Bryan T. Adey and Nam Lethanh

Infrastructure Maintenance Processes

Lecture notes

December 16, 2015

ETH Zurich

Foreword

This document has been compiled to provide students taking the above mentioned course in understanding the examples given in class and assignments. All questions presented have been prepared by the authors based to varying degrees on different sources. The sources are given for each example.
This document is being issued so that students taking the course may better prepare for the exam. As it has not yet undergone the complete rigorous control processes of IBI this document is to be considered a work in progress and is hence labeled as a DRAFT.

IBI-ETH, September 2015

Professor: Bryan T. Adey
Dr: Nam Lethanh

Contents

Chapter 1

Level of Service - Impact Hierarchy

Bryan T. Adey and Nam Lethanh

1.1 General

In order to determine what to do with infrastructure it is necessary to determine what is expected from the infrastructure, i.e. it is necessary to determine the required LOSs. This must be defined for a specific time period. It requires the determination of an impact hierarchy and the determination of the values of each impact indicator or impact value over the specified period of time [Adey et al., 2012].

1.2 Impact hierarchy - Theory

1.2.1 General

The best way to define LOS provided by infrastructure is to consider it to be identical to the amount that persons are affected by the functioning (or non-functioning) of infrastructure with respect to the required LOS. For example, on a road network, it may be expected that next year *x mus*[1] are to be spent on the execution of maintenance interventions and y mus worth of impacts are due to accidents. If more than *x mus* are to be spent on maintenance and more than *y mus* worth of impacts occur due to accidents than it can be considered that an inadequate LOS was provided. If less or equal, than it can be considered that an adequate LOS was provided.

If it is too difficult to quantify all impacts to be used to measure LOS in terms of directly comparable units (above they are given in *mus*) proxies can be used. For example, instead of stating that y mus worth of impacts due accidents are expected next year, the number of accidents may be stated.

The type of proxy used can be evaluated with respect to its usefulness in assessing the amount that persons are affected by the functioning (or non-functioning) of infrastructure with respect to the required LOS. For example, the number of accidents might be considered a relatively good proxy with respect to the amount persons are affected by accidents, whereas the number of potholes in the road, might not be. The number of accidents can be multiplied by the average impact of an accident to give the expected total impacts, whereas if the number of potholes is used to calculate the expected total impact, it must be multiplied with the expected number of accidents that occur when there are this number of and the average impact of an accident [2]. In general, the fewer the number of calculations that are required to determine the amount that persons are affected, the better.

[1] *mus* stands for monetary units

[2] the expected number of accidents when a road section has *x* numbers of potholes is often estimated as probabilistic function that takes into account of some factors such as numbers of potholes, traffic volume, vehicle speed, etc.

In any case, the determination of the values of the proxies to be used should be approximated based on a rigorous assessment of the impacts related to these values, and a comparison of these values with the other impact types used to define LOS. If not, an optimal performance of the infrastructure will not be targeted.

It is important to define the LOS so that it covers all ways in which persons can be affected by the functioning (or non-functioning) of infrastructure. If this is not done, then there is a significant probability that managers will concentrate on some impacts and ignore others, which may also be important. For example, a manager might choose to reduce the probability of occurrence of an accident in a tunnel to a very small amount by imposing a regulation to reduce vehicle speed (e.g. from 100 km/hour to 80 km/hour) in the tunnel if only owner costs and accident costs are considered. If, however, owner costs, the costs of travel time and the costs of accidents are taken into consideration he may determine that the best strategy is to renew the road surface and the lighting system more frequently, even though it would result in higher costs for the owner.

To ensure that all this can be done it is useful to:

- determine the objective of setting the LOS
- determine and classify the persons affected by the functioning (or non-functioning) of the infrastructure, i.e. stakeholders
- determine and classify the ways in which each of the stakeholders are affected, i.e. the impact types.
- determine the indicators to be used to determine the value of the impacts incurred by stakeholders,
- determine the value of a unit change in the indicators used in units that can be compared across all impact types.

The grouping of stakeholders, their impacts at each level, and the indicators to be used to measure these impacts are referred to as an impact hierarchy. In the development of an impact hierarchy it is important that it is:

1. **complete** - all non-negligible ways in which persons are affected should be included. If not, there will be an unwanted focus on the items included in the hierarchy to the detriment of those that are not, which may lead overall to sub-optimal decisions.
2. **decomposable** - each impact should be broken down, or decomposed, to a level that is reasonable quantifiable. If the hierarchy is not decomposable to a reasonable good quantifiable level persons will be required to place numbers on impacts relatively subjectively.
3. **operational** - the impacts should be defined in a way and a level which it is reasonable to quantify them in terms of effort. If a hierarchy is defined in too much detail the effort will be too large to conduct the required analysis or to collect the required information. If a hierarchy is defined in too little detail, there will be so much subjectivity in the assessment of the value of the impacts that the results will be relatively useless.
4. **non-redundant, or orthogonal** - the impacts should be defined in a way that ensures impacts are not double counted. To help ensure orthogonality it is sometimes helpful to make use of the pillars of sustainability, i.e. economic, environmental, and social, in the definition of the lowest level impacts in the hierarchy. If the hierarchy is not orthogonal than the double counting of some, but not all, of the impacts will result in overall sub-optimal decisions.

1.2.2 Objective

The stakeholders, impacts types and proxies to be considered depend on the objective of determining the LOS. For example, if a road manager wants to determine the work program that will maximize the net positive impact that society as a whole obtains from the infrastructure, only changes that are relevant to society as a whole are considered, and therefore positive impacts on one stakeholder that are negative impacts on another are not taken into consideration. For example, the income that is received from charging a toll on a public road would not be considered as it is a positive impact on the owner but a negative impact of equal magnitude on the user.

1.2.3 Stakeholder

There are many ways to classify these persons. It is, however, advantageous to group them keeping in mind how models will be built to assess how they are affected over time. For example, a person affected by noise who is driving a car may not be affected in the same way as a person who lives next to road on which cars pass.

Being a stakeholder is time dependent. For example, when a person is driving a vehicle on a road he is a user at that point in time, but when he is off of the road and in his house far from the road he is part of the indirectly affected public. Although he is the same person he is affected by the road when he is using it in one way, and when he is not using it in another way.

Depending on the objective it can be helpful to classify stakeholders as either first level, or second level or third level stakeholders. The first level stakeholders are those whose net positive impacts should be maximized. The second level stakeholders are those whose impacts are the outcome of the maximization of the net impacts of the first level stakeholders, and should be monitored. The third level of stakeholders are those whose impacts are of no concern.

1.2.4 Impact

Each stakeholder can be affected by the functioning (or non-functioning) of the infrastructure in different ways. For example, someone driving on a road can be affected by being in an accident or through the amount of time they spend traveling.

The impacts are attributed to the stakeholder who is most directly affected. For example, the money spent to hire persons to execute an intervention on a bridge should be attributed to the owner of the bridge rather than the tax payers.

The impact are to be grouped in a hierarchy and given to a level of detail in a way in which it can be reasonably expected that they can be quantified.

The impact hierarchy should include all non-negligible impacts to the principal stakeholders of public road infrastructure, i.e. those whose positive impacts should be maximized.

1.2.5 Indicator

In order to estimate the impacts it is necessary to determine impact indicators, which are considered to be representative of each impact type. It is only by estimating the values of these indicators over time, and attributing monetary values to each unit change in the indicators, that the impact on the stakeholders can be correctly evaluated. If the values of these indicators are used directly to set a LOS they are referred to as proxies.

1.2.6 Conclusions

Determination of the impact hierarchy is the precursor to setting the required LOS. It is important that sufficient time be spent to define it well, and to ensure that decision makers understand it. It is only with this understanding that the statement of the required LOS will make sense, which will set the stage for good decision making in the future.

Ideally the determination of the required LOS would be closely linked to the theoretically optimal LOS, i.e. the stakeholders as a whole receive the maximum net benefit from the infrastructure. As there are, however, many different stakeholders affected by the functioning (or non-functioning) of infrastructure, with many different opinions of the values to be associated with the unit changes in different types of impacts, the fixing of the required LOS is something that is associated with considerable negotiation between all involved stakeholders [Brauers, 2003].

1.3 Impact hierarchy - an example

1.3.1 Question A

Develop an impact hierarchy that will allow a public road manager to determine the work program that will maximize the net positive impact that society as a whole obtains from the infrastructure over the next 20 years. Be clear as to the objective of developing the impact hierarchy. Explain the main components of the hierarchy in general. Give the complete hierarchy. Explain how you would attempt to quantify each impact type at the lowest level.

1.3.2 Answer A

1.3.2.1 Objective

This impact hierarchy is developed for public roads assuming that a road manager wants to determine the work program that will maximize the net positive impact that society as a whole obtains from the infrastructure Therefore, only changes that are relevant to society as a whole are considered, i.e. positive impacts on one stakeholder that are negative impacts on another are not taken into consideration. For example, the income that is received from charging a toll on public road is not considered as it is a positive impact on the owner but a negative impact of equal magnitude on the user.

1.3.2.2 Components

The impact hierarchy is to consist of stakeholders, impact types and impact indicators.

1.3.2.2.1 <u>Stakeholders</u>

A stakeholder is considered as an individual, group, or organization, which is affected by changes to public roads. Being a stakeholder is time dependent, i.e. when a person is driving a vehicle on a road he is a user at that point in time. When he is off of the road and in his house far from the road he is part of the indirectly affected public. It is considered that all stakeholders can be grouped as either first level or second level stakeholders. The first level stakeholders are those whose net positive impacts should be maximized. The second level stakeholders are those whose impacts are the outcome of the maximization of the net impacts of the first level stakeholders, and should be monitored.

The four first level stakeholder groups are the owner, the user, the directly affected public (DAP), and the indirectly affected public (IAP). It is assumed that all impacts to be maximized can be attributed to one of these four stakeholder groups. The impacts are attributed to the stakeholder who is most directly affected. The definitions of each stakeholder group are given in later parts of this section.

The proposed impact hierarchy is to be used to take into consideration all non-negligible impacts to the "principal" stakeholders of public road infrastructure, i.e. those whose positive impacts should be maximized. It is considered that all people at a point in time can be divided into one of four stakeholder groups, the owner, the user, the directly affected public, and the indirectly affected public. The impact types in the hierarchy are given to a level of detail in which they could be reasonably expected to be quantified. To help ensure orthogonality the pillar of sustainability, i.e. economic, environmental, social, is stated for each of the lowest level impacts. An example of each impact type is given.

Table 1.1: Stakeholders

Stakeholders	Definition	Examples
Owner	the persons who are responsible for decisions with respect to physically modifying the infrastructure	a federal road authority
Users	the persons who are using the roads	a driver and passengers of a vehicle on a road.
Directly affected public	the persons who are in the vicinity of the road but are not using it	persons in a house next to the road that hear vehicles driving on the road.
Indirectly affected public	the persons who are not in the vicinity of the road but are affected by its use	persons in a house far away from the road that do not hear vehicles driving on the road, but are affected by a changing climate due to the emissions produced by vehicles driving on the road.

Second level stakeholder groups, such as contractors, financial institutions, and operators are not considered further than they are considered when considered as one of the above listed stakeholders. The impacts on these stakeholder groups are only the outcome of our efforts to maximize the positive net impact of the four first level stakeholders in the first level.

1.3.2.2.2 Impact types

The impacts on each stakeholder are grouped as impact types. The impact types are subdivided at increasingly fine levels until the impact of each type can be reasonably and objectively quantified and modelled. To help to ensure orthogonality in the impact hierarchy, each impact type, on the lowest defined level, is explained and classified as contributing to one of the pillars of sustainability (economic, societal, environmental). An example is given for each to help clarify its meaning.

1.3.2.2.3 Impact indicators

Impact indicators, which are considered to be representative of each impact type are given for each impact type. By estimating the values of these indicators over time, and attributing monetary values to each unit change in the indicators, it is possible to evaluate the impact of the stakeholders. Examples of models to be used to estimate how their values change over time per impact type are given in section 5. The impact hierarchy is given in Table 1.2 to Table 5 in the following sections.

1.3.2.3 Impact hierarchy

1.3.2.3.1 Owner

The impacts attributed to the owner are grouped as intervention costs (Table 1.2), i.e. the impact on the owner of executing interventions. The impact indicators are the amount of labour, the amount of equipment, and the amount of material to be used to execute interventions, I^o_{labor},,$I^o_{material}$ e.g. intervention w required x man-hours, y generator-hours and z kilograms of material[3].

The monetary value placed on the

- labor used represents the economic impact of persons performing tasks, i.e. the value from society's perspective of the person doing the intervention.
- material used represents the economic impact of people ensuring that materials are available for use, i.e. the value from society's perspective of persons preparing the materials for use.
- equipment used represents the economic impact of people ensuring that equipment is available for use, i.e. the value from society's perspective of persons preparing the equipment for use.

The estimation of the value of an intervention is often done using one of two approaches:

- a disaggregate approach where expenditures for each item or activity are estimated and summed. When this approach is used the work break down structure of the intervention project it is often used.
- an aggregate approach where sum of all expenditures is estimated directly. An aggregate approach often includes regression analysis and historical information.

Table 1.2: Owner impact types

Level 1		Level 2	
Label	Description	Label	Description
Intervention	the impact of executing interventions	Labor*	the economic impact of people performing tasks
		Material*	the economic impact of people ensuring that materials are available for use
		Equipment*	the economic impact of people ensuring that equipment is available for use

* These could be further subdivided based on the type of activity performed, e.g. administration, planning, etc...

1.3.2.3.2 User

During the interventions, users experience inconveniences such as traffic jams, bumping condition of temporarily roads (detour roads). The unfavorable conditions of traveling via the road sections and links subjected to intervention will eventually lead to higher possibility of accidents, various types of physical exhaustion and illness as comfort levels decrease, loss of travel time, and increases in fuel consumption and frequency of the maintenance of vehicles.

In between interventions, deterioration processes result in a worsening condition of any infrastructure object. The increasingly poor condition of infrastructure also results in a change in how stakeholders are affected e.g. increases in the number of accidents, travel time, the costs of operating and maintaining vehicles.

[3] Throughout the document, I represents "impact" and it is used to as mathematical notation in functions where appropriate.

The impacts attributed to the user are grouped as: safety, operation efficiency, operation quality, and environment preservation (Table 1.3).

Table 1.3: User impact types

Level 1		Level 2	
Label	Description	Label	Description
Safety	the impact on the user due to the user being involved in an accident	property damage	the economic impact of repairing the vehicle
		injury	the societal impact due to the injury
		death	the societal impact due to death
Operation efficiency	the impact of travel in terms of time required	work	the economic impact of wasting work time traveling
		leisure	the economic impact of wasting leisure time traveling
	the impact of travel in terms of vehicle costs	operation	the economic impact of people ensuring that fuel and oil is available for use
		maintenance	the economic impact of people repairing vehicles and ensuring that materials, e.g. tires and brake pads, are available for use
Operation quality	the impact of traveling on the user	physical	the societal impact of obtaining for example, bruises from an extremely bumpy ride
		psychological	the societal impact of having for example, anxiety due to a perceived increase in the probability of being involved in an accident, or of seeing things while traveling.
Environment preservation (reduction of noise)	the societal impact due to the user coming in contract with sound emissions		

1. Safety (Accidents)

   Accidents result in damage to the property of involved parties. The owners will often repair the damaged objects so as to provide adequate service to users after accidents (this would be attributed to the owner). The users, however, will be required to repair their vehicles, and will also be affected by any injury and of course, death, that may befall them. The safety impact type attributed to the user is subdivided into property damage, injury and death impact types.

2. Property damage

   The property damage impact type represents the economic impact of repairing the vehicle, i.e. of providing the user with a functioning mode of transport similar to the one being used before the accident, e.g. the costs of the labor, materials and equipment required to replace the bumper on a vehicle that has been in an accident. The impact indicators for property damage are the amount of labor and materials used to repair vehicles damaged in an accident, expressed as the property damage costs given by I^{u}_{acc-pd}. The value of this impact type can be approximated using the receipts from past repairs.

3. Injury and death

   The injury impact type and the death impact type attributed to the user represent the social impact due to injuries and deaths, respectively. They represent the change in interactions between persons that will occur because the user is injured or dead. It is not to be confused with the injury impact type and the death impact type attributed to the directly affected public (refer to section 1.3.2.3.3) or to the indirectly affected public (refer to section 1.3.2.3.4). The impact indicators are the number of injuries and deaths incurred in a specified time interval, given by$I^{u}_{acc-injury}$, $I^{u}_{acc-death}$, respectively

   The value of these impact types can be estimated by using the user's willingness to pay to avoid injury or death. Care must; however, be taken to avoid double counting with the injury impact type and the death impact type attributed to the indirectly affected public.

4. Operation efficiency

   The operation efficiency impact type represents the impact on the traveling of users and on the maintenance and operation of the vehicles.

5. Travel time

   The amount of time traveling on the road is determined by various factors such as: speed, road condition (drivers feel comfortable on a smooth road, and therefore drives faster than on a bumpy road), and the DTV (the daily traffic volume, especially in relation to the road capacity), and road geometry. Furthermore, in case of an intervention, a detour might be needed.

   The economic impact of wasting work and leisure time traveling may be thought of as the loss of productivity of the users due to time spent traveling. The impact indicators are the amounts of work and leisure time wasted while traveling and are given by $I^u_{tt-work}$, $I^u_{tt-leisure}$

   The value of travel time can be determined using willingness to pay surveys.

6. Vehicle operation and maintenance

   The vehicle operation and maintenance impact types represent the economic impact of people ensuring that fuel and oil is available for use, and the economic impact of people repairing vehicles and ensuring that materials, e.g. tires and brake pads, are available for use, respectively. The impact indicators are$I^u_{v-oper}I^u_{v-main}$.

   The value of the vehicle operation and maintenance impact types can be approximated using the receipts from fuel and vehicle service receipts.

7. Operation quality

   The operation quality impact type is subdivided into the physical and psychological impact types. The physical impact type represents the social impact of obtaining for example, bruises from an extremely bumpy ride. It represents the change in interactions between people that will occur because the physical change in the user due to the bumpy ride. The psychological impact type represents the impact of having for example, anxiety due to a perceived increase in the probability of being involved in an accident, or of seeing things while traveling, e.g. aesthetics. It is believed that any economic impacts relevant to society due to the physical and psychological impacts of traveling, such as the loss of productivity, are negligible, in developed countries.

   The impact indicators are the amounts of physical and psychological impacts of traveling, given by $I^u_{physical}$, $I^u_{psychological}$.

   The value of degrees of bumpiness could be determined through willingness to pay investigations.

8. Environment preservation (noise)

   The environment preservation impact type represents the social impact due to the user coming in contact with sound emissions. It is meant to capture the changes that occur in the interactions between people due to sound emissions, e.g. the inability to communicate between driver and passenger while driving.

   The impact indicator is the amount of sound emissions to which the directly affected public is exposed I^u_{noise}.

   The value of an amount of sound emissions can be determined through willingness to pay investigations.

1.3.2.3.3 Directly affected public (DAP)

The impacts attributed to the directly affected public are grouped as: safety, operation quality, and environment preservation (Table 1.4), similar to the impact types of the user. The reason they are handled separately is that the directly affected public is affected in fundamentally different ways than the user.

Table 1.4: Directly affected impact types

Level 1		Level 2	
Label	Description	Label	Description
"Safety (Accidents)	the impact on the directly affected public due being involved in an accident	property damage	the economic impact of repairing the vehicle
		injury	the societal impact due to the injury
		death	the societal impact due to death
Operation quality	the impact of traveling on the directly affected public	physical	the societal impact of experiencing vibrations in a building next to a road
		psychological	the societal impact of having for example, anxiety due to a perceived increase in the probability of being involved in an accident, or of seeing things while being next to a road.
Environment preservation (reduction of noise)	the societal impact due to the directly affected public coming in contract with sound emissions		

1. Safety (Accidents)

   The safety impact type is subdivided into property damage, injury and death impact types.

2. Property damage

   The property damage impact type represents the economic impact of repairing damaged property, to the condition it was prior to the occurrence of the accident, e.g. the costs of the labor, and materials required to repair a retaining wall that has been damaged in an accident. The impact indicator is the property damage cost, I^{dap}_{acc-pd}.

   The value of this impact type can be approximated using the receipts from past repairs.

3. Injury and death

   The injury impact type and the death impact type attributed to the directly affected public represent the societal impact due to injuries and deaths, respectively. They represent the change in interactions between persons that will occur because someone other than the user is injured or dead. It is not to be confused with the injury impact type and the death impact type attributed to the indirectly affected public. The impact indicators are the number of injuries and deaths incurred in a specified time interval, given by$I^{dap}_{acc-injury}$, $I^{dap}_{acc-death}$, respectively

   The value of these impact types can be estimated by using willingness to pay of the directly affected public to avoid injury or death. Care must; however, be taken to avoid double counting with the injury impact type and the death impact type attributed to the indirectly affected public.

4. Operation quality (Comfort)

   The operation quality impact type is subdivided into the physical and psychological impact types. The physical impact type represents the societal impact of obtaining for example, discomfort through vibrations that occur due to road use. It represents the change in interactions between persons that will occur because the physical change in the directly affected public.

   The psychological impact type represents the social impact of having for example, anxiety due to a perceived increase in the probability of being involved in an accident, or of seeing the infrastructure, e.g. aesthetics. It is believed that any economic impacts to be attributed to the directly affected public due to the physical and psychological impacts of others traveling, such as the loss of productivity, are negligible.

   The impact indicators are the amounts of physical and psychological impacts of traveling, given by $I^{dap}_{physical}$, $I^{dap}_{psychological}$.

The value of degrees of bumpiness can be determined through willingness to pay investigations.

5. Environment preservation (Noise)

   The environment preservation impact type represents the social impact due to the directly affected public coming in contact with sound emissions. It is meant to capture the changes that occur in the interactions between people due to sound emissions, e.g. the necessity to change where people meet due to excess noise.

   The impact indicator is the amount of sound emissions to which the directly affected public is exposed I^{dap}_{noise}.

   The value of an amount of sound emissions can be determined through willingness to pay investigations.

1.3.2.3.4 Indirectly affected public (IAP)

The indirectly affected public are those that are affected by roads through other mediums, e.g. a person who is affected by an increase in the temperature of the earth due to the CO_2emitted during the execution of an intervention on a road. The impacts attributed to the indirectly affected public are grouped as: safety, socio-economic activity, environment preservation, and environment consumption.

Table 1.5 and Table 1.6 list the most common impact types and indicators. It is noted that several impact types, such as gas and particle emission also directly affect the users and directly affected public group. However, it is assumed that these impacts are minimal.

Table 1.5: Indirectly affected public impact types (safety, socio-economic activity)

Level 1		Level 2		Level 3	
Label	Description	Label	Description	Label	Description
Safety (Accidents)	The impact on the indirectly affected public of accidents occurring on roads	injuries	the economic impact due to an injury		
		deaths	the economic impact due to a death		
Socio-economic activity	The impact on the ...due to changes in socio-economic development	Persons	the impact of not being able to transport people	Productivity	the economic impact due to not being able to transport people, e.g. a crop cannot be harvested as workers cannot be transported to the field
				Health	the societal impact due to injuries and deaths of not being able to get proper medical care
		Goods	the impact of not being able to move goods	Productivity	the economic impact due to not being able to deliver goods, e.g. because of not being able to work as planned
				Health	the societal impact due to not being able to deliver goods, e.g. due to deaths because of lack of food or medical supplies
		Employment	social and economic impact of interventions in terms of employing people		

Table 1.6: Indirectly affected public impact types (environment preservation, environment consumption)

Level 1		Level 2		Level 3	
Label	Description	Label	Description	Label	Description
"Environment preservation (Reduction of gas and particles emissions)	The impact on ... due to the environment being impacted by particle emissions	CO2	the impact due to the emissions	Production	the environmental impact of emissions emitted during the production of materials
				Material transport	the environmental impact of emissions emitted during the transport of materials to and from the construction site
				Person transport	the environmental impact of emissions emitted during travel
				Health	the societal impact due to emissions (human health)
		PM10	Same as for CO2		
		nitrogen			
		carbon monoxide			
		aldehydes			
		nitrogen dioxide			
		sulphur dioxide			
		polycyclic aromatic hydrocarbons			
		dust			
Environment consumption	The impact on ... due to the slower depletion of finite amounts of non renewable resources	Energy	the environmental impact due to the consumption of energy not related to emissions, e.g. depletion of finite amounts of non-renewable energy sources		
		Materials	the environmental impact of consuming materials, not related to emissions		
		Land	The land impact type represents the environmental impact due to the consumption of land not related to emissions		
		Culture	The culture impact type represents the societal impact of changing things important to our identity (of which heritage is part)		

1. Safety (Accidents) The injury impact type and the death impact type attributed to the indirectly affect public represents the economic impact due to injuries and deaths, respectively. They represent the loss in productivity due to injuries and deaths, respectively. It includes changes to human activity, such as a doctor's time in an emergency room and the time required ensuring that an insurance company conducts the required financial transactions. The impact indicator is the amount of work time lost, when compared to the reference case where the accident had not occurred given by $I^{iap}_{safety\_time}$.

   The value of each amount can be determined by estimating the loss of productivity of the person involved in the accident.

2. Socio-economic activity

   The socio-economic activity impact type represents the contribution of the road to socio-economic development. It is composed of persons, goods and employment impact types.

3. Persons

   The person impact type is further divided into a productiveness impact type and a health impact type.

   - Productivity

     The productiveness impact type represents the economic impact due to not being able to travel, e.g. a farmer cannot harvest his entire crop because he needs to spend a significantly larger portion of his time getting his goods to market. The impact indicator is the amount of lost work, given by $I^{iap}_{trans-p-wt}$, and expressed in units of time.

     The value of each amount can be determined by conducting through simulations of the performance of the region.

   - Health

     The health impact type represents the societal impact due to injuries and deaths that occur due to not being able to obtain standard medical care, due to a shortage of available persons. The impact indicators are the number of injuries and deaths incurred in a specified time interval, given by$I^{iap}_{trans-p-injury}$, $I^{iap}_{trans-p-death}$, respectively. The value of each amount can be determined by conducting through simulations of the region.

4. Goods

   The goods impact type is also further divided into a productiveness impact type and a health impact type.

   - Productivity

     The productiveness impact type represents the economic impact due to not being able to deliver goods, e.g. a farmer cannot plant his crop on time since fertilizer could not be delivered as planned. The impact indicator is the amount of lost work, given by $I^{iap}_{trans-g-wt}$.and expressed in units of time.

   - Health

     The health impact type represents the societal impact due to injuries and deaths due to goods such as food or medical supplies not being delivered as planned, e.g. the change in society that occurs due to the death of someone in a hospital that would not have died if medical supplies had been delivered as planned. The impact indicators are the number of injuries and deaths incurred in a specified time interval, given by$I^{iap}_{trans-g-injury}$, $I^{iap}_{trans-g-death}$, respectively.

5. Employment

The employment impact type represents the societal impact of executing interventions in terms of employing people that is not captured by the impact type attributed to the owner due to the execution of interventions or the user due to the maintenance of vehicles used for traveling. It includes economic development. The impact indicator is the amount of work provided given by I^{idap}_{work}.

The value can be estimated as using economic impact assessment models such as those proposed in CUBRC [2001]; Davis [2001]; Kumares and Samuel [2007], using predictions of business output, value added, employment level, wages and salaries, and wealth are made.

6. Environment preservation (emissions)

The environment preservation (emissions) impact type of the indirectly affected public represents the environmental and societal impact due emissions emitted during the production and transport of materials and persons.

It is subdivided by particle emitted, e.g. CO_2, PM10, NO, CO, aldehydes, NO_2, SO_2, polycyclic aromatic hydrocarbons and dust. The impact indicators are the amounts of each that are emitted, given by $I^{iap}_{CO_2}$., $I^{iap}_{PM_{10}}$, I^{iap}_{CO}, $I^{iap}_{aldehydes}$, $I^{iap}_{NO_2}$, $I^{iap}_{SO_2}$, I^{iap}_{pah}, I^{iap}_{dust} .

Each of these are further subdivided into

- the production impact type, which represents the <u>environmental</u> impact of emissions emitted during the production of materials

- the material transport impact type, which represents the <u>environmental</u> impact of emissions emitted during the transport of materials to and from the construction site

- the person transport impact type, which represents the <u>environmental</u> impact of emissions emitted during travel

- the health transport impact type, which <u>societal</u> impact due to emissions (human health). It is meant to capture the changes that occur in the interactions between people due the changes in the people, e.g. due to sickness.

The value of each amount can be determined by analyzing historical records or by conducting empirical studies using emission measurement tools and instruments.

7. Environment consumption

The environment consumption impact type represents the depletion of finite amounts of non-renewable resources. It is subdivided into the energy impact type, the material impact type, the land impact type and the culture impact type.

- Energy

  The energy impact type represents the environmental impact due to the consumption of energy not related to emissions, e.g. depletion of finite amounts of non-renewable energy sources. The impact indicator is a measure of the significance of a unit depletion of a finite resource, given by I^{iap}_{energy}.

- Material

  The material impact type represents the environmental impact of consuming materials, not related to emissions, e.g. the consumption of wood has an impact on woodland areas. The impact indicator is the amount of consumed material, e.g. kg of wood cut, given by $I^{iap}_{materials}$.

- Land

The land impact type represents the environmental impact due to the consumption of land not related to emissions, e.g. increased environmental damages due to floods. The impact indicator is the m^2 of land converted from a natural state to a built state, given by I^{iap}_{land}.

- Culture

  The culture impact type represents the societal impact of changing things important to our identity (of which heritage is part).The impact indicator is a measure of the significance of an object to our identity, given by $I^{iap}_{culture}$. The value can be estimated through willingness to pay investigations.

1.3.3 Question B

In reading such an impact hierarchy some people will imagine things that may seemingly not be in the list of impacts. To help clarify this, explain why your impact hierarchy does not include an impact type “access” or an impact type “natural hazard risk”.

1.3.4 Answer B

1.3.4.1 Access

The improvement of access is often stated as an impact [Kumares and Samuel, 2007]. Improved access, however, is a proxy for the many things that actually happen when access is improved. For example, if access to a highway is improved a user traveling from point A, to point B, who normally would not take the highway, would perhaps have a shorter travel time, reduced vehicle costs, and reduced accident costs. It is proposed that if it is desired to approximate the value of improved access that the changes in the relevant impact types for each stakeholder are calculated.

A more numeric example is as follows (Table 1.7). If there are two interventions being considered for a highway (intervention A with a medium level of access and intervention B with a high level of access), it is possible that intervention B will result in a shorter distance for users to travel to arrive at their destinations in comparison with intervention A. This in turn results in a reduction in:

- accident costs of 2 mus,
- comfort costs of 5 *mus*,
- travel time costs of 4 *mus*, and
- vehicle costs of 2 *mus*.

Table 1.7: Access: Example

Interventions	Accident	Comfort	Travel Time	Vehicle
A	10	15	14	8
B	8	10	10	6

This means that intervention B is 13 *mus* better than intervention A, or in other words, the improvement in access has a value of 13 *mus*.

1.3.4.2 Natural hazard risks

The reduction of natural hazard risks is often stated as an impact. The reduction of natural hazard risk is, however, a proxy for the many things that may actually happen when the risk of natural hazards is reduced. It is proposed that if it is desired to approximate the value of reduced natural hazard risk that the changes in the relevant impact types (probability of occurrence and values) for each stakeholder be estimated.

For example, if there are two interventions (Table 1.8) being considered for the modification of a section of highway, one of them (A) may make it so that it is much quicker to restore service following the interruption of traffic flow than the other (B). Based on exactly how the interventions are planned and executed this might result in increased risk to the user as follows:

- accident costs of 1.5 *mus* (0.1 x 15 *mus*),
- comfort costs of 2.4 *mus* (0.1 x 24 *mus*),
- travel time costs of 11.3 *mus* (0.1 x 113 *mus*), and
- vehicle costs of 5.7 *mus* (0.1 x 57 *mus*).

Table 1.8: Natural hazard risk: example

Interventions	Accident	Comfort	Travel Time	Vehicle
A	10	15	14	8
B	11.5	17.4	25.3	13.7

1.4 Changes in LOSs - Theory

1.4.1 General

It is important to note that neither the required nor the provided LOSs are constant over time. Provided LOSs change due to deterioration of the infrastructure. Required LOSs change due to changing needs for the infrastructure. This is illustrated in Figure 1.1.

An example to help the interpretation of Figure 1 is as follows. At t = 0, a road link consisting of a bridge and a road section starts to provide a LOS that is higher than the initially agreed upon required LOS (e.g. the travel time costs required per vehicle to go from A to B should be below x, and earthquake risks should be below y). To ensure this, the road is to have a roughness above 20 mm/m, and the bridge is designed to withstand an earthquake of magnitude 4 with a very low probability of damage. Over time both objects deteriorate and the provided LOS decreases, i.e. due to the worsening road condition vehicles on average take more time to go from A to B, and the earthquake risks increases because the probability of the bridge withstanding an earthquake of magnitude 4 with no damage decrease.

As the infrastructure manager is aware that the provided LOS is decreasing over time, he executes two interventions a few years apart to restore the provided LOS to the initially agreed required level. The first time due planning problems the provided LOS falls below the initially agreed required LOS. The second time this does not happen.

Following the second intervention, the government introduces a new law, which requires that all bridges be able to resist an earthquake of a magnitude of 8 with no interruption to traffic flow, i.e. the earthquake risks are to be reduced. The infrastructure manager then executes an intervention to bring the provided LOS up to the initially required LOS

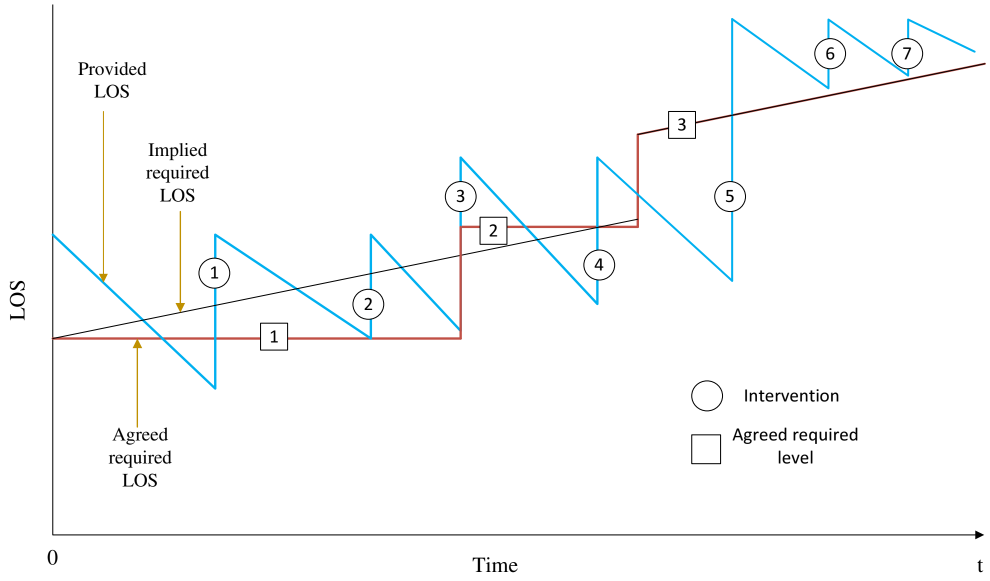


Fig. 1.1: Illustration of changes in required and provided LOS

for the road section and the bridge up to, and beyond, the new required LOS perfectly coordinated with the wishes of the government.

As deterioration of the road section and bridge continue, the infrastructure manager executes another intervention to restore the provided LOS to the new required LOS, but again due to planning problems later than he should have.

Immediately after the execution of the intervention, the public votes to have the speeds on the road increased from 100 to 120 kph, and there is a very short time frame given for the infrastructure manager to respond. It is also decided that the required LOS should be fixed not only with the traffic seen today but taking into consideration the growth in the future. The infrastructure manager responds with an intervention to expand the road at the same time improving the condition of both the road and the bridge and establishes a plan to execute interventions at appropriate times in the future so adequate LOSs are provided.

1.4.2 Ability to provide LOSs (deterioration)

Deterioration is the process of becoming worse. The deterioration of infrastructure is the process with which infrastructure loses its ability to provide a specific LOS.

From an infrastructure managers perspective, deterioration processes can be classified as either manifest or latent Lethanh et al. [2015]. Manifest processes are processes that are observable under the inspection strategies being followed, so that there is enough time for a manager to execute an intervention so as ensure that the infrastructure provides an adequate LOS.

Latent processes are processes that are unobservable under the inspection strategies being followed, so that there is enough time for a manager to execute an intervention so as ensure that the infrastructure provides an adequate LOS.

Although it does depend on the monitoring strategies in place, typical examples of manifest processes are pavement cracking, the leaching of mortar between masonry blocks and chloride induced corrosion, where the corrosion is expansive. Typical examples of latent processes are those that result in floods, earthquakes, and the overloading of infrastructure. Other examples are given in Table 1.9. An example of a process that is not as clear is the fatigue of metal.

Table 1.9: Examples of manifest deterioration processes

Process	Subprocess	Comments
Concrete surface wear	Abrasion	moving objects in contact with concrete which can wear away and often desired roughness
Aggregate deterioration	Alkali-silika reaction	a reaction between the hydroxyl ions in the alkaline cement pore solution in the concrete and reactive forms of silica in the aggregate form a gel is produced, which increases in volume by taking up water and so exerts an expansive pressure
	Freezing and thawing	under the humid conditions pores may be water-filled and under freezing conditions, water in these pores may freeze and thaw, which causes immense hydraulic pressure, which cracks concrete
Cement paste deterioration	Deterioration due to sulphates	exposure to industrial and agricultural chemicals can result in the break-down of the cement paste, the glue that holds concrete together
Reinforcement corrosion	Chloride induced corrosion of steel	Corrosion of steel can cause the production of a rust product that is significantly larger than the normal steel cracking the surrounding concrete and destroying the bond between concrete and steel. There may also be section loss of the steel reinforcement.
	Carbonation induced corrosion of steel	Carbonation, or neutralization, is a chemical reaction between carbon dioxide in the air with calcium hydroxide and hydrated calcium silicate in the concrete. The carbon dioxide in the air reacts with the alkali in the cement and makes the pore water more acidic, thus lowering the pH.

In order to understand how the provided LOS changes over time it is necessary to understand the deterioration processes themselves, how they affect the materials in which they are occurring, as well as how the changes being caused in the materials affect the functionality of the elements composed of these materials, the objects composed of these elements, and the networks composed of these objects.

1.4.2.1 Depth of process analysis

There is, of course, considerable depth with which deterioration processes can be analyzed. A few examples are given in Table 1.9. Even the lowest levels shown there, however, can be analyzed in more detail. For example, to analyze chloride induced corrosion of steel one can investigate, or model, the processes shown in Table 1.10. Much more information can be found in Elsener [2013] for the interested reader.

1.4.2.2 Relationship deterioration process - LOS

A simple example of how deterioration processes change the provided LOS is illustrated in Table 1.11 and Figure 1.2 for the chloride induced corrosion of reinforced concrete bridge. It can be seen that a change in the phase of a deterioration process does not necessarily mean that there is a change in the LOS. The relationship depends on the how the processes affect the materials, elements, objects and networks. As phases of deterioration are often expressed as condition states and LOSs are sometimes expressed as performance states, the relationship may also be considered as a condition state - performance state relationship. An example is given in Table 1.12 for a steel element.

Table 1.10: Examples of the subprocesses within the chloride induced corrosion of steel process for a reinforced concrete bridge

Subprocesses	Sub-subprocesses	Description
spreading processes		in which chlorides are spread on a concrete bridge
transport processes		in which the chlorides enter the bridge
	capillary absorption	in which substances in an aqueous solution get into the concrete near the surface
	diffusion	in which chloride ions due to different concentration of substances or different pressures within the concrete move within the concrete
	ion migration in electric fields	in which chloride ions move within the concrete
depassivation process		in which the steel in the concrete is depassivated which in turn allows corrosion to start if moisture and oxygen are present
corrosion process		In which RUST is formed
	cathodic (oxidation) process	in which oxidizing iron produces iron(II) ions and electrons , and electrons combine with water and oxygen to form hydroxide ions, $2H_2O+O_2+4e^- \rightarrow 4OH^-$
	anodic (reduction) process	hydroxide ions react with iron(II) ions to form Iron(II) hydroxide $2Fe^{2+}+4OH^- \rightarrow 2Fe(OH)_2$, and Iron(II) hydroxide reacts with oxygen to form $4Fe+3O_2 \rightarrow Fe_2O_3$rust

Table 1.11: Illustration of relationship between a deterioration process and provided LOSs

Deterioration process phases		Visual condition indications	Time	Provided LOSs
1	Penetration of water and de-icing salts into surface concrete	None	1	1
2	Breakdown of depassivation layer of top reinforcement	None	2	1
3	Rusting of top reinforcement	None	3	1
4	Cracking and spalling of surface concrete	Small pockmarks on the surface, often accompanied with a rust stain, and usually at a point corresponding to the location of a top reinforcing bar	5	1
5	Shifting of the compressive zone in the concrete deck	Spalling in multiple locations and some of these locations had significant amounts of exposed reinforcement Uniformly spaced cracks on the underside of the slab, beneath spalled areas	5	1
6	Over-stressing of the bottom reinforcement	No additional visual indications	6	1
7	Cracking of the bottom concrete	Spalling in numerous widely dispersed locations over relatively large areas, more intense under the truck lanes.	7	1
8	Punching failures	"Spalling in numerous widely dispersed locations over relatively large areas, extensive cracking was observed two punching failures of the deck slab occurred"	8	2
9	Collapse	-	?	3

1.4.3 Required LOSs

Changes in required LOSs are the result of changing human needs. They often result in existing infrastructure providing an inadequate LOS. For example, if a bridge was designed to have two lanes but a decision was made to expand the road to four lanes to accommodate changing human needs, then the bridge would no longer provide an adequate LOS.

From an infrastructure managers perspective, processes that result in changes in required LOSs, as deterioration processes, can be classified as either manifest or latent.

With respect to changes in required LOSs most processes are considered manifest if development strategies are well developed and triggers for the execution of interventions are determined. For example, the need to expand a road due to it reaching its carrying capacity in terms of number of vehicles in peak periods can be considered as a manifest

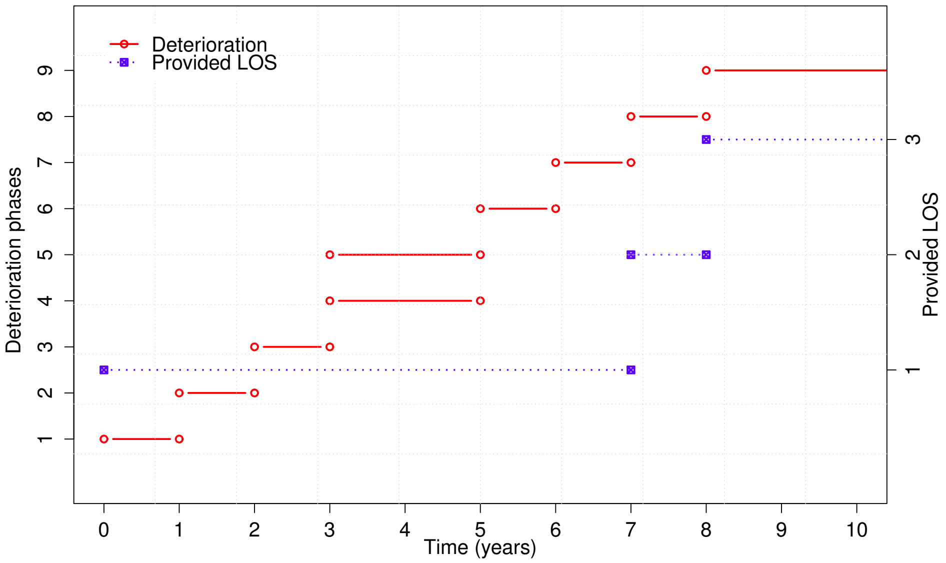


Fig. 1.2: Illustration of relationship between deterioration processes and the provided LOS

process if the number of vehicles per hour are recorded every day and a development strategy has been determined that states that once the road is at 80% its maximum capacity the road will be widened.

Most processes are considered as latent if development strategies are not well developed and no triggers for the execution of interventions are determined. For example, if the infrastructure manager is focused simply on maintaining the physical condition of the road it may come to him as a surprise when suddenly politicians request that a certain road be widened.

Some example of processes that can result in changes in the required LOS in buildings can be are shown in Table 1.12.

Table 1.12: Example changes for buildings and related processes

Change	Related process
Floor space layout	Demand for apartments suitable for growing elderly population
New windows	Demand for energy efficient houses and rises.
New heating system	Demand for less maintenance on the heating system
New control system	Demand for better control over the quality of the indoor air condition.

Changes to the required LOS can happen due the wishes of the owners of infrastructure or be imposed by external persons such as government. The government can also provide incentives for the owner to change the required LOS. Other stakeholders can exert influence on the owner to change the required LOS, e.g. tenants in an apartment building can threaten not to pay the rent unless an air conditioning system is installed. The development of new technologies can also instigate changes in the required LOS as new developments may make things possible that were not before, e.g. new heating systems.

1.4.4 Methodology

The five basic steps to modelling changes in the required LOS are (adapted from de Neufville and Scholtes [2011]):

- Step 1: Identify key performance drivers, i.e. determine the things that if changed will result in a change in the required LOS, e.g. if the number of vehicles per day pass x, then the road will need to be widen.

- Step 2: Analyse historical trends, i.e. obtain as much historical data as possible and model how it changed over time with the plan to use this model to predict the future values, e.g. obtain the evolution of vehicles on the road over the last 20 years.

- Step 3: Identify trend breakers, i.e. dig deeper and try to determine if there is something that can be used to help you predict if this trend will continue or might change suddenly, e.g. a new rail link on which passenger trains will circulate is about to be completed where there was none before.

- Step 4: Establish forecast accuracy, i.e. if enough data exists put yourself mentally in the past and see how could you have predicted the future for which you have data, e.g. using only traffic data from 1980 to 1999 how could would you be at predicting the traffic between 2000 and 2014.

- Step 5: Build a dynamic model, i.e. build a model that allows a range of predictions to be made using a sensible distribution.

An example for a hospital is provided in the work of de Neufville and Scholtes [2011] .

More detail on model development will be provided later.

1.5 Change in LOS-Example

1.5.1 Question C

Using the owner and user impacts of the impact hierarchy developed in section 3, develop simple models to be used to estimate how the provided LOS of a road will change over 20 years to the relatively slow deterioration of the pavement surface.

1.5.2 Answer C

1.5.2.1 Background

In order to develop simple models to quantify the values of the various impact types it is useful to think of the object under investigation as being in one of two states, i.e. when no intervention is being executed and when an intervention is being executed. In general, these can be represented by the functions $f^k(t,x)$ and $g^k(d,x)$, respectively (here, k represents intervention to be executed, e.g. renewal of the road section; x represents the performance indicator, t and d represent in between interventions and during intervention, respectively). It is assumed that they are expressed in *mus*. Both can be given by functions that vary over time. For example, two following generic exponential functions which can be used are:

$$f^k(t,x) = a^{k,x} + b^{k,x} \cdot \exp(\beta^{k,x} \cdot t) \tag{1.1}$$
$$g^k(d,x) = \bar{a}^{k,x} + \bar{b}^{k,x} \cdot \exp(\bar{\beta}^{k,x} \cdot d^k) \tag{1.2}$$

Where or $f^k(t,x)$ and $g^k(d,x)$ represent the impacts (costs) that are incurred as a function of time (t or d) and parameters a, b, and β.

If it is assumed that the change in condition of the road over time is known, this equation links the changes any impact on any stakeholder to the change in condition of the road over time.

An example of the evolution of vehicle operation cost (VoC) along with the evolution of roughness for 1 m length of road is shown in Figure 1.3. In the figure, Eq. (1) has parameter values a=500, b=800, and $\beta = 0.095$. These values are often estimated using empirical models. For example, in order to estimate the evolution of the VoC in between two consecutive interventions (from t_1 to t_2), Ouyang and Madanat [2004] used the following function.

$$\int_{t_1}^{t_2} f^k[s(t)]e^{-rt}dt = c\int_{t_1}^{t_2} s(t)e^{-rt}dt \tag{1.3}$$

In Eq. (1.3), $s(t)$ represents the roughness indicators; c is constant associated with cost; r is discount factor. The functional form of the roughness can be estimated with a simple exponential function as follows:

$$s(t) = s(t_0)\cdot \exp[\alpha(t-t_0)] \quad (for \ \ t \geq t_0) \tag{1.4}$$

Where $s(t_0)$ is initial roughness indicator (the initial LOS) and $\alpha$is a deterioration parameter.

When the time t is discrete, Ouyang and Madanat [2004] suggests to use following function

$$s(t+1) = [s(t)+\upsilon]\cdot \exp(\alpha) \tag{1.5}$$

With υ is the pre-defined model's parameter.

The cost curve (green curve in Figure 1.3) represents the increase of vehicle operation cost over time. The discount rate used is 2%. Evolution of roughness is estimated from using Eq. (1.5) with $\upsilon = 2$ and $\alpha = 0.0153$. Value of parameter $c$in Eq. (1.3) is 2.1.

Being generic, Eqs. (1.1) and (1.2) can be used to model the evolution of any impact that varies as a function of the condition of a road section, and allows the heterogeneity of each road section, e.g. the different deterioration rates, to be taken into consideration. For example, if it is desired to model the relationship between impacts on the user and the condition of a road section, the values of a, b, and β (e.g. for user) could be selected as non-negative values, ensuring that the function f will result in the values of the impact on the users that increase exponentially with time. This is similar to that of the user cost estimated using Eq. (1.3). Although the form of functions f and g are flexible, here, the exponential form is often used.

1.5.2.2 Owner

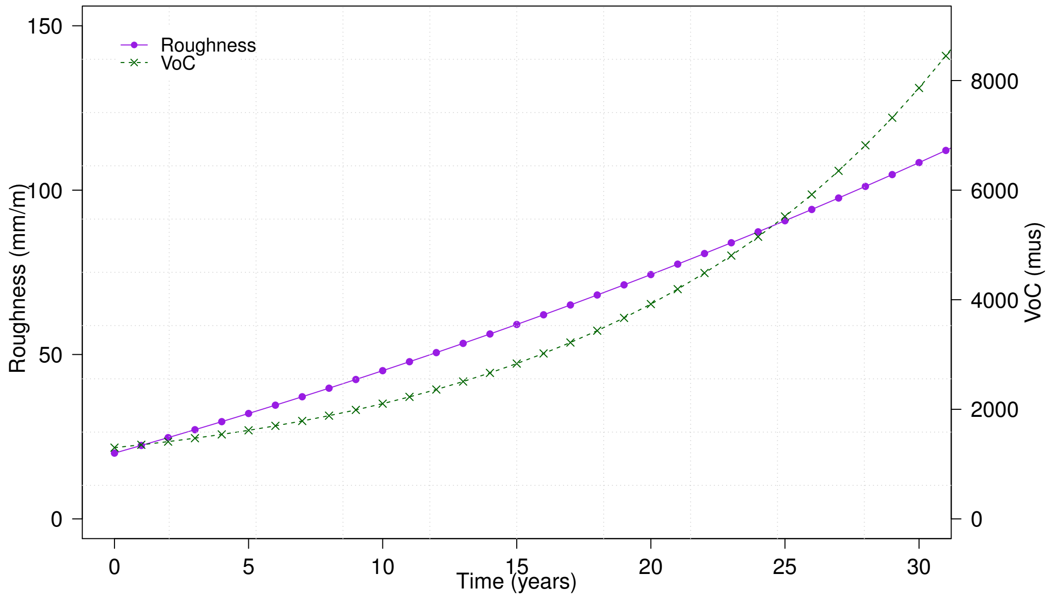


Fig. 1.3: Evolution of VoC and roughness indicator

$$f^k(t,x_i^o) = \sum_{i=1}^{3} I_i^o(t) \cdot c_i^o(t) \tag{1.6}$$

$$g^k(d,x_i^o) = \sum_{j=1}^{3} I_i^o(d) \cdot c_i^o(d) \tag{1.7}$$

Where:

i: is the index of impact types: labor, equipment, and materials

I_i^o: is the owner impact indicator for impact type i. In this case it is the number of person-hours, numbers of equipment use hours, and quantity of materials used for intervention

$c_i^o(t)$ and $c_i^o(d)$: are owner unit costs, which vary as a function of time t and d

1.5.2.3 Users

$$f^k(t,x_i^u) = \sum_{i=1}^{3} I_i^u(t) \cdot c_i^u(t) \tag{1.8}$$

$$g^k(d,x_i^u) = \sum_{i=1}^{3} I_i^u(d) \cdot c_i^u(d) \tag{1.9}$$

Where:

i: is the index of impact types

I_i^u: is the user impact indicator for impact type i

$c_i^u(t)$ and $c_i^u(d)$: are user unit costs, which vary as a function of time t and d.

1.5.2.3.1 Safety

The values of impact indicators ($I_i^u(t)$and $I_i^u(d)$) are generally estimated through regression analysis using empirical models. In empirical models, values of impact indicators associated with accidents can be estimated by using the "accident rate" multiplied by the cost of the accidents if they occur. Both depend on multiple factors such as the condition of the road, daily traffic volume, the physical condition of the driver of the vehicle and environmental factors such as, the occurrence of poor weather that can effect accident rate. For example, a popular way to estimate the value of impact indicators is by use of following equation [Lindenmann, 2008].

$$I_{property}^u = t \cdot DTV \cdot s \cdot \theta \cdot \omega \cdot \upsilon \tag{1.10}$$

Where:

t: is the number of days,

DTV: is the daily traffic volume,

s: is the length of the object,

θ: is the coefficient depending on the deterioration level of the civil infrastructure,

ω: is the correction factor for accident rate, and

υ: is the approximate numbers of vehicles involved in an accident.

According to Tarko et al. [2000], three ways commonly used to estimate the accident rate are

1. crash rate method, which recommends a rate of crash for a specific type of road, in specific location, and with specified level of traffic volume;
2. crash equation method, which relies on a functional form of traffic volume and engineering factors;
3. crash reduction factor method, which is an extension method of crash equation method, where the crash reduction factor is given by:

$$f(X) = k \cdot Y \cdot Q^{\gamma} \cdot e^{\beta x} \tag{1.11}$$

Where:

$f(X)$: is the crash reduction function,

k: is a constant,

Y, Q: are the exposure variables representing the temporal span of data and indicate the section length and traffic volume, respectively,

β: is the an unknown parameter associated with the variable X.

X: is the variable vector of engineering factors, including the performance condition of the road.

After obtaining the accident rate related to the execution of an intervention for each intervention type on an each object (i.e. between and during interventions), the value of impact indicators $I_i^u(t)$ and can be calculated.

For example, after the execution of an intervention, an asphalt road section of 20 km is renewed to be in condition state 1. The daily traffic volume is 600 vehicles/day. The accident rates during and after intervention are estimated to be 0.03% and 0.005% out of the total traffic volume in a day. The duration of the intervention is 30 days. It is assumed that the average number of vehicles involved in an accident are two. The expected total numbers of vehicles involved in accidents in a year after intervention will be:

$I^u_{property}(t = 365)\ days\ = 365 \times 600 \times 0.005\% \times 2 \approx 22$ vehicles

The expected total numbers of vehicles involved in accidents during intervention will be:

$I^u_{property}(d = 30)\ days\ = 30 \times 600 \times 0.03\% \times 2 \approx 11$ vehicles

The average unit cost $c^u_{\text{property}}(t)$per vehicle can be approximated from historical data Lindberg [1999].

1.5.2.3.2 Operation efficiency

The values of the travel time impact indicator can be approximated as a function of multiple factors including vehicle speed, amount of congestion, and curviness of the road.

The values of the vehicle operation and maintenance impact indicators can be approximated as a function of multiple factors including the total number of vehicle type j. (type j means type of vehicles such as car, truck, bus, etc),

$$f_{n_l}^{k_{n_l}}(t, x_i^u) = I^u_{i,k_{n_l}} = \sum_{i=1}^{2} \sum_{j=1}^{J} C^u_{ij}(t) \cdot VH_j(t) \tag{1.12}$$

$$g_{n_l}^{k_{n_l}}(d, x_i^u) = I^u_{i,k_{n_l}} = \sum_{i=1}^{2} \sum_{j=1}^{J} C^u_{ij}(d) \cdot VH_j(d) \tag{1.13}$$

Where:

I_i^u: is the impact indicator for the index i

i: is the index for vehicle maintenance cost and vehicle operation cost.

VH: is total number of vehicle type j. (type j means type of vehicles such as car, truck, bus, etc). The value of *VH* can be obtained from examining historical record on traffic volume and annual growth of traffic volume.

The value of $C^u_{ij}(t)$can be estimated by empirical study or regression analysis based on recorded numbers of bill paid for operation and maintenance respectively. To date, several models have been developed to relate such cost to deterioration or performance of the road. For example,OpusCL [1999] studied the relationship between vehicle operation cost and international roughness index. The value of *VH* can be obtained from examining historical record on traffic volume and annual growth of traffic volume.

1.5.2.3.3 Operational Quality

Impact indicators associated with and can be measured by using qualitative scale (e.g, scale from 1 to 5, with 1 is the best and 5 is the worst) or by carrying out an empirical study on the loss in effective working time if users travel on a target road link. The values of the differences between these states in these scales can be determined through willing to pay investigations.

Following proposed function can be used for estimating the values of impact indicators $BI^u_{physical,k_{n_l}}$ and $BI^u_{psychological,k_{n_l}}$.

$$BI^u_{i,k_{n_l}} = t \cdot U \cdot \mu_i \tag{1.14}$$

Where:

t: is the number of days,

U: is expected number of users per day, which can be approximated by means of daily traffic volume,

μ: is the mean value of amount of physical and psychological impacts. It can be obtained by carrying out surveys on a population sample of the users.

For example, a survey using qualitative method on 300 users, who travel on a certain road link of 20 km, reveals a mean value $\mu_{physical} = 3.4$ (from a scale of 5) per 1 km. The daily users have a factor of 2.5 times the daily traffic volume, which is 600 vehicles/day. Thus, in 30 days, the values of $I^u_{physical} = 3.4$ will be determined as:

$I^u_{physical} = 30 \times (600 \times 2.5) \times 3.4 = 153'000$ units

If each unit of physical impact is relatively cost 0.05 *mus*, then in total 30 days, the cost incurred to users due to physical impact will be $0.05(mus/unit) \times 153'000(units) = 7'650$ (mus).

1.5.2.3.4 Environment preservation (reduction of noise)

Impact indicators associated with I^u_{noise} can be approximated by use of following equation.

$$I^u_{noise} = t \cdot \overline{dBA} \cdot U \tag{1.15}$$

Where:

t: are the number of days in between interventions and during intervention,

$\overline{dBA}$: is the expected increase unit of noise (in dBA).

U: is the expected number of users within a specific period.

For example, an intervention is scheduled to last 30 days. The expected numbers of users could be estimated as 300 persons. If during intervention, the noise due to construction activities increases by 2 dBA compared to the normal day without intervention, and one additional dBA has a value of 100 mus. Using Eq. (1.15), the values of impact indicator will be

$I^u_{noise} = 30 \times 2 = 18'000 dBA - person - days$

and thus, the values of function $g^d(d, x^u_{noise})$ will be

$g^d(d, x^u_{noise}) = 18'000 \times 100 = 1'800'000\ mus$

1.6 Assignments

1.6.1 Problem A

In class an impact hierarchy was given to be used to determine optimal intervention strategies and work programs for a road network. Discuss the suitability of this impacts hierarchy if it was to be used to determine optimal intervention strategies and work programs for

a) a private road network, and

b) a public rail network.

1.6.2 Answer A-a

No specific answer is provided. Please submit your answer to obtain feedback.

1.6.3 Answer A-b

No specific answer is provided. Please submit your answer to obtain feedback.

1.6.4 Problem B

Water is one of the fundamental needs of millions of people living in a megacity. Water of sufficient quality and quantity must be provided around the clock. In order to fulfill this need, a city depends on its water distribution infrastructure, which includes pipes made of different materials and laid at different times. These pipes are affected by processes of different types and deteriorate at different rates. The consequences related to pipe failure vary significantly depending on the type of failure, e.g. a pipe break, or a leaking pipe, as does the reaction time required to fix the pipe. For example, if a pipe breaks, a corrective intervention must be executed immediately, and if it is noticed that there is progressive water loss over time then a preventive intervention can be planned before there is an inadequate level of service. Part of the water distribution network in mega-city Q is shown in Figure 1.4. The pipe characteristics are given in Table 1.13.

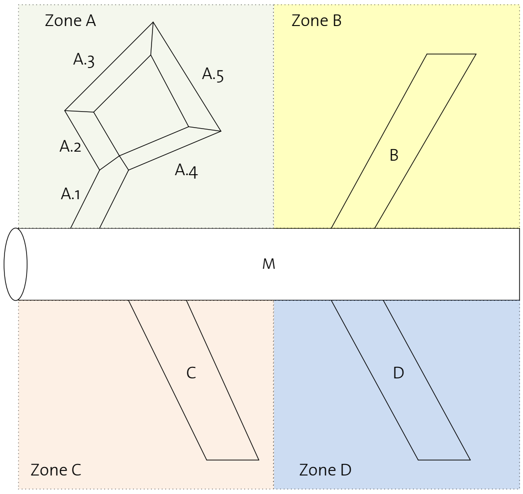


Fig. 1.4: Simplified network of water supply pipes

Table 1.13: Pipe and their attributes

Pipe	Material	Demand (m3/day)	Length (m)	Year of construction	Usage
M	Concrete		5’000	1998	Distribution only
A.1	PVC	250’000	600	2003	Distribution only
A.2	PVC		600	2003	Distribution only
A.3	PVC		600	2003	Distribution and connection to buildings
A.4	PVC		600	2003	Distribution only
A.5	PVC		600	1998	Distribution and connection to buildings
B	Cast iron type 1	200’000	2’500	1993	Distribution and connection to buildings
C	Cast iron type 2	100’000	1’600	1993	Distribution and connection to buildings
D	Cast iron type 3	350’000	4’000	1983	Distribution and connection to buildings

1.6.5 Question B

Make an impact hierarchy that you would use to measure the performance of the network. Be complete at the highest level, i.e. the stakeholders, and be detailed for one of the stakeholders besides the owner.

1.6.6 Answer B

No specific answer is provided. Please submit your answer to obtain feedback.

1.6.7 Problem C

1.6.8 Question C

Using the impact hierarchy for public roads in the script explain three processes that would result in a change in a required level of service defined using accident costs, travel time costs and maintenance costs.

1.6.9 Answer C

No specific answer is provided. Please submit your answer to obtain feedback.

References

Bryan T. Adey, Nam Lethanh, and Philippe Lepert. An impact hierarchy for the evaluation of intervention strategies for public roads. In *4th European Pavement and Asset Management Conference, 5-7 September, 2012, Malmo, Sweden*, September 2012.

W.K. Brauers. *Optimization Methods for a Stakeholder Society: A Revolution in Economic Thinking by Multi-objective Optimization*, volume 73 of *Nonconvex Optimization and Its Applications*. Kluwer Academic Publishers, 2003.

CUBRC. *Handbook of Economic Development*. CUBRC (Calspan-University of Buffalo Research Center) State University of New York, New York State Department of Transportation, 2001.

H. Craig Davis. *Regional economic impact analysis and project evaluation*. UBC Press, 2001.

Richard de Neufville and Stefan Scholtes. *Flexibility in Engineering Design*. MIT Press, 2011.

Bernhard Elsener. *Corrosion of Steel in Concrete: Materials Science and Technology*. Wiley, 2013.

C. S Kumares and L. Samuel. *Transportation decision making: principles of project evaluation and programming*. John Wiley and Sons, 2007.

Nam Lethanh, Bryan T. Adey, and Dilum N. Fernando. Optimal intervention strategies for multiple objects affected by manifest and latent deterioration processes. *Structure and Infrastructure Engineering*, 11(3):389–401, 2015. doi: 10.1080/15732479.2014.889178. URL http://dx.doi.org/10.1080/15732479.2014.889178.

Gunnar Lindberg. Calculating transport accident costs. Technical report, VTI, 1999.

H.P Lindenmann. Planning and measurement for preservation of road network (german name: Massnahmenplanung im erhaltungsmanagement von fahrbahnen). Technical Report 1259, Institute for Transport Planning and Systems, 2008.

OpusCL. *Review of VOC - Pavement Roughness Relationships Contained in Transfunds Project Evaluation Manual*. Opus Central Laboratories, Opus, Lower Hutt, New Zealand, rep. 529277.00 edition, 1999.

Y. Ouyang and S. Madanat. Optimal scheduling of rehabilitation activities for multiple pavement facilities: exact and approximate solutions. *Transportation Research Part A: Policy and Practice*, 38(5):347 – 365, 2004. ISSN 0965-8564. doi: DOI:10.1016/j.tra.2003.10.007. URL http://www.sciencedirect.com/science/article/B6VG7-4BSVNB4-1/2/a9f74a65d6101d9b289448f4e9bbe8fa. ref id=12.

A. Tarko, K.C. Shinha, S. Eranky, H. Brwon, E. Roberts, R. Scinteie, and S. Islam. Crash reduction factors for improvement activities in indiana. Technical report, Joint Transportation Research Program, Purdue University, West Lafayette, IN, 2000.

Chapter 2

Level of services indicators - reliability

Bryan T. Adey and Nam Lethanh

2.1 Introduction

In order to define an adequate LOS, to determine if an adequate LOS is provided, and to determine optimal intervention strategies and work programs, it would be best to focus on the maximizing the net benefit from the road infrastructure, as defined in an all-inclusive impact hierarchy over a specific time period. Due to the significant effort required to determine the values of some impact types, and collecting these values over time, proxies are often used. Four commonly used proxies are the reliability, availability, maintainability and safety of the infrastructure. Of these, reliability is discussed in the remaining part of this document. Availability and maintainability are discussed in the subsequent document. Safety refers to the likelihood of persons being injured or killed due to the infrastructure or use of the infrastructure. It is not discussed here further. The definitions of the three performance indicators discussed are:

- Reliability is the probability that an item to provide an adequate LOS.
- Availability is the proportion of time an item provides an adequate LOS.
- Maintainability is the ease with which an item can be maintained

The use of the word item in all definitions is to allow it to be replaced by any term that refers to physical infrastructure that is composed of smaller parts. For example, a network is comprised of objects, objects are comprised of elements, and elements are comprised of segments. In such a case, the word item in each of the three definitions could be replaced by either the word “network”, the word “object”, the word “element” or the word “segment”. For example, reliability is the ability of a network to provide an adequate level of service, or reliability is the ability of an object to provide an adequate level of service.

2.2 Reliability of an item

2.2.1 General

When managing infrastructure the reliability of an item is often a useful bit of information, i.e. how likely is it that an item is going to provide an adequate LOS for a specified period of time. If one speaks of the collapse of a bridge as being the first level of inadequate service, one often refers to structural reliability.

If the adequate LOS and the specified time period are defined, then one is interested in knowing the reliability of the item. If the adequate LOS is defined but not the specified time period then one is interested in knowing the amount of time required until an item provides an inadequate LOS. This is, often referred to as the service life of the item

It is important to realize that there is a tight connection between reliability, the definition of adequate LOS, and the service life of an item. For example, two identical items could have the same reliability, if one an adequate LOS is defined to be high and the service life short, and the other one had an adequate LOS defined to be low and the service life long.

2.2.2 Reliability of an item over a specified time period

The reliability of an item in the time interval from 0 to t is equivalent to the probability that the item provides an adequate LOS for a longer time than t, which can be expressed as:

$$R_i(t) = P(\tau_i > t) \tag{2.1}$$

Where:

τ: is the amount of time that an item provides an adequate LOS,

i: is the item identifier

When item i provides an adequate LOS for τ *tus,* which is a continuous positive random variable with c.d.f, τ_i represents the amounts of time that the items provide adequate LOSs.

Mathematically,

$$\begin{aligned} X_i(t) &= 1 \quad if \quad \tau_i > t \\ X_i(t) &= 0 \quad if \quad \tau_i \leq t \end{aligned} \tag{2.2}$$

When multiple items are being investigated $\boldsymbol{X(t)} = (X_1(t),\ldots X_n(t))$ is the item state vector at time t.

In the estimation of reliability of items over multiple time periods t, it is important to realize that it is the same items that survive each successive interval. In other words, it is assumed that the items provide an adequate LOS at t = 0, that there are no interventions executed on the items over the time period investigated, and that the probability of an item providing an inadequate LOS at t = 2 depends on it providing an adequate level of service until t = 2, i.e. it is conditional on the item surviving until t = 2.

2.2.3 Conditional probability of failure and failure rate

The probability of failure in the time interval $(t,\ t{+}\Delta]$, given that the item provides an adequate LOS at t, is the conditional probability of failure, as shown in Equation (2.3).

$$P(t \leq \tau \leq t + \Delta \,|\, \tau > t) = \frac{P(\text{failure appears in } (t, t+\Delta))}{P(\tau > t)} \tag{2.3}$$

If the τ is a positive random variable with d.f. $f(t)$ and c.d.f $F(t)$, then for small delta the conditional failure probability in $(t,\ t{+}\Delta]$ given the item is operational at t is, for small delta, approximately equal to the failure rate $f(t)\,/\,(1\text{-}F(t))$, *i.e.*

$$P(t \leq \tau \leq t + \Delta \mid \tau > t) = \frac{f(t) \cdot \Delta}{1 - F(t)} \qquad (2.4)$$

$$\text{Prob}(t \leq \tau \leq t + \Delta t | \tau \geq t) = \lambda(t)\Delta t = \frac{f(t)\Delta t}{\tilde{F}(t)} \qquad (2.5)$$

Where:

$f(t)$: is the probability that an item will not provide an adequate LOS at any instance from 0 to t,

$1 - F(t)$: is the probability that an item provides an adequate LOS from 0 to t. This is often referred to as the survival function.

The failure rate can then be expressed as:

$$\lambda(t) = \frac{f(t)}{\tilde{F}(t)} \qquad (2.6)$$

Note that $\lambda(t)$ is defined only for t such that $F(t) < 1$.

2.2.4 Failure rates

As can be seen from equation (2.5), the failure rate is a time dependent, i.e. it is not necessarily constant over time. For example, some engineering systems have failure rates that exhibit bathtub curve characteristics (Figure 2.1). It is common to divide these failure rates over time for these systems in three phases. In the first phase is there is a monotonic decrease in the failure rate, in the second phase the failure rate is more or less constant, and in the third phase the failure rate monotonically increases. When the failure can be considered to be constant, it can be

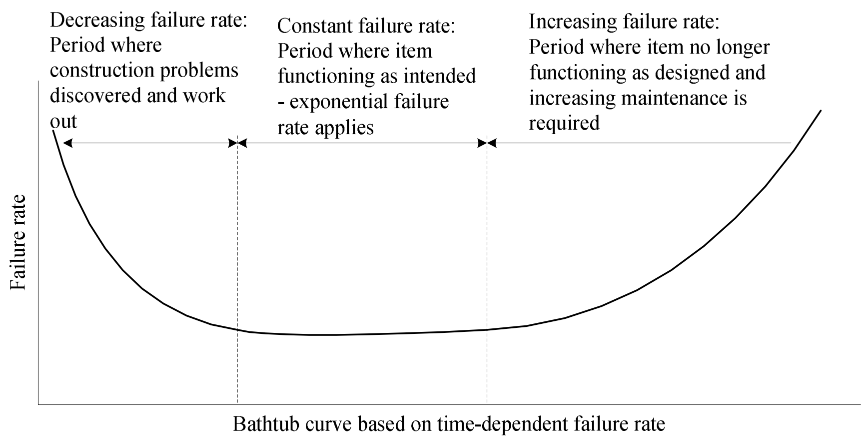


Fig. 2.1: Bathtub curve

approximated by assuming that the probability of failure can be modeled using an exponential distribution function. In this case,

$$\lambda(t) = \theta \qquad (2.7)$$

where the probability density function and the cumulative distribution function are:

$$f(t) = \theta \exp(-\theta \cdot t) \tag{2.8}$$
$$F(t) = 1 - \exp(-\theta \cdot t) \tag{2.9}$$

2.2.5 Failure rate and rate of occurrence of failures

Failure rate should not be confused with the rate of occurrence of failures. Failure rate is defined only for a random variable describing the lifetime of a non-renewable item. The rate of occurrence of failures is the derivative of the mean number of failures *[0,t]* with respect to *t* of a point process N*(t), t >= 0,* describes failures of a renewable item, i.e.

$$v(t) = \frac{E\left(N(t)\right)}{dt} \tag{2.10}$$

2.2.6 Reliability of an item with a time invariant failure rate

When an item has a constant failure rate, the reliability of the item, which is sometimes referred to as the survival probability of the item, is given as:

$$R(t) = \tilde{F}(t) = 1 - F(t) = \exp(-\theta \cdot t) \tag{2.11}$$

Example of how the reliability of the item changes over time for three different constant failure rates is given in Figure 2.2 and the corresponding probability density functions are given in Figure 2.3. Interestingly the reliability of items

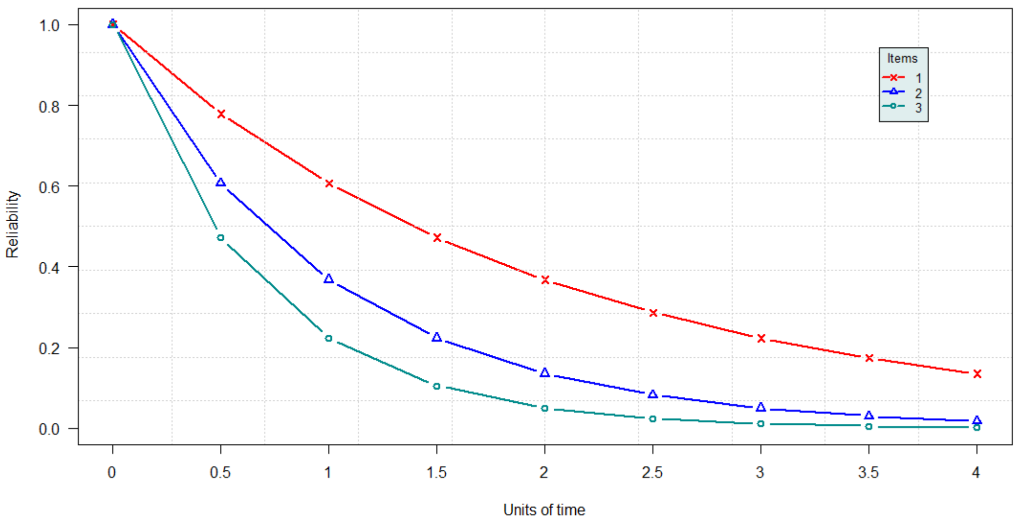


Fig. 2.2: Reliability

with constant failure rates, regardless of their values, is approximately 37% at the mean life of the item. This can be seen easily mathematically, since the length of time that an item provides an adequate LOS if it has a constant failure rate can be calculated as:

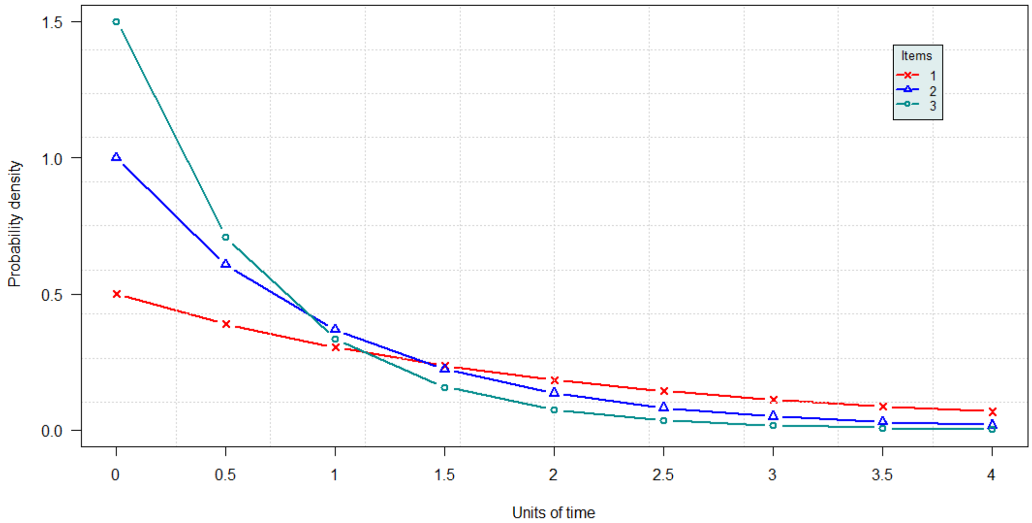


Fig. 2.3: Probability density function

$$\Theta = \int_0^\infty \tilde{F}(t)dt = \int_0^\infty \exp(-\theta \cdot t)dt = \frac{1}{\theta} \tag{2.12}$$

where θ is the failure rate and its reliability is then given by:

$$R(t) = R(\frac{1}{\theta}) = \exp(-\theta \cdot \frac{1}{\theta}) = \exp(-1) = 0.3678794 \approx 0.37 \tag{2.13}$$

No matter what the value of the failure rate, the reliability at the mean service life is 0.37, i.e. if the failure rate is 0.5, and, therefore, the mean service life is 2 *tus*, then the probability that the item provides an adequate LOS for longer than 2 *tus* is 0.37. This is illustrated graphically in Figure 2.4, for items with failure rates of 0.5, 1, 1.5, where the mean service lives are 2, 1 and 0.67. The codes for generating these graphs are given in Appendix 2.A.

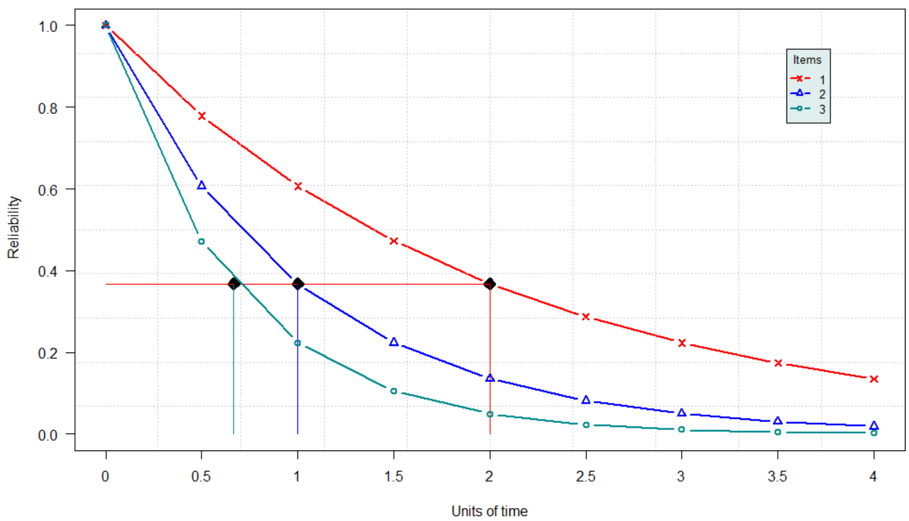


Fig. 2.4: Probable service life and reliability

2.2.7 Estimating the remaining amount of time an item will provide an adequate LOS

In many cases when it is know that an item has provided an adequate LOS until t it is desired to know how much longer it might continue to provide an adequate LOS. For example, if it known that an item provides an adequate LOS for 37 years and it has now been in service 37 years, what is the expected amount of time that it will continue to provide an adequate LOS. How to calculate this was explained in class.

2.2.8 Reliability of an item with time-variant failure rate

When an item has a time variant failure rate, the probability of failure cannot be modeled using an exponential distribution. If it is modeled with the Weibull function, a common selection, the reliability of the item can be estimated as:

$$R(t) = e^{-\alpha(t)^m} \tag{2.14}$$

where α and m are scale and shape parameters, respectively.

In this case the probability density function is:

$$f(t) = \alpha \cdot m \cdot t^{m-1} \cdot e^{-\alpha(t)^m} \tag{2.15}$$

And the failure rate

$$\lambda(t) = \alpha \cdot m \cdot t^{m-1} \tag{2.16}$$

Here, it can be see that if the value of $m = 1$, this equation is identical to the one used to estimate the constant failure rate.

2.3 Estimating reliability when LOS data available

The reliability of an item can be estimated in different ways. The way to be used depends on the availability of information and the composition or structure of the item (e.g. the item can be composed of many sub-items). If data is available it can be used as described above to estimate the failure rate of an item and then estimate the reliability of an item. An example is shown for a case where the failure is time invariant and a case where the failure is time variant.

2.3.1 Example 1: Reliability of an item with time invariant failure rate

A railway tests 10 switches, over 4’180 days of use under the same loads. It was observed that:

- Switch 1 failed in 75 days
- Switch 2 failed in 125 days

- Switch 3 failed in 130 days
- Switch 4 failed in 325 days
- Switch 5 failed in 525 days

The failure rate is estimated as

$$\theta = \frac{5}{4'180} = 0.001196$$

Assuming that this rate is constant, then the reliability of a switch of this type is estimated as:

$$R(t) = \exp(-\theta \cdot t) = \exp(-0.001196 \cdot t)$$

Which shows, for example, that the reliability of a switch, i.e. the probability that a switch will provide an adequate LOS for 1000 days, is 0.302.

2.3.2 Example 2: Reliability of an item with time-variant failure rate

The lights in street lamps frequently burn out. Regular inspections are done to see if the lights are function or not. The data shown in Table 2.1 has been collected. In the table, the column "time" is the units of time that the items have been in service. The column "status" refers to the state of the lighting, 1 is "burnt out" and 0 is "working". Using the

Table 2.1: Failure data of the lighting system

Item	time	status	Item	time	status	Item	time	status	Item	time	status	Item	time	status
1	9	1	6	28	0	11	161	0	16	12	1	21	33	1
2	13	1	7	31	1	12	5	1	17	16	0	22	43	1
3	13	0	8	34	1	13	5	1	18	23	1	23	45	1
4	18	1	9	45	0	14	8	1	19	27	1			
5	23	1	10	48	1	15	8	1	20	30	1			

Weibull distribution as described in equation (2.15), the value of θ and m are 0.027 and 1.097, respectively. These values are estimated using the Maximum likelihood estimation approach. The R code is given in Appendix 2.B.

With these values, the failure rate is given by:

$$\lambda(t) = 0.027 \times 1.097 \cdot t^{1.097-1}$$

And the probability density function and the reliability function are therefore:

$$f(t) = 0.027 \times 1.097 \cdot t^{1.097-1} \cdot e^{-(0.027 \cdot t)^{1.097}}$$

$$R(t) = e^{-(0.027 \cdot t)^{1.097}}$$

The evolution of reliability and failure rate over time are shown in Figure 2.5

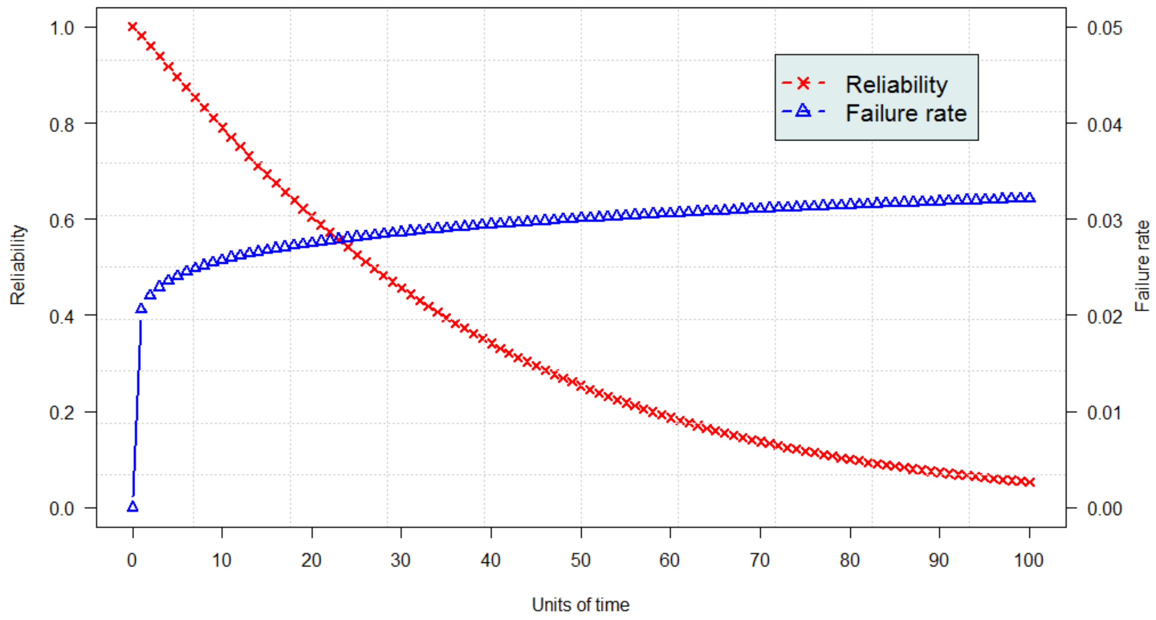


Fig. 2.5: Reliability and failure rate - Weibull distribution function

2.4 Estimating reliability when LOS data not available

There are many cases where the LOS data is not available, e.g. the failure of a bridge. One way to estimate the reliability of the item in this case is to model the uncertainty in the variables of the models used and compare the expected resistance to the expected loads.

2.4.1 Based on models and uncertainty of variables

In this case a model of the performance of the infrastructure object is selected and laboratory tests or simulations are done to determine probabilistic distributions to represent the uncertainties in key parameters in the models. For example, if one was to estimate the reliability of a bridge with respect to landslides, one almost certainly does not have enough data to say with any confidence the reliability of a bridge in the future, so one may embark upon determining the probability that the bridge would continue to provide an adequate LOS if it came in contact with specific volumes of rock. This is done by modeling the performance of the bridge if it came in contact with rock falls of different intensities, taking into consideration, for example, the strength of the concrete, the strength of the reinforcement, the location of the reinforcement, and how the elements in the bridge work together as a system. Such information is often included in fragility curves. An example is given in Figure 2.6 and more can be found in Schultz et al. [2010] and Lethanh et al. [2015] The fragility curves shown in Figure 2.6 represent the probability (vertical axis) of bridge deck failure, taking into consideration the initial condition of the bridge deck, i.e. the manifest deterioration state of the bridge deck (i = 1, 2, and 3) prior to hazard occurrence given the volume of rock fall that comes in contact with the deck. As can be seen from the figure, if 15 m^3 of rock come in contact with a bridge deck in state 1, there is a 0.089 probability (point A) that the bridge deck would fail. It can also be seen that the higher the initial state the higher the probability of bridge deck failure. e.g. when the bridge deck is initially in state 2 and state 3 the probability that the bridge deck would fail is 0.372 and 0.716, respectively).

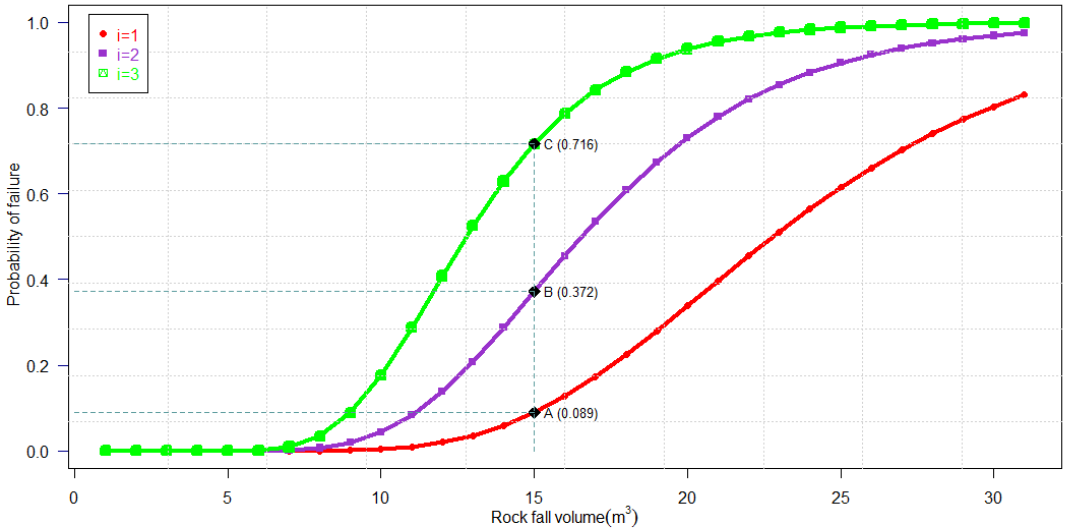


Fig. 2.6: Example bridge deck fragility curves

The probability of each failure state, given the intensity of the rock fall, which are indicated with different curves in Figure 2.6, are given by

$$\Omega_{s,i}^{l} = prob\left[CS > l|s,i\right] \tag{2.17}$$

where the probability of exceeding condition state is conditional on the intensity of the load effect event and the condition state of the object, immediately prior to occurrence of the load effect event.

Of course, to estimate the reliability of the bridge, in general, the probability of occurrence of the load effect event itself has to be estimated, i.e.

$$H_s(t) = Prob\left[\text{intensity} > s(0,t)\right] \tag{2.18}$$

2.4.2 Based on reliability of sub-items

When data does not exist on the LOS of the item itself, but it does on the sub-items of which it is composed, the sub-items can be combined in ways to allow the estimation of the reliability of the item. Four basic ways to combine things are as sub-items in series, sub-items in parallel, parallel sub-items connected in series, and series sub-items connected in parallel. These are explained in the following four sub-sections.

2.4.2.1 Sub-items in series

If the sub-items are in series (e.g. Figure 2.7) and the LOS provided by the item is equal to the LOS provided by the worst sub-item, then the LOS provided by the item is given by:

$$\Phi(\vec{x}) = \prod_{i=1}^{n} x_i = \min_{1 \le i \le n} x_i \tag{2.19}$$

where x_i is the state of sub-item i, i = $1,\dots n$, and it is represented as a binary variable, where 1 indicates the sub-item provides an adequate LOS and the 0 indicates that it does not.

The reliability that the item will provide an adequate LOS is, therefore, given by:

$$\mathbf{R}(t) = \prod_{n=1}^{N} R_n(t) \tag{2.20}$$

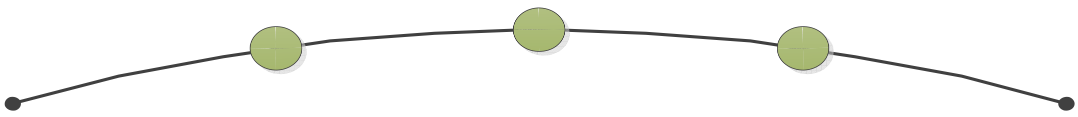


Fig. 2.7: Sub-items in series

for example, a bridge can be considered to provide an adequate LOS if and only if the abutment 1 and 2, and the deck provide adequate LOSs, and the probabilities of each providing an inadequate LOS in one year are 0.01, 0.02 and 0.03, respectively, and their failure rates are constant, then the reliabilities of the bridge over 5 years are:

$$\begin{aligned}\mathbf{R}(t=5) = \prod_{n=1}^{N} R_n(5) &= R_{abutment1}(5) \cdot R_{deck}(5) \cdot R_{abutment2}(5) \\ &= e^{-0.02\times 5} \cdot e^{-0.01\times 5} \cdot e^{-0.03\times 5} = e^{-(0.02+0.01+0.03)\cdot 5} \approx 0.74\end{aligned} \tag{2.21}$$

2.4.2.2 Sub-items in parallels

If the sub-items are in parallel (e.g. Figure 2.8), and the LOS provided by the item is equal to the LOS provided by the best sub-item, then the LOS provided by the item is given by:

$$\Phi(\overrightarrow{x}) = 1 - \prod_{i=1}^{n}(1-x_i) = \max_{1\leq i\leq n} x_i \tag{2.22}$$

The reliability that the item will provide an adequate LOS is, therefore, given by:

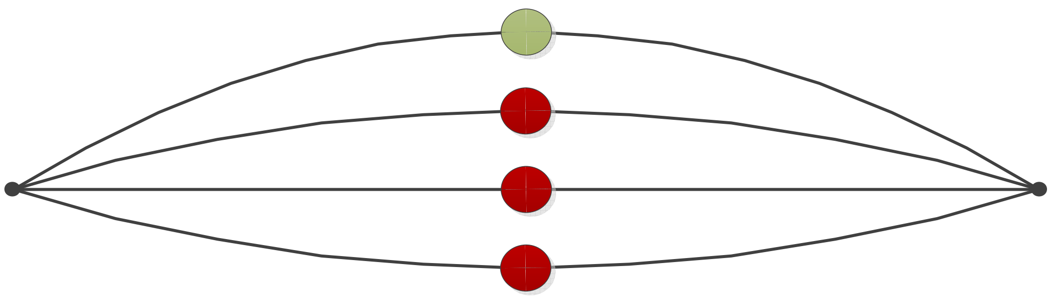


Fig. 2.8: Sub-items in parallel

$$\mathbf{R}(t) = 1 - \prod_{i=1}^{N} (1 - R_n(t)) \tag{2.23}$$

2.4.2.3 Parallel sub-items connected in series

In some cases items are comprised of sub-items that can be thought of as being groups of sub-items connected in parallel where the groups are then connected in series (e.g. Figure 2.9). In this case the LOS provided by the item is connected to the LOS provided by the sub-items as follows:

$$\Phi(\mathbf{x}) = [1 - (1 - x_1) \cdot (1 - x_2)] \cdot [1 - (1 - x_3) \cdot (1 - x_4)] \tag{2.24}$$

The reliability that the item will provide an adequate LOS is, therefore, given by:

Fig. 2.9: Parallel sub-items connected in series

$$\mathbf{R} = [1 - (1 - R_1) \cdot (1 - R_2)] \cdot [1 - (1 - R_3) \cdot (1 - R_4)] \tag{2.25}$$

2.4.2.4 Series sub-items connected in parallel

In some cases items are comprised of sub-items that can be thought of as being groups of sub-items connected in series where the groups are then connected in parallel (e.g. Figure 2.10). In this case the LOS provided by the item is connected to the LOS provided by the sub-items as follows:

$$\Phi(\mathbf{x}) = 1 - (1 - x_1 \cdot x_2) \cdot (1 - x_3 \cdot x_4 \cdot x_5) \tag{2.26}$$

The reliability that the item will provide an adequate LOS is, therefore, given by:

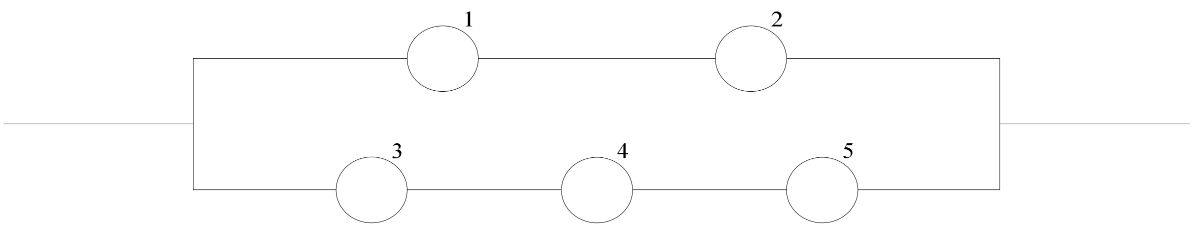


Fig. 2.10: Series sub-items connected in parallel

$$\mathbf{R} = 1 - (1 - R_1 \cdot R_2) \cdot (1 - R_3 \cdot R_4 \cdot R_5) \tag{2.27}$$

2.4.2.5 Complex configurations

As not all items can be thought of as neat combinations of series and parallel sub-items, it is useful to have a systematic way to determine the structure function for the item. Three ways are pivotal decomposition, minimal paths and minimal cuts.

2.4.2.5.1 Pivotal decomposition

Pivotal decomposition is a method to determine the reliability of an item by replacing unreliable sub-items with either 100% reliable sub-items and/or 0% reliable (i.e. failed) sub-items until a configuration is created for which the reliability is easily computable. The reliability of the item is then calculated as:

$$R(\overrightarrow{x}) = R_i \cdot R(1_i, \overrightarrow{x}) + (1 - R_i) \cdot R(0_i, \overrightarrow{x}) \tag{2.28}$$

Where $(\alpha_i, \overrightarrow{x})$ represents the vector of sub-items with $\overrightarrow{x}$ the sub-item i replaced by α_i. , which is either totally reliable and has the value of 1 or failed and has the value of 0.

2.4.2.5.2 Minimal path and minimal cuts

In the case of minimal cuts, it is necessary to consider

- the vector of sub-item states a *cut vector* if φ(x) = 0,
- the specific sub-items in the cut vector that are required to not provide an adequate LOS to mean that the item does not provide an adequate LOS, form the *cut set*, and
- a cut set in which there is no possibility to improve the LOS provided by a sub-item without improving the LOS provided by the item is a *minimal cut set*.

In the case of minimal paths, it is necessary to consider

- the vector of sub-item states a *path vector* if φ(x) = 1
- the specific sub-items in the path vector that are required to provide an adequate LIS to mean that the item provides an adequate LOS, form the *path set*, and
- a path set in which there is no possibility to reduce the LOS provided by a sub-item without reducing the LOS provided by the item is a *minimal path set*.

2.4.2.5.3 Example

Question

Define the structure function to be used to estimate the reliability of the item composed of 5 sub-items as shown in Figure 2.11.

Answer

Using the pivotal decomposition, the item structure function can be determined by substituting sub-item 3 with a sub-item with a reliability of 1 or 0, yielding either equation (2.29) or (2.30),respectively. These are illustrated in Figure

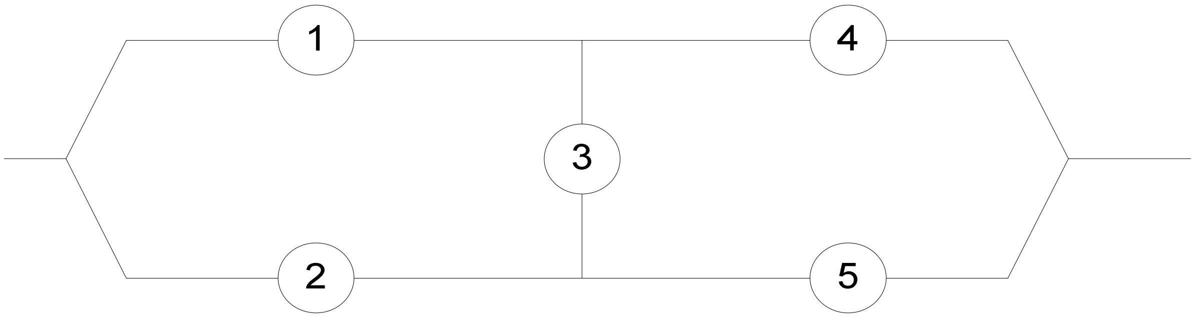


Fig. 2.11: Item with 5 sub-items in a complex structure

2.12 and Figure 2.13, respectively

$$R(1_3, x) = [1-(1-x_1)(1-x_2)]\cdot[1-(1-x_4)(1-x_5)] \tag{2.29}$$

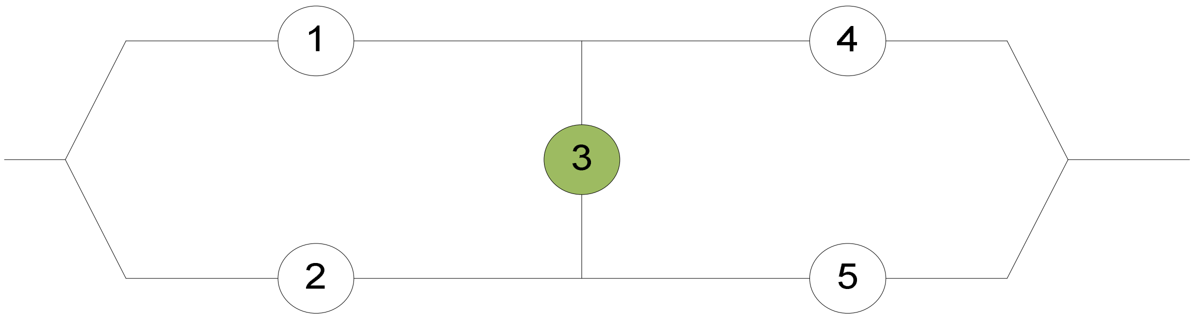


Fig. 2.12: Sub-item 3 provides an adequate LOS

$$R(0_3, \mathbf{x}) = [1-(1-x_1\cdot x_4)\cdot(1-x_2\cdot x_5)] \tag{2.30}$$

This means that the structure function of the item to be used to estimate the reliability of the item is:

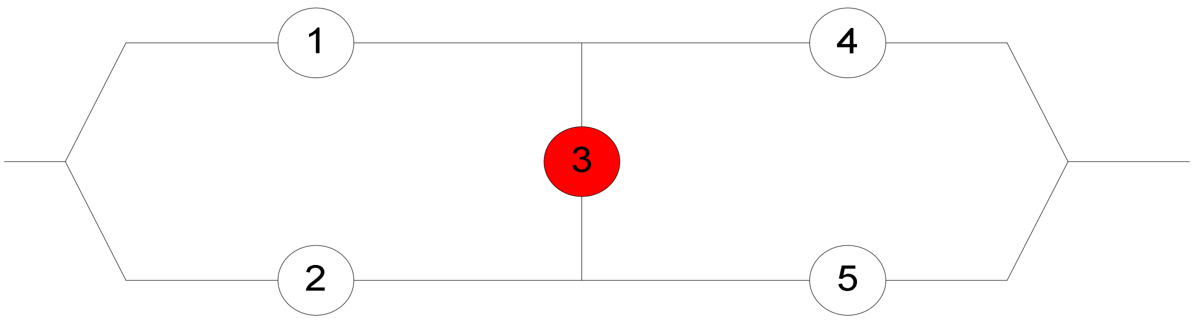


Fig. 2.13: Sub-item 3 does not provide an adequate LOS

$$R(\mathbf{x}) = R_3\cdot R(1_3, \mathbf{x}) + (1-R_3)\cdot R(0_3, \mathbf{x}) \tag{2.31}$$

$$R(\mathbf{x}) = R_1R_3R_4 + R_1R_3R_5 + R_2R_3R_4 + R_2R_3R_5 + R_1R_2R_4R_5 \\ - R_1R_2R_3R_4 - R_1R_3R_4R_5 - R_1R_2R_3R_5 - R_2R_3R_4R_5 \tag{2.32}$$

Question

Show that the minimal path method will give the same answer as pivotal decomposition.

Answer

The “minimal path sets” include

$$P_1 = (1,4)$$
$$P_2 = (1,3,5)$$
$$P_3 = (2,5)$$
$$P_1 = (2,3,4)$$

$$\begin{aligned}
\phi(\mathbf{x}) &= 1 - \prod_{s=1}^{S}\left(1 - \prod_{n\in P_s} x_n\right)\\
&= 1 - (1-x_1x_4)(1-x_1x_3x_5)(1-x_2x_5)(1-x_2x_3x_4)\\
&= x_1x_3x_4 + x_1x_3x_5 + x_2x_3x_4 + x_2x_3x_5 + x_1x_2x_4x_5\\
&\quad - x_1x_2x_3x_4 - x_1x_3x_4x_5 - x_1x_2x_3x_5 - x_2x_3x_4x_5
\end{aligned} \tag{2.33}$$

2.5 Conclusion

Once adequate LOSs are defined, one is often interested in estimating the likelihood that they will be provided. Without thinking of the specific consequences that might occur, reliability can be used as a performance indicator. Keep in mind though that the consequences related to the occurrence of inadequate services are almost never constant, and, therefore, focusing exclusively on reliability as a performance indicator will not allow the prioriotization of activities to maximize net benefits related to infrastructure. An example is included at the end of the next chapter.

When using reliability as a performance indicator for an item, there are multiple ways to estimate the reliability. When using reliability as a performance indicator for a sub-item one should not forget that the LOS it provides depends on how it is integrated into the item as a whole and the reliability of the other sub-items in the item. There are different ways to measure this that vary in terms of accuracy and effort.

2.6 Assignments

2.6.1 Problem A

2.6.1.1 Question A1

Determine the reliability of the road network shown in Figure 2.14, when the reliability of each object is as given in Table 2.2.

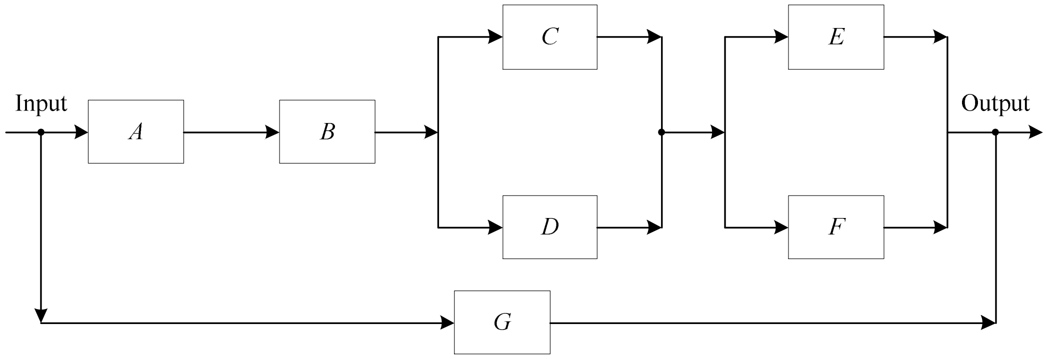


Fig. 2.14: A network of 7 objects

Table 2.2: Object reliabilities

Object	Reliability
A	0.95
B	0.97
C	0.92
D	0.94
E	0.90
F	0.88
G	0.98

2.6.1.2 Answer A1

Seeing that the network can be broken down into series and parallel networks and remembering that the structure function for series, and parallel objects are as given in Equations (2.20) and (2.23), respectively.

$$\mathbf{R}_C = 1 - [1 - (R_A \cdot R_B \cdot R_{CD} \cdot R_{EF})] \cdot [1 - R_G] \tag{2.34}$$

where R_{CD}, R_{EF} represents the reliability of the parallel components composed of item C, D and E, F, respectively. $\mathbf{R}_{Network}$ represents reliability of the network configuration.

$$\begin{aligned} &\mathbf{R}_{Network} = 1 - [1 - (R_A \cdot R_B \cdot R_{CD} \cdot R_{EF})] \cdot [1 - R_G] \\ &= 1 - [1 - (R_A \cdot R_B \cdot \{(1 - (1 - R_C) \cdot (1 - R_D)\} \cdot \{(1 - (1 - R_E) \cdot (1 - R_F)\})] \cdot [1 - R_G] \end{aligned} \tag{2.35}$$

Using the value of reliability for each item shown in Figure **??**, following result is obtained

$$\mathbf{R}_{Network} = 0.998121$$

2.6.2 *Problem B*

Imagine that the infrastructure organization offers bid to external contractors for construction and installation of its infrastructure network. The networks consist of 17 objects (Table 2.3) and, minimum reliability of the network reliability must be 70%. The A, B and C bid 57'000 CHF, 39'000 CHF, and 42'000 CHF, respectively and proposed the configuration shown in Figure 2.15, Figure 2.16, and Figure 2.16.

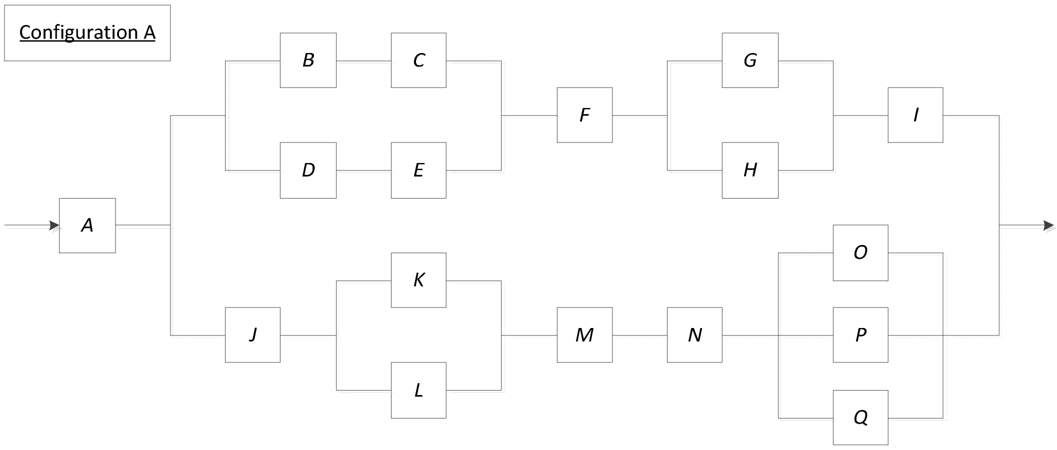


Fig. 2.15: Configuration A

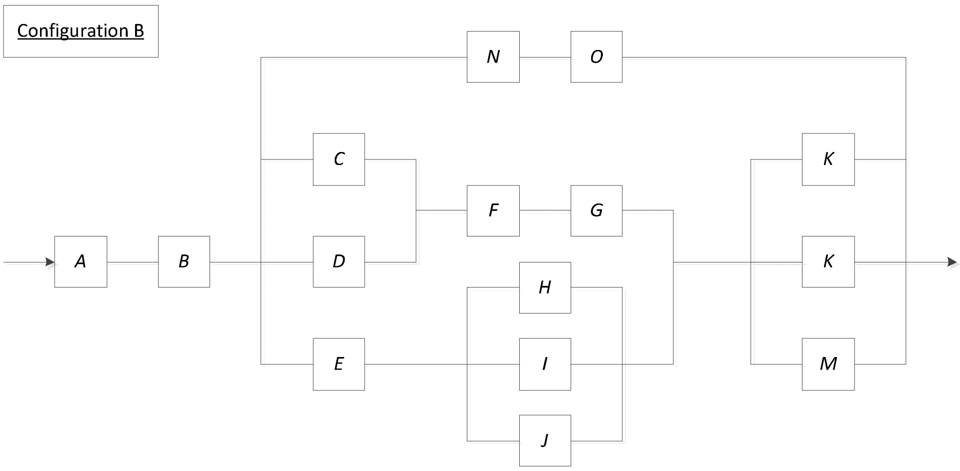


Fig. 2.16: Configuration B

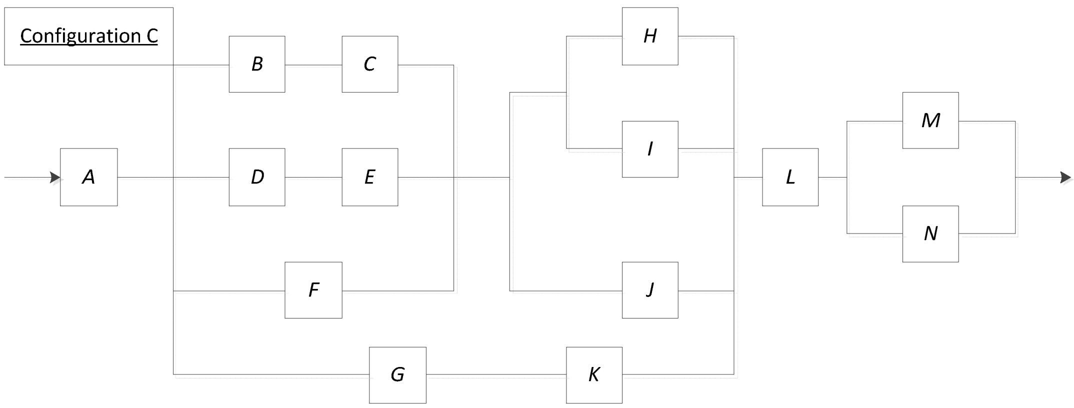


Fig. 2.17: Configuration B

2.6.2.1 Question B1

Given this information, determine the system reliability for each of the three configurations.

Table 2.3: Object reliabilities

Object	Reliability	Object	Reliability	Object	Reliability
A	0.840	G	0.870	M	0.830
B	0.860	H	0.880	N	0.850
C	0.890	I	0.890	O	0.840
D	0.860	J	0.860	P	0.890
E	0.870	K	0.850	Q	0.890
F	0.820	L	0.860		

2.6.2.2 Answer B1

Configurations A, B, and C are composed of both items connected in series and parallel. Equations (2.20) and (2.23) are used.

One possible formula to be used to determine the reliability of configuration A, B and C are:

Configuration A

- B and C are in the same link (Eq. (2.20))
- D and E are in the same link (Eq. (2.20))
- {B,C} and {D, E} are in parallel (Eq. (2.23))
- G and H are in parallel (Eq. (2.23))
- Set 1 = [{(B,C),(D,E)}, F, (G,H), and I] are in the same link (Eq. (2.20))
- K and L are in parallel (Eq. (2.23))
- O, P, and Q are in parallel (Eq. (2.23))
- Set 2 = [J, (K,L), M, N, and (O,P,Q)] are in the same link (Eq. (2.20))
- Set 1 and 2 are in parallel (Eq. (2.23))

$$\mathbf{R}_{configurationA} = R_A \cdot R_{B\to Q} \tag{2.36}$$

where $R_{B\to Q}$ represents the the reliability of the sub-system

$$R_{B\to Q} = 1-(1-R_{B\to I})\cdot(1-R_{J\to Q}) \tag{2.37}$$

$$\begin{aligned} R_{B\to I} &= R_{B\to E}\cdot R_F\cdot R_{G\to H}\cdot R_I \\ &= [1-(1-R_B\cdot R_C)\cdot(1-R_D\cdot R_E)]\cdot R_F\cdot[1-(1-R_G)\cdot(1-R_H)]\cdot R_I \end{aligned} \tag{2.38}$$

$$\begin{aligned} R_{J\to Q} &= R_J\cdot R_{K\to L}\cdot R_M\cdot R_N\cdot R_{O\to Q} \\ &= R_J\cdot[1-(1-R_K)\cdot(1-R_L)]\cdot R_M\cdot R_N\cdot[1-(1-R_O)\cdot(1-R_P)\cdot(1-R_Q)] \end{aligned} \tag{2.39}$$

Using the reliability values of items in Table 2.3, the reliability of the configuration A is

$$\mathbf{R}_{configurationA} = 0.729$$

Configuration B

$$\mathbf{R}_{configurationB} = R_A \cdot R_B \cdot R_{C\to O} \tag{2.40}$$

$$R_{C\to O} = 1-(1-R_{C\to J}\cdot R_{K\to M})\cdot(1-R_{N\to O}) \tag{2.41}$$
$$R_{C\to J} = 1-(1-R_C\cdot R_F\cdot R_G)(1-R_E\cdot R_{H\to J}) \tag{2.42}$$
$$R_{H\to J} = 1-(1-R_H)\cdot(1-R_I)\cdot(1-R_J) \tag{2.43}$$
$$R_{K\to M} = 1-(1-R_K)\cdot(1-R_L)\cdot(1-R_M) \tag{2.44}$$
$$R_{N\to O} = R_N\cdot R_O \tag{2.45}$$

Using the reliability values of items in Table 2.3, the reliability of the configuration B is

$$\mathbf{R}_{configurationB} = 0.714$$

Configuration C

$$\mathbf{R}_{configuration\mathrm{C}} = R_A \cdot R_{B\to K}\cdot R_L\cdot R_{M\to N} \tag{2.46}$$

$$R_{B\to K} = 1-(1-R_{B\to F}\cdot R_{H\to J})\cdot(1-R_{G\to K}) \tag{2.47}$$
$$R_{B\to F} = 1-(1-R_B\cdot R_C)\cdot(1-R_D\cdot R_E)\cdot(1-R_F) \tag{2.48}$$
$$R_{H\to J} = 1-(1-R_H)\cdot(1-R_I)\cdot(1-R_J) \tag{2.49}$$
$$R_{G\to K} = R_G\cdot R_K \tag{2.50}$$

Using the reliability values of items in Table 2.3, the reliability of the configuration C is

$$\mathbf{R}_{configurationC} = 0.702$$

Result

The reliability of configurations A, B and C are 0.729, 0.714, and 0.702 respectively. Based on reliability alone configuration A would be chosen.

2.6.2.3 Question B2

How high would the failure costs have to be before you would not choose the network with the lowest upfront costs? (Use a one year time period and assume that the network would not failure more than once).

2.6.2.4 Answer B2

If failure costs nothing than network configuration results in the lowest cost (Table 2.4).

Table 2.4: Total cost of networks

Failure cost (10^3 x CHF)	Network	Reliability	Construction cost (10^3 x CHF)"	Total cost in one year (10^3 x CHF)
0	A	0.729	57	57
	B	0.714	39	39
	C	0.702	42	42
900	A	0.729	57	301
	B	0.714	39	296
	C	0.702	42	310
1000	A	0.729	57	328
	B	0.714	39	325
	C	0.702	42	340

2.6.3 *Problem C*

You have a series network with 5 infrastructure objects, two pavement sections and three bridges. The probability of failure in each of the next 5 years of each bridge is 0.1 (before and after repair) and the probability of failure for each road section is 0.15 (before and after repair). Each bridge failure will take 60 days to repair and will cost 250'000 mus during which time the road will be unusable in which users will incur 300'000 mus. Each road section failure will take 20 days to repair and will cost 75'000 mus during which time the road will be unusable in which users will incur 100'000 mus. Failures are considered mutually exclusive and that only one failure can occur per year.

2.6.3.1 Question C1

Calculate the reliability of the road over the 5 year-period.

2.6.3.2 Answer C1

The reliability is defined (2.11). The hazard rate θ can be estimated.

$$ln[R(t)] = \ln[\exp(-\theta \cdot t)] \Leftrightarrow \ln[R(t)] = -\theta.t \tag{2.51}$$

$$\rightarrow \theta = -\frac{\ln[R(t)]}{t} \tag{2.52}$$

The reliability in one year of the network $R(t = 1)$is estimated as

$R(t = 1) = (1 - 0.1)^3 \cdot (1 - 0.15)^2 = 0.5267025$

The join hazard rate will be

$\rightarrow \theta = -\frac{\ln[0.5267025]}{1} = 0.6411194$

The reliability in 5 year of the network will be

$R(t=5) = \exp(-0.6411194 \times 5) = 0.0405347$

2.6.3.3 Question C2

Calculate the maintainability of the road (in terms of time and money).

2.6.3.4 Answer C2

The formulation to estimate the maintainability is

$$\bar{M} = \frac{\sum_{n=1}^{N} p_n \cdot N \cdot \mu_n^M}{\sum_{n=1}^{N} p_n \cdot N} \tag{2.53}$$

Where:

p_n: is the failure probability of item type n,

N: is the number of item type n

μ_n^M: is the mean time of impacts occur during the maintenance

Thus, in term of time, we have

$$\bar{M}^{time} = \frac{0.1 \times 3 \times 60}{0.1 \times 3 + 0.15 \times 2} + \frac{0.15 \times 2 \times 20}{0.1 \times 3 + 0.15 \times 2} = 40\text{days}$$

and in term of money, we have

$$\bar{M}^{money} = \frac{0.1 \times 3 \times (250'000 + 300'000)}{0.15 \times 2 + 0.1 \times 3} + \frac{0.15 \times 2 \times (75'000 + 100'000)}{0.15 \times 2 + 0.1 \times 3} = 362'500\text{mus}$$

2.6.3.5 Question C3

What is the impact on reliability and maintainability of repairing one bridge perfectly, i.e failure will not occur?

2.6.3.6 Answer C3

When one bridge is perfectly repaired, its reliability is 1 and thus the network can be considered to have only two bridges instead of 3. The same way of calculation for a new configuration of the network with only 2 bridges and 2 road sections is used to compute the reliability and maintainability.

The reliability in one year of the network $R(t=1)$ is estimated as

$R(t=1) = (1-0.1)^2 \times (1-0.15)^2 = 0.585225$

$\rightarrow \theta = -\dfrac{\ln[0.585225]}{1} = 0.5357589$

The reliability in 5 year of the network will be

$R(t=5) = \exp(-0.5357589 \times 5) = 0.06864586$

$\bar{M}^{time} = \dfrac{0.15 \times 2 \times 60}{0.15 \times 2 + 0.1 \times 2} + \dfrac{0.1 \times 2 \times 20}{0.15 \times 2 + 0.1 \times 2} = 36\text{days}$

and in terms of money, we have

$\bar{M}^{money} = \dfrac{0.15 \times 2 \times (250'000 + 300'000)}{0.15 \times 2 + 0.1 \times 2} + \dfrac{0.1 \times 2 \times (75'000 + 100'000)}{0.15 \times 2 + 0.1 \times 2} = 325'000\text{mus}$

2.6.3.7 Question C4

How much should we be willing to pay to execute such an intervention?

2.6.3.8 Answer C4

The amount of money that we are willing to pay depending on the expected impact due to failure that can be generated under the two above options. Following table describes the calculation of the amount of money that we are willing to pay. The option 1 and 2 are corresponding to the case one bridge is not perfectly repaired and another case when a bridge is perfectly repaired, respectively. The reliability and the expect cost to repair are calculated from previous steps. From the reliability, the failure probability can be straightforward estimated and it is used to compute the possible cost (expected cost) if failure happens in the future. As can be seen from the amount of possible cost under option 1 and 2, option 2 yields less amount of cost than option 1, and the willingness to pay is the difference of the twos, which is about 45.115 mus.

Table 2.5: Comparison of two options

Option	Reliability	Cost (mus)	Failure probability
1	0.0405347	362’500	0.9594653
2	0.06864586	325’000	0.93135414
	Willingness to pay = 45.115 mus		

2.6.3.9 Question C5

Discuss the assumptions mutual exclusiveness and that only one failure can occur per year.

2.6.3.10 Answer C5

The mutual exclusiveness assumed in this example means that only one failure can occur per year. This assumption can be used but it also bears a limitation that it might not perfectly reflect the increasing failure rate over time as the bridges and road sections are aging year by year. In addition, as the bridges and road sections are independent, they can actually failure to provide adequate level of services simultaneously in the same year. This possibility should be taken into consideration of calculating the reliability and maintainability in reality.

2.6.4 Problem D

A network is connected and represented as shown in the following Figure 2.18. All 6 objects are identical. In one month, on average it is recorded that 1 out of 6 objects fails and is renewed.

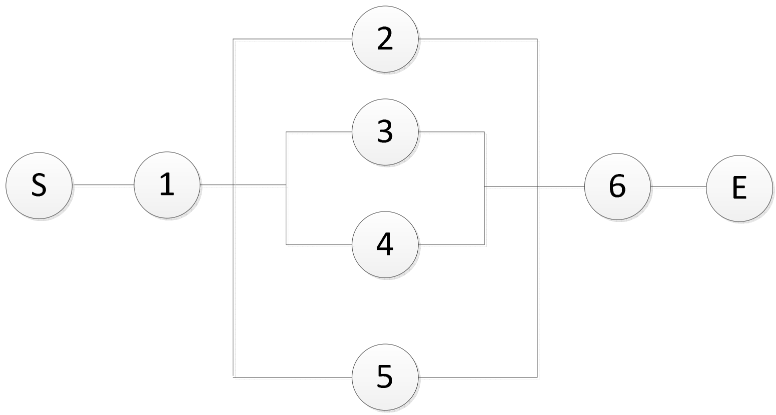


Fig. 2.18: Network

2.6.4.1 Question D1

Calculate the reliability of the network in 5 months, i.e. that it is possible to send one item from the start node (S) to the end note (E)? Note: you can round up the values in 2 digits.

2.6.4.2 Answer D1

One month failure probability for each object is 1/6. The reliability for the network in one month is defined as

$$\mathbf{R}^1_{network} = R^1_1 \cdot R^1_{2\to5} \cdot R^1_6 \quad (2.54)$$

$$R^1_{2\to5} = 1-(1-R^1_2)\cdot\left[\left\{1-(1-R^1_3)\cdot(1-R^1_4)\right\}\right]\cdot(1-R^1_5) \quad (2.55)$$

$$\mathbf{R}^1_{network} \approx 0.694$$

The reliability in 5 months of the network will be estimated based on the failure rate of individual item

$$\to \theta = -\frac{\ln[1-1/6]}{1} \approx 0.182$$

The reliability of each item in 5 months is

$$R^5_{item} = \exp(-0.183\times5) \approx 0.402$$

Using equations (2.54) and (2.55) for a period of 5 months, we have

$$\mathbf{R}^5_{network} = R^5_1 \cdot R^5_{2\to5} \cdot R^5_6 \tag{2.56}$$
$$R^5_{2\to5} = 1-(1-R^5_2)\cdot\left[\left\{1-(1-R^5_3)\cdot(1-R^5_4)\right\}\right]\cdot(1-R^5_5) \tag{2.57}$$
$$\mathbf{R}^5_{network} \approx 0.141$$

2.6.4.3 Question D2

If the network is not working, a loss of 30 *mus* is expected. In order to improve the overall reliability of the network over the next 5 months, manager proposes 2 options

Option 1: Replace link 1 and link 6 with links that are 5 % more reliable than the existing ones.

Option 2: Replace links 2, 3, 4, and 5 with the links that are 8% more reliable than the existing ones.

The differences in replacement costs are negligible and thus not considered.

Which option shall be chosen?

2.6.4.4 Answer D2

Option 1: Replace link 1 and link 6 with links that are 5 % more reliable than the existing ones.

Under this option, new reliability of link 1 and link 6 become

$$R^1_1 = R^1_6 = (1-\frac{1}{6}) + (1-\frac{1}{6})\cdot\frac{5}{100} = 0.875$$

Using the new value of reliability for item 1 and 6 and the values of reliability for item 2, 3, 4, and 5 ($R^1_2 = R^1_3 = R^1_4 = R^1_5 = (1-\frac{1}{6})$) for the equations (30) and (31), we have

$$\mathbf{R}^{1,option1}_{network} \approx 0.765$$

Option 2: Replace links 2, 3, 4, and 5 with the links that are 8% more reliable than the existing ones.

Under this option, new reliability of link 1 and link 6 become

$$R^1_2 = R^1_3 = R^1_4 = R^1_5 = (1-\frac{1}{6}) + (1-\frac{1}{6})\cdot\frac{8}{100} = 0.9$$

Using the new value of reliability for item 2, 3, 4 and 5 and the old values of reliability for item 1, and 6 ($R^1_1 = R^1_6 = (1-\frac{1}{6})$) for the equations (30) and (31), we have

$$\mathbf{R}_{network}^{1,option1} \approx 0.694$$

Similarly, for 5 months, we have

$$R_1^5 = R_6^5 \approx 0.402 + 0.402 \cdot \frac{5}{100} = 0.4221$$

Using the new value of reliability for item 1 and 6 and the values of reliability for item 2, 3, 4, and 5 ($R_2^5 = R_3^5 = R_4^5 = R_5^5 = 0.402$) for the equations (30) and (31), we have

$$\mathbf{R}_{network}^{5,option1} \approx 0.155$$

Option 2: Replace links 2, 3, 4, and 5 with the links that are 8% more reliable than the existing ones.

Under this option, new reliability of link 1 and link 6 become

$$R_2^2 = R_3^3 = R_4^4 = R_5^5 = 0.402 + 0.402 \cdot \frac{8}{100} = 0.434$$

Using the new value of reliability for item 2, 3, 4 and 5 and the old values of reliability for item 1, and 6 ($R_1^5 = R_6^5 = 0.402$) for the equations (30) and (31), we have

$$\mathbf{R}_{network}^{5,option2} \approx 0.145$$

Following table compares the expected impact under the two options As can be seen from the table, under both 1

Table 2.6: Comparison of two options

Time (months)	Options	Reliability	Failure probability	Unit impact (mus)	Total impact (mus)
1	1	0.765	0.235	30	7.05
	2	0.694	0.306	30	9.18
5	1	0.155	0.845	30	25.35
	2	0.145	0.855	30	25.65

month and 5 months the expected total impacts incurred by option 1 is smaller than that of option 2, thus, option 1 is preferable than option 2.

2.6.5 Problem E

You are evaluating two offers for replacement of an existing bridge on a highway. You have determined that the link to which the existing bridge belongs has a reliability of 0.9 over the time period of interest, if it is assumed that the bridge is perfectly reliable. The consequences of not being able to use the link due to failure of the bridge are 100 million CHF, principally due to lost travel time and economic impact.

- Offer A, for 15 million CHF, is a wide bridge that can accommodate two lanes of traffic in both directions and has reliability over the time period of interest of 0.95.

- Offer B, for 19 million CHF, is for two narrow bridges that can accommodate two lanes of traffic each, and each have a reliability of 0.95 over the same time period.

2.6.5.1 Question E1

If it is assumed that the only differences between the two offers in terms of costs and benefits are the cost of construction and the cost of not being able to use the link if failure occurs,

which offer should be chosen?

Back up your argument with exact numbers. In words explain how your decision would be affected if you discovered that you over-estimated the reliability of the rest of the link.

2.6.5.2 Answer E1

Option B without the link has a higher reliability than the option A without the link. Therefore if the reliability of the link is lower than assumed, the failure cost of both options will be increasing accordingly. Therefore, option B will remain as the best option. The calculations can be seen in Table 2.7.

Table 2.7: Calculations

Offers	Cost	Object reliability	Link reliability w/o the object	Link reliability w object	Link failure probability	Costs due to not being able to use the link if failure occurs	Expected costs due to not being able to use the link	Differential cost	Differential benefit	Net benefit
A	15	0.95	0.9	0.855	0.145	100	14.500	0	0	
B	19	0.95 0.95	0.9			100	10.225	4	4.275	0.275

2.6.6 Problem F

Water is one of the fundamental needs of millions of people living in a megacity. Water of sufficient quality and quantity must be provided around the clock. In order to fulfill this need, a city depends on its water distribution infrastructure, which includes pipes made of different materials and laid at different times. These pipes are affected by processes of different types and deteriorate at different rates. The consequences related to pipe failure vary significantly depending on the type of failure, e.g. a pipe break, or a leaking pipe, as does the reaction time required to fix the pipe. For example, if a pipe breaks, a corrective intervention must be executed immediately, and if it is noticed that there is progressive water loss over time then a preventive intervention can be planned before there is an inadequate level of service.

Part of the water distribution network in mega-city Q is shown in Figure 1.4. The pipe characteristics are given in Table 2.8.

Table 2.8: Pipe and their attributes

Pipe	Material	Demand (m^3/day)	Length (m)	Year of construction	Usage	Distribution of PS in year 2014
M	Concrete		5’000	1998	Distribution only	[0.9; 0.1; 0; 0; 0]
A.1	PVC	250’000	600	2003	Distribution only	[0.85; 0.1; 0.05; 0; 0]
A.2	PVC		600	2003	Distribution only	
A.3	PVC		600	2003	Distribution and connection to buildings	
A.4	PVC		600	2003	Distribution only	
A.5	PVC		600	2003	Distribution and connection to buildings	
B	Cast iron type 1	200’000	2’500	1993	Distribution and connection to buildings	[0.8; 0.2; 0; 0; 0]
C	Cast iron type 2	100’000	1’600	1993	Distribution and connection to buildings	[0.83; 0.1; 0.07; 0; 0]
D	Cast iron type 3	350’000	4’000	1983	Distribution and connection to buildings	[0.8; 0.1; 0.1; 0; 0]

When newly built the pipes to not leak but over time they do. The amount of water lost due to these leaks increases over time. In order to model the increase in leakage over time the pipes the performance states in Table 2.9 are used. The one year transition probabilities for pipe M and A are given in Table 2.10.

Table 2.9: Performance states used to model water leakage over time

Performance state	Water loss (%)
1	W < 5%
2	5% <= W < 10%
3	10% <= W < 15%
4	15% <= W < 20%
5	20% <= W

Table 2.10: One year transition probabilities for pipe M and A

Pipe	M					A				
PS	1	2	3	4	5	1	2	3	4	5
1	0.93	0.07	0	0	0	0.91	0.09	0	0	0
2	0	0.87	0.13	0	0	0	0.88	0.12	0	0
3	0	0	0.85	0.15	0	0	0	0.85	0.15	0
4	0	0	0	0.8	0.2	0	0	0	0.82	0.18
5	0	0	0	0	1	0	0	0	0	1

In addition to the pipes leaking they sometimes break. The processes that lead to these breaks are not the same as the ones that cause leakage. The reliability of the pipes affected by these processes are modeled with the Weibull function with the values of the parameters given in Table 2.11.

Table 2.11: Deterioration parameters and year of construction and intervention information

Zone	Section	Scale parameter α	Shape parameter m	Duration of corrective intervention (days)	Cost of corrective intervention (mus)
A	A.1	0.005	1.3	5	8’000
	A.2	0.004	1.4	2	12’000
	A.3	0.003	1.5	8	10’000
	A.4	0.004	1.5	6	20’000
	A.5	0.004	1.3	4	15’000
B	B	0.003	1.3	9	45’000
C	C	0.001	1.7	10	60’000
D	D	0.002	1.5	9	50’000
M	M	0.004	1.2	11	70’000

The types of intervention considered possible for the pipes, as well as their unit costs, are given in Table 2.12. If a pipe breaks, it will be replaced for a cost of 500’000 CHF/100 m. The types of inspection that can be included in monitoring strategies are given in Table 2.13.

Table 2.12: Interventions

Performance state	Intervention type	Unit cost 10^3 mus/100m
1,2,3,4	Do nothing	0
3	Rehabilitation	25
4	Rehabilitation	30
3,4,5	Replacement	100
1,2,3,4,5	Replacement	500

Table 2.13: Inspection

Intervention type	Unit cost 10^3 mus/100m
Ultra-sonic	3
Acoustic emission	2
Wireless sensor	4
Camera	5

2.6.6.1 Question F1

Estimate the *failure rate* for performance states 1-4 of pipe M

2.6.6.2 Answer F1

The failure rate for each performance state is estimated from the reliability of that state. The reliability of each state equals to the survival probability of that state. The value of the reliability of each state is therefore equivalent to the transition probability that the pipe will remains in that state. Values of transition probability in Table 2.10 will be used for computing the failure rate of each performance state.

Using equation (2.52), following results are shown.

$$\theta_1 = -\frac{\ln[R_1(t=1)]}{t=1} = -\frac{\ln[0.93]}{t=1} \approx 0.073$$

$$\theta_2 = -\frac{\ln[R_2(t=1)]}{t=1} = -\frac{\ln[0.87]}{t=1} \approx 0.139$$

$$\theta_3 = -\frac{\ln[R_3(t=1)]}{t=1} = -\frac{\ln[0.85]}{t=1} \approx 0.162$$

$$\theta_4 = -\frac{\ln[R_4(t=1)]}{t=1} = -\frac{\ln[0.8]}{t=1} \approx 0.223$$

2.6.6.3 Question F2

Estimate the expected amount of time that pipe M remains in each performance state

2.6.6.4 Answer F2

The amount of time that pipe M remains in each performance state can be directly estimated from the reliability function, that is

$$\Theta_i = \int_0^\infty \tilde{F}_i(t)dt = \int_0^\infty \exp(-\theta_i \cdot t)dt = \frac{1}{\theta_i} \tag{2.58}$$

In the case of exponential distribution, the amount of time is the invert value of the failure rate

$$\Theta_1 = \frac{1}{\theta_1} = \frac{1}{0.073} = 13.69\text{years}$$
$$\Theta_2 = \frac{1}{\theta_2} = \frac{1}{0.139} = 7.19\text{years}$$
$$\Theta_3 = \frac{1}{\theta_3} = \frac{1}{0.162} = 6.17\text{years}$$
$$\Theta_4 = \frac{1}{\theta_4} = \frac{1}{0.223} = 4.48\text{years}$$

2.6.6.5 Question F3

What is the probability that there will be an adequate level of service in Zone A in year 2015 (PS4, PS5 = inadequate level of service)? The suspected amount of water loss in 2014 is given with the following probabilities (PS1 = 0.85, PS2 = 0.10, PS3 = 0.05, PS4 = 0.00; PS5 = 0.00).

2.6.6.6 Answer F3

Using the following equation (in Chapter describing Markov model)

$$\overrightarrow{x^{t+1}} = \begin{bmatrix} x_1 \; x_2 \; x_3 \; x_4 \; x_5 \end{bmatrix} \cdot \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} \\ 0 & p_{22} & p_{23} & p_{24} & p_{25} \\ 0 & 0 & p_{33} & p_{34} & p_{35} \\ 0 & 0 & 0 & p_{44} & p_{45} \\ 0 & 0 & 0 & 0 & p_{55} \end{bmatrix} \tag{2.59}$$

to estimate the state probability in year 2015

$$\begin{aligned} \overrightarrow{x^{t+1}} &= \begin{bmatrix} 0.85 \; 0.1 \; 0.05 \; 0 \; 0 \end{bmatrix} \cdot \begin{bmatrix} 0.91 & 0.09 & 0 & 0 & 0 \\ 0 & 0.88 & 0.12 & 0 & 0 \\ 0 & 0 & 0.85 & 0.15 & 0 \\ 0 & 0 & 0 & 0.82 & 0.18 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0.7735 \; 0.1645 \; 0.0545 \; 0.0075 \; 0 \end{bmatrix} \end{aligned} \tag{2.60}$$

The probability that pipe in zone A will be in an adequate level of service is the sum of the state probability from CS 1 to CS 3, that is

$$Pr\,o^{2015}\,[\text{Zone A in adequate level of service}] = \sum_{i=1}^{3} x_i^{2015} = 0.7735 + 0.1645 + 0.0545 = 0.9925$$

2.6.6.7 Question F4

Calculate the probability of there being at least one pipe break in zone A in year 2020 (5 points).

2.6.6.8 Answer F4

Step 1: Calculate the reliability for each pipe A1, A2, A3, A4, A5 in Zone A. The failure probability of each pipe is independent from each other.

Example is done for pipe A1

$R^{A1}(t = 22) = e^{-(0.005 \cdot 22)^{1.3}} = 0.945$

the failure probability is

$F^{A1}(t = 22) = 1 - R^{A1}(t = 22) = 1 - 0.945 = 0.055$

Step 2: Since the break of pipe is not dependent, it is not joint probability and thus, it is simple the average failure probability, which equals to 0.025.

The calculations are shown in Table 2.14.

Table 2.14: Calculations (2020)

Zone	Section	Scale parameter α	Shape parameter m	Elapsed time till 2013 (years)	Elapsed time till 2020 (years)	Reliability ($R(t)$)	Failure probability ($F(t)$)
A	A.1	0.005	1.3	15	22	0.945	0.055
	A.2	0.004	1.4	10	17	0.977	0.023
	A.3	0.003	1.5	10	17	0.988	0.012
	A.4	0.004	1.5	10	17	0.982	0.018
	A.5	0.004	1.3	10	17	0.982	0.018

2.6.6.9 Question F5

When is it expected that the annual probability of pipe A1 breaking will exceed 0.3 given that it did not break earlier?

2.6.6.10 Answer F5

Value of time t can be estimated from from equation (2.14)

$R(t) = e^{-(\alpha \cdot t)^m} \Leftrightarrow \ln[R(t)] = \ln\left[e^{-(\alpha \cdot t)^m}\right] = -(\alpha \cdot t)^m$
$\rightarrow t^m = -\frac{\ln[R(t)]}{\alpha^m} = -\frac{\ln[0.945]}{0.005^{1.3}} = 55.45 \approx 56\text{years}$

2.6.6.11 Question F6

Calculate the reliability of being able to provide an adequate level of service to zone A in year 2014 (Keep in mind that water can flow in both directions through the pipes).

2.6.6.12 Answer F6

Step 1: Calculate the reliability of each pipe in Zone A using given reliability function. This is exactly the same calculation as shown in subsection 6.8, except the elapsed time till 2014 for each pipe is different. Table 2.15 shows the results of estimation for reliability and failure in year 2014.

Table 2.15: Calculations (2014)

Zone	Section	Scale parameter α	Shape parameter m	Elapsed time till 2013 (years)	Elapsed time till 2020 (years)	Reliability ($R(t)$)	Failure probability ($F(t)$)
A	A.1	0.005	1.3	15	16	0.963	0.037
	A.2	0.004	1.4	10	11	0.987	0.013
	A.3	0.003	1.5	10	11	0.994	0.006
	A.4	0.004	1.5	10	11	0.991	0.009
	A.5	0.004	1.3	10	11	0.991	0.009

Step 2: if for example pipe break at A.2 or A.4, the system is still functioning because the water can flow into both direction. And even A.3 break, it can still provide adequate level of service. In that situation, the system turns to be like a series network with independent reliability.

3 scenarios can be seen

- Scenario 1: If A.1 breaks, it affects the entire zone A
- Scenario 2: If A.2 breaks, it does not, because water can flow into A4
- Scenario 3: If A.4 breaks, it does not, because water can flow in to A2

Following is the calculation for reliability of each scenario

$R^{scenario1}(2014) = R^{A1}(2014) = 0.963$

$R^{scenario2}(2014) = R^{A1}(2014) \cdot R^{A3}(2014) \cdot R^{A4}(2014) \cdot R^{A5}(2014) = 0.940$

$R^{scenario3}(2014) = R^{A1}(2014) \cdot R^{A2}(2014) \cdot R^{A3}(2014) \cdot R^{A5}(2014) = 0.937$

As each scenario can occurs independently from other scenarios, the average reliability is then evaluated as the average reliability of the three scenario

$R^{\text{Adequatelevelofservice}}(2014) = \frac{R^{\text{scenario1}}(2014)+R^{\text{scenario2}}(2014)+R^{\text{scenario3}}(2014)}{3}$
$= \frac{0.963+0.940+0.937}{3} = 0.947$

2.6.7 Problem G

The details of the number of fan units used in road tunnels retired at different ages are given in Table 2.16. Value in

Table 2.16: Record of fan units failed

Item	time	status	Item	time	status	Item	time	status
1	15	0	13	14	0	25	12	1
2	11	1	14	11	1	26	13	0
3	10	0	15	16	0	27	13	0
4	12	0	16	11	0	28	14	0
5	18	0	17	14	1	29	19	0
6	12	0	18	18	1	30	11	0
7	16	1	19	17	0	31	20	1
8	19	0	20	17	0	32	12	1
9	12	0	21	16	1	33	10	0
10	13	0	22	20	1	34	14	1
11	20	1	23	12	1	35	14	1
12	16	0	24	15	1	36	16	0

the column “status” has binary value 1 and 0. Value 1 and 0 represent the “fail” or “not fail” of the tunnel fans at the inspection time.

2.6.7.1 Question G1

What is the average expected life of the units that have survived until year 4?

2.6.7.2 Answer G1

The observed data infers a case that the failure rate is not a constant value. The failure rate of the fan is a time-varying variable. The failure rate can be followed a Weibull function (Equation (2.16)).

The average life of the units that have survived until a certain year will be defined as

$$\Phi\left[\tau \geq t\right] = \left[\int_{t}^{\infty} R(\tau)d\tau\right] = \left[\int_{t}^{\infty} e^{-(\alpha\cdot\tau)^{m}} d\tau\right] \tag{2.61}$$

In this example, t=4 years. The unknown parameters are value of α and m. Values of these parameters can be estimated with the Maximum likelihood estimation approach (MLE) (see the R code in the Appendix).

Using the MLE, value of α and m are 0.0542 and 5.8582, respectively. The average life of the units that have survived after 4 years will be

$$\Phi\left[\tau \geq 4\right] = \left[\int_{4}^{\infty} e^{-(0.0542\cdot\tau)^{5.8584}} d\tau\right] \approx 13\text{years}$$

References

Nam Lethanh, Bryan T. Adey, and Dilum N. Fernando. Optimal intervention strategies for multiple objects affected by manifest and latent deterioration processes. *Structure and Infrastructure Engineering*, 11(3):389–401, 2015. doi: 10.1080/15732479.2014.889178. URL http://dx.doi.org/10.1080/15732479.2014.889178.

M.T. Schultz, B.P. Gouldby, J.D. Simm, and J.L. Wibowo. Beyond the factor of safety: Developing fragility curves to characterize system reliability. Technical report, US Army Corps of Engineers, July 2010.

2.A Reliability of an item with time invariant failure rate

```
#This program is coded by Nam Lethanh for use in the class IMP
#Purpose: To calculate reliability
TIME=4
T=seq(0,TIME, by =0.5) #investigate time
N=3 #Number of item
lambda<-matrix(double(1),nrow=1,ncol=N)
lambda<-c(0.5,1,1.5) # value of hazard rate
# define dimension of matrix
fail<-matrix(double(1),nrow=length(T),ncol=N)
survive<-matrix(double(1),nrow=length(T),ncol=N)
#reliability function for exponential distribution
failure<-function(t,lambda){lambda*exp(-lambda*t)}
survival<-function(t,lambda){exp(-lambda*t)}
for (i in 1:N){
for (t in (1:length(T))){
  fail[t,i]<-failure(T[t],lambda[i])
  survive[t,i]<-survival(T[t],lambda[i])
}
}
cat('PROBABILITY DENSITY \n')
print(cbind(T,fail))
cat('RELIABILITY \n')
print(cbind(T,survive))
cat('EXPECTED LIFE TIME \n')
theta=lambda
print(1/theta)
#Plotting for reliability
plot.new()
par(mar=c(5,4,4,6)+0.3)
limy=c(0,1)
limx=c(0,4)
plot(T,survive[,1],lwd=2,col="red",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=4)
axis(2,ylim=limy,col="black",las=1)
axis(1,c(seq(0,max(limx),by=0.5)),c(seq(0,max(limx),by=0.5)))
mtext(expression(paste('Reliability')),side=2,col="black",line=3)
mtext(expression(paste('Units of time')),side=1,col="black",line=3)
grid(10, 10, col = "lightgray", lty = "dotted",lwd = par("lwd"), equilogs = TRUE)
box()
par(new=TRUE)
plot(T,survive[,2],lwd=2,col="blue",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=2)
par(new=TRUE)
plot(T,survive[,3],lwd=2,col="cyan4",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=1)
colors=c("red","blue","cyan4")
legend("topright", inset=0.09, title="Items",col=colors,lty=2,lwd=2,legend=c(1:N),pch=c(4,2,1),bg="azure2",cex
    =0.8)
#Plotting for probability density function
plot.new()
par(mar=c(5,4,4,6)+0.3)
limy=c(0,1.5)
limx=c(0,4)
plot(T,fail[,1],lwd=2,col="red",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=4)
axis(2,ylim=limy,col="black",las=1)
axis(1,c(seq(0,max(limx),by=0.5)),c(seq(0,max(limx),by=0.5)))
mtext(expression(paste('Probability density')),side=2,col="black",line=3)
mtext(expression(paste('Units of time')),side=1,col="black",line=3)
grid(10, 10, col = "lightgray", lty = "dotted",lwd = par("lwd"), equilogs = TRUE)
box()
par(new=TRUE)
plot(T,fail[,2],lwd=2,col="blue",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=2)
par(new=TRUE)
plot(T,fail[,3],lwd=2,col="cyan4",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=1)
colors=c("red","blue","cyan4")
legend("topright", inset=0.09, title="Items",col=colors,lty=2,lwd=2,legend=c(1:N),pch=c(4,2,1),bg="azure2",cex
    =0.8)
#Plotting for reliability
plot.new()
par(mar=c(5,4,4,6)+0.3)
limy=c(0,1)
limx=c(0,4)
plot(T,survive[,1],lwd=2,col="red",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=4)
axis(2,ylim=limy,col="black",las=1)
axis(1,c(seq(0,max(limx),by=0.5)),c(seq(0,max(limx),by=0.5)))
mtext(expression(paste('Reliability')),side=2,col="black",line=3)
mtext(expression(paste('Units of time')),side=1,col="black",line=3)
grid(10, 10, col = "lightgray", lty = "dotted",lwd = par("lwd"), equilogs = TRUE)
box()
hazardrate=theta[1]
points(1/hazardrate,survival(1/hazardrate,hazardrate),pch=23,bg="black",lwd=5)
segments(0,survival(1/hazardrate,hazardrate),1/hazardrate,survival(1/hazardrate,hazardrate), col= 'red',lty=1,
    lwd=1)
segments(1/hazardrate,survival(1/hazardrate,hazardrate),1/hazardrate,0, col= 'red',lty=1,lwd=1)
par(new=TRUE)
plot(T,survive[,2],lwd=2,col="blue",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=2)
```

```
hazardrate=theta[2]
points(1/hazardrate,survival(1/hazardrate,hazardrate),pch=23,bg="black",lwd=5)
segments(0,survival(1/hazardrate,hazardrate),1/hazardrate,survival(1/hazardrate,hazardrate), col= 'red',lty=1,
    lwd=1)
segments(1/hazardrate,survival(1/hazardrate,hazardrate),1/hazardrate,0, col= 'blue',lty=1,lwd=1)
par(new=TRUE)
plot(T,survive[,3],lwd=2,col="cyan4",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=1)
hazardrate=theta[3]
points(1/hazardrate,survival(1/hazardrate,hazardrate),pch=23,bg="black",lwd=5)
segments(0,survival(1/hazardrate,hazardrate),1/hazardrate,survival(1/hazardrate,hazardrate), col= 'red',lty=1,
    lwd=1)
segments(1/hazardrate,survival(1/hazardrate,hazardrate),1/hazardrate,0, col= 'cyan4',lty=1,lwd=1)
colors=c("red","blue","cyan4")
legend("topright", inset=0.09, title="Items",col=colors,lty=2,lwd=2,legend=c(1:N),pch=c(4,2,1),bg="azure2",cex
    =0.8)
cat("THE END")
```

2.B Reliability of an item with time-variant failure rate

```
#This program is coded by Nam Lethanh for use in the class IMP-HS2014
#Purpose: To calculate reliability from Cox Model
data=read.csv("example2.csv",header=TRUE)
attach(data)
library(survival)
likelihood=survreg(Surv(time,status)~1,data=data,dist="weibull")
print(likelihood)

density<-function(t,theta,m){theta*m*(theta*t)^(m-1)*exp(-(theta*t)^m)}
rate<-function(t,theta,m){theta*m*(theta*t)^(m-1)}
reliability<-function(t,theta,m){exp(-(theta*t)^m)}
TIME=100
T=seq(0,TIME, by =1)
den<-matrix(double(1),nrow=T,ncol=1)
hazard<-matrix(double(1),nrow=T,ncol=1)
reli<-matrix(double(1),nrow=T,ncol=1)
#theta=0.912
#m=3.624

theta=1/exp(3.624)
m=1/0.912

for (t in (1:length(T))){
  den[t]<-density(T[t],theta,m)
  hazard[t]<-rate(T[t],theta,m)
  reli[t]<-reliability(T[t],theta,m)
}
results=cbind(den,hazard,reli)
print(results)

plot.new()
par(mar=c(5,4,4,6)+0.3)
limy=c(0,1)
limx=c(0,100)
plot(T,reli,lwd=2,col="red",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=4)
axis(2,ylim=limy,col="black",las=1)
axis(1,c(seq(0,100,by=10)),c(seq(0,100,by=10)))
mtext(expression(paste('Reliability')),side=2,col="black",line=3)
mtext(expression(paste('Units of time')),side=1,col="black",line=3)
grid(10, 10, col = "lightgray", lty = "dotted",lwd = par("lwd"), equilogs = TRUE)
box()
par(new=TRUE)
limy=c(0,0.05)
limx=c(0,100)
plot(T,hazard,lwd=2,col="blue",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=2)
axis(4,ylim=limy,col="black",las=1)
mtext(expression(paste('Failure rate')),side=4,col="black",line=3)
colors=c("red","blue")
legend("topright", inset=0.09, col=colors,lty=2,lwd=2,legend=c("Reliability","Failure rate"),pch=c(4,2),bg="
     azure2",cex=1.3)

cat("THE END")
```

2.C Maximum likelihood estimation for Weibull model

```
#This program is coded by Nam Lethanh for use in the class IMP
#Purpose: To calculate reliability from Cox Model
data=read.csv("IMP-A-fan.csv",header=TRUE)
attach(data)
library(survival)
likelihood=survreg(Surv(time,status)~1,data=data,dist="weibull")
print(likelihood)

density<-function(t,theta,m){theta*m*(theta*t)^(m-1)*exp(-(theta*t)^m)}
rate<-function(t,theta,m){theta*m*(theta*t)^(m-1)}
reliability<-function(t,theta,m){exp(-(theta*t)^m)}
TIME=10
T=seq(0,TIME, by =0.1)
den<-matrix(double(1),nrow=T,ncol=1)
hazard<-matrix(double(1),nrow=T,ncol=1)
reli<-matrix(double(1),nrow=T,ncol=1)
theta=1/exp(2.915419)
m=1/0.170694
for (t in (1:length(T))){
  den[t]<-density(T[t],theta,m)
  hazard[t]<-rate(T[t],theta,m)
  reli[t]<-reliability(T[t],theta,m)
}
results=cbind(den,hazard,reli)
print(results)

theta=1/exp(2.915419)
m=1/0.170694

a=integrate(reliability, lower = 4, upper = Inf,theta=theta,m=m)$value
print(a)
```

Chapter 3

Availability and maintainability

BRYAN T. ADEY AND NAM LETHANH

3.1 Maintainability of an item

When managing infrastructure the maintainability of an item is often a useful bit of information, i.e. how much effort is it to restore the item so that it once again provides an adequate LOS. For example, if the speeds of trains need to be reduced on a rail link due to deformations in a track, how much will it cost to fix the rail link in a way that the trains can once again run at normal speed, and how long will this take?

The maintainability of an item only has to do with the execution of interventions on the item. The maintainability and reliability together give the availability of an item.

3.1.1 Time

In order to estimate how long it will take to restore an item so that it provides an adequate LOS, it is necessary to estimate the steps to be taken from the instant that an inadequate LOS is not provided until it is restored. These steps include the determination that a failure has occurred, the determination of what exactly went wrong, the determination of how to restore the item, the execution of the intervention, the testing that the item again can provide an adequate LOS and then putting it back in operation so that it does provide an adequate LOS. An illustration of this for the restoration of a part of an electricity distribution network using Business Process Modelling Notation is shown in Figure 3.1.

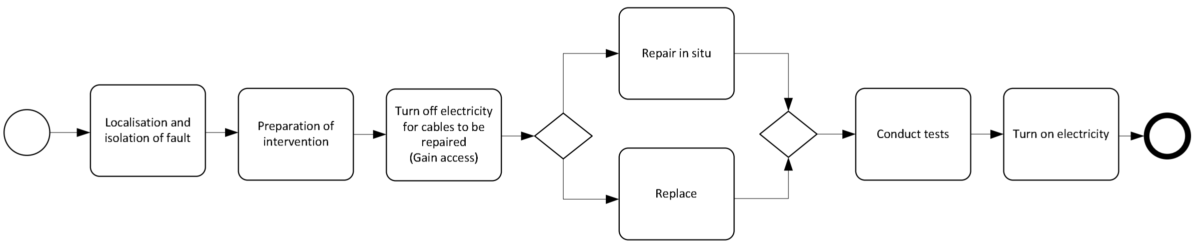


Fig. 3.1: Simplified process for the restoration of a part of an electricity distribution network

Once the process has been identified it is necessary to estimate the length of time that each of the activities will take. This depends on many factors, including the quantity and skill levels of persons required and available, the quantity of required replacement parts required and available, and the amount of time required to react to an indication that something is not working, sometimes referred to as responsiveness.

3.1.2 Indicators

Indicators that are often used to express maintainability are the mode[1], mean[2] and maximum[3] amount of duration or cost of interventions. Indicators are usually developed for specific types of interventions, e.g. the mean duration of a corrective intervention is 3 months, or the mean duration of a preventive intervention is 1 week.

3.2 Maintainability of an item comprised of sub-items

The maintainability of an item can be estimated by taking into consideration the structure of the sub-items in the item, and the maintainability of the sub-items. This is advantageous if there is little data on the interventions that have been executed on the item itself, e.g. a bridge, but abundant data on interventions that have been executed on sub-items similar to those of which the item is composed, e.g. a concrete abutment.

If it can be assumed that each intervention is only used to restore one sub-item so that it provides an adequate LOS, the durations of the interventions can be simply combined, e.g. if it takes 1 week to repair an abutment and 2 weeks to repair a deck and both happen once every 10 years, the average time spent to repair a bridge per year, i.e. its' maintainability, is 0.30 weeks ((1+2)/10). If, however, 50% of the time the abutment and deck are repaired at the same time which takes 2 weeks, 50% of the time they are repaired separately than the average time spent to repair a bridge per year, i.e. its maintainability, is 0.175 weeks ((0.5(2)+0.25(1+2))/10).

The basic steps to combine sub-items, if it can be assumed that each intervention is only used to restore one sub-item so that it provides an adequate LOS, are:

- divide all sub-items by type, so it can be assumed that each sub-item per sub-item type behaves in the same way,
- determine the rate of occurrence of the type of intervention of interest, e.g. corrective intervention, per sub-item type,
- estimate the rate of occurrence of the type of intervention, per sub-item type by multiplying the rate of occurrence of the type of intervention by the number of sub-items per sub-type,
- estimate the total duration of the type of intervention per sub-item type by multiplying the rate of occurrence of the type of intervention by the mean duration of the type of intervention per sub-item type
- estimate the rate of occurrence of an intervention on the item by summing the rates of occurrence of the types of interventions per sub-item
- estimate the expected intervention duration, or maintainability, of the item by summing the total duration of the type of intervention per sub-item.

[1] The duration or costs that occurs the most often
[2] The average of the sum of all durations or costs divided by the number of values
[3] The largest of all durations or costs observed or expected. This is also often given as a percentile of the largest of all durations or costs.

3.3 Availability of an item

The availability of an item is based on both the reliability of an item and the maintainability of an item. Simply, availability is:

$$A_i = \frac{E\left[t_i^{adequate}\right]}{E\left[t_i^{adequate}\right] + E\left[t_i^{inadequate}\right]} \tag{3.1}$$

Where:

$t_i^{adequate}$: the duration of time item i provides an adequate LOS,

$t_i^{inadequate}$: the duration of time item i provides an inadequate LOS

When investigating the availability of an item the exact definition of availability used is important. One should be explicit about

- the types of interventions being considered, e.g. preventive and corrective interventions or only corrective interventions
- the type of time being considered, e.g. is the time spent waiting for parts and the time on-site to be considered, or only the time on-site.

Different measure of availability serve different purposes.

3.3.1 Example

A road has been in service for 10 years (3’652 days). Over those 10 years, 6 interventions have been executed. Some of those have been due to rock falls, some due to scour and some due to chloride induced corrosion. The total amount of time that an adequate LOS was not provided was 156 days, i.e. traffic flow was restricted for 156 days in the last 3’652 days.

3.3.1.1 Question A

What was the availability of the road link?

3.3.1.2 Answer A

The total length of time when an adequate LOS was provided was 3’496 days (3’652 – 156). The rate of occurrence of interventions was 0.001716 interventions/day (6/3’496). The mean time between interventions is therefore 582.7 days (1/0.001716). The availability is then

$$A_i = \frac{E\left[t_i^a\right]}{E\left[t_i^a\right]+E\left[t_i^n\right]} = \frac{582.7}{582.7+156/6} = 0.9573 \tag{3.2}$$

3.3.1.3 Question B

Is this also the availability of the road next year?

3.3.1.4 Answer B

It depends on the types of failures and how they were repaired. In order to state that the availability of the road next year is also 0.9573, it would be necessary to be able to model the time between interventions as an exponential distribution, meaning that rate of occurrence of failure can be modelled as constant. This is essentially saying that no matter how many failure have occurred in the past there will be on average the same number in the future. This is in many cases not correct. For example, if there is an area in which a landslide damages a road, it may be that there will never be another landslide in this area. Or, if many interventions has to be executed due the corrosion of the reinforcement in deck slabs and waterproofing layers on all than the number of interventions per year in the future would change.

3.4 Availability of an item comprised of sub-items

In determining the availability of an item comprised of sub-items it is important to explicitly take into consideration the structure of the item. The availability of item i comprised of sub-items connected in series is given by:

$$A_i = \prod_j^J \left(\frac{E\left[t_j^a\right]}{E\left[t_j^a\right] + E\left[t_j^n\right]} \right) \tag{3.3}$$

3.4.1 Example

You own a water distribution network that is connected as shown in Figure 3.2. The five pipes in the network have life expectancies of 20 years. The life expectancies can be modelled using an exponential distribution. It takes 1 week to repair pipes 1 and 4, 2 weeks for pipes 2 and 5 and 3 weeks for pipe 3, once it breaks.

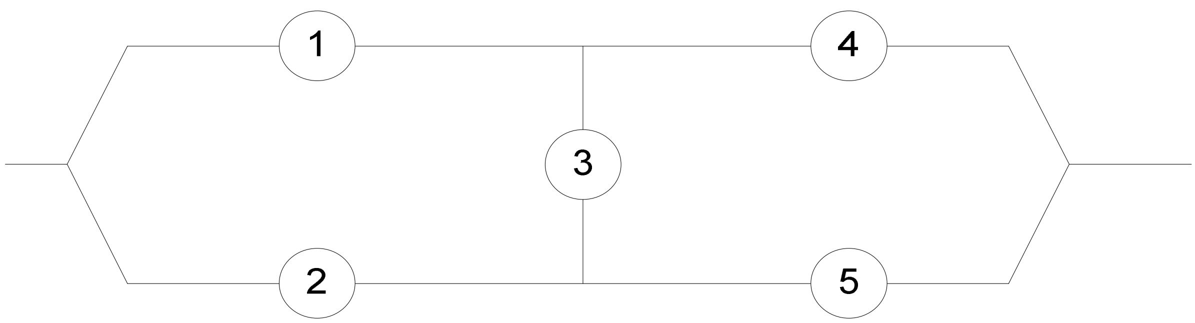


Fig. 3.2: Water distribution network

3.4.1.1 Question C

What is the reliability of the network?

3.4.1.2 Question D

What is the availability of the network?

3.4.1.3 Answer C

Using the pivotal decomposition method, the structure function of the network is

$$\begin{aligned}\phi(\mathbf{x}) &= x_1x_3x_4 + x_1x_3x_5 + x_2x_3x_4 + x_2x_3x_5 + x_1x_2x_4x_5 \\ &- x_1x_2x_3x_4 - x_1x_3x_4x_5 - x_1x_2x_3x_5 - x_2x_3x_4x_5\end{aligned} \tag{3.4}$$

The steps to arrive at the above function has been described in section 3.3 of the script of week 9.

In term of reliability, the structure function can be rewritten as

$$\begin{aligned}R_{network}(t) &= R_1(t)\cdot R_3(t)\cdot R_4(t) + R_1(t)\cdot R_3(t)\cdot R_5(t) + R_2(t)\cdot R_3(t)\cdot R_4(t) + R_2(t)\cdot R_3(t)\cdot R_5(t) \\ &+ R_1(t)\cdot R_2(t)\cdot R_4(t)\cdot R_5(t) - R_1(t)\cdot R_2(t)\cdot R_3(t)\cdot R_4(t) \\ &- R_1(t)\cdot R_3(t)\cdot R_4(t)\cdot R_5(t) - R_1(t)\cdot R_2(t)\cdot R_3(t)\cdot R_5(t) - R_2(t)\cdot R_3(t)\cdot R_4(t)\cdot R_5(t)\end{aligned} \tag{3.5}$$

The reliability of any item i at time t is defined in (2.11)

As the expected time until an inadequate LOS is provided is about 20 years, the "failure rate" can be estimated as follows:

$$\int_0^{+\infty} R_i(t)dt = \int_0^{+\infty} \exp(-\theta_i \cdot t)\, dt = \frac{1}{\theta_i} = 20\text{years}$$

$$\theta_i = \frac{1}{20} = 0.05$$

Using equations (2.11) the reliability of each sub-item, and the item, in 30 years can be estimated (Table 3.1). For example, the probability that sub-item 1 will provide an adequate LOS until the start of year 14 is 0.522 and the probability that the item will provide an adequate LOS until year 14 is 0.346. This is illustrated in Figure 3.3. Only one curve is used to represent the reliability of a sub-item, since the reliability for all sub-items is the same.

3.4.1.4 Answer D

The function form for calculating the availability is:

Table 3.1: Reliability

Time (years)	Reliability					
	1	2	3	4	5	Item
1	1	1	1	1	1	1
2	0.951	0.951	0.951	0.951	0.951	0.987
3	0.905	0.905	0.905	0.905	0.905	0.952
4	0.861	0.861	0.861	0.861	0.861	0.904
5	0.819	0.819	0.819	0.819	0.819	0.847
6	0.779	0.779	0.779	0.779	0.779	0.786
7	0.741	0.741	0.741	0.741	0.741	0.723
8	0.705	0.705	0.705	0.705	0.705	0.66
9	0.670	0.670	0.670	0.670	0.670	0.599
10	0.638	0.638	0.638	0.638	0.638	0.541
11	0.607	0.607	0.607	0.607	0.607	0.487
12	0.577	0.577	0.577	0.577	0.577	0.436
13	0.549	0.549	0.549	0.549	0.549	0.389
14	0.522	0.522	0.522	0.522	0.522	0.346
15	0.497	0.497	0.497	0.497	0.497	0.307
16	0.472	0.472	0.472	0.472	0.472	0.272
17	0.449	0.449	0.449	0.449	0.449	0.241
18	0.427	0.427	0.427	0.427	0.427	0.212
19	0.407	0.407	0.407	0.407	0.407	0.187
20	0.387	0.387	0.387	0.387	0.387	0.164
21	0.368	0.368	0.368	0.368	0.368	0.144
22	0.350	0.350	0.350	0.350	0.350	0.126
23	0.333	0.333	0.333	0.333	0.333	0.111
24	0.317	0.317	0.317	0.317	0.317	0.097
25	0.301	0.301	0.301	0.301	0.301	0.085
26	0.287	0.287	0.287	0.287	0.287	0.074
27	0.273	0.273	0.273	0.273	0.273	0.064
28	0.259	0.259	0.259	0.259	0.259	0.056
29	0.247	0.247	0.247	0.247	0.247	0.049
30	0.235	0.235	0.235	0.235	0.235	0.043

$$\begin{aligned}A_{network}(t) &= A_1(t)\cdot A_3(t)\cdot A_4(t)+A_1(t)\cdot A_3(t)\cdot A_5(t)+A_2(t)\cdot A_3(t)\cdot A_4(t)+A_2(t)\cdot A_3(t)\cdot A_5(t)\\ &+A_1(t)\cdot A_2(t)\cdot A_4(t)\cdot A_5(t)-A_1(t)\cdot A_2(t)\cdot A_3(t)\cdot A_4(t)\\ &-A_1(t)\cdot A_3(t)\cdot A_4(t)\cdot A_5(t)-A_1(t)\cdot A_2(t)\cdot A_3(t)\cdot A_5(t)-A_2(t)\cdot A_3(t)\cdot A_4(t)\cdot A_5(t)\end{aligned} \tag{3.6}$$

where $A_i(t)$is the availability of sub-item i.

The availability of each sub-item is calculated with equation (3.1), in which, it is necessary to calculate the mean time to repair or the mean time (duration) the sub-item is in adequate level of service. This calculation involves the value of reliability for each sub-item calculated in previous step.

$$E(t_i^n) = \Delta_i \cdot F_i(t) = \Delta_i \cdot [1 - R_i(t)] \tag{3.7}$$

In equation (3.7), $F_i(t)$is the failure probability at time t.

The availability of each sub-item in year t will be

$$A(t_i) = \frac{E\left[t_i^a\right]}{E\left[t_i^a\right]+E\left[t_i^n\right]} = \frac{\int\limits_0^{tu} R_i(t)dt}{\int\limits_0^{tu} R_i(t)dt + \Delta_i \cdot [1 - R_i(tu)]} \tag{3.8}$$

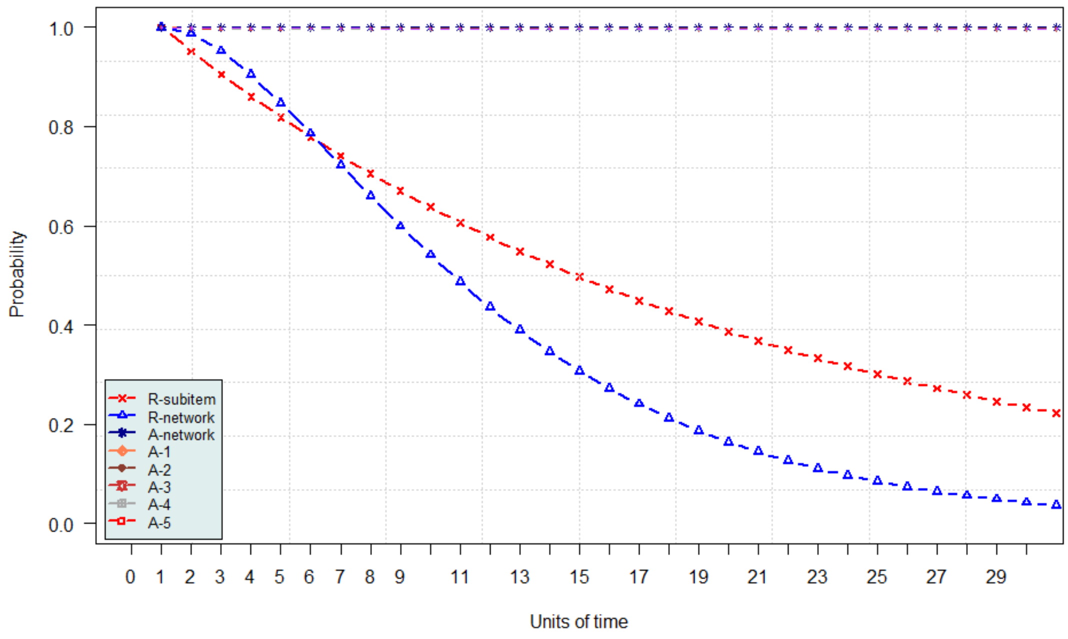


Fig. 3.3: Reliability and availability

The availability of each item and the network in 30 years.

Using equations (2.11) the availability of each sub-item, and the item, for periods of time up to 30 years can be calculated (Table 3.2). For example, the availability of sub-item 1 for a 14 year time period is 0.999090421, and the availability of the item for a 14 year time period is 0.999947505. This means that on average over periods of time that are 14 years in length that one can expect that sub-item 1 would provide an adequate LOS for 13.9872 years and an inadequate LOS for 0.01273411 years (4.647949 days), and that the item would provide an adequate LOS for 13.99927 years and an inadequate LOS for 0.00073 years (0.26645 days). The longer the period of time investigated the lower the availability because the reliability of the sub-items decreases with age, when there is a constant failure rate. This is illustrated in Figure 3.3.

The code to compute this example is given in the Appendix 3.A

Table 3.2: Availability

Time	Availability					
(years)	1	2	3	4	5	Item
1	1	1	1	1	1	1
2	0.999509	0.999018	0.997549	0.998038	0.998528	0.999985
3	0.999345	0.998691	0.996735	0.997386	0.998039	0.999973
4	0.999264	0.998528	0.996329	0.997061	0.997794	0.999966
5	0.999215	0.998431	0.996086	0.996867	0.997648	0.999961
6	0.999182	0.998366	0.995925	0.996737	0.997551	0.999958
7	0.999159	0.99832	0.995810	0.996645	0.997482	0.999955
8	0.999142	0.998285	0.995724	0.996576	0.99743	0.999953
9	0.999128	0.998258	0.995657	0.996523	0.99739	0.999952
10	0.999118	0.998237	0.995604	0.99648	0.997358	0.999951
11	0.999109	0.998219	0.995561	0.996445	0.997332	0.99995
12	0.999102	0.998205	0.995525	0.996416	0.99731	0.999949
13	0.999096	0.998193	0.995494	0.996392	0.997292	0.999948
14	0.99909	0.998182	0.995469	0.996372	0.997276	0.999948
15	0.999086	0.998174	0.995446	0.996354	0.997263	0.999947
16	0.999082	0.998166	0.995427	0.996338	0.997251	0.999947
17	0.999079	0.998159	0.995410	0.996325	0.997241	0.999946
18	0.999076	0.998153	0.995395	0.996313	0.997232	0.999946
19	0.999073	0.998148	0.995382	0.996302	0.997224	0.999945
20	0.999071	0.998143	0.99537	0.996293	0.997217	0.999945
21	0.999068	0.998139	0.995359	0.996284	0.99721	0.999945
22	0.999066	0.998135	0.99535	0.996276	0.997205	0.999945
23	0.999065	0.998131	0.995341	0.996269	0.997199	0.999945
24	0.999063	0.998128	0.995333	0.996263	0.997195	0.999944
25	0.999062	0.998125	0.995326	0.996257	0.99719	0.999944
26	0.99906	0.998123	0.99532	0.996252	0.997186	0.999944
27	0.999059	0.99812	0.995314	0.996247	0.997183	0.999944
28	0.999058	0.998118	0.995308	0.996243	0.99718	0.999944
29	0.999057	0.998116	0.995303	0.996239	0.997176	0.999944
30	0.999056	0.998114	0.995298	0.996235	0.997174	0.999943

3.5 Trade-offs

Indicators of performance are just that, indicators. One can minimize the amount of time required to execute interventions on an item, i.e. maximize maintainability, by constructing an expensive but easy to repair item. This would most likely increase the availability of the item but may have little effect on the reliability of the item.

One can maximize the reliability of an item by constructing an expensive but difficult to repair item. This would most likely decrease the maintainability of the item but may have little effect on the availability of the item.

Once can maximize the availability of an item by constructing an expensive item that is fairly reliable and fairly easy to repair.

In each of these cases, however, the most important question for the infrastructure manager is whether the improvements in reliability, availability or maintainability are worth the money spent. For example, if a road has very little traffic it may not be worthwhile to spend a large amount of money to shorten the durations of interventions, to increase the reliability of the bridge in the road, or to improve the availability of the road.

3.5.1 Example

You are the manager of a tunnel that is considering the construction of tunnel with five possible configurations. Each configuration is comprised of five components, but three of the components can be one of two different types. The configurations to be considered, as well as the mean time between failures for each component in each configuration, the initial cost of each configuration, and the mean intervention duration are given in Table 3.3.

Table 3.3: Tunnel components

Components		Option 1	Option 2	Option 3	Option 4	Option 5
No.	Description	mean time between failures (years)				
1	a power supply system	245	9’612	245	9’612	9’612
2	a ventilation system	4’819	4’819	36’955	4’819	36’955
3	a fire services system	1.1x106	1.1x106	1.1x106	1.1x106	1.1x106
4	a drainage system	1	1	1	2’194	2’194
5	a central monitoring and control system	2.4x105	2.4x105	2.4x105	2.4x105	2.4x105
	Initial cost (mus)	0	2	4	6	12
	Mean intervention duration (tu)	0.25	0.28	0.3	1.5	2

3.5.1.1 Question

Which configuration has the highest reliability? The highest maintainability? The highest availability? And which configuration results in the lowest overall net costs over a 20 *tu* time period?

3.5.1.2 Answer

3.5.1.2.1 Reliability

In the estimation of the reliability of the tunnel item one first has to recognize that the 5 sub-items are conceptually connected in series, i.e. if one sub-item fails the item fails. The next step is to calculate the failure rate of each component, which if it is assumed that the failures of each component follows can be modelled using an exponential distribution, is 1/the mean time between failures. These are shown in Table 3.4. θ_1 θ_2 θ_3 θ_4 θ_5 The mean time between failures for the tunnel is then 1 over the sum of the failure rates as shown in equation (3.9):

$$\frac{1}{\sum_{i}^{I} \theta_i} \tag{3.9}$$

Where i is the index of tunnel components.

And the reliability of the tunnel in one year is given by:

$$R = \exp\left[-\left(\sum_{i}^{I} \theta_i\right) \cdot 1\right] \tag{3.10}$$

These values for each of the options are shown in Table 3.4. It can be seen that option 5 is the most reliable, although not significantly more reliable than option 4.

Table 3.4: Failure rates of tunnel components, and mean time between failure and reliability of the tunnel

Option	Item	Components					MTBF of Tunnel	Reliability
		Power Supply System	Ventilation System	Fire Services System	Drainage System	Central monitoring and control system		
1	MTBF (years)	245	4’819	1.10E+06	1	2.40E+05	1	0.3663
	θ_1	4.08E-03	2.08E-04	9.09E-07	1.00E+00	4.17E-06		
2	MTBF (years)	9’612	4’819	1.10E+06	1	2.40E+05	1	0.3678
	θ_2	1.04E-04	2.08E-04	9.09E-07	1.00E+00	4.17E-06		
3	MTBF (years)	245	36’955	1.10E+06	1	2.40E+05	1	0.3664
	θ_3	4.08E-03	2.71E-05	9.09E-07	1.00E+00	4.17E-06		
4	MTBF (years)	9’612	4’819	1.10E+06	2’194	2.40E+05	1’295	0.9992
	θ_4	1.04E-04	2.08E-04	9.09E-07	4.56E-04	4.17E-06		
5	MTBF (years)	9’612	36’955	1.10E+06	2’194	2.40E+05	1’689	0.9994
	θ_5	1.04E-004	2.71E-005	9.09E-007	4.56E-004	4.17E-006		

3.5.1.2.2 Maintainability

In order to calculate the maintainability we can look simply at the average length of time it will take to restore each configuration when it fails. It is here clear that the configuration 1, which has an average down time of 0.25 months if it fails is the most maintainable. This can be seen in Table 3.5.

Table 3.5: Initial cost vs reliability, maintainability, availability and net benefit of the five configurations

Configuration	Initial Cost (mus)	Reliability	Maintainability (months)	Availability
1	0	0.3663	0.25	0.979506
2	2	0.3678	0.28	0.97759
3	4	0.3664	0.3	0.975512
4	6	0.9992	1.5	0.999903
5	12	0.9994	2	0.999901

3.5.1.2.3 Availability

The availability of each configuration is given by taking the mean time between failures, i.e. the mean time the item works, and dividing it by the average time it takes to restore the configuration if it fails plus the mean time it works, i.e. the average amount of time from starting to providing an adequate LOS to providing an adequate LOS again. One can also refer to this as the renewal period. The configuration with the highest availability is configuration 4 (Table 3.5).

3.5.1.2.4 Costs

The net costs are the initial costs plus the costs incurred over the 20 year period due to the costs due to lost service when an inadequate LOS is provided, which in this case are considered to include the cost of interventions when executed. Assuming that these costs are directly and uniformly related to the amount of time required to restore service, the annual maintenance costs for each configuration are given by:

$$C_{am} = \frac{E\left[t_i^n\right] \cdot I_{cr}/12}{E\left[t_i^a\right] + E\left[t_i^n\right]} \tag{3.11}$$

Where I_{cr} represents the cost of executing an intervention multiplied by the probability of executing an intervention

The annual maintenance costs as well as the net benefit excluding upfront costs and including upfront costs are shown for each configuration in Table 3.6. The net benefit of a configuration is the cost of configuration 1 minus the costs of the configuration being investigated. It can be seen that configuration has the highest net benefit over the 20 *tu* period.

$$NB^i = NB^i_c + NB^i_m \quad = \left(C^r_c - C^i_c\right) + \left(C^r_m - C^i_m\right) \tag{3.12}$$

NB^i: is the net benefit of configuration i,

C^r_m: is the maintenance costs of the reference configuration

C^i_m: is the maintenance costs of configuration i

Table 3.6: Initial cost versus total costs of the five configurations

Configuration	Cost					Net benefit excluding upfront costs	Net benefit including upfront costs
	initial (mus) C_c	Cost mus/month I_{cr}	Cost (mus) $E[t^n_i] \cdot I_{cr}$	Cost per tu C_{am}	Cost for 20 tus $C_{am} \cdot 20$	NB^i_m	NB^i
1	0	5	1.25	$7.79x10^{-1}$	15.5844	0	0
2	2	5	1.38	$8.50x10^{-1}$	17.0023	-1.418	-3.418
3	4	5	1.5	$9.31x10^{-1}$	18.6197	-3.035	-7.035
4	6	5	7.5	$4.47x10^{-6}$	0.0001	15.584	9.584
5	12	5	10	$3.50x10^{-6}$	0.0001	15.584	3.584

3.5.1.2.5 Summary

As can be seen in Table 3.7 that estimating the availability was the only performance indicator that allowed the optimal configuration to be determined. This will, however, not always be the case. This should be kept in mind when selecting performance indicators to evaluate performance. The R code for this example is given in Appendix 3.B

Table 3.7: Initial cost vs reliability, maintainability, availability and net benefit of the five configurations

Configuration	Initial Cost (mus)	Reliability	Maintainability (months)	Availability	Net benefit including upfront costs
1	0	0.3663	0.25	0.979506	0.000
2	2	0.3678	0.28	0.97759	-3.418
3	4	0.3664	0.3	0.975512	-7.035
4	6	0.9992	1.5	0.999903	9.584
5	12	0.9994	2	0.999901	3.584

3.6 Assignments

3.6.1 Problem A

An infrastructure manager is responsible for maintaining a number of facilities over a period of time. The length of time required for past corrective interventions are given in Table 3.8

Table 3.8: Corrective maintenance task times

Task time	Frequency
11	2
13	3
15	8
17	12
19	12
21	14
23	13
25	10
27	10
29	8
31	7
33	6
35	5
36	5
37	4
39	3
41	2
47	2

3.6.1.1 Question A1

What is the range of observations?

3.6.1.2 Answer A1

The range of observations is the difference between the smallest and the largest value in the data set. In this respect, the range of the above data set is

$$Range = \max(data) - \min(data) \tag{3.13}$$

$Range = 14 - 2 = 12$

3.6.1.3 Question A2s

Using a class interval of 4, determine the number of class intervals. Plot the data and construct a curve. What is the most likely type of distribution is indicated by the curve?

3.6.1.4 Answer A2

With a class interval of 4, one can construct a range of observations. This is equivalent to the class interval of 4.

From this table, a histogram of frequency can be constructed (Figure 3.4)

By observing the shape of the histogram, one can conclude that the data is distributed close to the density of the lognormal distribution function, which has a skew tail on the right side of the density function.

Table 3.9: Range of observation

Range	Frequency
9-12	2
13-16	11
17-20	24
21-24	27
25-28	20
29-32	15
33-36	16
37-40	7
41-44	2
45-48	2

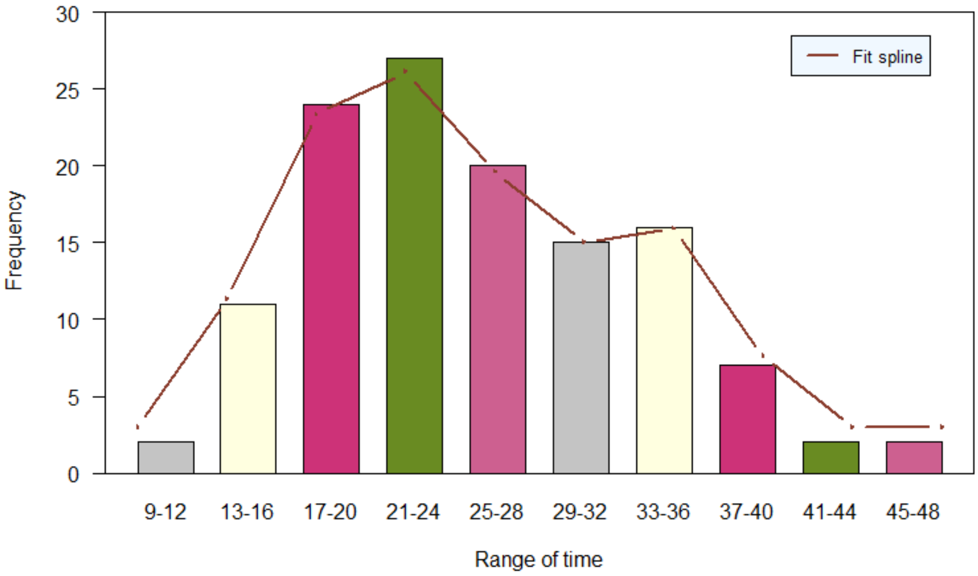


Fig. 3.4: Histogram of data

The lognormal distribution function has following properties

Probability density function

$$f_X(x,\mu,\sigma) = \frac{1}{x\cdot\sigma\sqrt{2\pi}}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, x > 0 \tag{3.14}$$

Cumulative distribution function

$$F_X(x,\mu,\sigma) = \frac{1}{2}+\frac{1}{2}erf\left[\frac{(\ln x-\mu)}{\sqrt{2\sigma}}\right] \tag{3.15}$$

where:

x: is the random variable of the distribution,

μ: is the mean of the variable's natural logarithm,

σ: is the standard deviation of the variable's natural logarithm,

erf: is the error function.

3.6.1.5 Question A3

What is the mean length of time for a corrective intervention, M_{ct}?

3.6.1.6 Answer A3

For estimating the mean corrective intervention duration

$$\overline{M}_{ct} = \frac{\sum (f_{ct_i})(M_{ct_i})}{\sum f_{ct_i}} \tag{3.16}$$

where:

f_{cti}: is the frequency of the i^{th} corrective intervention in interventions per system operating time unit,

M_{cti}: is the elapsed time required for the i^{th} corrective intervention.

The answer is 25.15 days

Table 3.10: Mean time of intervention duration

Time	Frequency	Mean time	Variance
(1)	(2)	(3)=(1)*(2)	$(4) = (2) * [(1) - \bar{M}_{ct}]$
11	2	22	400.49
13	3	39	442.93
15	8	120	824.31
17	12	204	797.23
19	12	228	453.99
21	14	294	241.21
23	13	299	60.14
25	10	250	0.23
27	10	270	34.2
29	8	232	118.53
31	7	217	239.49
33	6	198	369.66
35	5	175	485.03
36	5	180	588.53
37	4	148	561.61
39	3	117	575.4
41	2	82	502.39
47	2	94	954.78
Total	**126**	**3’169**	**7’650.13**

$$\overline{M}_{ct} = \frac{3'169}{126} = 25.15$$

3.6.1.7 Question A4

What is the geometric mean of the repair times?

3.6.1.8 Answer A4

For calculating geometric mean, refer to general statistical materials

$$\bar{\mu}_{geo} = e^{\mu} \tag{3.17}$$

Where μ is the mean that is calculated based on the mean $\overline{M}_{ct}$ in this case)

$$\mu = \ln(E[X]) - \frac{1}{2}\sigma^2 \tag{3.18}$$

$$\sigma^2 = \ln\left(1 + \frac{Var[X]}{(E[X])^2}\right) \tag{3.19}$$

from data, we have

$Var[X] = \sqrt{\frac{7'650.13}{126}} = 7.79$

$E[X] = \bar{M}_{ct} = 25.15$

$\mu = \ln(25.15) - \frac{1}{2}\left[\ln\left(1 + \frac{7.79^2}{25.15^2}\right)\right] = 3.179$

The geometric mean is then

$\bar{\mu}_{geo} = e^{3.179} = 24.02$

3.6.1.9 Question A5

What is the standard deviation of the sample data?

3.6.1.10 Answer A5

The standard deviation is 7.79, which has been calculated in previous steps.

3.6.1.11 Question A6

Assume 90% confidence level, what is the M_{max} value?

3.6.1.12 Answer A6

For obtaining the the M_{max} value with assumption of a 90% confidence level, refer to the formulation of the cumulative density function in equation (3.15) of the log-normal distribution and interpolate the data.

The Mmax is 27.

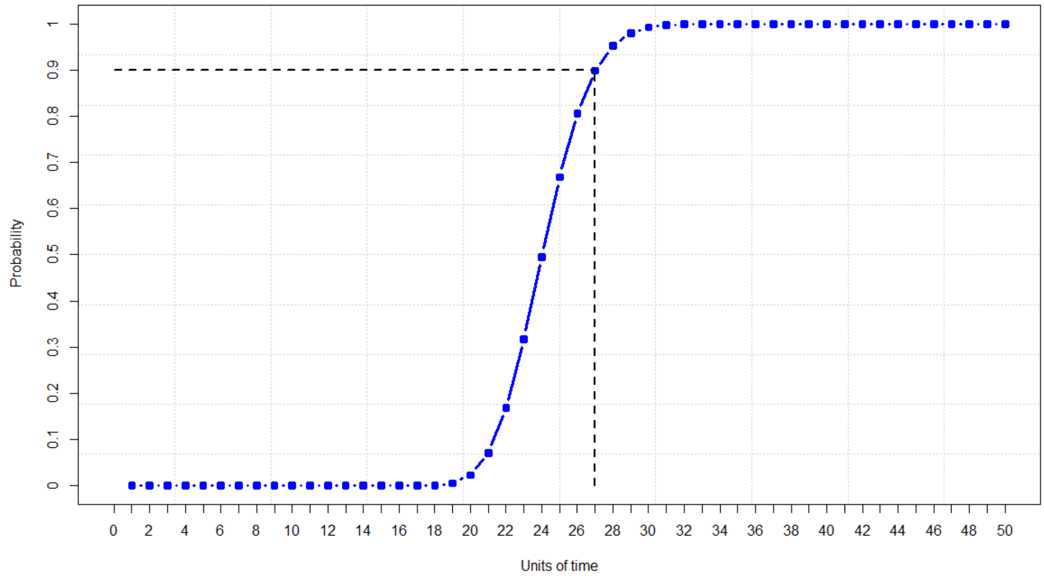


Fig. 3.5: Cumulative probability

R code is given in Appendix 3.D for this assignment

3.6.2 Problem B

A transportation network connecting city A to city B is shown in Figure 3.6. In order to go from city A to city B and vice versa, vehicles can travel in both directions of every link. Links are either a road section, a bridge or a tunnels. It is periodically affected by natural hazards such as flood, avalanches, and rockfall. The probability of the links not providing an adequate LOS due to the natural hazards can be modelled using the exponential or the Weibull distributions as shown in Table 3.11. The mean time required to restore an adequate LOS following the occurrence of a natural hazard on each of the links is also is included in Table 3.11.

Table 3.11: Characteristics of links in the network

Link object	Deterioration distribution	Weibul		Exponential	Mean time of intervention
		α	m	θ	(days)
Road A-1	Exponential	NA	NA	0.024	20
Road 2-3	Exponential	NA	NA	0.023	15
Road 3-4	Exponential	NA	NA	0.03	17
Road 3-5	Exponential	NA	NA	0.06	33
Road 4-5	Exponential	NA	NA	0.04	40
Bridge 1-2	Weibull	0.025	1.5	NA	30
Bridge 1-3	Weibull	0.03	2	NA	50
Bridge A-4	Weibull	0.025	2.3	NA	70
Tunnel B-2	Exponential	NA	NA	0.04	60
Tunnel B-5	Exponential	NA	NA	0.03	45

The impacts incurred when each link does not provide an adequate LOS, and all others do, are given in Table 3.12, in terms of monetary units (*mus*) incurred per day of inadequate service

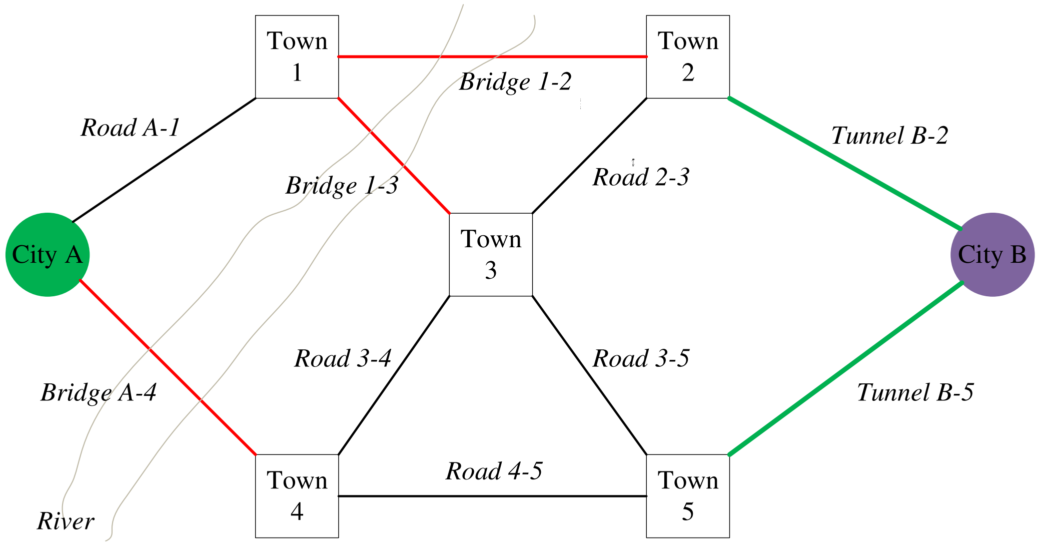


Fig. 3.6: A transportation network connecting city A to city B

Table 3.12: Impacts

Link	Impacts (mus)			
	Owner	Users	DAP	IAP
Road A-1	5	6	3	3
Road 2-3	7	5	4	2
Road 3-4	5	7	4	3
Road 3-5	4	6	4	2
Road 4-5	4	4	3	1
Bridge 1-2	7	6	4	3
Bridge 1-3	8	5	4	2
Bridge A-4	7	7	3	1
Tunnel B-2	6	4	4	2
Tunnel B-5	4	3	4	2

Note: DAP and IAP stand for directly affected public and indirectly affected public

3.6.2.1 Question B1

Estimate the reliability and availability of each link and of the entire network for each year in 20 years.

3.6.2.2 Answer B1

Step 1: Estimate the reliability of each link

The network can be simplified with the circles representing the links and the connecting point representing the cities and towns. The reliability of each link is given in Table 3.13 and illustrated in Figure 3.9.

Step 2: Estimate the reliability of the network

First, it is necessary to determine the structure function, which can be calculated using the pivotal decomposition method.

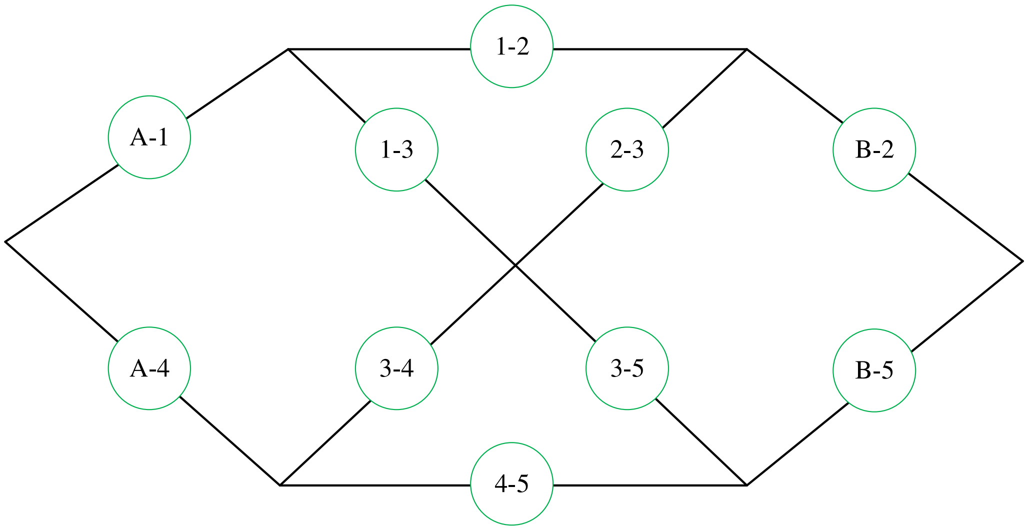


Fig. 3.7: Network simplification

Table 3.13: Reliability

Time (years)	Links										Networks
	Road sections					Bridge			Tunnels		
	A-1	2-3	3-4	3-5	4-5	1-2	1-3	A-4	B-2	B-5	
1	1	1	1	1	1	1	1	1	1	1	1
2	0.98	0.98	0.97	0.94	0.96	1	1	1	0.96	0.97	1
3	0.95	0.96	0.94	0.89	0.92	0.99	1	1	0.92	0.94	0.99
4	0.93	0.93	0.91	0.84	0.89	0.98	0.99	1	0.89	0.91	0.98
5	0.91	0.91	0.89	0.79	0.85	0.97	0.99	1	0.85	0.89	0.97
6	0.89	0.89	0.86	0.74	0.82	0.96	0.98	0.99	0.82	0.86	0.95
7	0.87	0.87	0.84	0.7	0.79	0.94	0.97	0.99	0.79	0.84	0.92
8	0.85	0.85	0.81	0.66	0.76	0.93	0.96	0.98	0.76	0.81	0.89
9	0.83	0.83	0.79	0.62	0.73	0.91	0.94	0.98	0.73	0.79	0.86
10	0.81	0.81	0.76	0.58	0.7	0.9	0.93	0.97	0.7	0.76	0.82
11	0.79	0.79	0.74	0.55	0.67	0.88	0.91	0.96	0.67	0.74	0.78
12	0.77	0.78	0.72	0.52	0.64	0.87	0.9	0.95	0.64	0.72	0.74
13	0.75	0.76	0.7	0.49	0.62	0.85	0.88	0.94	0.62	0.7	0.69
14	0.73	0.74	0.68	0.46	0.59	0.83	0.86	0.93	0.59	0.68	0.65
15	0.71	0.72	0.66	0.43	0.57	0.81	0.84	0.91	0.57	0.66	0.6
16	0.7	0.71	0.64	0.41	0.55	0.79	0.82	0.9	0.55	0.64	0.56
17	0.68	0.69	0.62	0.38	0.53	0.78	0.79	0.89	0.53	0.62	0.52
18	0.66	0.68	0.6	0.36	0.51	0.76	0.77	0.87	0.51	0.6	0.47
19	0.65	0.66	0.58	0.34	0.49	0.74	0.75	0.85	0.49	0.58	0.43
20	0.63	0.65	0.57	0.32	0.47	0.72	0.72	0.83	0.47	0.57	0.39

First, link 1-3 is selected as the pivot. When link 1-3 is perfect, the original network can be further simplified as shown in Figure 3.9, and when link 1-3 is not perfect, the original network diagram can be further simplified as shown in Figure 3.10

The structure function of the network can then be expressed as:

$$\phi(\mathbf{x}) = x_{1-3} \cdot r(1_{1-3}, \mathbf{x}) + (1 - x_{1-3}) \cdot r(0_{1-3}, \mathbf{x}) \tag{3.20}$$

However, even with the newly simplified network, still it is difficult to estimate the value of the sub-structure function $r(1_{1-3}, \mathbf{x})$ and $r(0_{1-3}, \mathbf{x})$directly. It is, therefore, necessary to further decompose the simplified sub-networks.

With respect to the network $r(1_{1-3}, \mathbf{x})$, the link 3-5 is selected to be the pivot. When the link 3-5 is 100% reliable, the network in Figure 6 becomes the one shown in Figure 3.11, and when the link 3-5 is 0% reliable, the network in Figure 6 becomes the one shown in Figure 3.12.

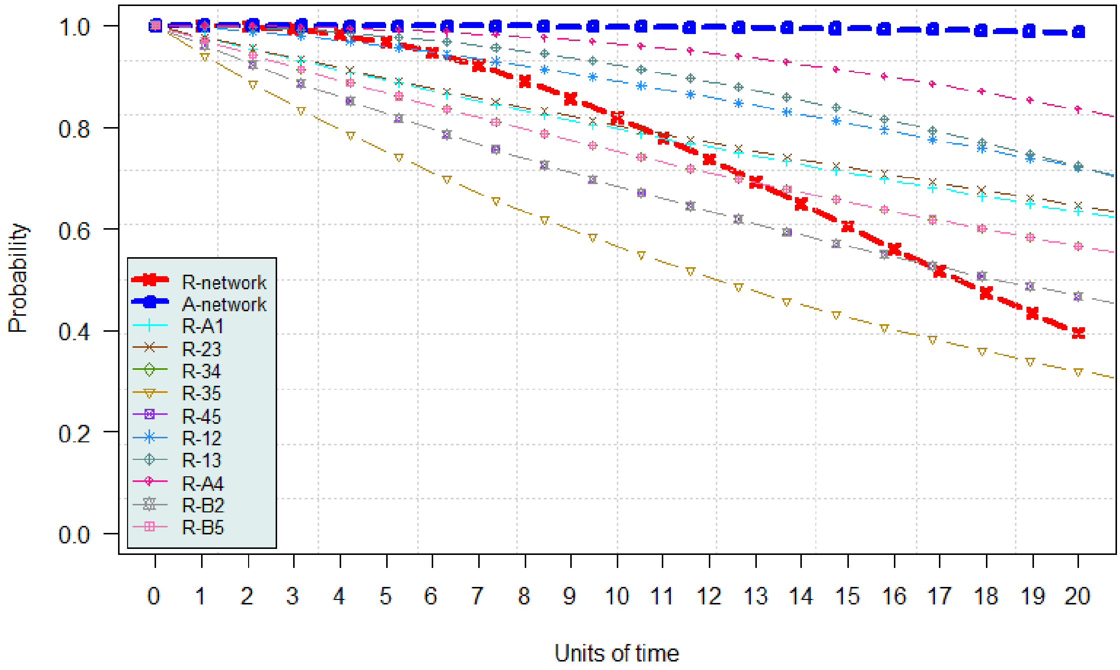


Fig. 3.8: Reliability of each link and network + network availability over 20 years

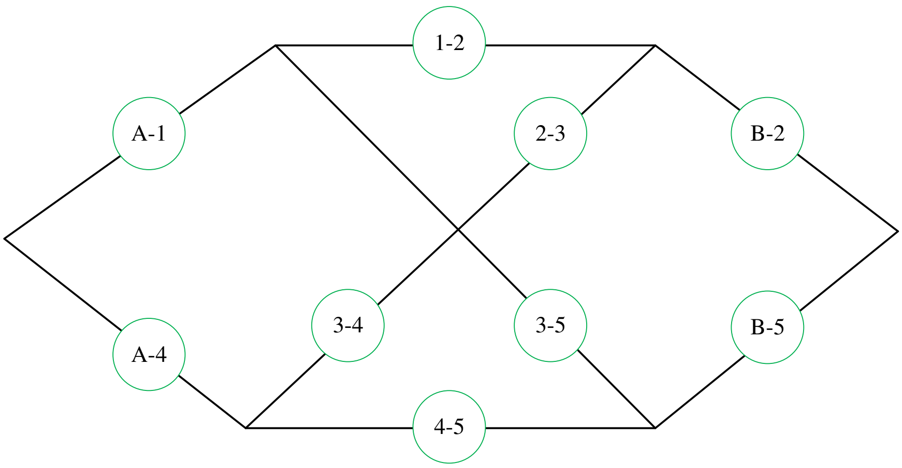


Fig. 3.9: Network simplification - decomposition when when $x_{1-3} = 1$

The structure function $r(1_{1-3}, \mathbf{x})$ is then:

$$r(1_{1-3}, \mathbf{x}) = x_{3-5} \cdot r^1_{1-3}(1_{3-5}, \mathbf{x}) + (1 - x_{3-5}) r^1_{1-3}(0_{3-5}, \mathbf{x}) \tag{3.21}$$

where $r^1_{1-3}(1_{3-5}, \mathbf{x})$ and $r^1_{1-3}(0_{3-5}, \mathbf{x})$refer to the sub-structure functions after link 1-3 and link 3-5 have been decomposed when link 1-3 is perfect.

The value of $r^1_{1-3}(1_{3-5}, \mathbf{x})$ and $r^1_{1-3}(0_{3-5}, \mathbf{x})$can now be directly computed as follows:

$$\begin{aligned} r^1_{1-3}(1_{3-5}, \mathbf{x}) = [1 - (1 - x_{A-1})(1 - x_{A-4} \cdot \{1 - (1 - x_{3-4}) \cdot (1 - x_{4-5})\})] \cdot \\ [1 - (1 - x_{B-5}) \cdot (1 - x_{B-2} \cdot \{1 - (1 - x_{1-2}) \cdot (1 - x_{2-3})\}] \end{aligned} \tag{3.22}$$

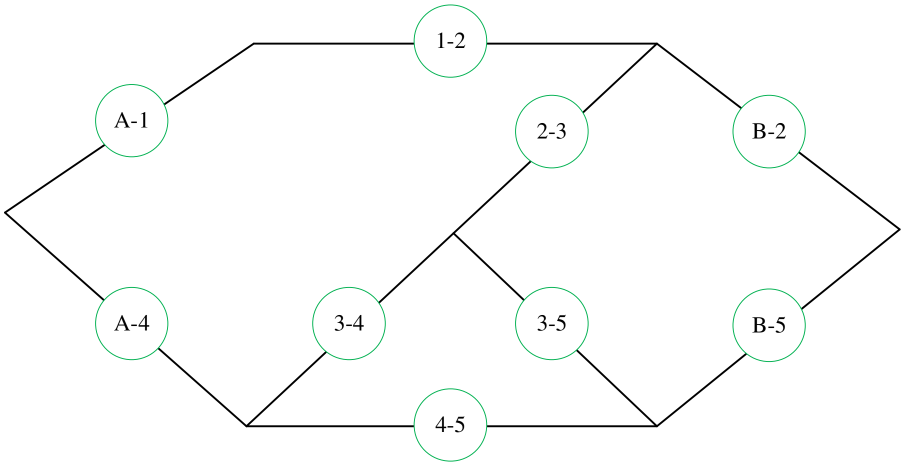


Fig. 3.10: Network simplification - decomposition when when $x_{1-3} = 0$

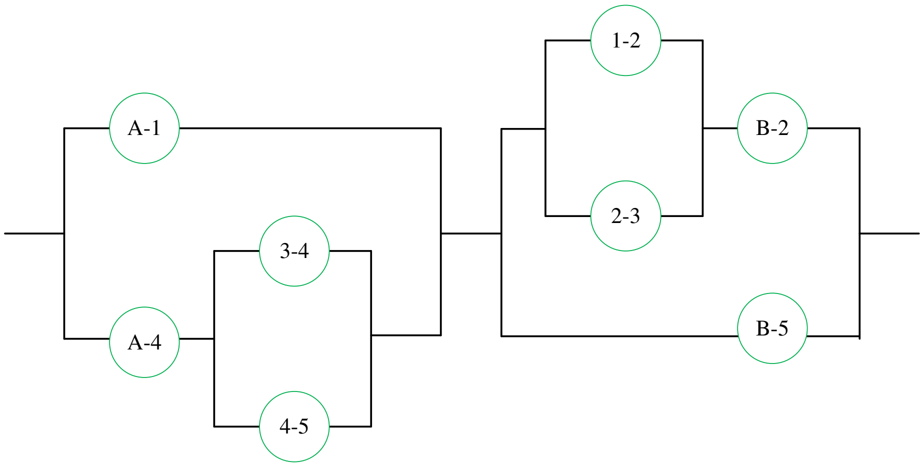


Fig. 3.11: Network simplification - decomposition when $x_{3-5} = 1$ and $x_{1-3} = 1$

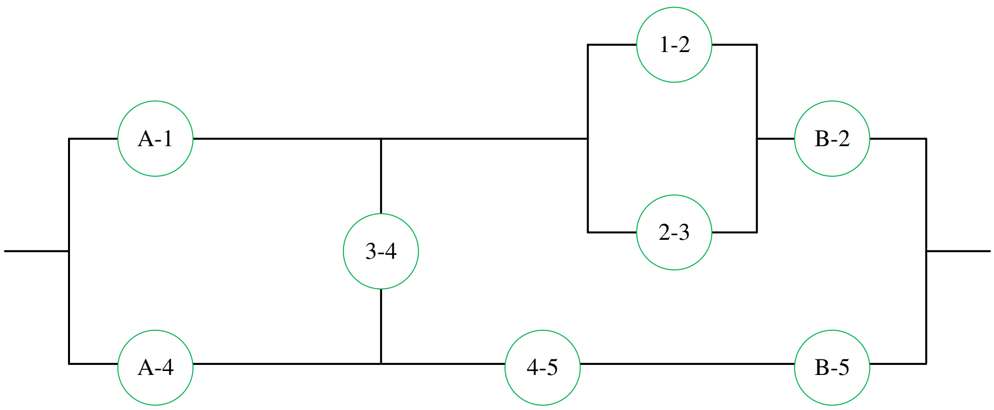


Fig. 3.12: Network simplification - decomposition when $x_{3-5} = 0$ and $x_{1-3} = 1$

$$
\begin{aligned}
r^1_{1-3}(0_{3-5},\mathbf{x}) &= x_{A-1}x_{3-4}\cdot x_{B-2}\cdot[1-(1-x_{1-2})\cdot(1-x_{2-3})] \\
&+x_{A-1}x_{3-4}x_{4-5}x_{B-5} \\
&+x_{A-4}x_{3-4}x_{B-2}\cdot[1-(1-x_{1-2})\cdot(1-x_{2-3})] \\
&+x_{A-4}x_{3-4}x_{4-5}x_{B-5} \\
&+x_{A-1}x_{A-4}x_{4-5}x_{B-5}x_{B-2}\cdot[1-(1-x_{1-2})\cdot(1-x_{2-3})] \\
&-x_{A-1}x_{A-4}x_{3-4}x_{B-2}\cdot[1-(1-x_{1-2})\cdot(1-x_{2-3})] \\
&-x_{A-1}x_{3-4}x_{B-2}\cdot[1-(1-x_{1-2})\cdot(1-x_{2-3})]x_{4-5}x_{B-5} \\
&-x_{A-1}x_{A-4}x_{3-4}x_{4-5}x_{B-5} \\
&-x_{A-4}x_{3-4}x_{B-2}\cdot[1-(1-x_{1-2})\cdot(1-x_{2-3})]x_{4-5}x_{B-5}
\end{aligned}
\tag{3.23}
$$

With respect to the network $r(0_{1-3},\mathbf{x})$, the link 3-5 is choosen to be the pivot. When link 3-5 is 100% reliable, the network in Figure 3.10 becomes that shown in Figure 3.13 and when the link 3-5 is 0% reliable, the network in Figure 3.10 becomes that in Figure 3.14

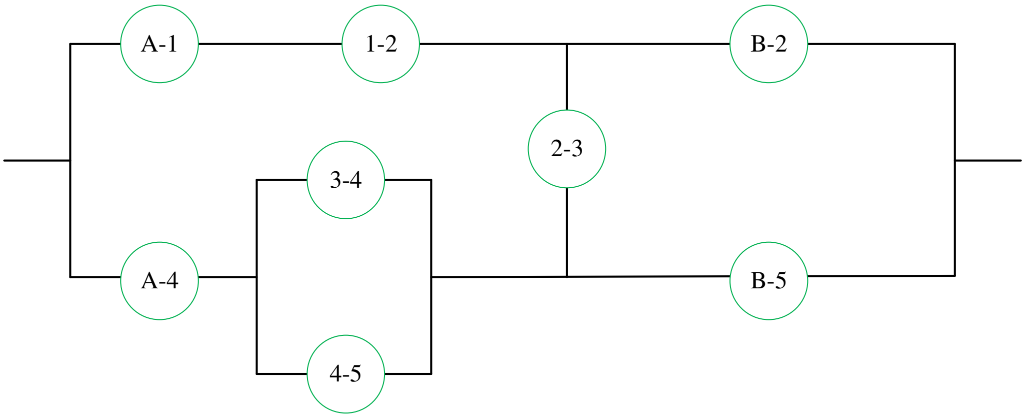


Fig. 3.13: Network simplification - decomposition $x_{3-5}=1$ and $x_{1-3}=0$

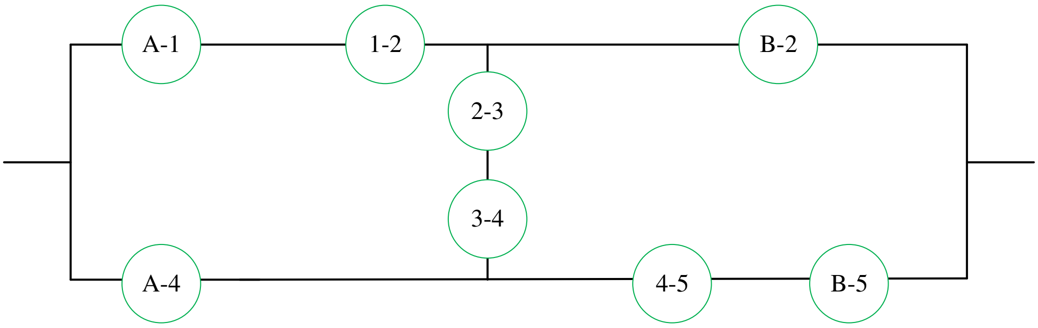


Fig. 3.14: Network simplification - decomposition $x_{3-5}=0$ and $x_{1-3}=0$

The structure function $r(0_{1-3},\mathbf{x})$ is:

$$r(0_{1-3},\mathbf{x}) = x_{3-5}\cdot r^0_{1-3}(1_{3-5},\mathbf{x}) + (1-x_{3-5})\cdot r^0_{1-3}(0_{3-5},\mathbf{x}) \tag{3.24}$$

where $r^0_{1-3}(1_{3-5},\mathbf{x})$ and $r^0_{1-3}(0_{3-5},\mathbf{x})$refer to the sub-structure function after link 1-3 and link 3-5 have been decomposed when link 1-3 is 0% reliable.

The values of $r^0_{1-3}(1_{3-5},\mathbf{x})$ and $r^0_{1-3}(0_{3-5},\mathbf{x})$ can now be directly computed

$$\begin{aligned}
&r^0_{1-3}(1_{3-5},\mathbf{x}) = x_{A-1}\cdot x_{1-2}\cdot x_{2-3}\cdot x_{B-2}\\
&+x_{A-1}\cdot x_{1-2}\cdot x_{2-3}\cdot x_{B-5}\\
&+x_{A-4}\cdot[1-(1-x_{3-4})\cdot(1-x_{4-5})]\cdot x_{2-3}\cdot x_{B-2}\\
&+x_{A-4}\cdot[1-(1-x_{3-4})\cdot(1-x_{4-5})]\cdot x_{2-3}\cdot x_{B-5}\\
&+x_{A-1}\cdot x_{1-2}\cdot x_{A-4}\cdot[1-(1-x_{3-4})\cdot(1-x_{4-5})]\cdot x_{B-5}\cdot x_{B-2}\\
&-x_{A-1}\cdot x_{1-2}\cdot x_{A-4}\cdot[1-(1-x_{3-4})\cdot(1-x_{4-5})]\cdot x_{2-3}\cdot x_{B-2}\\
&-x_{A-1}\cdot x_{1-2}\cdot x_{2-3}\cdot x_{B-2}\cdot x_{B-5}\\
&-x_{A-1}\cdot x_{1-2}\cdot x_{A-4}\cdot[1-(1-x_{3-4})\cdot(1-x_{4-5})]\cdot x_{2-3}\cdot x_{B-5}\\
&-x_{A-4}\cdot[1-(1-x_{3-4})\cdot(1-x_{4-5})]\cdot x_{2-3}\cdot x_{B-5}\cdot x_{B-2}
\end{aligned} \tag{3.25}$$

The set of equation from equation (3.20) to equation (3.26) can be used directly to calculate the reliability of the network given the reliability of the individual link every year. The reliability of each link and the network for periods of time ranging from 0 to 20 years are given in Table 3.13 and illustrated in Figure 3.9.

R code for this example is given in Appendix B.

Step 3: Estimate the availability of the network

Similar to reliability, the set of structure functions above can be used to compute the availability of the network based on the reliability of individual link. The availability of individual link is calculated based on following equation

$$A(t_i) = \frac{E\left[t_i^a\right]}{E\left[t_i^a\right] + E\left[t_i^n\right]} = \frac{\int_0^{tu} R_i(t)dt}{\int_0^{tu} R_i(t)dt + \Delta_i \cdot [1 - R_i(tu)]} \tag{3.26}$$

where $R_i(t)$ is the reliability of link i at time t and $\Delta_i$is the mean time of intervention for link i.

The availabilities of each link and the network for periods of time from 0 to 20 years are given in Table 3.14, and illustrated in Figure 3.9.

Table 3.14: Availability

Time (years)	Links										Networks
	Road sections					Bridge			Tunnels		
	A-1	2-3	3-4	3-5	4-5	1-2	1-3	A-4	B-2	B-5	
1	1	1	1	1	1	1	1	1	1	1	1
2	0.999335	0.999502	0.999431	0.998864	0.99865	0.999022	0.998376	0.99773	0.997976	0.998496	0.999995
3	0.999114	0.999336	0.99924	0.998459	0.998187	0.998707	0.997859	0.997011	0.997282	0.99799	0.999992
4	0.999003	0.999253	0.999142	0.998237	0.997945	0.998557	0.997615	0.996675	0.99692	0.997733	0.99999
5	0.998937	0.999204	0.999083	0.998088	0.997791	0.998471	0.99748	0.996492	0.996691	0.997575	0.999988
6	0.998892	0.999171	0.999042	0.997975	0.997682	0.998417	0.997397	0.996385	0.996527	0.997467	0.999987
7	0.998861	0.999148	0.999012	0.997883	0.997599	0.998382	0.997344	0.996319	0.996402	0.997388	0.999986
8	0.998837	0.999131	0.998988	0.997804	0.997531	0.998357	0.997308	0.996279	0.996301	0.997327	0.999986
9	0.998819	0.999118	0.998969	0.997734	0.997473	0.998339	0.997284	0.996255	0.996215	0.997277	0.999985
10	0.998804	0.999107	0.998954	0.99767	0.997424	0.998325	0.997266	0.996242	0.99614	0.997235	0.999985
11	0.998792	0.999099	0.99894	0.99761	0.997379	0.998315	0.997254	0.996236	0.996074	0.9972	0.999984
12	0.998782	0.999092	0.998929	0.997554	0.997339	0.998308	0.997244	0.996236	0.996014	0.997169	0.999984
13	0.998774	0.999086	0.998918	0.9975	0.997302	0.998302	0.997236	0.996238	0.995958	0.997142	0.999984
14	0.998766	0.999081	0.998909	0.997448	0.997268	0.998297	0.99723	0.996243	0.995907	0.997118	0.999983
15	0.99876	0.999076	0.998901	0.997398	0.997235	0.998293	0.997225	0.99625	0.995858	0.997096	0.999983
16	0.998755	0.999073	0.998893	0.997349	0.997204	0.998289	0.997219	0.996257	0.995812	0.997075	0.999983
17	0.99875	0.99907	0.998886	0.997302	0.997175	0.998286	0.997214	0.996264	0.995768	0.997057	0.999983
18	0.998746	0.999067	0.998879	0.997256	0.997147	0.998284	0.997208	0.996272	0.995726	0.997039	0.999983
19	0.998742	0.999064	0.998873	0.99721	0.99712	0.998281	0.997202	0.996279	0.995686	0.997023	0.999982
20	0.998738	0.999062	0.998867	0.997165	0.997093	0.998279	0.997196	0.996286	0.995647	0.997007	0.999982

3.6.2.3 Question B2

Calculate the total impacts due to links not being able to provide an adequate LOS and the network over a 20 year period. (Assume that link failures are mutually exclusive.)

3.6.2.4 Answer B2

The expected cost is estimated based on following equation

$$C_{am} = \frac{E\left[t_i^n\right] \cdot I_{cr}/12}{E\left[t_i^a\right] + E\left[t_i^n\right]} \tag{3.27}$$

The mean time between failures can be calculated with following equation:

$$E\left[t_i^a\right] = \Theta_i = \int_0^\infty r_i(t)dt \tag{3.28}$$

where $r_i(t)$ is the reliability of link i.

Table 3.15 shows the summary of total cost for each link and the total cost for the network

Table 3.15: Cost

Time (years)	Links										Total costs
	Road sections					Bridge			Tunnels		
	A-1	2-3	3-4	3-5	4-5	1-2	1-3	A-4	B-2	B-5	
Owner	18.27	18.56	18.39	44.18	42.18	29.96	57.86	43.03	94.78	38.87	406.08
User	21.93	13.26	25.74	66.27	42.18	25.68	36.16	43.03	63.19	29.15	366.59
DAP	10.96	10.61	14.71	44.18	31.64	17.12	28.93	18.44	63.19	38.87	278.64
IAP	10.96	5.3	11.03	22.09	10.55	12.84	14.46	6.15	31.59	19.43	144.42
Total	62.13	47.74	69.87	176.72	126.54	85.6	137.41	110.64	252.75	126.32	1195.73

References

3.A Example - Reliability & Availability

```
#

#This program is coded by Nam Lethanh for use in the class IMP
#Purpose: To calculate reliability
TIME=30
T=seq(0,TIME, by =1) #investigate time
N=6 #Number of item
lambda<-matrix(double(1),nrow=1,ncol=N)
lambda<-1/20 # value of hazard rate

delta<-matrix(double(1),nrow=length(T),ncol=N-1)

delta[1]<-1*7/365
delta[2]<-2*7/365
delta[3]<-5*7/365
delta[4]<-4*7/365
delta[5]<-3*7/365


# define dimension of matrix
fail<-matrix(double(1),nrow=length(T),ncol=N)
survive<-matrix(double(1),nrow=length(T),ncol=N)
availability<-matrix(double(1),nrow=length(T),ncol=N)


#reliability function for exponential distribution
failure<-function(t,lambda){lambda*exp(-lambda*t)}
survival<-function(t,lambda){exp(-lambda*t)}
#

for (t in 1:length(T)){
  for (n in 1:N){
    if (n < N){
      survive[t,n]<-survival(T[t],lambda)
      availability[t,n]<-integrate(survival,0,t,lambda)$value/(integrate(survival,0,t,lambda)$value+1*delta[n]*(1-survive[t,n]))
   #  availability[t,n]<-integrate(survival,0,t,lambda)$value/(integrate(survival,0,t,lambda)$value+delta[n]*integrate(failure,0,t,lambda)$value)
    } else {
      survive[t,n]=
        survive[t,1]*survive[t,3]*survive[t,4]+
        survive[t,1]*survive[t,3]*survive[t,5]+
        survive[t,2]*survive[t,3]*survive[t,4]+
        survive[t,2]*survive[t,3]*survive[t,5]+
        survive[t,1]*survive[t,2]*survive[t,4]*survive[t,5]-
        survive[t,1]*survive[t,2]*survive[t,3]*survive[t,4]-
        survive[t,1]*survive[t,3]*survive[t,4]*survive[t,5]-
        survive[t,1]*survive[t,2]*survive[t,3]*survive[t,5]-
        survive[t,2]*survive[t,3]*survive[t,4]*survive[t,5]

      availability[t,n]=
        availability[t,1]*availability[t,3]*availability[t,4]+
        availability[t,1]*availability[t,3]*availability[t,5]+
        availability[t,2]*availability[t,3]*availability[t,4]+
        availability[t,2]*availability[t,3]*availability[t,5]+
        availability[t,1]*availability[t,2]*availability[t,4]*availability[t,5]-
        availability[t,1]*availability[t,2]*availability[t,3]*availability[t,4]-
        availability[t,1]*availability[t,3]*availability[t,4]*availability[t,5]-
        availability[t,1]*availability[t,2]*availability[t,3]*availability[t,5]-
        availability[t,2]*availability[t,3]*availability[t,4]*availability[t,5]
    }
  }
}
cat("Reliability of each item and the network \n")
print(survive)
cat("Availability of each item and the network \n")
print(availability)


#Plotting

plot.new()
par(mar=c(5,4,4,6)+0.3)
limy=c(0,1)
limx=c(0,TIME)
plot(survive[,1],lwd=2,col="red",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=4,cex=0.8)
axis(2,ylim=limy,col="black",las=1)
axis(1,c(seq(0,TIME,by=1)),c(seq(0,TIME,by=1)))
mtext(expression(paste('Probability')),side=2,col="black",line=3)
mtext(expression(paste('Units of time')),side=1,col="black",line=3)
grid(10, 10, col = "lightgray", lty = "dotted",lwd = par("lwd"), equilogs = TRUE)
box()
par(new=TRUE)
```

```
limy=c(0,1)
limx=c(0,TIME)
plot(survive[,6],lwd=2,col="blue",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=2,cex=0.8)

par(new=TRUE)
limy=c(0,1)
limx=c(0,TIME)
plot(availability[,6],lwd=1,col="darkblue",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=8,
     cex=0.7)

par(new=TRUE)
limy=c(0,1)
limx=c(0,TIME)
plot(availability[,1],lwd=1,col="coral",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=9,cex
     =0.2)

par(new=TRUE)
limy=c(0,1)
limx=c(0,TIME)
plot(availability[,2],lwd=1,col="coral4",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=10,
     cex=0.2)

par(new=TRUE)
limy=c(0,1)
limx=c(0,TIME)
plot(availability[,3],lwd=1,col="brown3",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=11,
     cex=0.2)

par(new=TRUE)
limy=c(0,1)
limx=c(0,TIME)
plot(availability[,4],lwd=1,col="darkgrey",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch
     =12,cex=0.2)

par(new=TRUE)
limy=c(0,1)
limx=c(0,TIME)
plot(availability[,5],lwd=1,col="darkorchid1",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch
     =22,cex=0.2)

colors=c("red","blue","darkblue","coral","coral4","brown3","darkgrey")
legend("bottomleft", inset=0.009, col=colors,lty=2,lwd=2,legend=c("R-subitem","R-network","A-network","A-1","A
     -2","A-3","A-4","A-5"),pch=c(4,2,8,9,10,11,12,22),bg="azure2",cex=0.8)

file.remove("data.csv")
file.create("data.csv")
write.table(availability, file="data.csv", sep = ",", append = TRUE,col.names = FALSE)
```

3.B Example - HongKong tunnel

```
#This program is coded by Nam Lethanh for use in the class IMP
#Purpose: To calculate reliability, availability, and maintainability of tunnel

data <- read.csv("IMP-5-E-tunnel.csv",header=TRUE,sep=",") #
attach(data)
TIME=20
T=seq(0,TIME, by =1) #investigate time
N=5 #Number of item
M<-5 #number of configuration

cat("how data look like? \n")
print(data)
theta<-matrix(double(1),nrow=M,ncol=N)

MTBF<-matrix(double(1),nrow=M,ncol=1) #mean time between failure for each configuration
Reli<-matrix(double(1),nrow=M,ncol=1) #Reliability for each configuration
Avai<-matrix(double(1),nrow=M,ncol=1) #Availability for each configuration
Cost<-matrix(double(1),nrow=M,ncol=1) #Cost for each configuration in 20 years
Benefit_no_upfront<-matrix(double(1),nrow=M,ncol=1) #
Benefit_with_upfront<-matrix(double(1),nrow=M,ncol=1) #

for (i in 1:N){
  for (j in 1:M){
    theta[i,j]<-1/data[i,j+1]
  }
}
cat("Value of failure rate for each item i (column) in each configuration j (row) \n")
print(theta)

for (i in 1:N){
  MTBF[i]<-1/sum(theta[i,])
  Reli[i]<-exp(-1/MTBF[i]*1)
  Avai[i]<-MTBF[i]/(MTBF[i]+data$MID[i]/12)
  Cost[i]<-20*data$costcr[i]*data$MID[i]*(1-Reli[i])/(MTBF[i]+data$MID[i]/12)

}
cat("Value of mean time between failure (MTBF) of each configuration \n")
print(MTBF)

cat("Reliability of each configuration (Reli) \n")
print(Reli)

cat("Availability of each configuration \n")
print(Avai)
cat("Cost of each configuration in 20 years \n")
print(Cost)

for (i in 1:M){
  if (i==1){
    Benefit_no_upfront[i]<-Cost[i]-Cost[i]
    Benefit_with_upfront[i]<-Benefit_no_upfront[i]-data$costint[i]
  } else {
    Benefit_no_upfront[i]<-Cost[1]-Cost[i]
    Benefit_with_upfront[i]<-Benefit_no_upfront[i]-data$costint[i]
  }
}
cat("Net benefit without upfront cost \n")
print(Benefit_no_upfront)

cat("Net benefit with upfront cost \n")
print(Benefit_with_upfront)
```

3.C Assignment - Availability-log-normal distribution

```
#This program is coded by Nam Lethanh for use in the class IMP-HS2014
#Purpose: To calculate availability
data <- read.csv("IMP-availability.csv",header=TRUE,sep=",") #
attach(data)
T=50
cat("How data looks like?\n")
print(data)
cat("range of observations \n")
range<-max(fre)-min(fre)
print(range)
cat("range of observations - interval \n")

rangetime<-c("9-12","13-16","17-20","21-24","25-28","29-32","33-36","37-40","41-44","45-48")

fre<-c(2,11,24,27,20,15,16,7,2,2)
plot.new()
limy=c(0,1)
barplot(fre, beside = TRUE, axes = FALSE, col = sample(colours(), 5), space = 0.5, names.arg=rangetime,ylim=c
    (0,30))
axis(2, ylim = limy, col = "black", las = 1)
mtext(expression(paste('Frequency')), side = 2, col = "black", line = 3)
mtext(expression(paste('Range of time')), side = 1, col = "black", line = 3)
box()

library(graphics)
par(new=TRUE)
plot(fre,lwd=2,col="coral4",ylab="",xlab="",ylim=c(0,30),axes=FALSE,lty=1,type="b",pch=10,cex=0.2)
legend("topright", inset = 0.05, col = "coral4", lty = 1, lwd = 2, legend = c("Fit spline"), bg = "aliceblue",
    cex = 0.8)

#calculate the mean corrective intervention duration
MCID<-sum((data$time)*data$fre)/sum(data$fre)
cat("mean corrective intervention duration \n")
print(MCID)
#calculate the mean of sample
VarX<-(sum(data$fre*(data$time-MCID)^2)/sum(data$fre))^(0.5)
cat("value of standard deviation of sample \n")
print(VarX)

#calculate the geometric mean

lognormalmean<-log(MCID)-0.5*(log(1+VarX^2/(MCID)^2))
cat("value of mean of the lognormal distribution \n")
print(lognormalmean)
cat("value of the geometric mean \n")
geomean<-exp(lognormalmean)
print(geomean)

#calculate the time corresponding to 90% of cummulative distribution
cat("standard deviation of the lognormal \n")
Varx<-log(1+(VarX/MCID)^2)
print(Varx)

#define errors function
erf1=function(x) 2 * pnorm(x * sqrt(2)) - 1

library(pracma)

PDF<- matrix(double(1),nrow=T,ncol=1) # p.d.f
CDF1<-matrix(double(1),nrow=T,ncol=1) #c.d.f
CDF<-matrix(double(1),nrow=T,ncol=1) #c.d.f

x<-matrix(double(1),nrow=T,ncol=1) #c.d.f

pi=3.14159292035398230088888888888888888
for (t in 1:T){
  PDF[t]<-(1/(t*Varx*sqrt(2*pi)))*exp(-((log(t)-lognormalmean)/(2*Varx))^2)
  CDF1[t]<-0.5*(1+erf1((log(t)-lognormalmean)/(Varx*sqrt(2))))
  CDF[t]<-0.5*erfc(-(log(t)-lognormalmean)/(Varx*sqrt(2)))
}
print(cbind(CDF1,CDF))

plot.new()
limy=c(0,1)
limx=c(0,T)
plot(CDF1,lwd=3,col="red",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=12,cex=0.8)
axis(2,c(seq(0,1,by=0.1)),c(seq(0,1,by=0.1)))
axis(1,c(seq(0,T,by=1)),c(seq(0,T,by=1)))
mtext(expression(paste('Probability')),side=2,col="black",line=3)
mtext(expression(paste('Units of time')),side=1,col="black",line=3)
grid(10, 10, col = "lightgray", lty = "dotted",lwd = par("lwd"), equilogs = TRUE)
box()
```

```
par(new=TRUE)
plot(CDF,lwd=3,col="blue",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=12,cex=0.8)

segments(0, 0.9, 27, 0.9, col= 'black',lty=2,lwd=2)

segments(27, 0, 27, 0.9, col= 'black',lty=2,lwd=2)
```

3.D Assignment - availability, reliability, and impact estimation for the network

```
#This program is coded by Nam Lethanh for use in the class IMP-HS2014
#Purpose: To calculate availability
data <- read.csv("IMP-10-HS2014-A-availability-network.csv",header=TRUE,sep=",") #
attach(data)
TIME=20
T=seq(0,TIME, by =1) #investigate time
N=10
cat("How data looks like?\n")
print(data)

#define function
#reliability function for exponential distribution
failureexp<-function(t,lambda){lambda*exp(-lambda*t)}
survivalexp<-function(t,lambda){exp(-lambda*t)}
#reliability function for weibull distribution

failureweib<-function(t,theta,m){theta*m*(theta*t)^(m-1)*exp(-(theta*t)^m)}
rateweib<-function(t,theta,m){theta*m*(theta*t)^(m-1)}
survivalweib<-function(t,theta,m){exp(-(theta*t)^m)}
#define structure function
decompose<-function(x,y,z){
  x*y+(1-x)*z
}
#calculating the structure function when link 1-3 is perfect and link 3-5 is perfect
expression1<-function(xA1,xA4,x34,x45,xB5,xB2,x12,x23){
(1-(1-xA1)*(1-xA4*(1-(1-x34)*(1-x45))))*(1-(1-xB5)*(1-xB2*(1-(1-x12)*(1-x23))))
}

#calculating the structure function when link 1-3 is perfect and link 3-5 is not perfect

expression2<-function(xA1,xA4,x34,x45,xB5,xB2,x12,x23){
xA1*x34*xB2*(1-(1-x12)*(1-x23))+xA1*x34*x45*xB5+xA4*x34*xB2*(1-(1-x12)*(1-x23))+xA4*x34*x45*xB5+xA1*xA4*x45*xB5*xB2*(1-(1-x12)*(1-x23))-xA1*xA4*x34*xB2*(1-(1-x12)*(1-x23))-xA1*x34*xB2*(1-(1-x12)*(1-x23))*x45*xB5-xA1*xA4*x34*x45*xB5-xA4*x34*xB2*(1-(1-x12)*(1-x23))*x45*xB5
}

#calculating the structure function when link 1-3 is not perfect and link 3-5 is perfect
expression3<-function(xA1,xA4,x34,x45,xB5,xB2,x12,x23){
xA1*x12*x23*xB2+xA1*x12*x23*xB5+xA4*(1-(1-x34)*(1-x45))*x23*xB2+xA4*(1-(1-x34)*(1-x45))*x23*xB5+xA1*x12*xA4*(1-(1-x34)*(1-x45))*xB5*xB2-xA1*x12*xA4*(1-(1-x34)*(1-x45))*x23*xB2-xA1*x12*x23*xB2*xB5-xA1*x12*xA4*(1-(1-x34)*(1-x45))*x23*xB5-xA4*(1-(1-x34)*(1-x45))*x23*xB5*xB2
}

#calculating the structure function when link 1-3 is not perfect and link 3-5 is not perfect
expression4<-function(xA1,xA4,x34,x45,xB5,xB2,x12,x23){
xA1*x12*x23*x34*xB2+xA1*x12*x23*x34*x45*xB5+xA4*x23*x34*xB2+xA4*x23*x34*x45*xB5+xA1*x12*xA4*xB2*x45*xB5-xA1*x12*xA4*x23*x34*xB2-xA1*x12*x23*x34*xB2*x45*xB5-xA1*x12*xA4*x23*x34*x45*xB5-xA4*x23*x34*xB2*x45*xB5
}

#reliability of the link
rA1<-matrix(double(1),nrow=length(T)-1,ncol=1)
rA4<-matrix(double(1),nrow=length(T)-1,ncol=1)
r12<-matrix(double(1),nrow=length(T)-1,ncol=1)
r13<-matrix(double(1),nrow=length(T)-1,ncol=1)
r23<-matrix(double(1),nrow=length(T)-1,ncol=1)
r34<-matrix(double(1),nrow=length(T)-1,ncol=1)
r35<-matrix(double(1),nrow=length(T)-1,ncol=1)
r45<-matrix(double(1),nrow=length(T)-1,ncol=1)
rB2<-matrix(double(1),nrow=length(T)-1,ncol=1)
rB5<-matrix(double(1),nrow=length(T)-1,ncol=1)
#reliability of the decompose function when link 1-3 is perfect
R131<-matrix(double(1),nrow=length(T)-1,ncol=1)
#reliability of the decompose function when link 1-3 is not perfect
R130<-matrix(double(1),nrow=length(T)-1,ncol=1)
#reliability of the network
R<-matrix(double(1),nrow=length(T)-1,ncol=1)


#failure of each link
fA1<-matrix(double(1),nrow=length(T)-1,ncol=1)
fA4<-matrix(double(1),nrow=length(T)-1,ncol=1)
f12<-matrix(double(1),nrow=length(T)-1,ncol=1)
f13<-matrix(double(1),nrow=length(T)-1,ncol=1)
f23<-matrix(double(1),nrow=length(T)-1,ncol=1)
f34<-matrix(double(1),nrow=length(T)-1,ncol=1)
f35<-matrix(double(1),nrow=length(T)-1,ncol=1)
f45<-matrix(double(1),nrow=length(T)-1,ncol=1)
fB2<-matrix(double(1),nrow=length(T)-1,ncol=1)
fB5<-matrix(double(1),nrow=length(T)-1,ncol=1)
#availability of each link
aA1<-matrix(double(1),nrow=length(T)-1,ncol=1)
aA4<-matrix(double(1),nrow=length(T)-1,ncol=1)
a12<-matrix(double(1),nrow=length(T)-1,ncol=1)
```

```
a13<-matrix(double(1),nrow=length(T)-1,ncol=1)
a23<-matrix(double(1),nrow=length(T)-1,ncol=1)
a34<-matrix(double(1),nrow=length(T)-1,ncol=1)
a35<-matrix(double(1),nrow=length(T)-1,ncol=1)
a45<-matrix(double(1),nrow=length(T)-1,ncol=1)
aB2<-matrix(double(1),nrow=length(T)-1,ncol=1)
aB5<-matrix(double(1),nrow=length(T)-1,ncol=1)

#reliability of the decompose function when link 1-3 is perfect
A131<-matrix(double(1),nrow=length(T)-1,ncol=1)
#reliability of the decompose function when link 1-3 is not perfect
A130<-matrix(double(1),nrow=length(T)-1,ncol=1)
#reliability of the network
A<-matrix(double(1),nrow=length(T)-1,ncol=1)


for (t in 1:(length(T))){
  #reliability
  rA1[t]<-survivalexp(T[t],data$alpha[1])
  r23[t]<-survivalexp(T[t],data$alpha[2])
  r34[t]<-survivalexp(T[t],data$alpha[3])
  r35[t]<-survivalexp(T[t],data$alpha[4])
  r45[t]<-survivalexp(T[t],data$alpha[5])
  r12[t]<-survivalweib(T[t],data$alpha[6],data$m[6])
  r13[t]<-survivalweib(T[t],data$alpha[7],data$m[7])
  rA4[t]<-survivalweib(T[t],data$alpha[8],data$m[8])
  rB2[t]<-survivalexp(T[t],data$alpha[9])
  rB5[t]<-survivalexp(T[t],data$alpha[10])
  #availability
  aA1[t]<-integrate(survivalexp,0,t,data$alpha[1])$value/(integrate(survivalexp,0,t,data$alpha[1])$value+(data$
     mit[1]/365)*(1-rA1[t]))
  a23[t]<-integrate(survivalexp,0,t,data$alpha[2])$value/(integrate(survivalexp,0,t,data$alpha[2])$value+(data$
     mit[2]/365)*(1-rA1[t]))
  a34[t]<-integrate(survivalexp,0,t,data$alpha[3])$value/(integrate(survivalexp,0,t,data$alpha[3])$value+(data$
     mit[3]/365)*(1-rA1[t]))
  a35[t]<-integrate(survivalexp,0,t,data$alpha[4])$value/(integrate(survivalexp,0,t,data$alpha[4])$value+(data$
     mit[4]/365)*(1-rA1[t]))
  a45[t]<-integrate(survivalexp,0,t,data$alpha[5])$value/(integrate(survivalexp,0,t,data$alpha[5])$value+(data$
     mit[5]/365)*(1-rA1[t]))
  a12[t]<-integrate(survivalweib,0,t,data$alpha[6],data$m[6])$value/(integrate(survivalweib,0,t,data$alpha[6],
     data$m[6])$value+(data$mit[6]/365)*(1-rA1[t]))
  a13[t]<-integrate(survivalweib,0,t,data$alpha[7],data$m[7])$value/(integrate(survivalweib,0,t,data$alpha[7],
     data$m[7])$value+(data$mit[7]/365)*(1-rA1[t]))
  aA4[t]<-integrate(survivalweib,0,t,data$alpha[8],data$m[8])$value/(integrate(survivalweib,0,t,data$alpha[8],
     data$m[8])$value+(data$mit[8]/365)*(1-rA1[t]))
  aB2[t]<-integrate(survivalexp,0,t,data$alpha[9])$value/(integrate(survivalexp,0,t,data$alpha[9])$value+(data$
     mit[9]/365)*(1-rA1[t]))
  aB5[t]<-integrate(survivalexp,0,t,data$alpha[10])$value/(integrate(survivalexp,0,t,data$alpha[10])$value+(
     data$mit[10]/365)*(1-rA1[t]))

  #reliability of the decompose link
 R131[T[t]]<-decompose(r35[T[t]],expression1(rA1[T[t]],rA4[T[t]],r34[T[t]],r45[T[t]],rB5[T[t]],rB2[T[t]],r12[T[
     t]],r23[T[t]]),expression2(rA1[T[t]],rA4[T[t]],r34[T[t]],r45[T[t]],rB5[T[t]],rB2[T[t]],r12[T[t]],r23[T[t]])
     )

 R130[T[t]]<-decompose(r35[T[t]],expression3(rA1[T[t]],rA4[T[t]],r34[T[t]],r45[T[t]],rB5[T[t]],rB2[T[t]],r12[T[
     t]],r23[T[t]]),expression4(rA1[T[t]],rA4[T[t]],r34[T[t]],r45[T[t]],rB5[T[t]],rB2[T[t]],r12[T[t]],r23[T[t]])
     )
 #reliability of the network
  R[T[t]]<-decompose(r13[T[t]],R131[T[t]],R130[T[t]])

 #availability of the decompose link
 A131[T[t]]<-decompose(a35[T[t]],expression1(aA1[T[t]],aA4[T[t]],a34[T[t]],a45[T[t]],aB5[T[t]],aB2[T[t]],a12[T[
     t]],a23[T[t]]),expression2(aA1[T[t]],aA4[T[t]],a34[T[t]],a45[T[t]],aB5[T[t]],aB2[T[t]],a12[T[t]],a23[T[t]])
     )

 A130[T[t]]<-decompose(a35[T[t]],expression3(aA1[T[t]],aA4[T[t]],a34[T[t]],a45[T[t]],aB5[T[t]],aB2[T[t]],a12[T[
     t]],a23[T[t]]),expression4(aA1[T[t]],aA4[T[t]],a34[T[t]],a45[T[t]],aB5[T[t]],aB2[T[t]],a12[T[t]],a23[T[t]])
     )
 #availability of the network
 A[T[t]]<-decompose(a13[T[t]],A131[T[t]],A130[T[t]])

}

cat("value of reliability for each link and the network")
R_link<-data.frame(rA1,r23,r34,r35,r45,r12,r13,rA4,rB2,rB5)
print(R_link)
print(R)
cat("value of availability for each link and the network")

A_link<-data.frame(aA1,a23,a34,a35,a45,a12,a13,aA4,aB2,aB5)
print(A_link)
print(A)


#plotting
#Plotting
```

```
plot.new()
par(mar=c(5,4,4,6)+0.3)
limy=c(0,1)
limx=c(0,TIME)
plot(R,lwd=4,col="red",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=4,cex=0.8)
axis(2,ylim=limy,col="black",las=1)
axis(1,c(seq(0,TIME,by=1)),c(seq(0,TIME,by=1)))
mtext(expression(paste('Probability')),side=2,col="black",line=3)
mtext(expression(paste('Units of time')),side=1,col="black",line=3)
grid(10, 10, col = "lightgray", lty = "dotted",lwd = par("lwd"), equilogs = TRUE)
box()

par(new=TRUE)
limx=c(1,TIME)
plot(A,lwd=4,col="blue",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=14,type="b",pch=22,cex=0.8)

par(new=TRUE)
limy=c(0,1)
limx=c(1,TIME)
plot(rA1,lwd=1,col="cyan",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=3,cex=0.6)

par(new=TRUE)
limy=c(0,1)
limx=c(1,TIME)
plot(r23,lwd=1,col="chocolate4",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=4,cex=0.6)

par(new=TRUE)
limy=c(0,1)
limx=c(1,TIME)
plot(r34,lwd=1,col="chartreuse4",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=5,cex=0.6)

par(new=TRUE)
limy=c(0,1)
limx=c(1,TIME)
plot(r35,lwd=1,col="darkgoldenrod",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=6,cex=0.6)

par(new=TRUE)
limy=c(0,1)
limx=c(1,TIME)
plot(r45,lwd=1,col="blueviolet",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=7,cex=0.6)


par(new=TRUE)
limy=c(0,1)
limx=c(1,TIME)
plot(r12,lwd=1,col="dodgerblue1",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=8,cex=0.6)

par(new=TRUE)
limy=c(0,1)
limx=c(1,TIME)
plot(r13,lwd=1,col="darkslategray4",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=9,cex
    =0.6)


par(new=TRUE)
limy=c(0,1)
limx=c(1,TIME)
plot(rA4,lwd=1,col="deeppink2",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=10,cex=0.6)

par(new=TRUE)
limy=c(0,1)
limx=c(1,TIME)
plot(rB2,lwd=1,col="gray57",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=11,cex=0.6)

par(new=TRUE)
limy=c(0,1)
limx=c(1,TIME)
plot(rB5,lwd=1,col="hotpink1",ylab="",xlab="",xlim=limx,ylim=limy,axes=FALSE,lty=1,type="b",pch=12,cex=0.6)


colors=c("red","blue","cyan","chocolate4","chartreuse4","darkgoldenrod","blueviolet","dodgerblue1","
    darkslategray4","deeppink2","gray57","hotpink1")

items<-c("R-network","A-network","R-A1","R-23","R-34","R-35","R-45","R-12","R-13","R-A4","R-B2","R-B5")

linestyles<-c(1,1,1,1,1,1,1,1,1,1,1,1)
lineweights<-c(4,4,1,1,1,1,1,1,1,1,1,1)
pchstyles<-c(4,22,3,4,5,6,7,8,9,10,11,12)


legend("bottomleft", inset=0.009, col=colors,lty=linestyles,lwd=lineweights,legend=items,pch=pchstyles,bg="
    azure2",cex=0.8)

#mean time between failure
mA1<-integrate(survivalexp,0,Inf,data$alpha[1])$value
m23<-integrate(survivalexp,0,Inf,data$alpha[2])$value
```

```
m34<-integrate(survivalexp,0,Inf,data$alpha[3])$value
m35<-integrate(survivalexp,0,Inf,data$alpha[4])$value
m45<-integrate(survivalexp,0,Inf,data$alpha[5])$value
m12<-integrate(survivalweib,0,Inf,data$alpha[6],data$m[6])$value
m13<-integrate(survivalweib,0,Inf,data$alpha[7],data$m[7])$value
mA4<-integrate(survivalweib,0,Inf,data$alpha[8],data$m[8])$value
mB2<-integrate(survivalexp,0,Inf,data$alpha[9])$value
mB5<-integrate(survivalexp,0,Inf,data$alpha[10])$value

#cost of intervention for each link
cost<-matrix(double(1),nrow=4,ncol=10)
cost[1,1]<-cA1_owner<-20*(data$owner[1])*data$mit[1]*(1-R_link[21,1])/(mA1+data$mit[1]/365)
cost[2,1]<-cA1_user<-20*(data$user[1])*data$mit[1]*(1-R_link[21,1])/(mA1+data$mit[1]/365)
cost[3,1]<-cA1_dap<-20*(data$dap[1])*data$mit[1]*(1-R_link[21,1])/(mA1+data$mit[1]/365)
cost[4,1]<-cA1_iap<-20*(data$iap[1])*data$mit[1]*(1-R_link[21,1])/(mA1+data$mit[1]/365)
cA1<-cA1_owner+cA1_user+cA1_dap+cA1_iap

cost[1,2]<-c23_owner<-20*(data$owner[2])*data$mit[2]*(1-R_link[21,2])/(mA1+data$mit[2]/365)
cost[2,2]<-c23_user<-20*(data$user[2])*data$mit[2]*(1-R_link[21,2])/(mA1+data$mit[2]/365)
cost[3,2]<-c23_dap<-20*(data$dap[2])*data$mit[2]*(1-R_link[21,2])/(mA1+data$mit[2]/365)
cost[4,2]<-c23_iap<-20*(data$iap[2])*data$mit[2]*(1-R_link[21,2])/(mA1+data$mit[2]/365)
c23<-c23_owner+c23_user+c23_dap+c23_iap

cost[1,3]<-c34_owner<-20*(data$owner[3])*data$mit[3]*(1-R_link[21,3])/(mA1+data$mit[3]/365)
cost[2,3]<-c34_user<-20*(data$user[3])*data$mit[3]*(1-R_link[21,3])/(mA1+data$mit[3]/365)
cost[3,3]<-c34_dap<-20*(data$dap[3])*data$mit[3]*(1-R_link[21,3])/(mA1+data$mit[3]/365)
cost[4,3]<-c34_iap<-20*(data$iap[3])*data$mit[3]*(1-R_link[21,3])/(mA1+data$mit[3]/365)
c34<-c34_owner+c34_user+c34_dap+c34_iap

cost[1,4]<-c35_owner<-20*(data$owner[4])*data$mit[4]*(1-R_link[21,4])/(mA1+data$mit[4]/365)
cost[2,4]<-c35_user<-20*(data$user[4])*data$mit[4]*(1-R_link[21,4])/(mA1+data$mit[4]/365)
cost[3,4]<-c35_dap<-20*(data$dap[4])*data$mit[4]*(1-R_link[21,4])/(mA1+data$mit[4]/365)
cost[4,4]<-c35_iap<-20*(data$iap[4])*data$mit[4]*(1-R_link[21,4])/(mA1+data$mit[4]/365)
c35<-c35_owner+c35_user+c35_dap+c35_iap

cost[1,5]<-c45_owner<-20*(data$owner[5])*data$mit[5]*(1-R_link[21,5])/(mA1+data$mit[5]/365)
cost[2,5]<-c45_user<-20*(data$user[5])*data$mit[5]*(1-R_link[21,5])/(mA1+data$mit[5]/365)
cost[3,5]<-c45_dap<-20*(data$dap[5])*data$mit[5]*(1-R_link[21,5])/(mA1+data$mit[5]/365)
cost[4,5]<-c45_iap<-20*(data$iap[5])*data$mit[5]*(1-R_link[21,5])/(mA1+data$mit[5]/365)
c45<-c45_owner+c45_user+c45_dap+c45_iap

cost[1,6]<-c12_owner<-20*(data$owner[6])*data$mit[6]*(1-R_link[21,6])/(mA1+data$mit[6]/365)
cost[2,6]<-c12_user<-20*(data$user[6])*data$mit[6]*(1-R_link[21,6])/(mA1+data$mit[6]/365)
cost[3,6]<-c12_dap<-20*(data$dap[6])*data$mit[6]*(1-R_link[21,6])/(mA1+data$mit[6]/365)
cost[4,6]<-c12_iap<-20*(data$iap[6])*data$mit[6]*(1-R_link[21,6])/(mA1+data$mit[6]/365)
c12<-c12_owner+c12_user+c12_dap+c12_iap

cost[1,7]<-c13_owner<-20*(data$owner[7])*data$mit[7]*(1-R_link[21,7])/(mA1+data$mit[7]/365)
cost[2,7]<-c13_user<-20*(data$user[7])*data$mit[7]*(1-R_link[21,7])/(mA1+data$mit[7]/365)
cost[3,7]<-c13_dap<-20*(data$dap[7])*data$mit[7]*(1-R_link[21,7])/(mA1+data$mit[7]/365)
cost[4,7]<-c13_iap<-20*(data$iap[7])*data$mit[7]*(1-R_link[21,7])/(mA1+data$mit[7]/365)
c13<-c13_owner+c13_user+c13_dap+c13_iap

cost[1,8]<-cA4_owner<-20*(data$owner[8])*data$mit[8]*(1-R_link[21,8])/(mA1+data$mit[8]/365)
cost[2,8]<-cA4_user<-20*(data$user[8])*data$mit[8]*(1-R_link[21,8])/(mA1+data$mit[8]/365)
cost[3,8]<-cA4_dap<-20*(data$dap[8])*data$mit[8]*(1-R_link[21,8])/(mA1+data$mit[8]/365)
cost[4,8]<-cA4_iap<-20*(data$iap[8])*data$mit[8]*(1-R_link[21,8])/(mA1+data$mit[8]/365)
cA4<-cA4_owner+cA4_user+cA4_dap+cA4_iap


cost[1,9]<-cB2_owner<-20*(data$owner[9])*data$mit[9]*(1-R_link[21,9])/(mA1+data$mit[9]/365)
cost[2,9]<-cB2_user<-20*(data$user[9])*data$mit[9]*(1-R_link[21,9])/(mA1+data$mit[9]/365)
cost[3,9]<-cB2_dap<-20*(data$dap[9])*data$mit[9]*(1-R_link[21,9])/(mA1+data$mit[9]/365)
cost[4,9]<-cB2_iap<-20*(data$iap[9])*data$mit[9]*(1-R_link[21,9])/(mA1+data$mit[9]/365)
cB2<-cB2_owner+cB2_user+cB2_dap+cB2_iap

cost[1,10]<-cB5_owner<-20*(data$owner[10])*data$mit[10]*(1-R_link[21,10])/(mA1+data$mit[10]/365)
cost[2,10]<-cB5_user<-20*(data$user[10])*data$mit[10]*(1-R_link[21,10])/(mA1+data$mit[10]/365)
cost[3,10]<-cB5_dap<-20*(data$dap[10])*data$mit[10]*(1-R_link[21,10])/(mA1+data$mit[10]/365)
cost[4,10]<-cB5_iap<-20*(data$iap[10])*data$mit[10]*(1-R_link[21,10])/(mA1+data$mit[10]/365)
cB5<-cB5_owner+cB5_user+cB5_dap+cB5_iap

totalcost<-cA1+c23+c34+c35+c45+c12+c13+cA4+cB2+cB5
costinfomration<-data.frame(cA1,c23,c34,c35,c45,c12,c13,cA4,cB2,cB5,totalcost)
cat("cost of intervention on each link and on network for 20 year \n")
print(costinfomration)


file.remove("data.csv")
file.create("data.csv")
write.table(cost, file="data.csv", sep = ",", append = TRUE,col.names = FALSE)
#cost calculation
```

Chapter 4

Mechanistic-empirical models and regression models

BRYAN T. ADEY AND NAM LETHANH

4.1 General

4.1.1 Manifest and latent processes

Models are required to predict both the ability to provide a specified level of service (LOS) and the required LOS. The processes that result in changes to either of these can be classified as gradual (manifest) or sudden (latent), depending on how the processes are monitored, the speed with which the infrastructure manager can react and the defined LOS. The classification of a change in LOS as gradual or sudden depends on the state of preparedness and the perception of the infrastructure management owner (IMO).

A gradual change in the LOS provided is one that happens in a way that there is enough time to execute an intervention so as to ensure that the infrastructure continues to provide an adequate LOS. For example, a gradual change may be the increases in the amount of travel time due to increasingly congested roads, something that may happen at a rate of 2% per year.

A sudden change in the LOS provided is one that happens in a way that there is not enough time to execute an intervention so as to ensure that the infrastructure continues to provide an adequate LOS. A sudden change may be a new law that allows the axle loads to increase by 50%.

Typical models to be used to predict either the provided or the required LOS due to gradual processes are mechanistic-empirical models, regression models, Markov models, neural network models, and Bayesian networks.

Typical models to be used to predict with either the provided or the required LOS due to sudden processes are event trees and fault trees.

4.1.2 Continuous vs. discrete states

Condition of an infrastructure object changes over time. The range of condition in which an object can change before it is referred to as something else is referred to as a "state". An adjective or a noun in front of the word "state" refers to the factor being used to classify the object. For example, a "condition state" is defined based on the physical characteristics of an object, and a "performance state" is defined based on the ability of the object to perform, or provide a level of service.

As states are man-made descriptions of the object there are almost an infinite number of possible ways that states can be defined. How states are defined will affect the illustration of the movement of the object over time through the states. This can be considered in the simplification of the mathematical modeling of the behavior of infrastructure.

4.1.3 Deterministic vs. probabilistic models

Models of the behavior of infrastructure over time can be done using models that are classified as deterministic or probabilistic. Both have advantages and disadvantages.

The use of deterministic models gives the impression that the values of the performance indicators is known with certainty at every time t. They give very precise, but perhaps not very accurate predictions.

The use of probabilistic models gives the probability of having each value of the performance indicators at every time t. They give less precise predictions than deterministic models but may be more accurate.

4.1.4 Classification of models

The models that are typically used to make prediction of the future states of infrastructure objects can be classified in different ways. Some examples are given in Table 4.1. The mechanistic-empirical and regression models will be discussed in more detail in the next sections.

Table 4.1: Types of models for manifest processes

Type of model	Definition
Analytical	has a closed form solution, i.e. the solution to the equations used to describe changes in a system can be expressed as a mathematical analytic function.
Numerical	use some sort of numerical time-stepping procedure to obtain the models behaviour over time
Mechanistic	is based on an understanding of the behavior of the elements in a the system, often developed using the laws of physics
Empirical	is based on direct observation, measurement and extensive data records, often without understanding the physical processes at work
Extrapolation	can be used to estimate a value of a variable outside a known range from values within a known range
Interpolation	can be used to estimate a value of a variable within a known range from other values within this range
Regression	is developed using regression analysis, which is a statistical approach to predicting the change in the value of a dependent variable based on the changes in the values of an independent variables
Markovian	is stochastic and where the conditional probability distribution of future states of the process (conditional on both past and present values) depends only upon the present state, not on the sequence of events that preceded it

4.2 Mechanistic-empirical models

4.2.1 Theory

A mechanistic model is one that is based on an understanding of the behavior of the elements in the system, often developed using the laws of physics. An empirical model is one that is based on direct observation, measurement and extensive data records, often without understanding the physical processes at work. A mechanistic - empirical model is a combination of both.

These types of models are often used to predict the occurrence of very specific conditions, e.g. the amount of chlorides at a specific depth in a concrete beam. Coupled with the probabilistic distributions to represent the values of the variables and parameters in the models, they can be used to estimate the probable specific conditions.

They can also be used to predict a LOS, e.g. the amount of intervention costs in a specified unit of time. Or more specifically, the probability of occurrence of each LOS within each investigated time interval can be estimated using a Monte Carlo simulation or a random sampling algorithm, given the probabilistic distributions of each parameters, and the mechanistic-empirical model.

4.2.2 Example

You are an infrastructure manager that would like to predict the evolution of a reinforced concrete bridge slab whose main deterioration process is chloride-induced corrosion of the steel reinforcement over time (More information related to such processes can be found in Elsener [2013] and Bertolini et al. [2013] so that you can plan an intervention before the bridge slab provides an inadequate LOS. You know that the chloride-induced corrosion is composed of two deterioration phases: the first phase is called initiation phase, which covers the entire time from construction to the time the first visual crack appears. In this phase the chloride starts to penetrate into concrete cover, reach the reinforcement with a sufficient concentration to start corrosion and continue until there is enough corrosion to cause cracks to be visible. The second phase of deterioration then starts and includes the continued widening of the cracks in the concrete due to the continued corrosion of the reinforcement. The entire deterioration process can be graphically illustrated in Figure 4.1. According to DuraCrete [2000], chloride penetration into reinforced concrete is a complicated process,

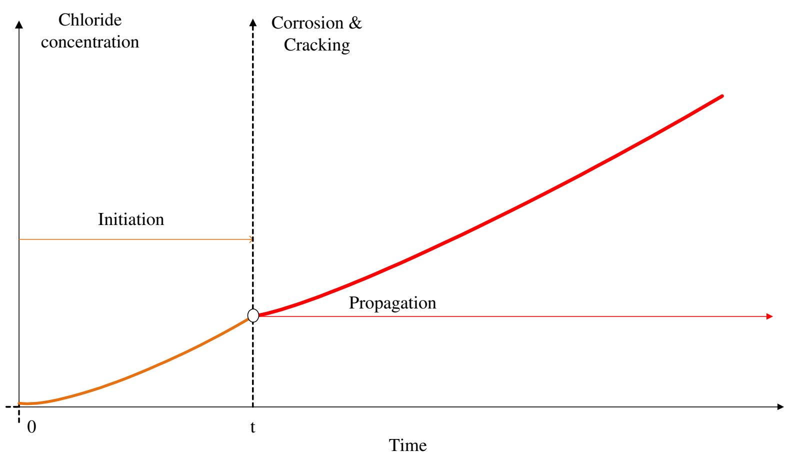


Fig. 4.1: Deterioration process of reinforced concrete due to chloride-induced corrosion

which involves inter alia ion diffusion and convection [Bertolini et al., 2013]. This complex transport mechanism can be represented by Fick's second law of diffusion.

$$\frac{\partial C_{cl}}{\partial t} = D_{cl}\frac{\partial^2 C_{cl}}{\partial x_{cl}^2} \tag{4.1}$$

Where:

C_{cl}: is the chloride ion concentration at the depth,

x_{cl}: from the surface of concrete after it exposes to chlorides till time t (here small notation cl denotes the abbreviation for chloride).

D_{cl}: is the chloride diffusion coefficient.

The solution for partial differential equation of Eq. (4.1) gives the following explicit form to calculate the chloride concentration as a function of depth $x_{cl}$and time t.

$$C_{cl}(x_{cl},t) = C_s\left(1 - erf\left(\frac{x_{cl}}{2\sqrt{D_{cl}t}}\right)\right) \tag{4.2}$$

where $erf\,()$denotes the error function[1].

The duration for the diffusion process of chloride ion to initiate corrosion, that is time t in Eq. (4.2), can be estimated by setting the value of C_{cl} to be equal to the chloride concentration. In other words, by setting the value of $C_{cl}$for each discrete condition state i, the time to arrive at that CS can be obtained by solving Eq. (4.2) with respect to time t and a certain depth of concrete cover from the re-bar.

For a bridge slab, value of variables C_{cl},C_s, and D_{cl}in the equation are can be considered as random, each one is associated with its own statistical distribution. Thus, estimation of time to arrive at certain value of chloride concentration is probabilistic in nature.

After the value of chloride concentration reaches a certain limit, corrosion starts on the reinforcing bars and after it reaches another higher limit the cracking process starts. The crack initiation phase is referred to here as the propagation phase and the width of the cracks over time can be determined by using following equations:

$$w(t) = w_0 + \beta \cdot (P(t) - P_0) \tag{4.3}$$

Where:

$w(t)$: width (mm) of crack over time,

β: the parameter that controls the propagation

w_0: width of crack when it is visible (˜0.05 mm)

P_0: the amount of loss of re-bar diameter (mm) when crack width is visible

$P(t)$: the amount of loss of re-bar diameter (mm) at t

The reinforcement loss function can be represented as:

$$P(t) = \int_0^t V_{corr} \cdot \alpha \cdot wet \cdot t dt \tag{4.4}$$

[1] Definition of error function at http://en.wikipedia.org/wiki/Error_function

Where:

V_{corr}: corrosion rate coefficient (mm/year)

wet: wet period in a year (equal to the ratio between total numbers of rainy day and 365 days)

α: pitting factor that takes non-uniform corrosion of the re-bars into consideration

Based on field tests and experiments, statistical values of models coefficients and parameters are given in Table 4.2

Table 4.2: Values of models parameters and coefficients

Parameters/ Coefficients	Units	Mean (μ)	Standard deviation (σ)	Distribution	Reference equations
C_s	$wt.-\%Cl-/binder$	2.565	0.405	Normal	Eq. (4.2)
x_{cl}	mm	20	NA	NA	Eq. (4.2)
D_{cl}	NA	100	0.5	Normal	Eq. (4.2)
w_0	mm	0.05	0.005	Normal	Eq. (4.3)
P_0	mm	5	0.5	Normal	Eq. (4.3)
β	NA	0.001	0.001	Normal	Eq. (4.3)
V_{corr}		0.012	0.006	Normal	Eq. (4.4)
wet	year	1	0.25	Normal	Eq. (4.4)
α	NA	14	4.04	Normal	Eq. (4.4)

4.2.2.1 Question

If you know that the chloride concentration C_{crit} which starts the propagation phase is $C_{cl} = C_{crit} = 0.48$, and you consider that once a crack width is 0.5 mm, that an intervention is required to ensure that an adequate LOS is provided, when do you need to plan an intervention to ensure that an adequate LOS is provided?

Hint: Construct a probabilistic deterioration model with discrete condition states for the entire chloride induced corrosion process of the bridge slab.

4.2.2.2 Answer

4.2.2.2.1 Condition states

Since the deterioration process of chloride induced corrosion is a continuous process with two distinct phases, it is useful to define ranges of condition states for each phase in order to construct the probabilistic deterioration model with discrete condition states. This is illustrated in Figure 4.2, when 5 condition states are used. Two are attributed to the initiation phase and two to the propagation phase. In Figure 4.2, five discrete condition states are used to illustrate the process. Their definitions are explained in Table 4.3.

4.2.2.2.2 Deterioration: Chloride concentration

Using Eq. (4.2), the concentration of chlorides at any time t can be deterministically estimated given the mean values of the parameters of the model. These are, however, probabilistic in nature, thus, an explicit and straightforward arithmetic calculation is not always desirable. At any time in the future, for example, the value of C_{cl} can be represented by using the mean and standard deviation. This concept is illustrated in Figure 4.3 for a normal distribution, where the mean value μ and standard deviation σ are 0.3 and 0.08, respectively

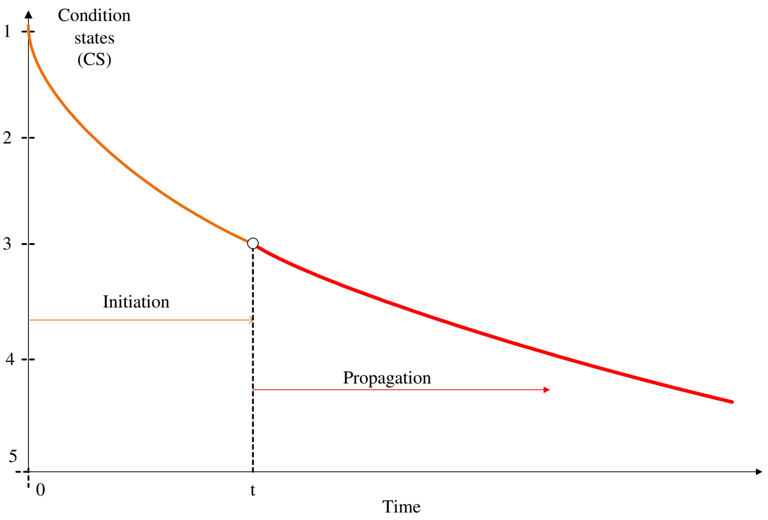


Fig. 4.2: Deterioration process with discrete condition states

Table 4.3: Definition of condition states

Phase	CS	Description	Indicator	Criteria
1	1	New/partial new	Amount of chlorides in the concrete at reinforcing bar level C_{cl}	$0 < C_{cl} \leq 0.24$
	2	Concrete contaminated		$0.24 < C_{cl} \leq 0.48$
2	3	Corrosion has initiated, no visible cracking has occurred	Width of crack (w)	$C_{cl} > 0.48, w \leq 0.25$
	4	Visible cracking has occurred		$0.25 < w \leq 0.5$
	5	Visible cracking has occurred and cover has spalled off		$w > 0.5$

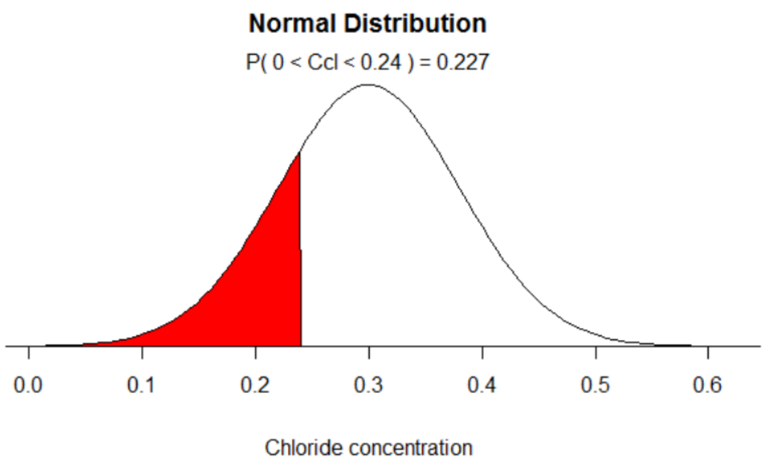


Fig. 4.3: Probability of condition state CS 1-initiation phase.

In order to determine when an intervention is required to ensure that an adequate LOS is provided, the probability of being in each condition state over time needs to be calculated.

Let assume that it is time t now, and the probability of the bridge slab being in CS 1 will be equivalent to the probability that its continuous value falls between 0 and 0.24. The equation used to estimate the red area is.

In order to determine when an intervention is required to ensure that an adequate LOS is provided, the probability of being in each condition state over time needs to be calculated.

Let assume that it is time t now, and the probability of the bridge slab being in $CS1$ will be equivalent to the probability that its continuous value falls between 0 and 0.24. The equation used to estimate the red area is.

$$P_1 = P_1^{cl} = Prob[CS = 1] = Prob[0 \leq C_{cl} \leq 0.24] = \int_0^{0.24} f(x) \cdot dx = 0.227 \quad (4.5)$$

where $f(x)$ is density function of the normal distribution.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4.6)$$

Similarly, the probability that the bridge element is in $CS2$ is given by:

$$\begin{aligned} P_2 = P_2^{cl} &= Prob[CS = 2] = Prob[0.24 \leq C_{cl} \leq 0.48] \\ &= \int_{0.24}^{0.48} f(x) \cdot dx = 0.761 \end{aligned} \quad (4.7)$$

The probability that the bridge slab is in $CS3$ is in a condition state greater than CS2 is then given by:

$$P_{>2} = P_{>2}^{cl} = 1 - Prob[CS = 2] - Prob[CS = 2] = 1 - 0.227 - 0.761 = 0.012 \quad (4.8)$$

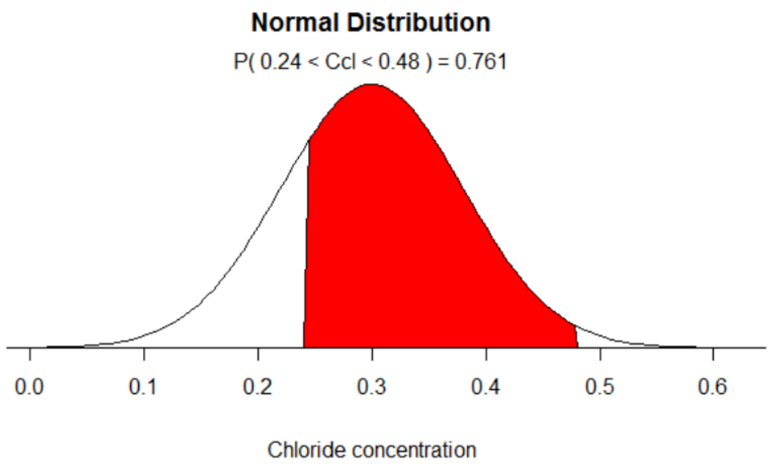


Fig. 4.4: Probability of condition state CS-2-initiation phase.

4.2.2.2.3 Deterioration: Corrosion propagation

If it is assumed that the crack width can be modeled using a normal distribution with $[\mu = 0.22, \sigma = 0.03]$ at time t (the same time that is used to compute the state probability of the initiation phase), then the probability of being in condition state 3 is given by:

$$\begin{aligned} P_3 = P_3^{cr} &= Prob[CS = 3] = Prob[w \leq 0.25] \cdot Prob[C_{cl} > 0.48] \\ &= 0.841 \cdot 0.012 = 0.010092 \end{aligned} \quad (4.9)$$

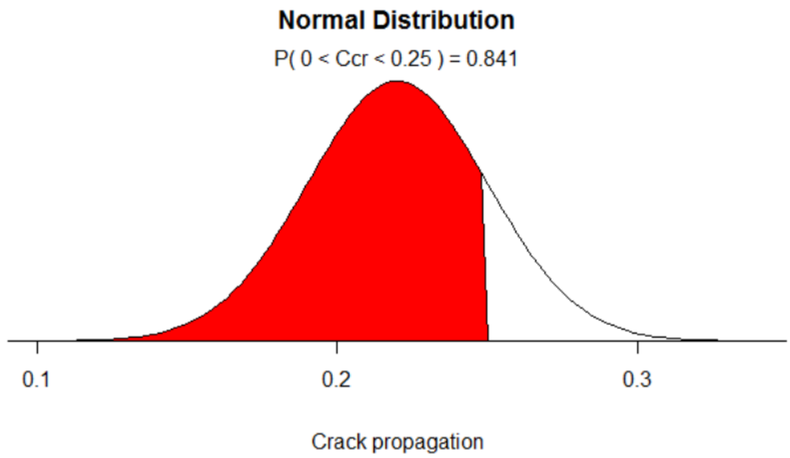


Fig. 4.5: Probability of condition state CS 3-propagation phase.

Eq. (4.9) shows a joint probability of the event that chloride concentration level has reached more than its critical level 0.48 and corrosion has progressed to a point where at least one crack of 0.05 mm has opened up. The width of crack is, however, less than 0.25 mm.

Similarly, following state transition probabilities are obtained

$$P_4 = P_4^{cr} = Prob[CS = 4] = Prob[0.25 < w \leq 0.5] \cdot Prob[C_{cl} > 0.48]$$
$$= 0.159 \cdot 0.012 = 0.001908 \quad (4.10)$$
$$P_5 = P_5^{cr} = Prob[CS = 5] = Prob[w \geq 0.5] \cdot Prob[C_{cl} > 0.48] = 0 \cdot 0.012 = 0 \quad (4.11)$$

The above steps of calculation are used for a specific case when the condition state is defined in Table 4.3. For a general case that mathematically describes the relationship and for sake of programing, following generic formulation is used.

$$P_k = \begin{cases} P_i^{cl} = \int_a^b f(x)dx \quad i = (1, ..., I) \\ P_I^{cl} \cdot P_j^{cr} \qquad where \qquad P_j^{cr} = \int_c^d g(x)dx \qquad j = (1, ..., J) \end{cases} \quad (4.12)$$

Where:

P_i^{cl}: is state probability of CS i of the initiation phase with total number CS of I

P_j^{cl}: is state probability of CS j of the propagation phase with total number CS of J

$f(x)$ and $g(x)$: are density function of the normal distribution for the initiation phase and propagation phase, respectively.

4.2.2.2.4 State distribution as proportional data

Once the mechanistic-empirical equation, and the probabilistic distributions of each parameter to be used are known than the probable value of both chloride concentration and crack width can be generated by running many simulations, e.g. 20000. Once enough points are generated at the point of time to be investigated, either the probabilistic distribution that best fits these points can be determined and from this distribution the probability of the bridge slab being in a specific condition state can be estimated, or the probability of the bridge being in each condition state can be estimated directly from the number of data points, e.g. 10000/20000 are in the specified range.

In previous steps of calculation, the example is shown only for a specific time t given the mean μ and the standard deviation σ for each process. However, these values have to be estimated at each time step given the probabilistic distribution of models parameters. Results of estimation are the state probabilities over a pre-defined number of years. This pool of state probabilities is proportional data and state probabilities of the bridge in previous year is not related to the following year.

4.2.2.2.5 Deterioration prediction

Once the mechanistic-empirical equation, and the probabilistic distributions of each parameter to be used are known than the probable value of both chloride concentration and crack width can be generated by running many simulations, e.g. 20’000. Once enough points are generated at the point of time to be investigated, either the probabilistic distribution that best fits these points can be determined and from this distribution the probability of the bridge slab being in a specific condition state can be estimated, or the probability of the bridge being in each condition state can be estimated directly from the number of data points, e.g. 10’000 / 20’000 are in the specified range.

In previous steps of calculation, the example is shown only for a specific time t given the mean μ and the standard deviation σ for each process. However, these values have to be estimated at each time step given the probabilistic distribution of models parameters (Table 4.2).

Results of the estimation are the state probability over a pre-defined number of years. In this example, calculation is done for 60 years. At each year, the probabilistic value of chloride concentration C_{cl} and the width of crack X are re-computed using a pool of 20’000 samples that were randomly generated using Monte Carlo simulation. Then, the state probability that the bridge slab arrives at each discrete CS was computed for every year. These are shown in graphically in Figure 4.6.

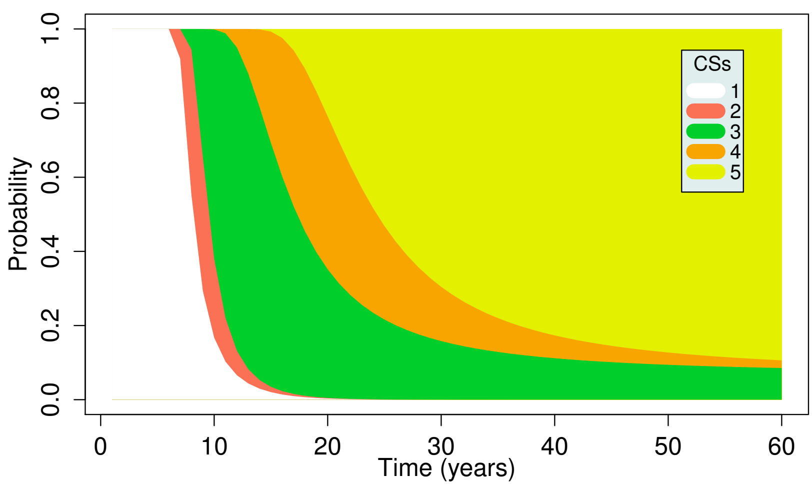


Fig. 4.6: Condition state distribution over 60 years

4.2.2.2.6 Results

As can be seen from Figure 4.6, in the first 7 years after the construction of the bridge slab, the chloride concentration is not expected to reach a value of 0.25. In other words, the bridge slab is expected to remain in CS 1. Then, the probability of the bridge slab being in CS 1 decreases, in the mean-time, of course, the probability that the bridge slab is in CS 2 increases. After 10 years, there is almost a 50% chance that the bridge slab will be in CS 2.

Between the 7^{th} and 10^{th} years, there is a non-zero probability that the critical chloride concentration has been reached (0.48), i.e. the propagation phase has started. This explains the appearance of CS 3, CS 4 and CS 5 in the chart.

After 20 years, there is a high probability that the critical chloride concentration has been reached and the propagation phase has started. There is also approximately a 40% chance that the surface of the slab will have at least one crack with a width between 0.00 mm and 0.25 mm. More or less the same chance that the crack width will be between 0.25 mm and 0.5 mm. There is, however, also a 20% chance that the crack width will be greater than 0.5 mm, and that an intervention will have to be executed to ensure that an adequate LOS is provided.

So, since I know that the chloride concentration C_{crit} that starts the propagation phase is $C_{cl} = C_{crit} = 0.48$, and if I consider that once a crack width is 0.5 mm that an intervention is required to ensure that an adequate LOS is provided, I will plan an intervention in year 15, when the probability of being in CS5 is still, according to my calculations is under 0.01 (Table 4.4). As I know, however, that CS5 does not necessarily mean that an inadequate LOS is provided,

Table 4.4: Evolution of state probability over 30 years

year	Condition states				
	1	2	3	4	5
1	1	0	0	0	0
2	1	0	0	0	0
3	1	0	0	0	0
4	1	0	0	0	0
5	1	0	0	0	0
6	1	0	0	0	0
7	0.9174	0.0826	0	0	0
8	0.5520	0.3907	0.0574	0	0
9	0.2977	0.3571	0.3452	0	0
10	0.1718	0.2123	0.6149	0.001	0
11	0.1065	0.1181	0.7640	0.0114	0
12	0.0693	0.0672	0.8153	0.0482	0
13	0.0466	0.0396	0.7965	0.1173	0.0001
14	0.0319	0.0242	0.7360	0.2069	0.0011
15	0.0221	0.0151	0.6584	0.2977	0.0068
16	0.0154	0.0097	0.5795	0.3720	0.0234
17	0.0107	0.0063	0.5073	0.4195	0.0561
18	0.0075	0.0041	0.4447	0.4387	0.1050
19	0.0052	0.0027	0.3918	0.4344	0.1659
20	0.0036	0.0018	0.3477	0.4139	0.2330
21	0.0025	0.0012	0.3110	0.3842	0.3011
22	0.0017	0.0008	0.2805	0.3506	0.3664
23	0.0012	0.0005	0.2551	0.3164	0.4268
24	0.0008	0.0004	0.2337	0.2839	0.4812
25	0.0006	0.0002	0.2157	0.2539	0.5296
26	0.0004	0.0002	0.2005	0.2269	0.5721
27	0.0003	0.0001	0.1874	0.2030	0.6093
28	0.0002	0.0001	0.1762	0.1819	0.6417
29	0.0001	0	0.1664	0.1634	0.6700
30	0.0001	0	0.1580	0.1472	0.6947

I can also treat it as a warning that an inadequate LOS is coming if I don't do anything, but that I have a bit of time before this will happen. In this case, I may plan the intervention to be in year 20, when the probability of being in CS5 is 0.233. If I decide to wait until there is 50% of being in CS5 than I will plan the intervention for either year 24 or 25. Whether or not this is a good decision depends on the probability of an inadequate LOS occurring once the slab reaches CS5 and the consequences if the adequate LOS happens, i.e. the risk associated with being in CS5.

The steps of numerical calculation can be tracked by reading the R script for this example provided at the Appendix 4.A.

4.3 Regression models

4.3.1 Theory

Regression models represent empirical relationships between multiple parameters and are developed using regression analysis. They are used to predict the change in the value of a dependent variable based on the changes in the values of independent variables. Each variable is described in terms of its mean and variance. The simplest form is a linear regression model as shown in equation 4.13. Other forms are possible.

$$Y_i = \alpha + \beta \cdot X_i + \varepsilon_i \tag{4.13}$$

Where:

- Y_i: is the dependent variable or output (e.g. state probability of *CS i*) that is observable
- X_i: is the characteristic or explanatory variable that is also observable (e.g. time, traffic volume, ambient temperature).
- α, β: are parameters that need to be estimated.
- ε: is the prediction error that can be assumed to follow a certain distribution (e.g. normal or log-normal distributions)

The mean or expected value of Y_i for each value of input X_i is

$$EY_i = \hat{\alpha} + \hat{\beta} \cdot X_i + \varepsilon_i \tag{4.14}$$

The overhead mark ˆ indicates the expected values.

The purpose of performing a regression analysis is to predict the expected values of the regression parameters $\hat{\alpha}, \hat{\beta}$ that minimize following objective function.

$$s = \sum_{i=1}^{N} \left[Y_i - \hat{Y}_i\right]^2 = \sum_{i=1}^{N} \left[Y_i - \hat{\alpha} - \hat{\beta} \cdot X_i\right]^2 \tag{4.15}$$

In words, the objective of Eq. (4.15) is to minimize the sum of the prediction errors. Here N is the total of data points.

In order to obtain the values of the regression parameters, one way of analytically solving Eq. (4.15) with respect to two variables $\hat{\alpha}, \hat{\beta}$ (hereafter refer as a vector set $\hat{\theta}$). This can be done by taking the first derivative of Eq. (4.15) and set it to be equal to 0, and then solve for $\hat{\theta}$.

$$\frac{\partial s}{\partial \hat{\theta}} = 0 \tag{4.16}$$

For the value of $\hat{\alpha}$

$$\frac{\partial s}{\partial \hat{\alpha}} = 0 \tag{4.17-a}$$

$$\Leftrightarrow \frac{\bar{Y}^2 - 2\cdot\bar{Y}(\hat{\alpha}+\hat{\beta}\cdot\bar{X}) + (\hat{\alpha}+\hat{\beta}\cdot\bar{X})^2}{\partial \hat{\alpha}} = 0 \tag{4.17-b}$$

$$\Leftrightarrow -2\cdot\bar{Y} + 2\cdot\hat{\alpha} + 2\cdot\hat{\beta}\cdot\bar{X} = 0 \tag{4.17-c}$$

$$\Leftrightarrow \hat{\alpha} = \bar{Y} - \hat{\beta}\cdot\bar{X} \tag{4.17-d}$$

Here, the bar $^-$ indicates the value of $\frac{\sum_{i=1}^{N} X_i}{N}$ and $\frac{\sum_{i=1}^{N} Y_i}{N}$, which are the mean values.

Similarly, for the value of $\hat{\beta}$

$$\frac{\partial s}{\partial \hat{\beta}} = 0 \tag{4.18-a}$$

$$\Leftrightarrow \frac{\bar{Y}^2 - 2\cdot\bar{Y}(\hat{\alpha}+\hat{\beta}\cdot\bar{X}) + (\hat{\alpha}+\hat{\beta}\cdot\bar{X})^2}{\partial \hat{\beta}} = 0 \tag{4.18-b}$$

$$\Leftrightarrow \beta\cdot\sum_{i=1}^{N} X_i^2 - \sum_{i=1}^{N} X_i\cdot Y_i + \alpha\sum_{i=1}^{N} X_i = 0 \tag{4.18-c}$$

$$\Leftrightarrow \frac{\sum_{i=1}^{N} X_i\cdot Y_i - n\cdot\bar{X}\cdot\bar{Y}}{\sum_{i=1}^{N} X_i^2 - n\cdot\bar{X}} = \frac{\sum_{i=1}^{N}(X_i-\bar{X})\cdot(Y_i-\bar{Y})}{\sum_{i=1}^{N}(X_i-\bar{X})^2} \tag{4.18-d}$$

Following figure graphically represents the regression equation and the calculation.

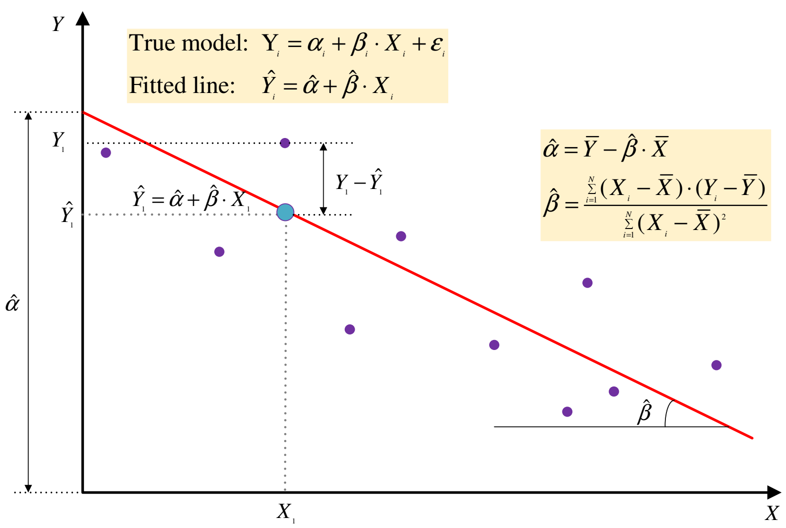


Fig. 4.7: Deriving linear regression coefficients

The goodness of fit of the regression line can be measured by the coefficient R^2, which measures the proportion of total variation about the mean $\bar{Y}$, which is explained by regression:

$$R^2 = \frac{\sum_{i=1}^{N}(\hat{Y}_i-\bar{Y})^2}{\sum_{i=1}^{N}(Y_i-\bar{Y})^2} = \frac{\beta^2\cdot\sum_{i=1}^{N}(X_i-\bar{X})^2}{\sum_{i=1}^{N}(\hat{Y}_i-\bar{Y})^2} \tag{4.19}$$

The prediction errors can be followed an independent type of distribution. In many practical situation, normal distribution is often used. In case of normal distribution $N(\mu,\sigma)$, its mean and standard deviation are defined as:

$$\mu_{\varepsilon} = Y_i - \hat{Y}_i \tag{4.20}$$

$$\sigma_{\varepsilon} = \sqrt{\frac{\sum_{i=1}^{N}(Y_i - Y_i)^2}{N-k-1}} \tag{4.21}$$

Where k is number of independent variable, which is equal to 1 in this example (as there is only one input X).

The above example is for the case of simple linear regression with only one independent variable. In many practical situations, linear regression models can be built for multiple variables, which often takes following form.

$$Y = \sum_{i=0}^{N} \beta^k \cdot X_i^k \tag{4.22}$$

Where $k = 1, ..., K$ is a vector set of numbers of independent variables.

4.3.2 Example

Assuming that you have many reinforced concrete slabs, as the one discussed in section 2.2 and have measured chloride concentrations and crack widths over the years, you may have the plots shown in Figure 4.8 and Figure 4.9[2]. (This data will be distributed in a .csv file in addition to this text).

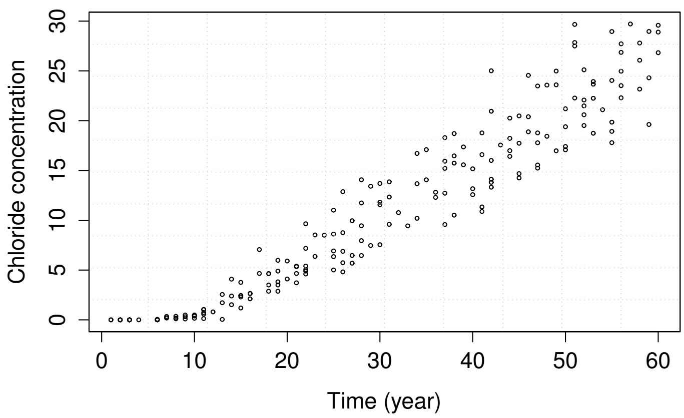


Fig. 4.8: Measured chloride concentrations in similar deck slabs over time

[2] Data used for this example is selected from the pool of randomly generated total data used in the previous example, taking into consideration of the mean and standard deviation at each time t.

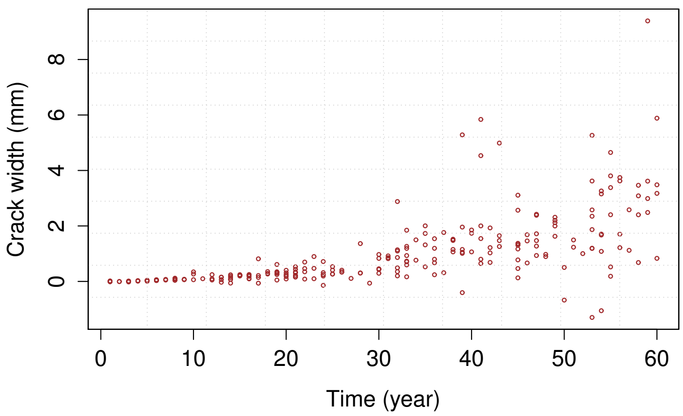


Fig. 4.9: Measured crack widths in similar deck slabs over time

4.3.2.1 Question

If you know that the chloride concentration C_{crit} which starts the propagation phase is $C_{cl} = C_{crit} = 0.48$, as in the previous example, and you consider that the once a crack width of 0.5 mm is reached, that an intervention is required to ensure that an adequate LOS is provided, when do you need to plan an intervention to ensure that an adequate LOS is provided? (Use linear regression analysis with the values of chloride concentration and crack width provided, estimate the functional form of deterioration).

4.3.2.2 Answer

As deterioration of the bridge slab is seen as happening in two different phases, separate models are developed using regression analysis for both phases; one for the initiation phase with chloride concentration and other for the propagation phase with the development of crack.

4.3.2.2.1 Condition states

The condition states are defined exactly as in Table 4.3.

4.3.2.2.2 Deterioration: Chloride concentration

The form of regression equation for chloride concentration, when it is assumed to be linear is:

$$Y_i^{cl} = \alpha_{cl} + \beta_{cl} \cdot X_i^{cl} + \varepsilon_i^{cl} \tag{4.23}$$

Where:

Y_i^{cl}: is the value of chloride concentration at year i.

X_i^{cl}: is the year i.

α_{cl}, β_{cl}: are regression parameters that need to be estimated.

ε_i^{cl}: is the prediction errors following a normal distribution

The expected values of regression parameters can be obtained using the R code in the Appendix 4.C on data created using the Appendix 4.B. The results are in listing 4.1.

```
Call:
lm(formula = mdatacl[, 3] ~ mdatacl[, 1])

Residuals:
    Min      1Q  Median      3Q     Max
-6.6339 -1.8083 -0.1265  1.7615  8.1523

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  -3.72861    0.38768  -9.618   <2e-16 ***
mdatacl[, 1]  0.49180    0.01065  46.164   <2e-16 ***
---
Signif. codes:  0   ***   0.001   **   0.01   *   0.05   .   0.1   1

Residual standard error: 2.741 on 198 degrees of freedom
Multiple R-squared:  0.915, Adjusted R-squared:  0.9146
F-statistic:  2131 on 1 and 198 DF,  p-value: < 2.2e-16
```

Listing 4.1: Results-Chloride concentration

The value of the regression parameters are: $\alpha_{cl} = -3.40275$ and $\beta_{cl} = 0.47975$. The value of R^2 is greater than 0.8959, which is nearly 10% far from 1. This means that the fitness of the linear model to this data set infers a small variation. This can also be seen from Figure 4.10, which shows fitted line of chloride concentration at the surface of the reinforcement as time goes by. The equation to be used for estimating the chloride concentration over time is then

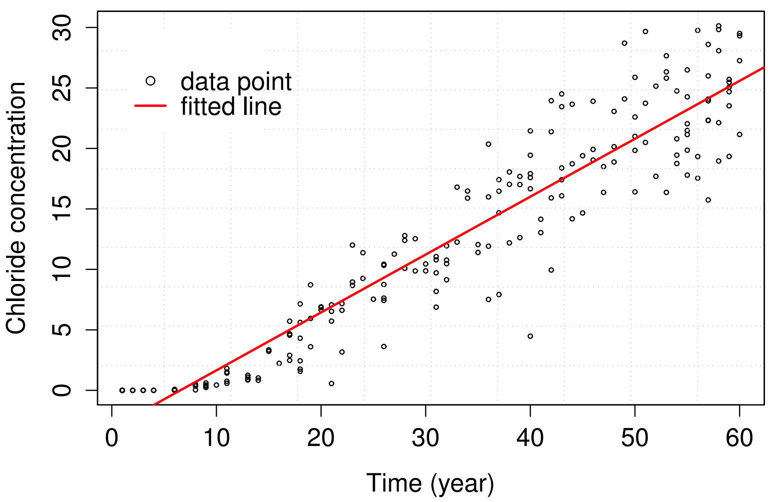


Fig. 4.10: Development of chloride concentration

$$Y_i^{cl} = -3.40275 + 0.47975 \cdot X_i^{cl} \tag{4.24}$$

Noted that the value of chloride concentration cannot be negative, therefore, it is necessary to consider a dummy variable to control the results of calculation

$$Y_i^{cl} = (-3.40275 + 0.47975 \cdot X_i^{cl}) \cdot \delta_i^{cl} \tag{4.25}$$

$$where \;\; \delta_i^{cl} = \begin{cases} 0 \text{ when } \mathrm{X}_i^{cl} < 7 \\ 1 \text{ when } \mathrm{X}_i^{cl} \geq 7 \end{cases} \tag{4.26}$$

Where 7 is the integer value closest to the ratio 3.40275/0.4975. An integer value is used here as the years are used as discrete units of time.

This result is obtained by running the R code provided in the Appendix 4.C on data created using the Appendix 4.B.

4.3.2.2.3 Deterioration: Crack propagation

To estimate the development of crack width using a regression model, a step similar to the one in section 4.3.2.2.2 is used. Following results (listing 4.2) are obtained (using the R code provided in the Appendix 4.C).

```
Call:
lm(formula = mdatacr[, 3] ~ mdatacr[, 1])

Residuals:
    Min      1Q  Median      3Q     Max 
-3.0200 -0.4295 -0.0560  0.2348  8.3059 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  -0.526209   0.129107  -4.076 6.18e-05 ***
mdatacr[, 1]  0.049676   0.003849  12.906  < 2e-16 ***
---
Signif. codes:  0    ***    0.001    **    0.01    *    0.05    .    0.1         1

Residual standard error: 1.004 on 248 degrees of freedom
Multiple R-squared:  0.4018,  Adjusted R-squared:  0.3994 
F-statistic: 166.6 on 1 and 248 DF,  p-value: < 2.2e-16
```

Listing 4.2: Results-Crack

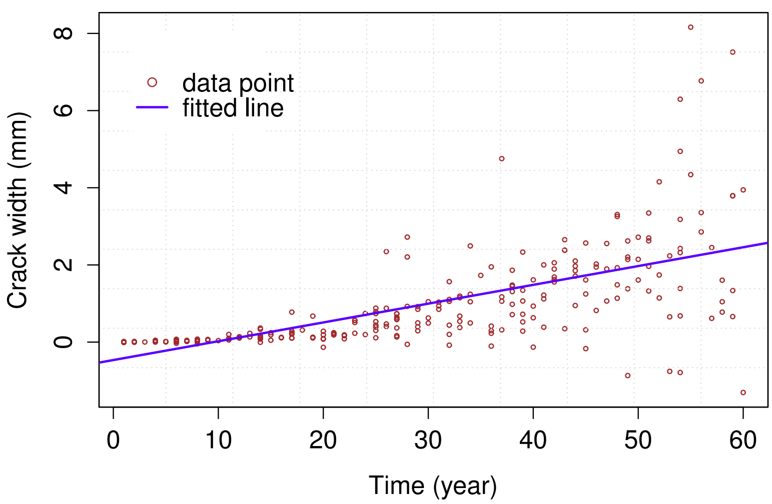


Fig. 4.11: Development of crack width

The equation to be used to predict crack width is then:

$$Y_i^{cr} = (-0.500755 + 0.053150 \cdot X_i^{cr}) \cdot \delta_i^{cr} \tag{4.27}$$

$$where \quad \delta_i^{cr} = \begin{cases} 0 & \text{when } \mathrm{X}_i^{cr} < 9 \\ 1 & \text{when } \mathrm{X}_i^{cr} \geq 9 \end{cases} \tag{4.28}$$

Where 9 is the integer value closest to the ratio 0.500755/0.053150. An integer value is used here as the years are used as discrete units of time.

4.3.2.2.4 Deterioration prediction

Using the regression equations given above, the predicted deterioration is as shown in Figure 4.12. The curve consists of three pieces. The first piece represents the period of time from new construction until the chloride concentration at the reinforcement level is thought to be equivalent to that determined by the regression equation. The second piece represents the period of time where the chloride concentration at the reinforcement can be represented by the regression equation to the point where corrosion has started to an extent to cause at least one crack wider than 0.5 mm, and the third piece shows the propagation phase, i.e. the extension of cracking once at least one crack is wider than 0.5 mm. This is represented by the second regression equation.

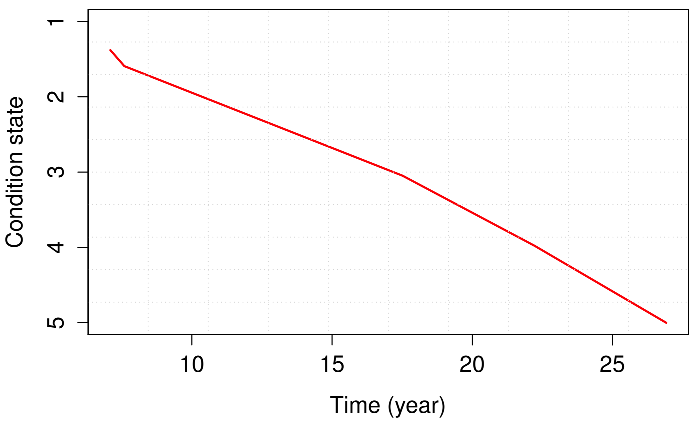


Fig. 4.12: Development of crack width

4.3.2.2.5 Results

As can be seen from the deterioration curve estimated by using regression model, by the time of 17 years, the chloride concentration certainly reach to a level that higher than 0.48. It means the bridge slab will enter CS 3 and the second phase of deterioration will start, which initiates cracking. Since the time to the worst condition state (CS 5) from the time propagation phase starts will be about 10 years, meaning that the total duration of deterioration from CS1 to CS 5 is about 27 years.

So, since I know that the chloride concentration C_{crit} that starts the propagation phase is $C_{cl} = C_{crit} = 0.48$, and if I consider that once a crack width is 0.5 mm that an intervention is required to ensure that an adequate LOS is provided, I will plan an intervention before year 27.

4.4 Comparison mechanistic-empirical models vs. regression models

Using both models, and planning an intervention when there is a 50% chance that the bridge slab will be in CS5 are the same (27 years).

Using the mechanistic model, in the way we did we could also determine the probability of the bridge slab being in CS5 in each time interval through-out the investigated time period. This made it possible to set other criteria on when to intervene, e.g. the probability of being in CS5 is negligible, or less than 20%. This was not possible using the regression model the way we did. If it was desired to take into consideration risk, using the regression analysis one would need to arbitrarily pick an earlier point in time to plan the intervention, e.g. in year 20 and not in year 25.

The use of the mechanistic-empirical model requires a lot of information on the specific values of the parameters to be used in the model, estimates of the uncertainty of the model, and the model. It is suitable in situations where there is not a lot of data on the performance of the deck slabs, and it is feasible to estimate the distributions to be used to represent the values of the parameters, develop the model, and estimate model uncertainty.

The use of the regression model requires a lot of information on the performance of similar deck slabs. It is suitable in situations where this information exists.

4.5 Assignments

4.5.1 Problem A - Erosion of concrete cover of waste water tanks

In a waste water treatment plant, reinforced concrete tanks are used to store the waste water before going through the cleaning process. Waste water is stored in the tanks and various chemical substances are poured into the tanks for the purposes of disinfection. The cement deteriorates due to this exposure.

The infrastructure manager is interested in predicting the condition of the tanks so that the maintenance interventions can be better planned over the next 30 years. She realizes that these can be estimated using the results of past inspections performed at 2 years intervals (Table 4.5). The first inspection time was executed 1 year after the begin use of the tanks. It is a general rule of thumb that the erosion should not be more than 12 mm.

Table 4.5: Inspection data

Tank	Depth of erosion (m)				
	inspection 1	inspection 2	inspection 3	inspection 4	inspection 5
1	0.2	1.1	2.4	3.8	4.9
2	0.5	0.9	1.7	3.5	5.2
3	2	2.7	3.2	4.5	5.5
4	1.5	3.1	3.6	5.5	6.3
5	1.1	1.8	1.9	3.1	4.9
6	0.7	1.5	2	3.8	5.1
7	0.8	0.9	1.8	4.2	5.4
8	1	1.9	2.6	3.9	5
9	0.9	1.7	3.6	4.3	7.1
10	0.8	0.9	1.6	4.1	6.7
11	0.5	1.8	2.3	3.7	5.8
12	0.6	1.2	2.2	3.6	6

She also realizes that this can be done using by constructing and testing tank walls in the laboratory. After having done this her engineers have come up with the following mechanistic deterioration model.

$$f(t+1) = f(t) \cdot e^{\beta} \tag{4.29}$$

Where:

$f(t+1)$: is the depth of corrosion.

β: is the deterioration coefficient taking its value of 0.1.

It is assumed that $f(0) = 0$ and $f(1) = 1$[3].

4.5.2 Question A1

What is the condition of each object in 30 years if the mechanistic-empirical model is used?

4.5.3 Question A2

What is the condition of each object in 30 years if predictions are made using a regression analysis?

4.5.4 Answer A1

Using Eq. (4.29), the evolution of erosion over 30 years is shown in Table 4.6 and Figure 4.13.

Table 4.6: Evolution of erosion (mechanistic-empirical model)

Time (years)	Erosion (mm)	Time (years)	Erosion (mm)	Time (years)	Erosion (mm)
1	1	11	2.7183	21	7.3891
2	1.1052	12	3.0042	22	8.1662
3	1.2214	13	3.3201	23	9.0250
4	1.3499	14	3.6693	24	9.9742
5	1.4918	15	4.0552	25	11.0232
6	1.6487	16	4.4817	26	12.1825
7	1.8221	17	4.9530	27	13.4637
8	2.0138	18	5.4739	28	14.8797
9	2.2255	19	6.0496	29	16.4446
10	2.4596	20	6.6859	30	18.1741

By the year 26, if the infrastructure manager has not already executed an intervention, the depth of erosion will reach 12 mm, which is beyond the critical level. In order to prevent the re-bars of tanks from the onset of corrosion, an intervention (e.g. renewal of cover layer of the tanks walls) should be executed before year 26.

[3] This assumption is used to prevent the value of function f to be equal to 0 at year t=0

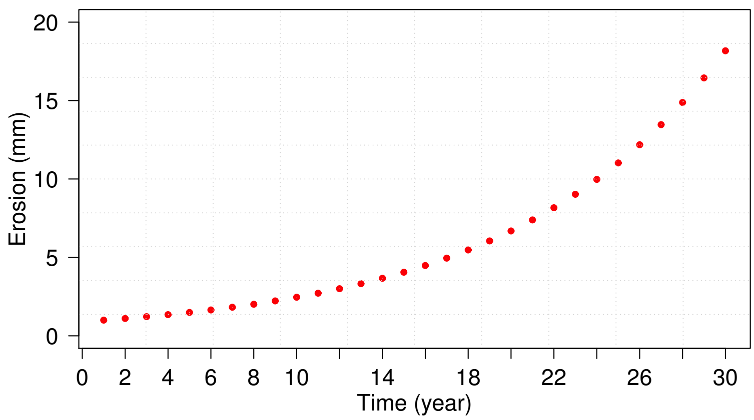


Fig. 4.13: Evolution of erosion mechanistic-empirical model

4.5.5 Answer A2

Assuming that the form of the regression is linear, the general regression equation is:

$$Y_i^{ero} = \alpha_{ero} + \beta_{ero} \cdot X_i^{ero} + \varepsilon_i^{ero} \tag{4.30}$$

Where:

Y_i^{ero}: is the value of erosion at year i.

X_i^{ero}: is the year i.

$\alpha_{ero}, \beta_{ero}$: are regression parameters that need to be estimated.

ε_i^{ero}: is the prediction errors following a normal distribution

Using the linear regression analysis (refer to section 4.3), and in the Appendix 4.D, following results are obtained (listing 4.3).

```
Call:
lm(formula = mdata[, 3] ~ mdata[, 1])

Residuals:
    Min      1Q  Median      3Q     Max 
-1.3223 -0.3891  0.0363  0.2446  1.8109 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.03630    0.12697  -0.286    0.776    
mdata[, 1]   0.59171    0.02421  24.438   <2e-16 ***
---
Signif. codes:  0    ***    0.001    **    0.01    *    0.05    .    0.1         1

Residual standard error: 0.6542 on 70 degrees of freedom
Multiple R-squared:  0.8951,  Adjusted R-squared:  0.8936 
F-statistic: 597.2 on 1 and 70 DF,  p-value: < 2.2e-16
```

Listing 4.3: Results

$$Y_i^{ero} = -0.03630 + 0.59171 \cdot X_i^{ero} \tag{4.31}$$

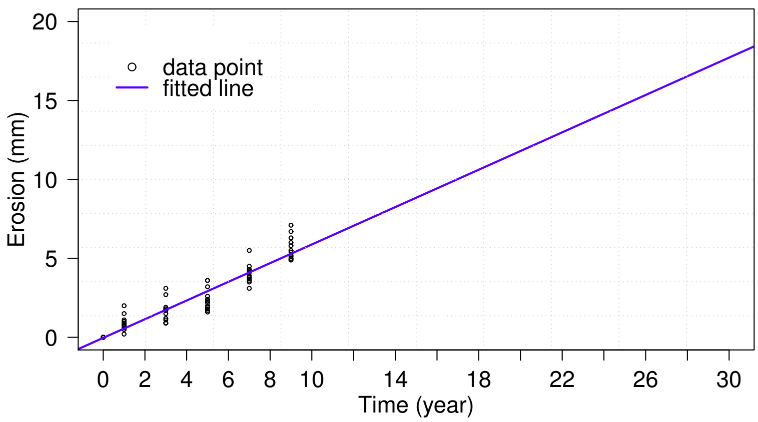


Fig. 4.14: Evolution of erosion - linear regression model

The fitted line plotted from Eq. (4.31) is shown in Figure 4.14. When the level of erosion becomes 12 mm, it corresponds to 20 years. If the infrastructure manager relies solely on this analysis she will plan the intervention to be executed before year 20.

4.5.6 Problem B - Bridge wear out

An infrastructure manager wants to make a work program for a bridge, i.e. when an intervention needs to be executed and what type of intervention. He normally models deterioration using the following mechanistic-empirical model:

$$I = \exp(\lambda \cdot t) - 1 \tag{4.32}$$

Where is $\lambda$is the deterioration coefficient, which is given by:

$$\lambda = \frac{\ln(2)}{Tc} \tag{4.33}$$

In Eq. (4.32), Tc represents the average life time of the structure elements of the bridge. It is modeled using a normal distribution with mean and standard deviation equal to 60 years and 7 years, respectively. A detail description of the model can be referred to the work of Brodsky et al. [2006]. The manager divides condition into five discrete states given in Table 4.7. The wear out of the bridge has its value from 0 to 1 as defined in Eq. (4.33).

Table 4.7: Definition of condition states

Condition states	Description	Corresponding values of I
1	New/Partially new	$I < 0.2$
2	Good	$0.2 \leq I < 0.4$
3	Average	$0.4 \leq I < 0.6$
4	Bad	$0.6 \leq I < 0.8$
5	Very bad	$I \geq 0.8$

4.5.7 Question

What is the probabilistic distribution of discrete condition states over the next 50 years for the bridge? What does this mean?

4.5.8 Answer

As the value of Tc is probabilistic in nature, the value of I is also probabilistic. A Monte Carlo simulation with 20'000 iterations is used to estimate the probability of I having values that correspond with each of the condition states over time. The results are shown in Figure 4.15. The R code in the Appendix 4.E is used to estimate the probability of the bridge being in each condition state over 50 years. As can be seen from the figure, in the first 10 years after the

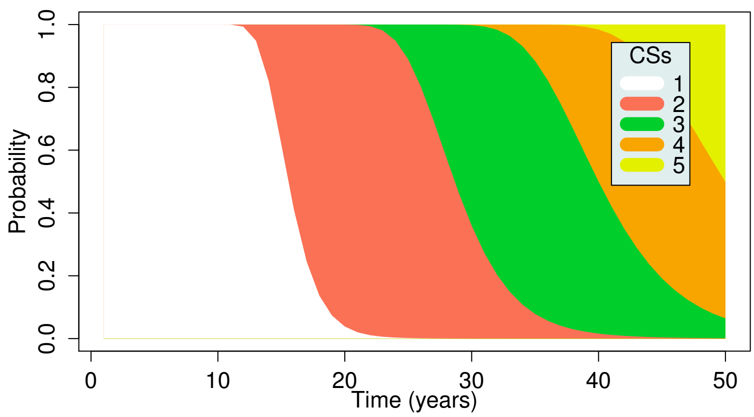


Fig. 4.15: Condition state distribution over 50 years

construction of the bridge, the bridge will remains in CS1. Then, the probability that the bridge is in CS1 decreases and the probability that it will be in CS2 increases. After 30 years, there is a 50% probability that the bridge will be in CS2.

After 30 years, there is a small probability that the bridge will be in CS3, and after 40 years, a 50% probability that the bridge will be in CS 4. The speed of degradation increases rapidly after 30 years. By year 50, the probability that the bridge is in CS 5 is more than 50%.

Based on this information the infrastructure manager might consider executing an intervention before year 40 to avoid an inadequate LOS from occurring. The steps of numerical calculation can be tracked by reading the R script for this example provided at the the Appendix 4.E.

References

Luca Bertolini, Bernhard Elsener, Pietro Pedeferri, Elena Redaelli, and Rob B. Polder. *Corrosion of Steel in Concrete: Prevention, Diagnosis, Repair*. Wiley-VCH Verlag GmbH & Co. KGaA, 2013. ISBN 9783527651696. doi: 10.1002/9783527651696.fmatter. URL http://dx.doi.org/10.1002/9783527651696.fmatter.

G. Brodsky, R. Muzykin, E. S. Brodskaia, Yu. A. Ponomarev, Yu. A. Yenyutin, and M. S. Vlasova. Analysis of parameters of structure deterioration models within the moscow bridge management system. *Structure and Infrastructure Engineering*, 2(1):13–21, 2006. doi: 10.1080/15732470500030950. URL http://www.tandfonline.com/doi/abs/10.1080/15732470500030950.

DuraCrete. *Statistical Quantification of the Variables in the Limit State Functions*. European Union, 2000.

Bernhard Elsener. *Corrosion of Steel in Concrete: Materials Science and Technology*. Wiley, 2013.

4.A R code for the mechanistic-empirical model-concrete bridge slab

```
#This program is coded by Nam Lethanh for use in the class IMP
#Purpose: To calculate the Markov transition probability based on Mechanistic-Empirical model of DuraCrete 2000
#.....STEP 1: Define model parameter
#Define the number of condition state
I=3 #number of CS related to chloride initiation, the last condition state related to failure state
J=3 #number of CS related to crack propagation
T=60 #years of investigation
limit_chrolide=0.48 #limit concentration of chloride
limit_crack=0.5 #limit crack width
a=seq(from=0, to=limit_chrolide,length.out=I) #range value of condition state for chloride
b=seq(from=0, to=limit_crack,length.out=J) #range value of condition state for crack
d=20 # distance from the concrete surface to rebar (in mm)
N=20000 #Maximum number used for random sampling with Monte Carlo or sampling techniques
mu_Dc=100 #mean value of diffusion coefficient Dc
sd_Dc=0.5 #standard deviation of the value of diffusion coefficient Dc
mu_Cs=2.565 #mean value of chloride concentration at the surface
sd_Cs=0.405 #standard deviation of the chloride concentration at the surface
mu_w0<-0.05  #: crack width when it is visible 0.05
sd_w0<-0.005
mu_beta<- 0.01   #: propagation controlling parameter
sd_beta<-0.001
mu_Px0 <- 5    #: attack penetration mm at when w0 is visible
sd_Px0<-0.5
mu_V0 <-  0.012   #: mean corrosion rate; sd=0.006 / mm/year
sd_V0<-0.006
mu_wet <-  1  #: wetness period (sum of all raining days in a year)  sd=0.25
sd_wet<-0.25
mu_alpha <- 14 #9.28 #: sd =4.04
sd_alpha<-4.04
epsilon<-0.005 #this is used to prevent value of simulated standard deviation to be less than 0.005, which gives a distribution with value less than 10^(-16), and in R, if value is less than 10^(-16), the PC cannot understand, basically when that situation happens the result returns to 0, which is not true.
#-----------------------------------------------
#.....STEP 2: Define functions to be used
#define error function
erf=function(x) 2 * pnorm(x * sqrt(2)) - 1
#Model parameters
#define the function to compute the time
timetocor<-function(d,Dc,Ccl,Cs){
  ((d^2)/(4*Dc))*(1/erf(1-Ccl/Cs))^(2)
}
#define the function of chloride concentration in initiation phase
Ccl<-function(Cs,x,Dc,t){
  Cs*(1-erf(x/(2*sqrt(Dc*t))))
}
#....Cs is the chloride concentration at the surface
#....x is the distance from surface to rebar (mm)
#....Dc is diffusion coefficient, which is probabilistic
#....t is the time (e.g. in year)
#---------------------------------------

#---define the probability in a certain range with lower bound and upper bound (for normal distribution)
pro=function(down,up,mu,sigmal){
  integrate(dnorm, down, up, mean = mu, sd = sigmal)
}
#define the dimension
pi_cl<-matrix(double(1),nrow=T,ncol=I) #state probability for initiation stage
pi_cr<-matrix(double(1),nrow=T,ncol=J) #state probability for propagation stage
mu_cl<-matrix(double(1),nrow=T,ncol=1) #mean value of chloride at time t
mu_cr<-matrix(double(1),nrow=T,ncol=1) #mean value of crack width at time t
sigmal_cl<-matrix(double(1),nrow=T,ncol=1) # standard deviation of chloride at time t
sigmal_cr<-matrix(double(1),nrow=T,ncol=1) # standard deviation of crack at time t
pi=matrix(double(1),nrow=T,ncol=(I+J-1)) # state probability of the entire process -1 mean we consider the last state I as failure state, therefore, the total state of the system is I+J-1
#
Cclt<-matrix(double(1),nrow=T,ncol=N)
#functions toi be used in propagation phase

# .....w(x) : crack width at time x (mm)
# .....w(0) : crack width when it is visible 0.05
# .....beta : propagation controlling parameter
# .....P(x) : attack penetration mm at time x
# .....P(x0) : attack penetration mm at when w0 is visible
# .....V0 : mean corrosion rate
# .....V : corrosion rate mm/year
# .....wet : wetness period (sum of all raining days in a year)
# .....alpha : pitting factor taking non-uniform corrosion of the rebars into account (normal distribution)
cor_penetrate<-function(V0,alpha,wet,t){
  V0*alpha*wet*t
} #this fucntion defines the corrosion penetration
crackwidth<-function(w0,beta,Px0,V0,alpha,wet,t){
  w0+beta*(integrate(cor_penetrate,0,t,V0,alpha,wet)$value-Px0)
} #this function defines the width of crack
```

```
#crackwidth(w0,beta,Px0,V0,alpha,wet,t)
#stop("debugg")
#integrate(dnorm, down, up, mean = mu, sd = sigmal)
Ccrt<-matrix(double(1),nrow=T,ncol=N)
#....................................
#.....STEP 3: ESTIMATION
#this part does random generation of positive value using rnorm function
repeat {
  #-----for parameter concerning the initiation
  x <- rnorm(N,mu_Dc,sd_Dc) #x represent for diffusion coefficient
  y <- rnorm(N,mu_Cs,sd_Cs) #y represent for surface Cl concentration
  #-----for parameter concerning the corrosion propagation
  r1 <- rnorm(N,mu_w0,sd_w0) #
  r2 <- rnorm(N,mu_beta,sd_beta) #
  r3 <- rnorm(N,mu_Px0,sd_Px0) #
  r4 <- rnorm(N,mu_V0,sd_V0) #
  r5 <- rnorm(N,mu_alpha,sd_alpha) #
  r6 <- rnorm(N,mu_wet,sd_wet) #
  if ((length(which(x<0)))==0){break}
  if ((length(which(y<0)))==0){break}
  if ((length(which(r1<0)))==0){break}
  if ((length(which(r2<0)))==0){break}
  if ((length(which(r3<0)))==0){break}
  if ((length(which(r4<0)))==0){break}
  if ((length(which(r5<0)))==0){break}
  if ((length(which(r6<0)))==0){break}
}
for (t in 1:T){
  #.....STEP 3.1: estimate the statistical properties of chloride Ccl at any time t using random generation (e.
     g. Monte Carlo simulation or random sampling)
  for (n in 1:N){
  Cclt[t,n]<-Ccl(x[n],d,y[n],t)
  Ccrt[t,n]<-crackwidth(r1[n],r2[n],r3[n],r4[n],r5[n],r6[n],t)
  }
 # estimate the statistical properties of corrosion propagation Ccr at any time t using random generation (e.g.
      Monte Carlo simulation or random sampling)
  mu_cl[t] <- mean(Cclt[t,])
  mu_cr[t] <-mean(Ccrt[t,])
# mu_cr[t] <- rnorm(1,mean=0.3,sd=0.005)
  sigmal_cl[t] <- sd(Cclt[t,])
  sigmal_cr[t] <- sd(Ccrt[t,])
#  sigmal_cr[t] <- rnorm(1,mean=0.03,sd=0.005)
  #.....STEP 3.2: Calculating the state probability for each phase
  #for chloride
  ##...............................................
  for (k in 1:I){
    if (k==1){
      if (sigmal_cl[t]<epsilon) {
        sigmal_cl[t]<-epsilon
      }
      pi_cl[t,k]=pro(-Inf,a[k+1],mu_cl[t],sigmal_cl[t])$value
    } else if (k<I) {
      pi_cl[t,k]=pro(a[k],a[k+1],mu_cl[t],sigmal_cl[t])$value
    } else if (k==I) {
      pi_cl[t,k]=pro(a[k],Inf,mu_cl[t],sigmal_cl[t])$value
    }
  }#...............................................
  #for propagation
  #...............................................
  for (k in 1:J){
    if (k==1){
      if (sigmal_cr[t]<epsilon) {
        sigmal_cr[t]<-epsilon #
      }
      pi_cr[t,k]=pro(-Inf,b[k+1],mu_cr[t],sigmal_cr[t])$value
    } else if (k<J) {
      pi_cr[t,k]=pro(b[k],b[k+1],mu_cr[t],sigmal_cr[t])$value
    } else if (k==J) {
      pi_cr[t,k]=pro(b[k],Inf,mu_cr[t],sigmal_cr[t])$value
    }
  }#........................................
  #.....STEP 3.3: calculating the state probability for the entire system
  #combining two probability
  for (k in 1:(I+J-1)){
    if (k<I){
      pi[t,k]=pi_cl[t,k]
    } else {
       pi[t,k]=pi_cl[t,I]*pi_cr[t,(k-I+1)]
    }
  }
}
#############################################
#print(mu_cl[t])
cat("state probability of the initiation phase \n")
print(pi_cl)
cat("state probability of the propagation phase \n")
print(pi_cr)
```

```
cat("state probability of the entire process \n")
print(pi)
#Step 4: Condition state distribution
plot.new()
stackedPlot <- function(data, time=NULL, col=1:length(data),...) {
  if (is.null(time)) {
    time <- 1:length(data[[1]])
    plot(0,0, xlim = range(time), ylim = c(0,max(rowSums(data))), t="n" , ...)
    for (i in length(data):1){
      #the summup to the current collumn
      prep.data <- rowSums(data[1:i]);
      # The polygon must have his first and last point on the zero line
      prep.y <- c(0, prep.data,0)
      prep.x <- c(time[1], time, time[length(time)] )
      polygon(prep.x, prep.y, col=col[i], border = NA);
    }
  }
}
#define evolution of conditison states over time into data frame
csstate<-matrix(double(1),nrow=T,ncol=I) #state probability in year
csstate=pi_cl*100 #turning probability distribution of condition states to percentage
csstate <- data.frame(csstate)
colors=c("white","coral1","limegreen")
stackedPlot(csstate,col = colors,xlab="",ylab="") #this is used for defaul colour
mtext(expression(paste("Percentage(%)")),side=2,line=2.2, adj = 0.5 )
mtext(expression(paste("Time (years)")),side=1,line=2.2, adj = 0.5 )
legend("topright", inset=0.09, title="CSs",col=colors,lty=2,lwd=13,legend=c(1:(I)),bg="azure2",cex=0.8)
#..........................................................
plot.new()
#define evolution of conditison states over time into data frame
csstate<-matrix(double(1),nrow=T,ncol=(J)) #state probability in year
csstate=pi_cr*100 #turning probability distribution of condition states to percentage
csstate <- data.frame(csstate)
colors=c("limegreen","orange","yellow2")
stackedPlot(csstate,col = colors,xlab="",ylab="") #this is used for defaul colour
mtext(expression(paste("Percentage(%)")),side=2,line=2.2, adj = 0.5 )
mtext(expression(paste("Time (years)")),side=1,line=2.2, adj = 0.5 )
legend("topright", inset=0.09, title="CSs",col=colors,lty=2,lwd=13,legend=c(1:(J)),bg="azure2",cex=0.8)
#..........................................................
plot.new()
#define evolution of conditison states over time into data frame
csstate<-matrix(double(1),nrow=T,ncol=(I+J-1)) #state probability in year
csstate=pi #turning probability distribution of condition states to percentage
csstate <- data.frame(csstate)
colors=c("white","coral1","limegreen","orange","yellow2")
stackedPlot(csstate,col = colors,xlab="",ylab="") #this is used for defaul colour
mtext(expression(paste("Probability")),side=2,line=2.2, adj = 0.5 )
mtext(expression(paste("Time (years)")),side=1,line=2.2, adj = 0.5 )
legend("topright", inset=0.09, title="CSs",col=colors,lty=2,lwd=13,legend=c(1:(I+J-1)),bg="azure2",cex=0.8)
file.remove("pi.csv")
file.create("pi.csv")
write.table(pi, file="pi.csv", sep = ",", append = TRUE,col.names = FALSE)
#THE END
cat("THE END")
```

4.B R code for generating data

```r
#This program is coded by Nam Lethanh for use in the class IMP
#Purpose: To generate the random data
#.....STEP 1: Define model parameter
#Define the number of condition state
I=3 #number of CS related to chloride initiation, the last condition state related to failure state
J=3 #number of CS related to crack propagation
T=60 #years of investigation
limit_chrolide=0.48 #limit concentration of chloride
limit_crack=0.5 #limit crack width
a=seq(from=0, to=limit_chrolide,length.out=I) #range value of condition state for chloride
b=seq(from=0, to=limit_crack,length.out=J) #range value of condition state for crack
d=20 # distance from the concrete surface to rebar (in mm)
N=100 #Maximum number used for random sampling with Monte Carlo or sampling techniques
mu_Dc=100 #mean value of diffusion coefficient Dc
sd_Dc=0.5 #standard deviation of the value of diffusion coefficient Dc
mu_Cs=2.565 #mean value of chloride concentration at the surface
sd_Cs=0.405 #standard deviation of the chloride concentration at the surface

mu_w0<-0.05  #: crack width when it is visible 0.05
sd_w0<-0.005
mu_beta<- 0.01   #: propagation controlling parameter
sd_beta<-0.001
mu_Px0 <- 5    #: attack penetration mm at when w0 is visible
sd_Px0<-0.5
mu_V0 <-  0.012   #: mean corrosion rate; sd=0.006 / mm/year
sd_V0<-0.006
mu_wet <-  1  #: wetness period (sum of all raining days in a year)  sd=0.25
sd_wet<-0.25
mu_alpha <- 14 #9.28 #: sd =4.04
sd_alpha<-4.04
epsilon<-0.005 #this is used to prevent value of simulated standard deviation to be less than 0.005, which
     gives a distribution with value less than 10^(-16), and in R, if value is less than 10^(-16), the PC cannot
      understand, basically when that situation happens the result returns to 0, which is not true.
#--
#.....STEP 2: Define functions to be used
#define error function
erf=function(x) 2 * pnorm(x * sqrt(2)) - 1
#Model parameters

#define the function of chloride concentration in initiation phase
Ccl<-function(Cs,x,Dc,t){
  Cs*(1-erf(x/(2*sqrt(Dc*t))))
}
#....Cs is the chloride concentration at the surface
#....x is the distance from surface to rebar (mm)
#....Dc is diffusion coefficient, which is probabilistic
#....t is the time (e.g. in year)
#-----------------------------------------
#---define the probability in a certain range with lower bound and upper bound (for normal distribution)
pro=function(down,up,mu,sigmal){
  integrate(dnorm, down, up, mean = mu, sd = sigmal)
}
#define the dimension
pi_cl<-matrix(double(1),nrow=T,ncol=I) #state probability for initiation stage
pi_cr<-matrix(double(1),nrow=T,ncol=J) #state probability for propagation stage
mu_cl<-matrix(double(1),nrow=T,ncol=1) #mean value of chloride at time t
mu_cr<-matrix(double(1),nrow=T,ncol=1) #mean value of crack width at time t
sigmal_cl<-matrix(double(1),nrow=T,ncol=1) # standard deviation of chloride at time t
sigmal_cr<-matrix(double(1),nrow=T,ncol=1) # standard deviation of crack at time t
pi=matrix(double(1),nrow=T,ncol=(I+J-1)) # state probability of the entire process -1 mean we consider the last
      state I as failure state, therefore, the total state of the system is I+J-1
Cclt<-matrix(double(1),nrow=T,ncol=N)
#functions toi be used in propagation phase
cor_penetrate<-function(V0,alpha,wet,t){
  V0*alpha*wet*t
} #this fucntion defines the corrosion penetration
crackwidth<-function(w0,beta,Px0,V0,alpha,wet,t){
  w0+beta*(integrate(cor_penetrate,0,t,V0,alpha,wet)$value-Px0)
} #this function defines the width of crack
#integrate(dnorm, down, up, mean = mu, sd = sigmal)
Ccrt<-matrix(double(1),nrow=T,ncol=N)
#...................................
#.....STEP 3: ESTIMATION
#this part does random generation of positive value using rnorm function
repeat {
  #-----for parameter concerning the initiation
  x <- rnorm(N,mu_Dc,sd_Dc) #x represent for diffusion coefficient
  y <- rnorm(N,mu_Cs,sd_Cs) #y represent for surface Cl concentration
  #-----for parameter concerning the corrosion propagation
  r1 <- rnorm(N,mu_w0,sd_w0) #
  r2 <- rnorm(N,mu_beta,sd_beta) #
  r3 <- rnorm(N,mu_Px0,sd_Px0) #
  r4 <- rnorm(N,mu_V0,sd_V0) #
```

```
  r5 <- rnorm(N,mu_alpha,sd_alpha) #
  r6 <- rnorm(N,mu_wet,sd_wet) #
  if ((length(which(x<0)))==0){break}
  if ((length(which(y<0)))==0){break}
  if ((length(which(r1<0)))==0){break}
  if ((length(which(r2<0)))==0){break}
  if ((length(which(r3<0)))==0){break}
  if ((length(which(r4<0)))==0){break}
  if ((length(which(r5<0)))==0){break}
  if ((length(which(r6<0)))==0){break}
}
for (t in 1:T){
  #.....STEP 3.1: estimate the statistical properties of chloride Ccl at any time t using random generation (e.
     g. Monte Carlo simulation or random sampling)
 # N=N*sample(1:10, 1)
  for (n in 1:N){
  Cclt[t,n]<-Ccl(x[n],d,y[n],t)
  Ccrt[t,n]<-crackwidth(r1[n],r2[n],r3[n],r4[n],r5[n],r6[n],t)
  }
 # estimate the statistical properties of corrosion propagation Ccr at any time t using random generation (e.g.
      Monte Carlo simulation or random sampling)
  mu_cl[t] <- mean(Cclt[t,])
  mu_cr[t] <-mean(Ccrt[t,])
#--write value into csv file
#syntac to convert continuous data into discrete condition state
}
file.remove("data.csv")
file.create("data.csv")
write.table(cbind(mu_cl,mu_cr), file="data.csv", sep = ",", append = TRUE,col.names = FALSE)

file.remove("measurmentpoints_Ccl.csv")
file.create("measurmentpoints_Ccl.csv")

file.remove("measurmentpoints_Ccr.csv")
file.create("measurmentpoints_Ccr.csv")
write.table(Cclt, file="measurmentpoints_Ccl.csv", sep = ",", append = TRUE,col.names = FALSE)

write.table(Ccrt, file="measurmentpoints_Ccr.csv", sep = ",", append = TRUE,col.names = FALSE)
cat("THE END")
```

4.C R code for regression analysis

```r
#This program is coded by Nam Lethanh for use in the class IMP-HS2014
#Purpose: To perform simple regression analysis for estimate the regression parameter
#....read data from file
datacl<-read.csv("measurmentpoints_Ccl.csv",header=FALSE,sep = ",")
attach(datacl)
datacl=data.frame(datacl) #making data to be in frame
library(reshape)
N=200
M=250
#.................................................
mdatacl <- melt(datacl, id=c("V1"))
mdatacl=mdatacl[sample(nrow(mdatacl),N),]
cat("regression for Chloride concentration \n")
Ccl=lm(mdatacl[,3]~mdatacl[,1])
#plotting
plot.new()
plot(mdatacl[,1],mdatacl[,3], ylab="Chloride concentration", xlab="Time (year)",type="p",cex=0.5)
grid(10, 10, col = "lightgray", lty = "dotted",lwd = par("lwd"), equilogs = TRUE)
box()
abline(Ccl,col = "red",lwd=2)
legend("topleft", inset=0.05, title="",col=c("black","red"),pch=c(1,NA),lty=c(0,1),lwd=c("1","2"),legend=c("data point","fitted line"),bg="white",cex=1,box.lwd = 0,box.col = "white")
cat("regression for crack width \n")
datacr<-read.csv("measurmentpoints_Ccr.csv",header=FALSE,sep = ",")
attach(datacr)
datacr=data.frame(datacr) #making data to be in frame
library(reshape)
mdatacr <- melt(datacr, id=c("V1"))
mdatacr=mdatacr[sample(nrow(mdatacr),M),]
Ccr=lm(mdatacr[,3]~mdatacr[,1])
#plot.new()
plot(mdatacr[,1],mdatacr[,3], ylab="Crack width (mm)", xlab="Time (year)", type="p",cex=0.5,col="brown")
grid(10, 10, col = "lightgray", lty = "dotted",lwd = par("lwd"), equilogs = TRUE)
box()
abline(Ccr,col = "blue",lwd=2)
legend("topleft", inset=0.05, title="",col=c("brown","blue"),pch=c(1,NA),lty=c(0,1),lwd=c("1","2"),legend=c("data point","fitted line"),bg="white",cex=1,box.lwd = 0,box.col = "white")
#plot.new() #plot the deterioration curve with discrete CS
cat("regression for Chloride concentration \n")
print(summary(Ccl))
cat("regression for crack width \n")
print(summary(Ccr))
#Plot the deterioration curve
CS<-matrix(double(1),nrow=5,ncol=2)
for (i in 1:5){
  CS[i,1]=i
  }
CS[1,2]=(3.40275)/0.47975 #this is from regression analysis
CS[2,2]=(3.40275+0.24)/0.47975 #this is from regression analysis
CS[3,2]=(3.40275+0.48)/0.47975 +(0.500755)/0.053150
CS[4,2]=(3.40275+0.48)/0.47975+(0.500755+0.25)/0.053150
CS[5,2]=(3.40275+0.48)/0.47975+(0.500755+0.5)/0.053150
smoothingSpline = smooth.spline(CS[,2],CS[,1], spar=0.35)
plot(CS[,2],CS[,1], ylab="Condition state", xlab="Time (year)", type="p",cex=0.0,col="black",ylim = rev(range(CS[,1])))
lines(smoothingSpline,col="red",lwd=2)
grid(10, 10, col = "lightgray", lty = "dotted",lwd = par("lwd"), equilogs = TRUE)
box()
cat("THE END")
```

4.D R code for estimation of erosion evolution over 30 years

```
#This program is coded by Nam Lethanh for use in the class IMP
#Purpose: To calculate the time to errosion of waster water tanks in the treatment plants.
#Input
#Question A1
T<- 30 #Total time of investigation
beta=0.1
#define the mechanistic function
errosion<-matrix(double(1),nrow=T,ncol=1) # standard deviation of crack at time t
for (t in 0:T){
  if (t==1){
    errosion[t]=1
  } else{
    errosion[t]=errosion[t-1]*exp(beta)
  }
}
#plot the errosion evolution curve
plot.new()
plot(errosion, axes=FALSE,ylim=c(0,20),ylab="", xlab="",type="p",cex=0.5,col="red",lwd="3")
#axis(2,c(seq(0,20,by=2)))
axis(2, ylim=c(0,20),col="black",las=1)  ## las=1 makes horizontal labels
mtext(expression(paste("Erosion (mm)")),side=2,col="black",line=2.2)
axis(1,c(seq(0,T,by=2)))

mtext(expression(paste("Time (year)")),side=1,col="black",line=2.5)
grid(10, 10, col = "lightgray", lty = "dotted",lwd = par("lwd"), equilogs = TRUE)
box()
cat("END OF QUESTION A1")
#Question A2 - REGRESSION ANALYSIS
#....read data from file
data<-read.csv("IMP-A-waste-water-tanks.csv",header=TRUE,sep = ",")
attach(data)
data=data.frame(data) #making data to be in frame
library(reshape)
mdata <- melt(data, id=c("time"))
cat("regression for Erosion \n")
errosionreg=lm(mdata[,3]~mdata[,1])
print(summary(errosionreg))
#plotting
plot.new()
plot(mdata[,1],mdata[,3],axes=FALSE,ylim=c(0,20), xlim=c(0,30),ylab="", xlab="",type="p",cex=0.5)
abline(errosionreg,col = "blue",lwd=2)
axis(2, ylim=c(0,20),col="black",las=1)  ## las=1 makes horizontal labels
mtext(expression(paste("Erosion (mm)")),side=2,col="black",line=2.2)
axis(1,c(seq(0,T,by=2)))
mtext(expression(paste("Time (year)")),side=1,col="black",line=2.5)
grid(10, 10, col = "lightgray", lty = "dotted",lwd = par("lwd"), equilogs = TRUE)
box()
legend("topleft", inset=0.05, title="",col=c("black","blue"),pch=c(1,NA),lty=c(0,1),lwd=c("1","2"),legend=c("
    data point","fitted line"),bg="white",cex=1,box.lwd = 0,box.col = "white")
```

4.E R code for chloride concentration

```
#This program is coded by Nam Lethanh for use in the class IMP
#Purpose: To calculate wear out of the bridge
#reference: Analysis of parameters of structure deterioration models within the Moscow BMS
#.....STEP 1: Define model parameter
#Define the number of condition state
CS=5 #number of CS
T=50 #years of investigation
limit_I=0.8 # correspond to CS4
a=seq(from=0, to=limit_I,length.out=CS) #range value of condition state for the wear out
N=20000 #Maximum number used for random sampling with Monte Carlo or sampling techniques
mu_Tc=60 #mean value of average time Tc
sd_Tc=7 #standard deviation of the Tc
epsilon<-0.005 #this is used to prevent value of simulated standard deviation to be less than 0.005, which gives a distribution with value less than 10^(-16), and in R, if value is less than 10^(-16), the PC cannot understand, basically when that situation happens the result returns to 0, which is not true.
#-----------------------------------------------
#.....STEP 2: Define functions to be used
#define the coefficient as a function of average life of a structure element
lambda<-function(Tc){log(2)/Tc}
#define the wear out function
I<-function(t,Tc){
  exp(lambda(Tc)*t)-1
}
#---define the probability in a certain range with lower bound and upper bound (for normal distribution)
pro=function(down,up,mu,sigmal){
  integrate(dnorm, down, up, mean = mu, sd = sigmal)
}
#define the dimension
pi<-matrix(double(1),nrow=T,ncol=CS) #state probability for the wear out
mu_I<-matrix(double(1),nrow=T,ncol=1) #mean value of wear out at time t
sigmal_I<-matrix(double(1),nrow=T,ncol=1) # standard deviation of chloride at time t
It<-matrix(double(1),nrow=T,ncol=N)
#this part does random generation of positive value using rnorm function
repeat {
  #-----for parameter concerning the initiation
  x <- rnorm(N,mu_Tc,sd_Tc) #x represent for the wear out
  if ((length(which(x<0)))==0){break}
  }
for (t in 1:T){
  #.....STEP 3.1: estimate the statistical properties of the wear out at any time t using random generation (e.g. Monte Carlo simulation or random sampling)
for (n in 1:N){
    It[t,n]<-I(t,x[n])
  }
# estimate the statistical properties of wear out at any time t using random generation (e.g. Monte Carlo simulation or random sampling)
  mu_I[t] <- mean(It[t,])
  sigmal_I[t] <- sd(It[t,])
  #.....STEP 3.2: Calculating the state probability
  ##...............................................
  for (k in 1:CS){
    if (k==1){
      if (sigmal_I[t]<epsilon) {
        sigmal_I[t]<-epsilon
      }
      pi[t,k]=pro(-Inf,a[k+1],mu_I[t],sigmal_I[t])$value
    } else if (k<CS) {
      pi[t,k]=pro(a[k],a[k+1],mu_I[t],sigmal_I[t])$value
    } else if (k==CS) {
      pi[t,k]=pro(a[k],Inf,mu_I[t],sigmal_I[t])$value
    }
  }#...............................................
}
##############################################
#print(mu_I[t])
cat("state probability \n")
print(pi)
#Step 4: Condition state distribution
plot.new()
stackedPlot <- function(data, time=NULL, col=1:length(data),...) {
  if (is.null(time)) {
    time <- 1:length(data[[1]])
    plot(0,0, xlim = range(time), ylim = c(0,max(rowSums(data))), t="n" , ...)
    for (i in length(data):1){
      #the summup to the current collumn
      prep.data <- rowSums(data[1:i]);
      # The polygon must have his first and last point on the zero line
      prep.y <- c(0, prep.data,0)
      prep.x <- c(time[1], time, time[length(time)] )
      polygon(prep.x, prep.y, col=col[i], border = NA);
    }
  }
}
```

```
#define evolution of conditison states over time into data frame
csstate<-matrix(double(1),nrow=T,ncol=CS) #state probability in year
csstate=pi #turning probability distribution of condition states to percentage
csstate <- data.frame(csstate)
colors=c("white","coral1","limegreen","orange","yellow2")
stackedPlot(csstate,col = colors,xlab="",ylab="") #this is used for defaul colour
mtext(expression(paste("Probability")),side=2,line=2.2, adj = 0.5 )
mtext(expression(paste("Time (years)")),side=1,line=2.2, adj = 0.5 )
legend("topright", inset=0.09, title="CSs",col=colors,lty=2,lwd=13,legend=c(1:(CS)),bg="azure2",cex=1)
#...............................................................
#THE END
cat("THE END")
```

Chapter 5

Event Tree and Fault Tree Analysis

BRYAN T. ADEY AND NAM LETHANH

5.1 Event tree analysis - Theory

5.1.1 Definition

Event tree analysis (ETA) is a way to analyse and display different discrete scenarios, their corresponding probability of occurrence and the resulting consequences. ETA uses Boolean logic to analyse a chronological sequence of events and estimate the probability of their occurrence. A key component of event tree analysis is the event tree. An event tree is built from a starting event and branches at each subsequent event based on the values of key parameters these key parameters were intensity measures. When the event tree is complete, it is a logical and visual representation of the set of scenarios that can occur, i.e. multiple possible futures. An example is shown in Figure 5.1, the initiating event is a 100-year precipitation, which may result in the occurrence of a flood with water depths over 30 cm, under 30 cm, or not at all. In this example, the flood can lead to different damage states related to an infrastructure element, which can result in interruptions to traffic flow on the road network. The consequences of the scenarios can then be quantified. Also the heavy rain itself can cause an interruption. ETA is often referred to as having forward logic as it starts from

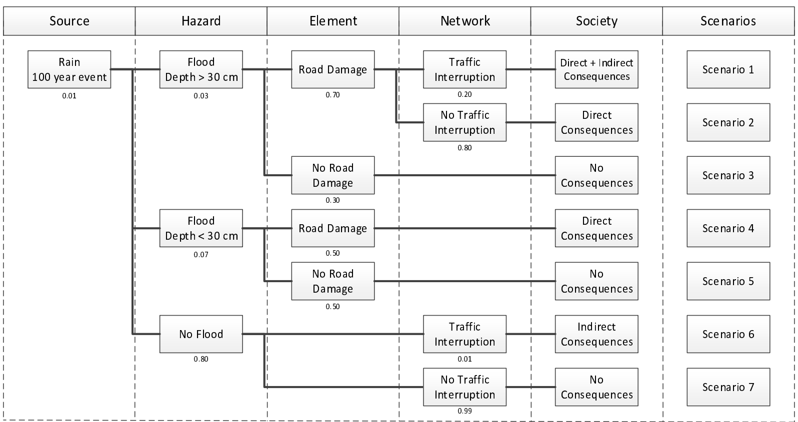


Fig. 5.1: Simplified event tree for a flood risk analysis

an initiating event and ends at a consequence. It is often used to analyze the risks associated with systems, and the consequences of functioning or failed systems given that an event(s) has (have) occurred. Simply the construction of an event tree can often be useful in decision situations as it brings to light many scenarios that decision makers otherwise would not have considered.

When conducting an ETA it is necessary to first determine the starting events that are to be considered and then to think through each of the next possible events to be considered. Each chain of events from initiating events to the end event (or consequence) is referred to as a path, and may be thought of as a scenario.

At each node in the event tree, the tree branches. There are often only two branches that emanate from a node but it is possible that no split occurs or that many splits occur. It is, however, of utmost importance to ensure that the split represents mutually exclusive possibilities, i.e. if one possibility happens the other cannot happen. The end events are often referred to as consequences.

The construction of an event tree requires knowledge about the system.

5.1.2 Steps

The basic steps to conduct an event tree analysis are:

1. Define static system: Determine what you would like to model and what is not to be considered in the model
2. Define dynamic system: Determine what you would like to model over time. For example, what hazards are of particular concern, e.g. flooding, what physical things may happen, e.g. bridge collapse, what consequences might occur, e.g. horrendous increases in travel times due to traffic jams
3. Identify specific initiating events: For example, heavy rain fall.
4. Identify specific intermediate events: For example, overtopping of a bridge
5. Construct the event tree
6. Calculate the probabilities of occurrence of each event: One possibility is using fault tree analysis.
7. Calculate the probability of occurrence of each scenario or path:
8. Calculate the overall risk:
9. Evaluate the acceptability of the risk associated with each scenario and overall:

Due to the almost infinite number of ways to represent reality and the almost infinite number of events that can occur, an appropriate system representation needs to be developed. A good starting point is often the system representation used to determine that there was a problem. It needs, however, to be verified that this representation is adequate for the evaluation of the candidate strategies. The necessary detail to be used depends on the specific question, the strategies to be evaluated and the level of detail desired.

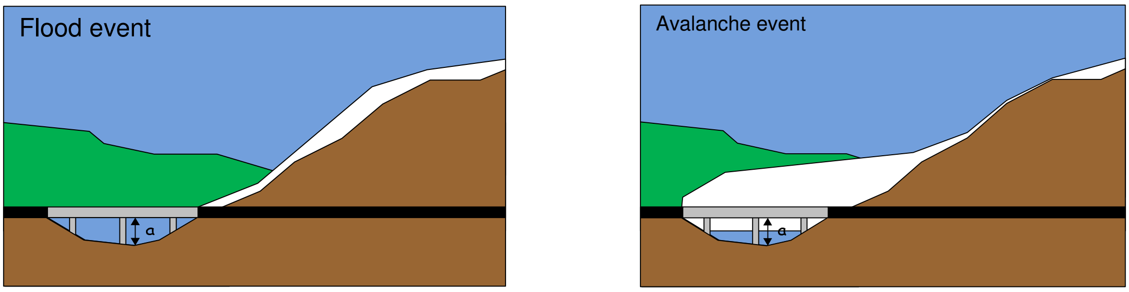
5.1.3 Example in literature

Please read the research work of Beim and Hobbs [1997], who constructed a event tree analysis for a lock gate system, as part of the script.

5.2 Event tree analysis - Example

5.2.1 Question A

You are the road manager responsible for a particularly risky section of road. It consists of a bridge that crosses a river and is just below a steep embankment, upon which it is known that avalanches sometimes occur. This is illustrated in Figure 5.2. You suspect that there are substantial infrastructure related risks due to both flooding and avalanches but



(a) Flood event (b) Avalanche event

Fig. 5.2: Illustration of risks

would like to quantify the risks. Develop an event tree to allow you to do this.

5.2.2 Answer A

One of the first things to do when developing an event tree to allow you to estimate the infrastructure related risks due to flooding and avalanches is to determine the what will cause the branching in the event tree[1]. The event tree is a part of the system representation[2]. The event tree in this case should have sufficiently good representations of the structures, hazards, and consequences, as well as the interaction between them so that it can be reasonably certain that there is an appropriate understanding of the system and that the risks. The event tree should consist of all possible things that can happen in a specified period of time, i.e. scenarios. In the event tree a scenario is represented by a path in the event tree. In the event tree proposed here the paths are divided into five components, 1) event occurrence, 2) effect of event, 3) physical changes in infrastructure, 4) direct consequences, and 5) indirect consequences. An example for a bridge exposed to flooding and avalanches is shown in Figure 5.4, assuming that flood and avalanche events are mutually exclusive. In this event tree the hazard consists of the event and the effects on the bridge object that are considered to have consequences. The effects on the bridge are represented by two levels of the key parameters, water depth $>= a$ and water depth $< a$, where a is the bridge clearance, and snow depth $>= b$ and snow depth $< b$, where b is the depth of snow on the bridge.

The consequence portion of the event tree is composed of the non-monetarisable and monetarisable direct and indirect consequences. For each level of the key parameter, assumptions are made with respect to the non-monetarisable direct consequences that can occur, e.g. the physical behavior of the structure when the specified level of the key parameter(s) occurs. The effect on the structure should be described with the fewest possible parameters. Assumptions are made with respect to the monetarisable direct consequences that can occur for each non-monetarisable direct consequence, e.g. the fatalities related to the collapse of the bridge. Assumptions are made with respect to the monetarisable indirect consequences that can occur for each monetarisable direct consequence, e.g. the additional user traffic time due to

[1] The answer is an adapted version of that used in Adey et al. [2009]

[2] The system representation is a model of the relevant part of reality used for the evaluation and consists of all possible realizations of stochastic processes within the investigated time period.

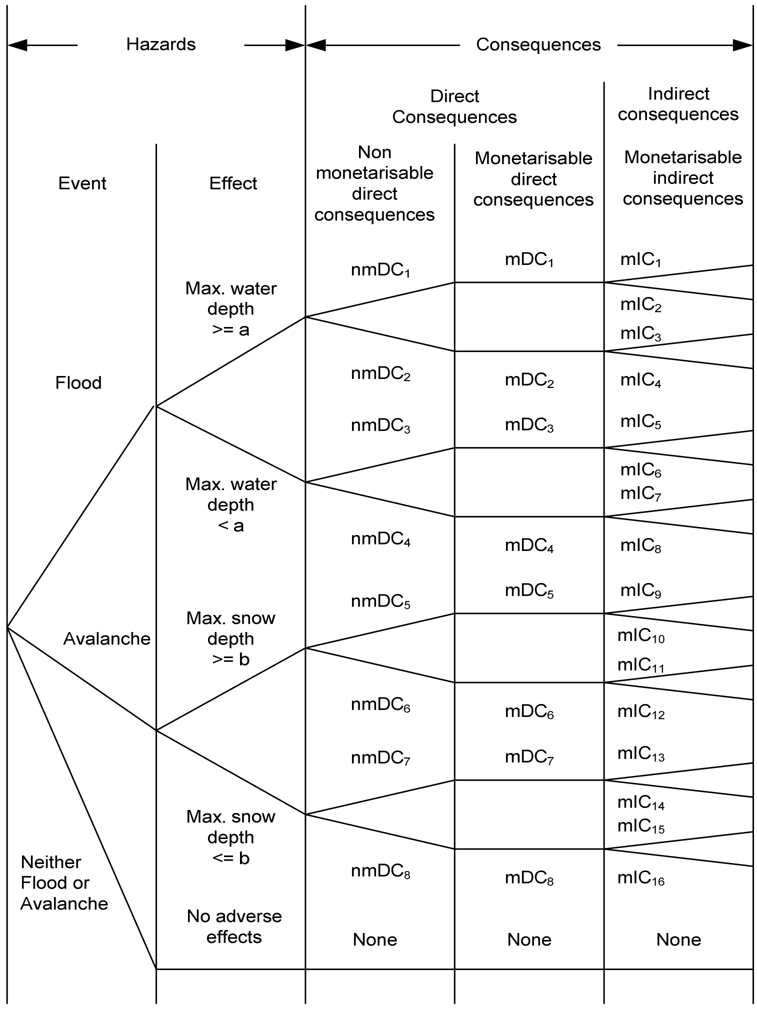


Fig. 5.3: Event tree with mutually exclusive flood and avalanche events.

a deviation. The monetarisable consequences can be attributed to those that will carry the consequences if failure occurs and can be grouped by consequences type. In the event tree (Figure 5.4), each effect has two possible non-monetarisable direct consequences ($nmDC_1$, $nmDC_2$). For example, when the water depth during a flood is $>= a$, than the two possible failure modes are the wash out of the abutments ($nmDC_1$) and the wash out of the columns ($nmDC_2$). The monetarisable direct consequences deterministically follow the non monetarisable direct consequences ($nmDC_1$). The monetarisable indirect consequences are branched into the monetarisable indirect consequences that would occur if the abutments were washed out during winter with low traffic volumes (mIC_1) and those that would occur if the abutments were washed out during summer with high traffic volumes (mIC_2)

5.2.3 *Question B*

Construct an event tree for a road segment and a bridge that may be affected by a flood or an avalanche, but not at the same time.

5.2.4 Answer B

The branches in the event tree are comprised of a chain of possible events and monetarisable direct and indirect consequences (Fig. 5.4 and Fig. 5.6). A slightly more detailed example of some of the avalanche scenarios is given in Table 5.1, where the effect on the physical object is listed simply as level. It is easy to imagine how detailed this can become.

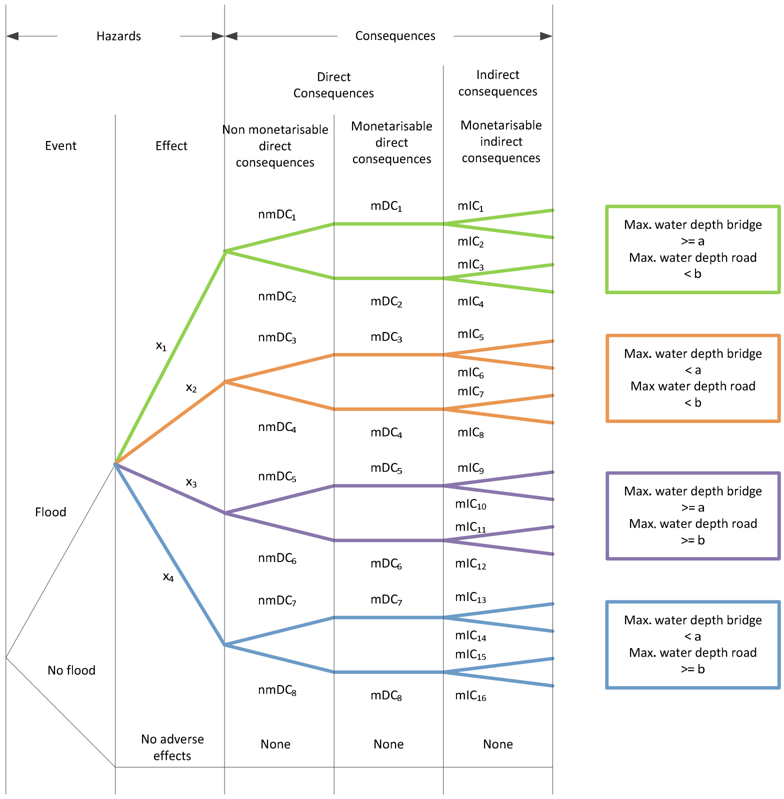


Fig. 5.4: Event tree with mutually exclusive flood event.

Table 5.1: Examples of scenarios

Nr.	Effect	Non-monetarisable direct consequences	Monetarisable direct consequences
1	On column level 1	Mild deformation of road surface	Level 1
2			Level 2
3		Severe deformation of road surface	Level 1
4			Level 2
5		Collapse of object	Level 2
6			Level 3
-	-	-	-
-	-	-	-
-	-	-	-
31	Horizontal loading of deck level 3	Mild displacement of deck	Level 1
32			Level 2
33		Severe displacement of deck	Level 1
34			Level 2
35		Removal of superstructure	Level 2
36			Level 3

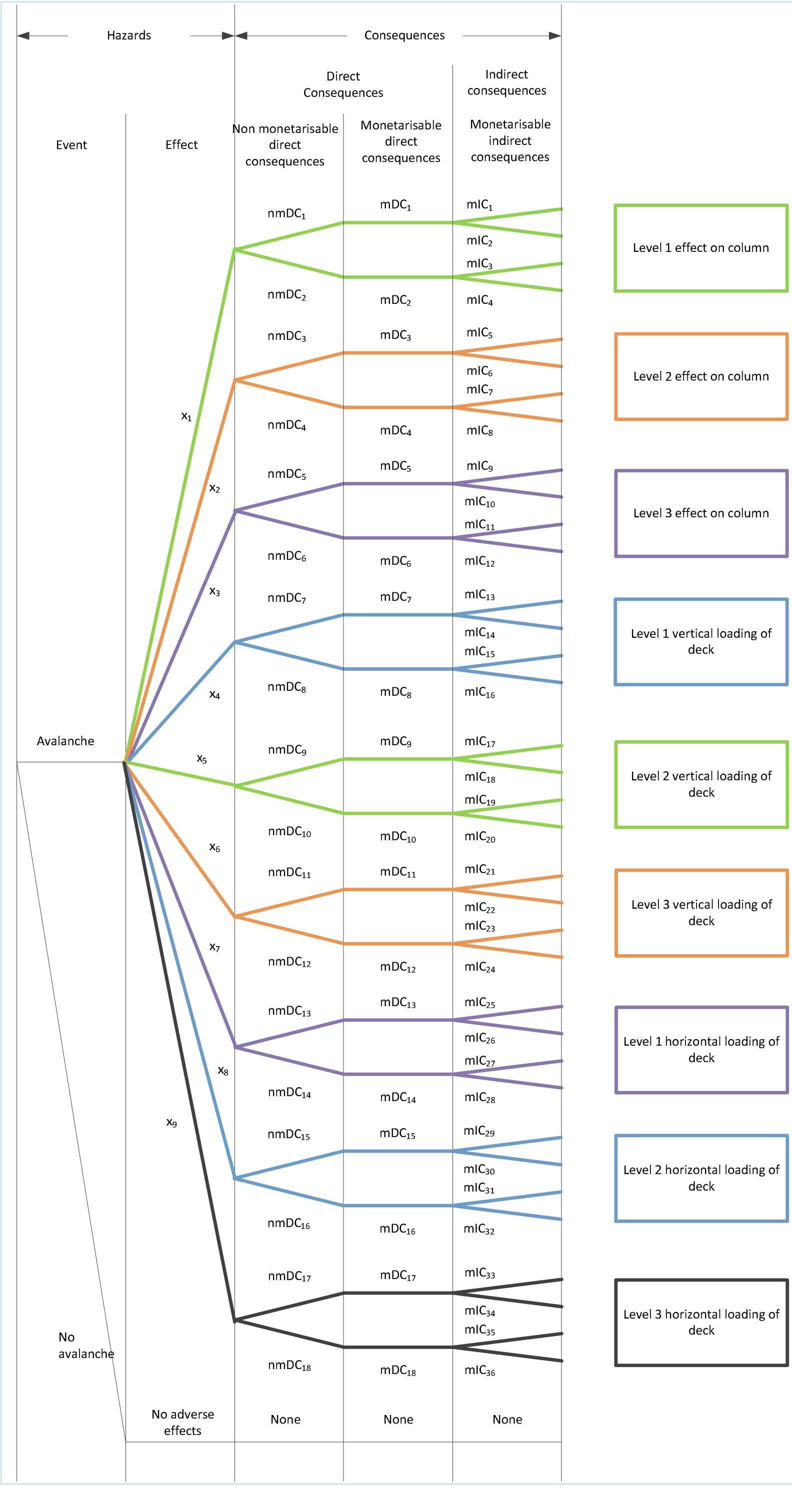


Fig. 5.5: Event tree with mutually exclusive avalanche event.

5.2.5 *Question C*

As a road manager you are considering the construction of avalanche protection to reduce the infrastructure related risk.

You know that with respect to the probabilities of occurrence of events that

- if you do nothing,the probability of occurrence of an avalanche decreases from 0.5 in 2009 to 0.3 in 2029 due to ever decreasing snow fall.
- If you construct avalanche protection, the probability of occurence of an avalanche will be 0 for the next 10 years but will increase after that

You know that the conditional probabilities of the effects on the loading of the column, the horizontal loading of the superstructure and the vertical loading of the superstructure are constant and as shown in Table 5.2. You know that conditional probability of physical changes due to horizontal loading of the columns and deck increases by 0.001 per year to take into account normal deterioration (Table 5.3 and Table 5.4) and due to vertical loading of deck is 1 and constant (Table 5.5). You also know that the due to changes in traffic that the conditional probabilities of consequences are as shown in Table 5.6. You have also estimate the consequences of each consequence level to be those as shown in Table 5.7.

Is it worth it if only a 20 year horizon is analysed? (Use a discount factor of 2%. Assume that the probability of multiple failures of the same infrastructure object are negligible. Show the yearly costs and benefits as well as the totals. For simplicity, assume that there is only one level of monetarisable consequences per level of physical damage.)

Table 5.2: Conditional probabilities of the effects

Year	Loading of the column			Horizontal loading of the superstructure			Vertical loading of the superstructure		
	Level 1	Level 2	Level 3	Level 1	Level 2	Level 3	Level 1	Level 2	Level 3
2009	0.3	0.4	0.3	0.3	0.4	0.3	0.3	0.4	0.3
2010	0.3	0.4	0.3	0.3	0.4	0.3	0.3	0.4	0.3
2011	0.3	0.4	0.3	0.3	0.4	0.3	0.3	0.4	0.3
2012	0.3	0.4	0.3	0.3	0.4	0.3	0.3	0.4	0.3
2013	0.3	0.4	0.3	0.3	0.4	0.3	0.3	0.4	0.3
-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-
2025	0.3	0.4	0.3	0.3	0.4	0.3	0.3	0.4	0.3
2026	0.3	0.4	0.3	0.3	0.4	0.3	0.3	0.4	0.3
2027	0.3	0.4	0.3	0.3	0.4	0.3	0.3	0.4	0.3
2028	0.3	0.4	0.3	0.3	0.4	0.3	0.3	0.4	0.3
2029	0.3	0.4	0.3	0.3	0.4	0.3	0.3	0.4	0.3

Table 5.3: Conditional probabilities of the occurrence of the physical changes due to loading of the column

Year	Effect		
	Loading of column		
	Mild deformation of road surface	Severe deformation of road surface	Collapse of object
2009	0.2	0.2	0.2
2010	0.201	0.201	0.201
2011	0.202	0.202	0.202
2012	0.203	0.203	0.203
-	-	-	-
-	-	-	-
2027	0.218	0.218	0.218
2028	0.219	0.219	0.219
2029	0.220	0.220	0.220

5.2.6 Answer C

The breakdown of the relative costs and benefits is given in Table 5.8. The yearly costs and benefits are given in Figure 5.6-Figure 5.10. The avalanche protection should be built.

Table 5.4: Conditional probabilities of the occurrence of the physical changes due to horizontal loading of the deck

Year	Effect		
	Horizontal loading of deck		
	Mild deformation of road surface	Severe deformation of road surface	Collapse of object
2009	0.2	0.2	0.2
2010	0.201	0.201	0.201
2011	0.202	0.202	0.202
2012	0.203	0.203	0.203
-	-	-	-
-	-	-	-
2027	0.218	0.218	0.218
2028	0.219	0.219	0.219
2029	0.220	0.220	0.220

Table 5.5: Conditional probabilities of the occurrence of the physical changes due to vertical loading of the deck

Year	Effect		
	Vertical loading of deck		
	Slightly yielded reinforcement	Plasticized reinforcement	Collapse of object
2009	1	1	1
2010	1	1	1
2011	1	1	1
2012	1	1	1
-	-	-	-
-	-	-	-
2027	1	1	1
2028	1	1	1
2029	1	1	1

Table 5.6: Conditional probabilities of monetarisable (direct and indirect) consequences

Year	Loading of column (L1, L2, L3) Horizonal loading of deck				Vertical loading of deck (L1, L2)		Vertical load of deck (L3)	
	Failure mode							
	Deformation of road surface		Collapse of object		Slightly yielded / plastized rebar		Collapse of object	
	Monetarisable direct and indirect consequences							
	Level 1	Level 2	Level 1	Level 2	Level 1	Level 2	Level 2	Level 3
2009	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
2010	0.499	0.501	0.499	0.501	0.499	0.501	0.499	0.501
2011	0.498	0.502	0.498	0.502	0.498	0.502	0.498	0.502
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
2027	0.482	0.518	0.482	0.518	0.482	0.518	0.482	0.518
2028	0.481	0.519	0.481	0.519	0.481	0.519	0.481	0.519
2029	0.480	0.520	0.480	0.520	0.480	0.520	0.480	0.520

Table 5.7: Consequences

Effect	Physical change	Level 1	Owner		User		Public	
			Inter-vention	Travel time	Nr. injuries	Nr. fatal-ities	Nr. injuries	Nr. fatal-ities
			106 CHF	106 CHF	106 CHF	106 CHF	106 CHF	106 CHF
On column level 1, 2 and 3	None	None	0	0	0	0	0	0
	Light deformation of road surface	1	0	0	0	0	0	0
		2	0.5	0.5	0.5	0.5	0.5	0.5
	Mild deformation of road surface	1	1.5	1.5	1.5	1.5	1.5	1.5
		2	2.5	2.5	2.5	2.5	2.5	2.5
	Collapse of object	2	5	5	5	5	5	5
		3	10	10	10	10	10	10
		-	-	-	-	-	-	-
Vertical loading of deck Stufe 3	Collapse of object	2	1	1	1	1	1	1
		3	2.5	2.5	2.5	2.5	2.5	2.5

Table 5.8: Costs, benefits and effectiveness

Intervention	Relative costs (10^6 mu)			Relative benefits (106 mu)			Effectiveness (B-C)
	Planned interventions	Restoration following failure	Total	Reduction of material damage risks	Reduction of failure risks	Total	
Do nothing	0	0	0	0	0	0	0
Avalanche protection	4.00	-1.00	3.00	2.00	3.00	5.00	2.00

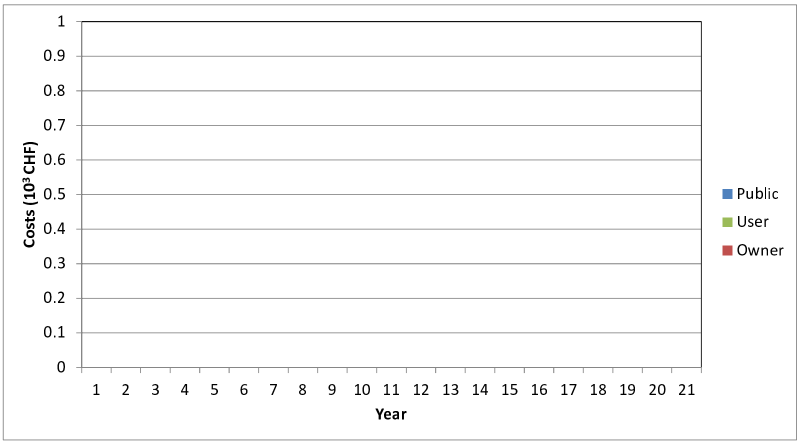


Fig. 5.6: Discounted costs of intervention per year for the do nothing intervention

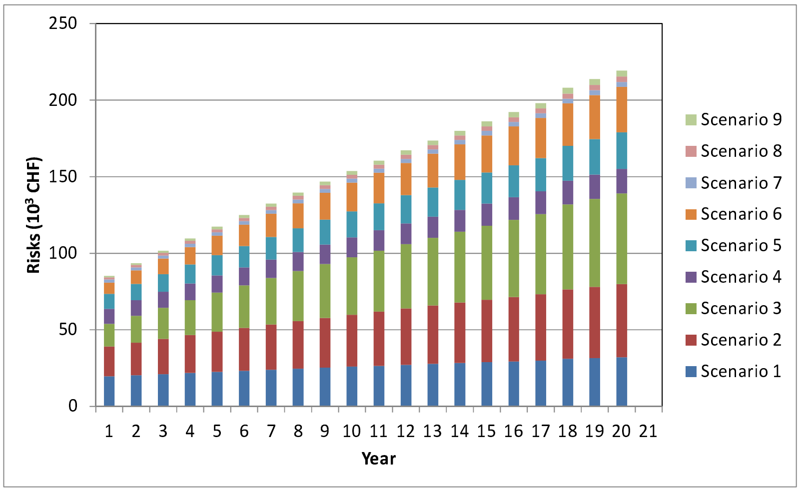


Fig. 5.7: Discounted risk per year for the do nothing intervention

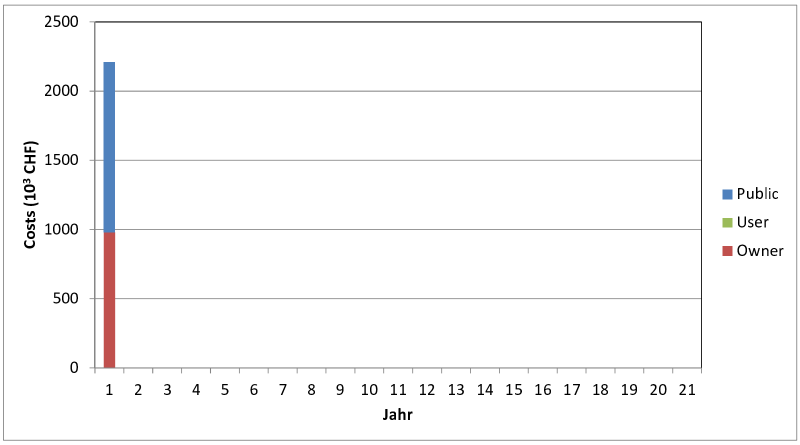


Fig. 5.8: Discounted costs of intervention per year for the avalanche protection intervention

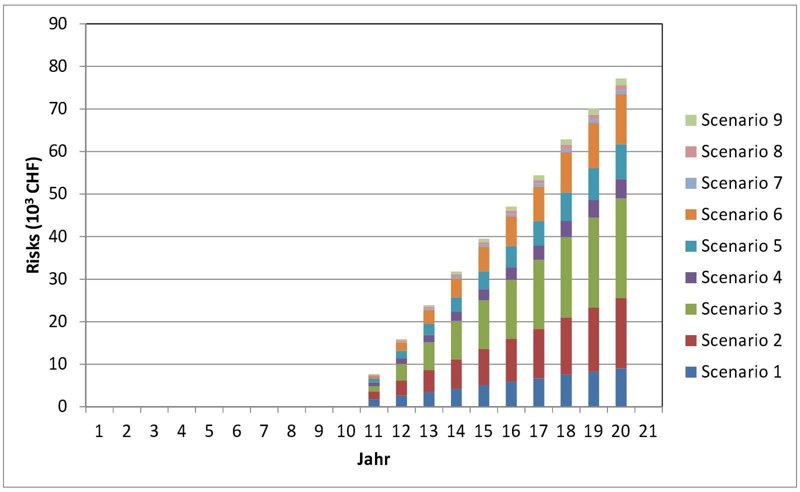


Fig. 5.9: Discounted risk per year for the avalanche protection intervention

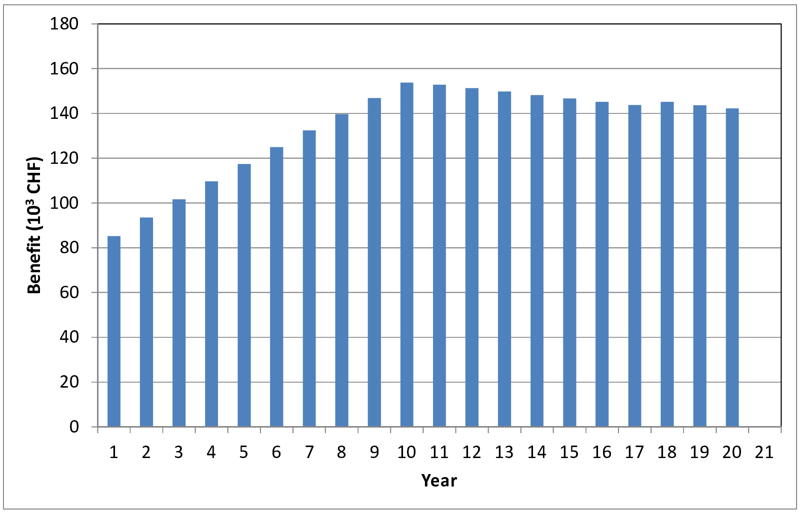


Fig. 5.10: Discounted benefit (in terms of risk reduction) per year for the avalanche protection intervention

5.3 Fault tree analysis – Theory

5.3.1 *Definition*

Fault tree analysis is often used when it is desired to focus on the particular ways a system can fail. It is particularly well suited to determine the probability of a specific event happening when this event can happen due to different precursor events. Fault trees are often used in the evaluation of large safety-critical systems [Andrews, 2012].

Fault tree analysis is often referred to as having backward logic as specific consequences are first determined and then the causes of these consequences are worked out, i.e. working backwards in time. Fault tree analysis is also sometimes seen as being top down, i.e. how can the top event be determined from a combination of lower level events.

In the development of a fault tree one normally starts by defining a top event and then developing logical expressions of sub-events that may lead to the occurrence of the top event. The last events in the tree are base events, or initiating events. Once a fault tree is constructed and the probability of occurrence of the base events are estimated the probability of occurrence of the top event can be estimated. When constructing fault trees it is necessary to ensure that that the events are binary events e.g. failure or no failure, that they are statistically independent, and that the relationships between events and causes can be represented by means of logical gates.

Fault trees can also be used to assess the infrastructure related risk due to natural hazards. Their focus, however, is more on a specific type of consequence, which may happen due to hazard of different types, e.g. a bridge may collapse due to a flood or an earthquake.

A separate fault tree is required for every separate top event, or in other words every failure mode.

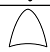
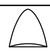
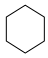
5.3.2 *Symbols*

A fault tree makes use of events and logic gates. These are shown in Table 5.9 and Table 5.10.

Table 5.9: Fault tree events

Event label	Event description	Event symbol
Basic event	failure or error in a subitem	
House event	The event is known to be true or false	
Undeveloped event	An event about which insufficient information is available, or which is of no consequence	
Conditional event	conditions that restrict or affect logic gates	
Intermediate event	can be used immediately above a primary event to provide more room to type the event description	

Table 5.10: Fault tree logic gates

Gate label	Gate description	Gate symbol
OR	the output occurs if any input occurs	
AND	the output occurs only if all inputs occur (inputs are independent)	
Exclusive OR	the output occurs if exactly one input occurs	
Priority AND	the output occurs if the inputs occur in a specific sequence specified by a conditioning event	
Inhibit	the output occurs if the input occurs under an enabling condition specified by a conditioning event	
Voting	output event occurs if at least m of the input occur	

5.3.3 Steps

FTA analysis involves five steps:

1. Define the undesired event, i.e. the top event.
2. Obtain an understanding of the system, i.e. determine all possible ways that this undesired event may happen. This requires experts who understand the functioning of the system well. If something is overlooked it will not be included in the fault tree and this may result in less than optimal decision being made to reduce the probability of occurrence of the undesired event.
3. Construct the fault tree:
4. Evaluate the fault tree: this requires the determination or estimation of the probability of occurrence of all base events and then the estimation of the probability of occurrence of the top event. As the determination of the probability of occurrence of all base events may take considerable time and effort it is worthwhile to keep in mind the exactness required so that a decision maker can come to the optimal decision. This estimation of the probability of the occurrence of the top event. This is explained in week 9 “Performance indicators: Reliability”.

5.3.4 Examples in literature

FTA has been widely used in management of many infrastructure systems. Some of examples are with gas system Andrews [2012], Tian et al. [2012]; bridges Davis-McDaniel et al. [2013], Johnson [1999], Sianipar and Adams [1997], LeBeau and Wadia-Fascetti [2007]; tunneling You and Tonon [2012], and urban transport Ma et al. [2014]. Students are requested to read these cited works as part of the handout.

5.4 Fault tree analysis – Example B

5.4.1 Question B1

As the manager of the bridge shown in Figure 5.11 and Figure 5.12, you are interested in having an overview of all the ways it might fail so you can determine the most effective way for you to reduce your risks. Construct a failure tree covering the following hazards 1) corrosion of the steel reinforcement, 2) spalling of concrete, 3) corrosion of the steel beams, 4) fatigue of the steel beams, 5) fire, 6) train impact, 7) over loading due to vehicles, 8) earthquakes, 9) scour. Assume the top event to be bridge collapse.

Fig. 5.11: Getwing bridge

5.4.2 Answer B1

A possible fault tree for the case of bridge collapse can be constructed and shown in following figures.

In Figure 5.13, the first level of the fault tree is shown. The bridge can collapse due to the collapse of either the superstructure (beams, girders) or the substructures (abutment and piers). The failure of both superstructure and the substructures are developed in the next figures.

Figure 5.14 shows the cause of the failure of superstructure can be due to the failure of the beams or girders (elements). Furthermore, the failure of each element can be, either due to manifest deterioration processes or latent deterioration processes.

Similar to Figure 5.14, Figure 5.15 shows the fault tree for the substructure, which includes the piers, abutments and bearing system. For the sake of simplicity, only three branches under sub-structure are shown. The full diagram would include many more branches. For example, instead of considering the failure of abutments as one event, the failure of each abutment may be considered as a separate event, i.e. failure of abutment 1 and failure of abutment 2, and therefore an illustration of this would include two more branches. Similarly, the development of the fault tree for piers could include the failure of each pier as a separate branch. The level of detail to be considered depends on the detail required in the analysis and the amount of effort to be invested.

Figure 5.16 shows the last step in development of the tree for the beam that is affected by manifest processes, e.g. corrosion, cracking, and fatigue under the normal traffic condition and no error in the design phase. There are a set of mathematical models to capture these manifest deterioration process (see script of week 4). It is noted that in this last step of tree development. The round circle or rectangular is used. This step often involves with the estimation of probability for the event to occur. For example, under the "failure due to corrosion of steels", there are two direct event that can be consider: 1) the first event is the concentration of chloride reaching a critical level of 0.48 (see script of week 4); 2) and other event could be the age of the bridge.

The complete fault tree for this example is given in OpenFTA extension file. Students are adviced to install an open source fault tree analysis software called "OpenFTA" from following link http://www.openfta.com/.

5.4.3 Question B2

Expand the fault tree to investigate the failure due to scour of the foundation of column 6.

5.4.4 Answer B2

This figure shows the extension of the fault tree, which includes the scouring process affecting the abutments and the piers of the bridges. The scouring process can be due to four process: 1) local scouring; 2) contraction; 3) widening; and 4) lateral migration. To understand these terminology, students are advised to read the work of Johnson [1999], which is provided as the class handout.

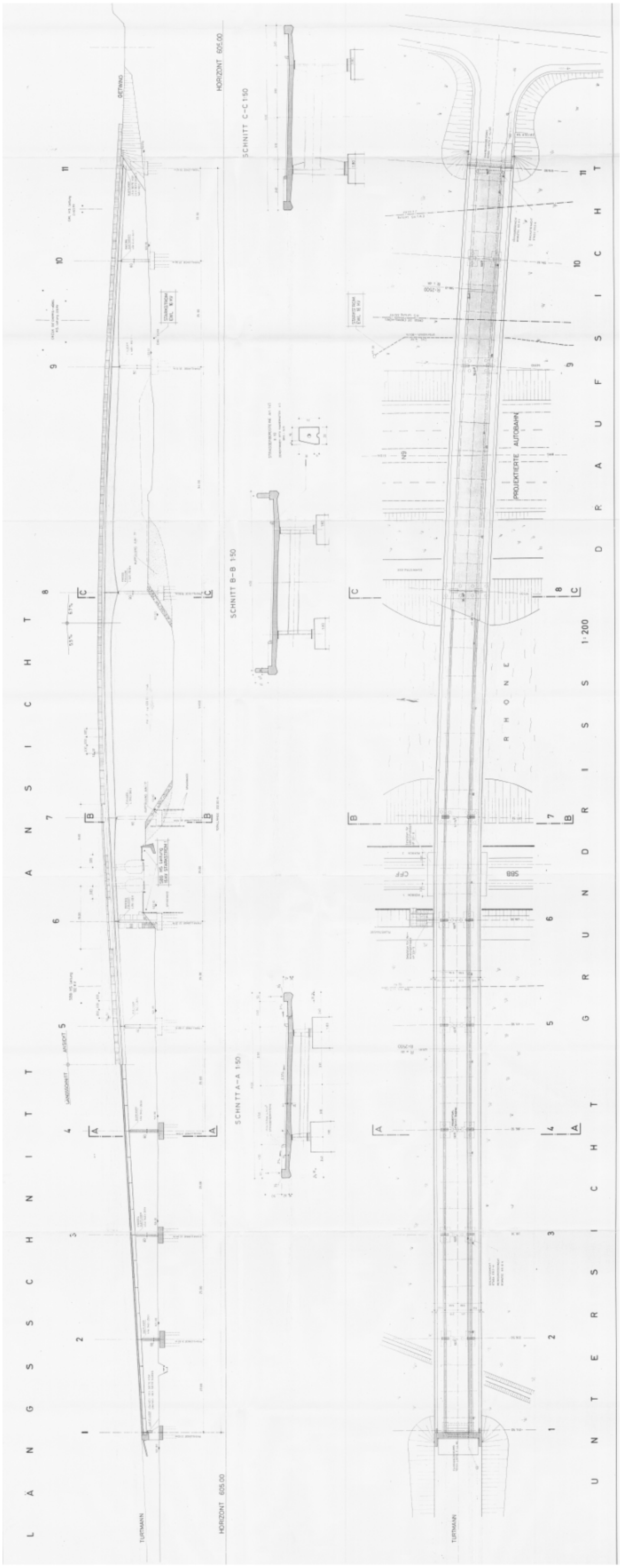


Fig. 5.12: Design of Getwing bridge

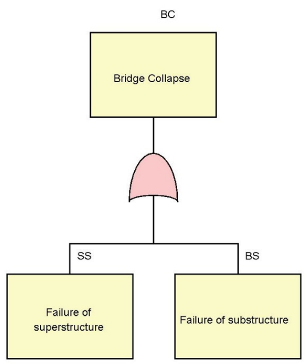


Fig. 5.13: First level of the fault tree

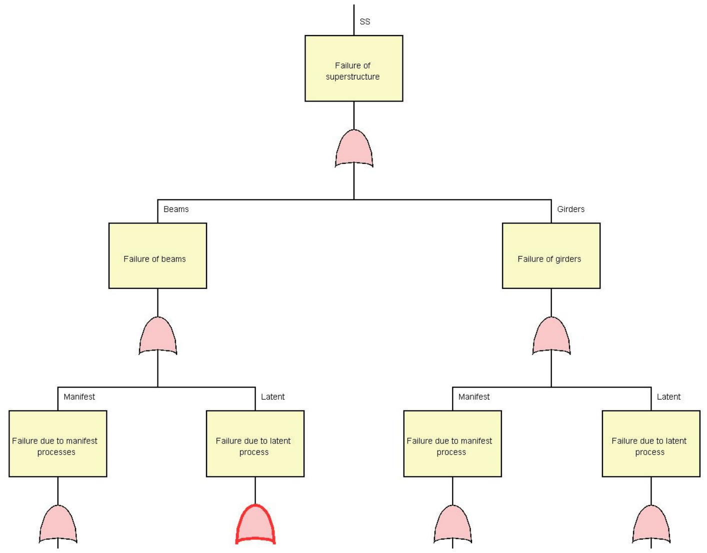


Fig. 5.14: Second level of the fault tree (superstructure)

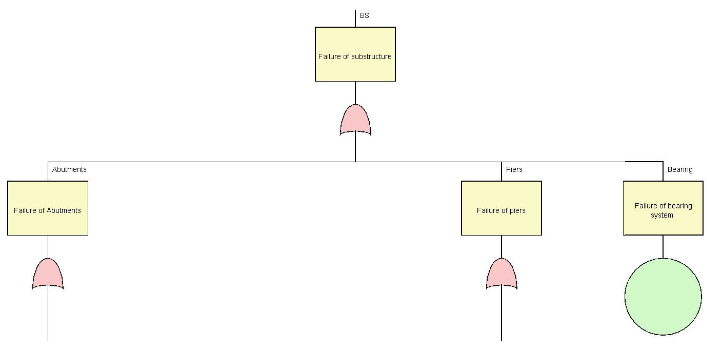


Fig. 5.15: Second level of the fault tree (sub-structures)

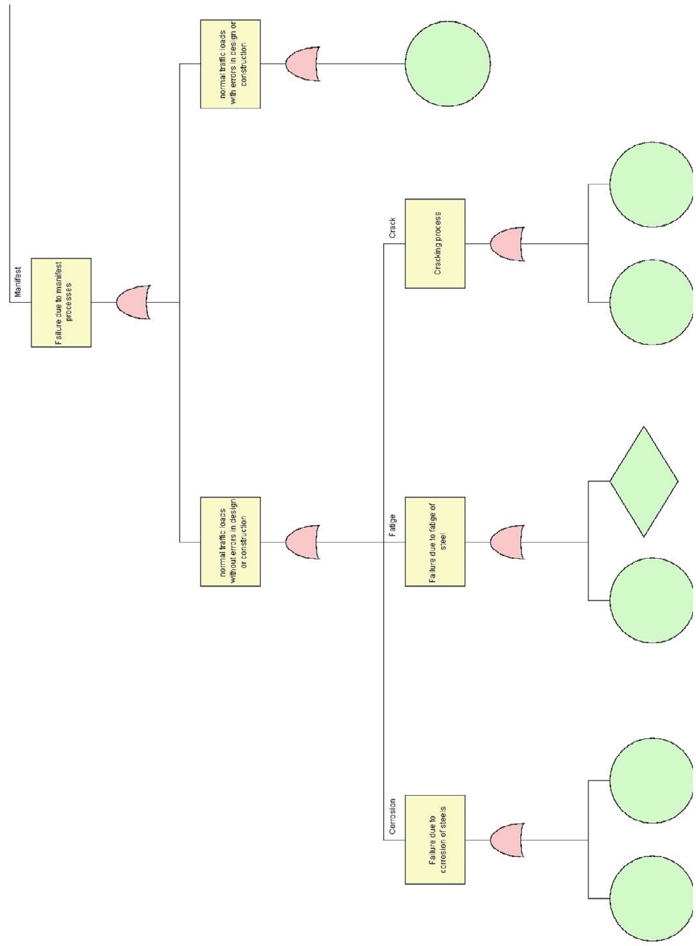


Fig. 5.16: Third level of the fault tree (superstructure-beam- manifest process)

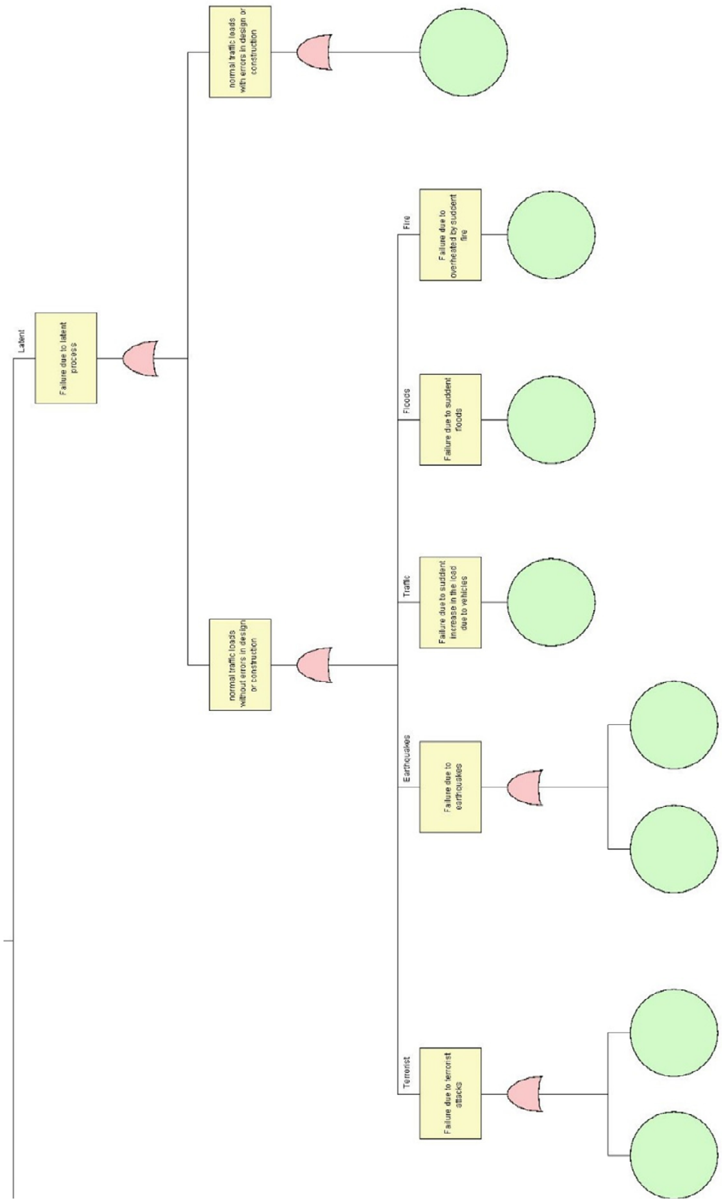


Fig. 5.17: Third level of the fault tree (superstructure- latent process)

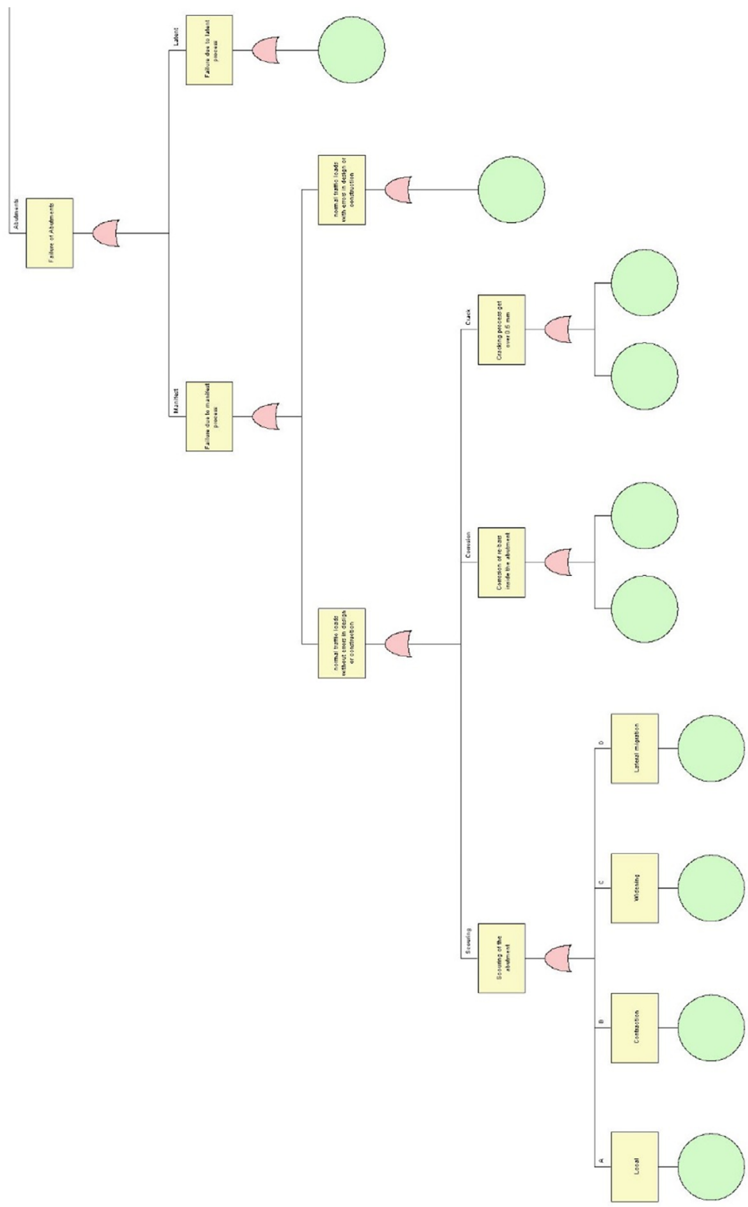


Fig. 5.18: Sub-structure (abutment)

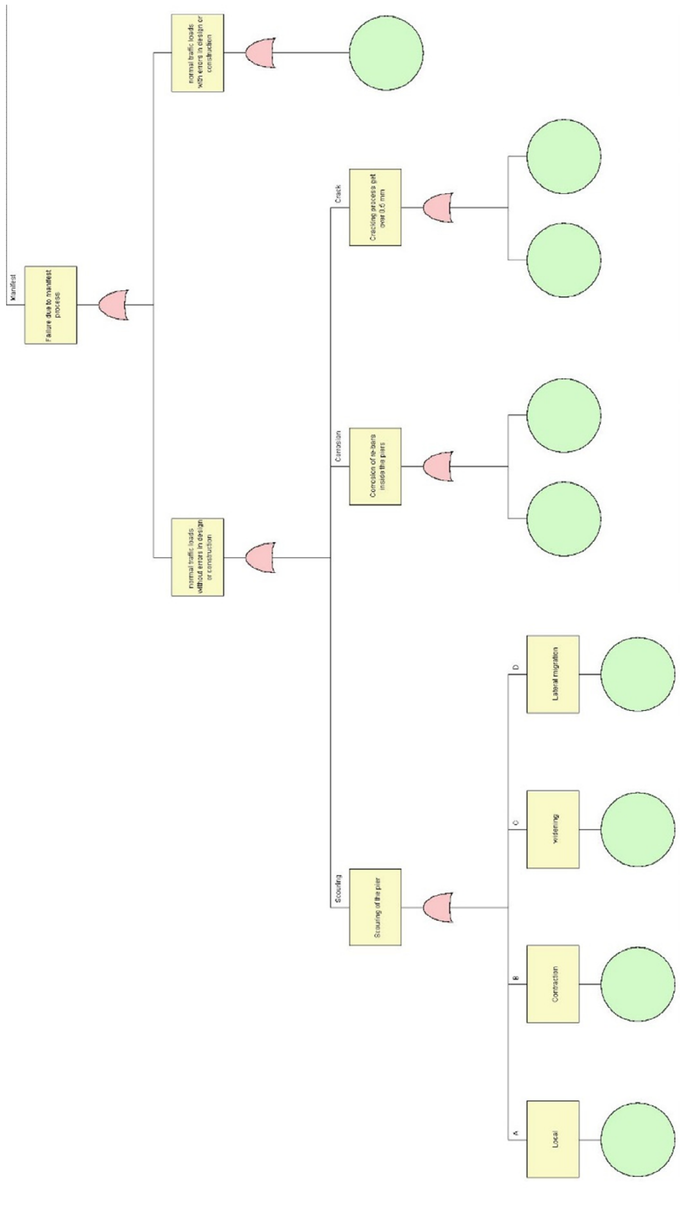


Fig. 5.19: Fault tree for scouring of the pier

5.5 Assignment

5.5.1 Problem A: Event tree analysis

Water is one of the fundamental needs of millions of people living in a megacity. Water of sufficient quality and quantity must be provided around the clock. In order to fulfill this need, a city depends on its water distribution infrastructure, which includes pipes made of different materials and laid at different times. These pipes are affected by processes of different types and deteriorate at different rates. The consequences related to pipe failure vary significantly depending on the type of failure, e.g. a pipe break, or a leaking pipe, as does the reaction time required to fix the pipe. For example, if a pipe breaks, an inadequate LOS occurs immediately and a corrective intervention is to be executed very quickly, and if it is noticed that there is progressive water loss over time then a preventive intervention can be planned before there is an inadequate LOS. Part of the water distribution network in mega-city Q is shown in Figure 1.4. The pipe characteristics are given in Table 1.13.

5.5.1.1 Question A1

Construct an event tree for the pipe network if an earthquake occurs? The event tree should include a sensible classification of events and consequences and it should be realistic to determine the probability of occurrence of each scenario. For the example, you should assume a probability of occurrence of each event in your event tree, and then estimate the probability of occurrence of each scenario. At whatever level you define your event tree be sure that it is complete!

5.5.1.2 Answer A1

No specific answer is provided. Please submit your answer to obtain feedback. A possible source event is the occurrence of an earthquake. Possible intermediary events are the zones in which a pipe breaks occur if an earthquake occurs. Possible network events are the loss of water and the loss of water pressure to serviced areas.

5.6 Problem B1

The manager of the pipeline network as shown in Figure 1.4 expects to understand how a large amount of water loss could occur. Water loss can be due to many reasons, which includes the manifest and latent deterioration process.

5.6.1 Question B1

Construct a fault tree for the pipe network (Figure 1.4) if there is a massive amount of water loss?

5.6.2 Answer B1

No specific answer is provided. Please submit your answer to obtain feedback. Keep in mind though that water loss may happen in one or more zones, and that these zones could be subdivided into specific pipes. Keep in mind that there are many different ways for pipes to fail, e.g. holes in pipes can occur due to differential soil movements, due to the penetration of roots into the pipes, the corrosion of reinforcement, sabotage, mistakes on construction sites, etc.

References

B.T. Adey, R. Hajdin, and J. Birdsall. Methodology to determine optimal intervention strategies for structures adversely affected by latent processes. In *Proceeding of the Annual Conference of The Transportation Association of Canada*, Vancouver, Canada, October (19-22) 2009.

John Andrews. Introduction to fault tree analysis. In *Proceeding of the Annual Reliability and Maintainability Symposium (RAMS)*, NV, USA, 23 Jan - 26 Jan 2012.

G. Beim and B. Hobbs. Event tree analysis of lock closure risks. *Journal of Water Resources Planning and Management*, 123(3):169–178, 1997. doi: 10.1061/(ASCE)0733-9496(1997)123:3(169). URL http://dx.doi.org/10.1061/(ASCE)0733-9496(1997)123:3(169).

C. Davis-McDaniel, M. Chowdhury, W. Pang, and K. Dey. Fault-tree model for risk assessment of bridge failure: Case study for segmental box girder bridges. *Journal of Infrastructure Systems*, 19(3):326–334, 2013. doi: 10.1061/(ASCE)IS.1943-555X.0000129. URL http://dx.doi.org/10.1061/(ASCE)IS.1943-555X.0000129.

P. Johnson. Fault tree analysis of bridge failure due to scour and channel instability. *Journal of Infrastructure Systems*, 5(1):35–41, 1999. doi: 10.1061/(ASCE)1076-0342(1999)5:1(35). URL http://dx.doi.org/10.1061/(ASCE)1076-0342(1999)5:1(35).

K. LeBeau and S. Wadia-Fascetti. Fault tree analysis of schoharie creek bridge collapse. *Journal of Performance of Constructed Facilities*, 21(4):320–326, 2007. doi: 10.1061/(ASCE)0887-3828(2007)21:4(320). URL http://dx.doi.org/10.1061/(ASCE)0887-3828(2007)21:4(320).

J. Ma, Y. Bai, J. Shen, and F. Zhou. Examining the impact of adverse weather on urban rail transit facilities on the basis of fault tree analysis and fuzzy synthetic evaluation. *Journal of Transportation Engineering*, 140(3):04013011, 2014. doi: 10.1061/(ASCE)TE.1943-5436.0000630. URL http://dx.doi.org/10.1061/(ASCE)TE.1943-5436.0000630.

P. Sianipar and T. Adams. Fault-tree model of bridge element deterioration due to interaction. *Journal of Infrastructure Systems*, 3(3):103–110, 1997. doi: 10.1061/(ASCE)1076-0342(1997)3:3(103). URL http://dx.doi.org/10.1061/(ASCE)1076-0342(1997)3:3(103).

Hui Tian, Weidong Qu, Pengfei Liu, Jun He, and Yijin Huang. *Application of Fault Tree Analysis in the Reliability Analysis of Oil-Gas Long Pipeline*, chapter 147, pages 1436–1447. ASCE, 2012. doi: 10.1061/9780784412619.147. URL http://ascelibrary.org/doi/abs/10.1061/9780784412619.147.

Xiaomin You and Fulvio Tonon. Event tree and fault tree analysis in tunneling with imprecise probabilities. In *GeoCongress 2012*, pages 2885–2894, 2012. doi: 10.1061/9780784412121.295. URL http://ascelibrary.org/doi/abs/10.1061/9780784412121.295.

Chapter 6

Bayesian Computation and Bayesian Network Models

NAM LETHANH

6.1 Introduction

Recently, Bayesian statistics has gained its ground in the research domain of the IAM and transportation engineering domain. It is considered as a powerful approach that can be used to solve many problems in the IAM. It is more advantageous than the traditional frequentist statistics [1] such as least squared estimation or regression analysis with Maximum likelihood estimation (MLE) approach.

A simple example on estimation of accident rate on a road link herewith to provide readers a basic difference between the traditional frequentist statistics and Bayesian statistics.

In the past 10 years, it is known that the probability of having an accident on a specific road section is 0.001 in a day. This value is considered as an expected mean probability, which is calculated by simply based on observed data. If someone asks you a question "what will be the probability of having an accident on that road tomorrow?", the correct answer is no longer with the value of 0.001, but it is different. The probability of having an accident tomorrow is a conditional probability that is calculated based on its dependency on observation of accident today, weather condition, number of daily traffic volume that are expected to be tomorrow. Frequentist statistics cannot provide answer to this question, but Bayesian statistics can because in the heart of Bayesian statistics, it is Bayes rule that allows to take into consideration of conditional probability and degree of belief based on various observed parameters such as accident today, weather condition, numbers of daily traffic volume, as well as opinions from experts [Andrew et al., 2006].

In briefly, Bayesian statistical approach can be used in replacement of frequentist statistical approaches for all of traditional statistical methods such as regression analysis, least squared estimation, and Maximum likelihood estimation approach. In addition, it can be used to solve complex and hierarchical models that methods used in traditional statistics cannot.

In the course of parameter estimation of statistical models, traditional statistics use analytical way to derive the solution. This analytical way faces difficulties when models are in complex hierarchical structures. In addition, traditional statistics is largely dependent on data. If data is insufficient or missing, the likelihood of estimation certainly involves greater bias and lower confidence [Andrew et al., 2006, Jeff, 2006]. Analytical method itself requires, for example, derivative approach for each model's parameters. For example, when using the MLE approach, it is necessary to obtain the first and second derivatives for each parameters. In mathematically term, it is necessary to derive the Jacobian matrices and Hessian matrices, then use Newton methods to obtain the values of model's parameters given data. However, in many cases, it is impossible to derive the Jacobian or Hessian matrices [2], leaving the problem unsolvable with the MLE approach. In this situation, the Bayesian approach can be used as it does not require the analytical solution, i.e. there is no need to obtain the Jacobian and Hessian matrices.

[1] the terminologies "frequentist statistics" and "Bayesian statistics" are used commonly in various statistical textbooks

[2] Jacobian and Hessian matrices are matrices of the first and second derivatives of model's parameters

In this chapter, the basic elements of the Bayesian inferential approach are introduced through examples related to prediction models for the IAM.

6.1.1 Bayes's theorem

Although literature on Bayesian statistics has been documented in a countless number, we would prefer to briefly present the Bayes's theorem in this section for the convenience of the readers and to serve as connection to the examples presented in this chapter.

In Bayesian statistics, the probability $P(A)$ of an event A is formulated as a degree of belief that A will occur. This degree of belief is subsequently referred as a prior probability which can be assumed relying on a subjective experience (theoretical assumption) or based on data of experiments or observations (frequentistic, empirical). Bayesian inference combines prior information and current information to derive the probability $P(A)$. This approach is considered as a process of fitting a probabilistic model to a set of data and summarizing the result by a probability distribution on the parameters of the model and on unobserved quantities such as predictions for new observations Geman and Geman [1984].

In the Bayesian statistics, the posterior distribution of parameters is estimated by using the likelihood function, which is defined by the prior distribution of parameters and the observed data. Here, the likelihood function is represented by $\mathscr{L}(\beta|\xi)$. β and ξ denote the unknown parameter vector and the observed data, respectively. It is assumed that β is a random variable, and is subjected to the prior probability density function $p(\beta)$. Under these conditions and according to Baye's law, when the observed data ξ is given, the posterior probability density function $p(\beta|\xi)$ of the unknown parameters β is defined as:

$$p(\beta|\xi) = \frac{\mathscr{L}(\beta|\xi)p(\beta)}{\int_{\Theta}\mathscr{L}(\beta|\xi)p(\beta)d\beta} \tag{6.1}$$

where Θ represents the parameter space. At this time, $p(\beta|\xi)$ can be expressed as follows:

$$p(\beta|\xi) \propto \mathscr{L}(\beta|\xi)p(\beta) \tag{6.2}$$

The symbol $\propto$ denotes 'be proportional to'. The denominator of the right-hand side of Eq. (6.1):

$$m(\xi) = \int_{\Theta}\mathscr{L}(\beta|\xi)p(\beta)d\beta \tag{6.3}$$

is called the normalization constant of $p(\beta|\xi)$, or the prior predictive distribution.

In short, the Bayes's theorem is summarized in following notation

$$\textbf{\textit{posterior}} \quad \propto \textbf{\textit{prior}} \times \textbf{\textit{likelihood}}$$

In general, the procedures of the Bayesian estimation can be summarized in following steps:

- Step 1: the prior probability distribution function $p(\beta)$ is specified, based on the prior information;
- Step 2: the likelihood function $\mathscr{L}(\beta|\xi)$ is defined based on the newly obtained data ξ;
- Step 3: the prior probability density function is modified in accordance with the Bayes' theorem, and the posterior probability density function $p(\beta|\xi)$ regarding the parameters β is updated.

6.1.2 MCMC Method

In statistic with Bayesian inference, prior and posterior probability are employed with aim to estimate the values of model's parameters. However, in actual analysis, it is hard to define a prior probability distribution [Ibrahim et al., 2001]. Methods to overcome the problems in the assumption of prior probability distribution often require numerical analysis with multi-dimensional integration, and thus remaining as a limitation in Bayesian estimation.

In recent years, an appealing solution to the problems in Bayesian estimation has been proposed, with the application of Markov Chain Monte Carlo (MCMC) simulation. The MCMC simulation technique does not require a high level of derivative and multi-dimensional integration of model's objective functions [Robert, 1996]. As a result, estimation results, in a great number of applied statistic research, have been improved through the combination of Bayesian estimation and MCMC simulation.

In MCMC simulation, Gibbs sampling and Metropolis Hastings (Metropolis-Hastings or MH) techniques have been extensively discussed [Robert, 1996]. Reference to research on image restoration is a good example of MCMC simulation [Geman and Geman, 1984]. Of that study, the algorithm of Gibbs sampling was used to estimate the posterior distribution in Bayesian estimation. In MH law, the iterative parameter β is defined by repeatedly generating random numbers through the conditional probability density function.

6.1.3 Example into Bayesian thinking

An infrastructure manager is interested in learning the accident risks on his/her highway network. He/she is given an information from a consultant that the numbers of accident per 100 *km* of road section in a year should be less than 8 cases.

What proportion of all road sections in the entire network get at least 8 accidents in a year?

Here, it is thought that the entire road network is composed of hundreds of road sections. Let p represents the proportion of this population which has a maximum level of 8 accidents per year. It is the interest of the infrastructure manager to understand the value of p, which is unknown.

As earlier discussed, in Bayesian inference, the belief of the manager on the proportion p can be represented by a probability distribution. This distribution is referred as prior knowledge that the manager believe in about the value of p.

Certainly, the infrastructure manager might not have all data for the entire network, therefore, he/she decides to carry out a sample that focuses on only a number of road sections. However, before that, he/she does some initial investigation on this proportion by reading some scientific research papers. This will help him/her in formulating the prior distribution used to compute the value of p.

First based on the sample report, he/she comes to know that there are, on average, 6 accidents occurs in a year for that pool of sample. Second, he/she does reading some technical reports in other regions on 100 samples, it is concluded in that technical report that there are approximately 70% of road sections have a numbers of accident per year falling in between 5 and 6 cases, 28% of road sections have 7 to 8 accidents, and about 2% of road section have more than 9 accident cases.

Based on these information, the manager believes that the numbers of accident per year for the entire network generally less than 8 cases per 100 *km* of road is having a proportion smaller than 0.5. After some reflection, her/his best guess value of p is 0.3. However, this is a very likely that the value of p can fall arbitrarily between 0 and 0.5.

He/she decides to carry out an investigation with 27 samples (e.g. road sections). Among 27 samples, 11 records have at least 8 accidents.

Following assumptions are used:

- the prior density for p is denoted by $g(p)$
- the good case is when the number of accident is less than 8 cases and the not good is when the number of accident is greater or equal than 8 cases, then we can write down the likelihood function in equation (6.4)

$$L(p) \propto p^n \cdot (1-p)^g \tag{6.4}$$

where g and n represent good and not good

The posterior density for p, with Bayesian inference, is obtained, up to a proportionality constant, by multiplying the prior density by the likelihood (equation (6.1)).

$$g(p|data) \propto g(p) \cdot L(p) \tag{6.5}$$

the word "data" in equation (6.5) infers the notation "ξ" in equation (6.1)

To obtain the posterior value of p, one can use discrete prior distribution as explained below

A simple approach to represent the prior value of p is to write it down as a list of reasonable proportion value and then assigns weights to these values. For example, following vector

0.05 0.15 0.25 0.35 0.45 0.55 0.65 0.75 0.85 0.95

represents possible value of p discretely. Based on her/his belief, following weights to this vector are assigned

Range of p	0.05	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95
Weight	1	5.2	8	7.2	4.6	2.1	0.7	0.1	0	0

Using this information, the discrete prior distribution can be shown graphically (Fig. 6.1) In this example, it is known

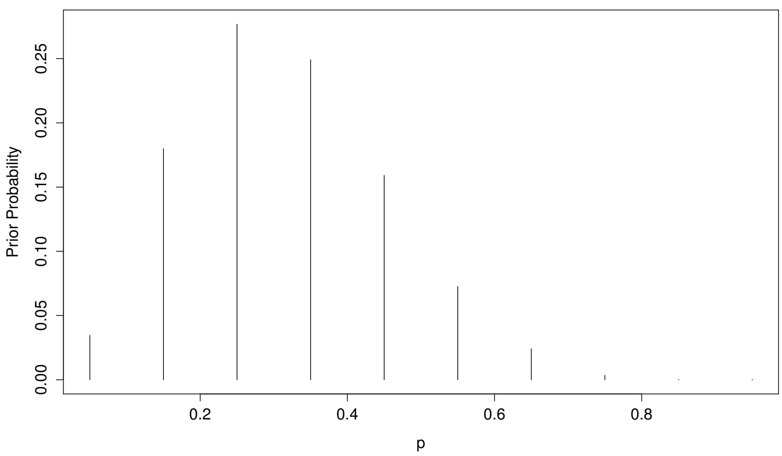


Fig. 6.1: Prior distribution for a proportion of p

that there are 11 samples out of 27 samples have numbers of accidents equal or greater than 8, therefore the value of n=11 and g=16, the likelihood function in equation (6.4) can be then written as The posterior density for p, with Bayesian inference, is obtained, up to a proportionality constant, by multiplying the prior density by the likelihood (equation (6.1)).

$$L(p) \propto p^{11} \cdot (1-p)^{16} \tag{6.6}$$

Using R code for this problem (Appendix 6.A), following posterior probability of p can be estimated Fig. 6.2 shows

Table 6.1: Prior and posterior of p

No.	p	Prior	Posterior
1	0.05	0.03	0
2	0.15	0.18	0
3	0.25	0.28	0.13
4	0.35	0.25	0.48
5	0.45	0.16	0.33
6	0.55	0.07	0.06
7	0.65	0.02	0
8	0.75	0	0
9	0.85	0	0
10	0.95	0	0

the distribution of the prior and posterior of p The posterior value of p, as can be seen from the table and the figure, is

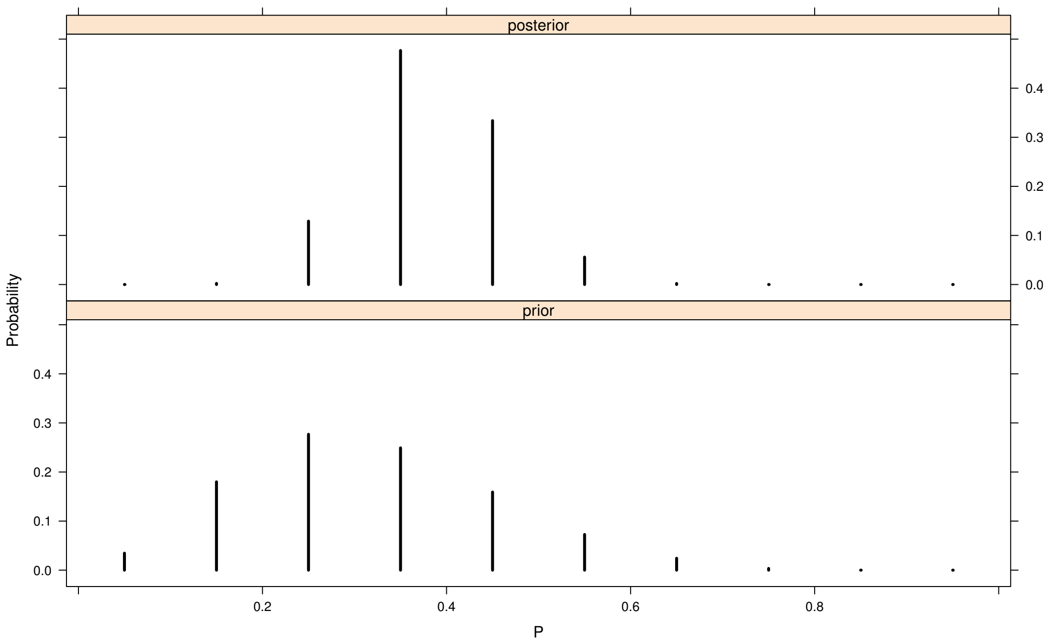


Fig. 6.2: Prior and posterior distribution of p

about 95% concentrating on the range from 0.25 to 0.45 . This infers that the probability that numbers of accidents per year for 100 *km* road reaches equal or greater than 8 cases will likely to happen only in the range from 0.25 to 0.45.

6.2 Bayesian inference for simple linear regression model

In section 4.3 of Chapter 4, a simple linear regression model is presented along with the squared root estimation approach. This chapter elaborates the Bayesian inference for solving this simple linear regression to show how it can be applied in a simplest model. A complete description of the approach is referred to Bolstad [2007].

In a simple regression model (e.g. Eq. (4.13)), we expect to model a relationship between two variables, x and y. Both variables are observed data. x is referred characteristic variables (sometimes also referred as predictor variable or independent variable). Example of x in the context of data in the IAM are traffic volume, thickness of road sections, ambient temperature, etc. These data are observable. y is often referred as outcome (e.g. condition state of road sections, percentage of cracks on a bridge deck), which is also observable. The mission is to model the relationship between x and y in a linear form given a data set of n ordered pairs of sample (x_i, y_i) for $i = 1, \cdots, n$.

In order to construct such a relationship, as shown in section 4.3 of Chapter 4, a derivative method is used. The method is referred as *least squares* estimation.

Using this linear model, in fact, we can solve many problems in practices involving not only linear form but also nonlinear form of models. For example, in following equation describes a growth model in exponential form.

$$y = e^{\alpha+\beta\cdot x} \tag{6.7}$$

This equation is used often in practice, also for the deterioration prediction with exponential form [Lethanh and Adey, 2013, Lethanh et al., 2015], for estimating impacts such as vehicle operating cost [OpusCL, 1999] and accident rate [Kumares and Samuel, 2007].

The equation (6.7) can be easily transformed into a linear form if we take logarithm for both sides of it.

$$log(y) = \alpha + \beta \cdot x \tag{6.8}$$

Using the linear form in equation (6.8), again, we can use the least squares method to obtain the model's parameters α and β given data.

In reference to the Bayes' theorem, for this regression model, it is necessary to define the likelihood and select the prior for the model's parameters α and β.

6.2.1 The joint likelihood for α and β

For a sample i^{th}, the joint probability density function as a joint function of parameters α and β for the paired observed data (x_i, y_i) is

$$likelihood_i(\alpha,\beta) \propto e^{-\frac{1}{2\sigma^2}[y_i-(\alpha+\beta(x_i-\bar{x}))]^2} \tag{6.9}$$

The likelihood of the entire data is the join probability and therefore is expressed as

$$likelihood_{data}(\alpha,\beta) \propto \prod_{i=1}^{n} e^{-\frac{1}{2\sigma^2}[y_i-(\alpha+\beta(x_i-\bar{x}))]^2} \tag{6.10}$$

$$\propto -\frac{1}{2\sigma^2}\sum_{i=1}^{n}[y_i-(\alpha+\beta(x_i-\bar{x}))]^2 \tag{6.11}$$

The polynomial inside the bracket of equation (6.11) equals to

$$\sum_{i=1}^{n}[y_i-\bar{y}+\bar{y}+(\alpha+\beta(x_i-\bar{x}))]^2 \tag{6.12}$$

which equals to

$$\sum_{i=1}^{n}(y_i-\bar{y})^2+2\sum_{i=1}^{n}(y_i-\bar{y})(\bar{y}-(\alpha+\beta(x_i-\bar{x})))+\sum_{i=1}^{n}(\bar{y}-(\alpha+\beta(x_i-\bar{x})))^2 \tag{6.13}$$

This is simplified as

$$\varepsilon_y-2\beta\varepsilon_{xy}+\beta^2\varepsilon_x+n(\alpha-\bar{y})^2 \tag{6.14}$$

where $\varepsilon_y=\sum_{i=1}^{n}(y_i-\bar{y})^2$, $\varepsilon_{xy}=\sum_{i=1}^{n}(y_i-\bar{y})(x_i-\bar{x})$, and $\varepsilon_x=\sum_{i=1}^{n}(x_i-\bar{x})^2$.

Finally, the join likelihood can be described as

$$likelihood_{data}(\alpha,\beta)\propto e^{-\frac{1}{2\sigma^2}\left[\varepsilon_y-2\beta\varepsilon_{xy}+\beta^2\varepsilon_x+n(\alpha-\bar{y})^2\right]} \tag{6.15}$$

$$\propto e^{-\frac{1}{2\sigma^2}\left[\varepsilon_y-2\beta\varepsilon_{xy}+\beta^2\varepsilon_x\right]}\times e^{-\frac{1}{2\sigma^2}\left[n(\alpha-\bar{y})^2\right]} \tag{6.16}$$

$$\propto e^{-\frac{1}{2\sigma^2/\varepsilon_x}\left[\beta-\frac{\varepsilon_{xy}}{\varepsilon_x}\right]^2}\times e^{-\frac{1}{2\sigma^2/n}\left[(\alpha-\bar{y})^2\right]} \tag{6.17}$$

In equation (6.17), it is obvious to recognize that the ratio $B=\frac{\varepsilon_{xy}}{\varepsilon_x}$ is the slop of the least squared line and $\bar{y}$ is the intercept of the vertical line $A=\bar{y}$. Evidently, the join likelihood of the two parameter has been factored .

$$likelihood_{data}(\alpha,\beta)\propto likelihood_{data}(\alpha)\times likelihood_{data}(\beta) \tag{6.18}$$

where

$$likelihood_{data}(\beta)=e^{-\frac{1}{2\sigma^2/\varepsilon_x}[\beta-B]^2} \tag{6.19}$$

and

$$likelihood_{data}(\alpha)=e^{-\frac{1}{2\sigma^2/n}\left[(\alpha-A)^2\right]} \tag{6.20}$$

It can be seen from equation (6.19) and (6.20) of β and α that the two likelihood functions are normal distributions $\mathcal{N}(B,\frac{\sigma^2}{\varepsilon_x})$ and $\mathcal{N}(A,\frac{\sigma^2}{n})$, respectively.

6.2.2 The joint prior for α and β

The joint prior for α and β is the multiplication of the two priors

$$g(\alpha,\beta)=g(\alpha)\times g(\beta) \tag{6.21}$$

The selection for the distribution of $g(\alpha)$ and $g(\beta)$ can be normal distribution, which is a conjugate distribution to the likelihood.

6.2.3 The join posterior for α and β

From the definition of Bayes's theorem ($posterior \propto prior \times likelihood$), we can have the joint posterior as follows

$$g(\alpha,\beta|data) \propto g(\alpha,\beta) \times likelihood_{data}(\alpha,\beta) \propto g(\alpha|data) \times g(\beta|data) \tag{6.22}$$

6.2.4 Example - Accident data on road

Table 6.2 shows the data on numbers of accident occurred on different road sections during the period of 2 year. In order to understand the functional relationship between the numbers of accident and the characteristic variables such as condition state (CS), daily traffic volume (DTV), slope, and speed of vehicle at the time of accident, investigator use a linear model (refer to equation (4.13)). In Figure 6.3, scatter plots are shown based on the data given in the table.

Table 6.2: Accident data on a road link

Sections	Numbers of accident	CS	DTV	Slope	Speed
1	5	2	9'746	19	52
2	2	1	6'878	1	59
3	4	2	5'758	0	50
4	5	1	6'903	17	52
5	10	2	9'114	20	54
6	8	3	7'890	8	80
7	10	4	9'539	12	65
8	2	1	7'062	3	72
9	3	5	8'846	9	54
10	2	5	7'383	3	79
11	6	5	9'755	18	77
12	10	3	7'351	10	51
13	3	3	8'376	17	63
14	7	2	7'646	15	72
15	1	5	7'002	10	74
16	5	1	8'644	19	78
17	2	1	9'719	7	69
18	5	2	8'851	13	54
19	5	3	9'353	19	67
20	10	4	7'824	18	68
21	0	2	7'092	10	50
22	4	1	7'383	9	52
23	1	3	7'716	8	51
24	10	2	6'810	1	63
25	3	1	8'506	18	54
26	7	1	5'073	0	79
27	7	2	9'331	6	63
28	8	2	9'516	10	65
29	9	5	6'495	7	64
30	1	2	7'968	19	74

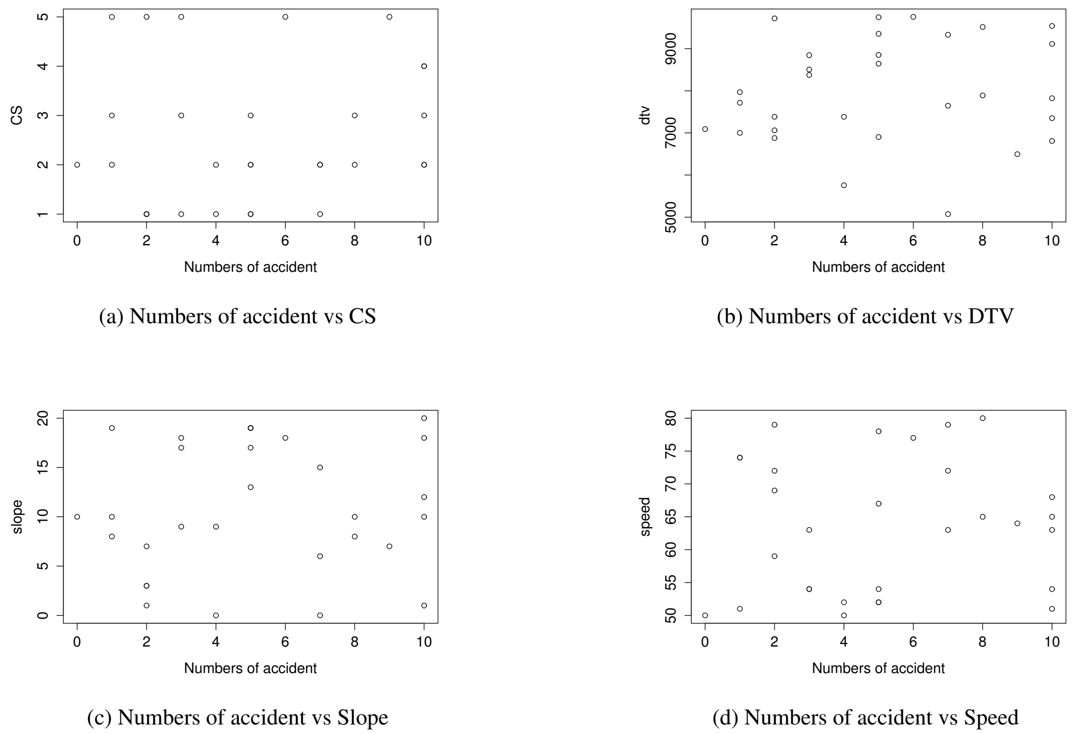


Fig. 6.3: Scatter plot of numbers of accident vs observed characteristic variables

6.2.4.1 Question A

Use the least squared method to obtain the model's parameter?

6.2.4.2 Answer to question A

The linear model shown in equation (4.13) can be simplified as follows

$$No.of accident = 1 \cdot \beta_0 + CS \cdot \beta_1 + DTV \cdot \beta_2 + Slope \cdot \beta_3 + Speed \cdot \beta_4 + \varepsilon \tag{6.23}$$

Using the R code provided in the Appendix 6.B, results are obtained and shown in listing 6.1.

6.2.4.3 Question B

Use the Bayesian method to obtain the model's parameter? and give a discussion on the values of model's parameters in comparison with that values obtained by using the least squared method.

```
Call:
lm(formula = noaccident ~ CS + dtv + slope + speed, data = data, x = TRUE, y = TRUE)

Residuals:
    Min      1Q  Median      3Q     Max
-4.7784 -2.5178 -0.3127  2.7137  5.4096

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 2.3884110  5.8412861   0.409    0.686
CS          0.2620322  0.4604536   0.569    0.574
dtv         0.0001349  0.0006295   0.214    0.832
slope       0.0275909  0.1193644   0.231    0.819
speed       0.0116089  0.0630220   0.184    0.855

Residual standard error: 3.382 on 25 degrees of freedom
Multiple R-squared:  0.02812, Adjusted R-squared:  -0.1274
F-statistic: 0.1808 on 4 and 25 DF,  p-value: 0.9462
```

Listing 6.1: Results of least squared fit method

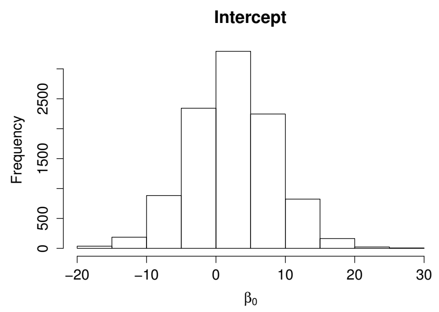
6.2.4.4 Answer to question B

Using the model presented in subsection 6.2 with 10'000 numbers of iterations, values of models parameters can be obtained. As this process involves highly computational efforts, the R code shown in the Appendix 6.B is used.

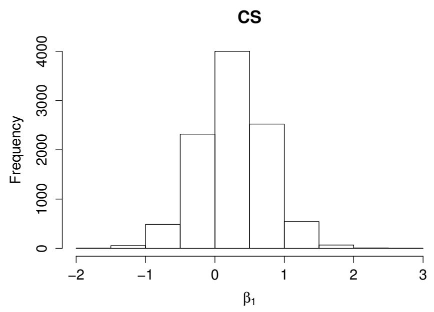
A summary of the mean values of model's parameters are shown in listing 6.2. This listing also presents the values of quantiles at the boundaries 5% and 95%.

As can be seen in Figure 6.4, after 10'000 iterations of sampling, the distributions of model's parameters become normal distributed with their mean values almost equal to that values obtained by the least squared method in the previous section. This is exactly the highlight of difference between the traditional frequentist estimation technique and Bayesian technique for a simple example.

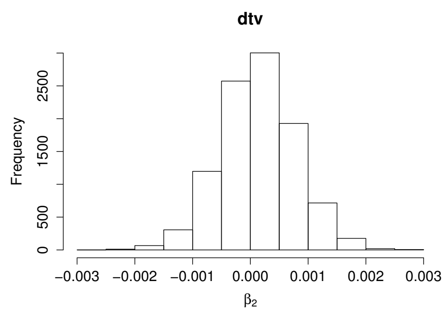
With respect to the statistical inferences on the values of model's parameters, it can be state that all characteristic variables are positively attributed to the occurrence of accident. However, contributions are different among them. It is likely that the condition state exposes to have the highest impact on the occurrence of accident as its value is 0.27, which is far more greater than that of other characteristic variables. The slope is the second important factor contributing to the occurrence of accident (value of 0.0255). Interesting, the DTV seems to have less impacts on the occurrence of accident for this road link.



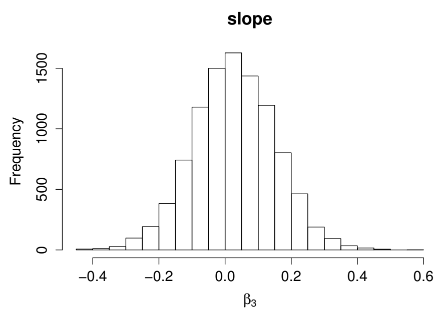
(a) Intercept



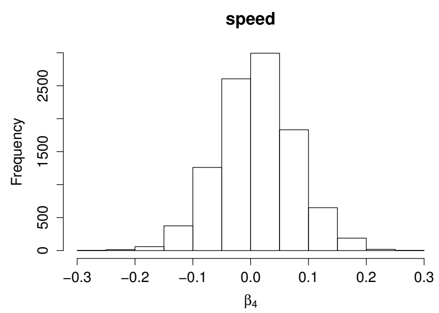
(b) Condition state



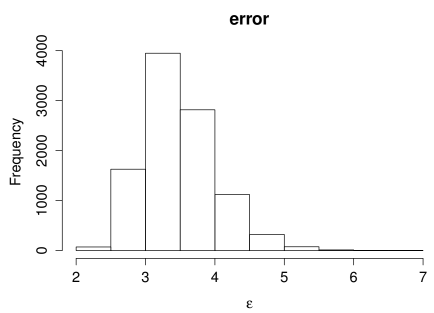
(c) Daily traffic volume



(d) Slope



(e) Speed



(f) Error term

Fig. 6.4: Distribution of models parameters

```
RESULTS OF BAYESIAN ESTIMATION
Confident interval
    X(Intercept)        XCS          Xdtv      Xslope      Xspeed Error
5%     -7.721114 -0.5087041 -0.0009349268 -0.17878857 -0.09732576 2.766584
50%     2.171526  0.2748703  0.0001513297  0.02557782  0.01239987 3.429186
95%    12.147611  1.0522804  0.0012215059  0.23095838  0.12003851 4.432573
```

Listing 6.2: Confident interval

6.3 Bayesian inference for a simple Weibull model

In sub-section 2.2.8 of Chapter 2, we learn that failure rate can be time-variant, and so does the reliability of an item. This type of time-variant failure and reliability models is commonly observed in reality. To address this time-variant characteristics of failure and reliability, a commonly used model is the Weibull proportional model.

In Weibull proportional model, the reliability function, density function, and failure rate, are shown in equations (2.14), (2.15), and (2.16), respectively. For convenience of reading, they are rewritten as follows:

Reliability function

$R(t) = e^{-\alpha \cdot t^m}$

Probability density function (p.d.f)

$f(t) = \alpha \cdot m \cdot t^{m-1} \cdot e^{-\alpha \cdot t^m}$

Failure rate

$\lambda(t) = \alpha \cdot m \cdot t^{m-1}$

where α and m are scale and shape parameters, respectively. These are parameters that need to be estimated given data.

6.3.1 Maximum likelihood estimation for the Weibull model

In order to estimate the values of the models parameters α and m, the maximum likelihood estimation (MLE) can be used. The MLE is used extensively in frequentist statistic as a de-factor estimation method.

Let assume that the vector of parameter and data are θ and $\bar{t}$, respectively. The likelihood function has the following general form.

$$L(t_1, t_2, \cdots, t_N, \theta) = \prod_{n=1}^{N} f(t_n, \theta) \tag{6.24}$$

In this equation, f represents the density function of a random variable t and characterized by θ. n is the sample and N is total numbers of samples.

The likelihood function (6.24) can be expressed in following form taking into consideration of data.

$$L(t_1, t_2, \cdots, t_N, \theta) = \prod_{n=1}^{N} [f(t_n, \theta)]^{\delta_n} \cdot [1 - F(t_n, \theta)]^{1-\delta_n} \tag{6.25}$$

In equation (6.25), $F(t_n, \theta)$ is the cumulative distribution function (c.d.f). Obviously, the second polynomial inside the product is the reliability function $(R = 1 - F)$ that we learn in Chapter 2. δ_n is dummy variable taking its value of 1 when a sample (or item) has been observably failed and 0 otherwise.

Using the Weibull function for the p.d.f and the c.d.f, we can expressed equation (6.25) as follows

$$L(\bar{t}, \theta) = \prod_{n=1}^{N} [\alpha \cdot m \cdot t_n^{m-1} \cdot exp(-\alpha \cdot t_n^m)]^{\delta_n} \cdot [exp(-\alpha \cdot t_n^m)]^{1-\delta_n} \tag{6.26}$$

The logarithmic form of the likelihood function will be

$$ln[L(\bar{t},\theta)] = \sum_{n=1}^{N} \delta_n \cdot [ln(\alpha) + ln(m) + (m-1)ln(t_n) - \alpha \cdot t^m] + \sum_{n=1}^{N} (1-\delta_n) \cdot [-\alpha \cdot t^m] \tag{6.27}$$

In order to obtain models parameters, analytically, following derivative has to be solved for each parameter.

$$\frac{\partial ln[L(\bar{t},\theta)]}{\partial \theta_i} = 0 \tag{6.28}$$

where the index i represent the running index of models parameters. It is important to note that the scale parameter α can be further expressed as a function of characteristic variables in a linear form such as

$$\alpha = \sum_{k=1}^{K} x_k \cdot \beta_k \tag{6.29}$$

where, k is the index of characteristic variable such as traffic volume, ambient temperature, etc and β_k turns to be the models parameters associated with each characteristic variable x_k.

6.3.2 Bayesian inference

Using the Bayes law as shown in equation (6.1), we can express the posterior p.d.f of the model as follows

$$p(\alpha, m|data) = \frac{L[\bar{t}|\alpha, m] \cdot p(\alpha) \cdot p(m)}{\int_0^\infty \int_0^\infty L[\bar{t}|\alpha, m] \cdot p(\alpha) \cdot p(m) d\alpha dm} \tag{6.30}$$

where $p(\alpha)$ and $p(m)$ are prior distributions for α and m, respectively. As can be seen from this formula, the prior distributions for for α and m infer that we can make use of knowledge that we might have known for these two parameters. This turns out to be extremely useful when dealing with small set of data. The choice of using which prior distribution to fit with α and m can be varied depending on each modeler and knowledge gained from practice. This section leave out a detailed mathematical formulation of Bayesian inference for the Weibull model as it involves a great number of mathematic ingredients. For curious readers, who expect to fully capture the formulation of the model, research works of Dodson [2006], Nassar and Eissa [2005], and Soliman et al. [2006] are recommended.

6.3.3 Example: Underground pipeline deterioration

A water distribution network of a mega urban city encompasses a large numbers of pipelines, which were laid underground. Pipelines were constructed in different time. Overtime, pipelines deteriorate, leakages or cracks might occur that lead to the lost of water in different part of the city. In some cases, pipelines were broken suddenly, resulting in a great amount of negative impacts to stakeholders (e.g. owners, users, directly affected public, and indirectly affected public). As pipelines were laid underground, it is not easy to perform frequent inspections. Moreover, inspection activities are also expensive. Most of data to date that is available for analysis is failure of pipelines over the entire city. The data is described in Table 6.3.

As can be seen in the table, there are 26 segments of pipelines in total. The column “Time” indicates the time in time units (tus) that each segment has been used since it was constructed. The column “Status” indicates the failure status

Table 6.3: Data on pipeline failure p

Segment	Time (tus)	Status	Water lost (normalized)	Depth (cm)
1	156	1	0.5283	40
2	1040	0	0.9542	45
3	59	1	0.6773	47
4	421	0	0.7279	41
5	329	1	0.1481	42
6	769	0	0.4236	47
7	365	1	0.4951	47
8	770	0	0.9857	42
9	1227	0	0.1728	48
10	268	1	0.0712	42
11	475	1	0.5424	41
12	1129	0	0.1673	43
13	464	1	0.4240	50
14	1206	0	0.9539	50
15	638	1	0.2992	48
16	563	1	0.0558	46
17	1106	0	0.3183	42
18	431	1	0.4762	44
19	855	0	0.4101	42
20	803	0	0.9394	45
21	115	1	0.9573	43
22	744	0	0.8402	46
23	477	0	0.3166	47
24	448	0	0.7257	46
25	353	1	0.6178	41
26	377	0	0.9832	45

of the pipe segment to date. The status receives value of 1 and 0. If the segment of pipe has failed, it receives value 1 and if not, the value is 0. The column “Water lost” indicates a normalized water lost recorded before the failure time for segments have failed and before inspection time for segments that have not failed. The last column indicates the depth in cm of pipe underneath the road surface.

6.3.3.1 Question

Use the MLE method and Bayesian method to estimate the models parameters assuming that the time to failure is dependent on water lost and depth. Interpret the meaning in the values of models parameters.

6.3.3.2 Answer

Using the MLE method, following results are obtained. The results (listing 6.3) were obtained by using the R code in the Appendix 6.C

From this result, the scale parameter and shape parameter of the distribution can be obtained [3]

$\alpha = 2.7736 + 0.8820 \times mean(waterlost) + 0.0873 * mean(depth)$

$\alpha = 2.7736 + 0.8820 \times 0.5466 + 0.0873 * 44.61538 = 7.1110381$

$m = 1/Scale = 1/0.885 = 1.129402$

[3] Scale and shape parameter of R function for Weibull model is described in https://stat.ethz.ch/R-manual/R-devel/library/survival/html/survreg.html

```
Call:
survreg(formula = Surv(time, status) ~ 1 + waterlost + depth,
    dist = "weibull")
              Value Std. Error      z     p
(Intercept)  2.7736     4.2633  0.651 0.515
waterlost    0.8820     0.9088  0.971 0.332
depth        0.0873     0.0961  0.908 0.364
Log(scale)  -0.1217     0.2520 -0.483 0.629

Scale= 0.885

Weibull distribution
Loglik(model)= -96.9   Loglik(intercept only)= -98
  Chisq= 2.09 on 2 degrees of freedom, p= 0.35
Number of Newton-Raphson Iterations: 5
n= 26
```

Listing 6.3: Confident interval

Using these parameters, the reliability function can be drawn

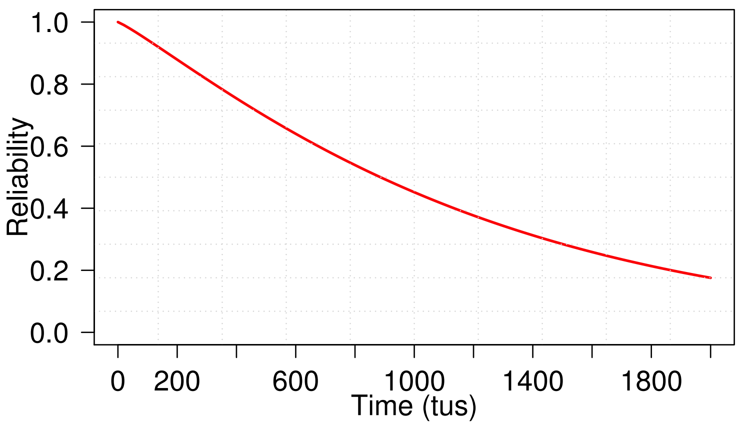


Fig. 6.5: Reliability over time

Using Bayesian method, the model's parameters can be summarized in distributions, with their means are the expected values of models parameters (Figure 6.6).

6.3.4 Conclusion

Using Bayesian statistics, once can solve many models that cannot be solved with traditional frequentist statistics. These models allow us to have better prediction results given incomplete data set. Examples of using Bayesian statistics in the field of infrastructure asset management are not so many. Some of the recent research works in this direction are the works of Hong and Prozzi [2006], Gao et al. [2012], Kobayashi et al. [2012], Kobayashi et al. [2014], and Lethanh et al. [2015].

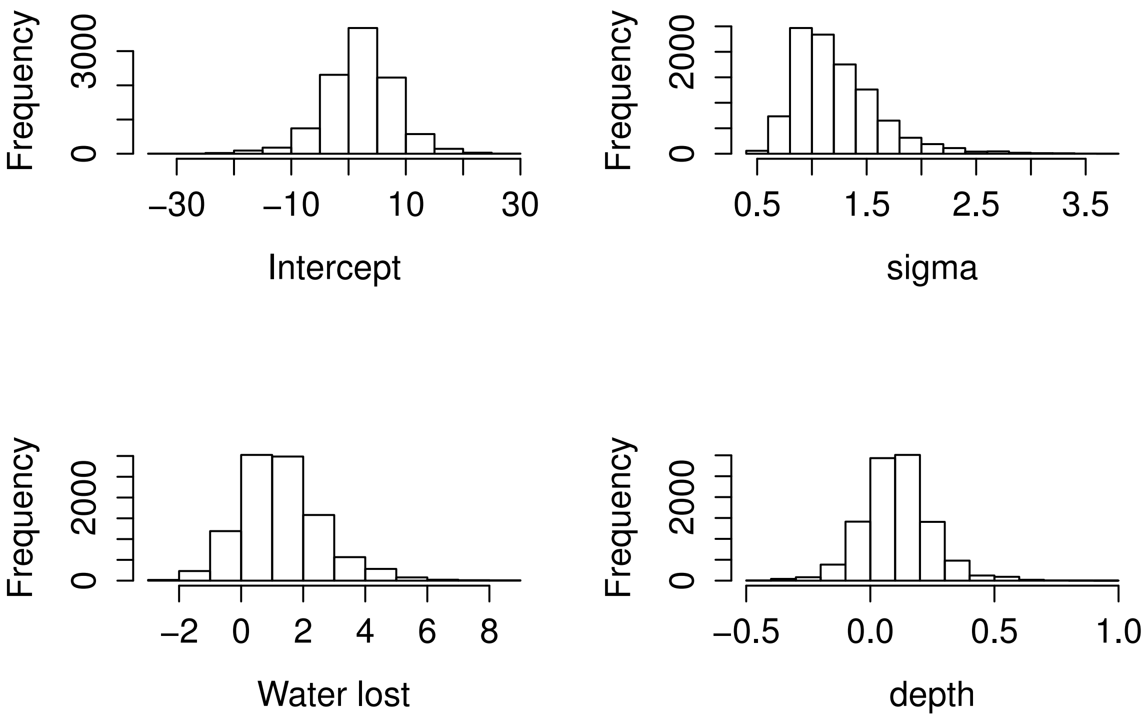


Fig. 6.6: Distribution of model parameters

6.4 Bayesian network model

Bayesian network modeling approach has recently gained its popularity in the fields of transportation engineering, structural analysis, and infrastructure asset management. It is a useful approach to quantify and measure risks related to transportation and infrastructure systems. Some examples of using Bayesian network models are with the works of Bayraktarli and Faber [2011] for quantifying risks occurred in a mega city after an earthquake; the works of Bateni et al. [2007] for prediction of scour deterioration process of bridge piers; the works of Schubert et al. [2012], Wang et al. [2013], and Deublein et al. [2013] for predicting accident rates on a transportation network; and the works of Straub and Kiureghian [2010] for analysing structural reliability of an infrastructure object such as concrete beams and columns in a building or piers and decks of a bridge.

The principle concept to formulate a Bayesian network model is with the Bayes law that has been described in section 6.1.1 with equation (6.1) and the graphical probabilistic method. There are many types of Bayesian network models, which really depend on the nature of observed data and assumptions of modelers.

This section aims to give a gentle introduction on Bayesian network modeling through a number of examples.

6.4.1 General concept

A Bayesian network (BN) is a graphical structure used to describe the interdependent of variables and parameters of a probabilistic model. The network with its notation $D = (V, E)$ is referred as a direct acyclic graph (DAG) where V and E representing vectors of finite sets of nodes and edges, respectively. The node $v \in V$ in a BN represent a random variable X_v, thus the entire node vector V represents a set of random variables $\mathbf{X} = X_1, X_2, \cdots, X_n$. The edge e (sometimes refer to as arcs or links) connects pairs of nodes X_i and X_j, i.e. $X_i \rightarrow X_j$, which represents the direct dependencies between variables.

In this context, the numbers of nodes are discrete and finite, thus the set of random variables is also discrete and finite. Each variable is characterized by a probability distribution. As nodes are undirectly or directly connected, they represent the dependencies. Thus, the probability of having one node (e.g. the event of a random variable to receive a value of x or y) is a conditional probability with its value dependent on other nodes. To represent the dependencies in a graphical model, two terminologies are used: “parent nodes” and “child nodes”. Parent nodes are nodes being considered as sources that lead to the occurrence of child nodes. A node, can be either “parent” or “child” node. For example, in Figure 6.7, node B is the parent node of node C but it is the child node of nodes A and D.

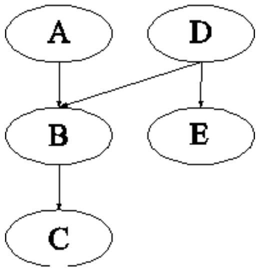


Fig. 6.7: An example of a simplified BN model

To understand how a BN model is structured, let go through a simple example

6.4.2 Introductory example - reliability of a bridge

It is observed that over the last 15 years a bridge located in a mountainous regions had to close a numbers of times due to its condition that was unable to provide adequate level of services (LOS). The inadequate LOS is defined by manager as the levels of reliability and it is classified into three categories: “low” is when the reliability is less than 0.6; “medium” is when the reliability is between 0.6 and 0.8; and “high” is when the reliability is greater than 0.8. The reliability can be measured by performing stress tests.

Not providing adequate LOS might be due to latent process such as avalanche, rockfall, floods, and its manifest deterioration such as cracking of bridge decks, weakness of its bearing system, and bad quality of joints. The latent deteriorate process can be triggered by an earthquake, amount of snow accumulated in the starting zone of avalanches, and volume of run-off due to rains.

In each year, discrete values of these six variables were recorded.

- **Avalanche**: the avalanche, recorded as “yes” if there was at least one occurrence of avalanches in that year and the avalanches came to contact with the bridge, and “no” if there was no avalanche occurred.
- **Rockfall**: the rockfall, recorded as “yes” if there was at least one occurrence of rockfalls in that year and the rockfalls came to contact with the bridge, and “no” if there was no rockfall occurred.
- **Floods**: the floods, recorded in three categories: “low”, “medium”, and “high” indicating levels of floods happened.
- **Cracks of bridge decks**: the cracks appeared on bridge decks indicates serious problems on the deterioration of the decks. The cracks were recorded in three categories: “small” is when the percentage of cracks is less than 5%

in the total amount of bridge decks areas; "moderate" is when the percentage of cracks is between 5% and 10% and "high" is when the percentage of cracks is higher than 10%.

- **Weakness of the bearing system**: the bearing system is made of rubber pads. The condition of the system is recorded as "good" and "bad".
- **Quality of joints**: Quality of joints were recorded as "good" and "bad", which indicates the condition state of joint. It is good when the joint has no defect. It is bad when either top or bottom layers or both of the two layers are in bad condition.
- **Earthquake**: the earthquake is recorded as "yes" or "no".
- **Snow**: the snow is recorded as "low" and "high" depending on the amount of snow accumulated on the starting zone of avalanches
- **Rain**: the rain is recorded as "low", "medium", and "high" depending on the volume of run-off recorded at the weather station nearby the bridge.
- **Daily traffic volume (DTV)**: the daily traffic volume was recorded in numbers but it can also be categories into 3 discrete scales: "low", "medium", and "high".

The reliability of the bridge is dependent on both latent process and manifest process. Therefore, by study the structure of this system, it might be useful to calculate the reliability of bridge more accurate and eventually beneficial to come up with better management for the bridge.

6.4.3 Graphical representation and values of nodes

The way the variables were recorded in the inspection data, and more in general in two categories (latent and manifest) that each variable belongs to, infers the interdependency among them. Some of them have direct relationship such as the reliability of the bridge is directly related to a rockfall; and some of them have indirect relationship such as the reliability of the bridge is indirectly related to the volume of run-off.

Both these kinds of relationship can be graphically modeled using a direct graph as earlier mentioned in section 6.4.1. In the graph, nodes (V) and edges (E) are used to systematically represent the relationship. In a very simple representation, a partial graphical representation of the network looks like the one shown in Figure 6.8.

This figure infers that the reliability of the bridge is dependent on two factors: avalanche and rockfall. Those factors are dependent on natural hazard such as earthquake. Also in the figure, there are conditional probabilities table (cpt). Values shown in the tables indicate the probabilities of that variables (or event) receiving value "yes" or "no". For example, if there is a earthquake, there is 40% chance that there will be rockfall. Certainly, the probability for the bridge not providing adequate LOS is conditional dependent on the probability of occurring avalanches or rockfall, which, in turn, dependent on the probability of earthquake occurrence.

With this rule of conditional probability in a chain, the joint probability of all nodes in this BN can be computed by using following formula.

$$Pr(E,A,R,B) = Pr(E) \times Pr(A|E) \times Pr(R|E) \times Pr(B|E,A,R) \quad (6.31)$$

$$Pr(E,A,R,B) = Pr(E) \times Pr(A|E) \times Pr(R|E) \times Pr(B|A,R) \quad (6.32)$$

where, E, A, R, and B stand for Earthquake, Avalanche, Rockfall, and Bridge.

Let say, we are now interested in calculating the likelihood that Avalanche will causes the bridge to be in condition that it will not provide an adequate LOS. Using the Bayes rule as shown in (6.1), we can formulate the posterior probability

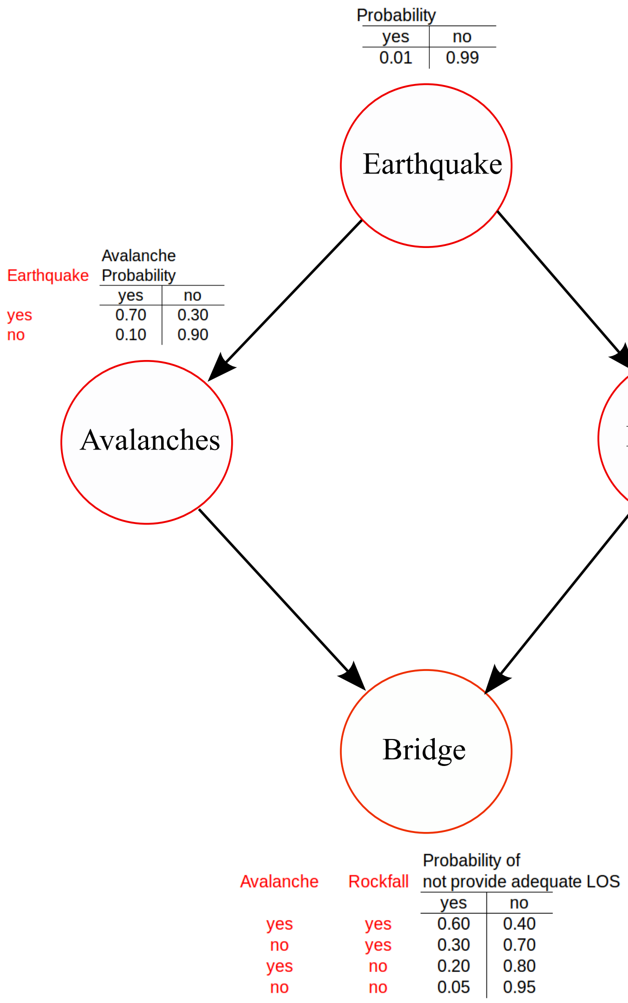


Fig. 6.8: Partial representation of the Bayesian network for the bridge - Latent process

for the avalanche.

$$Pr(A = yes|B = yes) = \frac{Pr(A = yes, B = yes)}{Pr(B = yes)} = \frac{\sum_{E,R} Pr(E, A = yes, R, B = yes)}{Pr(B = yes)} \tag{6.33}$$

The nominator $\sum_{E,R} Pr(A = yes, B = yes, E)$ infers that in order to calculate this probability, we have to consider the probability of earthquake occurrence and the probability of rockfall as well. Using equation (6.32), the nominator in equation (6.33) becomes

$$\sum_{E,R} Pr(E, A = yes, R, B = yes) =$$

$$= Pr(E = yes, A = yes, R = yes, B = yes) \tag{6.34}$$
$$+Pr(E = no, A = yes, R = yes, B = yes) \tag{6.35}$$
$$+Pr(E = yes, A = yes, R = no, B = yes) \tag{6.36}$$
$$+Pr(E = no, A = yes, R = no, B = yes) \tag{6.37}$$
$$= Pr(E = yes) \times Pr(A = yes|E = yes) \times Pr(R = yes|E = yes) \times Pr(B = yes|A = yes, R = yes) \tag{6.38}$$
$$+Pr(E = no) \times Pr(A = yes|E = no) \times Pr(R = yes|E = no) \times Pr(B = yes|A = yes, R = yes) \tag{6.39}$$
$$+Pr(E = yes) \times Pr(A = yes|E = yes) \times Pr(R = no|E = yes) \times Pr(B = yes|A = yes, R = no) \tag{6.40}$$
$$+Pr(E = no) \times Pr(A = yes|E = no) \times Pr(R = no|E = no) \times Pr(B = yes|A = yes, R = no) \tag{6.41}$$

$$\sum_{E,R} Pr(E, A = yes, R, B = yes) =$$
$$\begin{aligned}&= 0.01 \times 0.70 \times 0.40 \times 0.60\\ &+0.99 \times 0.10 \times 0.05 \times 0.60\\ &+0.01 \times 0.70 \times 0.60 \times 0.20\\ &+0.99 \times 0.10 \times 0.95 \times 0.20\\ &= 0.0243\end{aligned}$$

Similarly using the (6.31) for extending the formulation of the denominator of equation (6.32), its value is obtained

$$Pr(B = yes) = \sum_{E,A,R,B} Pr(E, A, R, B = yes) = 0.0804375 \tag{6.42}$$

thus the likelihood of avalanche to be occurred if the bridge is not providing adequate LOS is

$$Pr(A = yes|B = yes) = \frac{Pr(A = yes, B = yes)}{Pr(B = yes)} = \frac{0.0243}{0.0804375} = 0.302097$$

From this example, it is clear that the joint probability of variables in a BN model can be computed based on the rule of conditional probability. In general mathematical expression, that rule is described as follows

$$Pr(V) = \prod_{v \in V} Pr(v|pa(v)) \tag{6.43}$$

The notation pa infers the parent nodes of node v and $Pr(v|pa(v))$ is a function defined on $(v, pa(v)$. This function satisfies the condition that the sum of probability over a space equals to 1, i.e. $\sum_{v*} Pr(v = v * |pa(v)) = 1$. It can be seen that a set values of $Pr(v|pa(v))$ is actually the conditional probabilities table (the cpt in short). In this view, it can be said that a BN is a complex stochastic model built up by forming together simple components. In other words, it is referred as hierarchical Bayesian model.

The graph and result of this example can be easily drawn and calculated by hand. For convenient of readers, R code is provided in Appendix 6.E

The above example is a simplest assumption possible that we can think of. In the example, only latent deterioration process is considered. However, in practice, latent deterioration is not only the cause of deterioration. Manifest deterioration should also be considered in calculating the reliability of the bridge. In addition, both latent and manifest processes can be further expanded (like the event tree). A more realistic BN for the same example can be visualized in Figure 6.9.

Figure 6.9 shows that conceptually a BN model can be expanded to include many layers of variables possible. Also, values describing each variable can be also continuous and be assumed to follow a parametric distribution (e.g. Poisson, Weibull, Gamma) depending on the behaviors of data.

The calculation for this BN becomes more difficult than the earlier one as the size of the cpt exponentially growth with numbers of variables. The R code in Appendix 6.F is used to generate the BN shown in Figure 6.9. It includes also a random generated cpt for each node of the BN model and perform calculation of marginal probability for each node. Noted that in this extended example, the reliability of the bridge is recorded in three categories “low”, “medium”, and “high” instead of of only two states in original example.

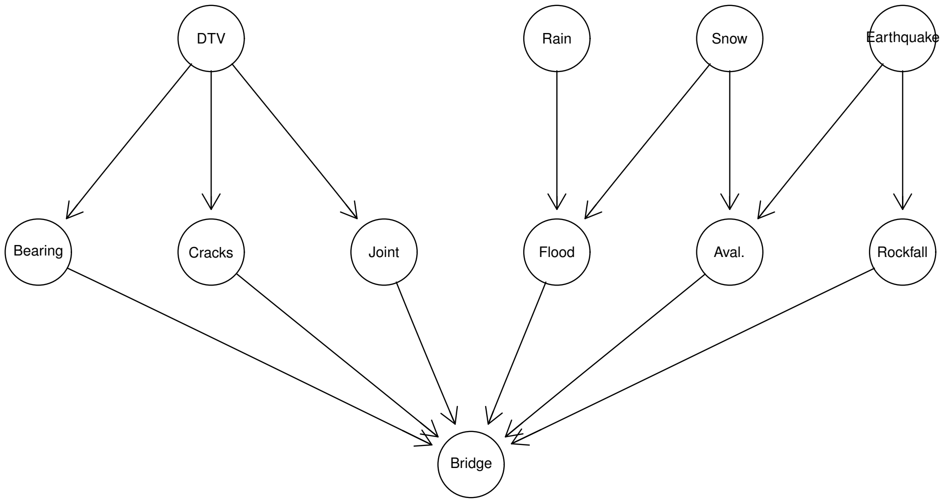


Fig. 6.9: Partial representation of the Bayesian network for the bridge - Latent process

6.4.4 Estimating the cpt from data

In earlier section, the cpt is presented along with the graph of the BN models. Values of the properties in the cpt can be estimated from various source of information such as expert opinions, simulation, or data. Estimation the cpt from data is common in most cases. Typically, data is stored in various extension files (e.g. txt, tex, csv) or in database systems (e.g. MySQL, MsAccess, PostgreSQL). An example of how data looks like for the above example is listed in Table 6.4 and Table 6.5.

Table 6.4 shows a simplified data that includes only 4 variables (Earthquake, Rockfall, Avalanche, Bridge) and for each variable, there is only two discrete value "yes" and "no". The conditional probability for each variable can be computed using the data shown in this table. Since values of variables are discrete, one way to do it is to use following equation

$$Pr(B = yes|R = yes) = \frac{Pr(B = yes, R = yes)}{Pr(R = yes)} \tag{6.44}$$

$$\frac{\text{Number of observations for which B=yes and R=yes}}{\text{Number of observations for which R=yes}} \tag{6.45}$$

This equation yields exact the conditional probability for this event to occur. This is corresponding to the traditional *frequentist* and *maximum likelihood* estimates.

The R codes for computing the cpt based on these two tables are given in Appendix 6.E and 6.F and the estimated cpt for the data shown in Table 6.4 and Table 6.5 are given in Appendix 6.G and 6.H, respectively.

6.4.5 Forming a BN model from data and the curse of dimensionality

In the above example, it is given that the structure of the BN model is known. That is we know the relationship between nodes in the graphical model. In many practical situations, data is collected that includes many different

Table 6.4: Simplified observed data

Earthquake	Avalanche	Rockfall	Bridge
yes	yes	no	yes
yes	no	no	yes
no	no	no	yes
no	no	no	no
no	no	no	no
no	no	yes	no
yes	no	no	no
no	no	no	no
no	yes	yes	no
yes	no	yes	no
no	yes	no	yes
no	no	no	no
no	yes	yes	no
no	yes	no	yes
no	no	no	yes
yes	yes	yes	no
no	yes	no	yes
no	no	yes	no
no	yes	yes	no
yes	no	yes	no
no	no	no	yes
no	no	yes	no
yes	yes	yes	yes
yes	no	no	yes
yes	yes	yes	no
no	no	no	no
no	no	no	yes
yes	no	yes	yes
yes	yes	no	yes

Table 6.5: Extended observed data

Earthquake	Snow	Rain	DTV	Aval	Rockfall	Flood	Cracks	Bearing	Joint	Bridge
yes	high	high	low	yes	no	no	moderate	good	yes	high
yes	low	high	high	no	no	yes	moderate	good	no	high
no	low	medium	high	no	no	no	high	good	no	high
no	low	low	low	no	no	no	high	good	no	medium
no	high	low	low	no	no	no	high	bad	yes	low
no	high	low	high	no	yes	yes	high	bad	yes	low
yes	high	medium	low	no	no	yes	moderate	good	yes	low
no	low	low	low	no	no	yes	moderate	good	no	medium
no	low	low	medium	yes	yes	yes	high	bad	no	medium
yes	low	high	medium	no	yes	no	moderate	bad	yes	medium
no	low	low	medium	yes	no	yes	moderate	bad	no	high
no	high	medium	low	no	no	yes	moderate	good	yes	low
no	low	high	high	yes	yes	no	small	bad	no	medium
no	high	medium	low	yes	no	yes	high	good	yes	high
no	low	low	high	no	no	yes	moderate	bad	yes	high
yes	high	low	low	yes	yes	no	small	good	yes	low
no	high	low	low	yes	no	yes	small	good	no	high
no	low	high	high	no	yes	no	moderate	good	no	low
no	low	low	low	yes	yes	yes	moderate	bad	yes	medium
yes	high	low	high	no	yes	yes	small	bad	yes	low
no	low	medium	low	no	no	yes	small	good	yes	high
no	high	medium	medium	no	yes	no	small	bad	yes	low
yes	high	low	medium	yes	yes	no	moderate	bad	no	high
yes	high	medium	high	no	no	yes	moderate	bad	yes	high
yes	high	medium	medium	yes	yes	no	high	good	yes	medium
no	high	high	high	no	no	yes	high	bad	yes	low
no	high	low	low	no	no	no	small	good	yes	high
yes	high	high	medium	no	yes	no	high	bad	no	high
yes	low	high	high	yes	no	no	high	bad	no	high

types of variables and attributes that we do not precisely know their relationship. Fortunately, given the data, it is possible to learn from the behaviors of data in order to create a set of BN models that suit the data the most. Following example of data on accident risks illustrate the idea of how to form a BN model from data.

A city transportation department wants to understand the accident risks that occurs on the city road network when three intervention strategies (IS)[4] is pilot tested. These pilot tests can be executed in different parts of the city network. The effectiveness of these tests are measured by observing the outcome as the numbers of accidents occurs in two consecutive weeks. A description of the data is given in following table 6.6.

Table 6.6: Data on pipeline failure p

Time	IS	week 1	week 2
day	1	5	6
day	1	7	6
day	1	9	9
day	1	5	4
day	2	9	12
day	2	7	7
day	2	7	6
day	2	6	8
day	3	14	11
day	3	21	15
day	3	12	10
day	3	17	12
night	1	7	10
night	1	8	10
night	1	6	6
night	1	9	7
night	2	7	6
night	2	10	13
night	2	6	9
night	2	8	7
night	3	14	9
night	3	14	8
night	3	16	12
night	3	10	5

The network has its primary property as node as graphically represented in Figure 6.10. The cell in the table represents each entry or node of the network. Each entry contains information attached with each node (e.g. day or night, different numbers of accidents).

In this simple network, the input nodes are the time (day or night) and the intervention strategies (IS), and the outcome are the numbers of accident occurs in each consecutive weeks. The infrastructure manager wants to identify the probability that the accident occurs due to time and intervention strategies. In addition, as the week 2 is after week 1, the node representing week 1 can also become a parent node of week 2, i.e. the likelihood of accidents occurs on week 2 is dependent also on the numbers of accidents that have had occurred in week 1.

It is assumed that the joint probability for this problem is Gaussian distribution (normal distribution). The Bayesian network allows to compute such joint probability for every combination of nodes. For example, the joint probability of combination $(Time \to IS \to W1 \to W2)$ or $(Time \to W1)$ or $(IS \to W1 \to W2)$.

Following graphs show the outcome of running a Bayesian network for this example (the R code used to compute the example is given in Appendix 6.I).

As can be seen from the figure Figure 6.11 to Figure 6.16, there are many possible combination of nodes to form a network. The arrow in each of the graph represents the dependency. In total, with only 4 nodes, there are 144 networks being analyzed.

[4] Intervention strategies can be the setup of traffic flows or the installation of traffic signals

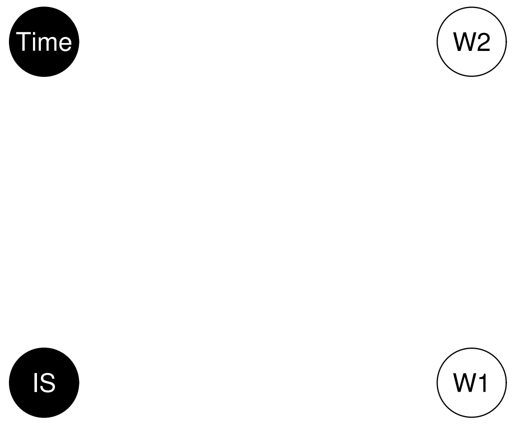


Fig. 6.10: Graphical representation of the network for accidents

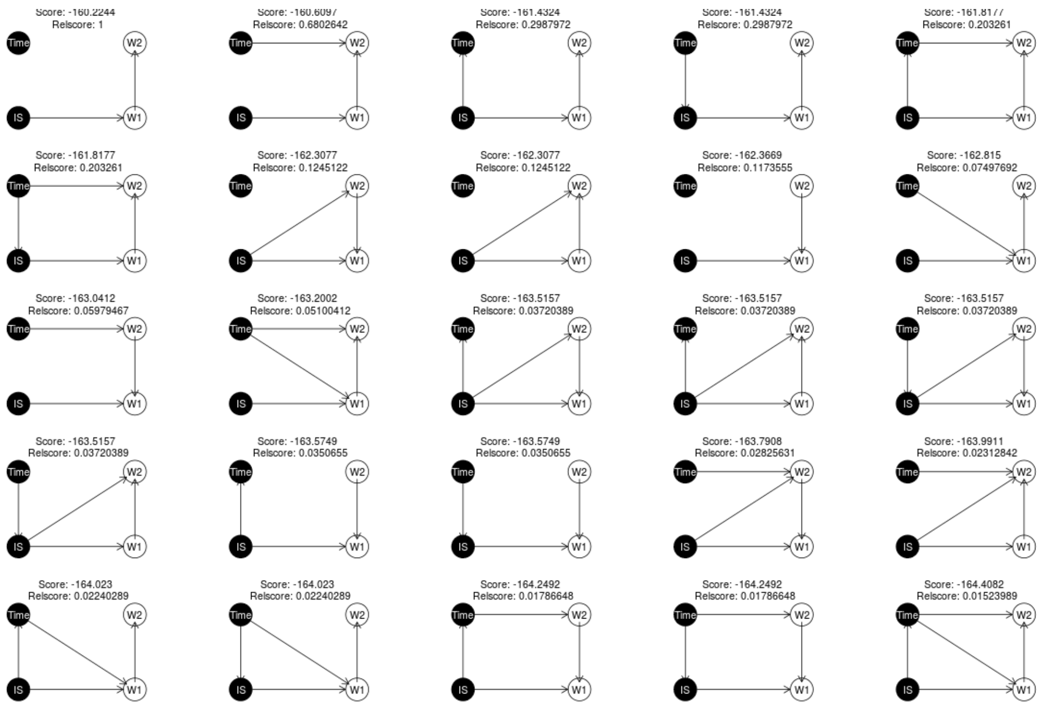


Fig. 6.11: Possible BN networks (1)

Table 6.7 lists a possible numbers of networks that can be formed given a certain numbers of nodes with their lower bound and upper bound values [Heckerman, 1996].

In these graphs, the statistic score $S(D)$ are displayed along with their rescaling score. The score is computed for every combination of nodes based on following equation

$$S(D) = p(D, data) = p(data|D) \cdot p(D) \tag{6.46}$$

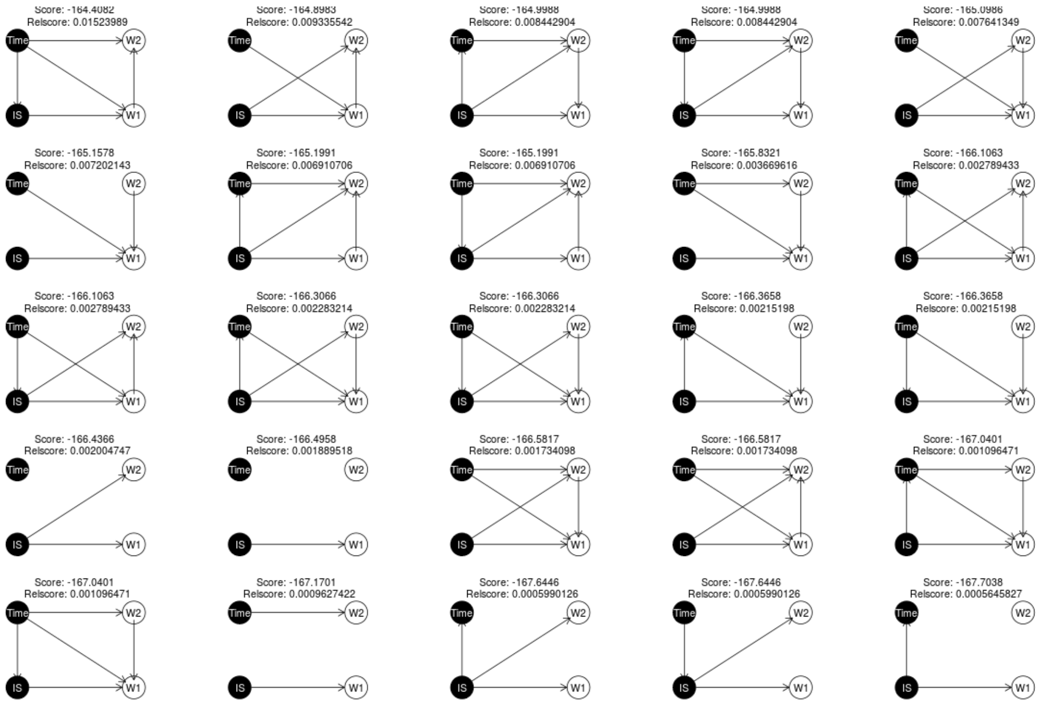


Fig. 6.12: Possible BN networks (2)

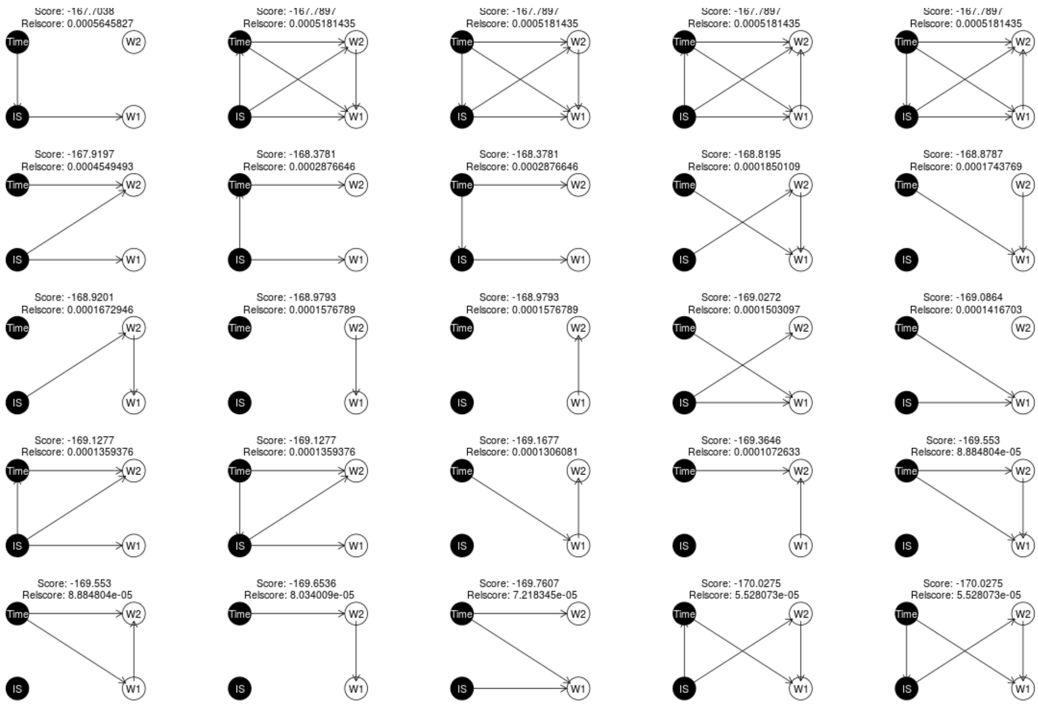


Fig. 6.13: Possible BN networks (3)

This score is used to assess the likelihood for that combination to be realized with certain probability. In this example, the network $IS \rightarrow W1 \rightarrow W2$ (Figure 6.17) will be most likely to be realized with highest probability. It means that the numbers of accidents that will occurs in week 2 will have high correlation with the intervention strategies and the numbers of accidents in week 1.

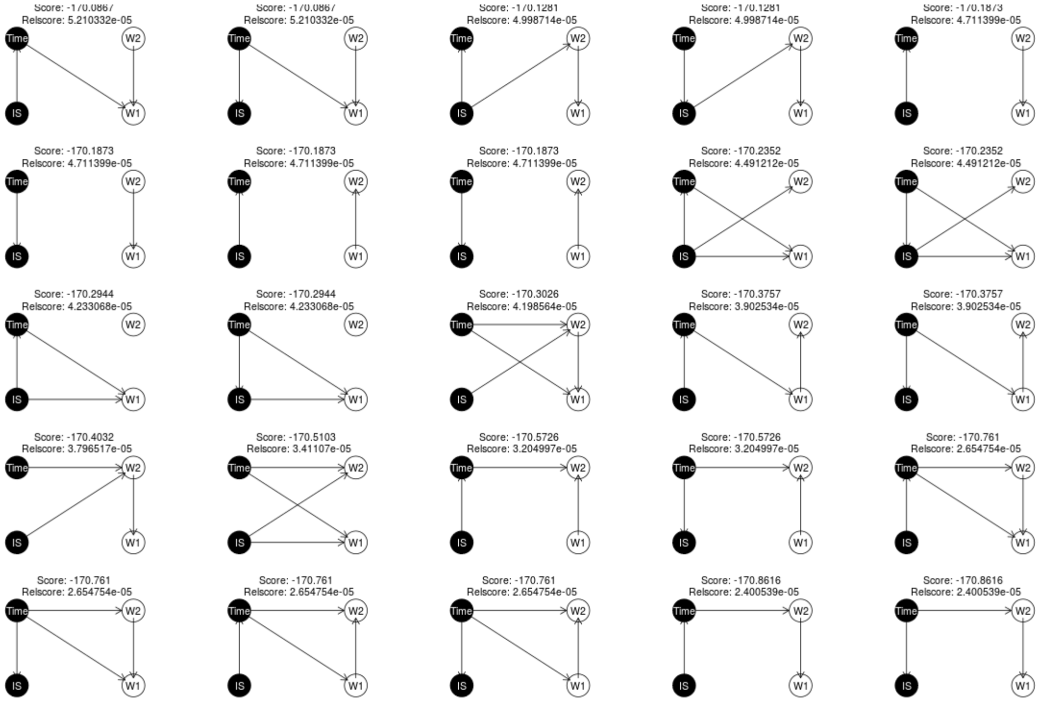


Fig. 6.14: Possible BN networks (4)

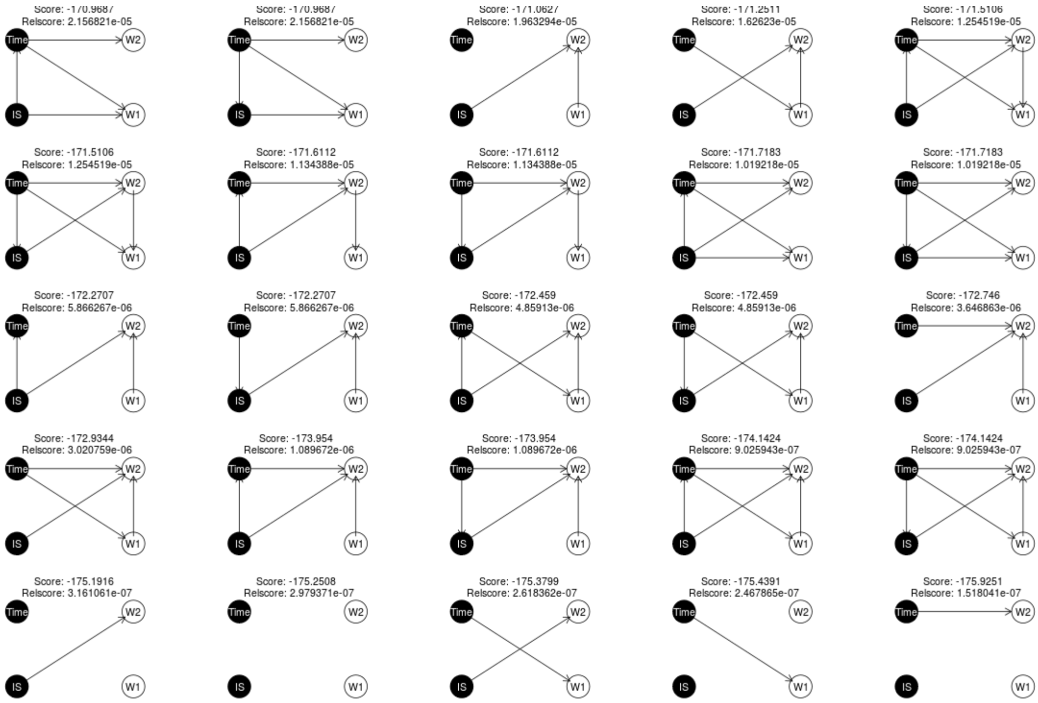


Fig. 6.15: Possible BN networks (5)

The above example discusses a general Bayesian network frame, in which, data model can be formulated with different combination and interaction. As can be shown, there are many scenarios that are not necessary to be analyzed for the final result. For example, the scenario that the arrow is only connecting between time and week 2. This scenario is not likely to give a good answer as it is known from engineering background that the IS should be involved. By using engineering background, the numbers of combinations between nodes can be reduced significantly.

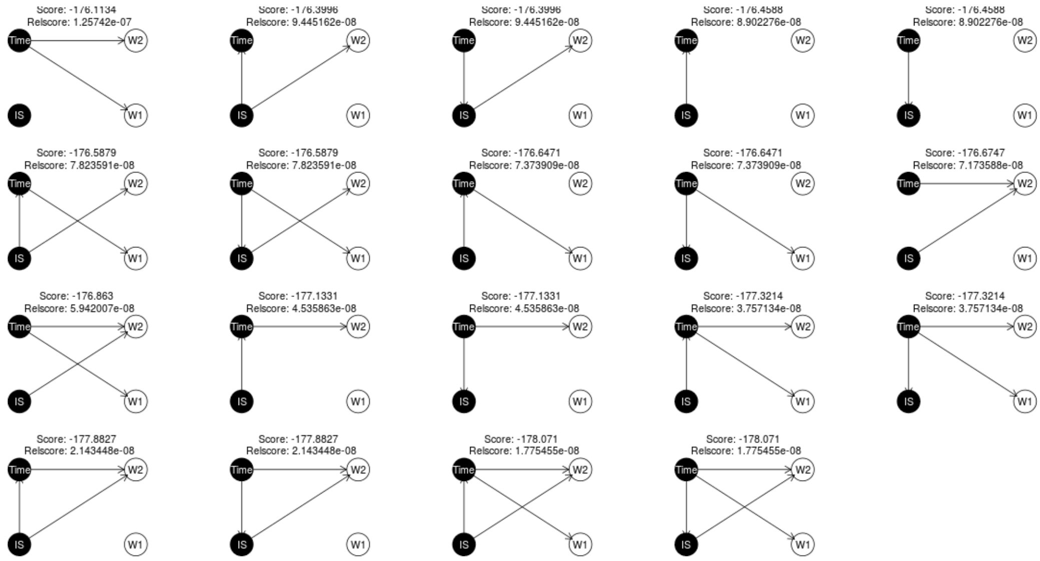


Fig. 6.16: Possible BN networks (6)

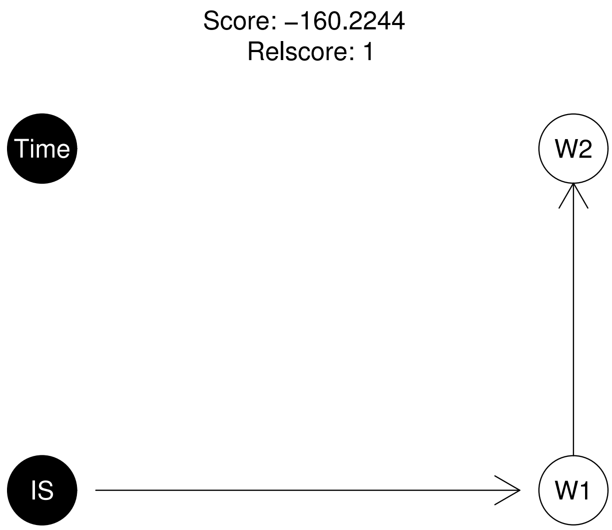


Fig. 6.17: The optimal network structure

Table 6.7: Numbers of networks

#Node	*#networks*
1	1
2	2-3
3	12-25
4	144-543
5	4’800-29’281
6	320’000-3’781’503

References

G. Andrew, B. C. John, S. S. Hal, and B. R. Donald. *Bayesian Data Analysis*. Chapman and Hall/CRC, 2006.

S. Mohyeddin Bateni, Dong-Sheng Jeng, and Bruce W. Melville. Bayesian neural networks for prediction of equilibrium and time-dependent scour depth around bridge piers. *Advances in Engineering Software*, 38(2): 102 – 111, 2007. ISSN 0965-9978. doi: http://dx.doi.org/10.1016/j.advengsoft.2006.08.004. URL http://www.sciencedirect.com/science/article/pii/S0965997806001177.

Yahya Y. Bayraktarli and Michael H. Faber. Bayesian probabilistic network approach for managing earthquake risks of cities. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 5(1):2–24, 2011. doi: 10.1080/17499511003679907. URL http://www.tandfonline.com/doi/abs/10.1080/17499511003679907.

William M. Bolstad. *Introduction to Bayesian Statistics*. Wiley, 2007.

Markus Deublein, Matthias Schubert, Bryan T. Adey, Jochen Khler, and Michael H. Faber. Prediction of road accidents: A bayesian hierarchical approach. *Accident Analysis & Prevention*, 51:274 – 291, 2013. ISSN 0001-4575. doi: http://dx.doi.org/10.1016/j.aap.2012.11.019. URL http://www.sciencedirect.com/science/article/pii/S0001457512004101.

Bryan Dodson. *The Weibull Analysis Handbook*. Quality Press, American Society for Quality, 2006.

Lu Gao, Jose Pablo Aguiar-Moya, and Zhanmin Zhang. Bayesian analysis of heterogeneity in modeling of pavement fatigue cracking. *Journal of Computing in Civil Engineering*, 26(1):37–43, 2012. doi: 10.1061/(ASCE)CP.1943-5487.0000114. URL http://link.aip.org/link/?QCP/26/37/1.

S. Geman and D. Geman. Stochastic relaxation, gibbs distributions and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984.

David Heckerman. A tutorial on learning with bayesian networks. Technical Report MSR-TR-95-06, Microsoft Research, Advanced Technology Division, Microsoft Corporation, 1996.

F. Hong and J.A. Prozzi. Estimation of pavement performance deterioration using bayesian approach. *Journal of Infrastructure System*, 12(2):77–86, June 2006.

J.G. Ibrahim, C. Ming-Hui, and D. Sinha. *Bayesian Survival Analysis*. Springer Series in Statistics, 2001.

G. Jeff. *Bayesian Methods: A Social and Behavioral Sciences Approach*. Chapman and Hall/CRC Taylor and Francis Group, 2 edition, 2006.

Kiyoshi Kobayashi, Kiyoyuki Kaito, and Nam Lethanh. A bayesian estimation method to improve deterioration prediction for infrastructure system with markov chain model. *International Journal of Architecture, Engineering and Construction*, 1(1):1–13, April 2012.

Kiyoshi Kobayashi, Kiyoyuki Kaito, and Nam Lethanh. A competing markov model for cracking prediction on civil structures. *Transportation Research Part B: Methodological*, 68(0):345 – 362, 2014. ISSN 0191-2615. doi: http://dx.doi.org/10.1016/j.trb.2014.06.012. URL http://www.sciencedirect.com/science/article/pii/S0191261514001192.

C. S Kumares and L. Samuel. *Transportation decision making: principles of project evaluation and programming*. John Wiley and Sons, 2007.

Nam Lethanh and Bryan T. Adey. Use of exponential hidden markov models for modeling pavement deterioration. *International Journal of Pavement Engineering*, 14(7):645–654, 2013. doi: 10.1080/10298436.2012.715647.

Nam Lethanh, Kiyoyuki Kaito, and Kiyoshi Kobayashi. Infrastructure deterioration prediction with a poisson hidden markov model on time series data. *Journal of Infrastructure Systems*, 21(3):04014051, September 2015. doi: 10.1061/(ASCE)IS.1943-555X.0000242. URL http://dx.doi.org/10.1061/(ASCE)IS.1943-555X.0000242.

M. M. Nassar and Fathy H. Eissa. Bayesian estimation for the exponentiated weibull model. *Communications in Statistics - Theory and Methods*, 33(10):2343–2362, 2005. doi: 10.1081/STA-200031447. URL http://dx.doi.org/10.1081/STA-200031447.

OpusCL. *Review of VOC - Pavement Roughness Relationships Contained in Transfunds Project Evaluation Manual*. Opus Central Laboratories, Opus, Lower Hutt, New Zealand, rep. 529277.00 edition, 1999.

C.P. Robert. *Mixtures of Distributions: Inference and Estimation, in: Gillks, W.R., Richardson, S. and Spiegelhalter, D.J. (eds.) Markov Chain Monte Carlo in Practice*. Chapman and Hall, 1996.

Matthias Schubert, Niels Peter Hj, Arild Ragny, and Harald Buvik. Risk assessment of road tunnels using bayesian networks. *Procedia - Social and Behavioral Sciences*, 48(0):2697 – 2706, 2012. ISSN 1877-0428. doi: http://dx.doi.org/10.1016/j.sbspro.2012.06.1239. URL http://www.sciencedirect.com/science/article/pii/S1877042812029825. Transport Research Arena 2012.

Ahmed A. Soliman, A.H. Abd Ellah, and K.S. Sultan. Comparison of estimates using record statistics from weibull model: Bayesian and non-bayesian approaches. *Computational Statistics & Data Analysis*, 51(3):2065

– 2077, 2006. ISSN 0167-9473. doi: http://dx.doi.org/10.1016/j.csda.2005.12.020. URL http://www.sciencedirect.com/science/article/pii/S016794730600003X.

Daniel Straub and Armen Der Kiureghian. Bayesian network enhanced with structural reliability methods: Methodology. *Journal of Engineering Mechanics*, 136(10):1248–1258, 2010.

Yan Fu Wang, Min Xie, Kwai-Sang Chin, and Xiu Ju Fu. Accident analysis model based on bayesian network and evidential reasoning approach. *Journal of Loss Prevention in the Process Industries*, 26(1):10 – 21, 2013. ISSN 0950-4230. doi: http://dx.doi.org/10.1016/j.jlp.2012.08.001. URL http://www.sciencedirect.com/science/article/pii/S0950423012001301.

6.A Example into Bayesian thinking

```
# This program was coded by  Nam Lethanh (lethanh@ibi.baug.ethz.ch)
#Bayesian statistic example - Accident risk analysis
library(LearnBayes) #this is dependency package
p = seq(0.05, 0.95, by = 0.1)
prior = c(1, 5.2, 8, 7.2, 4.6, 2.1, 0.7, 0.1, 0, 0)
prior = prior/sum(prior)
plot(p, prior, type = "h", ylab="Prior Probability")
data = c(11, 16)
post = pdisc(p, prior, data)
#post=round(cbind(p, prior, post),2)
print(post)
library(lattice)
PRIOR=data.frame("prior",p,prior)
POST=data.frame("posterior",p,post)
names(PRIOR)=c("Type","P","Probability")
names(POST)=c("Type","P","Probability")
data=rbind(PRIOR,POST)
print(xyplot(Probability~P|Type,data=data,layout=c(1,2), type="h",lwd=3,col="black"))
#THE END
cat("THE END")
```

6.B Bayesian estimation - Linear model

```
################################
# example of Bayesian estimation for a linear model
# Coded by Nam Lethanh (lethanh@ibi.baug.ethz.ch)
################################
library(LearnBayes) #requires R package to perform Bayesian estimation
data=read.csv("accidentdata.csv", header=TRUE) #Data
attach(data)
plot(noaccident,CS,xlab="Numbers of accident")
plot(noaccident,dtv,xlab="Numbers of accident")
plot(noaccident,slope,xlab="Numbers of accident")
plot(noaccident,speed,xlab="Numbers of accident")
##### Least-squares fit
fit=lm(noaccident~CS+dtv+slope+speed,data=data,x=TRUE,y=TRUE)
summary(fit)
cat("RESULTS OF LEAST SQUARED FIT METHOD \n")
print(summary(fit))
##### Sampling from posterior
N=10000 #sampling numbers
theta.sample=blinreg(fit$y,fit$x,N) #Sampling from Posterior function
hist(theta.sample$beta[,1],main="Intercept",
     xlab=expression(beta[0]))
hist(theta.sample$beta[,2],main="CS",
     xlab=expression(beta[1]))
hist(theta.sample$beta[,3],main="dtv",
     xlab=expression(beta[2]))
hist(theta.sample$beta[,4],main="slope",
     xlab=expression(beta[3]))
hist(theta.sample$beta[,5],main="speed",
     xlab=expression(beta[4]))
hist(theta.sample$sigma,main="error",
     xlab=expression(epsilon))
cat("RESULTS OF BAYESIAN ESTIMATION \n")
cat("Confident interval \n")
quantile1<-apply(theta.sample$beta,2,quantile,c(.05,.5,.95))
quantile2<-quantile(theta.sample$sigma,c(.05,.5,.95))
print(quantile1)
print(quantile2)
cat("THE END")
```

6.C Maximum likelihood estimation estimation - Weibull model

```
#This program is coded by Nam Lethanh @IBI, ETH Z?rich.
# ESTIMATE THE MODEL'S PARAMETER and Draw the graph
library(MASS)
library(survival)
data <- read.csv("pipelinedata.csv",header=TRUE,sep=",")
attach(data)
sr.fit = survreg(Surv(time,status)~1+waterlost+depth,dist='weibull')
summary(sr.fit)
cat('model parameters \n')
print(summary(sr.fit))
scale=exp(sr.fit$icoef[1])
shape=1/sr.fit$scale[1]
YearMax=2000 # this is the time frame of investigation
plot.new()
#par(mar=c(5, 4, 4, 1) + 0.1)
x=seq(0,YearMax,length=200)
curve(pweibull(x, shape=shape,scale=scale,  lower.tail=FALSE), from=0, to=YearMax, col='red', lwd=2, ylab='',
     xlab='',bty='n',ylim=c(0,1),axes=FALSE)
axis(2, ylim=c(0,1),col="black",las=1)  ## las=1 makes horizontal labels
mtext(expression(paste("Reliability")),side=2,col="black",line=2.2)
axis(1,c(seq(0,YearMax,200)))
mtext(expression(paste("Time (tus)")),side=1,col="black",line=2.2)
#draw the grid on the chart
grid(10, 10, col = "lightgray", lty = "dotted",lwd = par("lwd"), equilogs = TRUE)
box()
```

6.D Bayesian estimation - Weibull model

```
#This program was coded by Nam Lethanh for use in IMP class
data=read.csv("pipelinedata.csv", header=TRUE) #Data
attach(data)
library(LearnBayes) #R package for Bayesian statistics
library(splines)
library(survival) #R package for regression analysis with MLE approach
#BEGIN Bayesian estimation approach on data
#-----function to define posterior likelihood
weibullregpost=function (theta, data)
{
  logf = function(t, c, x, sigma, mu, beta) {
    z = (log(t) - mu - x %*% beta)/sigma
    f = 1/sigma * exp(z - exp(z))
    S = exp(-exp(z))
    c * log(f) + (1 - c) * log(S) }
  k = dim(data)[2]
  p = k - 2
  t = data[, 1]
  c = data[, 2]
  X = data[, 3:k]
  sigma = exp(theta[1])
  mu = theta[2]
  beta = array(theta[3:k], c(p, 1))
  return(sum(logf(t, c, X, sigma, mu, beta)))
}
#-----------------------------
start=c(-.5,9,.5,-.05)
d=cbind(time,status,waterlost,depth)
fit=laplace(weibullregpost,start,d)
proposal=list(var=fit$var,scale=1.5)
bayesfit=rwmetrop(weibullregpost,proposal,fit$mode,10000,d) #This function rwmetrop is a standard function in
     LearnBayes R package describing the Metro-Polish Hasting Algorithm used to generate data based on MCMC.
par(mfrow=c(2,2))
sigma=exp(bayesfit$par[,1])
mu=bayesfit$par[,2]
beta1=bayesfit$par[,3]
beta2=bayesfit$par[,4]
hist(mu,xlab="Intercept", main="")
hist(sigma,xlab="sigma",main="")
#hist(mu,xlab="mu",main="")
hist(beta1,xlab="Water lost",main="")
hist(beta2,xlab="depth",main="")
#END of Bayesian estimation approach for the data
cat("RESULTS OF BAYESIAN ESTIMATION \n")
cat("Confident interval \n")
beta<-c(mu,beta1,beta2)
cat("THE END")
mean(beta1)
```

6.E Bayesian network - bridge reliability (simplified)

```r
#Created by Nam Lethanh for the IMP class
library(bnlearn) #pls install the bnlearn package from CRAN

#------------A simplified BN example-----------

#creating an empty a network in abstract level

dag1 <- empty.graph(nodes = c("Earthquake","Avalanche","Rockfall","Bridge"))
#creating arcs for the network
dag1 <- set.arc(dag1, from = "Earthquake", to = "Avalanche")
dag1 <- set.arc(dag1, from = "Earthquake", to = "Rockfall")
dag1 <- set.arc(dag1, from = "Avalanche", to = "Bridge")
dag1 <- set.arc(dag1, from = "Rockfall", to = "Bridge")
dag1 <- model2network("[Earthquake][Avalanche|Earthquake][Rockfall|Earthquake][Bridge|Avalanche:Rockfall]")
plot(dag1) #plot the structure of the network
graphviz.plot(dag1) ##plot the structure of the network but make it more appealling
library(gRain)
#Define the conditional probability table for the network CPT
yn <- c("yes","no")
c <- cptable(~Earthquake, values=c(0.01,0.99), levels=yn)
s.c <- cptable(~Avalanche | Earthquake, values=c(0.7,0.3,0.1,0.9), levels=yn)
r.c <- cptable(~Rockfall | Earthquake, values=c(0.4,0.6,0.05,0.95), levels=yn)
w.src <- cptable(~Bridge | Avalanche:Rockfall,values=c(0.6,0.4,0.3,0.7,0.2,0.8,0.05,0.95),levels=yn)
cpt.list <- compileCPT(list(c,s.c,r.c,w.src))
print(cpt.list)
#Create a graphical representation
bnet <- grain(cpt.list)
plot(bnet)
# Compile network (details follow)
bnet <- compile(bnet)
# Query network to find marginal probabilities of diseases

p<-querygrain(bnet, nodes=c("Avalanche","Rockfall","Bridge"))
print(p)

# Estimating CPT from data
survey=read.csv("bayesiannetwork-bridgereliabilitya.csv", header=TRUE) #Data
attach(survey)
#-------------Model from data file -------------------
#define sets of variables for each nodes
Earthquake.state <- c("yes", "no")
Aval.state <- c("yes", "no")
Rockfall.state <- c("yes", "no")
Bridge.state <- c("yes", "no")

#using the Maximum likelihood method
bn.mle <- bn.fit(dag1, data = survey, method = "mle")
print(bn.mle)
#prop.table(table(survey[, c("O", "E")]), margin = 2)
#using Bayesian method
bn.bayes <- bn.fit(dag1, data = survey, method = "bayes",iss = 10)
print(bn.bayes)
```

6.F Bayesian network - bridge reliability (Extended)

```r
#Created by Nam Lethanh for the IMP class
library(bnlearn) #pls install the bnlearn package from CRAN

#------------A simplified BN example-----------

#creating an empty a network in abstract level

dag1 <- empty.graph(nodes = c("Earthquake","Avalanche","Rockfall","Bridge"))
#creating arcs for the network
dag1 <- set.arc(dag1, from = "Earthquake", to = "Avalanche")
dag1 <- set.arc(dag1, from = "Earthquake", to = "Rockfall")
dag1 <- set.arc(dag1, from = "Avalanche", to = "Bridge")
dag1 <- set.arc(dag1, from = "Rockfall", to = "Bridge")
dag1 <- model2network("[Earthquake][Avalanche|Earthquake][Rockfall|Earthquake][Bridge|Avalanche:Rockfall]")
plot(dag1) #plot the structure of the network
graphviz.plot(dag1) ##plot the structure of the network but make it more appealling
library(gRain)
#Define the conditional probability table for the network CPT
yn <- c("yes","no")
c <- cptable(~Earthquake, values=c(0.01,0.99), levels=yn)
s.c <- cptable(~Avalanche | Earthquake, values=c(0.7,0.3,0.1,0.9), levels=yn)
r.c <- cptable(~Rockfall | Earthquake, values=c(0.4,0.6,0.05,0.95), levels=yn)
w.src <- cptable(~Bridge | Avalanche:Rockfall,values=c(0.6,0.4,0.3,0.7,0.2,0.8,0.05,0.95),levels=yn)
cpt.list <- compileCPT(list(c,s.c,r.c,w.src))
print(cpt.list)
#Create a graphical representation
bnet <- grain(cpt.list)
plot(bnet)
# Compile network (details follow)
bnet <- compile(bnet)
# Query network to find marginal probabilities of diseases

p<-querygrain(bnet, nodes=c("Avalanche","Rockfall","Bridge"))
print(p)






#-------------Extended Example for more complex structure

#creating an empty a network in detail
dag <- empty.graph(nodes = c("Aval", "Rockfall", "Flood", "Cracks", "Bearing", "Joint", "Earthquake", "Snow", "
     Rain", "DTV", "Bridge"))

#creating arcs for the network
dag <- set.arc(dag, from = "Earthquake", to = "Aval")
dag <- set.arc(dag, from = "Earthquake", to = "Rockfall")
dag <- set.arc(dag, from = "Snow", to = "Flood")
dag <- set.arc(dag, from = "Snow", to = "Aval")
dag <- set.arc(dag, from = "Rain", to = "Flood")
dag <- set.arc(dag, from = "DTV", to = "Cracks")
dag <- set.arc(dag, from = "DTV", to = "Bearing")
dag <- set.arc(dag, from = "DTV", to = "Joint")
dag <- set.arc(dag, from = "Aval", to = "Bridge")
dag <- set.arc(dag, from = "Rockfall", to = "Bridge")
dag <- set.arc(dag, from = "Flood", to = "Bridge")
dag <- set.arc(dag, from = "Cracks", to = "Bridge")
dag <- set.arc(dag, from = "Bearing", to = "Bridge")
dag <- set.arc(dag, from = "Joint", to = "Bridge")
plot(dag) #plot the structure of the network
graphviz.plot(dag) ##plot the structure of the network but make it more appealling

#define sets of variables for each nodes
Earthquake.state <- c("yes", "no")
Snow.state <- c("low", "high")
Rain.state <- c("low", "medium","high")
DTV.state <- c("low", "medium","high")
Joint.state <- c("good", "bad")
Bearing.state <- c("good", "bad")
Cracks.state <- c("small", "moderate","high")
Aval.state <- c("yes", "no")
Rockfall.state <- c("yes", "no")
Flood.state <- c("yes", "no")
Bridge.state <- c("low", "medium","high")


#Assign conditional probability table for each node

#read the data of conditional probability
Earthquake.prob <- array(c(0.01, 0.99), dim = 2,dimnames = list(Earthquake = Earthquake.state))

```

```
#this part creates a random conditional probability
nrow=1
ncol=2
N=10
Snow.sam<-matrix(double(1),nrow=nrow,ncol=ncol)
for (t in 1:nrow){
  sample=sample(c(1:N),ncol,replace=TRUE)
  sample=sample/sum(sample)
  Snow.sam[t,]=sample
}
Snow.sam=as.vector(t(Snow.sam))
Snow.prob <- array(Snow.sam, dim = 2,dimnames = list(Snow = Snow.state))

#this part creates a random conditional probability
nrow=1
ncol=3
N=10
Rain.sam<-matrix(double(1),nrow=nrow,ncol=ncol)
for (t in 1:nrow){
  sample=sample(c(1:N),ncol,replace=TRUE)
  sample=sample/sum(sample)
  Rain.sam[t,]=sample
}
Rain.sam=as.vector(t(Rain.sam))

Rain.prob <- array(Rain.sam, dim = 3,dimnames = list(Rain = Rain.state))

#this part creates a random conditional probability
nrow=1
ncol=3
N=10
DTV.sam<-matrix(double(1),nrow=nrow,ncol=ncol)
for (t in 1:nrow){
  sample=sample(c(1:N),ncol,replace=TRUE)
  sample=sample/sum(sample)
  DTV.sam[t,]=sample
}
DTV.sam=as.vector(t(DTV.sam))

DTV.prob <- array(DTV.sam, dim = 3,dimnames = list(DTV = DTV.state))

#this part creates a random conditional probability
nrow=2*2*2
ncol=2
N=10
Aval.sam<-matrix(double(1),nrow=nrow,ncol=ncol)
for (t in 1:nrow){
  sample=sample(c(1:N),ncol,replace=TRUE)
  sample=sample/sum(sample)
  Aval.sam[t,]=sample
}
Aval.sam=as.vector(t(Aval.sam))

Aval.prob <- array(Aval.sam, dim = c(2,2,2),dimnames = list(Aval = Aval.state, Snow = Snow.state,Earthquake=
     Earthquake.state))

#this part creates a random conditional probability
nrow=2*2
ncol=2
N=10
Rockfall.sam<-matrix(double(1),nrow=nrow,ncol=ncol)
for (t in 1:nrow){
  sample=sample(c(1:N),ncol,replace=TRUE)
  sample=sample/sum(sample)
  Rockfall.sam[t,]=sample
}
Rockfall.sam=as.vector(t(Rockfall.sam))

Rockfall.prob <- array(Rockfall.sam, dim = c(2,2),dimnames = list(Rockfall = Rockfall.state, Earthquake=
     Earthquake.state))

#this part creates a random conditional probability
nrow=2*2*3
ncol=2
N=10
Flood.sam<-matrix(double(1),nrow=nrow,ncol=ncol)
for (t in 1:nrow){
  sample=sample(c(1:N),ncol,replace=TRUE)
  sample=sample/sum(sample)
  Flood.sam[t,]=sample
}
Flood.sam=as.vector(t(Flood.sam))
Flood.prob <- array(Flood.sam, dim = c(2,2,3),dimnames = list(Flood = Flood.state, Snow = Snow.state,Rain=Rain.
     state))

#this part creates a random conditional probability
nrow=3*3
```

```
ncol=3
N=10
Cracks.sam<-matrix(double(1),nrow=nrow,ncol=ncol)
for (t in 1:nrow){
  sample=sample(c(1:N),ncol,replace=TRUE)
  sample=sample/sum(sample)
  Cracks.sam[t,]=sample
}
Cracks.sam=as.vector(t(Cracks.sam))
Cracks.prob <- array(Cracks.sam, dim = c(3,3),dimnames = list(Cracks = Cracks.state, DTV = DTV.state))

#this part creates a random conditional probability
nrow=2*3
ncol=2
N=10
Bearing.sam<-matrix(double(1),nrow=nrow,ncol=ncol)
for (t in 1:nrow){
  sample=sample(c(1:N),ncol,replace=TRUE)
  sample=sample/sum(sample)
  Bearing.sam[t,]=sample
}
Bearing.sam=as.vector(t(Bearing.sam))
Bearing.prob <- array(Bearing.sam, dim = c(2,3),dimnames = list(Bearing = Bearing.state, DTV = DTV.state))
#this part creates a random conditional probability
nrow=2*3
ncol=2
N=10
Joint.sam<-matrix(double(1),nrow=nrow,ncol=ncol)
for (t in 1:nrow){
  sample=sample(c(1:N),ncol,replace=TRUE)
  sample=sample/sum(sample)
  Joint.sam[t,]=sample
}
Joint.sam=as.vector(t(Joint.sam))
Joint.prob <- array(Joint.sam, dim = c(2,3),dimnames = list(Joint = Joint.state, DTV = DTV.state))

#this part creates a random conditional probability
nrow=2*2*2*2*3*2
ncol=3
N=10
Brige.sam<-matrix(double(1),nrow=nrow,ncol=ncol)
for (t in 1:nrow){
  sample=sample(c(1:N),ncol,replace=TRUE)
  sample=sample/sum(sample)
  Brige.sam[t,]=sample
}
Brige.sam=as.vector(t(Brige.sam))
Bridge.prob <- array(Brige.sam, dim = c(3,2,2,2,2,3,2), dimnames = list(Bridge = Bridge.state, Aval = Aval.state, Rockfall=Rockfall.state,Flood=Flood.state,Joint=Joint.state,Cracks=Cracks.state,Bearing=Bearing.state))

#re-arrange the network so it looks nicer
dag1 <- model2network("[Earthquake][Snow][Rain][DTV][Aval|Earthquake:Snow][Rockfall|Earthquake][Flood|Snow:Rain][Cracks|DTV][Bearing|DTV][Joint|DTV][Bridge|Aval:Rockfall:Flood:Cracks:Bearing:Joint]")
plot(dag1) #plot the structure of the network
graphviz.plot(dag1) #plot the structure of the network but make it more appealling
#compiling all conditional probability table into one network table
cpt <- list(Earthquake=Earthquake.prob,Snow=Snow.prob,Rain=Rain.prob,DTV=DTV.prob,Aval=Aval.prob,Rockfall=Rockfall.prob,Flood=Flood.prob,Cracks=Cracks.prob,Bearing=Bearing.prob,Joint=Joint.prob,Bridge=Bridge.prob) #list all individual condition table into a network conditional table
bn1 <- custom.fit(dag1, cpt) #calculate the marginal probability distribution based on the defined network and conditional probability table
#-------------Model from data file -------------------
survey=read.csv("bayesiannetwork-bridgereliabilityb.csv", header=TRUE) #Data
attach(survey)
#using the Maximum likelihood method
bn.mle <- bn.fit(dag1, data = survey, method = "mle")
print(bn.mle)
#prop.table(table(survey[, c("O", "E")]), margin = 2)
#using Bayesian method
bn.bayes <- bn.fit(dag1, data = survey, method = "bayes",iss = 10)
print(bn.bayes)
```

6.G cpt of the BN simplified example

```
Bayesian network parameters
  Parameters of node Avalanche (multinomial distribution)
Conditional probability table:
         Earthquake
Avalanche        no       yes
      no  0.6666667 0.5454545
      yes 0.3333333 0.4545455
  Parameters of node Bridge (multinomial distribution)
Conditional probability table:
, , Rockfall = no
      Avalanche
Bridge        no       yes
   no  0.5000000 0.0000000
   yes 0.5000000 1.0000000
, , Rockfall = yes
      Avalanche
Bridge        no       yes
   no  0.8333333 0.8333333
   yes 0.1666667 0.1666667
  Parameters of node Earthquake (multinomial distribution)
Conditional probability table:
        no       yes
0.6206897 0.3793103

  Parameters of node Rockfall (multinomial distribution)

Conditional probability table:

        Earthquake
Rockfall        no       yes
     no  0.6666667 0.4545455
     yes 0.3333333 0.5454545


  Bayesian network parameters

  Parameters of node Avalanche (multinomial distribution)

Conditional probability table:

         Earthquake
Avalanche        no       yes
      no  0.6304348 0.5312500
      yes 0.3695652 0.4687500

  Parameters of node Bridge (multinomial distribution)

Conditional probability table:

, , Rockfall = no

      Avalanche
Bridge        no       yes
   no  0.5000000 0.1666667
   yes 0.5000000 0.8333333

, , Rockfall = yes

      Avalanche
Bridge        no       yes
   no  0.7352941 0.7352941
   yes 0.2647059 0.2647059


  Parameters of node Earthquake (multinomial distribution)

Conditional probability table:
        no       yes
0.5897436 0.4102564

  Parameters of node Rockfall (multinomial distribution)

Conditional probability table:

        Earthquake
Rockfall        no       yes
     no  0.6304348 0.4687500
     yes 0.3695652 0.5312500
```

6.H cpt of the BN extended example

```
Bayesian network parameters
  Parameters of node Aval (multinomial distribution)
Conditional probability table:
, , Snow = high
     Earthquake
Aval          no       yes
  no  0.7500000 0.5000000
  yes 0.2500000 0.5000000
, , Snow = low
     Earthquake
Aval          no       yes
  no  0.6000000 0.6666667
  yes 0.4000000 0.3333333
  Parameters of node Bearing (multinomial distribution)
Conditional probability table:
       DTV
Bearing      high       low    medium
   bad  0.7000000 0.1666667 0.8571429
   good 0.3000000 0.8333333 0.1428571
  Parameters of node Bridge (multinomial distribution)
Conditional probability table:
, , Bearing = bad, Cracks = high, Flood = no, Joint = no, Rockfall = no
        Aval
Bridge   no yes
  high      1.0
  low       0.0
  medium    0.0
, , Bearing = good, Cracks = high, Flood = no, Joint = no, Rockfall = no
        Aval
Bridge    no yes
  high   0.5
  low    0.0
  medium 0.5
, , Bearing = bad, Cracks = moderate, Flood = no, Joint = no, Rockfall = no
        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = good, Cracks = moderate, Flood = no, Joint = no, Rockfall = no

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = bad, Cracks = small, Flood = no, Joint = no, Rockfall = no

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = good, Cracks = small, Flood = no, Joint = no, Rockfall = no

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = bad, Cracks = high, Flood = yes, Joint = no, Rockfall = no

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = good, Cracks = high, Flood = yes, Joint = no, Rockfall = no

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = bad, Cracks = moderate, Flood = yes, Joint = no, Rockfall = no

        Aval
```

```
Bridge   no yes
  high      1.0
  low       0.0
  medium    0.0

, , Bearing = good, Cracks = moderate, Flood = yes, Joint = no, Rockfall = no

        Aval
Bridge    no yes
  high   0.5
  low    0.0
  medium 0.5

, , Bearing = bad, Cracks = small, Flood = yes, Joint = no, Rockfall = no

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = good, Cracks = small, Flood = yes, Joint = no, Rockfall = no

        Aval
Bridge   no yes
  high      1.0
  low       0.0
  medium    0.0

, , Bearing = bad, Cracks = high, Flood = no, Joint = yes, Rockfall = no

        Aval
Bridge    no yes
  high   0.0
  low    1.0
  medium 0.0

, , Bearing = good, Cracks = high, Flood = no, Joint = yes, Rockfall = no

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = bad, Cracks = moderate, Flood = no, Joint = yes, Rockfall = no

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = good, Cracks = moderate, Flood = no, Joint = yes, Rockfall = no

        Aval
Bridge   no yes
  high      1.0
  low       0.0
  medium    0.0

, , Bearing = bad, Cracks = small, Flood = no, Joint = yes, Rockfall = no

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = good, Cracks = small, Flood = no, Joint = yes, Rockfall = no

        Aval
Bridge    no yes
  high   1.0
  low    0.0
  medium 0.0

, , Bearing = bad, Cracks = high, Flood = yes, Joint = yes, Rockfall = no

        Aval
Bridge    no yes
  high   0.0
  low    1.0
  medium 0.0

, , Bearing = good, Cracks = high, Flood = yes, Joint = yes, Rockfall = no

```

```
        Aval
Bridge   no yes
  high      1.0
  low       0.0
  medium    0.0

, , Bearing = bad, Cracks = moderate, Flood = yes, Joint = yes, Rockfall = no

        Aval
Bridge    no yes
  high   1.0
  low    0.0
  medium 0.0

, , Bearing = good, Cracks = moderate, Flood = yes, Joint = yes, Rockfall = no

        Aval
Bridge    no yes
  high   0.0
  low    1.0
  medium 0.0

, , Bearing = bad, Cracks = small, Flood = yes, Joint = yes, Rockfall = no

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = good, Cracks = small, Flood = yes, Joint = yes, Rockfall = no

        Aval
Bridge    no yes
  high   1.0
  low    0.0
  medium 0.0

, , Bearing = bad, Cracks = high, Flood = no, Joint = no, Rockfall = yes

        Aval
Bridge    no yes
  high   1.0
  low    0.0
  medium 0.0

, , Bearing = good, Cracks = high, Flood = no, Joint = no, Rockfall = yes

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = bad, Cracks = moderate, Flood = no, Joint = no, Rockfall = yes

        Aval
Bridge   no yes
  high      1.0
  low       0.0
  medium    0.0

, , Bearing = good, Cracks = moderate, Flood = no, Joint = no, Rockfall = yes

        Aval
Bridge    no yes
  high   0.0
  low    1.0
  medium 0.0

, , Bearing = bad, Cracks = small, Flood = no, Joint = no, Rockfall = yes

        Aval
Bridge   no yes
  high      0.0
  low       0.0
  medium    1.0

, , Bearing = good, Cracks = small, Flood = no, Joint = no, Rockfall = yes

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = bad, Cracks = high, Flood = yes, Joint = no, Rockfall = yes
```

```
        Aval
Bridge   no yes
  high      0.0
  low       0.0
  medium    1.0

, , Bearing = good, Cracks = high, Flood = yes, Joint = no, Rockfall = yes

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = bad, Cracks = moderate, Flood = yes, Joint = no, Rockfall = yes

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = good, Cracks = moderate, Flood = yes, Joint = no, Rockfall = yes

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = bad, Cracks = small, Flood = yes, Joint = no, Rockfall = yes

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = good, Cracks = small, Flood = yes, Joint = no, Rockfall = yes

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = bad, Cracks = high, Flood = no, Joint = yes, Rockfall = yes

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = good, Cracks = high, Flood = no, Joint = yes, Rockfall = yes

        Aval
Bridge   no yes
  high      0.0
  low       0.0
  medium    1.0

, , Bearing = bad, Cracks = moderate, Flood = no, Joint = yes, Rockfall = yes

        Aval
Bridge    no yes
  high   0.0
  low    0.0
  medium 1.0

, , Bearing = good, Cracks = moderate, Flood = no, Joint = yes, Rockfall = yes

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = bad, Cracks = small, Flood = no, Joint = yes, Rockfall = yes

        Aval
Bridge    no yes
  high   0.0
  low    1.0
  medium 0.0
```

```
, , Bearing = good, Cracks = small, Flood = no, Joint = yes, Rockfall = yes

        Aval
Bridge   no yes
  high      0.0
  low       1.0
  medium    0.0

, , Bearing = bad, Cracks = high, Flood = yes, Joint = yes, Rockfall = yes

        Aval
Bridge    no yes
  high   0.0
  low    1.0
  medium 0.0

, , Bearing = good, Cracks = high, Flood = yes, Joint = yes, Rockfall = yes

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = bad, Cracks = moderate, Flood = yes, Joint = yes, Rockfall = yes

        Aval
Bridge   no yes
  high      0.0
  low       0.0
  medium    1.0

, , Bearing = good, Cracks = moderate, Flood = yes, Joint = yes, Rockfall = yes

        Aval
Bridge   no yes
  high
  low
  medium

, , Bearing = bad, Cracks = small, Flood = yes, Joint = yes, Rockfall = yes

        Aval
Bridge    no yes
  high   0.0
  low    1.0
  medium 0.0

, , Bearing = good, Cracks = small, Flood = yes, Joint = yes, Rockfall = yes

        Aval
Bridge   no yes
  high
  low
  medium


  Parameters of node Cracks (multinomial distribution)

Conditional probability table:

          DTV
Cracks          high       low    medium
  high     0.4000000 0.2500000 0.4285714
  moderate 0.4000000 0.4166667 0.4285714
  small    0.2000000 0.3333333 0.1428571

  Parameters of node DTV (multinomial distribution)

Conditional probability table:
      high       low    medium
0.3448276 0.4137931 0.2413793

  Parameters of node Earthquake (multinomial distribution)

Conditional probability table:
        no       yes
0.6206897 0.3793103

  Parameters of node Flood (multinomial distribution)

Conditional probability table:

, , Snow = high

     Rain
Flood      high       low    medium
```

```
  no  0.6666667 0.5714286 0.3333333
  yes 0.3333333 0.4285714 0.6666667

, , Snow = low

     Rain
Flood      high       low    medium
  no  0.8000000 0.1666667 0.5000000
  yes 0.2000000 0.8333333 0.5000000


  Parameters of node Joint (multinomial distribution)

Conditional probability table:

     DTV
Joint      high       low    medium
  no  0.5000000 0.2500000 0.5714286
  yes 0.5000000 0.7500000 0.4285714

  Parameters of node Rain (multinomial distribution)

Conditional probability table:
      high       low    medium
0.2758621 0.4482759 0.2758621

  Parameters of node Rockfall (multinomial distribution)

Conditional probability table:

        Earthquake
Rockfall        no       yes
     no  0.6666667 0.4545455
     yes 0.3333333 0.5454545

  Parameters of node Snow (multinomial distribution)

Conditional probability table:
      high       low
0.5517241 0.4482759
```

6.I Bayesian network - Learning structure of a BN from data

```
# This program was coded by  Nam Lethanh (lethanh@ibi.baug.ethz.ch)
#Bayesian network example
library(deal) #this is dependency package
data <- read.csv("BayesianNetwork-dataaccident.csv",header=TRUE)
accident<-network(data)
plot(accident)
accident.nd <- nodes(accident)# the list of nodes
accident.j <- jointprior(accident)
accident.prior <- jointprior(accident)
#accident.prior <- jointprior(accident,12)
accident <- learn(accident,data,accident.prior)$nw
nodes(accident)$Time$condprior
nodes(accident)$Time$condposterior
nodes(accident)$Time$loglik
accident$score
allaccident <- networkfamily(data,accident,accident.prior)
allaccident <- nwfsort(allaccident$nw)
print(allaccident)
plot(allaccident)
accident.s <- autosearch(accident,data,accident.prior)$nw
plot(accident.s)
#THE END
cat("THE END")
```