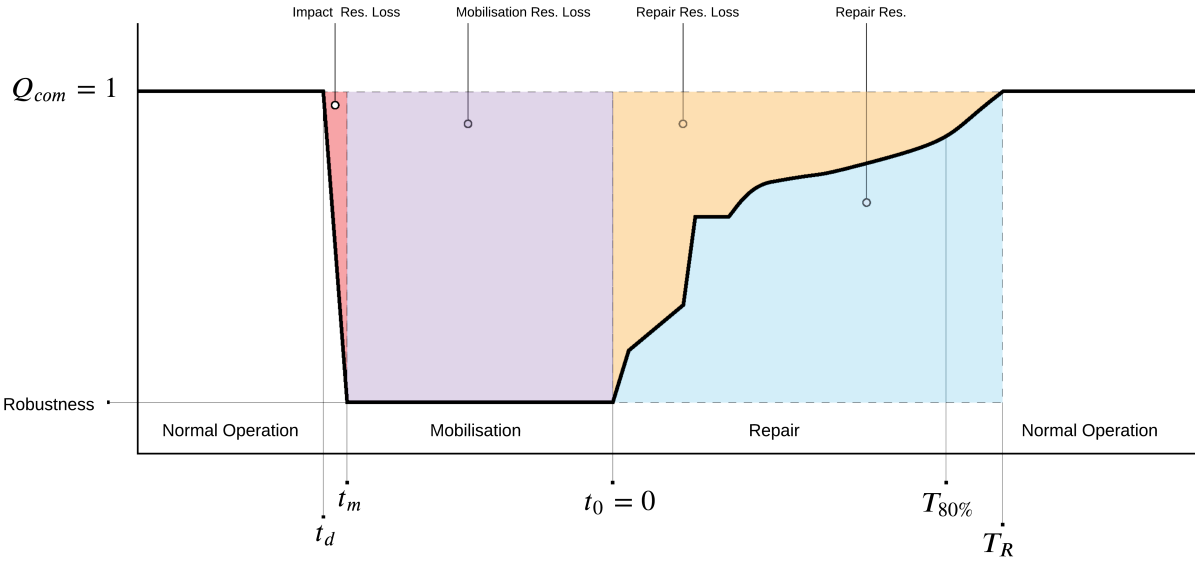


Figure 3.8: Illustration of Community Functionality-time curve with markers on key times used in this methodology and their associated loss areas

3.6. Objective Function and Reward

The reward function is defined as a function of the change in community functionality over time steps. Community functionality is the objective function and is defined as an aggregate measure of sub-system functionalities. Community functionality Q_{com} is defined as:

$$Q_{\text{com}}(t) = Q_{\text{econ}} + Q_{\text{crit}} + Q_{\text{health}} \tag{3.11}$$

$$Q_{\text{com}}(t) = w_{\text{econ}} \cdot q_{\text{econ}} + w_{\text{crit}} \cdot q_{\text{crit}} + w_{\text{health}} \cdot q_{\text{health}}, \tag{3.11}$$

where q_{econ}, q_{crit}, and q_{health} are the functionalities of the economic, critical infrastructure, and healthcare subsystems respectively, and w_{econ}, w_{crit}, and w_{health} are their corresponding weights reflecting relative importance in the overall community functionality. Weights in the reward formulation, including the ones mentioned here are seen as a hyper-parameters, which do not affect the learning performance, but rather the nuance and adaptability of learning to different resilience objectives.

Economic functionality $q_{econ}(t)$ is defined generally as:

$$q_{\text{econ}(t)} = 1 - EL(t) \tag{3.14}$$

where $EL(t)$ is the total economic loss at time t and is defined as a weighted sum of economic component losses:

$$EL(t) = \sum w_i L_i(t) \quad \forall i \in \{Inc, Rep, Reloc, Traffic\} \tag{3.15}$$

where i denotes an economic component, in this case income, repair, relocation and traffic costs. Thus, economic functionality is formulated as:

$$q_{\text{econ}}(t) = 1 - \left[w_{\text{inc}} * \frac{IL_B(t)}{I_B(t_d)} + w_{\text{rep}} * \frac{C_{B,\text{rep}}(t) + C_{R,\text{rep}}(t)}{C_{B,\text{repl}} + C_{R,\text{repl}}} + w_{\text{reloc}} * \frac{C_{B,reloc}(t)}{C^*_{B,\text{reloc}}} + w_{\text{traffic}} * \frac{C_{TD}(t)}{I_B(t)} \right] \tag{3.16}$$

where:

- $IL_B(t)$ is the sum of all buildings' income losses at time t and $I_B(t_d)$ is the sum of all buildings' income right at the time of the earthquake.
- $C_{B,\text{rep}}(t)$ and $C_{R,\text{rep}}(t)$ are the sums of building and road repair costs at time t respectively.
- $C_{B,\text{repl}}$ and $C_{R,\text{repl}}$ are the sums of building and road replacement costs, which are stationary.
- $C_{B,\text{reloc}}(t)$ is the sum of all buildings' relocation costs and $C^*_{B,\text{reloc}}$ is the maximum sum of all buildings' relocation costs.
- $C_{TD}(t)$ is the traffic delay cost at time t

Income is taken from $HAZUS$ Earthquake Model Inventory Technical Manual Table 6-16 [34] .The total cost at time t, denoted $H(t)$, is given by:

Building repair costs in dollars, $C_{B,\text{rep}}(t)$ follow the formula from [35] eq. 11-1, (note that in the manual its defined per occupancy class, whereas in this thesis its defined per building):

$$C_{B,\text{rep}}(t) = \sum_{ds=1}^{4} \left[BRC(b|occ_b) * P(ds_b) * RCS(b|occ_b) \right], \tag{3.17}$$

where:

- $BRC(b|occ_b)$ is the building replacement cost of building b given occupancy type occ_b which is found in Tables 6-2 and 6-3 of [34],
- $P(ds_b)$ is the probability of building b being in damage state ds which is stochastically predicted as explained above when the environment is reset
- $RCS(b|occ_b)$ is the structural repair cost ratio, given as a percentage of building replacement cost, which is found in Table 11-2 of [35].

Road repair costs include road and bridge repair costs. Specifically, roads use data from FHWA's Highway Investment Analysis Methodology [32]. The manual specifies costs for resurface and a cost for reconstruction per unit area of a certain type of road. These are used to define the costs in dollars as follows:

$$C_{R,\text{rep}}(t) = d_r \cdot w_r \cdot \begin{cases} cost_{RS} & \text{if } 0 < ds_r < 3 \\ (cost_{RS} + cost_{RC}) & \text{otherwise} \end{cases} \tag{3.18}$$

where:

- d_r is the length of road r in miles.
- w_r is width of road r in number of lanes. HAZUS $HRD1$ roads have 6 lanes and HAZUS $HRD2$ have 4 lanes [35],
- $cost_{RS}$ is the resurfacing cost,
- $cost_{RC}$ is the reconstruction cost,
- ds_r is the randomly rampled damage state of road r

For HAZUS's HRD1 road designation, values from FHWA's *Freeway / Interstate : Major Urbanized* road designation are taken [32]. For HAZUS's HRD2 road designation, values from FHWA's *Other Principal Arterial : Large Urbanized* road designation are taken [32]. Traffic delay costs are calculated using FHWA's "Work Zone Road User Costs - Concepts and Applications", Chapter 2 [3]. The cost calculation is stochastic as it sampled a buisness-personal ratio every time it is called. The values used to sample from range from $91\% - 9\%$ to $96\% - 4\%$. While the differences made are not drastic it still allows for a non-deterministic traffic delay cost, which can more realistically reflect real traffic conditions. Traffic delay cost is calculated as a yearly delay cost. Considering MTT is calculated in hours for a day's traffic pattern, traffic delay costs are defined as:

$$C_{TD,base}(t) = \left[MTT(t) - MTT(t_d)\right] * \tau_{drive} * 365, \quad s.t \quad MTT(t) \geq MTT(t_d) \tag{3.19}$$

$$\tau_{drive} = \begin{bmatrix} r_{business} & r_{personal} \end{bmatrix} \cdot \begin{bmatrix} \tau_{business} \\ \tau_{personal} \end{bmatrix}, \quad s.t \quad r_{buisness} + r_{personal} = 1.0 \tag{3.20}$$

$$C_{TD}(t) = \min \left[\sum_{b \in C_{\mathcal{B}}} I_b(t_d),\ C_{TD,base}(t) \right] \tag{3.21}$$

where:

- $C_{TD,base}(t)$ is the base traffic delay.
- $MTT(t) - MTT(t_d)$ is the difference in mean travel time at time t after the earthquake and before the earthquake. $MTT(t)$ can only be bigger than $MTT(t_d)$,
- τ_{drive} is the average time value of commuters travelling in the network in dollars per hour,
- $r_{buisness}$ and $r_{personal}$ are the randomly sampled buisness and personal travel ratios.
- $\tau_{buisness}$ and $\tau_{personal}$ are the time values of buisness and personal time travel which are calculated following the methodology in [3]
- $I_b(t_d)$ is the income of building b before the earthquake.

Traffic delay costs are capped in this study due to the potential for unrealistically large values when road capacities approach zero. In such cases, traffic delays can grow exponentially, which may not accurately reflect real-world traffic behaviour. Specifically, if a roadway becomes nearly or completely blocked, it is assumed that drivers would forgo commuting altogether rather than experience indefinite delays. Under this assumption, the complete obstruction of the transportation network results in a complete loss of income due to the inability of individuals to reach their workplaces or carry out travel-related economic activities.

This assumption implies that drivers after an earthquake continue to pursue the same travel patterns as they did prior to the earthquake, irrespective of changes in the desired destinations of drivers after an earthquake. As a result, when road capacities are significantly diminished, the model generates disproportionately high delay costs due to unadjusted route choices and travel demand.

While this approach abstracts away the complex dynamics of adaptive driver behaviour, it provides a computationally efficient baseline for estimating post-earthquake traffic. These simplifications also highlight key areas for future research, such as incorporating dynamic traffic assignment or elastic demand models to more realistically simulate post-earthquake driver responses.

Relocation costs for the residential population are calculated based on the methodology outlined in $HAZUS$ [35], specifically Equation 11-14. Although $HAZUS$ defines these costs by occupancy class, this thesis adapts the calculation on a per-building basis.

$$C_{B,\text{reloc}}(t) = \sum_{b \in C_{\mathrm{l}}} \left[A_b * (1 - \%OO_{occ,b}) * \sum_{ds=3}^{5} \left[P(ds) * (DC_{occ,b} * RENT_{occ,b} * RT_{ds}) \right] \right] \tag{3.22}$$

where:

- A_b is the total floor area of building b,
- $\%OO_{occ,b}$ is the percent owner-occupied of building b given its occupancy class occ_b and is defined in [34],
- $P_b(ds)$ is the probability of building b being damage state ds,
- $DC_{occ,b}$ is the disruption costs for building b given its occupancy class occ_b and are defined in [34],
- $RENT_{occ,b}$ is the rental cost of building b given occupancy occ_b and is defined is [34],

- RT_{ds,occ_b} is the recovery time for building b given its damage state ds and its occupancy type occ_b, this is defined in [35] and includes both repair time and post-repair functional downtime.

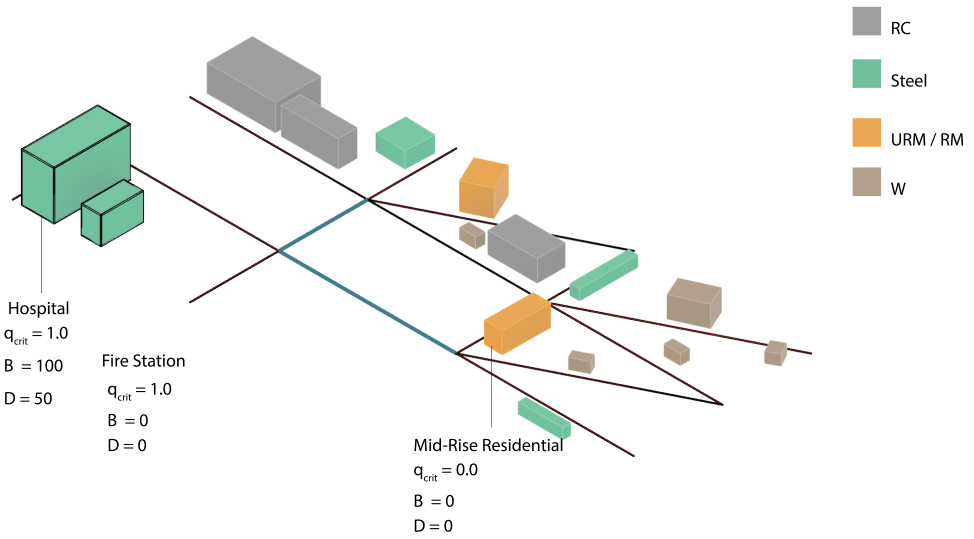


Figure 3.9: Critical functionality is only applicable to essential buildings such as fire stations or hospitals. These have a un-disrupted critical functionality of 1, with all other buildings having a value of 0. Likewise, only hospitals and medical facilities have values for B and D which are the number of beds and doctors respectively.

Critical infrastructure functionality $q_{\text{crit}}(t)$ is defined as:

$$q_{\text{crit}}(t) = \frac{\sum_{c \in \mathcal{C}} q_{\text{crit}}(t)}{\sum_{c \in \mathcal{C}} q_{\text{crit}}(t_0 - \varepsilon)} \tag{3.14}$$

where $q_{\text{crit}}(t)$ is 1 if component c is critical at time t, and 0 otherwise. The numerator sums over the critical components' functionality at time t, while the denominator is the sum of the functionalities of critical components just before the event (at $t_0 - \varepsilon$).

Health functionality $q_{\text{health}}(t)$ is defined as:

$$q_{\text{health}}(t) = w_{\text{bed}} \cdot q_{\text{bed}}(t) + w_{\text{doc}} \cdot q_{\text{doc}}(t), \tag{3.15}$$

where: - $q_{\text{bed}}(t)$ is the functionality of the health subsystem's beds, and is calculated as the ratio of current available beds to the initial number of beds:

$$q_{\text{bed}}(t) = \frac{\sum_{b \in \mathcal{C}_{\text{l}}} B(t)}{\sum_{b \in \mathcal{C}_{\text{l}}} B(t_0 - \varepsilon)} \tag{3.16}$$

where $B(t)$ is the number of beds available at time t, and B_0 is the initial number of beds before the event.

- $q_{\text{doc}}(t)$ is the functionality of the health subsystem's doctors, and is calculated as the ratio of current available doctors to the initial number of doctors:

$$q_{\text{doc}}(t) = \frac{\sum_{d \in \mathcal{C}_{\text{f}}} D(t)}{\sum_{d \in \mathcal{C}_{\text{f}}} D(t_0 - \varepsilon)}, \tag{3.17}$$

where $D(t)$ is the number of doctors available at time t, and D_0 is the initial number of doctors before the event. The weights w_{bed} and w_{doc} reflect the relative importance of beds and doctors in the overall

healthcare functionality. Critical functionality is illustrated in figure 3.9, where the environment is in a pre-disaster state of no damage. The only buildings that have a critical functionality of 1.0 are the hospital and fire station, all other buildings are non-critical and have no doctors or hospital beds. This follows *HAZUS's* essential facility designation which includes other government offices too such as police stations [35].

This formulation of community functionality is made so that it aligns with formulations of resilience seen in the general literature, given the calculated losses presented in this methodology. The principal aim is to use post-earthquake community functionality to compute its finite time-integral as the recovery resilience. Furthermore, considering similar formulations from Yang et al. [96], the reward is formulated as cost. In doing so, the reward is the negative instantaneous resilience loss at each time-step.

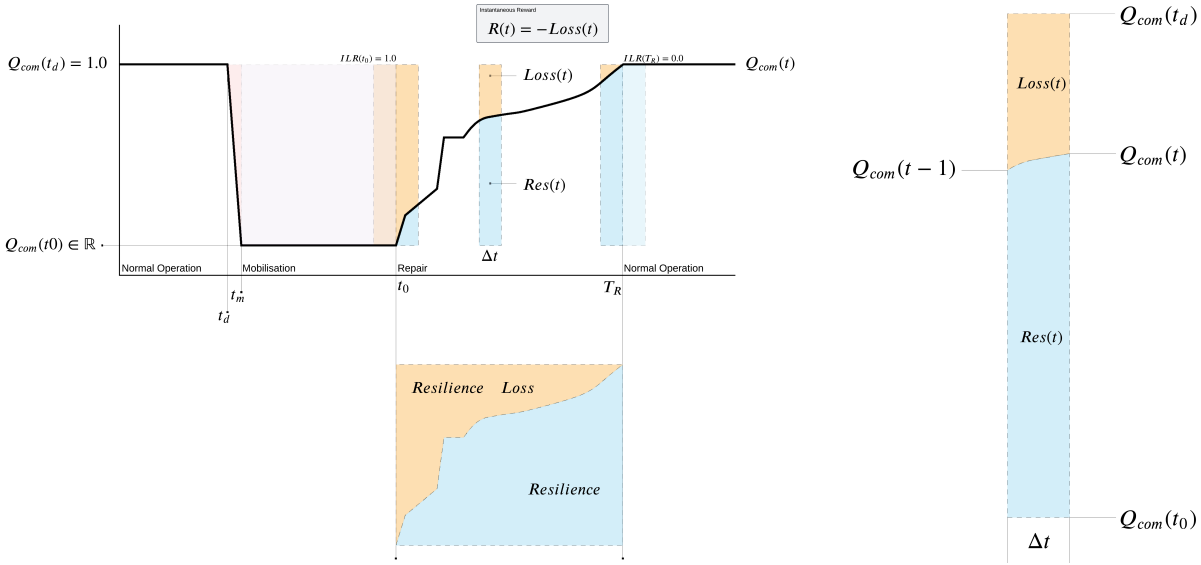


Figure 3.10: The reward at each time-step is the negative instantaneous loss of resilience $-Loss(t)$ which is the tranche of resilience loss associated with that time-step. The total resilience loss is then the sum of rewards, or the returns of a realisation

The cost at time t is the resilience loss, $Loss(t)$ and $Res(t)$ is the resilience at time t:

$$Res(t) = \frac{1}{2}\Delta t * [Q_{com}(t) + Q_{com}(t-1] \tag{3.23}$$

$Loss(t)$ is calculated using the sum of the two:

$$Res(t) + Loss(t) = \Delta t * [Q_{com}(t_d) - Q_{com}(t_0)] \tag{3.24}$$

$$Loss(t) = \left[\Delta t * \left(Q_{com}(t_d) - Q_{com}(t_0)\right)\right] - \left[\frac{1}{2}\Delta t * \left(Q_{com}(t) + Q_{com}(t-1)\right)\right] \tag{3.25}$$

Thus, reward R is defined as the state-reward at each time-step:

$$R(t) = R(s,t) \tag{3.26}$$

$$R(s,t) = -Loss(t) \tag{3.27}$$

3.7. Environment Dyamics

While above sections describe the attributes of the environment that contribute to states actions and rewards, this section predominantly focuses on how these aspects come together to define the dynamics of the environment. Considering that the environment is defined as fully observable, the dynamics describes the interaction of building and road agents taking actions and how their interaction affects what action they take. Fig. X shows a high level overview of these interactions and how they lead to agents receiving reward.

A rollout of the environment is a trajectory from an initial state to some termination or truncation state. The environment is terminated if all components are repaired, and truncated if the finite time horizon is reached. At each time step, a *joint action* $\mathbf{A}_t \in \{0,1\}^{\mathcal{M}}$ is selected, where:

- $|\mathcal{M}| = |C_{\mathcal{B}}| + |C_{\mathcal{R}}|$, the number of agents is the same as the total number of components (buildings and roads),
- $a_t^c = 1$ indicates a *repair action* for component c, and $a_t^c = 0$ indicates a *do-nothing action*.

The action vector is partitioned as:

$$\mathbf{a}_t = [\mathbf{a}_t^b, \mathbf{a}_t^r] \quad \text{where } \mathbf{a}_t^b \in \{0,1\}^{n_b}, \quad \mathbf{a}_t^r \in \{0,1\}^{n_r} \tag{3.28}$$

At each time step, a random ranking L_t over the components is generated:

$$L_t : \{1, 2, \ldots, m\} \to \{1, 2, \ldots, m\} \tag{3.29}$$

Considering the number of crews as n_{crews} as time t, the set of requested repair actions at time t is as follows:

$$NA_{rep}(t) = \left\{c \in \{1, \ldots, m\} \mid a_t^c = 1\right\} \tag{3.30}$$

If $|NA_{rep}| \leq n_{\text{crews}}$, all requested repairs are allowed. Otherwise, repairs are restricted to the top-ranked n_{crews} components:

$$a_t^i = \begin{cases} 1 & \text{if } c \in NA_{rep} \text{ and } L_t(c) \leq k \\ 0 & \text{otherwise} \end{cases} \quad \text{where } k = \min(n_{\text{crews}}, |NA_{rep}(t)|) \tag{3.31}$$

Ranking components is done at random as it is not the aim to influence *MARL* agents directly by heuristically guiding their actions. Rather the goal is for the agents to learn to navigate through this *noise* by bounding their joint repair effort to satisfy the number of repair crews available. In doing so, the random ranking should on average perform badly as the top *n-crews* components after randomly ranking are expected to not be the components that should be repaired. Therefore, agents will try to navigate around this by choosing actions such that this random ranking is not applied. This is a key aspect of the dynamics of the environment as it allows for budgetary constraints in terms of repair-crew availability, without giving explicit directions to agents.

Considering the formulation of reward, there are no explicit restrictions or penalties placed on agents for going *over-budget*. Costs are only considered as total remaining costs and not instantaneous repair costs, thus they are not seen as a budgetary constraint. Crew prioritisation is then seen to act as a budgetary constraint; while not being directly measured as a monetary constraint, the number of available crews is a proxy for the available repair funds.

Considering a building and a road at *Complete* damage states, figure 3.12 shows the profile of their attributes as they get *optimally* repaired; that is, they are repaired at every time-step. The phases of repair are illustrated as:

- **Yellow**: Removal of a building's debris, only applies to the debris of a single building, not any of its neighbours. This phase only affects the imposed capacity reduction a building has on its adjacent road, it does not affect any of the associated losses.

- **Red**: Repair of a building or a road. This phase reduces the repair time and costs. In the case of buildings the relocation costs are also reduced but not the income losses. In the case of roads, the capacity reduction is also reduced.
- **Green**: Recovery phase. This phase only applies to commercial buildings that can generate income. It is the phase after repair has completed and is when the income losses are reduced, following a quadratic increase.

The performance attributes shown are normalised relative to the initial damage state. This is done purely for visualisation purposes. It is important to note that here income shown as 0 is not 0 income, but is instead the immediate post-disaster income considering the maximum income loss. The three buildings shown are a mid-rise residential building with *HAZUS* occupancy designation *RES3A*, a commercial centre with designation *COM2* and an agriculture building such as a grain storage silo or farmhouse with designation *AGR1*.Residential buildings are not considered to generate any income in this thesis, which HAZUS's general recommendations [35]. However, residential buildings along with other lodging- or accommodation-related occupancies are the only occupancy types which have relocation costs. Relocation costs follow the damage states with a *lag*, i.e when a building is in *Slight* damage state the relocation cost drops to 0. Conversely, the commercial building's repair profile shows a zero relocation cost but a positive income. Income and repair costs decay and grow quadratically to reflect a more realistic recovery scenario. In doing so, the assumption made is that most of the repair costs are spent when beginning to repair and as repair progresses the instantaneous repair costs reduce. Likewise, as companies and employees begin to occupy a freshly repaired commercial building it can be expected that at first they generate little to no income, but as time progresses they are able to reach pre-disaster income levels faster.

Furthermore, it can be seen that the agricultural building has a much longer debris clearing phase and a shorter repair phase. This can be intuitively understood as farm structures are usually very large in scale and thus might produce a lot of debris; however, they can often be simply constructed, having less mechanical services than a commercial building. A non-optimal profile of repair for a building is shown in figure 3.11, where usually the repair phase is longer than expected. The debris removal phase is also usually a bit longer, but the recovery phase remains the same as it is not dependant on the type of action chosen.

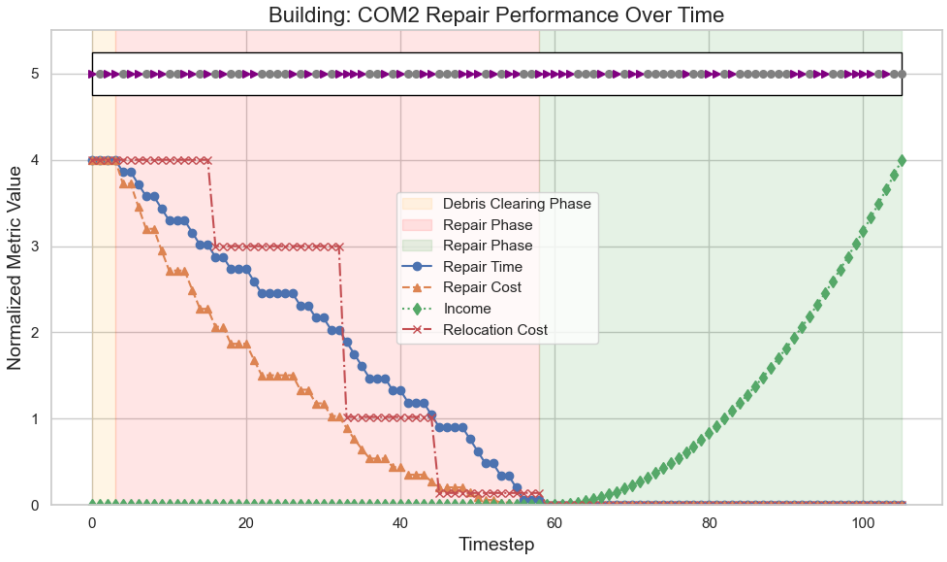
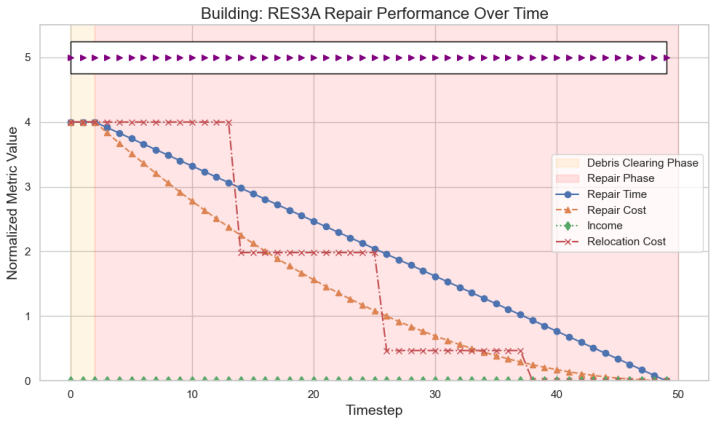
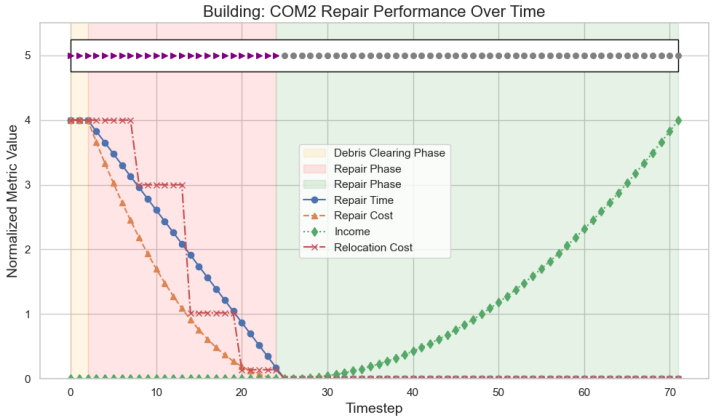


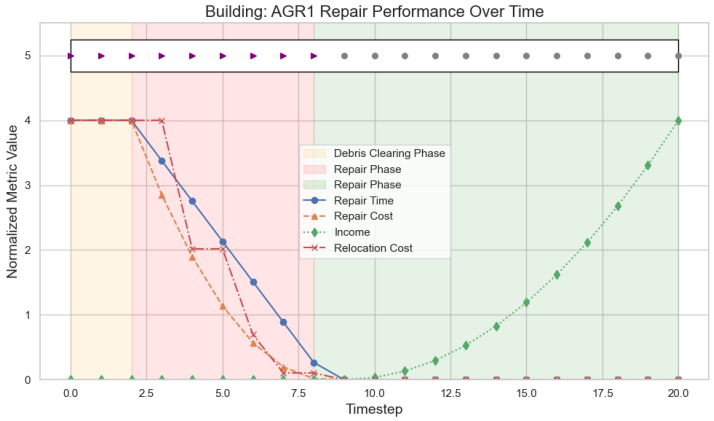
Figure 3.11: d) Repair and Recovery of Mid-Rise Commercial Building while taking random actions.



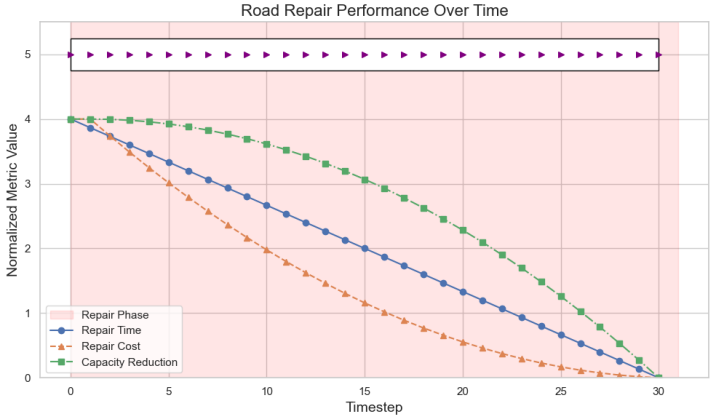
a) Optimal Repair and Recovery of Mid Rise Residential Building



b) Optimal Repair and Recovery of Mid Rise Commercial Building (Non-Essential)



c) Optimal Repair and Recovery of Agricultural building



d) Optimal Repair of arterial Highway Road

Figure 3.12: Optimal repair sequences of different road and building designations as per $HAZUS$, [35]

3.8. Environment Solvers

This thesis uses three principal solvers. MARL, random and importance-based strategies. Importance-based and random policy solvers are seen as *baseline* methods and MARL is seen as the *benchmarked* method. Specifically VDN (Value Decomposition Network) is principally used for testing as it can be robust and reliable method to test and debug [92]. $DCMAC$ and $QMIX-PS$ are also used. $QMIX-PS$ is similar to $VDN-PS$ but includes a mixer network instead just a summation over individual Q-values. $DCMAC$ is a case of $CTCE$ (Centralised Training with Centralised Execution), this means that both actor and critic networks have input feature vectors that are the size of the joint state space instea of the individual state spaces. $CTCE$ methods tends to perform better than other DRL algorithms but face scaling challenges as the size of the features of the networks grows with the size of the joint state space, even if the individual state space stays constant.

Random policy is self explanatory and involves the agents taking random actions at each time-steps without using any observations or reward to guide their decisions. Importance-based repair scheduling is a rudimentary ranking-based, custom solver including rule-based value, or rank computation. It involves ranking roads and buildings in ascending order by considering an aggregate metric *repair value*; doing so by repairing the top n_{crews} number of repair-value-sorted components. Fig 3.13 shows the flow chart for importance-based repair scheduling.

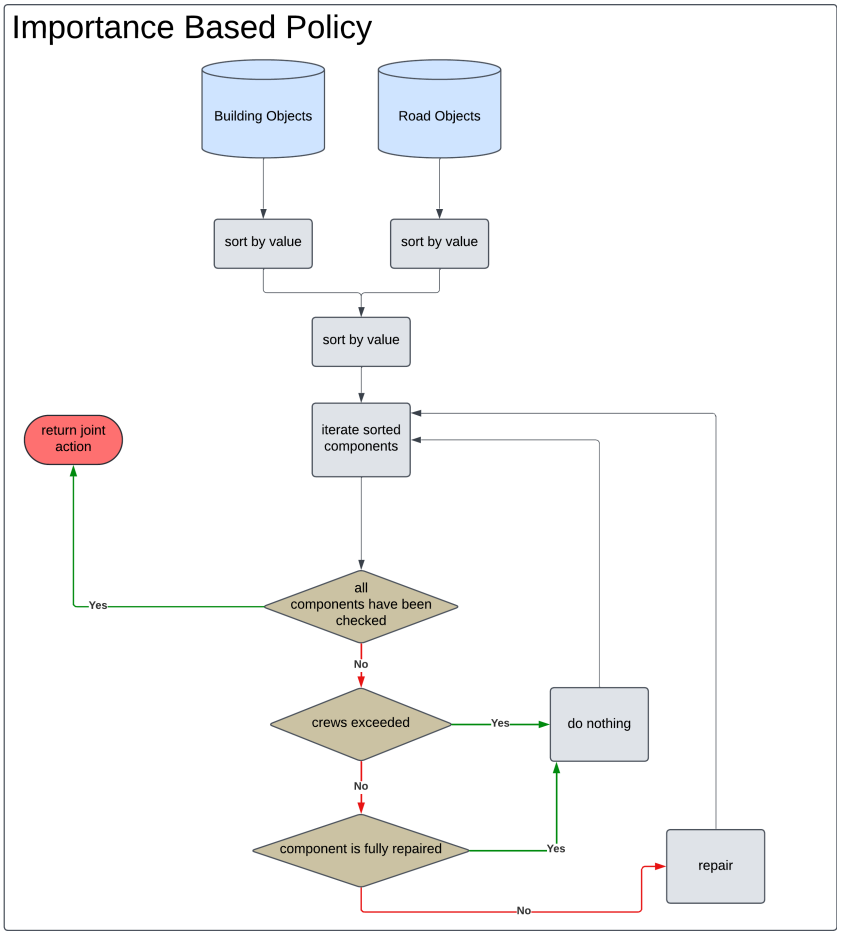


Figure 3.13: Flow chart of the importance-based ($IMPB$) algorithm

The algorithm uses the value V of components to rank them and keep the number of repair actions below the number of repair crews available. Value for buildings V_b is defined as:

$$V_b(t) = \left(\frac{I_b(t)}{I_b^{\text{nom}}}\right) \cdot \left(\frac{ds_b(t)}{ds_{max}(t)}\right) \cdot \left(\frac{A_b}{A_b^{\text{nom}}}\right) \cdot \begin{cases} 1, & \text{if } q_{\text{crit}}(t_d) = 1 \\ 0.5, & \text{if } q_{\text{crit}}(t_d) = 0 \end{cases} \quad \forall b \in C_{\mathcal{B}} \tag{3.32}$$

where:

- $I_b(t)$ is the income of building b at time t,
- I_b^{nom} is the nominal pre-disaster income, $\arg\max_{b \in C_{\lfloor}} [I_b(t0 - \epsilon)]$
- $ds_b(t)$ is the damage state of building b at time t,
- $ds_{max}(t)$ is the nominal building damage state at time t, $\arg\max_{\lfloor \in C_{\mathcal{B}}}(ds_b(t))$

Likewise, the value of each building at each timestep t is defined as:

$$V_r = \frac{u_r(t) \cdot ds_r(t)}{u_{nom}(t) \cdot ds_{nom}(t)} \quad \forall r \in C_{\mathcal{R}} \tag{3.33}$$

where:

- $u_r(t)$ is the capacity of road r at time t,
- $ds_r(t)$ is the damage state of road r at time t,
- $u_{nom}(t)$ is the nominal capacity of the network at time t, $u_{nom}(t) = \arg\max_{r \in C_{\mathcal{R}}}(u_r(t))$
- $ds_{nom}(t)$ is the nominal damage state at time t

The two methods for assigning value to components are combined to get a joint sorting of values in the network, which is then used to select a joint action. The importance based algorithm is shown in figure 3.13. First, the algorithm receives updated building and road objects after a reset or step of the environment. Then, buildings and roads are sorted, components are iterated, if the component is fully repaired, "do nothing" action is selected, otherwise repair action is selected.

This key attributes of the solver lie in the value formulation of buildings and roads. Component value is formulated as a product of normalised attribute functionalities, with time-dependant normalisation constants.Along with the overlying assumptions of the environment, importance-based, post-earthquake repair scheduling as defined here makes the principal assumption that the prioritisation of actions depends only on the parameters included, e.g $I_b(t)$, $ds_r(t)$. This method does not consider other factors that might play a role in reducing resilience loss, such as the spatial qualities of the combined networks or the cascading effects of debris on neighbouring buildings. It is nevertheless, an online algorithm that uses the observations of agents during a realisation to rank components. Random policy selects a joint action with probability $\frac{1}{M}$.

The principal MARL algorithm used in this thesis is VDN (Value Decomposition Network) with parameter sharing and is highly reliable and robust. It was developed specifically for co-operative MARL and can scale well as it involes only one shared agent network and sharing of parameters[92]. Furthermore, the q-value mixer used is simply a summation over agent q-values which makes it computationally cheaper than other methods. Furthermore, experience replay is used during training, that is the storage of previous trajectories and random sampling of mini batches during training to use as input into the Q-Network NN. This is standard practice for MARL research. The reason for using experience replay is to combat the problem of auto-correlation during training. Given the recurrent nature of data flowing into the network, if experience replay is not used successive datapoints can be simillar, and thus the data is said to be autocorrelated, this makes training difficult. In essence, this makes the practice of training the network closer to supervised learning, which is preferrable as NNs, all things considered, perform better when large amounts of highly-diverse data are used as input. The architecture of the network involves 2 hidden layers of 64 neurons each, with a MSE loss function, adam optimiser and relu activation function. The implimentation of the algorithm is developed by Prateek Bhustali and includes a few small modifications that allow use in this thesis, [11].

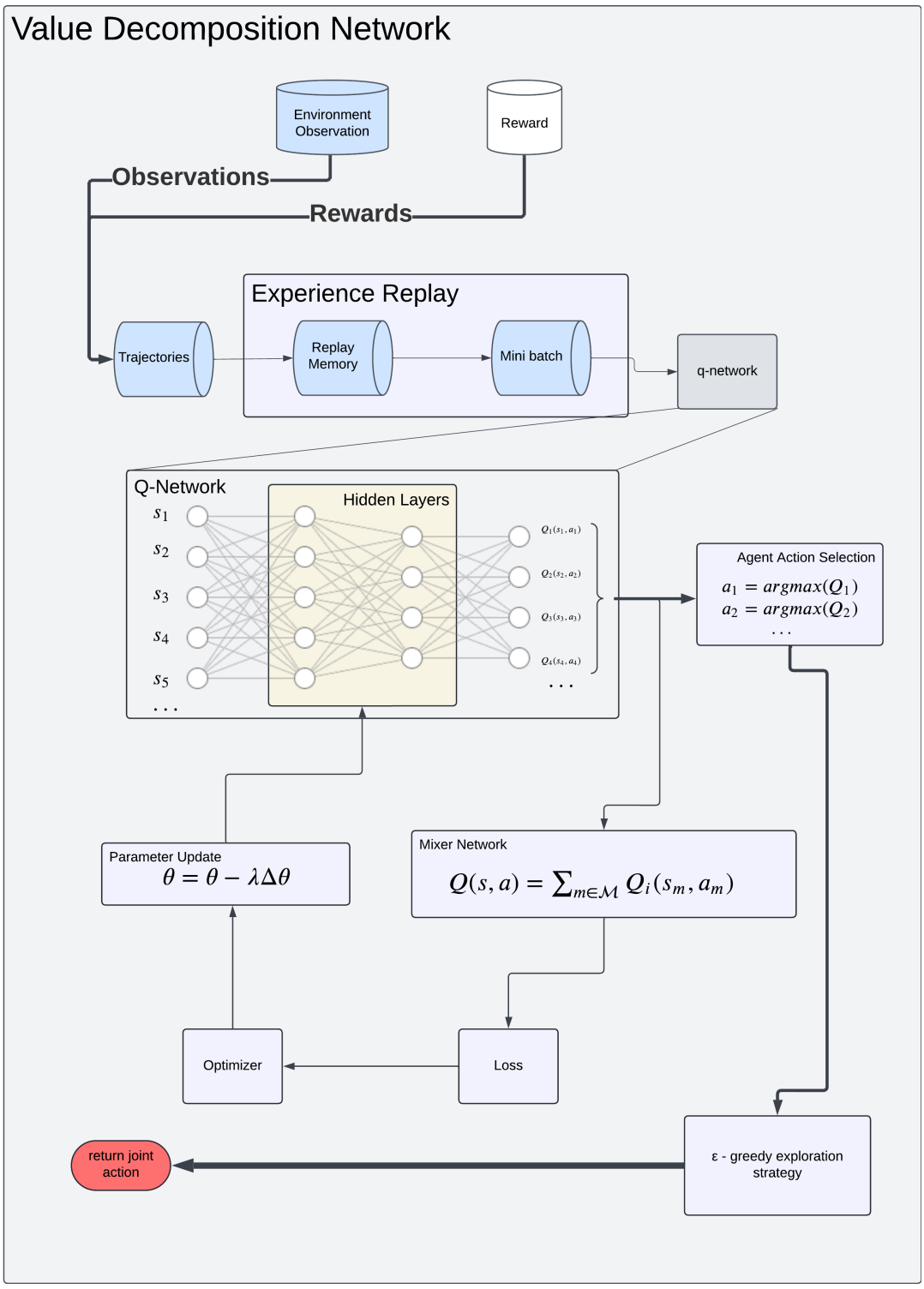


Figure 3.14: Simplified architecture of Value Decomposition Network with Parameter Sharing.

4

Case Studies and Results

4.1. Environment Setup

The proposed methodology involves developing a simulatation of a post-earthquake repair scheduling decision process as an MDP; it then aims to use MARL algorithms for finding favourable repair schedules. The developement of the simulation is made using two principal testebeds *toy-city-4* and *toy-city-30*. They have 4 and 30 components respectively with equal numbers of buildings and roads. The size of the two environments is principally governed by the associated runtime of an MARL experiment. It is noted that MARL experiments in similar research usually converge after completing 10^5 to 10^6 rollouts. Thus, the computation of a single rollout is a key limiting factor and requires careful consideration. While the developed tool allows for the development of environments using data from NSI, OSM and NBI, it is not used in MARL as the computation time of larger cities or towns can take up to a minute.

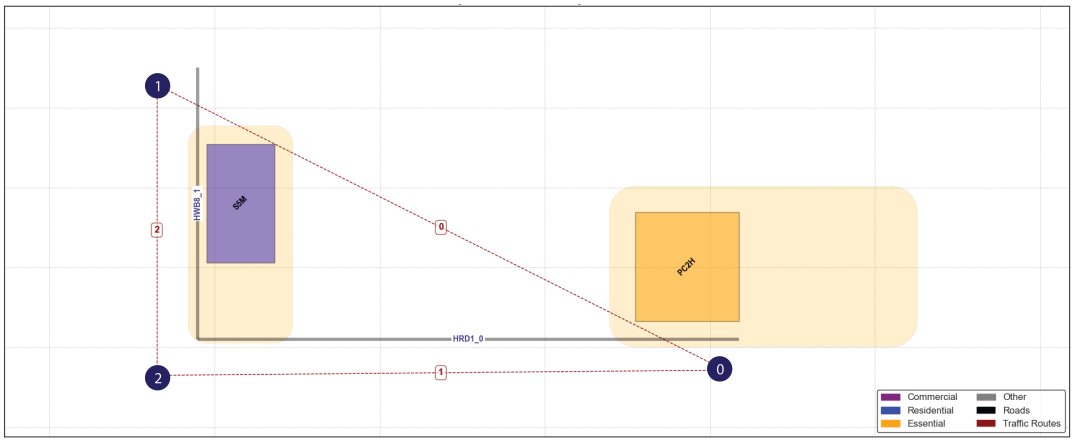
The two test beds along with the rest of the experiment are simulated using Python 3.9.20 in a conda environment [51]. A conda environment allows reliable and reproducible dependency management. The testbeds are passed into the environment as *GeoDataframes* using the package *pandas*, [60]. While many other packages are used in the implementation, only principal packages will be referenced in this thesis. Snippets of important code segments are referenced in the Appendix; a repository of the working code for the project is available as a GitHub repository at `https://github.com/Antonios-M/qres_marl`.

Conceptually, the two tested environments have two different aims. The smaller environment is aimed at being a baseline environment which can be easy to validate and test, while not focusing too much on making interesting spatial arrangements between components. Conversely, given the increase in the number of components, the environment with 30 components is aimed at isolating the critical facilities of the community from the commercial / residential zones by using two bridges. This is done so that the recovery scenario introduces certain context-specific conditions that might produce interesting recovery results. Some attribute metrics of *toy-city-4* are listed below; full tables of attributes for both environments can be found in the appendix. The two occtypes of the buildings correspond to a mid-rise residential and hospital buildingd respectively. *E-Facility* is an essential facility boolean.

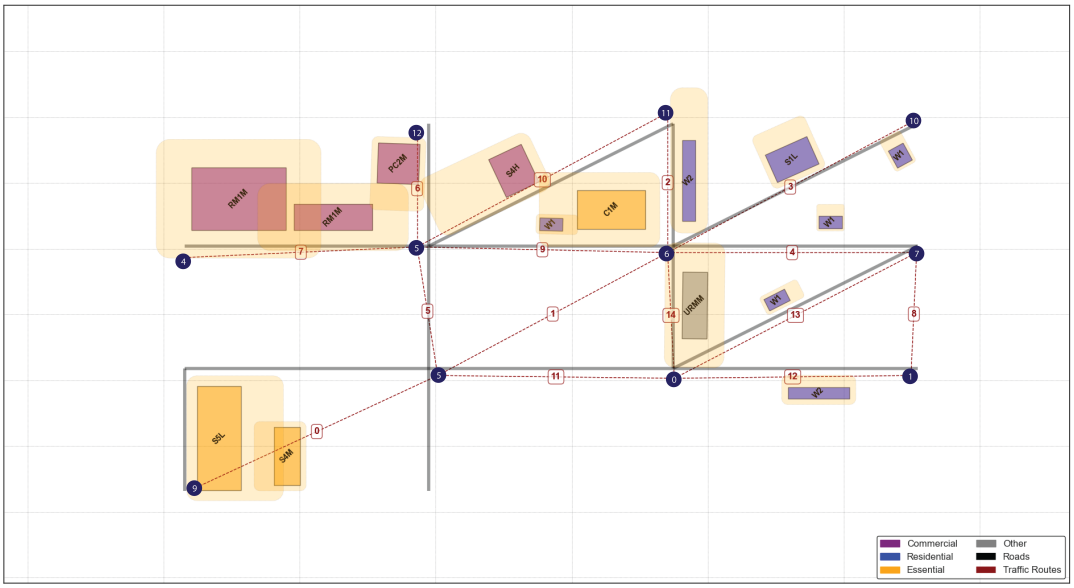
Index	Occ. Type	Dwell Units	Struct. Typ.	Appr. Value	E-Facility	Sq. Ft.
0	RES3E	74	S5M	55.2 Mill.	FALSE	59417
1	COM6	0	PC2H	102.8 Mill	TRUE	110683

Table 4.1: Attributes of buildings (excluding geometry)

An important nuance of making the testbeds is generating trip data for the associated traffic network. Basline link data such as free-flow-speed and capacity are taken from Ben Stabler's networks and matched to the environment's lengths. Trip generation is usually made using trip-generation data and methods such as ITE's trip generation handbook [93]. However, given the combined scope of



(a) *toy-city-4*, an environment with 4 components, 2 roads and 2 buildings, made to provide the practically smallest environment formulation for validation and testing.



(b) *toy-city-30*, an environment with 30 components, 15 roads and 15 buildings, with two bridges connecting the essential facilities to the rest of the community.

Figure 4.1: Case Study testbeds used to test MARL for post-earthquake repair scheduling, Author's Own Work..

Index	Length (mi)	Length (km)	Hazus R	Unit Cost
0	0.11	0.17	HRD1	6992
1	0.056	0.09	HRD1	6992

Table 4.2: Road segment attributes: length, Hazus type, and costs

environment modelling and MARL this thesis instead proposes the generation of trips manually, using reasonable assumptions about daily trips. Conceptually node 0 is the origin of trips of people leaving the hospital, either patients or health workers. Trips originating from node 1 are people leaving their apartment to go to the hospital or any other part of the network. That is, assuming traffic could be originated in nodes which are not included in the environment. Thus the following daily trips are defined:

Init Node	Term Node	Demand
0	1	600
0	2	400
1	2	200
1	0	400

Table 4.3: *toy-city-4* O-D trip table

The interdependencies between buildings and roads are first found, following by the division of the two networks in component-level objects; these are Python classes *Building* and *Road*. They are implemented to run repair actions on single road or building components and then combine the resulting component states and functionality metrics into joint states and rewards. The interdependency of roads to buildings in *toy-city-4* is matrix $A_{r,b}$:

$$A_{r,b} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{4.1}$$

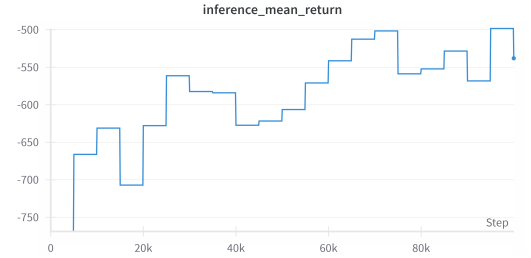
The resulting environment consists of a set of road and building $Python$ objects and csv files for the traffic network and demand. They are all passed in as arguments to a custom *Gymnasium* environment, which acts as the MDP. $Gymnasium$ is an open-source library for Single- and Multi-Agent DRL and is standard for modelling DRL environments. Each time the environments *.reset()* function is called, a new earthquake is randomly chosen and all losses are initiated, e.g income loss, repair time etc.. Actions are passed into the environment's *.step()* method which checks if the total repair effort exceeds the number of crews, in which case a random ranking assigns the top *n-crews* repair actions. These actions are passed through individual *Building* and *Road* objects, where they are processed considering the dynamics of the environment. For instance, if a repair actions is requested on a road which has debris on it due to interdependent buildings, then a *do-nothing* action is taken instead. The environment's *.step()* function returns a tuple of states, reward, information, termination and truncation conditions. Termination is *True* if all components are repaired and truncation is *True* if the time horizon is reached.

Considering the complex nature of the environment, many aspects of it were not acting as expected when conducting $MARL$ training. Thus, *toy-city-4* was iterated over 8 versions to *debug* and fix certain environment aspects that hindered training. The table below shows the differences between environment versions:

Parameter	v1	v2	v3	v4	v5	v6	v7	v8	v9
Number of Agents	4	4	4	4	4	4	4	4	30
Number of Crews	4	4	4	4	4	4	2	2	10
Time Horizon	20	20	20	20	20	20	20	50	100
Time-step Duration	20	20	20	20	20	20	20	40	30
Compute Debris	false	false	false	false	true	true	true	true	true
Trucks per Bldg per day (debris)	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Bldg Stochastic DS	false	false	true	true	false	true	true	true	true
Bldg Stoch. RT	false	false	false	true	false	true	true	true	true
Bldg Stoch. RC	false	true	true	true	false	true	true	true	true
Bldg Stoch. Inc. Loss	false	true	true	true	false	true	true	true	true
Bldg Stoch. LOF Time	false	true	true	true	false	true	true	true	true
Bldg Stoch. Reloc Cost	false	true	true	true	false	true	true	true	true
Compute Debris (Capacity Red.)	false	false	false	false	false	true	true	true	true
Road Stoch. DS	false	false	false	false	false	true	true	true	true
Road Stoch. RT	false	false	false	false	false	true	true	true	true
Road Stoch. RC	false	true	true	true	false	true	true	true	true

Table 4.4: Progressive changes in environment configuration from v1 to v9. All cells start light blue; darker cells indicate changes. DS = damage state, RT = repair time, LOF = Loss of Function time, Reloc Cost = Relocation Cost, Stoch. = Stochastic

(a) Episodic returns during training for environment version 9, logged every 100 training episodes



(b) Episodic mean returns over 500 inference steps for inference during training for environment version 9

Figure 4.2: Training plots for environment version 9. Y-axis shows cumulative resilience losses (negative) per episode (episodic returns). Training episodic returns show that learning was slow and that returns were only beginning to increase after training time-step 90000, however returns during inference show a clear upwards trajectory. Given that this run took 35 hrs to complete, it was exceedingly difficult to run training for more timesteps.

4.2. Results

Value Decomposition Network with Parameter Sharing ($VDN-PS$) was tested on all environment versions and Q-Mixer with Parameter sharing ($QMIX-PS$) as well as Deep Centralised Multi-Agent Actor Critic ($DCMAC$) were tested on version 8. Following this experiment on *toy-city-4*, $DCMAC$ was tested on *toy-city-30* as it yielded the best results for *toy-city-4*. While specific experiment attributes are given in the results and the appendix, all the experiments were run for $100K$ training time-steps. The experiments were run on a laptop with an available GPU. Training included parallel inference of 500 time-steps every 1000 training time-steps with a total of 5000 inference timesteps. The GPU used is a NVIDIA GeForce RTX 4060 Laptop GPU with a boosted speed of 2300 MHz and 8GB of available memory. Parallel inference was run using 2 workers as more workers caused *MemomoryError* exceptions. While most MARL algorithms showed learning, none of them were able to converge within the given configurations. This is because $100K$ time-steps were the highest number that was feasible to simulate and took no less than 30 hours of run-time. In essence, this is to say that the algorithms did show monotonic improvement but require more training time-steps to fully converge; given the available hardware this was not possible.

The configurations of of the three tested algorithms was constant across both environments and is described in the following table:

4.2.1. Environment I: 4 components

The resulting learning performance for $VDN-PS$, $QMIX-PS$ and $DCMAC$ on *toy-city-4* was tested by running 1000 inference rollouts using the three algorithms and importance-based ($IMPB$) as well as

Parameter	VDN_PS	QMIX_PS	DCMAC
Num Episodes	100,000	100,000	100,000
Num Inf. Episodes	500	500	500
Inference Freq.	1000	1000	1000
Max Memory Size	10,000	1,000	500,000
Batch Size	64	64	64
Discount Factor (γ)	0.99	0.99	0.99
Network Hidden Layers	[64, 64]	[64, 64]	Actor: [32, 32], Critic: [64, 64]
Optimizer	Adam	Adam	Actor: Adam, Critic: Adam
Learning Rate	0.001	0.001	Actor: 0.0001, Critic: 0.005
LR Scheduler	Linear	Linear	Linear
LR Scheduler Iters	10,000	10,000	20,000
Start Factor / End Factor	1 / 0.1	1 / 0.1	1 / 0.1
Exploration Strategy	Epsilon-Greedy	Epsilon-Greedy	Epsilon-Greedy
Epsilon Max / Min	1 / 0.005	1 / 0.005	1 / 0.005
Epsilon Decay Episodes	10,000	10,000	20,000

Table 4.5: Hyperparameters for VDN_PS, QMIX_PS, and DCMAC Algorithms

random policies. The results are shown below:

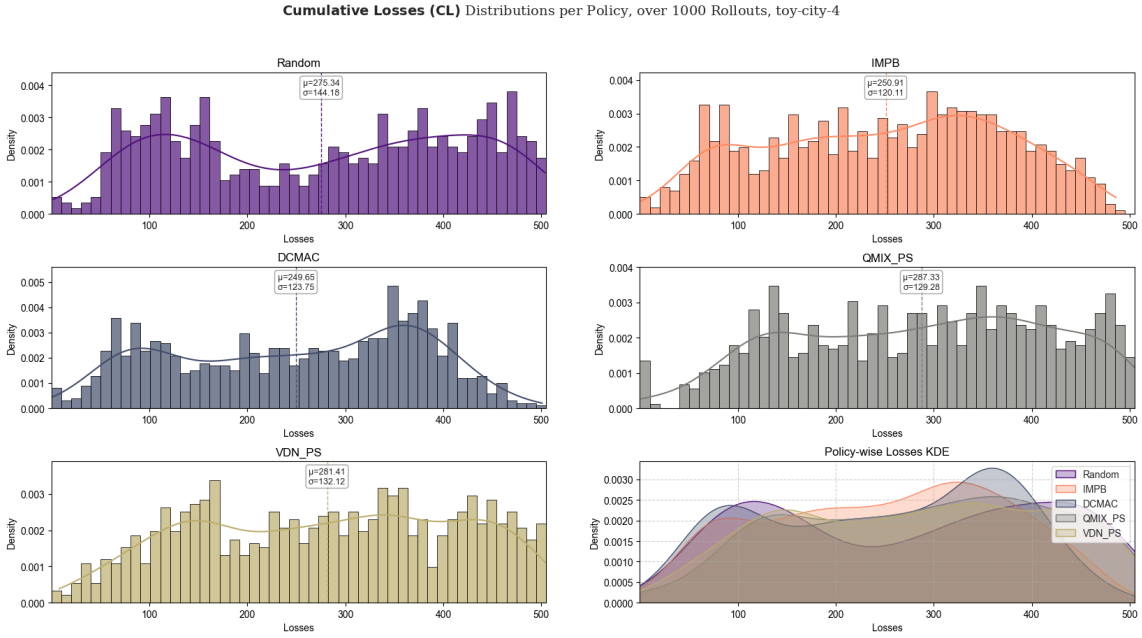


Figure 4.3: Cumulative Losses to full recovery per policy for 1000 rollouts on *toy-city-4* (lower is better)

The charts show histograms of cumulative losses on the x-axis and relative density on the y-axis for the 1000 rollouts over the 5 tested policies. Density is chosen as the y-variable as it allows comparison between policies. The results show expected relative results between $DCMAC$ and the other two algorithms as $DCMAC$ performs better than the other two. Conversely, $QMIX - PS$ performs slightly worse than $VDN - PS$ which is not expected; however the difference is negligible and could just be due to sampling frequency differences.

On average, $DCMAC$ essentially matches the performance of importance-based scheduling. This doesn't say much on the specific recovery actions chosen, however it is a starting point to gauge the overall performance of the policy. Additionally, $DCMAC$ tends to show higher frequency over the extremes, i.e more high-performing instances but also more low-performing instances, and less average-performing

instances. This is most likely due to the fact that learning did not fully converge. Interestingly, $VDN-PS$ and $QMIX-PS$ both perform worse than random policies, this is also unexpected. However, given that there are only two possible actions and only 4 components, a random policy might in fact be quite performative as the action-space is quite narrow. Additionally, there are no direct penalties for choosing more repair actions than there are available repair crews. Considering that all components usually have a positive repair time at initialisation due to the earthquakes being above 6.0 M , a random action is then almost certainly going to result in at least one successful repair action as the only case it wouldn't is if all four actions chosen are *do-nothing*, which on average should happen only 6.25% of the time (0.5^4). Likewise, even if the random policy results in more repair actions than there are available repair crews, the agents incur no penalties for doing so.

Looking closer into the two high-performing policies, certain conclusions about the optimality of DRL as opposed to rule-based repair scheduling policies can be drawn. Firstly, $IMPB$ is illustrated in figure 4.4, showing a functionality-time curve for the whole network and loss-time curves for each component. For illustration purposes the various losses shown in the component-level graphs are normalised to the damage state extremes of that component. For instance, the residential building starts at a damage state of 4 and so all its loss attributes begin at 4 as well and reduce to 0 as it is repaired.

Considering the environment has 2 available repair crews at each time-step, there is a clear prioritisation of the highway road and the hospital to be repaired before the other two components, which is seen in the repair actions taken up until time-step 15. This is primarily as a consequence of the interdependency of the hospital to the road and the residential building to the bridge. In this case, both buildings are at a damage state above 3, which is the threshold for which debris is generated. Thus, neither of the roads can begin to be repaired until their interdependent buildings are clear of debris. Additionally, the hospital is chosen to be repaired before the residential building because of its contribution to both critical as well as healthcare functionalities. Once the hospital's debris is cleared at time-step 4, repair of the road begins. The road is repaired until time-step 15 when the bridge is instead chosen to be repaired due to now having a higher damage state than the road. However, this results in *do-nothing* actions to be taken on the road as the mid-rise residential building is not yet clear of debris. This is a limitation of $IMPB$ as it does not consider the interdependencies as an explicit rule. Once the hospital is repaired, the residential building and bridge are chosen to be repaired at time-step 27. The effect of the bridge repair can be clearly seen in the traffic delay cost curve, which drops drastically from time-step 27 to time-step 33. The bridge is then completely repaired at time step 40 and the residential building at time-step 48, after which the hospital keep gaining its remaining income until time-step 65 when the rollout terminates.

Some key takeaways from this policy can be drawn. First, the traffic seems to solely rely on the bridge and not at all on the road. This is due to the $O-D$ matrix used for traffic assignment:

init_node	term_node	demand
0	1	500
0	2	150
1	2	150
1	0	800

Table 4.6: Node demand table

Considering the roads and the traffic perofrmance, out of a total of 1600 trips, only 150 of them rely only on traffic nodes 1 and 2. The trips going from nodes 0 and 2 also partially rely on the traffic link between 1 and 2; however, the vast majority of trips are from 0 to 1 and back, specifically 1300 trips. The decision to model the traffic as such was made to weight the road connecting the hospital to the residential building more heavily than the bridge. However, it can be said that this was perhaps exaggerated. Nevertheless, while this is easy to point out in such a small network, the task of constructing a custom $O-D$ matrix for a custom traffic network is exceedingly difficult.

Conversely, when considering the two buildings, there is a clear bias to repair the hospital first. This can be intuitive considering the community functionality weights used in this simulation:

- Economic Functionality Weight, $w_{econ} = 0.5$

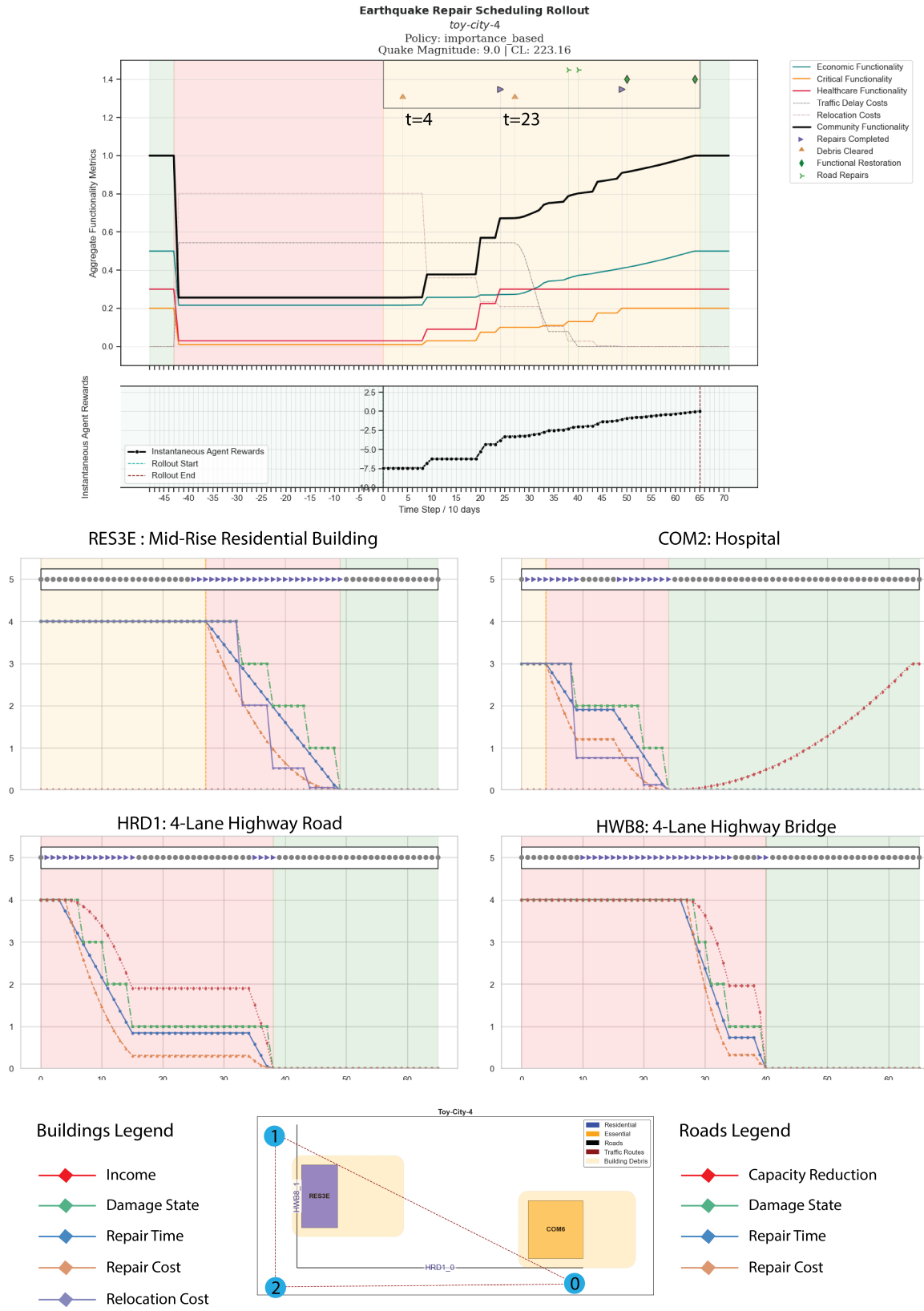


Figure 4.4: One rollout of using *IMPB* on *toy-city-4* with an earthquake of 8.5 M

- Critical Functionality Weight, $w_{crit} = 0.2$
- Healthcare Functionality Weight, $w_{health} = 0.3$

The residential building only contributes to the economic functionality and not the other two, while the hospital contributes to all sub-system functionalities. This makes the clear bias for the hospital clear to understand. This bias is reasonable and also favourable as hospitals should generally be repaired before residential buildings. However, this weighting is only appropriate as the share of critical to non-critical facilities is $50 - 50$, which is not the case for most metropolitan areas. For instance, a metropolitan area in the US and other parts of the world might have non-essential to essential building ratios closer to $90\% - 10\%$.

When comparing this policy to the one learnt by $DCMAC$, it is clear to see the discrepancies outlined above. Figure 4.5 shows a rollout for an 8.5 M earthquake using $DCMAC$ as the solver. The resulting losses are more than the ones yielded from $IMPB$ for an initial residual community functionality of around 0.21 after earthquake impact. This shows that $IMPB$ does perform better in this case. However, this is principally due to the fact that the $DCMAC$ policy does not attempt to repair the residential building or the bridge at all. The reason for this behaviour is not easy to pin-point, however it is likely a combination of not achieving full training convergence and incurring less significant rewards for repairing the residential building and the bridge.

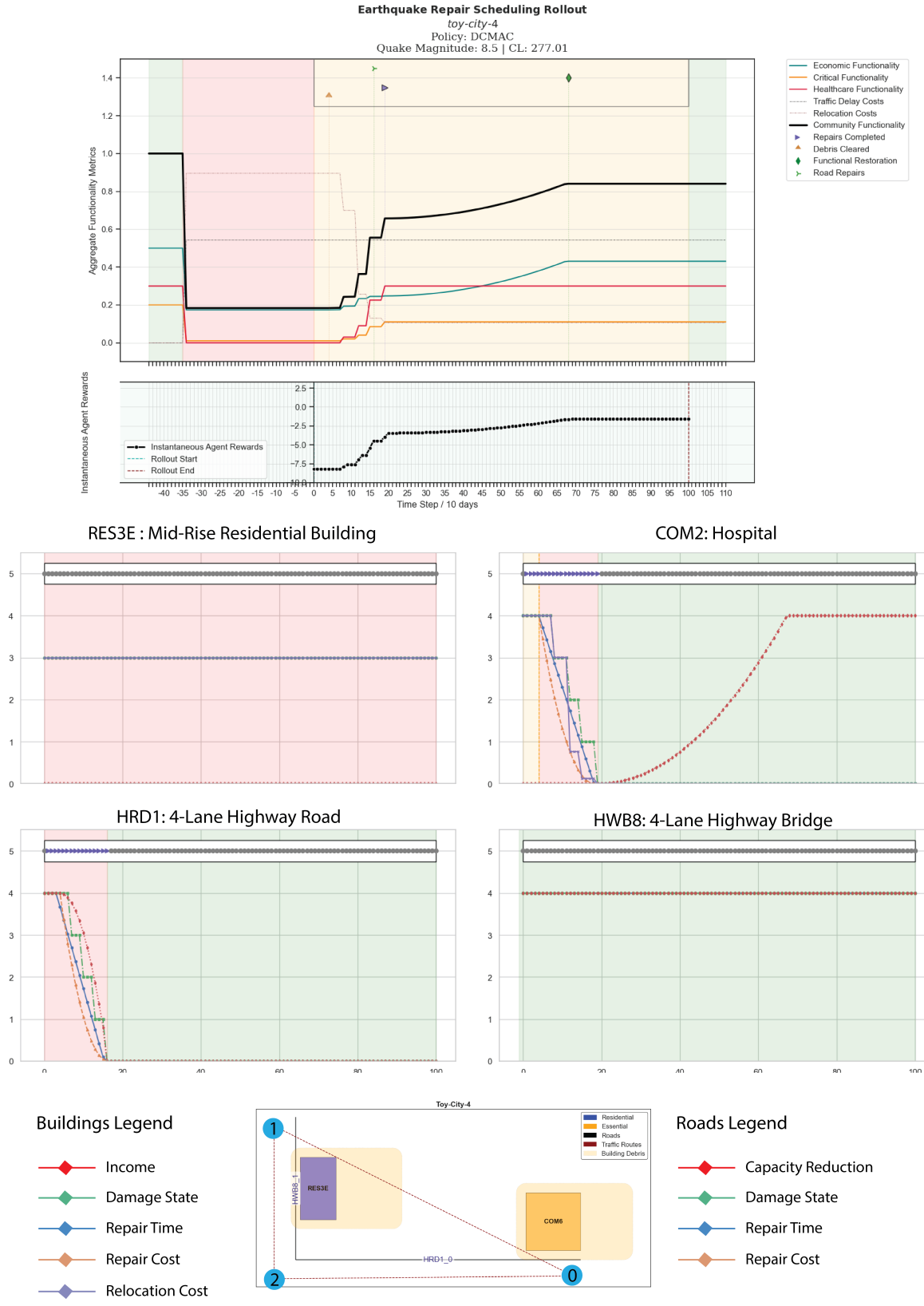


Figure 4.5: One rollout of using *DCMAC* on *toy-city-4* with an earthquake of 8.5 M

Regardless of the convergence of training, *DCMAC* is able to recover approximately 74% (0.21 → 0.81)

of community functionality (CF), while not repairing half of the network components at all. This prioritisation and lack of action on the residential building and bridge can be seen in the inability of the algorithm to fully repair functionality, reaching 0.81, and flat-lining until truncation at time-step 100. Nevertheless, the profile of recovery is clearly steeper, as at around time-step 20, the $DCMAC$ policy has recovered from 0.21 to 0.75 CF, while $IMPB$ has recovered from 0.21 to 0.5 for the same time increment. It should be noted that when looking at the illustrated results, the repair times for the hospital and the road might seem shorter in $DCMAC$; however, in both $IMPB$ and $DCMAC$ the two components are repaired within 20 to 25 time-steps, which is around 200 to 250 days. The apparent skewness of the repair profile is due to the $DCMAC$ rollout continuing for 100 time-steps, while the $IMPB$ rollout terminating at 65 time-steps.

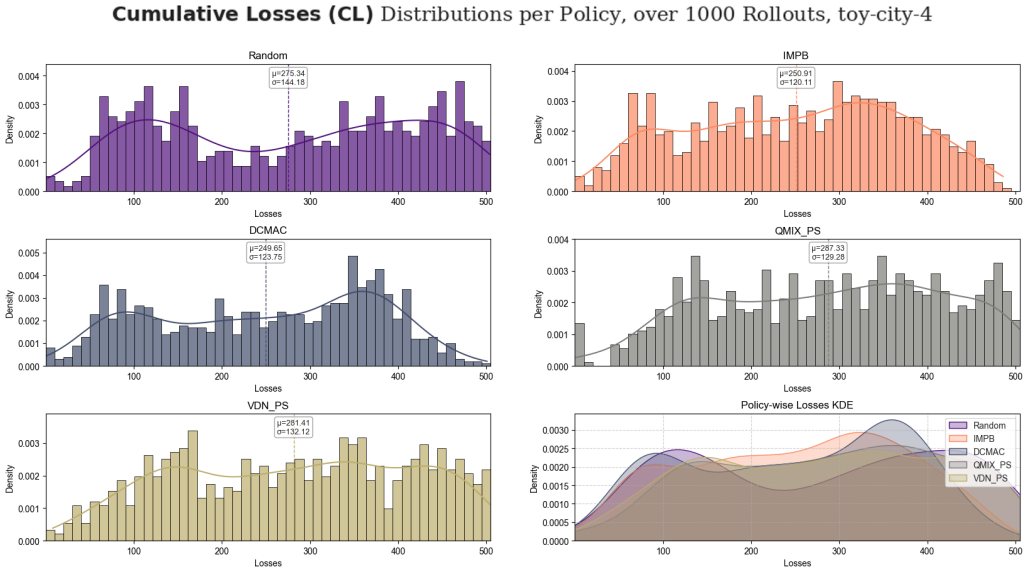


Figure 4.6: Density-Loss histograms for toy-city-4 of the tested policies when considering all incurred losses (lower is better)

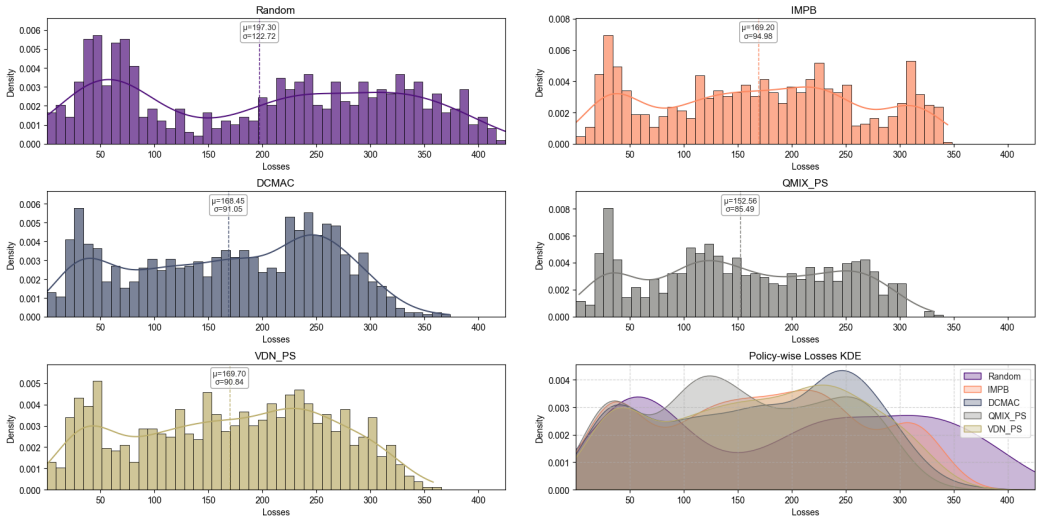


Figure 4.7: Density-Loss histograms for toy-city-4 of the tested policies when considering losses incurred until 70% recovery is achieved (lower is better)

One can see the general benefit of using DRL by comparing the $N-\%$ recovery when DRL against $IMPB$. Figures 4.6 and 4.7 show histograms of cumulative losses and 70% recovery losses respectively for the tested policies. This confirms the observation that $DCMAC$ generates a steeper early recovery curve. Interestingly, $QMIX-PS$ is the best performing algorithm here; however, it should be noted that $QMIX-PS$ and $VDN-PS$ were trained for $100K$ time-steps, while this version of $DCMAC$ was only trained for $40K$ due to time constraints. Furthermore, another interesting takeaway from both sets of results is that both $DCMAC$ and $QMIX-PS$ seem to show a lower density of *average* performing recovery curves, with higher densities of both low and high performing results.

4.2.2. Environment II: 30 Components

Following the promising results on *toy-city-4*, only $DCMAC$ is tested on *toy-city-30*, the environment with 30 components: 15 roads and 15 buildings. The decision to test $DCMAC$ over $QMIX-PS$ is made regardless of $QMIX-PS$ having a seemingly higher performance to $DCMAC$. This is because $DCMAC$ should generally perform better as is seen across literature; the relatively low performance of $DCMAC$ to $QMIX-PS$ in *toy-city-4* could just be due to training $DCMAC$ for less time-steps.

Looking at the general performance of $DCMAC$,it falls short of $IMPB$ when comparing cumulative losses for full network recovery as can be seen in figure 4.8, where $DCMAC$ has a mean loss of recovery of 359 and $IMPB$ having a mean of 311. In that case, $DCMAC$ only performs slightly better than a random policy when looking at full recovery. Conversely, when looking at 70% recovery in figure 4.9, $DCMAC$ shows a mean recovery loss of 187 with $IMPB$ having a mean of 265, which is actually higher than a random policy at 250. The KDE (Kernel Density Estimation) plot illustrates these discrepancies clearly; $IMPB$ shows a higher density of results around $180-200$ mean losses, but it also shows a higher density of results above 500 losses when compared to a random policy. $DCMAC$ outperforms both of them and the KDE plot shows a clear increse in density for losses below 300.

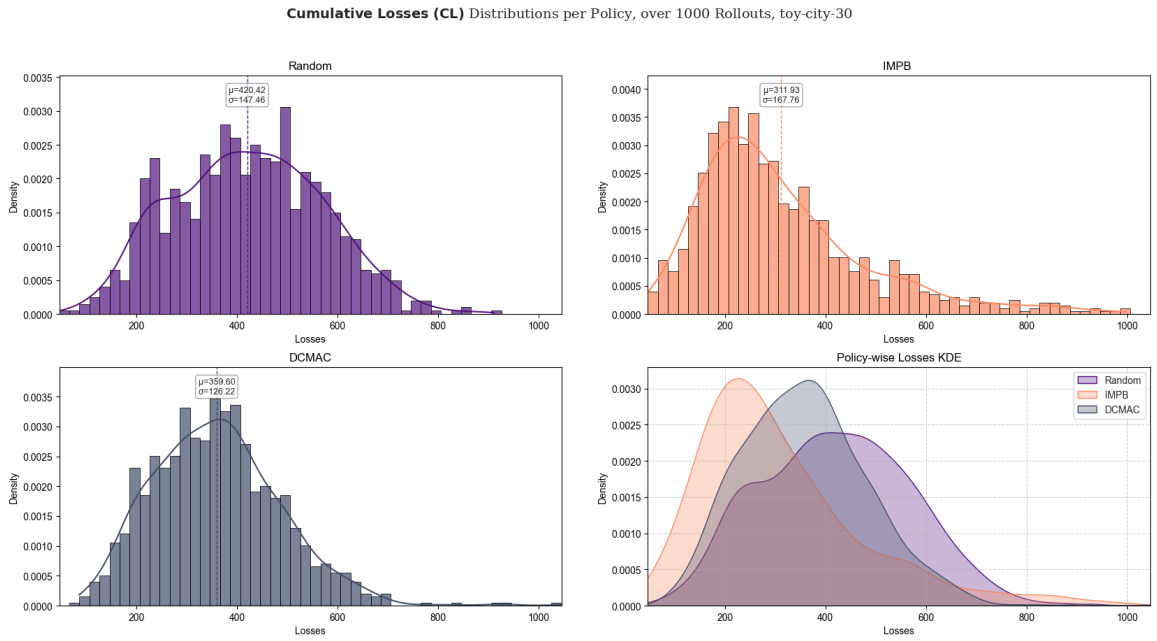


Figure 4.8: Density-Loss histograms for toy-city-30 of the tested policies when considering full recovery losses(lower is better)

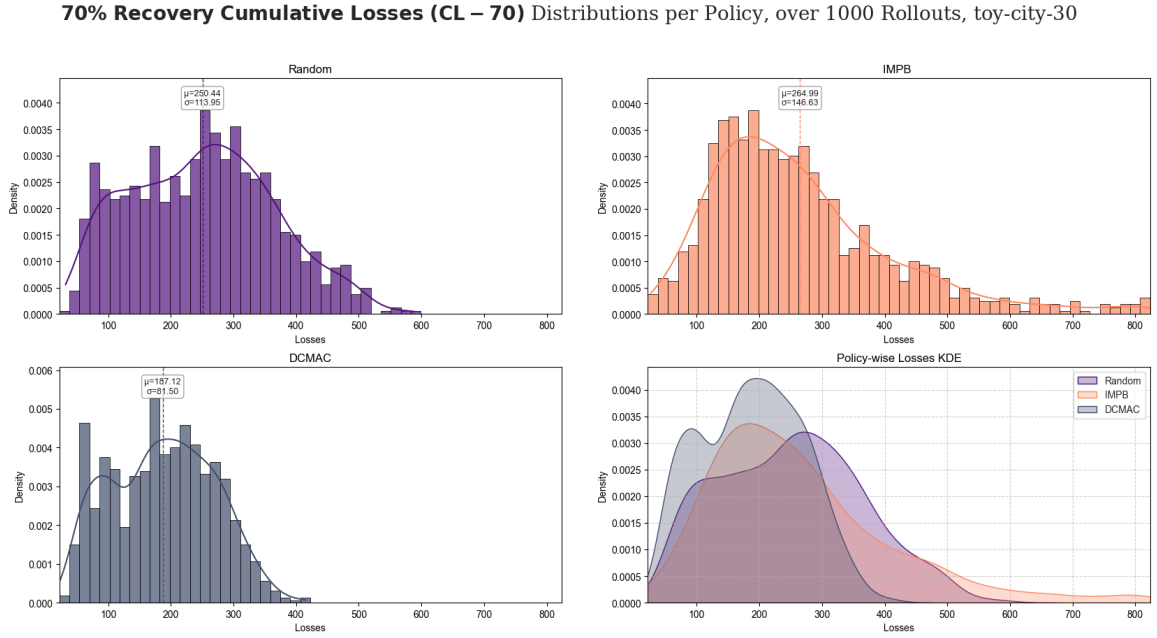


Figure 4.9: Density-Loss histograms for toy-city-30 of the tested policies when considering losses incurred until 70% recovery is achieved (lower is better)

This confirms the ability of $DCMAC$ to prioritise early recovery as is seen from the equivalent results from *toy-city-4*. It also points out that $IMPB$ is exceedingly inadequate at effective early recovery as the network increases in size and complexity. This behaviour is primarily due to $IMPB$ not considering the interdependencies of buildings to roads explicitly and attempting to repair roads that have debris on them before clearing the debris itself. This can be illustrated by looking at a sample rollout when using $IMPB$. Figure 4.10 shows a rollout for an earthquake of 8.0 M and an initial CF impact at 0.32. The recovery curve shows a very slow initial recovery until time-step 8 and then a swift trajectory until full recovery at time-step 29. In this case $IMPB$ only recovers the functionality from $0.32 \rightarrow 0.59$ in 10 time-steps. Conversely, when looking at $DCMAC$ in figure 4.11 for an identical initial impact at $CF = 0.32$, $DCMAC$ is able to recovery to 0.7 within 10 time-steps, showing a quicker early recovery. $DCMAC$ incurs total losses of 254 while $IMPB$ incurs losses of 238; however, it is clear to see that the majority of the losses when using $DCMAC$ are incurred after time-step 20.

In both cases, 10 repair crews were available at any one time-step. In the case of $IMPB$, during the first 10 time-steps the algorithm correctly repairs building 8, which is the fire-station ($GOV2\_8$). However, it spends a lot of the repair effort on attempting to repair roads $0, 10, 11, 9$; these roads are all interdependent to buildings that have debris on them ($DS \geq 3$), namely buildings $6, 5, 13, 4$. Interestingly, the relocation costs of building 0 are constant for all damage states apart from $DS = 0$ as defined in $HAZUS$. Building 5 is defined a college dorm residential building. $IMPB$ is then seen to have a clear inadequacy and leaves a lot of room for improvement in future work. Certain rules could be incorporated to account for these interdependencies; however, it points to an interesting limitation in rule-based decision making. The limitation lies in the fact that rules on how interdependencies should be treated need to be explicitly defined; as the number of considered networks grows and the interdependencies become more complex, this can quickly become a non-trivial problem to solve as the exact cascading effects of one component are hard to predict using explicit rules.

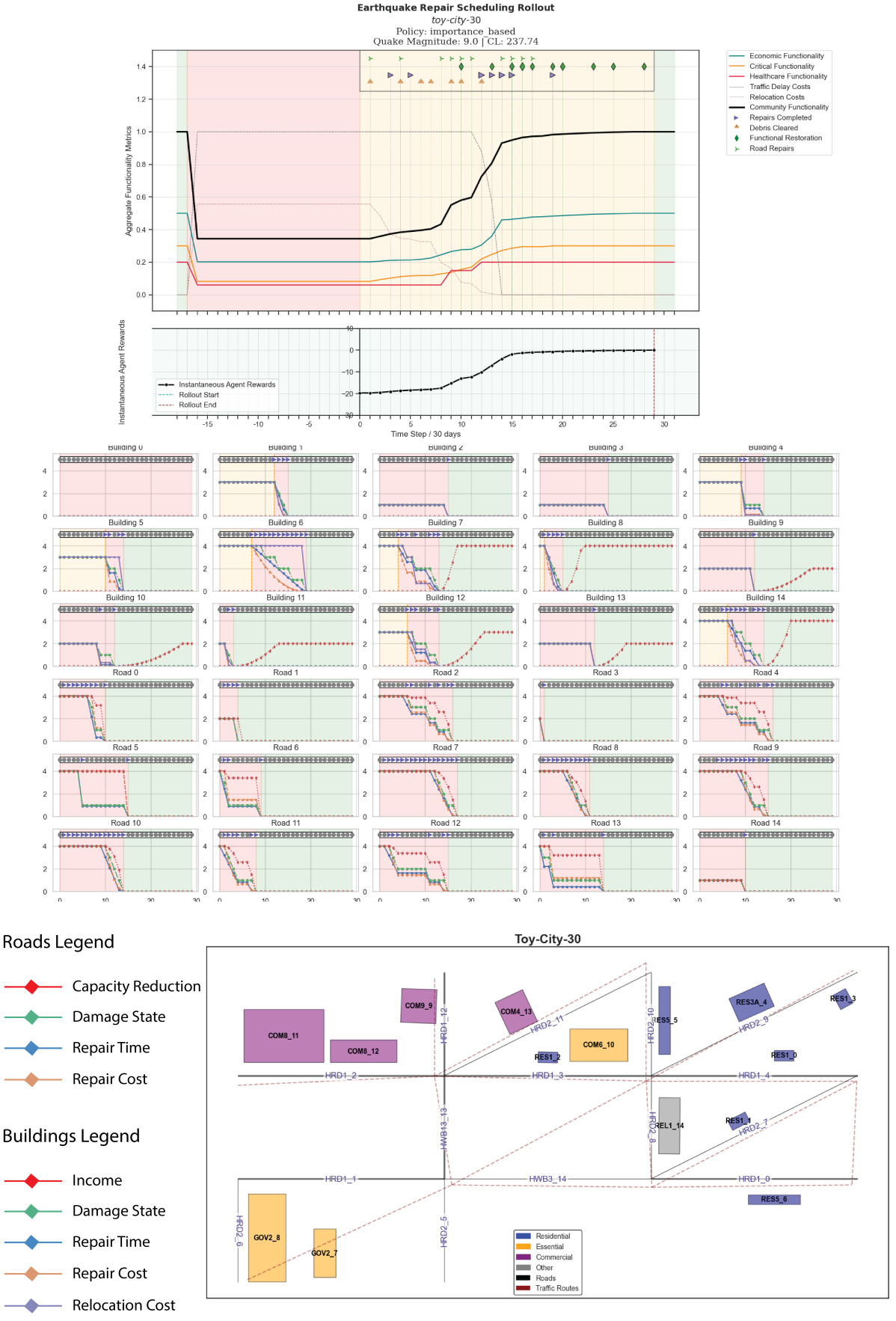


Figure 4.10: One realisation of *IMPB* on *toy-city-30* with an earthquake of 9.0 M

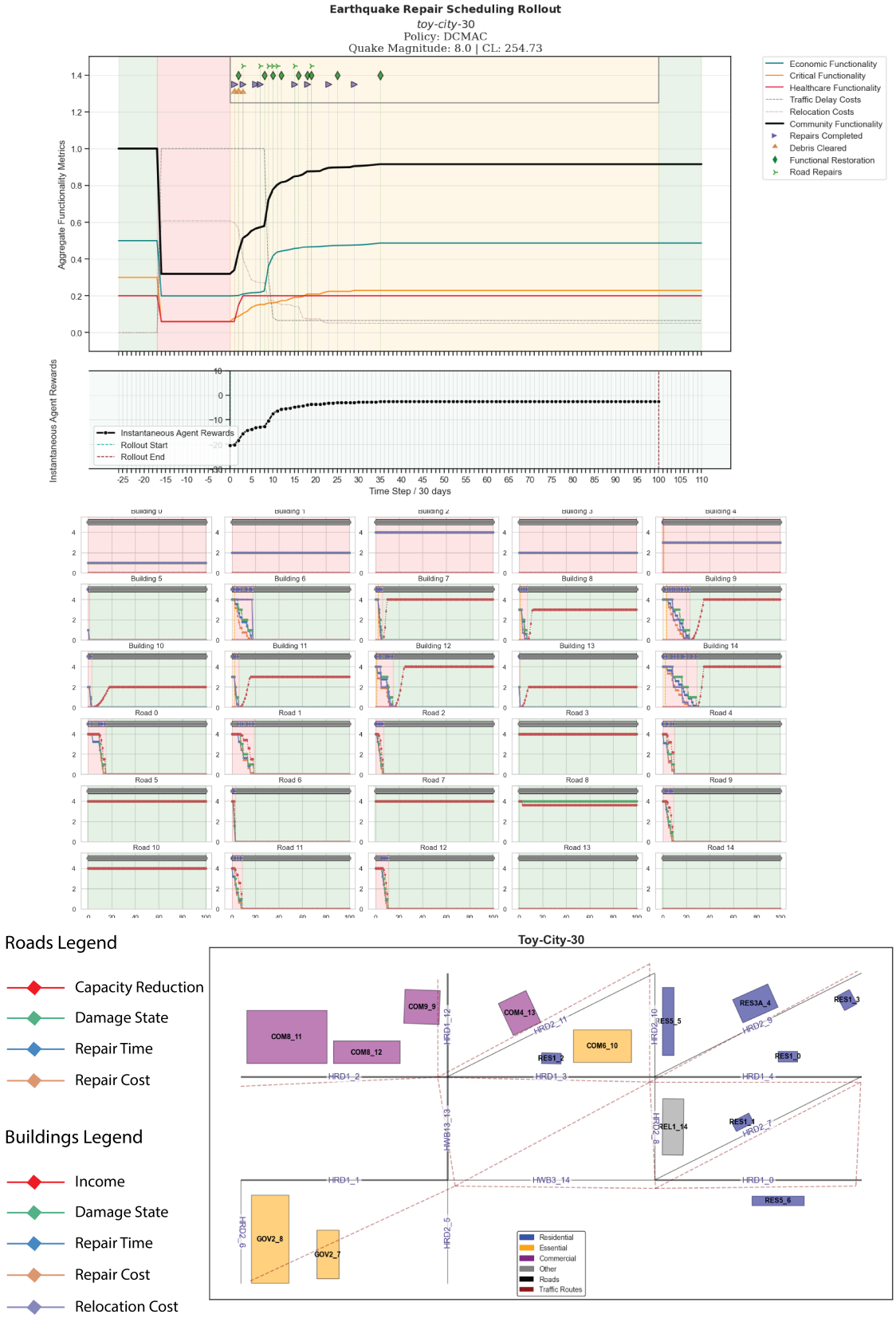


Figure 4.11: One realisation og *DCMAC* on *toy-city-30* with an earthquake of 8.0 M

Contrasting $IMPB$, $DCMAC$ chooses actions that result in more effective early recovery. Specifically, buildings 7, 8 and 9: the police station, the fire station and the hospital are repaired first along with the large commercial and office buildings: 9, 11 and 12. Interestingly, when considering the traffic network, $DCMAC$ makes an effort to clear the debris of the church: building 14 so that the capacity of road 8 increases from 0.0 to 0.9. Following that $DCMAC$ does not make an effort to repair the road itself. Furthermore, the bulk of the traffic delay cost is recovered at time step 10. This is as a result from the repair of roads $0, 1, 2, 4, 9, 11$ and 12. These are the roads that contribute the most to the traffic delay cost as after their full repair the cost drops from $1.0 \rightarrow 0.1$. The lack of convergence to full recovery is then as a result of the dis-repair on buildings: $0, 1, 2, 3$ and 4 as well as roads: $3, 5, 7, 8, 10$. The 5 buildings are the smallest of the residential buildings and only contribute to the relocation costs of economic functionality; therefore, their repair yields very little rewards. Considering the roads, road 5 is not involved in traffic assignment itself as it does not lie in any of the shortest paths between roads that map to the traffic links seen in dashed red lines. Road 5 was modelled to gauge this behaviour and whether $DCMAC$ would ignore it, which is the preferred policy and is what it seen in this rollout. Roads $3, 7, 10$ are central to the residential areas and if they are not repaired the trips between residential and commercial as well as essential areas are hindered. All though these roads have 0.0 capacity, these trips are facilitated through road 8, whose capacity is partially recovered early on by repairing the debris of the church, which in itself does not generate enough income to be otherwise prioritised. The alternative of this is aiming to repair the roads around the residential area by clearing the debris of the small residential buildings. This is clearly not chosen by $DCMAC$ as these buildings remain in dis-repair for the remainder of the rollout.

This points at the discrepancy of the action space. Only two actions are available to the agents: *repair* or *do-nothing*, this makes the agents less able to differentiate between only clearing the debris of the residential buildings and functionally repairing them. Furthermore, the reward for eradicating the remainder of the traffic delay cost is not enough to incentivise the collective debris clearance of the small residential buildings. While an addition of a debris clearance action would probably help the agents differentiate between the two actions, the traffic delay cost curve also points at a poor $O - D$ matrix. That is to say that there is not enough traffic delay cost being generated when the roads around the residential buildings are obstructed. For a detailed view on the $O - D$ matrix, the defined trips can be found in Appenix A, Table 7.4. More than 90% of the traffic delay cost is recovered when repairing the roads around the offices, commercial buildings, hospitals and fire and police stations. While this could be true for a real community, the ratio of traffic between the commercial / essential areas and the residential area would probably not be that drastic.

Conclusively, several key takeaways can be drawn from the results shown for both *toy-city-4* and *toy-city-30*. The takeaways are related both to the comparison between $IMPB$ and the three DRL tested algorithms, but also to the increase in network size between *toy-city-4* and *toy-city-30*. The takeaways are given below with a breakdown of the key result metrics given in **??** and illustrated in figure 4.12.

- DRL is exceedingly hardware-hungry and requires prolonged training time for full convergence, which is not reached in the trained agents tested in this thesis. However, partial training results in interesting and performative results when compared to $IMPB$.
- DRL performs poorly in terms of achieving full community recovery when compared to $IMPB$. Additionally, certain algorithms such as $VDN - PS$ and $QMIX - PS$ perform particularly poorly at achieving full recovery and even rank lower than a random policy.
- DRL performs better than $IMPB$ when comparing partial recovery. This is tested by measuring the recovery to 70% of initially impacted losses, where DRL outperforms $IMPB$ across the 2 environments and across 4 different algorithms: $DCMAC$, $QMIX - PS$ and $VDN - PS$, with $DCMAC$ performing better than the other two.
- $IMPB$ becomes increasingly worse at swift early recovery as the network size increases. Conversely, DRL becomes increasingly better as the network size increases

Num. Components	Policy	CL μ	CL-70 μ	CL σ	CL-70 σ	CL $\pm$ CI	CL-70 $\pm$ CI
4	Random	275.34	197.30	144.18	122.72	±8.94	±7.61
4	IMPB	250.91	169.20	120.11	94.98	±7.44	±5.89
4	VDN-PS	281.41	169.70	132.12	90.84	±8.19	±5.63
4	QMIX-PS	287.33	152.56	129.28	85.49	±8.01	±5.30
4	DCMAC	249.65	168.45	123.75	91.05	±7.67	±5.65
30	Random	420.42	250.44	147.42	113.95	±9.13	±7.06
30	IMPB	311.93	264.99	167.76	146.63	±10.39	±9.09
30	DCMAC	359.60	187.12	126.22	81.50	±7.83	±5.05

Table 4.7: Performance metrics by policy and number of components, including mean, standard deviation, and 95% confidence interval ($n = 1000$).

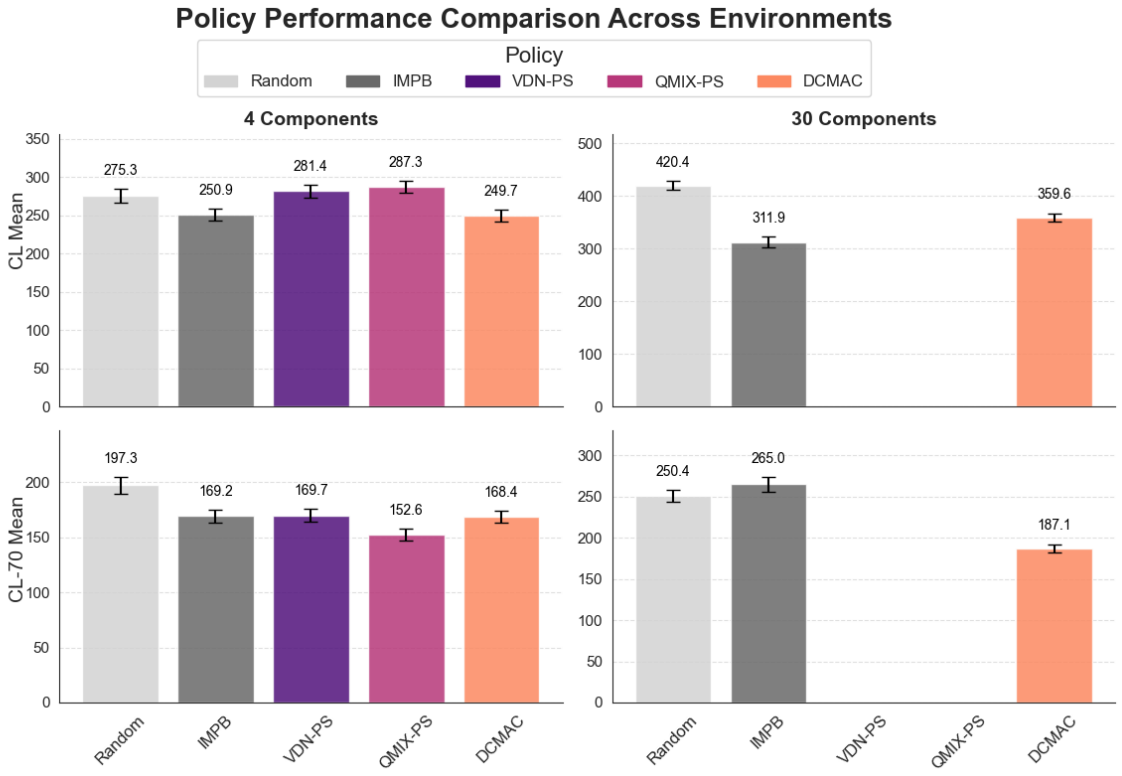


Figure 4.12: Conclusive results of VDN-PS (Value Decomposition Network with Parameter Sharing), QMIX-PS (Q-Mixer with Parameter Sharing), DCMAC (Deep Centeralised Multi-Agent Actor Critic), IMPB (Importance Based) and Random policies. Numbers above error bars are means and error bars are the 95^{th}% Confidence Interval. CL = Cumulative Losses to full recovey, CL-70 = Cumulative Losses to 70 % recovery (lower is better)

4.3. Discussion of Results

The results reveal a trade-off in the performance of DRL, given the non-convergence of learning. The agents consistently demonstrate high efficacy in prioritizing actions that lead to a rapid, early-stage recovery of network functionality. However, this focus on immediate reward maximization comes at the cost of achieving full, system-wide restoration, a goal more reliably met by heuristic baselines. The primary significance of this research lies not in presenting DRL as a superior solution, but in its utility as a powerful tool for auditing and challenging the assumptions embedded in human-designed, rule-based policies. Furthermore, while the issue of achieving full learning convergence can be mitigated using better hardware, the long training times for small networks indicate the need for such frameworks to be vectorised by using powerful numerical engines, like Google's JAX.

The results on the toy-city-4 network show that the Importance-Based ($IMPB$) heuristic, while effective in consistently achieving full recovery, may not represent the most efficient recovery pathway. Through DRL is is demonstrated that the majority of the early resilience loss can be minimised more swiftly by focusing on a select few components. This behaviour is valuable because it can reveal biases in the way we model environments as engineers and decision makers. While DRL methods can have their own biases, the approach in this thesis is designed to be minimally biased, as agents were given no explicit instructions on repair prioritization beyond the constraints of the environment itself. They learned to identify and exploit bottlenecks, such as the random ranking of components when multiple repair actions are chosen, thereby developing intelligent scheduling policies.

This emergent DRL behaviour can also be used as a diagnostic tool for the environment model. The agents' tendency to halt repairs after reaching approximately 70% recovery did not signify a failure to learn; rather, it signified a successful exploitation of a modelling discrepancy where the incentive for completing the final repairs was insufficient. Furthermore, the agents independently identified that certain components—such as a specific hospital and a key road link—governed a disproportionate amount of the network's overall performance. While this is true in any real network, the model allowed these components to be singled out almost exclusively. This highlights a fundamental challenge in resilience engineering: the difficulty of appropriately weighting the functionality of diverse community subsystems and ensuring that holistic resilience metrics do not inadvertently oversimplify complex interdependencies.

In a practical sense, as was seen from the results of *toy-city-30*, the agents were able to recover the majority of the traffic delay cost and other loss without repairing small single-family residential homes and several roads. Firstly, this points to a poor trip generation methodology as some roads did not contribute to overall traffic delay cost. Secondly, even though the small residential houses do not contribute much to the aggregate of community functionality, in a real disaster scenario they need to be repaired as well. Thus, the recommendation is to exploit the use of DRL for early-stage recovery because of the apparent effective prioritisation, but maintain a critical view of the recovery process by comparing and contrasting repair scheduling policies against existing importance-based approaches.

4.4. Comparison to State of the Art

A direct comparison of the numerical results of this work with existing literature is challenging due to two key areas of novelty. Firstly, the majority of research applying DRL to infrastructure management focuses on gradual deterioration over time, with very little work addressing the specific problem of post-earthquake repaier scheduling. Secondly, the work that does exist in the post-disaster context often frames the objective function and agent rewards in direct monetary terms. In contrast, this thesis introduces a dimensionless community resilience metric derived from several dimensioned, non-monetary and monetary subsystem metrics (e.g., number of available hospital beds, traffic flow capacity). However, some key points can be made; firstly, relating the work to Yang et al. which most closely resembles this research, this thesis makes several improvements in modelling a stochastic seismic hazard scenario, but also repair costs and losses, [96]. Conversely, the importance-based algorithm presented in this thesis was heavily influenced by the specifics of the environment and is thus difficult to directly compare it to literature. However, the general approach of $IMPB$ aligns with general ranking-based repair scheduling. Their results show a quicker early recovery and slower late recovery when using DRL, which matches this thesis's results. Interestingly, DRL was able to make decisions on repairing relatively *non-economically* important buildings such as the church just to increase the capacity of its interdependent road. This shows that DRL was able to, at least partially, perceive the interdependency between roads and buildings without having explicit instructions to do so via a graph or another relational state definition. In contrast, $IMPB$ was seen to perform increasingly worse for the larger environment as it attempted to make road repairs even if the debris of dependent buildings was not cleared, effectively delaying the repair process unnecessarily. This is principally due to $IMPB$ being developed with few rules, the work of Sediek et al. employs such heuristic algorithms that include a far deeper heuristic decision tree which, if implemented, could provide a stronger comparative baseline against which DRL could be tested, [82].

When looking at the general research in infrastructure management using DRL, this thesis confirms that $CTCE$ approaches such as $DCMAC$ tend to perform better than $QMIX-PS$ and $VDN-PS$ but are hard to scale to larger environments. A principal point of difference between this research and existing literature is the observability of the MDP. In most work that looks at deteriorating infrastructure environments, the agents hold a belief over possible states and the true state is *hidden*. Conversely, this thesis uses a fully observable MDP. While this might be viewed as more appropriate for post-disaster scheduling than for maintenance planning, the use of a $POMDP$ can generally describe real infrastructure management scenarios better. Furthermore, using a $POMDP$ leverages the advantages of DRL, as it can be expected to perform better with higher levels of uncertainty than rule-based approaches.

4.5. Limitations and Future Work

The limitations of the presented research lie principally in seismic hazard modelling, environment formulation, and the computational complexity of DRL. Specifically, considering that DRL can handle approximately 10^1 or 10^2 agents, it is exceedingly difficult to use it as a decision-making tool for real communities. Thus, a custom environment has to be used. Additionally, considering that the objective function is an aggregate of sub-system functionalities, several limitations arise in modelling a custom community such that it represents a realistic community.

Considering the methodology for seismic hazard assessment,this thesis simulates a list of earthquake magnitudes given a maximum and minimum range and uses to generate IM values at study sites. Probabilistic Seismic Hazard Assessment (PSHA) was tested in the later stages of this thesis, but found to be too computationally expensive. Given that overall simulation time was already hindered by DRL, it was not possible to integrate it into the methodology.However, if careful consideration of the faults around a community is made, then a methodology for modelling seismic hazard probabilistically as a function of an annual probability of exceedance can be formed. This would allow for the seed of a given rollout to be the annual probability of exceedance (APE) and not the earthquake magnitude, which, in resilience loss terms, is more valuable. This would allow the analysis of different policies on how they perform for certain APA values. This process has to be modelled as a Poisson process, which is the commonly used statistical model for stochastically predicting seismic hazard curves. While this requires the combined simulation of many rupture scenarios for the deduction of one hazard curve, other approaches can be found that use single-fault events at each DRL realisation. Given the large number of simulation steps in a DRL experiment, this could be adapted to approximate a Poisson process, regardless of the single-realisation analysis being non Poissonian. Additionally, the spatial correlation of IM levels should be considered, especially as the environment grows in size. This is because the sampled mean for two study sites can vary if the sampling technique is not correlated.

Traffic assignment relies on trip data that is either generated or collected from field research. In this thesis, trips were generated manually by using reasonable assumptions about driver behaviour. However, this proved to induce the majority of the post-disaster delay costs on a handful of roads. While this can be realistic for some communities, the limitation of this approach lies in realistically representing driver behaviour. Thus, trip generation for a custom environment can be improved by considering $TAZs$ (Traffic Assignment Zones) that relate to the land use of specific clusters within a community. For instance, existing data on trips to and from residential, commercial and essential zones can be used to generate trips. This would make the resulting contribution of each traffic link to overall traffic performance more realistic. Additionally, the use of a global traffic-related cost metric could be changed to contribute to specific sub-systems. For instance, the delay of trips to and from the hospital can be integrated into the healthcare functionality metrics. This will make the meaning of the traffic delay more sub-system specific. Another aspect of traffic modelling that can be improved is in post-disaster driver behaviour. Drivers and community users in general will tend to have different driving patterns after an earthquake as many facilities might be closed and users will be distraught or injured and might not be able to take any trips at all. For example, the trips from residential buildings to a shopping mall are assumed to stay constant in pre- and post-disaster scenarios. This is most probably not true after a disaster. At last, the algorithm and general methodology used to solve traffic assignment can be strengthened by considering dynamic route choice. This thesis considers static assignment such

that once drivers observe the reduction in capacity due to debris, they chose a route, which they take until completion, regardless of the effect of dynamic congestion. In reality, a driver is very likely to chose a non-congested road when being on a congested road.

Concerning the definition of community resilience, the thesis presents a basic approach in aggregating economic, healthcare and critical functionalities. Future improvements on this approach can be made by conducting a sensitivity analysis on the weighting of different sub-system functionalities. This would allow researchers to gauge the role that each sub-system plays in overall community performance and choose the weighting profile in a more informed way. Additionally, more indirect losses can be incorporated to deal with different hazards and different communities. For instance, if this approach is to be applied a multi-hazard setting that gauge the combined risk from many different disruption types, the types of losses that apply to any one hazard are very different from the rest. For example, when looking at the losses caused by a wildfire, indirect environmental losses are of high interest as vast areas of woodland can be lost.

5

Conclusion

5.1. Summary

This thesis researches the use of MARL for post-earthquake repair scheduling of interdependent infrastructure. It does so by collecting relevant fragility and vulnerability data on US-based infrastructure systems and modelling two custom test beds with 4 and 30 components respectively. MARL is tested as a solution concept by developing a Python environment for simulating the decision making scenario and running MARL experiments. MARL is benchmarked against baseline importance-based and random solvers. Scenario-based earthquake modelling is used to randomly sample an earthquake from a dataset of 9 earthquake magnitude (6.0 to 9.0) with 100 per-magnitude instances at every initialisation of a rollout. Intensity measure data for the earthquakes is retrieved from INCORE, and HAZUS fragility functions are used to sample component damage state probability distributions. Post-earthquake community functionality is formulated as a weighted sum of economic, healthcare and critical functionalities. Economic functionality includes the community's cumulative income against costs like building and road repair costs, traffic delay cost and relocation costs. Healthcare functionality is a measure of the number of beds and doctors available before and after the earthquake. Critical functionality is a measure of the functioning essential facilities before and after the earthquake. The aggregate community functionality metric is used to compute the instantaneous resilience of the community at each time-step as the shared reward among agents. Constraints on the environment include a budgetary constraint in the form of the number of available repair crews at each timestep. Agents are allowed to act on the environment as long as there are available repair crews. The objective and research questions are repeated below:

Problem Statement

- Natural Disasters are increasing in frequency and are causing exceeding economic and human losses to communities around the world. Decision-makers have little access to tools that can help them reduce their communities' disaster risk before, during and after disastrous events.

Research Questions

1. **Principal Research Question:**
   (a) How effective is Reinforcement Learning when used as a decision-making tool for post-earthquake repair scheduling of interdependent infrastructures and when compared to baseline methods?
2. **Sub-Questions:**
   (a) How accurate is the computational modelling of earthquakes for different locations?
   (b) What are the factors contributing to community functionality before and after an earthquake disaster?

(c) How can different community functionality metrics be distilled into an aggregate community functionality metric?

(d) How can an aggregate community functionality metric be used to describe the resilience of a community in terms of its ability to rebound after an earthquake?

MARL was tested using three algorithms, Value Decomposition Network with Parameter Sharing (VDN-PS), Q-Learning with a Mixer Network and Parameter Sharing (QMIX-PS) and Deep Centralised Mult-Agent Actor Critic (DCMAC). Through iteration and testing of eight environment versions, DCMAC was found to be the most performative. However, training did not result in clear convergence as the total number of training timesteps was kept at 100K. Given that these runs already took in excess of ten hours to complete, it was difficult to conduct thorough testing for so many environment versions with more timesteps. Nonetheless, inference of DCMAC showed that the specific advantage of using DRL lies in the early recovery phase. $DCMAC$ performed better than the other two algorithms and only matched $IMPB$ when considering the recovery of all losses. In contrast, when looking at recovering 70% of the impact losses $DCMAC$ performed much better than $IMPB$, which performed worse than a random policy for the larger environment. From this, three key conclusions are drawn:

- $DCMAC$ can effectively prioritise early recovery better than importance-based scheduling, $VDN-PS$ and $QMIX-PS$ by effectively learning which components contribute to aggregate community functionality the most. It exhibits this performance while not being given any explicit information about how important components are, even when put under budgetary constraints that apply a random ranking if the joint actions exceeds the number of available repair crews at each decision step. However, It fails to fully recover the community as it does not receive large enough reward for doing so, especially later on in the recovery process, where the discount factor is more effective.
- $IMPB$, as defined this thesis, exhibits poor performance for larger environments as it cannot effectively abstract the interdependencies of the network. This is due to the specific rules of the algorithm, which were kept simplistic. All though not tested in this thesis, the behaviour of an importance-based algorithm for larger, stochastic environments with more types of interdependences is expected to further deteriorate in relation to DRL.
- DRL is exceedingly resource-hungry and requires expensive hardware for effective learning. Thus, the computational implementation of $MDPs$ should vectorised with numerical computation libraries such as JAX that can allow parallel training instances.

Overall, the principal research question was answered by testing multiple environment versions and three MARL algorithm. MARL can then be effective and preferable in post-earthquake early-phase repair scheduling when compared with baseline importance-based and random solvers. Answering the sub-questions, earthquake modelling for different location is still a topic of ongoing research and current approaches can be accurate if proper boundary conditions are used. Specifically, a simplified scenario-based approach is used which can effectively attenuate ground motion, but does not say much about the risk associated with all the ruptures that might affect a given location. Concerning resilience, the factors contributing to community functionality as considered in this thesis are repair costs, income and income loss, repair times and relocation costs. However, this thesis recognises that these metrics are used given the formulation of the environment. It should be the common goal of resilience formulations to include many different aspects of resilience and consider a holistic post-disaster recovery. Given these conclusions, the goal is to then promote the use of DRL for post-earthquake recovery to inform existing heuristic approaches, such as importance-based scheduling. This can allow engineers and decision makers to make smarter early-phase recovery decisions by combining expert judgment as well as state-of-the-art DRL and heuristic approaches.

5.2. Reflection

Having completed this thesis, I believe the main contribution to the field lies in the ability of intelligent agentic systems to assist decision makers in strategising for post-disaster community recovery. The urgency of the problem and the computational and technical complexity of the implementation made

researching and compiling this thesis highly valuable. As someone with training in Architecture and little to no experience in infrastructure modelling or *DRL*, I found some of the topics incredibly difficult to understand, and even more so to implement in a tool. However, because of this, I think I learnt a lot that I otherwise wouldn't have, if I had chosen a thesis topic that most well aligned with my previous skill set. I took the decision to undertake this research project as I saw the urgency in mitigating hazard exposure in the built environment and because I enjoyed the systems-level thinking associated with infrastructure management. I also maintained interest in the topic, even when my experiments were failing, which happened more time than not. Considering that the work is quite novel, there was few tools available that could *fix* parts of the experiment for me; in this way I had to debug most of the issues myself, which was frustrating but very rewarding. Conclusively, I think, personally, the thesis offered me a way to develop my technical and software skill set, but also allowed me to think differently about infrastructure and hazard risk. In the broader field of Building Technology, I think the research provides a starting point in using cutting edge tools for mitigating hazard risk by means of *DRL*.

However, I think certain parts of the research leave much to be desired. Considering the applied nature of the topic, I spent the majority of the time developing and testing code. This allowed me to develop a rich environment that describes the various indirect losses such as income, traffic, relocation etc. Reading and compiling literature was done sporadically, when I think I could have spent a little more time early on to learn from existing literature. This would have both saved me time in amending errors later on in the code, but also provided a richer literature review, which is the weakest, I believe, part of the thesis. Likewise, I think I could have begun *DRL* testing earlier so that I cant test various reward functions, action spaces and different budgetary constraints. I was naive to think that the integration of *DRL* into the research would be manageable in the time-frame left between P3 and P5. However, even having a well documented and tested *DRL* framework from Prateek Bhustaki, who also was very kind to meet with me regularly, I was faced with many errors that were extremely hard to spot, never mind solve. However, it is hard to say whether I would say the same if I had an environment that was less rich in engineering terms but had more a more thorough *DRL* methodology. For me, this indicates the trade-offs one can expect in this type of research; considering that this research is done writhing engineering and architecture and not computer science, the principal goal is always to make a rich and realistic infrastructure environment over a complete *DRL* methodology. Nevertheless, I think these challenges allowed me to stay constantly engaged with the topic, even though the execution of certain chapters could be improved. I hope the work inspires other researchers or future Building Technology students to look into infrastructure management.

References

[1] Norman Abrahamson and Walter Silva. "Summary of the Abrahamson &; Silva NGA Ground-Motion Relations". In: *Earthquake Spectra* 24.1 (2008), pp. 67–97. DOI: 10.1193/1.2924360.

[2] Hojjat Adeli. "Neural Networks in Civil Engineering: 1989–2000". en. In: *Computer-Aided Civil and Infrastructure Engineering* 16.2 (2001). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/0885-9507.00219, pp. 126–142. ISSN: 1467-8667. DOI: 10.1111/0885-9507.00219. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/0885-9507.00219 (visited on 06/20/2025).

[3] Federal Highway Administration. *Work Zone Road User Costs - Concepts and Applications*. May 2022. URL: https://ops.fhwa.dot.gov/wz/resources/publications/fhwahop12005/sec2.htm.

[4] Stefano V Albrecht, Filippos Christianos, and Lukas Schäfer. *Multi-Agent Reinforcement Learning: Foundations and Modern Approaches*. en. The MIT Press, 2024. URL: https://www.marl-book.com/download/marl-book.pdf.

[5] Irasema Alcántara-Ayala. "Geomorphology, natural hazards, vulnerability and prevention of natural disasters in developing countries". In: *Geomorphology*. Geomorphology in the Public Eye: Political Issues, Education, and the Public 47.2 (Oct. 2002), pp. 107–124. ISSN: 0169-555X. DOI: 10.1016/S0169-555X(02)00083-1. URL: https://www.sciencedirect.com/science/article/pii/S0169555X02000831 (visited on 06/20/2025).

[6] C.P. Andriotis and K.G. Papakonstantinou. "Deep reinforcement learning driven inspection and maintenance planning under incomplete information and constraints". en. In: *Reliability Engineering & System Safety* 212 (Aug. 2021), p. 107551. ISSN: 09518320. DOI: 10.1016/j.ress.2021.107551. URL: https://linkinghub.elsevier.com/retrieve/pii/S095183202100106X (visited on 05/13/2025).

[7] C.P. Andriotis and K.G. Papakonstantinou. "Managing engineering systems with large state and action spaces through deep reinforcement learning". en. In: *Reliability Engineering & System Safety* 191 (Nov. 2019), p. 106483. ISSN: 09518320. DOI: 10.1016/j.ress.2019.04.036. URL: https://linkinghub.elsevier.com/retrieve/pii/S0951832018313309 (visited on 06/08/2025).

[8] Jack Baker, Brendon Bradley, and Peter Stafford. *Seismic Hazard and Risk Analysis*. en. 1st ed. Cambridge University Press, Oct. 2021. ISBN: 978-1-108-34815-7 978-1-108-42505-6. DOI: 10.1017/9781108425056. URL: https://www.cambridge.org/core/product/identifier/9781108348157/type/book (visited on 05/22/2025).

[9] Jack W. Baker. "Efficient Analytical Fragility Function Fitting Using Dynamic Structural Analysis". en. In: *Earthquake Spectra* 31.1 (Feb. 2015), pp. 579–599. ISSN: 8755-2930, 1944-8201. DOI: 10.1193/021113EQS025M. URL: https://journals.sagepub.com/doi/10.1193/021113EQS025M (visited on 06/16/2025).

[10] Matteo Bettini. *Static traffic assignment using user equilibrium and system optimum*. GitHub. 2021. URL: https://github.com/MatteoBettini/Traffic-Assignment-Frank-Wolfe-2021.

[11] Prateek Bhustali. *Inspection and Maintenance Planning using Reinforcement Learning (IMPRL)*. GitHub. 2024. URL: https://github.com/omniscientoctopus/imprl.

[12] Prateek Bhustali and Charalampos P. Andriotis. "Assessing the Optimality of Decentralized Inspection and Maintenance Policies for Stochastically Degrading Engineering Systems". en. In: *Artificial Intelligence and Machine Learning*. Ed. by Frans A. Oliehoek, Manon Kok, and Sicco Verwer. Vol. 2187. Series Title: Communications in Computer and Information Science. Cham: Springer Nature Switzerland, 2025, pp. 236–254. ISBN: 978-3-031-74649-9 978-3-031-74650-5. DOI: 10.1007/978-3-031-74650-5_13. URL: https://link.springer.com/10.1007/978-3-031-74650-5_13 (visited on 06/08/2025).

[13] Huseyin Bilgin. "Fragility-based assessment of public buildings in Turkey". In: *Engineering Structures* 56 (Nov. 2013), pp. 1283–1294. ISSN: 0141-0296. DOI: 10.1016/j.engstruct.2013.07.002. URL: https://www.sciencedirect.com/science/article/pii/S014102961300326X (visited on 05/13/2025).

[14] Agnieszka Blokus. "Reliability of aging multistate dependent systems". en. In: *Multistate System Reliability with Dependencies*. Elsevier, 2020, pp. 59–136. ISBN: 978-0-12-821260-8. DOI: 10.1016/B978-0-12-821260-8.00003-8. URL: https://linkinghub.elsevier.com/retrieve/pii/B9780128212608000038 (visited on 06/08/2025).

[15] Stephen D Boyles, Nicholas E Lownes, and Avinash Unnikrishnan. *Transportation Network Analysis*. National Science Foundation, 2019.

[16] David S. Brookshire et al. "Direct and Indirect Economic Losses from Earthquake Damage". In: *Earthquake Spectra* 13.4 (Nov. 1997), pp. 683–701. ISSN: 8755-2930. DOI: 10.1193/1.1585975. URL: https://doi.org/10.1193/1.1585975 (visited on 06/20/2025).

[17] T. Brown, J. Hörsch, and D. Schlachtberger. "PyPSA: Python for Power System Analysis". In: *Journal of Open Research Software* 6.4 (1 2018). DOI: 10.5334/jors.188. eprint: 1707.09913. URL: https://doi.org/10.5334/jors.188.

[18] Lucian Buşoniu, Robert Babuška, and Bart De Schutter. "Multi-agent Reinforcement Learning: An Overview". en. In: DOI: 10.1007/978-3-642-14435-6_7. URL: https://link.springer.com/chapter/10.1007/978-3-642-14435-6_7 (visited on 06/20/2025).

[19] H. Jithamala Caldera and S. C. Wirasinghe. "A universal severity classification for natural disasters". En. In: *Natural Hazards* 111.2 (Nov. 2021). Number: 2 Publisher: Springer, pp. 1533–1573. ISSN: 1573-0840. DOI: 10.1007/s11069-021-05106-9. URL: https://link.springer.com/article/10.1007/s11069-021-05106-9 (visited on 06/20/2025).

[20] Kennetg W Campbell and Yousef Borzognia. *Campbell-Bozorgnia NGA Ground Motion Relations for the Geometric Mean Horizontal Component of Peak and Spectral Ground Motion Parameters*. May 2007. URL: https://apps.peer.berkeley.edu/ngawest/documents/campbell-bozorgnia_nga_report_files/PEER_2007_02_Campbell_Bozorgnia.pdf.

[21] Lin Chen and Qiang Bai. "Optimization in Decision Making in Infrastructure Asset Management: A Review". en. In: *Applied Sciences* 9.7 (Apr. 2019). Number: 7 Publisher: Multidisciplinary Digital Publishing Institute, p. 1380. ISSN: 2076-3417. DOI: 10.3390/app9071380. URL: https://www.mdpi.com/2076-3417/9/7/1380 (visited on 06/20/2025).

[22] S.-J. Brian Chiou and R. Robert Youngs. *NGA Model for Average Horizontal Component of Peak Ground Motion and Response Spectra*. Nov. 2008. URL: https://apps.peer.berkeley.edu/products/CY-Program/PEER_Report_2008_09.pdf.

[23] Anil K. Chopra. *Dynamics of structures: theory and applications to earthquake engineering*. en. Fifth edition. Prentice-Hall international series in civil engineering and engineering mechanics. Hoboken, NJ: Pearson, 2017. ISBN: 978-0-13-455512-6.

[24] European Commission. *Eurocode 8: Design of structures for earthquake resistance*. EN 1998. URL: https://eurocodes.jrc.ec.europa.eu/EN-Eurocodes/eurocode-8-design-structures-earthquake-resistance.

[25] United States Army Corps. *National Structures Inventory*. URL: https://www.hec.usace.army.mil/confluence/nsi/technicalreferences/latest/technical-documentation.

[26] Helen Crowley et al. "Developing fragility and consequence models for buildings in the Groningen field". en. In: *Netherlands Journal of Geosciences* 96.5 (Dec. 2017), s247–s257. ISSN: 0016-7746, 1573-9708. DOI: 10.1017/njg.2017.36. URL: https://www.cambridge.org/core/product/identifier/S0016774617000361/type/journal_article (visited on 04/19/2025).

[27] Carlo Del Gaudio et al. "Seismic fragility for Italian RC buildings based on damage data of the last 50 years". En. In: *Bulletin of Earthquake Engineering* 18.5 (Dec. 2019). Number: 5 Publisher: Springer, pp. 2023–2059. ISSN: 1573-1456. DOI: 10.1007/s10518-019-00762-6. URL: https://link.springer.com/article/10.1007/s10518-019-00762-6 (visited on 05/07/2025).

[28] Marco Di Ludovico et al. "Relationships between empirical damage and direct/indirect costs for the assessment of seismic loss scenarios". En. In: *Bulletin of Earthquake Engineering* 20.1 (Oct. 2021). Number: 1 Publisher: Springer, pp. 229–254. ISSN: 1573-1456. DOI: 10.1007/s10518-021-01235-5. URL: https://link.springer.com/article/10.1007/s10518-021-01235-5 (visited on 06/20/2025).

[29] Alassane Drabo and Linguère Mously Mbaye. "Natural disasters, migration and education: an empirical analysis in developing countries | Environment and Development Economics". en. In: *Cambridge Core* (2014). DOI: 10.1017/S135577 0X14000606. URL: https://www.cambridge.org/core/journals/environment-and-development-economics/article/natural-disasters-migration-and-education-an-empirical-analysis-in-developing-countries/2BA25870881F4995A4E8C8D30DE4A4A1 (visited on 06/20/2025).

[30] M. Altuğ Erberik. "Fragility-based assessment of typical mid-rise and low-rise RC buildings in Turkey". In: *Engineering Structures* 30.5 (May 2008), pp. 1360–1374. ISSN: 0141-0296. DOI: 10.1016/j.engstruct.2007.07.016. URL: https://www.sciencedirect.com/science/article/pii/S0141029607002799 (visited on 05/13/2025).

[31] Xudong Fan et al. "A deep reinforcement learning model for resilient road network recovery under earthquake or flooding hazards". En. In: *Journal of Infrastructure Preservation and Resilience* 4.1 (Feb. 2023). Number: 1 Publisher: SpringerOpen, pp. 1–19. ISSN: 2662-2521. DOI: 10.1186/s43065-023-00072-x. URL: https://link.springer.com/article/10.1186/s43065-023-00072-x (visited on 06/20/2025).

[32] Federal Highway Administration. *Appendix A: Highway investment analysis methodology*. Nov. 2019. URL: https://www.fhwa.dot.gov/policy/23cpr/appendixa.cfm.

[33] Alexander Fekete and Frank Fiedrich, eds. *Urban Disaster Resilience and Security: Addressing Risks in Societies*. en. The Urban Book Series. Cham: Springer International Publishing, 2018. ISBN: 978-3-319-68605-9 978-3-319-68606-6. DOI: 10.1007/978-3-319-68606-6. URL: http://link.springer.com/10.1007/978-3-319-68606-6 (visited on 01/16/2025).

[34] FEMA. *Hazus 6.0 Inventory Technical Manual*. Hazus 6.1. Federal Emergency Management Agency. 2024. URL: https://www.fema.gov/sites/default/files/documents/fema_hazus-inventory-technical-manual-6.1.pdf.

[35] FEMA. *Hazus Earthquake Model Technical Manual*. Hazus 6.1. Federal Emergency Management Agency. 2024. URL: https://www.fema.gov/sites/default/files/2020-10/fema_hazus_earthquake_technical_manual_4-2.pdf.

[36] FEMA. *Seismic Performance Assessment of Buildings*. 2024. URL: https://www.fema.gov/sites/default/files/documents/fema_national-disaster-recovery-framework-third-edition_05062025_0.pdf.

[37] Edward H. Field et al. "A Synoptic View of the Third Uniform California Earthquake Rupture Forecast (UCERF3)". en. In: *Seismological Research Letters* 88.5 (Sept. 2017), pp. 1259–1267. ISSN: 0895-0695, 1938-2057. DOI: 10.1785/0220170045. URL: https://pubs.geoscienceworld.org/srl/article/88/5/1259-1267/354096 (visited on 06/14/2025).

[38] Imogen Foulkes. "The Swiss village wiped off the map by a glacier". In: *The Visual Journalism Team* (2025). URL: https://www.bbc.co.uk/news/resources/idt-c7f929de-96a9-45e5-b1bb-31de82fce72d.

[39] Hiroyuki Fujiwara et al. "DEVELOPMENT OF GROUND MOTION CHARACTERIZATION MODEL AT THE IKATA SITE BASED ON GUIDELINES FOR SSHAC LEVEL 3". In: *Journal of Japan Association for Earthquake Engineering* (2024).

[40] Marco Gaetani d'Aragona et al. "Aftershock collapse fragility curves for non-ductile RC buildings: a scenario-based assessment". en. In: *Earthquake Engineering & Structural Dynamics* 46.13 (2017). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/eqe.2894, pp. 2083–2102. ISSN: 1096-9845. DOI: 10.1002/eqe.2894. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/eqe.2894 (visited on 05/07/2025).

[41] Firas Gerges et al. "A perspective on quantifying resilience: Combining community and infrastructure capitals". In: *Science of The Total Environment* 859 (Feb. 2023), p. 160187. ISSN: 0048-9697. DOI: 10.1016/j.scitotenv.2022.160187. URL: https://www.sciencedirect.com/science/article/pii/S0048969722072874 (visited on 01/24/2025).

[42] Nafiseh Ghorbani-Renani et al. "Protection-interdiction-restoration: Tri-level optimization for enhancing interdependent network resilience". In: *Reliability Engineering & System Safety* 199 (July 2020), p. 106907. ISSN: 0951-8320. DOI: 10.1016/j.ress.2020.106907. URL: https://www.sciencedirect.com/science/article/pii/S0951832019308191 (visited on 01/14/2025).

[43] Camilo Gomez and Jack W. Baker. "An optimization-based decision support framework for coupled pre- and post-earthquake infrastructure risk management". en. In: *Structural Safety* 77 (Mar. 2019), pp. 1–9. ISSN: 01674730. DOI: 10.1016/j.strusafe.2018.10.002. URL: https://linkinghub.elsevier.com/retrieve/pii/S0167473017303739 (visited on 06/16/2025).

[44] Andrés D. González et al. "The Interdependent Network Design Problem for Optimal Infrastructure System Restoration". en. In: *Computer-Aided Civil and Infrastructure Engineering* 31.5 (2016). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/mice.12171, pp. 334–350. ISSN: 1467-8667. DOI: 10.1111/mice.12171. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/mice.12171 (visited on 01/14/2025).

[45] Roberto Guidotti et al. "Modeling the resilience of critical infrastructure: the role of network dependencies". In: *Sustainable and Resilient Infrastructure* 1.3-4 (Nov. 2016). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/23789689.2016.1254999, pp. 153–168. ISSN: 2378-9689. DOI: 10.1080/23789689.2016.1254999. URL: https://doi.org/10.1080/23789689.2016.1254999 (visited on 11/13/2024).

[46] Jürgen Hackl, Bryan T. Adey, and Nam Lethanh. "Determination of Near-Optimal Restoration Programs for Transportation Networks Following Natural Hazard Events Using Simulated Annealing". en. In: *Computer-Aided Civil and Infrastructure Engineering* 33.8 (2018). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/mice.12346, pp. 618–637. ISSN: 1467-8667. DOI: 10.1111/mice.12346. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/mice.12346 (visited on 06/20/2025).

[47] Zachary Hamida and James Alexandre Goulet. "Maintenance planning for bridges using hierarchical reinforcement learning". en. In: (2023).

[48] Xiaotian Hao et al. *Breaking the Curse of Dimensionality in Multiagent State Space: A Unified Agent Permutation Framework*. en. arXiv:2203.05285 [cs]. Oct. 2022. DOI: 10.48550/arXiv.2203.05285. URL: http://arxiv.org/abs/2203.05285 (visited on 06/20/2025).

[49] P. C. Haritha and M. V. L. R. Anjaneyulu. "Comparison of topological functionality-based resilience metrics using link criticality". In: *Reliability Engineering & System Safety* 243 (Mar. 2024), p. 109881. ISSN: 0951-8320. DOI: 10.1016/j.ress.2023.109881. URL: https://www.sciencedirect.com/science/article/pii/S0951832023007950 (visited on 06/20/2025).

[50] M. Izzat Idriss. *EMPIRICAL MODEL FOR ESTIMATING THE AVERAGE HORIZONTAL VALUES OF PSEUDO-ABSOLUTE SPECTRAL ACCELERATIONS GENERATED BY CRUSTAL EARTHQUAKES*. Jan. 2007. URL: https://apps.peer.berkeley.edu/ngawest/documents/Idriss_NGA/Idriss_NGA_Report_Jan_19_2007.pdf.

[51] Anaconda Inc. *Anaconda Software Distribution*. Version Vers. 2-2.4.0. 2020. URL: https://docs.anaconda.com/.

[52] Omar Kammouh et al. "Multi-system intervention optimization for interdependent infrastructure". In: *Automation in Construction* 127 (July 2021). arXiv:2203.01411 [eess], p. 103698. ISSN: 09265805. DOI: 10.1016/j.autcon.2021.103698. URL: http://arxiv.org/abs/2203.01411 (visited on 06/16/2025).

[53] Leonidas Alexandros S. Kouris and Andreas J. Kappos. "Fragility Curves and Loss Estimation for Traditional Timber-Framed Masonry Buildings in Lefkas, Greece". In: *Seismic Assessment, Behavior and Retrofit of Heritage Buildings and Monuments* (Jan. 2015). DOI: 10.1007/978-3-319-16130-3_8. URL: https://link.springer.com/chapter/10.1007/978-3-319-16130-3_8 (visited on 05/07/2025).

[54] Elisavet-Isavela Koutsoupaki et al. "Seismic Fragility Analysis of Retaining Walls Dependent on Initial Conditions". en. In: *Geosciences* 14.1 (Dec. 2023). Number: 1 Publisher: Multidisciplinary Digital Publishing Institute, p. 2. ISSN: 2076-3263. DOI: 10.3390/geosciences14010002. URL: https://www.mdpi.com/2076-3263/14/1/2 (visited on 05/07/2025).

[55] Li Lai et al. "Synergetic-informed deep reinforcement learning for sustainable management of transportation networks with large action spaces". In: *Automation in Construction* 160 (Apr. 2024), p. 105302. ISSN: 0926-5805. DOI: 10.1016/j.autcon.2024.105302. URL: https://www.sciencedirect.com/science/article/pii/S0926580524000384 (visited on 06/20/2025).

[56] Nelson Lam. "A review of stochastic earthquake ground motion prediction equations for stable regions". En. In: *International Journal of Advances in Engineering Sciences and Applied Mathematics* 15.1 (Jan. 2023). Number: 1 Publisher: Springer, pp. 1–14. ISSN: 0975-5616. DOI: 10.1007/s12572-022-00325-0. URL: https://link.springer.com/article/10.1007/s12572-022-00325-0 (visited on 04/18/2025).

[57] Pascal Leroy et al. *IMP-MARL: a Suite of Environments for Large-scale Infrastructure Management Planning via MARL*. arXiv:2306.11551 [cs]. Oct. 2023. DOI: 10.48550/arXiv.2306.11551. URL: http://arxiv.org/abs/2306.11551 (visited on 01/14/2025).

[58] Huangbin Liang et al. "Seismic risk analysis of electrical substations based on the network analysis method". en. In: *Earthquake Engineering & Structural Dynamics* 51.11 (2022). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/eqe.3695, pp. 2690–2707. ISSN: 1096-9845. DOI: 10.1002/eqe.3695. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/eqe.3695 (visited on 06/20/2025).

[59] Xuan Lyu et al. "On Centralized Critics in Multi-Agent Reinforcement Learning". In: *Journal of Artificial Intelligence Research* 77 (2023), pp. 295–354. DOI: 10.1613/jair.1.14386. URL: https://doi.org/10.1613/jair.1.14386.

[60] Wes McKinney et al. "Data structures for statistical computing in python". In: *Proceedings of the 9th Python in Science Conference*. Vol. 445. Austin, TX. 2010, pp. 51–56.

[61] David B Moore and Gail M Atkinson. *Boore-Atkinson NGA Ground Motion Relations for the Geometric Mean Horizontal Component of Peak and Spectral Ground Motion Parameters*. May 2007. URL: https://peer.berkeley.edu/publications/peer_reports/reports_2007/web_PEER701_BOOREatkinson.pdf.

[62] P. G. Morato et al. "Inference and dynamic decision-making for deteriorating systems with probabilistic dependencies through Bayesian networks and deep reinforcement learning". In: *Reliability Engineering & System Safety* 235 (July 2023), p. 109144. ISSN: 0951-8320. DOI: 10.1016/j.ress.2023.109144. URL: https://www.sciencedirect.com/science/article/pii/S0951832023000595 (visited on 06/20/2025).

[63] John Nash. "Non-Cooperative Games". PhD thesis. Princeton University, 1951.

[64] National Institute for Standards and Technology. *IN-CORE Python API Documentation*. https://incore.ncsa.illinois.edu/doc/incore/pyincore.html. Accessed: 2025-05-06. n.d.

[65] Mohammad Hossein Oboudi et al. "A Systematic Method for Power System Hardening to Increase Resilience Against Earthquakes". In: *IEEE Systems Journal* 15.4 (Dec. 2021), pp. 4970–4979. ISSN: 1937-9234. DOI: 10.1109/JSYST.2020.3032783. URL: https://ieeexplore.ieee.org/document/9250524/ (visited on 06/20/2025).

[66] Martin J. Osborne. *An Introduction to Game Theory*. Version: 2000/11/6. Comments to martin.osborne@utoronto.ca. Toronto, Canada: Oxford University Press, 2000. URL: https://mathematicalolympiads.wordpress.com/wp-content/uploads/2012/08/martin_j-_osborne-an_introduction_to_game_theory-oxford_university_press_usa2003.pdf.

[67] Min Ouyang and Kun Yang. "Does topological information matter for power grid vulnerability?" en. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 24.4 (Dec. 2014), p. 043121. ISSN: 1054-1500, 1089-7682. DOI: 10.1063/1.4897268. URL: https://pubs.aip.org/cha/article/24/4/043121/928364/Does-topological-information-matter-for-power-grid (visited on 06/20/2025).

[68] Guillermo Owen. *Game Theory*. en. Google-Books-ID: yeVbAAAAQBAJ. Emerald Group Publishing, Aug. 2013. ISBN: 978-1-78190-508-1.

[69] Mehmet Palanci, Sevket Murat Senel, and Ali Kalkan. "Assessment of one story existing precast industrial buildings in Turkey based on fragility curves". En. In: *Bulletin of Earthquake Engineering* 15.1 (June 2016). Number: 1 Publisher: Springer, pp. 271–289. ISSN: 1573-1456. DOI: 10.1007/s10518-016-9956-x. URL: https://link.springer.com/article/10.1007/s10518-016-9956-x (visited on 05/13/2025).

[70] Konstantinos Papatheodorou et al. "Rapid Earthquake Damage Assessment and Education to Improve Earthquake Response Efficiency and Community Resilience". en. In: *Sustainability* 15.24 (Dec. 2023). Number: 24 Publisher: Multidisciplinary Digital Publishing Institute, p. 16603. ISSN: 2071-1050. DOI: 10.3390/su152416603. URL: https://www.mdpi.com/2071-1050/15/24/16603 (visited on 05/07/2025).

[71] Craig Poulin and Michael Kane. "Infrastructure Resilience Curves: Performance Measures and Summary Metrics". In: *Reliability Engineering & System Safety* 216 (Dec. 2021). arXiv:2102.01009 [cs], p. 107926. ISSN: 09518320. DOI: 10.1016/j.ress.2021.107926. URL: http://arxiv.org/abs/2102.01009 (visited on 04/11/2025).

[72] Sima Rastayesh et al. "Development of Stochastic Fatigue Model of Reinforcement for Reliability of Concrete Structures". en. In: *Applied Sciences* 10.2 (Jan. 2020). Number: 2 Publisher: Multidisciplinary Digital Publishing Institute, p. 604. ISSN: 2076-3417. DOI: 10.3390/app10020604. URL: https://www.mdpi.com/2076-3417/10/2/604 (visited on 06/20/2025).

[73] Hannah Ritchie, Pablo Rosado, and Max Roser. *Natural Disasters*. https://ourworldindata.org/natural-disasters. Accessed: 2025-06-22. Our World in Data, Global Change Data Lab, 2022.

[74] F. Rosenblatt. "The perceptron: A probabilistic model for information storage and organization in the brain". In: *Psychological Review* 65.6 (1958). Place: US Publisher: American Psychological Association, pp. 386–408. ISSN: 1939-1471. DOI: 10.1037/h0042519.

[75] Annalisa Rosti, Maria Rota, and Andrea Penna. "Empirical fragility curves for Italian URM buildings". En. In: *Bulletin of Earthquake Engineering* 19.8 (Apr. 2020). Number: 8 Publisher: Springer, pp. 3057–3076. ISSN: 1573-1456. DOI: 10.1007/s10518-020-00845-9. URL: https://link.springer.com/article/10.1007/s10518-020-00845-9 (visited on 05/07/2025).

[76] M Rota et al. "DIRECT DERIVATION OF FRAGILITY CURVES FROM ITALIAN POST-EARTHQUAKE SURVEY DATA". In: *World Conference on Earthquake Engineering*. 2008.

[77] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. "Learning representations by back-propagating errors". en. In: *Nature* 323.6088 (Oct. 1986). Publisher: Nature Publishing Group, pp. 533–536. ISSN: 1476-4687. DOI: 10.1038/323533a0. URL: https://www.nature.com/articles/323533a0 (visited on 06/20/2025).

[78] David E. Rumelhart, Bernard Widrow, and Michael A. Lehr. "The basic ideas in neural networks". en. In: *Communications of the ACM* 37.3 (Mar. 1994), pp. 87–92. ISSN: 0001-0782, 1557-7317. DOI: 10.1145/175247.175256. URL: https://dl.acm.org/doi/10.1145/175247.175256 (visited on 06/20/2025).

[79] M. Saifullah et al. *Multi-agent deep reinforcement learning with centralized training and decentralized execution for transportation infrastructure management*. arXiv:2401.12455 [cs]. Jan. 2024. DOI: 10.48550/arXiv.2401.12455. URL: http://arxiv.org/abs/2401.12455 (visited on 05/13/2025).

[80] A. Sandoli et al. "Seismic vulnerability assessment of minor Italian urban centres: development of urban fragility curves". En. In: *Bulletin of Earthquake Engineering* 20.10 (Mar. 2022). Number: 10 Publisher: Springer, pp. 5017–5046. ISSN: 1573-1456. DOI: 10.1007/s10518-022-01385-0. URL: https://link.springer.com/article/10.1007/s10518-022-01385-0 (visited on 05/07/2025).

[81] Omar A. Sediek, Sherif El-Tawil, and Jason McCormick. "Dynamic Modeling of In-Event Interdependencies in Community Resilience". EN. In: *Natural Hazards Review* 21.4 (Nov. 2020). Publisher: American Society of Civil Engineers, p. 04020041. ISSN: 1527-6996. DOI: 10.1061/(ASCE)NH.1527-6996.0000413. URL: https://ascelibrary.org/doi/10.1061/%28ASCE%29NH.1527-6996.0000413 (visited on 11/14/2024).

[82] Omar A. Sediek, Sherif El-Tawil, and Jason McCormick. "Modeling Interdependencies between the Building Portfolio, Transportation Network, and Healthcare System in Community Resilience". EN. In: *Natural Hazards Review* 23.1 (Feb. 2022). Publisher: American Society of Civil Engineers, p. 04021060. ISSN: 1527-6996. DOI: 10.1061/(ASCE)NH.1527-6996.0000538. URL: https://ascelibrary.org/doi/10.1061/%28ASCE%29NH.1527-6996.0000538 (visited on 11/08/2024).

[83] Sevket Murat Senel and Ali Haydar Kayhan. "Fragility based damage assesment in existing precast industrial buildings: A case study for Turkey". en. In: *Structural Engineering and Mechanics* 34.1 (Jan. 2010), pp. 39–60. DOI: 10.12989/SEM.2010.34.1.039. URL: https://doi.org/10.12989/SEM.2010.34.1.039 (visited on 05/13/2025).

[84] Neetesh Sharma, Armin Tabandeh, and Paolo Gardoni. "Regional resilience analysis: A multiscale approach to optimize the resilience of interdependent infrastructure". en. In: *Computer-Aided Civil and Infrastructure Engineering* 35.12 (2020). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/mice.12606, pp. 1315–1330. ISSN: 1467-8667. DOI: 10.1111/mice.12606. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/mice.12606 (visited on 01/16/2025).

[85] Shi Shen et al. "Spatial distribution patterns of global natural disasters based on biclustering". En. In: *Natural Hazards* 92.3 (Mar. 2018). Number: 3 Publisher: Springer, pp. 1809–1820. ISSN: 1573-0840. DOI: 10.1007/s11069-018-3279-y. URL: https://link.springer.com/article/10.1007/s11069-018-3279-y (visited on 06/20/2025).

[86] Vitor Silva et al. "Development of the OpenQuake engine, the Global Earthquake Model's open-source software for seismic risk assessment". En. In: *Natural Hazards* 72.3 (Mar. 2013). Number: 3 Publisher: Springer, pp. 1409–1427. ISSN: 1573-0840. DOI: 10.1007/s11069-013-0618-x. URL: https://link.springer.com/article/10.1007/s11069-013-0618-x (visited on 05/07/2025).

[87] United States Geologicl Society. *Building damage states used to classify buildings affected by the 2018 lower East Rift Zone lava flows by damage severity*. Nov. 2022. URL: https://www.usgs.gov/media/images/building-damage-states-used-classify-buildings-affected-2018-lower-east-rift-zone-lava.

[88] BBC Sofia Ferreira Santos. "Spain floods: Before and after images show devastation". In: *The Visual Journalism Team* (2024). URL: https://www.bbc.com/news/articles/cz7wvpyewxlo.

[89] Nikolaos Soulakellis et al. "Post-Earthquake Recovery Phase Monitoring and Mapping Based on UAS Data". en. In: *ISPRS International Journal of Geo-Information* 9.7 (July 2020). Number: 7 Publisher: Multidisciplinary Digital Publishing Institute, p. 447. ISSN: 2220-9964. DOI: 10.3390/ijgi9070447. URL: https://www.mdpi.com/2220-9964/9/7/447 (visited on 05/07/2025).

[90] National Insitute for Standards et al. *INCORE Software Package*. 2024. URL: https://incore.%20ncsa.illinois.edu/.

[91] J. K. Summers et al. "Observed Changes in the Frequency, Intensity, and Spatial Patterns of Nine Natural Hazards in the United States from 2000 to 2019". en. In: *Sustainability* 14.7 (Mar. 2022). Number: 7 Publisher: Multidisciplinary Digital Publishing Institute, p. 4158. ISSN: 2071-1050. DOI: 10.3390/su14074158. URL: https://www.mdpi.com/2071-1050/14/7/4158 (visited on 06/01/2025).

[92] Peter Sunehag et al. *Value-Decomposition Networks For Cooperative Multi-Agent Learning*. arXiv:1706.05296 [cs]. June 2017. DOI: 10.48550/arXiv.1706.05296. URL: http://arxiv.org/abs/1706.05296 (visited on 05/08/2025).

[93] Institute of Transportation Engineers. *ITETripGen Web-Based App*. Educational Single User License. Jan. 2025. URL: https://www.itetripgen.org/.

[94] Alan M. Turing. "Computing Machinery and Intelligence". en. In: Nov. 2007. DOI: 10.1007/978-1-4020-6710-5_3. URL: https://link.springer.com/chapter/10.1007/978-1-4020-6710-5_3 (visited on 06/20/2025).

[95] Ryan J. Williams, Paolo Gardoni, and Joseph M. Bracci. "Decision analysis for seismic retrofit of structures". In: *Structural Safety*. Risk Acceptance and Risk Communication 31.2 (Mar. 2009), pp. 188–196. ISSN: 0167-4730. DOI: 10.1016/j.strusafe.2008.06.017. URL: https://www.sciencedirect.com/science/article/pii/S0167473008000714 (visited on 06/20/2025).

[96] Sen Yang et al. “Multi-agent deep reinforcement learning based decision support model for resilient community post-hazard recovery”. In: *Reliability Engineering & System Safety* 242 (Feb. 2024), p. 109754. ISSN: 0951-8320. DOI: 10.1016/j.ress.2023.109754. URL: https://www.sciencedirect.com/science/article/pii/S0951832023006683 (visited on 10/24/2024).

6

Environment I: 4 Components

Property	Feature 1	Feature 2
Longitude	-118.24292153078737	-118.24156906361358
Latitude	34.05356058157683	34.05340222441406
year_built	1999	1980
no_stories	6	8
occ_type	RES3E	COM6
dwell_unit	74	0
struct_typ	S5	PC2
str_typ2	S5M	PC2H
appr_bldg	55200000	102830000
cont_val	27500000	51400000
efacility	FALSE	TRUE
sq_foot	59417	110683

Table 6.1: Toy-City-4 Building Data

Property	Feature 1	Feature 2
highway	primary	primary
width	12	12
unit_cost	6992	6992
repl_cost	793	396
linknwid	0	1
fromnode	0	0
tonode	1	2
len_mile	0.11	0.056
len_km	0.17	0.09
hazus_r	HRD1	HRD1
hazus_b	None	HWB8
K3D_A	None	0.09
K3D_B	None	1.0
i_shape	None	0
skew_angle	None	10
num_spans	0	3

Table 6.2: Toy-City-4 Road Data

Property	Row 1	Row 2	Row 3	Row 4	Row 5	Row 6
init_node	0	2	1	1	0	2
term_node	1	0	2	0	2	1
capacity	207.8	185.8	94.8	207.8	185.8	94.8
length	703	629	321	703	629	321
free_flow_time	0.145248	0.129959	0.066322	0.145248	0.129959	0.066322
b	0.15	0.15	0.15	0.15	0.15	0.15
power	4	4	4	4	4	4
speed	4840	4840	4840	4840	4840	4840
toll	0	0	0	0	0	0
link_type	1	1	1	1	1	1

Table 6.3: Toy-City-4 Traffic Network Data

Property	Row 1	Row 2	Row 3	Row 4
init_node	0	0	1	1
term_node	1	2	2	0
demand	500	150	150	800

Table 6.4: Toy-City-4 Traffic O-D Matrix

7

Environment II: 30 Components

ID	Year	Stories	Occ. Type	Appr. Bldg	eFacility	Cont. Val	Dwell Units	Str. Typ2	Struct Typ	Sq. Foot
1	1999	2	RES1	978498.52	FALSE	489249.26	1.305	W1	W1	526.62
2	1980	2	RES1	978498.52	FALSE	489249.26	1.305	W1	W1	526.62
3	1968	2	RES1	978498.52	FALSE	489249.26	1.305	W1	W1	526.62
4	2009	2	RES1	1082100.00	FALSE	541025.95	1.443	W1	W1	582.35
5	1965	4	RES3A	9117100.00	FALSE	4558500.00	24.31	S1L	S1	2453.39
6	1975	2	RES5	3568200.00	FALSE	1784100.00	9.52	W2	W2	1920.36
7	2000	2	RES5	2341000.00	FALSE	1170500.00	6.24	W2	W2	1259.91
8	1982	5	GOV2	12716000.00	TRUE	6358000.00	0.00	S4M	S4	2737.50
9	1971	9	GOV2	69018000.00	TRUE	34509000.00	0.00	S5L	S5	8254.42
10	1961	3	COM9	8378100.00	FALSE	4189100.00	0.00	PC2M	PC2	3006.03
11	1998	4	COM6	39329000.00	TRUE	19665000.00	0.00	C1M	C1	4728.69
12	1961	5	COM8	49553000.00	FALSE	24776000.00	0.00	RM1M	RM1	10667.57
13	1960	5	COM8	17373000.00	FALSE	8686300.00	0.00	RM1M	RM1	3739.95
14	1986	5	COM4	12559000.00	FALSE	6279700.00	0.00	S4H	S4	2703.75
15	1979	4	REL1	11081000.00	FALSE	5540400.00	0.00	URMM	URMM	2981.81

Table 7.1: Toy-city-30 Building Data

ID	u	v	Highway	Width	Unit Cost	Repl Cost	From	To	Len (km)	Hazard	K3D_A	K3D_B	Skew	Spans	I-Shape
1	0	1	primary	12	6992	434.46	0	1	0.10	HRD1	–	–	–	–	–
2	2	3	primary	12	6992	434.46	2	3	0.10	HRD1	–	–	–	–	–
3	4	5	primary	12	6992	434.46	4	5	0.10	HRD1	–	–	–	–	–
4	5	6	primary	12	6992	434.46	5	6	0.10	HRD1	–	–	–	–	–
5	6	7	primary	12	6992	434.46	6	7	0.10	HRD1	–	–	–	–	–
6	8	2	residential	6	2557	79.44	8	2	0.05	HRD2	–	–	–	–	–
7	3	9	residential	6	2557	79.44	3	9	0.05	HRD2	–	–	–	–	–
14	2	5	primary	12	6992	217.23	2	5	0.05	HRD1	0.09	1.0	5	2	0
15	2	0	residential	6	2557	79.44	2	2	0.05	HRD2	0.25	1.0	5	2	1

Table 7.2: Toy-City-30 Road Data

Table 7.3: Toy-City-30 Traffic Network Data

Init Node	Term Node	Capacity	Length (m)	Free Flow Time (h)	b	Power	Speed (m/h)	Toll	Link Type
0	1	105.81	358	0.07397	0.15	4	4840	0	1
1	2	103.41	350	0.07231	0.15	4	4840	0	1
2	3	55.46	188	0.03884	0.15	4	4840	0	1
2	4	36.78	373	0.16955	0.15	4	2200	0	1
2	5	32.94	334	0.15182	0.15	4	2200	0	1
1	6	51.61	175	0.07955	0.15	4	2200	0	1
6	7	46.17	156	0.03223	0.15	4	4840	0	1
6	8	30.81	313	0.14227	0.15	4	2200	0	1
5	9	48.76	165	0.03409	0.15	4	4840	0	1
6	2	99.41	336	0.06942	0.15	4	4840	0	1
3	6	112.73	381	0.07872	0.15	4	4840	0	1
1	10	93.89	318	0.06570	0.15	4	4840	0	1
10	9	93.90	318	0.06570	0.15	4	2200	0	1
10	5	108.29	366	0.07562	0.15	4	2200	0	1
1	0	105.81	358	0.07397	0.15	4	4840	0	1
2	1	103.41	350	0.07231	0.15	4	4840	0	1
3	2	55.46	188	0.03884	0.15	4	4840	0	1
4	2	36.78	373	0.16955	0.15	4	2200	0	1
5	2	32.94	334	0.15182	0.15	4	2200	0	1
6	1	51.61	175	0.07955	0.15	4	2200	0	1
7	6	46.17	156	0.03223	0.15	4	4840	0	1
8	6	30.81	313	0.14227	0.15	4	2200	0	1
9	5	48.76	165	0.03409	0.15	4	4840	0	1
2	6	99.41	336	0.06942	0.15	4	4840	0	1
6	3	112.73	381	0.07872	0.15	4	4840	0	1
10	1	93.89	318	0.06570	0.15	4	4840	0	1
9	10	93.90	318	0.06570	0.15	4	2200	0	1
5	10	108.29	366	0.07562	0.15	4	2200	0	1

Table 7.4: Toy-City-30 Traffic O-D Matrix

Init Node	Term Node	Demand
0	2	60
0	4	60
0	7	80
2	4	80
2	7	90
2	8	60
5	4	60
5	7	80
5	8	55
8	2	40
8	5	30
10	2	20
10	5	25
10	6	20
10	5	35
5	10	10
4	2	50
9	7	100
9	10	50
6	8	50
1	7	100
1	10	45
1	4	90