NATO Science for Peace and Security Series - C:
Environmental Security

Transport of Dangerous Goods

Methods and Tools for Reducing the Risks of Accidents and Terrorist Attack

Edited by
Emmanuel Garbolino
Mohamed Tkiouat
Natalia Yankevich
Dalanda Lachtar

Springer

NATO OTAN
This publication is supported by: The NATO Science for Peace and Security Programme

Transport of Dangerous Goods

NATO Science for Peace and Security Series

This Series presents the results of scientific meetings supported under the NATO Programme: Science for Peace and Security (SPS).

The NATO SPS Programme supports meetings in the following Key Priority areas: (1) Defence Against Terrorism; (2) Countering other Threats to Security and (3) NATO, Partner and Mediterranean Dialogue Country Priorities. The types of meeting supported are generally “Advanced Study Institutes” and “Advanced Research Workshops”. The NATO SPS Series collects together the results of these meetings. The meetings are co-organized by scientists from NATO countries and scientists from NATO’s “Partner” or “Mediterranean Dialogue” countries. The observations and recommendations made at the meetings, as well as the contents of the volumes in the Series, reflect those of participants and contributors only; they should not necessarily be regarded as reflecting NATO views or policy.

Advanced Study Institutes (ASI) are high-level tutorial courses to convey the latest developments in a subject to an advanced-level audience

Advanced Research Workshops (ARW) are expert meetings where an intense but informal exchange of views at the frontiers of a subject aims at identifying directions for future action

Following a transformation of the programme in 2006 the Series has been re-named and re-organised. Recent volumes on topics not related to security, which result from meetings supported under the programme earlier, may be found in the NATO Science Series.

The Series is published by IOS Press, Amsterdam, and Springer, Dordrecht, in conjunction with the NATO Emerging Security Challenges Division.

Sub-Series

A.	Chemistry and Biology	Springer
B.	Physics and Biophysics	Springer
C.	Environmental Security	Springer
D.	Information and Communication Security	IOS Press
E.	Human and Societal Dynamics	IOS Press

http://www.nato.int/science
http://www.springer.com
http://www.iospress.nl

Series C: Environmental Security

Transport of Dangerous Goods

Methods and Tools for Reducing the Risks of Accidents and Terrorist Attack

edited by

Emmanuel Garbolino
MINES ParisTech, CRC
Sophia Antipolis, France

Mohamed Tkiouat
Université Mohammed V-Agdal
Rabat, Morocco

Natalia Yankevich
National Academy of Sciences of Belarus
Minsk, Belarus

and

Dalanda Lachtar
MINES ParisTech, CRC
Sophia Antipolis, France

Springer

Published in Cooperation with NATO Emerging Security Challenges Division

Proceedings of the NATO Advanced Research Workshop on
Risk Prevention for Environment and Human Society against Dangerous
Goods Transport Accidents and Malicious Intents: Methods and Tools
Paris, France
5–9 July 2010

Library of Congress Control Number: 2011945772

ISBN 978-94-007-2686-4 (PB)
ISBN 978-94-007-2683-3 (HB)
ISBN 978-94-007-2684-0 (e-book)
DOI 10.1007/978-94-007-2684-0

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

Foreword

**Emmanuel Garbolino, Mohamed Tkiouat,
Natalia Yankevich, and Dalanda Lachtar**

The transport of goods is significantly increasing because of the globalization of exchanges across the world. This increase also concerns the Dangerous Goods Transport (DGT) that could be in different modes like roads, boats, trains, airplanes, pipelines, etc. Although the probability of accident in the DGT is very low, the consequences can be qualified as "major" when the event affects a great number of people, buildings, infrastructures, ecosystems and environmental resources. Some examples of such accidents, concerning hazardous materials, are boat accidents (Torrey Canion in 1967; Amoco Cadiz in 1978; Erika in 1999 etc.), road accidents (Los Alfaques in 1978) or train accidents (Viareggio in 2009). DGT accidents can be also caused by malicious acts such as trucks' explosions (Kandahâr 2003; Kabul 2009; Moussayib 2005; Mossul 2007; Bagdad 2005; Djerba 2002 etc.).

The risk related to the transportation of dangerous goods is a delicate risk, even a complex one, because it is distributed in the whole network and it depends on multiple factors such as the traffic concentration, the weather conditions, the occurrence of undesirable events (accidents on the infrastructure, natural hazards etc.), the state of the transport infrastructure, the behaviour of the driver, the

E. Garbolino • D. Lachtar
MINES ParisTech, Crisis and Risk Research Centre (CRC),
1 rue Claude Daunesse, BP 207, 06904 Sophia Antipolis, France
e-mail: emmanuel.garbolino@mines-paristech.fr; dalanda.lachtar@mines-paristech.fr

M. Tkiouat
LERMA, Ecole Mohammadia d'Ingénieurs,
Université Mohamed V Agdal, Rabat, Morocco
e-mail: tkiouat@emi.ac.ma

N. Yankevich
Laboratory of Working Processes and Safety of Technical Systems, Joint Institute
of Mechanical Engineering of National Academy of Sciences of Belarus,
12, Academicheskaya str, 220072 Minsk, Belarus
e-mail: lab_12@tut.by

diversification of the means of transport, the political situation of the state, the international level, etc.

A DGT event can thus occur on very different territories according to the nature of the stakes that are exposed to the consequences of the event, such as highly urbanized, industrialized or touristic areas, or even low anthropized surroundings (natural parks, Natura 2000 sites, etc.). Besides, the diversity of the transported hazardous material brings forth an additional level of complexity in the apprehension of this risk. The phenomena that may happen (fire, explosion, toxic atmospheric release, etc.) depend on the characteristics of the hazardous material.

One of the difficulties of risk governance of DGT concerns the complexity of "DGT system". It is expressed through the relationships among the various stakeholders contributing to different levels of responsibility in the prevention of risks and crisis management. It is thus possible to distinguish between the following main groups of stakeholders with well defined activities and objectives. The objectives of the groups could be complementary or even antagonistic for some of them:

– The supply chain of hazardous material composed of the industries who produce the hazardous material, the ones who load the material and the shipping companies that convey the material, the industrialists who use, and the vendors who distribute the material. The priority is given to financial interests;
– The public authorities and the local governments that have to ensure the viability of the networks and to protect the citizens and the environment from the consequences of the DGT accidents. Their mission is to define the plans of intervention, the possible local regulation measures to avoid or limit the situations at risk, and the measurements of urbanization's control;
– Residents located near the transport infrastructures, the people in charge of the Establishments Receiving Public (ERP – hospitals, sports centers, supermarkets, etc.), the users and the citizens who must be preserved from the consequences of the DGT accidents;
– The stakeholders who are not directly involved in the DGT risks, but play a role at various levels: insurance companies, media, trade unions, associations, information centers on risks, and finally, the researchers who produce methods, tools and analyses, which may concern the various stakeholders.

In the context of DGT risk and crisis management challenges, this book encompasses the complexity of the "DGT system" in order to provide information and methods regarding the following issues:

– The global strategy to regulate and control the DGT flows, and provide a reflection frame to secure theses flows;
– The risk definition and the estimation methodologies of DGT events on a territory with reference to the specific elements that characterize the territory, the transport means, the infrastructure, etc.;
– The assessment methodology for analyzing the consequences on the environment and the population in case of a DGT accident or malicious act;

– The route optimization and planning in order to reduce the accident occurrence and to mitigate the consequences on the environment and the population;
– The supply chain risk management dedicated to DGT;
– The crisis planning and decision support for DGT accident or malicious acts.

The editors are very grateful to the contributors of this book, who shared their original and practical points of view on this subject. We hope that this knowledge will be useful to the public and private decision makers, to help them taking into account the risks associated with the DGT. The editors and the authors also want to express their acknowledgements to the NATO "*Science for Peace and Security Programme*" committee for its support and for its fruitful suggestions.

Contents

Chapter 1
Dangerous Goods Transportation and Biophysical Vulnerability: The Contribution of GIS and Simulation Softwares

Emmanuel Garbolino and Dalanda Lachtar

Abstract The transport of dangerous goods by road can be the cause of major accidents. This type of transport can also be hijacked by terrorists. They are therefore one of the factors that increase the vulnerability of the territories they cross. The risk associated with transport of dangerous goods is still underestimated in Europe. Therefore this chapter presents methods and tools to help policy-makers and the private sector to assess the vulnerability of areas under their control, with a specific focus on toxic releases and explosions. These methods and tools provide models and data for the implementation of a Spatial Decision Support System (SDSS). To illustrate the approach, two case studies are described. One concerns the transport of ethylene oxide and the other liquid petroleum gas (LPG) in the cross-border corridor (Nice – Imperia – Savona) between France and Italy.

1.1 Introduction

This chapter aims to raise awareness of the risks involved in the transport of dangerous goods (TDG) by road. It outlines a method to measure the range of potentially dangerous effects, which when coupled with Geographic Information Systems (GIS) can be used to identify exposed assets. Although accidents or malicious acts involving hazardous materials transported by road are very uncommon in Europe, the potential consequences can have a particularly serious impact on health, human activities and the stability and integrity of ecosystems.

To illustrate this point, this chapter includes a brief summary of TDG accidents in France and around the world. It highlights (according to the substances involved)

E. Garbolino (✉) • D. Lachtar
MINES ParisTech, Crisis and Risk Research Centre (CRC), 1 rue Claude Daunesse, BP 207, 06904 Sophia Antipolis, France
e-mail: emmanuel.garbolino@mines-paristech.fr; dalanda.lachtar@mines-paristech.fr

E. Garbolino et al. (eds.), *Transport of Dangerous Goods*, NATO Science for Peace and Security Series C: Environmental Security, DOI 10.1007/978-94-007-2684-0_1,

the main associated hazards, (e.g. explosion or release of toxic gas) and the various consequences. The mechanisms of these hazards are then explained. This is both to provide a factual description of how these situations can develop and also to identify the key parameters that must be taken into account for modelling.

The description of these phenomena, combined with feedback from such accidents represents a first step in understanding risk prevention issues associated with TDG. It aims to define the vulnerability of an area faced with a particular hazard, in order to develop a set of models aimed at reducing this vulnerability.

1.2 The Transport of Dangerous Goods by Road: A General Framework

According to the French *Ministère de Ecologie* (Environment Ministry), more than 3,000 designated hazardous substances circulate on French soil each year. The Seveso II Directive (1996) points out that the 'hazardous' can either be an intrinsic property of a substance or of a physical situation, which can cause damage to human health and/or the environment.

Regulatory documents regarding the transport of dangerous goods such as the ADR (the European Agreement concerning the International Carriage of Dangerous Goods by Road) divide dangerous goods into nine classes (Table 1.1). Hazardous materials are classed according to their physico-chemical characteristics and by their principal hazardous property.

According to the French *Ministère de l'Ecologie* risk can be considered as 'a confrontation between a hazard (whose probability and intensity of occurrence is known, or sought), and an asset (whose vulnerability and resilience to the hazard in question should be characterized)' [1]. In the context of the transport of dangerous goods, risk can be considered as the probability of occurrence of a transport accident or a malevolent act involving a hazardous material. The accident or act occurs in an area where people, property, infrastructure, networks, or the environment is exposed to the consequences of phenomena resulting from this dangerous event.

The risk related to the transport of dangerous goods by road is particularly subtle and difficult to understand because it is distributed across the entire road network and depends on diverse factors such as traffic density, weather conditions, adverse events (e.g. accidents, natural hazards), the transport infrastructure, the behaviour of the driver, etc. In addition, an accident or a malicious act involving TDG may occur in areas that vary widely in terms of the nature of assets exposed to the consequences of an accident – on the one hand highly urbanized or industrial areas, on the other hand, tourist centres or natural environments (e.g. national parks, reserves).

Moreover, in its turn, the diversity of dangerous goods transported brings another layer of complexity to understanding the risk. This is due to the variety of phenomena that can occur (e.g. fire, explosion, and toxic release into the air). For example, the transport of fuels such as gasoline or LPG can result in serious fires or the explosion of the tanks that carry them, causing heat, overpressure and missile effects.

Table 1.1 Classification of dangerous substances according to the ADR 2011

No	Class of substance	Principal hazardous property
1.	Explosive substance and articles	Explosive
2.	Gasses (liquid or compressed)	Gaseous state
3.	Flammable liquids	Flammable
4.1.	Flammable solids	Flammable
4.2.	Substances liable to spontaneous combustion	Flammable
4.3.	Substances which, in contact with water, emit flammable gasses	Flammable
5.1.	Oxidising substances	Flammable
5.2.	Organic peroxides	Toxic
6.1.	Toxic substances	Toxic
6.2.	Infectious substances	Infection
7.	Radioactive material	Radioactive
8.	Corrosive substances	Corrosive
9.	Miscellaneous dangerous substances and articles	Toxic, temperature-related, various

LPG in particular is usually transported through the centre of highly urbanized areas to the point-of-sale (service stations), which creates a particularly high risk for local residents. The city of Nice (France) has 52 such service stations located in densely urbanized areas where there are also hospitals, shopping centres, places of worship etc. This has required the local authorities to implement prevention measures, in consultation with public authorities. Other substances have properties which can lead to the formation of a cloud of toxic gas. This may be the result of a leakage caused by accidental or intentional piercing of the tank, or a technical failure of the container (e.g. corrosion, abnormal pressure). Finally, the proximity and quality of emergency resources available to the authorities, civil protection and infrastructure managers in turn determines the level of risk, particularly in terms of handling the immediate consequences in a crisis and for the eventual return to normality.

It is important to recognize that the risk of TDG remains, in general, poorly studied and still not given sufficient weight by governments and local authorities. This is in large part due to the difficulty of understanding the problem. This point underlines the value of addressing the issue in order to provide decision-makers with scientific, objective and generalizable data to support their risk prevention policy guidelines in the areas concerned.

Figure 1.1 illustrates the complexity of the TDG system. It shows the key stakeholders involved in the challenge of TDG risk. This figure distinguishes four main groups of actors, each with specific activities and objectives. Some of these are complementary, but in other cases they conflict:

- The hazardous substances logistics chain. This consists of the industries that produce hazardous materials, the transportation companies that carry them, the industries that use them, and retail outlets that distribute them. Financial interests have priority.

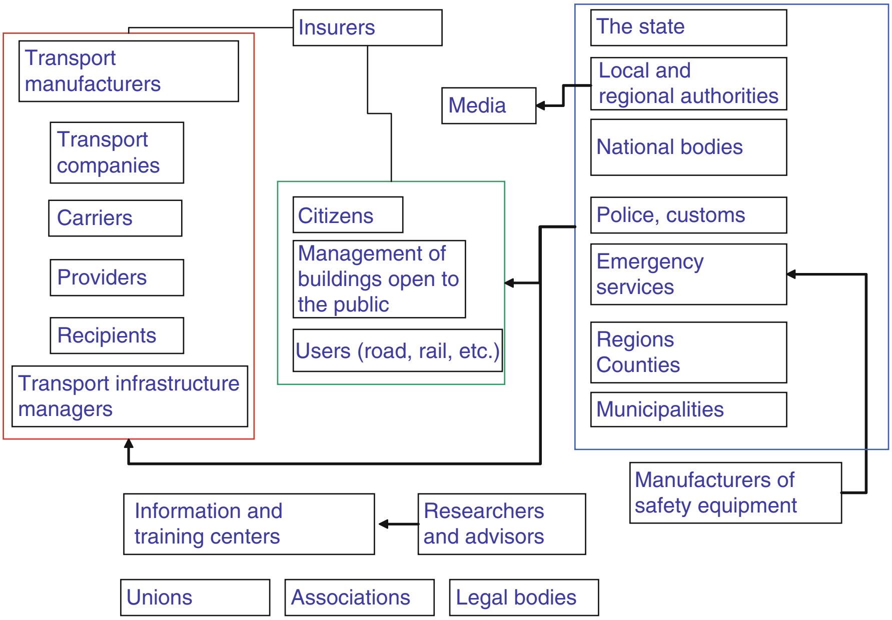


Fig. 1.1 The main stakeholders involved in TDG risk

- Governments and local authorities. These bodies have a duty to ensure the viability of transport networks and to protect both citizens and the environment from the consequences of TDG accidents. Their mission is based on the definition of action plans and the consequent regulation of local traffic flow to avoid or minimize risk situations.
- People living in the vicinity of transport infrastructure, the management of establishments open to the pubic (hospitals, sports centres, supermarkets etc.), road users and citizens. These people must be protected from the consequences of TDG hazards.
- 'Crossover' actors who are not necessarily directly affected by TDG risk but who have a role to play at various levels. This comprises insurers, the media, unions, associations, risk information centres, etc. It also includes researchers, who produce methods, tools, statistical analysis, etc. that can be of interest to the various parties shown in the diagram.

In addition to the interactions between the various actors involved in risk management and prevention of TDG (which itself adds complexity to the system), it is also necessary to analyse TDG flow data. Research in the field of TDG road transport highlights the lack of information available on traffic flows. It also highlights a particular observation on the bilateral flow between France and Italy (a finding that can be extended to other European countries).

Neither national nor local authorities, nor the managers of transport infrastructure (such as companies managing motorways) know precisely the nature, number and range of hazardous substances transiting their territory. Despite surveys by public authorities and highway operators in France and Italy, the numbers are insufficient to ensure sound risk management. Moreover, because of the diversity of hazards associated with the many different dangerous materials transported (explosion, toxic release, corrosion, radioactivity, etc.) and the assets potentially exposed to these phenomena (population, environment, infrastructure, etc.), risk prevention and crisis management must be based on accurate and comprehensive information made available to policy-makers.

1.3 TDG Road Accidents and Associated Phenomena

According to the French governmental website dealing with the transport of hazardous substances, accidents involving dangerous goods vehicles fall into two broad categories:

- Type ‘T’ (Traffic) accidents. These are the accidents in which there is little, or no release of hazardous substances.
- Type ‘D’ (Dangerous substances) accidents. These are characterized by:
  - injuries due to the hazardous substance,
  - substance spillage of more than 100 l,
  - a gas leak, regardless of size, and
  - a total or partial explosion of the shipment of dangerous goods.

It is type ‘D’ accidents that are of interest here, as they have serious consequences for local populations and the environment. Experience gained from previous road accidents of this type is a rich source of information as it highlights situations, phenomena and consequences that have already occurred but would have been difficult to predict. Among the databases that record such events related to technological risks, the French ARIA database (*Analyse, Recherche et Information sur les Accidents*) and BARPI (*Bureau d’Analyse des Risques et Pollutions Industrielles*) are good reference sources for French data, and to a lesser extent, for the rest of the world.

The ARIA database was established in February 1993 by BARPI which was itself created in 1992. It operates under the auspices of the French *Ministère de l’Ecologie.* The main objective of this database is to gather accurate and detailed information about all types of industrial and technological incidents and accidents in France and worldwide, in order to provide feedback and support the implementation of risk analyses. This database currently lists 33,000 accidents (80% in France) related to industrial facilities, transportation of hazardous materials, agricultural pollution, etc.

BARPI collects information from many sources; not only the inspection of classified installations, but also national and international media and the Internet (for notification of events), as well as central government, national security agencies,

Table 1.2 Examples of TDG road accidents and their consequences

Dangerous event	Country	Region or municipality	Year	Substance
BLEVE	Spain	Los Alfaques	1978	Propylene
BLEVE	Thailand	Bangkok	1990	LPG
Explosion	France	Port-Sainte-Foy-Et-Ponchapt	1997	Heating oil, diesel, high-octane petrol
Explosion	Italy	Palerme	1996	LPG
UVCE	France	Saint-Amand-Les-Eaux	1972	Propane
UVCE	Mexico	Xilatopec	1978	LPG
Explosion	Iran	Nosratabad	2004	Petrol
Toxic release	USA	Houston	1976	Ammonia
Toxic release	China	–	2005	Chlorine

maritime agencies, international organizations, etc. The information collected is initially analysed and assimilated by BARPI members, and updated if necessary. This information is then inserted into the ARIA database which is organised into around 100 tables containing about 2,300 fields and their relations. The aim is to characterize events, affected buildings, and measures taken during and after the event, etc. The information is then sent for a final check to classified facilities and certain professional bodies, such as the French organization GESIP (*Groupe d'Étude de Sécurité des Industries Pétrolières*). The information provided covers the following points:

- A general description of the site (its location, legal and administrative arrangements, type of economic activity, technical characteristics of the facility, and the human context).
- A description of the circumstances of the event (weather conditions at the time of the incident, its type, operating conditions, any equipment involved, analysis of failures and their causes, main products involved, officials directly involved, response and emergency measures).
- The consequences of the event (the effect on humans, material damage, loss of use, damage to the natural environment, other property damage, damage to fauna and flora, degradation of biological ecosystem indicators, the economy).
- Follow-up (administrative, legal and technical), cleaning and decontamination, environmental recovery and rehabilitation, corrective actions, etc.

Part of this information (33,000 accidents summaries from 1998 to 2007) as well as numerous reports and factsheets are also directly available to the public on the ARIA website. The analysis of TDG road accidents using BARPI data shows that the hazards that have caused the greatest number of deaths (apart from fire) are: BLEVE-type (Boiling Liquid Expanding Vapour Explosion), UVCE-type (Unconfined Vapour Cloud Explosion) explosions and the release of toxic substances into the atmosphere. The following incidents (Table 1.2) which have occurred in various countries, illustrate this finding.

This section has presented a typology of potential TDG accidents. The following section describes the hazards that arise as a result of a TDG accident.

1.4 The Hazards Arising from a TDG Accident

The previous section demonstrated that type 'D' road accidents can generate three main categories of hazardous phenomena (excluding fire):

- The BLEVE: effects are primarily related to heat flow.
- The UVCE: consequences depend on the effects of overpressure.
- The release of toxic substances which impacts human health and the environment.

Whether these phenomena occur depends on: the type of transport used (tank, hazardous material under pressure, etc.), the substance being transported (flammable, toxic, etc.) and how the accident occurred (the presence of a source of ignition, low wind speed, etc.). Furthermore, potential domino effects should be investigated (effects that may propagate and affect other facilities) when these events occur in heavy traffic and close to sites for the storage, production and distribution of high-risk substances.

1.4.1 Explosions

The impact of explosions has been evaluated, with particular respect to heat flows (a BLEVE event) and overpressures (a UVCE event).

1.4.1.1 The BLEVE Explosion

The BLEVE (Boiling Liquid Expanding Vapour Explosion) is a scenario analogous to an explosion caused by the rapid expansion of flammable vapours originating in a gaseous substance stored under pressure, as a liquid, in a confined environment. The consequences of this event may include overpressure and thermal irradiation (a fireball) on people and buildings. This type of event has three main dangers: the shock wave, heat flux and projectiles (flying debris). The shock wave is generated by a sudden change in pressure and occurs in two phases: the overpressure wave and the negative, or under-pressure wave.

Flying debris is a result of the explosion of the tank. Studies on various types of tanks with a volume of 39 m^3 show that:

- 80% of fragments travel around 250 m.
- 10% of fragments travel around 400 m.
- The maximum distance observed is about 1,200 m.

1.4.1.2 The UVCE Explosion

The UVCE (Unconfined Vapour Cloud Explosion) is an accident scenario created by the release and dispersal in an unconfined area of a flammable vapour or gaseous substance. Should this substance ignite, it will have thermal and overpressure effects on humans, buildings and the environment. The exact nature of these effects depends on local conditions in particular, the weather and the gas mix.

The thermal effects on humans are mainly caused by the fireball, which can have lethal effects on anyone in its path. The effect on buildings is generally limited to surface damage, although small cracks can sometimes be seen in metal structures.

A UVCE explosion generally follows five stages:

- The atmospheric release of a flammable product as a gas or a liquid which vaporises.
- The mixing of this product with a combustible agent (such as oxygen from the air) to form a flammable substance.
- The dilution and movement of the gas cloud. The cloud remains partially flammable up to a certain distance which depends on wind speed and the relief of the environment immediately surrounding the source of the release (e.g. the presence or absence of barriers and their height).
- Ignition of the cloud when it encounters a heat source.
- The propagation of a wall of flame within the cloud associated with expansion of the burning gas. This wall of flame acts like a piston on the surrounding colder gas and can be the cause of the formation of a detonation wave.

The effects of overpressure (due to the magnitude of the pressure wave generated) are directly proportional to the speed with which the wall of flame spreads: the faster this propagates, the greater its acceleration and the higher the overpressure. The effects of overpressure are of course inversely proportional to the distance from the ignition point, but they vary greatly depending on whether there are obstacles in the environment.

Threshold pressure values, both for people and buildings are given later in this chapter as they provide useful information in evaluating the biophysical vulnerability of humans and property when faced with a UVCE-type explosion.

1.4.2 Toxic Atmospheric Releases

Atmospheric dispersion corresponds to the physical phenomena of a gaseous product (the toxic cloud) mixing with air and its movements in time and space [2]. The conditions surrounding the atmospheric dispersion of a toxic cloud depend mainly on the following: the characteristics of the source of the emission (the molecular mass of the substance, its reactivity with water or the sun, closed or open storage, spillage, leakage, monophasic or biphasic pollutant, etc.), the weather (temperature, humidity, wind direction and strength, presence of an inversion, etc.) and the topography of the environment immediately surrounding the source of the emission (e.g. the relief, buildings, agricultural fields, forests). Depending on these characteristics, dispersion can be completely different from one product to another, from one area to another, and from one time period to another. In addition to this, various models are used to estimate the propagation distances of toxic clouds. The simulations used in this study use the ALOHA software [3], which is based on the Gaussian model and a global model of the atmospheric dispersion of pollutants.

1.4.2.1 The Source of the Release

The source of the release comprises all the elements involved in the creation of the toxic cloud. They therefore encompass all of the following:

- The hazardous material. It is important to know its atmospheric reactivity in order to assess whether its chemical composition will change and if it will react violently in combination with air or moisture, etc. The method used to store the material is another important parameter, particularly with respect to the state the hazardous substance is maintained in (liquid, gas, a mixture of liquid and gas) which affects both the kinetics of the formation of the toxic cloud and its dispersal.
- The storage hardware; for example, tanks and bottles. The capacity, possible break points, pressure and storage temperature, and the physical structure of the tank must be known.
- How the hazardous material is released into the atmosphere. Conventionally, models consider two types of discharge. The first type describes an instantaneous release that corresponds to the total destruction of the container which brings the entire hazardous substance into contact with the atmosphere. This is an unlikely event that most simulations do not take into account. The second scenario is a continuous release caused by a leak from a pipe or hole (of various sizes and positions) in the container. This release has an elongated form and propagates in a direction determined by the position of the hole and the wind. This type of release has been simulated as it is closest to the actual conditions surrounding the formation of a toxic cloud in a TGD accident on the road.
- How the cloud travels and its movements in the atmosphere. The velocity of the cloud and its dilution in the air affects the risk level for potential targets. The orientation of the hole is fundamental in estimating the spread of the toxic cloud (top, bottom, side, etc.). The storage pressure is also essential to determine the displacement and concentration of the cloud. In the case of clouds heavier than air, gravity will force the cloud to slide along the ground where it is subject to friction, which slows its progress and may increase the concentration.

1.4.2.2 The Weather

The basic weather conditions that come into play during the creation of a toxic cloud include:

- Wind speed. Depending on the density of the cloud, this will influence the speed of propagation, atmospheric dispersal and friction against the features of the terrain.

- Air temperature. This can influence the state (liquid, gas) of the hazardous material released. The temperature also affects whether, and how quickly molecules rise (thermal stratification of the atmosphere).
- Temperature inversion. This will cause the lower part of the cloud to stagnate in the atmosphere. This can aggravate the exposure of vulnerable assets as it prevents the rapid dispersal of the cloud, especially when it consists of light gases.
- Air humidity. This plays a role as toxic substances can react with water vapour. Conversely, they can be captured by water droplets which reduce the rate at which the clouds move through the atmosphere.
- Cloud cover. Some chemicals react with sunlight (a photochemical reaction). The extent of cloud cover is taken into account in simulations.
- Precipitation. Whether in the form of rain, snow, etc. precipitation acts as a barrier limiting the movement of the toxic cloud through the atmosphere.

This information on weather conditions is integrated into the simulation software either manually by users of the system, or automatically using weather stations with appropriate sensors.

1.4.2.3 The Immediate Environment

The concept of the environment immediately surrounding an atmospheric dispersal includes the following elements:

- The relief. For example, the presence or absence of obstacles, varying in shape, size and nature, occupying different places in space. The progression of the toxic cloud will vary depending on whether it is in an open area without obstacles (e.g. large plains), or an enclosed area containing buildings, trees, etc.
- Topography. This influences the flow of the cloud depending on the steepness of slopes and the wind speed. The presence of slope winds (warm air rising up a slope, and cooler air descending) also disturbs the flow of a toxic cloud.
- Thermal effects of the landscape. The main types of land use (open ground, agricultural fields, forests, paved roads, built-up areas, water bodies, etc.) have a thermal influence which can be seen in phenomena such as sea breezes, or disturbances caused by thermal columns.

Only models that take into account three-dimensional space make it possible to accurately represent the influence of topography and obstacles on the flow of a toxic cloud. Conventional two-dimensional models use weighted values to represent the influence of the immediate environment.

The simulations carried out in this study use ALOHA software. This is founded on a Gaussian model of neutral gases with a density close to that of air, and a comprehensive model for other gases. The software also makes it possible to calculate the range of heat flows and overpressures in the case of a BLEVE or UVCE-type explosion. These simulations will be run in the various case studies discussed later in this chapter.

The next section introduces the concept of territorial vulnerability as applied to the problem of TDG road transport.

1.5 The Concept of Vulnerability

The term 'vulnerability' originates in the late Latin word '*vulnerabilis'* meaning 'likely to injure' or 'wounding'. Mainly used in medical science and nature, the concept of vulnerability has been gradually introduced into the field of natural and technological risk. There is therefore a significant body of international literature concerned with the concept of vulnerability, especially in geography and in the science of risks and hazards.

D'Ercole and Metzger [4] suggest that:

> territorial vulnerability refers to the idea that there is, within any area, identifiable elements likely to generate and disseminate their vulnerability to the entire territory, causing effects that can disrupt, compromise or discontinue its operation and development. In this sense, the analysis of territorial vulnerability aims primarily to identify, characterize and prioritize the areas from which territorial vulnerability is created and diffused. It can therefore enable the definition of areas for which risk prevention measures are highly efficient, thereby taking an approach opposite to that of routine interventions aimed at reducing risk, which are often just ad-hoc contingency measures.

These authors also state that to study the vulnerability of an area is *a fortiori* to take into account the vulnerability of key assets, such as life support networks (also called critical infrastructure). They place specific emphasis on the vulnerability of these key assets as these are the elements essential to the functioning of the territory, regardless of the type of hazard. This position highlights the social space as the object of study.

However, we believe it is appropriate to consider the vulnerability of key assets, and the nature, frequency and intensity of hazards, in order to understand the vulnerability of the territory as a whole. In other words, the vulnerability of the territory as whole is inseparable from the vulnerability of key assets (which for the most part depend on each other). The origin of hazards should be understood, as all assets are not vulnerable in the same way to hazards. Therefore, it appears that the use of the term vulnerability is actually inseparable from the concept of risk. In this regard Brooks [5] provides a reminder of some references on the definition of risk (Table 1.3), particularly that of Crichton which introduces the term 'vulnerability'.

In the same vein, the ISDR (International Strategy for Disaster Reduction) of the UN (United Nations) defines risk as follows [6]. :

> The probability of harmful consequences, or expected losses (deaths, injuries, property, livelihoods, economic activity disrupted or environment damaged) resulting from interactions between natural or human-induced hazards and vulnerable conditions. Conventionally risk is expressed by the notation
>
> $$\text{Risk} = \text{Hazards} \times \text{Vulnerability}.$$
>
> Some disciplines also include the concept of exposure to refer particularly to the physical aspects of vulnerability. Beyond expressing a possibility of physical harm, it is crucial to recognize that risks are inherent or can be created or exist within social systems. It is important to consider the social contexts in which risks occur and that people therefore do not necessarily share the same perceptions of risk and their underlying causes.

Table 1.3 Definitions of the concept of risk [5]

Author(s)	Risk definition
Smith 1996 [7] (p5)	Probability × loss (probability of a specific hazard occurrence) *Hazard = potential threat*
IPCC 2001 [8] (p21)	Function of probability and magnitude of different impacts
Morgan and Henrion 1990 [9] (p1)/ Random House 1966 [10]	"Risk involves an 'exposure to a chance injury or loss'"
Adams 1995 [11] (p8)	"a compound measure combining the probability and manitude of an adverse affect"
Jones and Boer 2003 [12]; (also Helm 1996 [13])	Probability × consequence *Hazard: an event with the potential to cause harm, e.g. tropical cyclones, droughts, floods, or conditions leading to an outbreak of disease-causing organisms.*
Downing et al. 2001 [14]	Expected losses (of lives, persons injured, property damaged, and economic activity disrupted) due to a particular hazard for a given area and reference period *Hazard: a threatening event, or the probability of occurrence of a potentially damaging phenomenon within a given time period and area.*
Downing et al. 2001 [14]	Probability of hazard occurrence *Hazard = potential threat to humans and their welfare*
Crichton 1999 [15]	"Risk" is the probability of a loss, and depends on three elements, hazard, vulnerability and exposure."
Stenchion 1997 [16]	"Risk might be defined simply as the probability of occurrence of an undesired event [but might] be better described as the probability of a hazard contributing to a potential disaster … importantly, it involves consideration of vulnerability to the hazard."
UNDHA 1992 [17]	"Expected losses (of lives, persons injured, property damaged, and economic activity disrupted) due to a particular hazard for a given area and reference period. Based on mathematical calculations, risk is the product of hazard and vulnerability."

In fact, if these definitions of risk are combined with those currently used by the French *Ministère de l'Ecologie*, risk can be defined as the confrontation between a hazard and a particular asset in a given area, having its own dynamics:

$$\text{Risk} = \text{Hazard} \times \text{Asset}$$

Where the hazard is characterized by its probability of occurrence and intensity, and the asset is characterized by its vulnerability and resilience to the hazard.

Brooks [5] suggests that two types of vulnerability can be distinguished for a given territory:

- Biophysical vulnerability. The definition of biophysical vulnerability is related to the extent to which assets are damaged, be they human or material. It therefore relates to the physical impact of the hazard on assets, both in terms of intensity and frequency. This vulnerability is analogous to the 'sensitivity' of the system to the hazard. Thresholds enable the characterisation of lethal effects i.e. the biophysical vulnerability of the population in a given territory.
- Social vulnerability. This represents the ability of a system to cope with a hazardous event, which in this case is similar to the definition of resilience. A system is therefore more or less vulnerable and resilient in the extent to which it is able, at least in part, to face adverse circumstances. Social vulnerability therefore differs from biophysical vulnerability in that it does not depend solely on the frequency and intensity of the hazard, but rather on the properties of the system itself which makes it more or less vulnerable. Property insurance is an example of a factor that reduces the vulnerability of a system because it can compensate for the losses caused by a hazard.

Following a fairly extensive literature review of the concept of vulnerability, Füssel [18] proposes a conceptual framework for applying this concept to the problem of global climate change. The framework should, according to the author, include the following six dimensions:

- A definition of the system in question (identification of the system elements and the links between them).
- The hazard that may affect the system or its components.
- The spatial scale of the study being studied and its vulnerability.
- The time scale, in order to understand system dynamics and events.
- The domain affected (biophysical, socio-economic or both).
- The exposure or sensitivity of the system and its components to the hazard in question.

This framework remains very much a classic definition of vulnerability, which is that it should be seen as a function of a particular hazard whose frequency and intensity may affect the system. This proposal seems appropriate in an attempt to characterize the vulnerability of a territory, but unfortunately the author offers few concrete examples.

The approach proposed in this chapter is primarily concerned with determining the sensitivity of assets in terms of specific thresholds related to the physical and chemical effects of potential hazards (heat flux, overpressure, toxic concentrations). These thresholds represent elements that characterize the biophysical vulnerability of assets and can be used by emergency services or town planning legislators. Moreover, our approach is also interested in the precise characterization of the territory, including the probability of occurrence of TDG accidents and meteorological parameters which make it possible to simulate hazards in order to evaluate the range of their effect on assets. The case studies that follow focus mainly on the biophysical vulnerability of affected areas. These sensitivity thresholds define the level of exposure of local assets when faced with the hazards phenomena that may result following a TDG accident. These threshold values are described in the following section.

1.6 Biophysical Vulnerability Thresholds

The characterization of biophysical vulnerability (for example, of the population or buildings) depends on their sensitivity to hazardous phenomena. This sensitivity is evaluated using feedback from technological accidents, laboratory experiments (particularly materials or live entities such as cells or animals) or from mathematical modelling.

1.6.1 Vulnerability Thresholds: Temperature

The phenomenon causing the greatest damage during a BLEVE is the formation of a fireball. Heat radiation from a fireball can exceed many tens of metres and last for several seconds. The thermal effects on human beings (Fig. 1.2) are characterized as:

- Thermoregulation disturbance. This creates a significant increase in body temperature in order to remove excess heat through sweating. This physiological phenomenon can cause the death of the person if the temperature exceeds 42°C;
- The destruction of respiratory tract tissue (the throat, trachea, bronchi, alveoli) brought about by drying and burning;
- Skin lesions of various degrees (first, second and third degree burns). It is estimated that the ability of skin to withstand heat (without tissue damage) is 2.5 kW/m^2 over a period of 30 min.

Typically, thermal effects are studied for time periods of less than and greater than 2 min and they are expressed respectively in kW/m^2 and $[kW/m^2]^{4/3}$. In the context of a BLEVE incident, assets are exposed to heat flow for a very short period, typically a few seconds, but the level of radiated energy is often very high. The sudden and high increase in temperature which characterises this event makes it is possible to calculate the thermal effects using the following equation:

$$E = \phi^n \times T$$

Where:

E = thermal effect (lethal, 1st, 2nd degree burns, etc.);
T = exposure time in seconds;
Φ = heat flow expressed in kW/m^2;
n = a constant (equal to 4/3 in the aforementioned context).

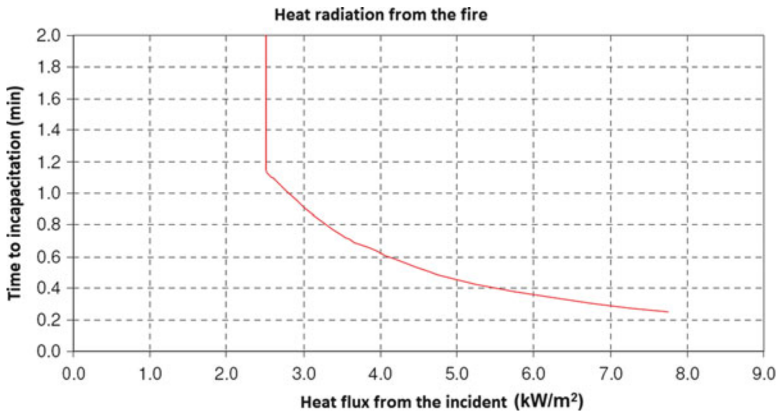


Fig. 1.2 Thermal radiation effects on human beings (following Stoll and Chianta 1969 [19], cited in [20])

Table 1.4 Thresholds of human vulnerability to heat flux [21]

Thermal flux and effects on humans	
Threshold expressed in kW/m^2	Effects
3	Irreversible effects (i.e. 600 $[kW/m^2]^{4/3}$)
5	Potentially lethal (i.e. 1,000 $[kW/m^2]^{4/3}$)
8	Probably lethal (i.e. 1,800 $[kW/m^2]^{4/3}$)

Table 1.5 Thresholds of material vulnerability to heat flux [21]

Thermal flux and effects on materials	
Threshold expressed in kW/m^2	Effects
5	Widespread damage to windows
8	Domino effects caused by the spread of fire (e.g. paint blistering, plastic deformation)
16	Resistance of most buildings to prolonged exposure (excluding concrete). Wood and plastics ignite or are destroyed.
20	Resistance of concrete for a period of several hours.
200	Destruction of concrete within half an hour.

Tables 1.4 and 1.5 show the vulnerability thresholds of human and materials to thermal flux.

Determining the consequences of heat flux on humans (as with pressure or toxicity) does not take into account individual characteristics such as age, general physical condition, etc. However, the elderly, children and those who are already ill are more susceptible to these phenomena than people in other age groups or who are in good physical condition. These characteristics, together with the age and sex of those exposed will be discussed in the chapter concerning the study of territorial vulnerability in the event of crisis management.

1.6.2 Vulnerability Thresholds: Overpressure

Overpressures particularly occur during a UVCE-type explosion. The pressure wave propagated through the atmosphere is due to the release of a gas which is initially at an almost constant volume. This pressure wave has various consequences on buildings and humans depending on the amount of combustible gas, the relief of the surrounding environment, wind speed, etc. It can be thought of as a sequence of shocks of varying intensity, and can also be termed a shock wave.

These overpressures are the cause of three types of injury to humans [22]. Injuries vary mainly depending on the distance of the victim from the source of the explosion:

- First-order injuries are caused by the direct effect of the shock wave on limbs and organs. One part of the energy of the originating wave is reflected while another portion propagates through the tissues. This latter part of the wave is divided into a pressure wave that penetrates tissues and another spalling wave (also called a stress wave). The impact of the shock wave on the chest causes damage to the lungs. Injuries include distension and rupture of the alveoli, haemorrhagic lesions of the pleura and alveoli, gas embolism caused by rupture of the alveoli, etc. The shock wave can cause lesions on the gastrointestinal tract (bleeding and hematoma of the colon, liver, etc.), perforation of the eardrum, broken bones, cardiovascular problems (lowering of the heart rate for several minutes, drop in blood pressure, etc.), brain damage (haemorrhage, lesions, etc.). Depending on their impact on the organs and limbs, these injuries have consequences that range from the immediate death of the victim, through unconsciousness, to respiratory disorders, etc.
- Second-order injuries are caused by flying debris hitting the victim. It can result in limbs being effectively amputated and/or multiple perforations of limbs and organs. Depending on the size and position of the wounds, this can in the most severe cases cause the death of the victim.
- Third-order injuries correspond to the victim being thrown by the force of the explosion. These injuries are not caused by the shock wave itself, but by the blast produced by the movement of air in contact with the wave. Death is most often caused in this case by a serious head injury.
- Fourth-order injuries are also the consequence of the blast. They correspond to victims being crushed by falling buildings and the post-traumatic stress experienced by victims who survived the explosion.

Tables 1.6 and 1.7 show the effects of overpressure on humans and buildings respectively, using thresholds measured in millibars.

The consequences of overpressures on human health may be significantly affected by the general condition of the victim.

Table 1.6 Thresholds of human vulnerability to overpressures [20]

Overpressure and its effects on humans	
Threshold in mbar	Effects
20	Irreversible effects caused by breaking glass (class II injuries).
50	Significant danger to human life with irreversible effects (class II and III injuries).
140	Serious danger to human life with possible lethal effects for 1% of the population (class II, III and IV injuries).
200	Probable lethal effects for at least 1% of the population (class I, II, III and IV injuries).

Table 1.7 Thresholds of material vulnerability to overpressures [20]

Overpressure and its effects on materials	
Threshold in mbar	Effects
20	At least 10% of windows destroyed, slight damage to houses.
50	At least 50% of windows destroyed, slight structural damage to buildings and window frames.
140	Serious damage to buildings: partial destruction of walls, roof tiles, minor damage to metallic structures.
200	Domino effects: breakage of metallic structures, movement of foundations, slight splits in steel tanks (e.g. fuel tanks), damage to machinery, partial damage to buildings made of brick, damage to walls made of breeze-blocks.
300	Serious damage to buildings: destruction of light industrial buildings, steel-framed buildings, storage tanks, deformation of pipework.

1.6.3 Vulnerability Thresholds: Toxic Substances

The toxicity thresholds for hazardous substances depend on the nature of the substance. International research in this field has led to the creation of a large number of databases which reflect the wide variety of substances. Classically, vulnerability thresholds are defined according to the length of time the subject is exposed and the extent of lethal or irreversible effects. IDLH (Immediately Dangerous to Life or Health) values can be used to form a lower threshold for irreversible effects. It is defined as maximal exposure for a period of 30 min without serious damage to health.

1.6.3.1 Ethylene Oxide

The case study carried out in the area surrounding Nice (south-east France) on the risks created by the transportation of dangerous goods by road focused on ethylene oxide. According to the *Direction Départementale de l'Equipement* (a regional organisation responsible for public infrastructure), 'a comparative risk analysis of all hazardous material transportation showed that the transport of ethylene oxide comprises 5% of TDG traffic volume on the Nice bypass, but contributes 43% of TDG risk' [23].

Consequently, local authorities regulated road transportation of ethylene oxide. In 2007 its road transportation was banned and currently only rail transport is allowed. It should be noted that the road transport of ethylene oxide is still possible in the case of a strike by rail workers.

Ethylene oxide (C_2H_4O) is a sterilising agent used to kill bacteria, moulds and fungi, particularly in hospitals and in the food industry (e.g. to sterilize spices). It is also used in industry for the synthesis of ethylene glycol. Ethylene glycol is used as antifreeze, for polyester and detergent production, etc. Ethylene oxide takes the form of a colourless gas. It has a boiling point of 10.6°C, it is heavier than air (vapour density = 1.49), and it has a sweet odour which is detectable in concentrations of about 300 ppm at 20°C and 1,013 mbar or 1 atm [23]. It reacts readily with other compounds and can be extremely explosive at high temperature or if contaminated by acids, bases, salts, combustible materials, oxidants, iron chlorides, aluminium, boron, tin and iron oxides (rust).

The risks to humans and animals of ethylene oxide are primarily associated with its inhalation. Depending on the concentration ethylene oxide causes irritation of the mucous membranes and eyes, respiratory disorders that can lead to the formation pulmonary oedema, digestive disorders (nausea, vomiting, diarrhoea) and neurological disorders (headache, drowsiness, muscle weakness, lack of coordination, coma). Ethylene oxide has also mutagenic (chromosomal aberrations in lymphocytes) and carcinogenic (leukaemia and lymphomas) effects following short-term exposure at high concentrations or long-term exposure at low concentrations. Finally, this substance can have reproductive effects, including miscarriage, premature birth or a prolongation of pregnancy.

The website of the National Institute for Occupational Safety and Health (NIOSH) provides the following reference values: the threshold for irreversible effects is set at 800 ppm for 30 min exposure (IDLH) and for lethal effects at 2,900 ppm for 1 h of exposure. This corresponds to LC50 when tested on laboratory rats. LC50 is the concentration at which 50% of the exposed population dies. NIOSH also highlights the fact that the lower limit of flammability of ethylene oxide is 30,000 ppm [24–25]: at this concentration a cloud of gas can ignite and cause a UVCE-type explosion. The three thresholds described here are used to simulate the consequences of a TDG accident.

1.7 Case Study of a Vulnerable Area: Nice – Imperia – Savona

The area chosen for this case study is the border area between Italy and France. It covers specifically the provinces of Savona and Imperia in Italy and the region of Alpes-Maritimes (particularly the coastal areas and the city of Nice) in France Fig. 1.3.

Overpasses, tunnels and the extensive urbanization of surrounding region make the area sensitive to risk and increase congestion on cross-border transportation channels. Factors such as these, together with the huge expansion of the transport

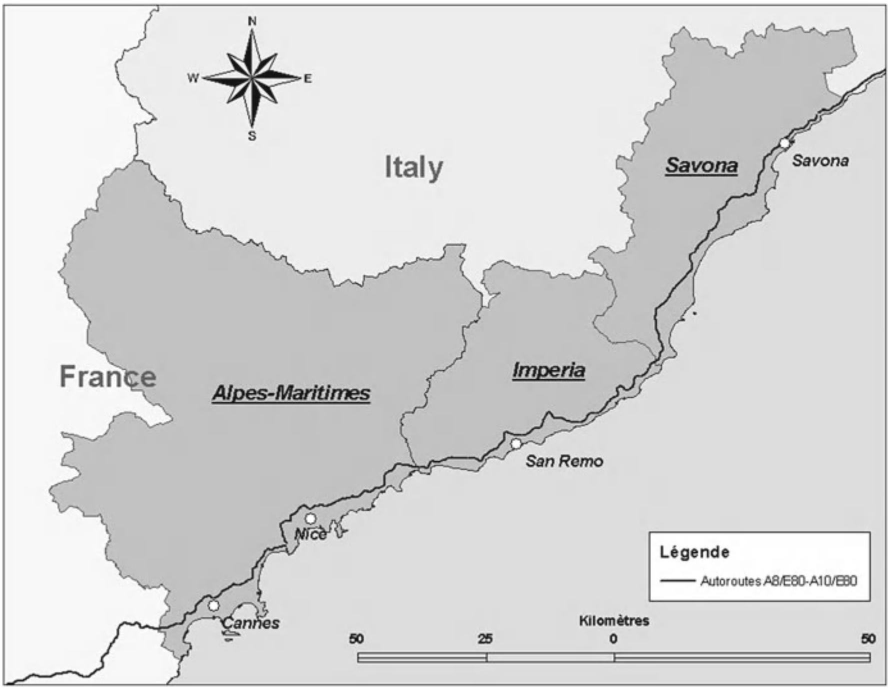


Fig. 1.3 Area covered by the case study

sector in Europe (particularly in border regions) will make it necessary in coming years to introduce and improve joint TDG planning policies in the countries concerned. It will also be necessary to improve the safety of transport infrastructure through the integration of methodologies and information and communications technologies.

A complete risk assessment requires a good knowledge of the area concerned, and the compilation of information on potential sources of TDG risk [26]. It requires an analysis of the landscape and the transport infrastructure, which takes into account sources of risk such as incoming, circulating and exiting traffic flows, and any buildings located in the vicinity of the transport infrastructure.

1.7.1 Savona

The Province of Savona comprises 69 municipalities covering 1,545 km^2. It has around 282,548 inhabitants [27], and a population density of around 182 people/km^2. In terms of the road network, the province is crossed by two motorways both carrying heavy traffic. The A10 motorway joins the cities of Genoa, Savona and Ventimiglia, while the Piedmont area, (in particular the province of Cuneo) is linked by the A6 motorway which runs between Turin and Savona. Savona is home to one of the most important tourist and commercial ports in Italy (after the city of Genoa in the region of Liguria).

The port of *Vado Ligure* (a municipality adjoining Savona) houses one of the largest deposits of hazardous materials in the area. It is a multi-purpose distribution centre where goods from the bordering regions and other Mediterranean ports are handled. It is also a major logistical node for the distribution of petroleum products and hazardous materials. While goods are transported throughout the Ligurian region, the cross-border corridor sees a considerable flow of traffic.

1.7.2 Imperia

The Province of Imperia (also in Italy) covers an area of 1,156 km^2 with 217,037 inhabitants and a population density of 188 inhabitants/km^2 [27]. It is crossed by the A10 motorway between Genoa and Ventimiglia and forms the border area with France.

1.7.3 Alpes-Maritimes

The French department of Alpes-Maritimes is in the extreme south-east of France. It covers an area of 4,299 km^2 with a population of 1,011,326 inhabitants and a population density of 235 inhabitants/km^2 (INSEE, http://www.insee.fr/en/default.asp). The A8 motorway runs through the department to the Italian border. The city of Nice, with 347,900 inhabitants, is the largest city in the department and is located 40 km from the Italian border. The principle activities of the city of Nice and the surrounding French Riviera are tourism and high technology. The industrial base is small compared to the nearby departments of Vaucluse and particularly Bouches-du-Rhone. However, because of its location (between the two major industrial centres of Liguria and Bouches-du-Rhone), Alpes-Maritimes is continuously crossed by TDG vehicles.

In terms of natural environment, the coastal areas close to the Mediterranean Sea feature hills, mountains and some flat plains. There is a complex river system caused by the contrasting geo-morphological features. The fauna and flora are particularly interesting and have a taxonomic diversity amongst the highest in Europe, featuring numerous endemic and rare species (orchids, algae, moths, beetles, etc.). The area houses several national parks and various other protected areas which promote the preservation of local biodiversity.

1.7.4 Traffic Flows

The Italian market for hazardous materials totalled exports of 23,058,000 tonnes and imports of 181,613,000 million tonnes in 2002 [28]. For the case study, data on traffic flows in Liguria and PACA (*Provence-Alpes-Cote d'Azur*, the French

Table 1.8 Annual transportation of dangerous materials by road from France to Italy (in millions of tonnes per km and millions of vehicles per km)

	Tonnes per km (millions)	Vehicles per km (millions)
Explosives	20	8
Gas—compressed, liquefied or dissolved under pressure	1,073	121
Flammable liquids	5,459	383
Flammable solids	52	4
Substances liable to spontaneous combustion	/	/
Substances which emit flammable gases	/	/
Oxidising materials	237	14
Organic peroxides	/	/
Toxic substances	54	4
Infectious substances	23	16
Radioactive substances	40	2
Corrosive substances	561	36
Mixtures of hazardous substances	530	32
Total hazardous substances	8,070	622

Table 1.9 Annual transportation of dangerous materials by road from Italy to France (in millions of tonnes per km and millions of vehicles per km)

	Tonnes per km (millions)	Vehicles per km (millions)
Explosives	32	2
Gas—compressed, liquefied or dissolved under pressure	1,712	176
Flammable liquids	6,584	338
Flammable solids	88	6
Substances liable to spontaneous combustion	24	3
Substances which emit flammable gases	/	/
Oxidising materials	170	9
Organic peroxides	12	2
Toxic substances	183	12
Infectious substances	143	14
Radioactive substances	/	/
Corrosive substances	1,209	62
Mixtures of hazardous substances	186	15
Total hazardous substances	10,361	640

region encompassing Alpes-Maritimes) was gathered. Data from Eurostat 2005 on national flows of hazardous materials between France and Italy is shown (by class) in Tables 1.8 and 1.9. In the PACA region a periodic survey was made of motorway vehicle traffic in order to create a database on TDG flows. Monitoring was carried out from September 2004 to February 2005 and flows were classified according to their direction (France to Italy or Italy to France) at various times and days of the week.

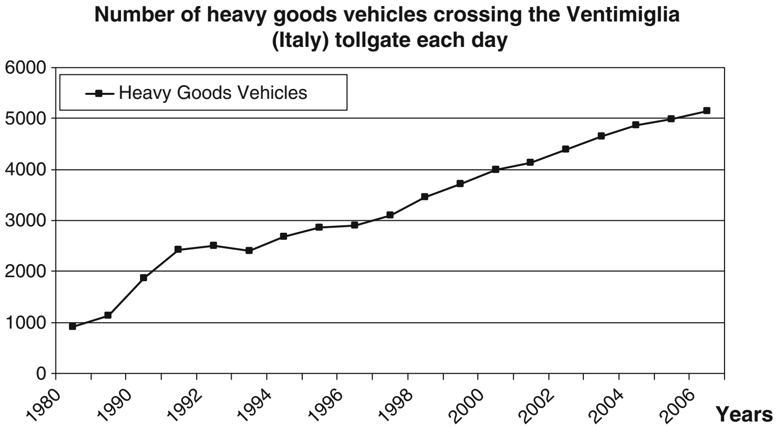


Fig. 1.4 Evolution of the daily traffic of heavy goods vehicles between France and Italy at the Ventimiglia tollgate. These figures show a steady growth in the flow of heavy goods vehicles for the past 20 years (website of the *Union Routière de France*, 2008)

However, this official data is highly aggregated and only makes it possible to establish the broad categories of hazardous materials transported. In addition, the available statistics come from surveys conducted at various times of the year and by individual actors such as motorway companies and local or state authorities. It indicates that dangerous goods transported by road represent about 5% of the total flow of merchandise transported by heavy goods vehicles.

Figure 1.4 shows the statistics generated by a survey carried out by the French highway company ESCOTA on the number and categories of vehicles passing through the toll station at Ventimiglia (Italy). These figures show that more than 5,000 heavy vehicles cross the tollgate each day and the trend has been rising for over 20 years.

Assuming that 5% of these vehicles are carrying hazardous materials (the average percentage suggested by the statistics), the number of TDG vehicles which crossed the tollgate at Ventimiglia can be estimated to be about 250 per day, 6,500 per month or 78,000 per year, excluding Sunday, when heavy goods traffic is not allowed on the road (see Fig. 1.5).

The final destinations of vehicles travelling from Italy to France show some interesting characteristics. 39% of heavy goods vehicles (all categories) end their journey in France (the PACA region) or the neighbouring region of Languedoc-Roussillon, while 57% of vehicles travel to Spain and Portugal. There is virtually no flow towards northern Europe. However, it is important to point out that these figures are estimates derived from surveys made over very short time periods (e.g. 7 days) and only a few hours (e.g. from 9 a.m. to 11 a.m. on 8th September, 2003, and from 10 p.m. to 1.15 a.m. on 9th September, 2003). These surveys only capture a very small part of TDG flows which makes it difficult to build up an overview of TDG risk in the area.

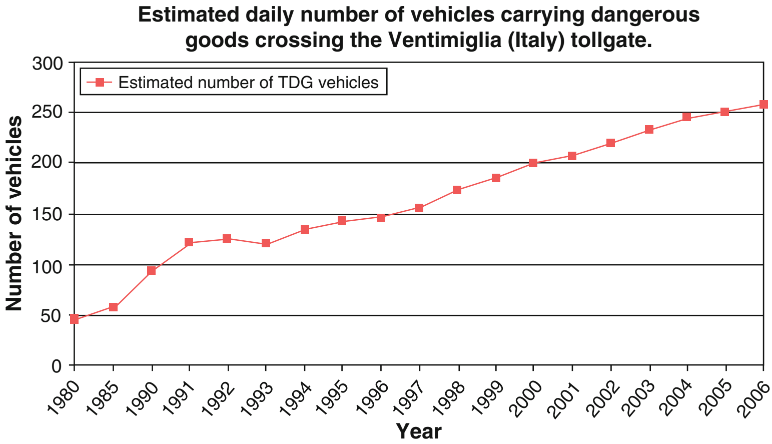


Fig. 1.5 Estimated daily number of TDG vehicles crossing the Ventimiglia (*Italy*) tollgate from 1980 to 2005. In this region the number of TDG vehicles has constantly grown to reach 250 per day

The next section presents a method for estimating the vulnerability of assets when exposed to the consequences of an accident or malicious act involving TDG road transport.

1.8 A Proposed Model for Damage Estimation

In order to synthesise the various features of the areas studied, the damage estimation model draws information from four main sources:

- Identification of areas where there is a high probability of an accident. This is based on historical information (accident rates) or on modelling of the accident risk.
- Characterization of assets in terms of their nature and their spatial distribution. This takes account of their potentially heterogeneous distribution across the exposed area.
- Identification of the nature and, to the greatest extent possible, flows of dangerous goods in the area.
- Simulation of dangerous phenomena. The aim is to estimate the extent of their effects in terms of distance. For simulations of toxic air releases this must also take into account the values of environmental variables such wind speed and direction.

Simulations of hazards caused by TDG accidents aim to provide information that is useful for authorities and rescue services. This information can help both in the preparation of contingency plans and in crisis management. It can also provide valuable input for local authorities and authorities responsible for land management. Data is presented as a measurement of the distances affected by dangerous effects such as:

- heat flux in a BLEVE-type explosion;
- overpressure waves in a UVCE-type explosion; and
- toxic concentrations in the case of a release of atmospheric pollutants.

These distances are calculated using the thresholds for lethal effects, irreversible effects and structural damage. Knowledge of dangerous goods flows makes it possible to select some examples to use as the basis for these (type 'D') accident scenarios. An ALOHA software simulation of dangerous effects can then determine propagation distances. These distances (or hazard perimeters) are integrated into a Geographic Information System (GIS) in order to estimate the number of people, buildings and establishments open to the public that are exposed to danger, using the corresponding thresholds. These estimates are then implemented in various models (developed in partnership with the University of Genoa). The results bode well in providing assistance for:

- The preparation of contingency plans based on scenarios which take into account the type of dangerous goods, the weather, the distribution of exposed assets in the area, etc.
- Public planning of TDG routes which must take particular account of the potential exposure of assets.
- Crisis management. The simulation of the possible effects of a dangerous situation as it develops (e.g. fire in an LPG tank) helps to prepare for an intervention by emergency services, to identify hazard perimeters, etc.
- Post-crisis management. An event which has already occurred can be simulated using feedback elements. This approach enables areas to be identified where use restrictions may need to be imposed on the population (drinking water consumption etc.) or health measures must be applied (medical toxicology follow-up, drug administration, etc.). This information is also useful for optimizing sampling strategies to detect damage, to implement pollution control devices and to provide insurers with information about particularly badly damaged sectors, etc.

1.8.1 The Damage Estimation Model

The impact of a TDG accident or malicious act that affects the local population must be assessed. This is both in order to estimate the level of risk that assets may be exposed to, and also to help governments, local authorities, infrastructure managers and the private sector in their approach to risk prevention, land management and planning of response capabilities.

The damage estimation model (Fig. 1.6; [29, 30]) uses information gathered by an optical recognition prototype for TDG detection and identification, together with data provided by sensors embedded in TDG vehicles. This information is transmitted

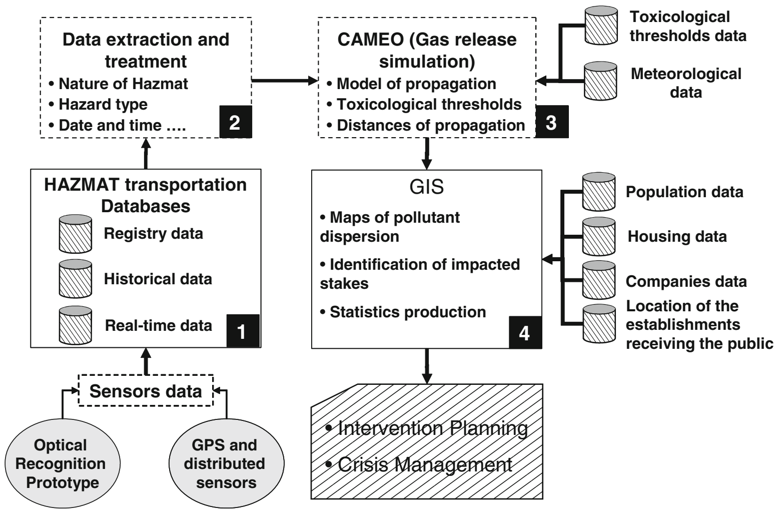


Fig. 1.6 The damage estimation model following a TDG accident

to a common database (1) that collects historical data, real-time data and data on the identity of vehicles on the road equipped with embarked systems.

A set of database queries (2) is implemented to extract information useful for policy makers. For example, it is possible to see, for a given period, which dangerous materials have transited the territory concerned and their associated hazard. In the case study, transits of ethylene oxide were identified from information stored in the common database and retrieved through the query module.

Knowledge of the movements of transported materials enables possible accident scenarios to be drawn up. Further analysis with ALOHA software makes it possible both to determine the atmospheric dispersion of toxic substances (in the form of a gas at room temperature and atmospheric pressure), and to estimate the distance of the thermal flux and overpressure effects caused by different types of explosion (3). The ALOHA software is described in detail in the next section which discusses the simulation of atmospheric dispersion.

The results of the simulations of atmospheric dispersion of toxic substances and explosive phenomena are then integrated into a GIS to provide a geographical representation of the territory affected (4). Quantum GIS (QGIS) was used to integrate spatial data set in order to generate maps and to estimate the number of assets exposed.

According to ESRI France, 'A Geographic Information System is a computerized tool to represent and analyse all the things that exist on earth and all events that occur there. GIS harnesses all the power of databases (such as the ability to run queries and statistical analysis). The major challenges we face today (environment, demography, public health, etc.) are all closely related to geography'.

Territorial modelling uses software to manipulate the geographic information. It integrates spatial data derived from national bodies, government ministries, local authorities or private companies into a GIS. The area can then be characterized in terms of the nature and location of assets such as population centres, private properties, buildings open to the public, waterways, transport infrastructure, conservation areas, etc. The integration of this information enables an approach to risk analysis that frames the vulnerability of the area in terms of the vulnerability its assets.

To implement this vision, a GIS must provide several modules for the display and manipulation of geographic information; a geo-database module, a geo-visualization module and a geo-manipulation module. With this in mind, the GIS software chosen for the study was Quantum GIS.

The Quantum GIS project began in earnest in May 2002 when the first lines of code were written. The idea was born when Gary Sherman attempted to find a GIS viewer for Linux that was fast enough and supported a wide range of formats. Initially, QGIS was hosted on the SourceForge platform (from July 2002) and the first version was published on 19th July. The first version only supported PostGIS software and was barely functional. The name Quantum GIS has no real significance other than that it starts with a Q and QGIS uses the cross-platform application framework known as Qt (from Nokia). Quantum GIS software is published under the GNU Public License. QGIS is one of the official projects of the Open Source Geospatial Foundation (OSGeo). It runs on Linux, UNIX, Mac OS X and Windows. It supports many vector and raster formats and is compatible with multiple databases. Quantum GIS has an ever-increasing list of features via a powerful extension system. It is possible to view, manage, edit and analyse data and create printable maps.

The GIS used in the study currently includes the following characterization of the area from information made available by the city of Nice for the purposes of this study:

- The local population is grouped into clusters. This data comes from the ISLOTS database run by INSEE (the National Institute of Statistics and Economic Studies) and represents the most densely populated urban areas as a block. Blocks (or small groups of blocks) are divided, where communal or cantonal boundaries make it necessary.
- Communication channels (highway, roads, etc.). This data comes from the 'ways' theme of the mapping reference database (BD CARTO®) maintained by the French National Geographic Institute (IGN).
- All buildings. This data comes from the 'buildings' theme of the same IGN source.
- 'ICPE' (Classified Installations for the Protection of the Environment) installations. This data comes from the ICPE theme (source as above) and indicates the location of ICPE installations identified by the city of Nice.
- Buildings open to the public. This data comes from the 'ERP' theme (source as above) and shows the location of buildings open to the public (schools, hospitals, churches, stadiums, museums, etc.) identified by local authorities.

This information can also be supplemented by data from other agencies such as:

- Various types of protected natural areas (e.g. the Natura 2000 network, Natural Areas of Ecological Interest, areas under bio-topic protection orders, nature reserves).

- Rivers appearing in the CARTHAGE database (a database for the thematic mapping of water agencies) provided by SANDRE (the French Service for the National Administration of Reference Data on Water).

The overlay of maps showing the dispersal of toxic substances and explosive phenomena with maps showing the distribution of assets helps to identify the proportion of people, property and natural environments exposed. Information on assets (similar to that provided by INSEE) was used to estimate the assets exposed to BLEVE and UVCE-type explosions in Italy, particularly the district of Savona in Liguria.

The following sections describe, respectively, the results of a simulation of toxic air releases and explosions (thermal effects and overpressure waves) on highly urbanized areas in the city of Nice and the city of Savona. These results describe both the extent of effects and the assets exposed.

1.8.2 Simulation of the Atmospheric Dispersion of a Toxic Substance in the City of Nice

The northern outskirts of Nice are crossed from east to west by the A8, which is transited by a large number of TDG vehicles (250–500 per day). The nature of the goods transported varies. To provide input for motorway managers, public authorities and local authorities on emergency handling procedures, the event simulation concerned the atmospheric release of ethylene oxide in a highly urbanized area of the city.

The area selected for the simulation was chosen using statistics describing the frequency of road accidents involving heavy goods vehicles during the period 1998–2005. The selected area is in the north of Nice, and has a population of over 60,000 divided into several districts: *le Ray*, *Borriglione*, *Saint Sylvestre*, *Gorbella*, *le Vallon des Fleurs*, *le Rouret*, *Las Planas*, *Gairault*, and *Saint Pancrace*.

This area (Fig. 1.7) is densely urbanized (1). It comprises large-scale public housing installations located on hillsides where access is difficult (2). Private housing is also located in the vicinity. Also in proximity to the A8 is a football stadium with a capacity of more than 15,000.

In addition, *Las Planas* (situated a few meters from the exit of the A8) is the location for the Nice Tramway marshalling yard. Finally, shopping centres, local shops, schools, and public gardens and parks, etc. are distributed throughout the area.

1.8.2.1 General Approach to Toxic Waste Simulation

The approach focuses on providing snapshots which take into account the most probable characteristics of the territory (particularly the direction and speed of the wind) as preparation for disaster management. Meteorological information was

Fig. 1.7 Photographs of the main assets located in northern Nice

gathered, which provided useful values for the parameters used to model a toxic atmospheric release in the area.

The toxic release simulation was carried out using the ALOHA module of the CAMEO software developed by NOAA (the National Oceanic and Atmospheric Administration). This was initially developed to model the atmospheric dispersion of pollutants and has since been developed to enable modelling of the effects of a BLEVE and UVCE-type explosion. The software is recognized by the US Department of Homeland Security for its consistency in estimating the diffusion of pollutants. It has been used for the modelling of chemical or industrial accidents [30–32].

In France, the *Ecole des Mines d'Alès* recommended the ALOHA software in a comparative study of pollution simulation software [33]. Moreover, an evaluation of the ALOHA module by INERIS [34] highlighted its use in an operational framework (following an accident) by emergency services, for the following reasons:

- The ergonomics of the user interface make it easy for the user to understand.
- The short amount of time it takes to enter parameters and run calculations.
- Results are consistent.

The software enables the main parameters for the simulation of toxic air releases to be taken into account. These parameters are listed below, together with the values used by government services in the context of urban planning control measures:

- The type of product: ethylene oxide.
- Atmospheric parameters: these values are based on information provided by French meteorological services (temperature and rainfall monitoring) and the work of Kessler and Chambraud [36] in Alpes-Maritimes. They include:
  - Wind speed: equal to 1 m/s. This corresponds to the upper-bound scenario for atmospheric dispersion of toxic substances.
  - Wind direction: the direction chosen is the north.
  - The height of the body of the toxic cloud: the default height is 3 m.
  - The relief of the terrain: urban or forest.

  - Cloud cover percentage: ‘partly cloudy’ was chosen as the default.
  - Air temperature: 20°C.
  - Atmospheric stability: Class B (the default) was selected (moderate wind during the day) from the six classes defined by Pasquill and Smith [37].
  - The presence of a temperature inversion at a definite height: this scenario was not used in the study.
  - Humidity expressed as a percentage: the value of 50% was selected.

- Type of container: the software can represent a TDG tank as a horizontal cylinder 6 m long and 2 m in diameter. The theoretical volume is estimated by the software and corresponds to 18.8 m^3. It is possible to specify the state of the material being transported according to the temperature of the container. In this case the material is a liquid at 20°C. The software allows the user to define the level of the container contents. In this case the container is considered to be full (100% of the theoretical capacity).
- Type of leak: the software takes into account the shape (circular or rectangular) and size of the hole through which the leak occurs. In the reference scenario the hole is circular, and 10 cm in diameter. The source of the leak is 10 cm from the base of the tank.

After incorporating these parameters, propagation distances can be calculated depending on the concentration of the hazardous substance (expressed in ppm). It is then possible to assess the risk of a UVCE-type explosion, and the thresholds for lethal and irreversible effects.

1.8.2.2 Crisis Management: Assets Exposed to Instantaneous Air Releases of Ethylene Oxide

Data on wind speed and direction (Fig. 1.8) was provided by the French meteorological service. Data was recorded by a weather station located at Nice airport in the years 1950–1989 and is characterised by the following values [37]:

- A night-time northerly wind with a speed of 4 m/s.
- A day-time southerly wind at 3 m/s (a sea breeze).

While the airport lies to the west of Nice, this data also applies to area north of Nice as the city lies in a north-south valley which determines the flow of air in the area.

The simulation therefore uses the following values which can be taken as typical for the simulation of airborne toxic releases Table 1.10.

In these two specific scenarios (Figs. 1.9 and 1.10), which use the most probable values for wind speed and direction during the day and at night, the propagation distances of toxic clouds differ quite sharply when compared to the upper-bound scenario (which uses a wind speed of 1 m/s). It should be noted that these differences are in fact somewhat less significant as the gas is diluted by the wind. This is due to the small amount of gas leaving the hole.

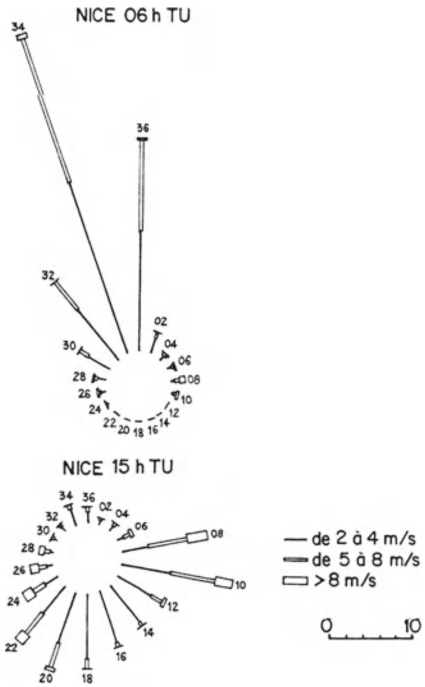


Fig. 1.8 Wind rose showing data from the weather station at Nice airport for the period 1950–1989 [38]. The figure shows a predominance of winds from the north at night and the south during the day

Table 1.10 Parameters and values used for the simulation of a toxic release of ethylene oxide

Parameter	Value used in the simulation
Type of product	Ethylene oxide in liquid form at 20°C
Wind speed	4 and 3 m/s
Wind direction	North (at 4 m/s) and south (at 3 m/s)
Height of the body of the toxic cloud	3 m
Relief of the terrain	Urban or forest
Cloud cover	Partly cloudy
Air temperature	20°C
Atmospheric stability	Class B (moderate day-time wind)
Presence of a temperature inversion	No
Humidity	50%
Type of container	Tank of 18.8 m^3 capacity
Type of leak	A circular hole, 10 cm in diameter and 10 cm from the base of the container

It is important to take into account the shape of the discharge plume, which reflects both the speed and the probable direction of the emission. The higher the wind speed, the narrower the cloud.

Moreover, Table 1.11 shows that biophysical vulnerability threshold perimeters are reached much faster than in the case of wind speed of 1 m/s. These results highlight the importance of meteorological parameters (particularly wind) in estimating hazard perimeters. This finding should be taken into account when emergency services and crisis management teams draw up an intervention strategy.

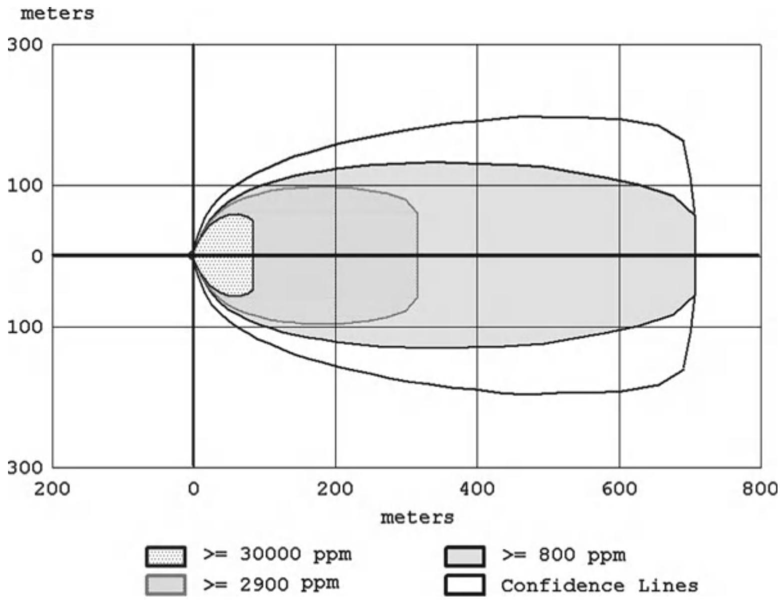


Fig. 1.9 Shape and propagation distances of a toxic cloud at a wind speed of 4 m/s. 30,000 ppm is the UCVE threshold; 2,900 ppm is the threshold for lethal effects and 800 ppm is the threshold for irreversible effects

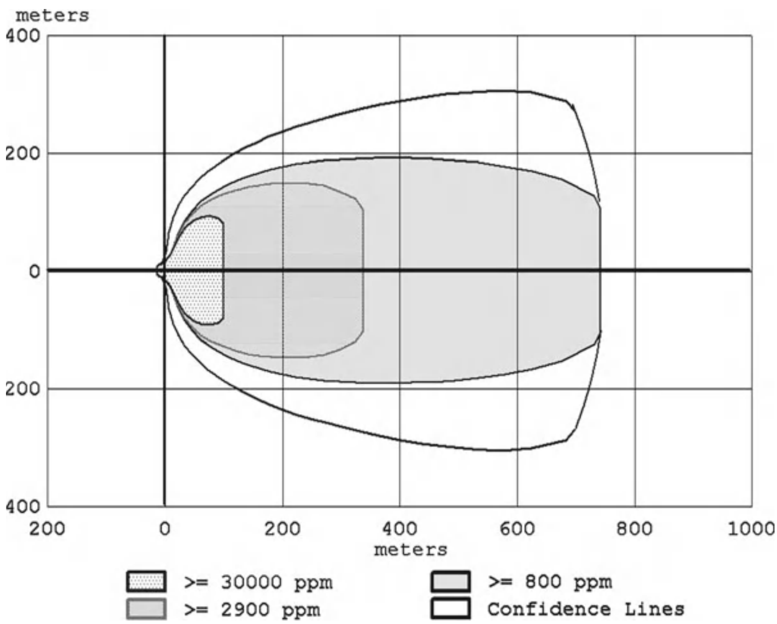


Fig. 1.10 Shape and propagation distances of a toxic cloud at a wind speed of 3 m/s. As in the previous figure, 30,000 ppm is the UCVE threshold; 2,900 ppm is the threshold for lethal effects and 800 ppm is the threshold for irreversible effects

Table 1.11 Distances and kinetics of toxic cloud formation using ethylene oxide reference thresholds for wind speeds of 1, 3 and 4 m/s

	Distances (m) and time to reach that distance (min)			
Dangerous substance	Wind speed (m/s)	UVCE	Lethal effects threshold	Irreversible effects threshold
Ethylene oxide	1	225 – 15	525 – 25	849 – 40
Ethylene oxide	3	101 – 2	340 – 5	744 – 9
Ethylene oxide	4	84 – 1	316 – 3	708 – 8

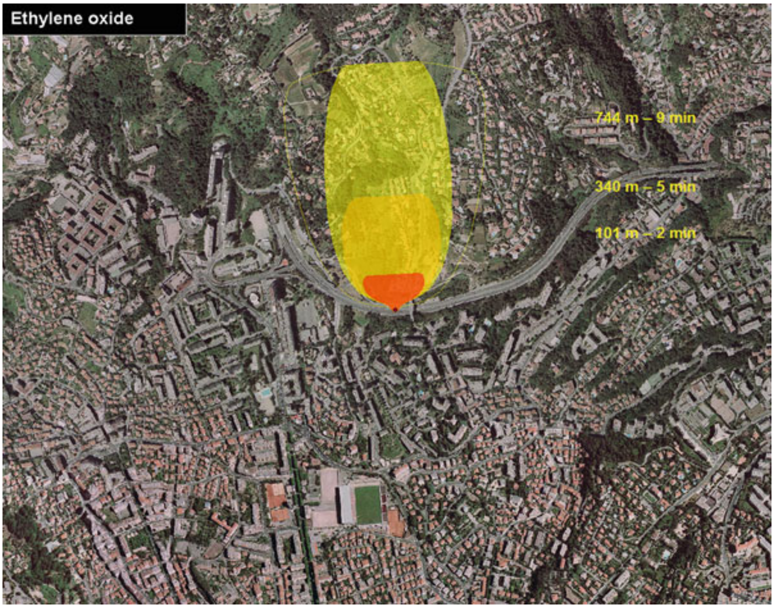


Fig. 1.11 Map of the of ethylene oxide cloud with a southerly day-time wind at 3 m/s (mapping support: IGN BD ORTHO, CRIGE PACA, Nice)

Figure 1.11 shows that with a diurnal southerly wind at 3 m/s, residential dwellings are mainly exposed to the toxic release.

This observation is confirmed by the analysis in Table 1.12. In this scenario no public buildings are affected by the toxic discharge, but private property is.

However the situation is completely different in the scenario involving a night-time northerly wind at 4 m/s (Fig. 1.12 and Table 1.13).

In this case, because of the amount of public housing exposed, the number of people affected by the discharge is multiplied by 5 (for the lethal effects threshold) and 3 (for the irreversible effects threshold). In the case of toxic concentrations having irreversible effects, the public buildings exposed correspond to schools, polling stations and other facilities including a sports stadium. It should be noted that should an incident occur when this stadium is full, about 18,000 people would be added to the population exposed to irreversible effects.

Table 1.12 Estimate of the assets issues exposed to a toxic atmospheric release with a 3 m/s southerly wind

Ethylene oxide. Southerly wind at 3 m/s

Threshold	Population exposed	Deaths	Buildings exposed	Housing exposed	Buildings open to the public exposed
UVCE	41	41	15	18	0
Lethal effects threshold	272	3	92	124	0
Irreversible effects threshold	2,105	0	510	1,117	0

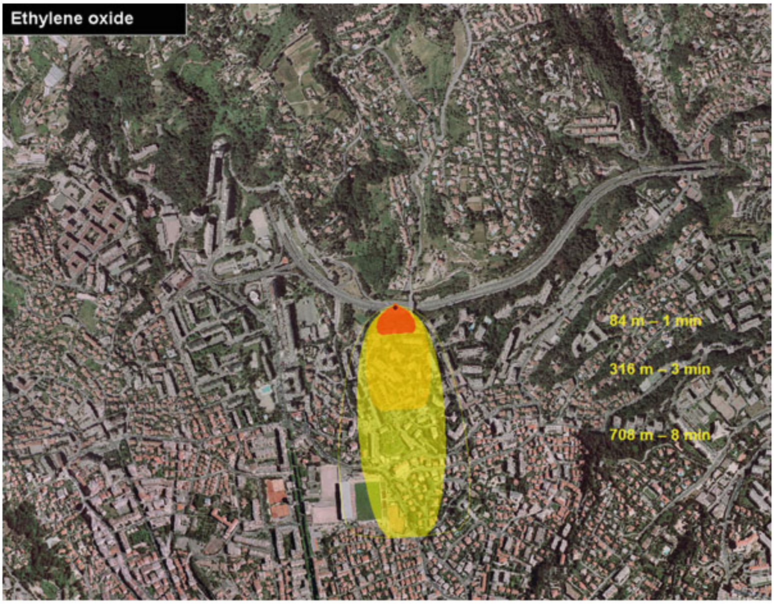


Fig. 1.12 Map of the ethylene oxide cloud with a northerly night-time wind at 4 m/s (mapping support: IGN BD ORTHO, CRIGE PACA, Nice)

This approach, based on the most probable values for wind speed and direction, seeks above all to take into account the particular features of the landscape. Timing is a key element for successful relief operations, both for the evacuation of the affected population and for the implementation of a containment strategy. These results underline the fact that emergency services have very little time to react—propagation distances for biophysical vulnerability thresholds are reached in less than 10 min. This makes it worthwhile to consider putting in place warning systems to raise an alert at the earliest possible opportunity. Moreover, these results also emphasize the value of information campaigns for government and local residents. In districts where the local population is mainly composed of immigrants, brochures should be prepared in the most common languages.

Analysis of these results demonstrates why the Alpes-Maritimes authorities wanted to ban the transport of ethylene oxide on the A8 motorway. However, other

Table 1.13 Estimate of the assets issues exposed to a toxic atmospheric release with a 4 m/s northerly wind

Ethylene oxide. Northerly wind at 4 m/s					
Threshold	Population exposed	Deaths	Buildings exposed	Housing exposed	Buildings open to the public exposed
UVCE	307	307	3	121	0
Lethal effects threshold	1,480	15	30	780	0
Irreversible effects threshold	6,255	0	201	3,351	18

substances, which are just as dangerous (particularly in terms of their toxicity), continue to use the A8. Amongst these, toluene diisocyanate (TDI) and sulphuric acid are two substances which are highly toxic to humans. It is therefore appropriate to question whether it is logical to prohibit the transit of one substance (albeit a particularly dangerous one) while other substances continue to be allowed; or if it is not better to concentrate on other prevention measures such increased TDG monitoring (with data sent directly to emergency services). These questions remain to be resolved, and the monitoring necessitated by future regulatory requirements will certainly call for the use of new communication and information technologies.

From an environmental perspective, the northern Nice scenario has no direct impact on protected sites and major rivers. However, these results should be put in context as, depending on weather conditions, the type of dangerous goods and the characteristics of the event (accident or malicious act), propagation distances may be much larger and extend to sensitive areas.

We have discussed one type of hazardous substance and one form of hole. Should the container rupture or be totally destroyed, propagation distances of the lethal and irreversible effects thresholds (in the probable wind conditions) would most likely increase tenfold, to, respectively nearly 1 km for lethal effects and 2 km for irreversible effects. In this case, the number of assets affected would of course be higher and conservation areas located further north would be affected.

In addition to modelling atmospheric dispersion, the software can also estimate the consequences of explosive phenomena such as a BLEVE or UVCE. The following section describes the approach and results of modelling an incident in the city of Savona.

1.8.3 Simulation of a BLEVE and UCVE-Type Explosion in Savona

This section simulates the crash of a fuel tanker in an area near the port of Savona. Savona is host to many fuel storage depots, and the city is traversed by numerous vehicles carrying fuels such as gasoline, fuel oil and LPG to various outlets.

Using information on the nature and routes of vehicles carrying fuel, two explosion scenarios are described: the BLEVE (Boiling Liquid Expanding Vapour Explosion) and the UVCE (Unconfined Vapour Cloud Explosion). These types of accidents generate high magnitude heat flows and overpressure waves, the consequences of which affect both the environment and humans. The two cases presented use the following working hypotheses:

- Product: hydrocarbon.
- Product temperature: 293 K (20°C).
- Air temperature: 293 K (20°C).
- Wind speed: 1 m/s.

The product mass is varied in order to assess the effects according to the load involved in the accident. The following quantities were used in simulations: 30,000, 15,000, 7,500, 3,000 and 1,500 kg. These correspond respectively to 100%, 50%, 25%, 10% and 5% of tanker capacity.

When this type of goods are moved, BLEVE and UVCE scenarios have a high probability of occurrence and a high probability of creating serious thermal and overpressure effects. For this reason baselines were established for each of the two scenarios and associated with their corresponding effects on humans and buildings; the simulation was carried out using the ALOHA software.

1.8.3.1 The BLEVE Scenario

Starting from the moment of the accident, distances can be calculated according to the quantity of product and simulated heat flux thresholds. The results are shown in Table 1.14.

Figure 1.13 shows a map of a simulated BLEVE caused by a fully loaded vehicle. The accident occurs in Savona, near the port. The map shows five different perimeters indicating the exposed areas. Each perimeter relates to a particular simulation scenario corresponding to the following values:

- The red perimeter indicates the 200 kW/m^2 threshold.
- The orange perimeter indicates the 35 kW/m^2 threshold.
- The yellow perimeter indicates the 27 kW/m^2 threshold.
- The green perimeter indicates the 8 kW/m^2 threshold.
- The blue perimeter indicates the 5 kW/m^2 threshold.

Table 1.15 shows the vulnerability of humans when exposed to these different levels of thermal flows.

In addition to the BLEVE scenario, the other main problem concerns thermal radiation. Fuel transport can be the source of a UVCE explosion, characterized by strong overpressure waves. This scenario is discussed below.

1.8.3.2 The UCVE Scenario

The UVCE simulation scenario provides distances according to the values of pressure thresholds corresponding to a particular quantity of TNT (trinitrotoluene).

Table 1.14 Distance of heat flux showing various kW/m^2 thresholds and fireball characteristic

	Heat threshold					Fireball	
Fuel mass	200 kW/m^2	35 kW/m^2	27 kW/m^2	8 kW/m^2	5 kW/m^2	Diameter (m)	Duration (s)
30,000 kg (100%)	215 m	515 m	586 m	1,015 m	1,392 m	185	12
15,000 kg (50%)	152 m	364 m	414 m	718 m	963 m	148	10
7,500 kg (25%)	108 m	257 m	293 m	507 m	681 m	118	9
3,000 kg (10%)	68 m	163 m	185 m	321 m	431 m	87	7
1,500 kg (5%)	48 m	115 m	131 m	227 m	304 m	70	6

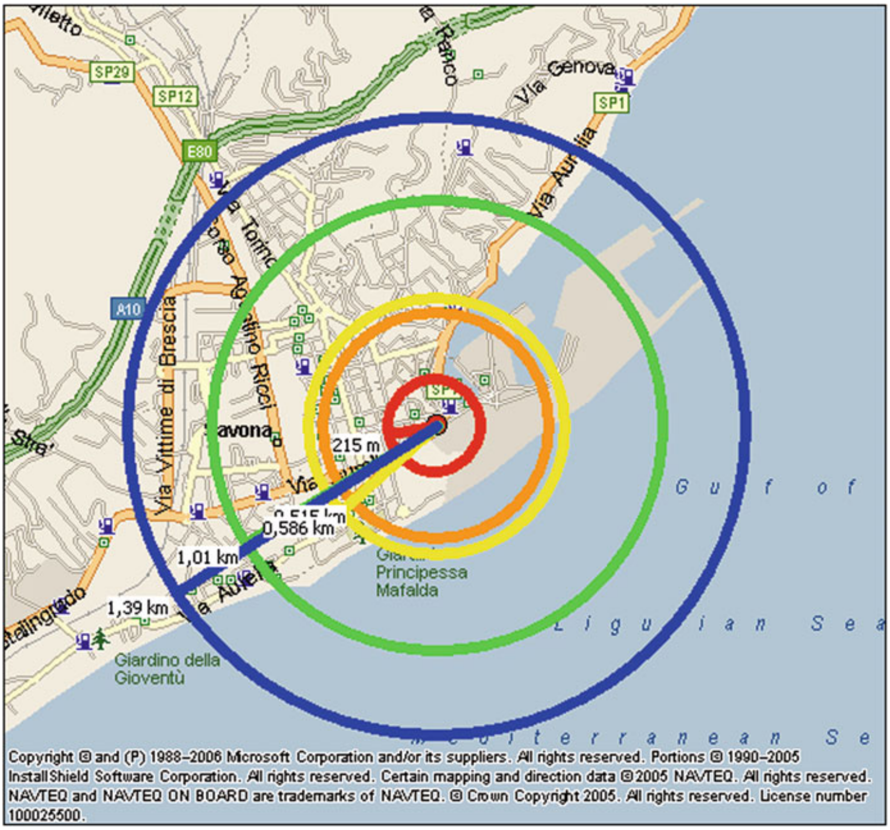


Fig. 1.13 Map of the area exposed to a BLEVE explosion caused by 30,000 kg of fuel

These thresholds are correlated with pressure levels of the shock wave produced by the explosion and their effects on buildings and humans:

- 200 mbar: significant lethal effects for at least 1% of the population (type I, II, III and IV injuries); domino effects including damage to metal structures, displacement of foundations, etc.
- 140 mbar: initial lethal effects created by a serious threat to human life (type II, III and IV injuries) for 1% of the population; serious damage to buildings including partial collapse of walls, tiles, minor damage to metallic structures.
- 50 mbar: irreversible effects brought about by a significant danger to human life (type II and III injuries);
- 20 mbar: reversible effects caused by breaking glass (type II injuries).

The results of the simulation are shown in Table 1.16.

Figures 1.14 and 1.15 demonstrate the worst-case scenario in terms of potential damage, i.e. a fully-loaded vehicle carrying 30,000 kg of fuel.

Table 1.15 Population exposed (PE) and distances depending on product quantity and heat threshold

	Heat threshold (kW/m^2) and population exposed									
	PE	200 kW/m^2	PE	35 kW/m^2	PE	27 kW/m^2	PE	9 kW/m^2	PE	5 kW/m^2
30,000 kg (100%)	423	215 m	1,377	515 m	1,542	586 m	3,088	1,015 m	3,671	1,392 m
15,000 kg (50%)	331	152 m	916	364 m	1,041	414 m	1,905	718 m	2,692	963 m
7,500 kg (25%)	204	108 m	557	257 m	634	293 m	1,318	507 m	1,714	681 m
3,000 kg (10%)	118	68 m	359	163 m	403	185 m	800	321 m	1,218	431 m
1,500 kg (5%)	78	48 m	236	115 m	224	131 m	465	227 m	801	304 m

Table 1.16 Distances to pressure thresholds and population exposed (PE) depending on product quantity and equivalent mass of TNT

	PE	30,000 kg (100%)	PE	15,000 kg (50%)	PE	7,500 kg (25%)	PE	3,000 kg (10%)	PE	1,500 kg (5%)
200 mbar	272	120,68 m	224	95,80 m	139	76,06 m	106	56,06 m	69	44,50 m
140 mbar	425	217,29 m	363	172,50 m	308	136,95 m	201	100,93 m	117	80,13 m
50 mbar	802	322,30 m	551	255,87 m	401	203,13 m	262	149,71 m	242	118,85 m
20 mbar	2,009	753,62 m	1,594	598,29 m	1,303	474,97 m	927	350,07 m	576	277,91 m

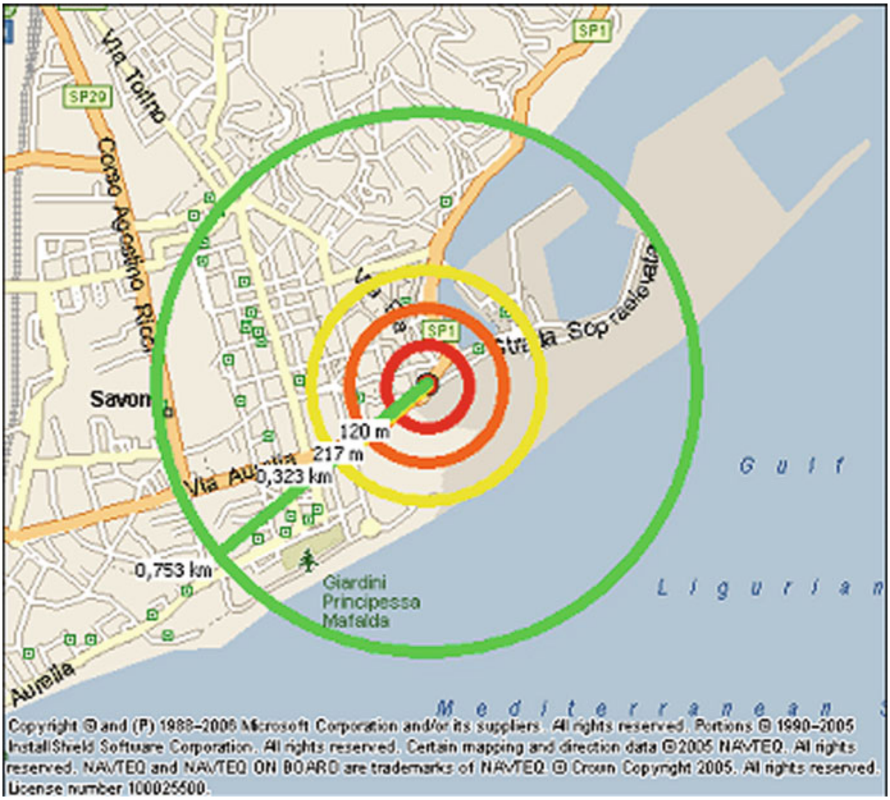


Fig. 1.14 Map of the area affected by a UVCE caused by 30,000 kg of fuel

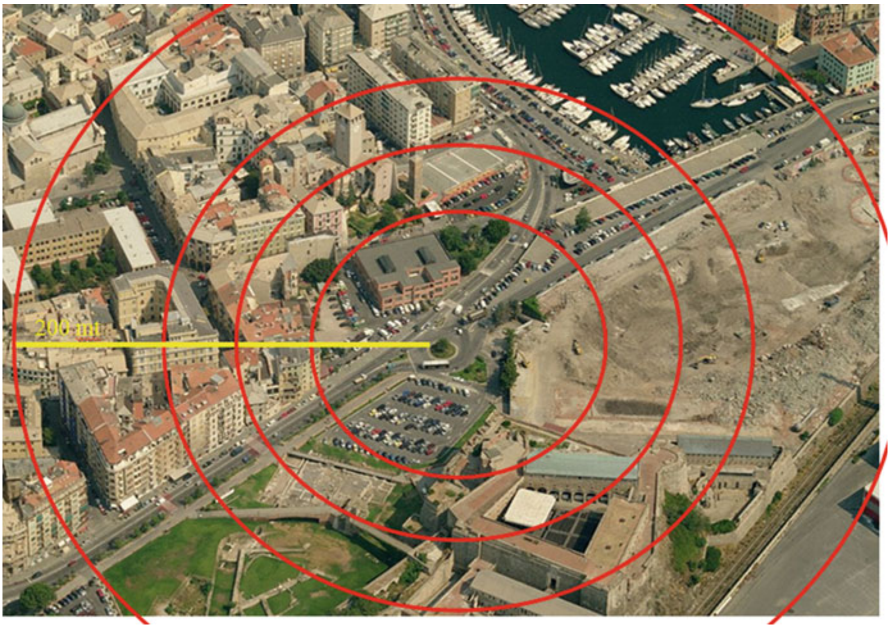


Fig. 1.15 Example of possible exposure in central Savona in the case of a UVCE

These simulations highlight the danger of an explosion caused by an accident involving a fuel tanker in a city centre. While fuel delivery is a commonplace and necessary activity in big cities, it requires particular vigilance from the actors and policy makers involved.

Some of the measures available to policy makers include local regulation of service stations, so that fuel is only delivered at certain times. This measure alone reduces the risk of accidents, especially when delivery schedules are timed for off-peak hours. Other cities have gone further and have defined secure routes which take account of the number of available road lanes, speed limits and proximity of emergency services.

1.9 Conclusion

The development of a model based on an estimate of exposed assets can demonstrate the vulnerability of particular sectors when faced with particularly dangerous accidental or intentional phenomena. The use of accident scenarios and modelling of the effects of shock waves, heat flux and the atmospheric concentration of pollutants provides a sound scientific basis for the preparation of emergency plans and land management by infrastructure managers, and public and local authorities who have to deal with the flow of dangerous goods

This approach also provides material which can assist decision-makers in their approach to crisis management, including the anticipation of consequences for people, property and the environment. The model enables the identification of the potential extent of a hazard on the surrounding area and, therefore, promotes the organization and implementation of measures to protect assets.

It also provides information for post-crisis management. The results of simulations complement field identification of the affected sectors and assets, and provide policy makers with an opportunity to:

- Optimize field studies that measure soil pollution (for example sampling strategies).
- Implement health measures in relation to the affected population (restricted use of drinking water etc.).
- Inform insurance companies of priority sectors for claim investigations.

References

1. Ministry of Ecology (http://www.developpementdurable.gouv.fr/)
2. MEDAD (Ministère de l'Ecologie et du Développement Durable) (2006) Compréhension des phénomènes et modélisation: la dispersion atmosphérique. 12 p
3. NOAA and EPA (National Oceanic and Atmospheric Administration and the U.S. Environmental Protection Agency) (2007) ALOHA: user's manual. 195 p

4. D'Ercole R, Metzger P (2009) La vulnérabilité territoriale : une nouvelle approche des risques en milieu urbain. Cybergeo: European Journal of Geography, Dossiers Vulnérabilités urbaines du sud, document 447. Available at: http://www.cybergeo.eu/index22022.html. Accessed 31 Mar 2009
5. Brooks N (2003) Vulnerability, risk and adaptation: a conceptual framework. Tyndall Centre for Climate Change Research Working Paper 38, 16 p
6. UNISDR (http://www.unisdr.org/)
7. Smith K (1996) Environmental hazards, Routeledge, London, 389 p
8. IPCC (2001) Climate change 2001: Impacts, adaptation and vulnerability, summary for policymakers, WMO
9. Morgan MG, Henrion M (1990) Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis, Cambridge University Press, 332 p
10. Random House (1966) The random house dictionary of the english language, Stein, J. (ed.), Random House, New York
11. Adams J (1995) Risk, university college London Press, London, 228 p
12. Jones R, Boer R (2003) Assessing current climate risks Adaptation policy framework: a guide for policies to facilitate adaptation to climate change, UNDP, in review, see http://www.undp.org/cc/apf-outline.htm)
13. Helm P (1996) Integrated risk management for natural and technological disasters, Tephra, 15(1): 4–13
14. Downing TE, Butterfield R, Cohen S, Huq S, Moss R, Rahman A, Sokona Y, Stephen L (2001) Vulnerability indices: climate change impacts and adaptation, UNEP Policy Series, UNEP, Nairobi
15. Crichton D (1999) The risk triangle, in Ingleton, J. (ed.), Natural disaster management, Tudor Rose, London, pp 102–103
16. Stenchion P (1997) Development and disaster management, Australian Journal of Emergency Management, 12(3):40–44
17. UNDHA (1992) Internationally agreed glossary of basic terms related to disaster management, United Nations Department of Humanitarian Affairs, Geneva
18. Füssel HM (2007) Vulnerability: a generally applicable conceptual framework for climate change research. Glob Environ Change 17:155–167
19. Stoll AM, Chianta MA (1969) Method and rating system for evaluation of thermal protection. Aerospace Medicine, 40:1232–1238.
20. Guillaume E (2006) Effets du feu sur les personnes. Synthèse bibliographique. LNE, 163 p
21. MEDAD (Ministère de l'Ecologie et du Développement Durable) (2004) Guide technique relatif aux valeurs de référence de seuils d'effets des phénomènes accidentels des installations classées, 29 p
22. Debien B, Leclerc T, Clapson P, Perez JP, Lenoir B, Pats B (2006) Lésions par explosion. Mise Au Point Anesthésie Réanimation, pp 537–555
23. DDE06 (Direction Départementale de l'Equipement des Alpes-Maritimes) (2003), http://www.alpes-maritimes.equipement.gouv.fr/
24. INRS (Institut National de Recherche et de Sécurité) (2006) Oxyde d'éthylène. Fiche toxicologique n°70. 8 p
25. NIOSH (National Institute for Occupational and Safety and Health) (2000) Preventing worker injuries and deaths from explosions in industrial ethylene oxide sterilization facilities. ALERT, NIOSH, EPA, and EOSA, Cincinnati, 32 p
26. Nicolet-Monnier M, Gheorghe AV (1996) Quantitative risk assessment of hazardous materials transport systems, topics in safety, risk, reliability and quality, ETH. Kluwer Academic Publishers, Dordrecht/Boston, 343 p
27. ISTAT (2010) Il trasporto merci su strada. Anni 2006-2007. Statistiche in breve. 7 p
28. Ministero delle Infrastrutture e dei Trasporti (2002) Conto nazionale delle infrastrutture e dei trasporti, anno 2002, 329 p
29. Garbolino E, Sacile R, Olampi S, Bersani C, Tomasoni A, Alexandre N, Trasforini E, Benza M, Giglio D (2007a) A spatial decision support system prototype for assessing road HAZMAT

accident impacts on the population in a dense urban area: a case study of the city of Nice, French Riviera. Chemical Engineering Transactions, Icheap-8 conference, 24–27 June, Ischia
30. Garbolino E, Sacile R, Olampi S, Tomasoni A, Bersani C, Alexandre N, Trasforini E, Benza M, Giglio D (2007b) Definition of a spatial decision support system for public authorities, civil protection and highway companies dedicated to the crisis management in the case of a hazmat transportation road accident in a dense urbanized area in the French Riviera. AIRO Genoa, 5–8 September 2007
31. Bellasio R, Bianconi R (2005) On line simulation system for industrial accidents. Environ Model Manage 20(3):329–342
32. Martin PH, LeBoeuf EJ, Daniel EB, Dobbins JP, Abkowitz MD (2004) Development of a GIS-based spill management information system. J Hazard Mater B112:239–252
33. Mundy D (2002) Local emergency planning committees: working together to increase chemical safety in our communities. Chem Health Saf 9(6):31–34
34. Tixier J, Dusserre G, Rault-Doumax S, Ollivier J, Bourely C (2002) OSIRIS: Software for the consequence evaluation of transportation of dangerous goods accidents. Environ Model Software 17:627–637
35. Lacombe JM, Vincent G, Baulig A, Kordek MA, Fontaine F, Tissot S (2006) Examen de l'utilisation du logiciel ALOHA-CAMEO en situation d'urgence. INERIS, Direction des Risques Accidentels, 41 p
36. Kessler J, Chambraud A (1990) Météo de la France. Tous les climats localité par localité. JC Lattès, 391 p
37. Pasquill F, Smith FB (1983) Atmospheric diffusion, 3rd edn. Wiley, New York, 440 p
38. Carrega P (1990) Vents et échelles de contraintes géographiques ; exemples en région niçoise. Publications de l'Association Internationale de Climatologie 2:83–88

Chapter 2
Legislative Context and Governance Principles for Dangerous Goods Transportation (DGT) Integrated Risk Management

Gianmarco Baldini*, Carmelo Di Mauro*, Jean Pierre Nordvik*, and Vincent Mahieu

Abstract Transportation of dangerous goods or hazardous material is a sensible security issues in the transportation sector. Dangerous goods can be transported by various means or modes (maritime, air freight, road and railway) but transportation of dangerous goods by road and rail are the most common modes. Road accidents involving dangerous goods can have critical impacts on population and environment present around the incident area. The risk of civilian casualties is increased in urban areas because of the high density of population. The danger of cascading effects in presence of industrial plants deserves also to be considered.

It is therefore highly important to identify any action or technology, which can reduce accidents and improve the safety.

This chapter is first reviewing the current state of the legislation, its rules and its limitations. It is then considering how ICT/ITS new technologies can improve the DGT monitoring, for a better dangerous goods transportation management. This assessment is based on past experiences of the Joint Research Centre involved in various demonstration and pilot projects. That first part concluding that, even if available and reliable, the advanced technologies for tracking and tracing are still not deployed, the second part will analyze what is missing for large scale deployment, and assess what are the challenges of an harmonized DGT governance at EU level.

---

*Disclaimer: the views expressed are those of the author and cannot be regarded as stating an official position of the European Commission.

G. Baldini (✉) • J.P. Nordvik • V. Mahieu
JRC Joint Research Centre, TP 361 Institute for the Protection and Security of the Citizen, Via E.Fermi, 2749 I - 21027 Ispra (VA), Italy
e-mail: gianmarco.baldini@jrc.ec.europa.eu; jeanpierre.nordvik@jrc.ec.europa.eu; vincent.mahieu@jrc.ec.europa.eu

C. Di Mauro
Risk Governance Solutions - RGS, Via Fratelli d'Italia 7, 21052 - Busto Arsizio (VA), Italy
e-mail: cdimauro.rgs@tiscali.it

E. Garbolino et al. (eds.), *Transport of Dangerous Goods*, NATO Science for Peace and Security Series C: Environmental Security, DOI 10.1007/978-94-007-2684-0_2,

2.1 Introduction

Why are we interested in the transportation of Dangerous Goods?

The first objective the European Commission is pursuing in that context is to improve the road safety. Several programs have been devoted to those objectives: 2003–2010 European Road Safety Action Programme, the European Road Safety Charter, eSafety initiative launched in 2002 by the Commission, a finally the new Road Safety Action plan to be launched autumn 2010.

For what the transport of dangerous material it is in general concerned, in Europe it is mostly performed by road and increasing steadily. In 2005 road transport of dangerous goods reached 78 billion tonne-kilometres (tKm) for the EU-25 and Norway, with flammable liquids accounting for 60% (from [1]). In that trend of continuous increase of road traffic and freight volume, the transportation of dangerous goods increases accordingly, and involves risks and potential harm to the drivers of trucks, to the road users, to the population and the environment (Fig. 2.1).

The transportation sector is a key element in the economy of a nation. It is responsible for moving millions of passengers and millions tons of goods through infrastructures deployed for thousands of kilometres both inside and across nations. The way of life of the citizens and their safety is highly dependent on transportation and the growth of a nation may be limited in the absence of adequate transportation infrastructures (Table 2.1).

The transportation sector includes many different modes of transport (maritime, road, railway, inland waters, aviation and pipeline) and a wide range of applications.

Some of these applications are quite relevant to the security and safety of the society and the general public. For example, in case of natural disasters, transportation is essential to bring in time the essential material to face a crisis.

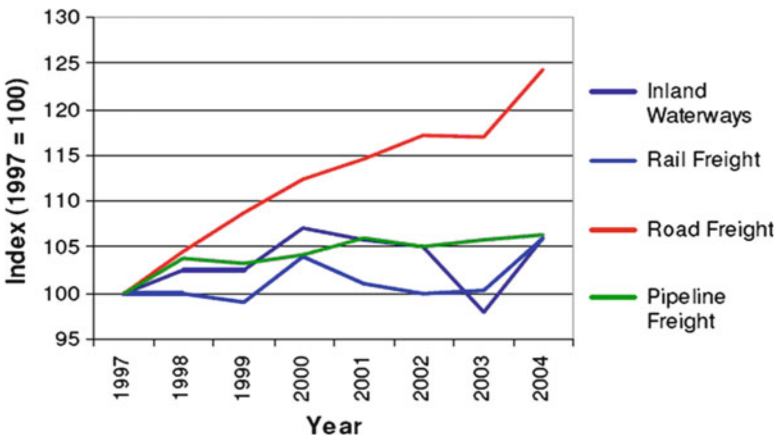


Fig. 2.1 Growth in tonne-kms by mode in the EU-15 (*Source: EUROSTAT*)

Table 2.1 Land freight transport and dangerous goods transport by mode in EU15

	Transport (billion tonne-kms)	Share of national transport (billion tonne-kms)		
Transport mode	Freight total	Dangerous goods	Freight total	Dangerous goods
Road	1,100	64	990	58
Rail	218	28	131	19
Inland waterway	99	19	29	6
Total	1,417	111	1,150	84

Source: EC – Council of the European Union – SEC(2006) 1726 – Data year 2002

Transportation relies on technologies and infrastructures to provide efficient services. In particular the transportation sector has been increasingly dependent on Information and Communication Technologies (ICT) components and services to improve the exchange of information, implement more sophisticated applications for the transportation users and increase the synergy among the various elements of the transportation chain. The growing dependence of the transportation sector on ICT including geo-location services like Global Positioning Systems (GPS) has been presented in [2, 3]. The various transport applications are examined, under the following three headings: operation and management of networks (all modes), information and guidance to the users (of the transport systems), operation and management of freight transport systems.

Geo-location services are particularly important in the transportation sector because of the mobility of its components. The knowledge of the position of the transportation vectors is an essential functionality to increase the overall security of the supply chain. This functionality is also called “tracking and tracing”. “Tracking & Tracing” targets a class of services for land applications, requiring the knowledge of the position of assets or persons, as provided by satellite navigation technologies. While GPS or the Russian GLONASS (GLObal NAvigation Satellite System) is still predominant, new navigation systems are already active or they are planned for deployment in the next years. They include augmentation systems like the American WAAS (Wide Area Augmentation System), the European EGNOS (European Geostationary Navigation Overlay Service), the Indian GAGAN (GPS Aided Geo Augmented Navigation) or new full navigation systems like the European Galileo or the Chinese COMPASS. In the rest of the chapter, we will use the term Global Navigation Satellite Systems (GNSS), to describe all the existent or planned satellite navigation systems, which provide geo-location services.

The role of GNSS in the transportation sector has grown considerably in the recent years. From maritime transport, the use of GNSS has expanded to road transport and aviation. GNSS is used to control the position of the maritime vessels in the sea, the position of trucks in fleet-management applications and it will have an important role in aviation thanks to improved accuracy and integrity of new GNSS augmentation systems like EGNOS and WAAS.

In this chapter, we will describe also the use of GNSS for the application of safe road transport of dangerous materials. GNSS technology is presently used for tracking & tracing functions and remote control of the dangerous material shipment.

Dangerous goods, because of their physical, chemical (physicochemical) or acute toxicity properties, present an immediate hazard to people, property or the environment. Types of substances classified as dangerous goods include explosives, flammable liquids and gases, nuclear or radioactive material, corrosives, chemically reactive or acutely (highly) toxic substances.

To reduce these risks, it is important that fleet managers and public safety organizations know the position of the transported dangerous goods at any time to prevent an emergency crisis or to improve the crisis response phase. Technologies like Global Navigation Satellite Systems (GNSS) and wireless telecommunications can be used to track and trace the transport vectors (e.g. trucks) of dangerous goods on the roads. These technologies have vulnerabilities, which may affect the reliability and security of the tracking and tracing system. This chapter provides an overview of the vulnerabilities and related countermeasures, which can implemented through technologies, procedures or good practices. Finally that introduction describes the results of the FP6 projects, where some of the described countermeasures were applied to a real system to provide reliable tracking and tracking of dangerous goods on the road.

Specifically on nuclear transport, about 20 million packages of all sizes containing radioactive materials (which may be either a single package or a number of packages sent from one location to another at the same time) are routinely transported worldwide annually on public roads, railways and ships (from [4]).

Between 1990 and 2002, the amount of moved dangerous goods has increased by 13%, from 98.3 to 111.1 billion tonnes-km per year in EU. The transportation of dangerous goods is regulated by a number of policies at national, European and global level.

Regulations have the purpose to minimize the risk of incidents and guarantee an effective response. They establish rules for the shipment of dangerous goods, and they fix constraints for their storage and transit in sensitive areas.

In Europe, the disappearance of the national ground borders thanks to the Schengen treaty has internationalized the issue of safety especially for the transport of dangerous goods. Real-time tracking of goods across nations is an important enabler to ensure control and safety for the society and the citizens.

The European Commission has financed a number of projects on the use of GNSS for tracking and tracing of goods. The authors have participated to the MENTORE (iMplemENtation of GNSS tracking & tracing Technologies fOR Eu regulated domains) project managed by the European GNSS Supervisory Authority (GSA) through the sixth Framework Programme funds (see [5]). In the pilot projects, GNSS (GPS/EGNOS enabled) devices are mounted on the vehicle, carrying dangerous goods to collect information on position, date and time. The information is then transmitted through a commercial communication system (e.g. GSM/GPRS/UMTS or satellite) to a control centre, where these data are compared against a predefined path. Any deviation generates an automatic alarm to notify law-enforcement agencies and the fleet managers. The system can be integrated by a risk analysis of the territory, where the most vulnerable areas in relation to the type of transported dangerous goods are identified.

Security aspects are an essential element in the transport of dangerous goods. In this paper, we will focus on the security aspects and vulnerabilities of the GNSS and communication components of the system, which have the task to collect and distribute the location information of the transport vectors. For example: GNSS or telecommunication systems signals could be jammed to cause Denial of Service (DoS) or spoofed to provide a false position.

2.2 The Current Regulatory Context

The definition of a regulatory framework is particularly important for the transport of dangerous goods, because of the associated risks for the populations and physical assets. Regulations are needed to prevent accidents to persons or properties or negative impacts to the environment.

One major challenge is the presence of different regulations for each country and for different transport applications. Dangerous goods are also subject to a number of regulations not directly related to transport like work safety regulations, consumer protection regulations, storage regulations and environment protection regulations.

To ensure consistency among various hazard classification criteria and hazard communication tools, the United Nations has started initiatives like GHS or Globally Harmonized System of Classification and Labeling of Chemicals, which is described in [6]. The purpose of GHS is to ensure that information on physical hazards and toxicity from chemicals is readily available in order to enhance the protection of human health and the environment during the handling, transport and use of these chemicals.

The European regulatory framework is part of a more global regulatory framework for a number of reasons including:

- Member states of Europe conduct trading outside the boundary of the European Union.
- The EU directives are based or they are dependent on international regulations. See for example the European agreement concerning the International Carriage of Dangerous Goods by Road (Ref. [7]).
- Member states may participate independently to regulatory bodies on road transportation.

An overview of the main road applications and correspondent regulations for nuclear transport and dangerous goods is presented in Table 2.2 both at international and European level.

At European level, we have different types of legislative acts:

- Regulations are legislative acts, which define compulsory rules to be applied to all member states in Europe.
- Directives are legislatives acts, which must be translated to European laws.

Table 2.2 Regulation for transportation of dangerous goods by road

	International regulation	European regulation
Nuclear transport	INFCIRC/225/Rev.4 on the Physical Protection of Nuclear Material and Nuclear Facilities (requirements during transport)	DIR. 1994/55/EC (on the approximation of the laws of the Member States with regard to the transport of dangerous goods by road) emended by DIR 1996/86/EC, DIR. 1999/47/EC, DIR. 2000/61/EC;
	IAEA safety guidance TS – R –1 "Regulations for the safe transport of radioactive material",	DIR. 2001/7/EC and DIR 2003/28/EC; Commission Decision 2002/886/EC; DIR 1995/50/EC (on uniform procedures for checks on the transport of dangerous goods by road), emended by DIR. 2001/26/EC. Consider also Council Regulation (EEC) No 2219/89 of 18 July 1989 on the special conditions for exporting foodstuffs and feeding stuffs following a nuclear accident or any other case of radiological emergency.
Dangerous goods	UN Recommendations on the Transport of Dangerous Goods. Convention on Civil Liability for Damage Cause during Carriage of Dangerous Goods by Road, Rail and Inland Navigation Vessels (CRTD)	DIR. 1994/55/EC (on the approximation of the laws of the Member States with regard to the transport of dangerous goods by road) emended by Directive 1996/86/EC, Directive 1999/47/EC, Directive 2000/61/EC and by 2004/112/EC; Amended also by 2003/28/EC and 2006/89/EC DIR. 2001/7/EC and DIR. 2003/28/EC; Commission Decision 2002/886/EC; DIR. 1995/50/EC (on uniform procedures for checks on the transport of dangerous goods by road), emended by DIR. 2001/26/EC.

- Recommendations are not compulsory but they can evolve into proposals for regulation or directives.

Even if they are not strictly legislatives acts, the following documents can affect the regulation process:

- Best practices are related to policies and/or to voluntary regulations (self-regulations) aimed at the collective benefits (normally for social and welfare interest).

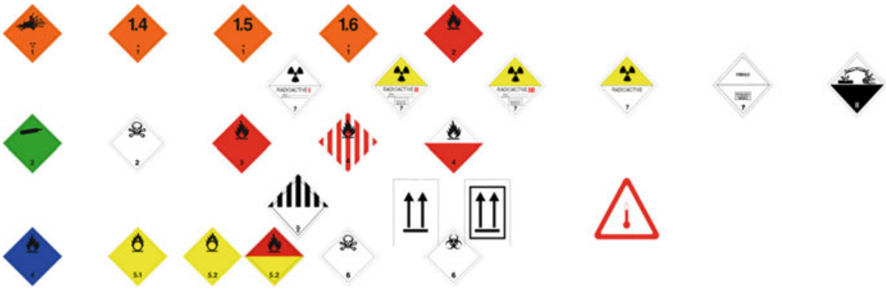


Fig. 2.2 Typical placarding

- White Papers and Green Papers describe a new strategy or approach and they are the EU instruments to start a regulation process.

In the context of dangerous goods, the main legislation act is the ADR (Accord européen relatif au transport international des marchandises Dangereuses par Route) or "European Agreement concerning the International Carriage of Dangerous Goods by Road", which governs the transnational transport of dangerous material.

This history of ADR goes back to 1957, under the aegis of the United Nations' Economic Commission for Europe and it started to be applied from 1968.

Recently, a new set of amendments were approved and entered in force from 1st January 2009 (see [7]) as ECE/TRANS/202, Vol. I and II or ADR 2009.

2.2.1 *Spotting – Placarding – Marking*

Currently dangerous goods-vehicles have to be placarded and marked for identification purposes.

The nature of the cargo loaded must be clear and visible. It means that a DG-vehicle must be identified by placards of 250×250 mm which correspond to the label(s) required for the dangerous good carried.

Placards are to be affixed to each side and at the rear of the vehicle, and to all four sides of containers (Fig. 2.2).

Marking means that a DG-vehicle/container must carry orange-coloured plates/self-adhesive sheets, with the hazard identification number and the four-digit UN number for the dangerous good loaded.

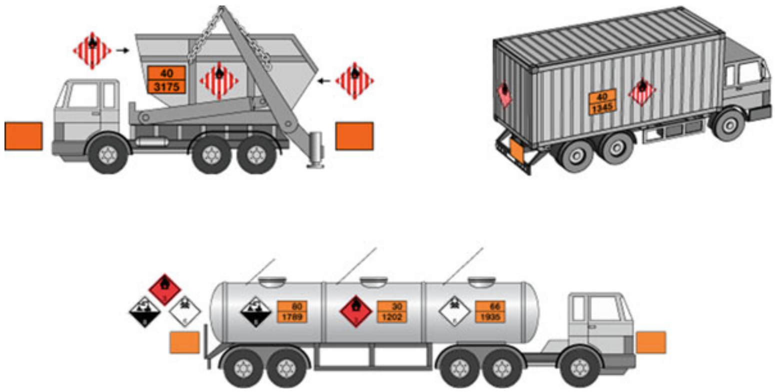


Fig. 2.3 Typical marking

The size is 300×400 mm, black border is 15 mm, numbers are 100 mm high.

Plates must resist a fire of 15 min, and stay affixed even if vehicle rolls over (Fig. 2.3).

2.2.2 *Limitations of the Current Legislation*

Even if largely adopted and efficient, the current legislation presents limitations:

- France recently reported the arrest of a "ghost" truck transporting old automotive batteries, full of acid and leaks, from UK to Spain… with no marking and dangerous goods documents at all…
- More dangerous situations have been reported, where a wrong marking induced wrong actions from rescue and fireman brigades, starting to pump a roll-over tanker and evacuate the injured driver, while the correct action would have been to stay away, because the air temperature conditions of the day and the explosive liquid transported where close to combine to blast in the air…
- With only packages or mixed loads of packages containing different dangerous goods, the vehicle just needs to carry two orange-coloured 'empty' plates at the front and at the back
- From the outside, the only thing that is known is that there is/are DG on board the vehicle, but not what it is (no hazard identification code), nor where onboard the vehicle it is loaded (rear? middle? front? side? top? bottom?)
- Small states, e.g. Luxemburg, are crossed by thousands of dangerous goods trucks a day and they simply do not know: very difficult for preparedness, anticipation, proportionality investment, intelligent monitoring…

2.2.3 Tracking and Tracing

Chapter 1.10 in Annex 2, is specifically focused on the security aspects for the transportation of dangerous goods and specifies the security requirements.

In the section of provision for high consequence dangerous goods the use of tracking devices is recommended to monitor the movement of high consequence dangerous goods.

The UN Recommendations on the Transport of Dangerous Goods [8] also recommends the use of tracking devices, when already installed, in clause 7.2.4.2. "When appropriate and already fitted, the use of transport telemetry or other tracking methods or devices shall be used to monitor the movement of high consequence dangerous goods".

Recently in occasion of the MENTORE workshop held in Brussels on June 2009 [9], the EC announced that one of the present priorities of the EU agenda is to draft a proposal for an EC directive for real-time tracking of dangerous materials.

In the wider context of the road transport sector, the European Commission identified in [10] the need for technical harmonization, standardization activity, creation of publicly available specifications and protocols for the use of telematics.

The conclusion is that tracking and tracing is an important regulatory requirement in the transportation of dangerous goods.

The next chapter highlights this topic relating projects and demonstrations of dangerous goods tracking and tracing operations.

2.3 Transportation of Dangerous Goods Based on GNSS

The categories of dangerous goods usually transported on the road can be quite numerous and they include substances and articles that have explosive, flammable, toxic, infectious or corrosive properties. A classification of the dangerous goods is presented in table A of the ADR (see Ref. [7]).

The transportation of goods by road presents the risk of traffic accidents. If the goods are present in the list defined in the ADR, the risk for the safety of the citizen may increase dramatically. Hazards like fire, explosion, chemical burn or environmental damage can be a direct consequence of a road accident involving a truck carrying a dangerous good. A road accident of this type can become a serious emergency crisis if it happens in a densely populated area like a city. Another aspect to consider is the possibility of hazards due to the presence of materials, which are not considered dangerous but they may become so in combination with a specific hazard. For example: two different trucks may carry chemical substances, which are not considered dangerous on their own but they can harm or poison other people on the road or the inhabited surroundings if combined as a consequence of a truck accident. Food substances can also become dangerous in the presence of a fire

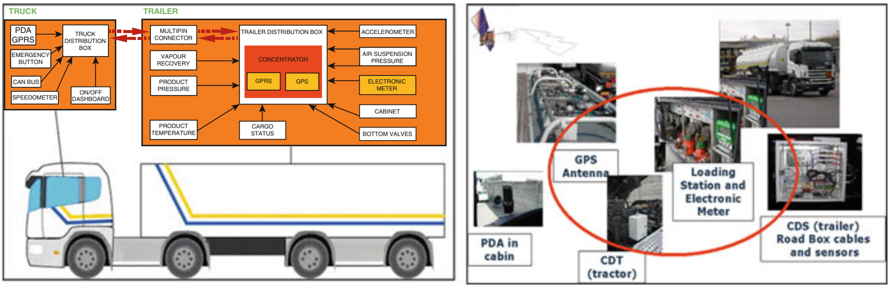


Fig. 2.4 Equipped truck with OBU and sensors

hazard as in the disaster of 24 March 1999 in the Mont Blanc Tunnel (see [11]), where a truck carrying flour and margarine caught fire in the tunnel.

The fire quickly propagated to other trucks carrying inflammable materials. The knowledge of the position of the transportation vector can be an essential element in preventing or resolving, in an efficient way, an emergency crisis of this kind.

Truck can be equipped with OBU (On Board Unit or Platforms), collecting the position information via GNSS techniques (now GPS) and transmitting the right data to a central server, through a wireless communication system. Additional information can also be collected by the OBU (status of the truck or conditions of the material) (Fig. 2.4).

This is the functionally commonly called "Vehicle Tracking" or "Tracking and Tracing" where the position of a vehicle can be monitored through a navigation device.

Tracking and Tracing can be passive or active:

Passive Tracking where the tracking device stores the vehicles location, through a positioning device (i.e. GNSS terminal), and other data (i.e. vehicle condition or container status) and stores this information in a data storage systems. At the end of the trip, the data can be collected and examined.

Active Tracking, where the tracking device stores and the vehicle location, through a positioning device (i.e. GNSS terminal), and send it through a wireless communication system to a control room for real-time update and monitoring.

The passive tracking may store a large number of information on the condition of the truck or the goods during the trip, but it does not provide real-time monitoring to the control centre. Consequently, it may be of limited help to support the resolution of a road accident even if the collected data may help the identification of the possible causes of the accident.

Active tracking can be used both to prevent an accident and to improve the response and resolution of the emergency crisis in the aftermath of an accident. In case of road accident, fleet managers and public safety responders (e.g. firefighters) can intervene in a more efficient and faster way using the knowledge of the last position of the truck.

Active tracing based on GNSS may also prevent a road accident by cross-checking the position of the truck with the conditions of the road, weather conditions or traffic jams or even status of the truck or the driver, which could increase the risk of a road accident.

There are a number of examples, where the location of the truck together with other information can be helpful:

The tank-pressure of a truck carrying chemical or inflammable materials can be monitored in real-time to detect potential leakage, especially when the truck is crossing an high densely populated area.

The status of the engine and its overheating, which could cause a fire to the vehicles transporting inflammable material.

The status of the driver could be monitored or checked on the basis of the trip itself. For example is the driving time is too long.

Unsafe road or roadwork in proximity of the truck.

Alert on bad weather conditions based on the position of the truck.

The control and coordination centre can analyze the received data and notify the driver through a commercial telecommunication system (e.g. GSM) if a risk-threshold is overcome

In active tracking, the position of a truck is frequently transmitted to a control centre to inform the fleet managers of the position of the various trucks on the road. The position of the truck is determined through a GNSS receiver mounted on the vehicle. Until now, GNSS receivers were predominantly based on the United States GPS system, which may have a limited availability in urban areas and lacks integrity assurance. Future tracking and tracing systems could take advantage of the improved availability and integrity of navigation systems like the European EGNOS or Galileo. The position of the truck is transmitted to the control centre through various wireless communication systems like cellular networks (GSM/GPRS/UMTS), satellite networks or even WLAN networks in urban areas based on WiFi or WiMAX (Worldwide Interoperability for Microwave Access). The evolution of the Intelligent Transportation System (ITS) paradigm may provide additional types of communication systems and a closer integration between transportation systems and localization systems. The information to be transmitted to the centre is usually limited to few 100 bytes. If the information is transmitted every minute from the On Board Unit (OBU) to the control centre, a link with a data rate of 9.6 kbits provides enough capacity to maintain the flow of data. Modern commercial communication systems can easily provide data links with this capacity, but they may not be able to provide adequate coverage. The combination of telecommunication systems to collect the location data and to process it, is also called "telematics" from the combination of telecommunication and informatics. Commercial systems are usually designed and deployed on the basis of business considerations. Therefore their coverage is higher in populated areas and limited or absent in inhabited areas. The only communication system with full coverage over a large geographical area is the satellite communications technology, but it may cost more than GSM/UMTS.

The Fig. 2.5 describes the generic architecture for tracking and tracing.

Another challenge is to guarantee when the reliability of the received GPS signal. Currently GPS is not able to provide such information to the terminal, but this is

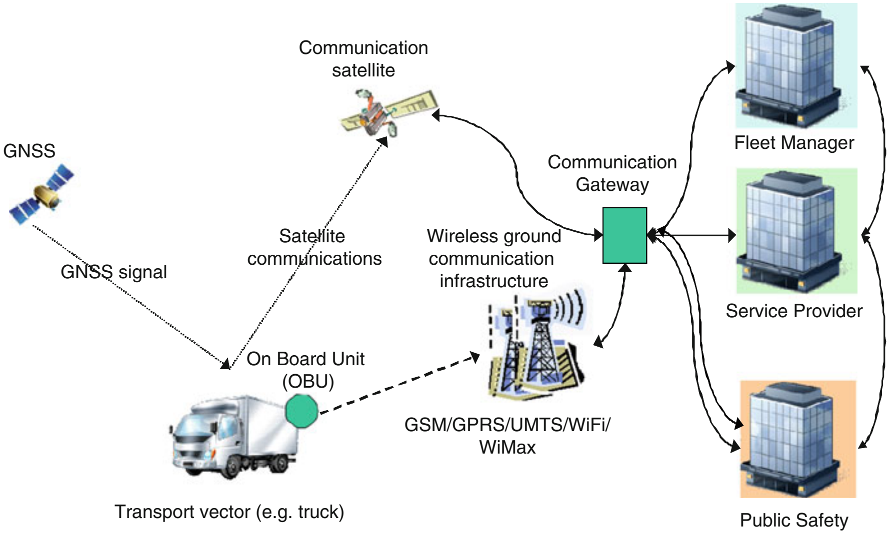


Fig. 2.5 Generic architecture for tracking and tracing of dangerous goods

essential to provide the adequate level of service to a fleet manager responsible to transport dangerous goods.

The implementation of reliable tracking and tracing for the transport of dangerous goods may generate additional costs and financial burdens for the fleet managers. On the other side, tracking and tracing may provide a number of benefits to the community including a reduction of the risks associated to the transport of dangerous goods. There is the need to define a complete business model, which does not impact negatively the fleet managers and which include the entire potential stakeholder in this application domain, including public safety organizations, government authorities, service providers and the owners of the cargo (i.e. the dangerous goods). In this business model, the regulatory framework is an essential component, but other elements are also important including the definition of standards for the transfer of data from the truck to the control centre to ensure interoperability, reliability and security of the transmitted data.

A number of tracking & tracing systems for dangerous goods have been implemented in Europe in recent years. One of the pioneer projects was SIMAGE [12] whose main objective was to implement a pilot information system for the near real-time monitoring of vehicles carrying hazardous freight and provide early alert warnings to the competent authorities in the case of emergency associated with the dangerous goods transport.

In [13], the authors describe a number of projects for land transportation (road and railway) of goods, which uses GPS for tracking and tracing. The paper presents the use of GPS in conjunction with geographic information system technology (GIS). Safety aspects are considered but not in relation to the transport of hazardous materials.

Another project MENTORE [5] targeted the use of the EU GNSS, EGNOS (European Geostationary Navigation Overlay Service) and Galileo, for "Regulated tracking and tracing" services.

The added value of EGNOS in comparison to GPS relies in the capability to provide "guaranteed positioning", thanks the exploitation of its integrity function. Thus, the present availability of EGNOS over Europe, in view of Galileo over the world, enables tracking & tracing services based on guaranteed positioning. These services can support the implementation of regulations and policies requiring precise and reliable localisation: the so-called "Regulated tracking & tracing" services.

Regulated Tracking & Tracing services address a wide range of applications in the transport and personal mobility markets. The most relevant benefits of these services are public and social; however they also imply commercial interests.

For the transport of dangerous goods, MENTORE successful identified the advantages resulting from the use of both EGNOS OS and EGNOS CS (based on the EGNOS data disseminated by EDAS, as mentioned before).

Another example of this type of projects is PROMIT financed as part of the Framework Program 6, whose objective is to "contribute to a faster improvement and implementation of intermodal transport technologies and procedures and to help promoting intermodal logistics and mode shift by creating awareness on innovations, good practices and intermodal transport opportunities for potential users as well as for politicians and for the research community" (from [14]).

In [15], the authors describe the results of the MITRA project, which is focused on the monitoring and intervention for the transportation of dangerous goods. The objective of the MITRA project was to specify and prototype an innovative operational platform for European civil security centres aimed at monitoring dangerous goods transportation in Europe and supporting intervention in the event of an accident. The MITRA project conducted a number of field tests across Europe and the paper provides a description of the field test and results. The paper does not address the vulnerabilities of the systems and possible countermeasures or good practices, which is the focus of this paper.

Another example is described in [16], where is described a remote identification system for dangerous goods safe transportation in Lithuania.

An interesting "case study" in Australia is presented in [17] where exceptional transports need a specific authorization to travel especially in urban areas. The operating model is mainly based on four actors: the transport operators, the government authorities, service providers, and auditors.

Restricted access is usually managed either by permanent labelling forbidding e.g. excess of weight, height, width, or by provisional notice due to exceptional circumstances (road works, accident, floods, event, etc.).

The main idea is to implement a flexible and intelligent system for the provision of permits, combining the monitoring of the compliance with the use of GNSS technologies. The process starts when an operator asks for an Intelligent Access permit. In the first phase, the contacted authority verifies which type of vehicle and goods are considered, where is the destination, and the supposed time of transit or arrival.

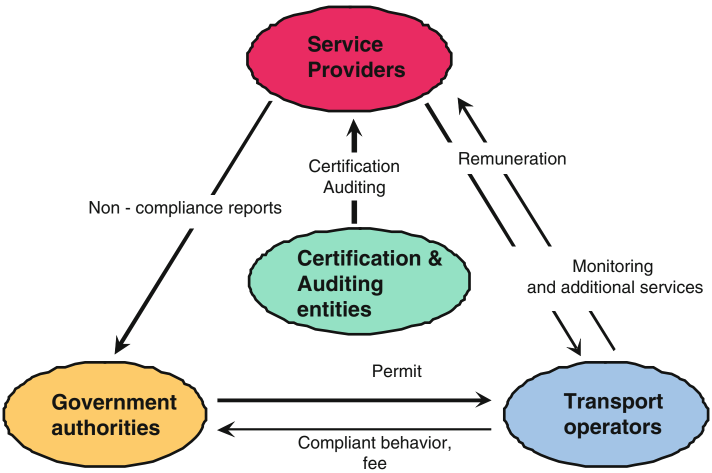


Fig. 2.6 Relationships among stakeholders in the Australian case study

After verification of the feasibility of the transport, the authority releases the requested permits under specific conditions, which include: requirements for the type of vehicle, the allowed travel routes, the time table and the authorized maximum speed. The permit includes also the monitoring or tracking/tracing by a service provider. The monitoring is performed using on-board platforms (i.e. OBU) and GNSS technologies, with the help of traffic telematics, by a service provider to be selected amongst the accredited service companies.

The role and functions of the service providers, paid by the transport operator, are to:

- Install a certified on board platform for telematics services
- Perform a constant monitoring of the vehicle concerning compliance with the transport requirement
- Report to the authority, in particular every non-compliant behaviour

The certification of devices and auditing of service providers is administered by a steering group under the authority of the government.

The relationships among the various stakeholders are shown in Fig. 2.6.

The Australian case study presents a well defined business model, which includes all the relevant stakeholders. Since, Australia is a single nation, it does not present the European challenges of cross-border tracking and different national agencies.

The described projects have identified a number of issues to be resolved for tracking & tracing in the transport of dangerous goods including:

- Monitoring the transport of dangerous substances is not only a technological challenge, but also and, perhaps more importantly, a participatory challenge, where various stakeholders from government, industry and research should work together for a common solution.

- Different competing technologies and architectures are currently available. Standardization is needed to converge on a single and efficient solution. A primary objective is the standardization of messaging structures for data exchange between the transportation vectors (e.g. trucks) and the control centres.
- There is a need for "sustainable business model", featuring: collaboration between all interested public and privates stakeholders, adoption of existing standards, service-oriented approach, interoperability and possibly integration with fleet management, new business opportunities and spin-off business services (e.g. customer services, anti-fraud, environmental management).
- Security and privacy of the tracking data is an important asset. The data storage and security is normally managed by the service provider, and only the authorities receive data concerning non-compliant behaviours.
- Because of the risk and dangers associated to the transportation of dangerous goods, it is important to guarantee well defined service level agreements. A complete tracking and tracking application of dangerous goods must include specification and processes for certification of the equipment, the transport vectors and the service providers.
- The existing GNSS like GPS or GLONASS may not provide the needed levels of guaranteed positioning to support the service level agreements. The navigation signal should be secure, available, stable and accurate.
- Finally, the system which implements the transportation of dangerous goods may have vulnerabilities. Beyond the conventional physical vulnerabilities, which are present in all the transport applications, "tracking and tracing" may have technical vulnerabilities in relation to the navigation and communications services used to determine and transmit the position of the truck.

Regarding specifically the tracking and tracing technologies, representing the future in term of dangerous goods monitoring and governance at EU level, one can conclude:

- Transportation of dangerous goods on the road is a key business sector that also implies citizens' safety aspects.
- Transportation of dangerous goods is increasingly based on tracking and tracing functionality to identify the position and status of the vehicle at any time. Tracking and tracing is based on telematics and navigation services, which may have vulnerabilities against intentional or unintentional attacks. It is important to define appropriate countermeasures, which can be based on innovative technologies.
- Future navigation systems like EGNOS and Galileo can play an important role as a tool for enhancing transport efficiency and ensuring the citizen safety… A major issue for tracking and tracing for the transport of dangerous goods is the presence of various proprietary solutions in Europe. Such fragmentation decreases the overall efficiency and creates interoperability barriers at European level, which are especially significant for cross-border transport routes. Thus, future developments of tracking and tracing need to address technical standardization to provide a secure, efficient and interoperable information flows from the vehicle to the control centres and among the control centres.

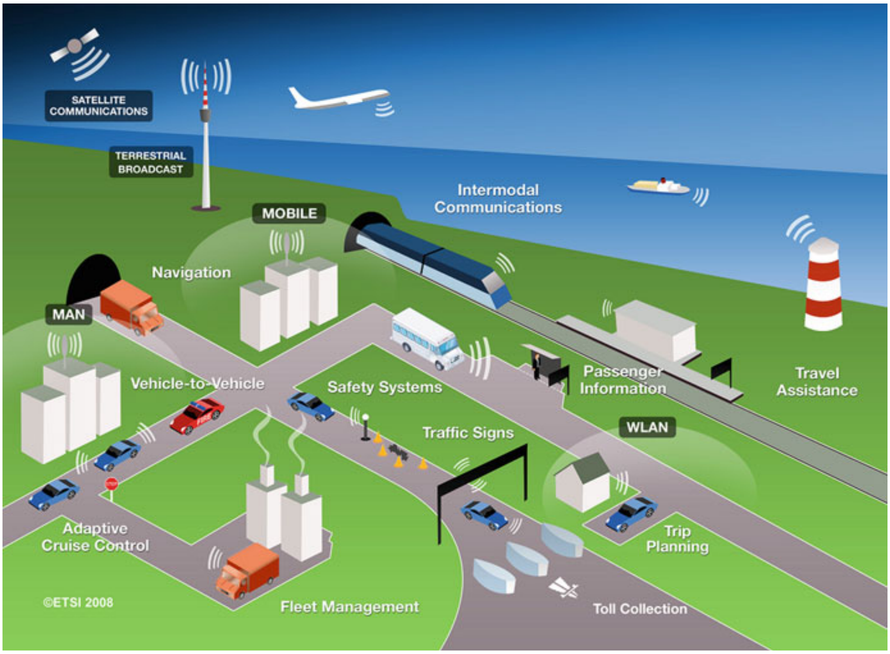


Fig. 2.7 ITS generalize the communication between various modes

2.4 The Future in an Increasing 'ITS' Environment

ITS (Intelligent Transport System) intends to apply ICT (Information and Communications Technology) to transport infrastructure (see Fig. 2.7) and vehicles in an effort to enhance e.g.:

- Mobility efficiency (fluidity, fuel consumption, car-sharing, E-toll, Road charging, Smart parkings)
- Security (rapid mass evacuation after a disaster, road and railways surveillance, transport of dangerous goods tracking, air and maritime security)
- Safety (E-call, V2V communications, Automatic speed limiter, live traffic messaging, collision avoidance, pedestrians detection, night vision, electric vehicles charging assistance)
- also more Commercial applications, from entertainment to user aids (navigation and local maps, anti-theft systems and stolen vehicles, remote doors unlocks, on-board internet, freight and fleet management systems FMS…)

2.4.1 On Board Platform Concepts

It is quite surprising that parallels ITS initiatives in commercial and/or private vehicles are developing with very little synergy even when their needs are similar

(e.g. transport of dangerous goods and live animals, Digital Tachograph, E-toll and E-Call)

Indeed, these ITS components or systems are generally confronted to similar barriers regarding their <u>implementation</u> or <u>enforcement</u> in vehicles, such as e.g.:

- Legal aspects: security/privacy, data protection and authentication;
- Migration: progressive retrofitting of existing fleet and infrastructure with new technologies;
- Harmonisation needs: if left to subsidiarity, large variability between Member States' implementation of such applications
- Lack of standards and technical specifications: interoperability, security, protection against fraud/abuse…
- International aspects regarding non-EU countries and agreements/conventions

There is most probably a space for the development of an integrated on-board platform, relevant for ITS policy applications, and addressing these issues.

But the introduction of an open-platform, mixing enforced applications with more commercial ones is still stuck with various resistances and barriers to overcome.

Technology is ready. Some automotive suppliers have all the possible applications ready (eCall, smart-parking, eToll, tracking and tracing, fleet management…) and no one is offering an integrated platform.

The major resistance is coming from vehicle manufacturers. It is already very difficult for them to manage the security of their electronic systems, sometimes hacked simply using the infrared doors locker knob as entry point. So, having a platform multi-owner, connected through Wi-Fi, GPRS, GPS or short range communication systems is currently too risky.

The second barrier regards apparently the users, fearing that an open platform, mixing enforced and non enforced applications would ease the legislator work for the introduction of new rules and enforcements.

2.5 Conclusions Regarding Tracking & Tracing in an Increasing Intelligent Transport System (ITS) Environment

At the end of the day, the technical solutions for tracking and tracing are available and reliable. While all the various demonstration projects had positive outcomes, no large scale applications ever developed.

The real challenge is elsewhere:

- Governance: who manage what, who own what, who is responsible for?
- Participation: collaboration and dialog between public and private actors
- Training and awareness: users and enforcers
- Sustainable business and operating models: service providers
- Frauds mitigation: safe and secure

- Harmonization-Interoperability-Continuity of services: fragmented standardization efforts
- Regulatory Framework: legal aspects, data privacy, competition, long term visibility

All this call for an integrated approach at EU level including advanced **governance strategies**.

This is quite a challenge that the next paragraphs will assess.

2.6 Need for an Integrated Risk Governance Approach: Definition of the Governance

In the domain of risk management as well as other relevant issues related to the sustainable development of modern societies, it is becoming increasingly clear that the integration of economic, social, cultural, political, and ecological factors is a stringent need. It requires a constructive articulation of the top-down approaches to development with the bottom-up initiatives. It requires the simultaneous consideration of the local and the global dimensions and of the way they interact. And it requires broadening the space and time horizons to accommodate the need for intergenerational as well as intergenerational equity. In other words, what is needed is nothing less than a fundamental shift in the way we approach development and the relations between society and nature. Therefore, the quest for sustainable development poses new, deep challenges to the ways we define problems, identify solutions, and implement actions.

Economic growth, rapid technological change, and the expansion of scientific knowledge have made societies more and more confident in their abilities to "manage" hazards and to minimize risks. A paradigm based on planning for efficiency, standardizing for easier social control, and reducing variability has come to pervade bureaucratic practices. Safety and security problems are framed as technical and administrative challenges of politics. People need to be informed about the risk appraisal and, and penalized if they do not follow the right management practices and rules.

Considering the technological evolution and the availability of information occurred during the last decades, it could argue that with good information and technical skills, the future risk management can be certainly improved. Nevertheless, it still requires the definition of common rules among societies, public actors and stakeholders to drive the exploitation of such resources.

Strengthening the capacity of societies to manage risk and to increase the resilience is critical to effectively pursuing sustainable development. This pursuit is a dynamic challenge in which many different conflicting interest or synergic actions can lead to different approaches and solutions.

However, who decides when to intervene and identifies the desirable system configurations? Who decides what portfolio of challenges the system should be made resilient to and which are of priority interest? How are those decisions made? Who controls implementation? What are the consequences of alternative courses of action for different stakeholder groups?

These are fundamentally questions about the politics of managing resilience and vulnerability. The central question we address in this paper is: How do certain attributes of governance function in society to enhance the capacity to manage the transportation of dangerous goods and to increase the resilience?

Governance can be defined as the structures and processes by which societies share power, shapes individual and collective actions [18]. Governance includes laws, regulations, discursive debates, negotiation, mediation, conflict resolution, public consultations, protests, and other decision-making processes. Governance is not the sole purview of the state through government, but rather emerges from the interactions of many actors, including the private sector and not-for-profit organizations.

2.7 Risk Governance of Transportation of Dangerous Goods

As it has been illustrated on the previous chapter, governance refers to the process whereby elements in society wield power and authority, and influence and enact policies and decisions concerning public life, and economic and social development. Governance involves interaction between these formal institutions and those of civil society and it has no automatic normative connotation.

Now the question is in terms the concept of "risk governance" can be interpret and described.

Risk is becoming a central concept in environmental policy and practise, but this does not mean that it is an easy or well-defined concept. Risk is a very complex and multi-faceted concept. Furthermore, risk means different things to different people. Despite the widespread use of the word, no single definition of risk can claim to be universal. In the literature the word "risk" is used in many different senses. A multitude of different kinds of risks are discussed (e.g. business risk, social risk, economic risk, safety risk, environmental risk, etc.) and the ongoing debate is mostly about the measurability or predictability of risk, between scientific and technological risk assessors on the one hand and social scientists and psychologists on the other [19]. The committee of the Society for Risk Analysis laboured for 4 years trying to define the appropriate meaning of the word "risk". After that period it gave up, saying in its final report that maybe it's better not to define risk. "Let each author define it his own way, only please each should explain clearly what way that is" [20].

In order to capture the intuitive idea of risk, it can be noted that it has be related to the concept of hazard, namely a source of danger. However, risk is something different than simply danger because it also related to the assessment of the consequences posed by the hazard. Thus, the notion of risk involves both some kind of loss or damage and the uncertainty of the consequences.

Kaplan and Garrick argued that when one asks (1981): What is the risk? One is really asking three questions:

- What can happen?
- How is the likelihood of it happening?
- If it does happen, what are the consequences?

In other terms risk can be defined as the probability of an undesired event.

The risk associated to the transportation of dangerous good can be described considering the related supply chain of dangerous materials. A supply chain may be defined as an integrated process wherein a number of various business entities work together in an effort to:

- Acquire raw materials;
- Convert these raw materials into specified final products;
- Deliver these final products to retailers.

This chain is traditionally characterised by a forward flow of material and backward flow of information.

Therefore, "risk governance" involves the translation of the substance and the core principles of governance to the context of risk and risk-related decision-making [21].

It could be argued that theoretically probability of occurrence of a hazard and the related consequences can be measured. In reality the problem associated to the definition of the risk is not its measurement but how it is socially defined and perceived. As has been emphasised by Slovic, that "danger is real, but risk is socially constructed" [22], and in this sense perception of risk and its communication play a crucial role. Therefore, risk governance includes, but also extends beyond, aspects like risk assessment risk management, and risk communication. It requires consideration of the legal, institutional, social and economic context in which a risk is evaluated, and involvement of the actors and stakeholders who represent them. Risk governance looks at the complex network of actors, rules, processes and mechanisms concerned with how relevant risk information is collected, analysed and communicated and how management decisions are taken. Encompassing the combined risk-relevant decisions and actions of both governmental and private actors, risk governance is of particular importance in situations where there is no single authority to take a binding risk management decision, but where instead, the nature of risk requires the collaboration of, and the coordination between a range of different stakeholders [23].

Considering the flows of hazardous materials and the related flow of information that every day travel around the world, and in particular the tide relationship between economic system of production and international markets, over the past decades, two separate and apparently opposite trends in government and risk governance have been implemented [24].

National Scale: governance describes structures and process for collective decision-making process involving governmental and non-governmental actors;

Global Scale: governance embodies a horizontally organised structure of functional self-regulation encompassing state and non-state actors bringing about collectively binding decisions without superior authority;

In reality the European Union can be identified a third approach that is, to a certain extent, the combination of the two approaches. The Treaty of Maastricht, signed on 7 February 1992 introduced the principle of subsidiary in the EU law. The present formulation, contained in Article 5(3) of the Treaty on European Union (consolidated version following the Treaty of Lisbon, which entered into force on 1 December 2009) consider that in areas which do not fall within its exclusive competence, the Union shall act only if and in so far as the objectives of the proposed action cannot be sufficiently achieved by the Member States, either at central level or at regional and local level, but can rather, by reason of the scale or effects of the proposed action, be better achieved at Union level.

The application and the related implication of the subsidiarity principle in the fields of transportation of dangerous goods is demonstrated by the Council Directive 94/55/EC of 21 November 1994 (Directive 94/55/EC is repealed by Directive 2008/68/EC from 30 June 2009). The directive foreseen that Member States may:

- Retain existing national legislation for national transport operations performed by vehicles registered within their own territory;
- Regulate or prohibit the transport of certain dangerous goods for reasons such as national security or environmental protection;
- Apply more stringent provisions regarding transport carried out by vehicles registered or put into circulation within their territory, but not regarding construction requirements;
- Adopt more restrictive measures in emergencies, subject to Commission approval;
- Retain national provisions for certain classes of substances;
- Authorise transport by road of certain dangerous goods for transport operations that include carriage by sea or air;
- Retain, after consulting the Commission, provisions less stringent than those set out in Annexes A and B of the Directive for the transport of small quantities of certain dangerous goods;
- Grant temporary derogations for the purpose of carrying out, within their territories, any trials necessary for amending these Annexes in order to adapt them to technological and industrial developments.

Vehicles registered or put into circulation in third countries are authorised to carry out international transport operations of dangerous goods within the Community if the transport complies with the European Agreement concerning the International Carriage of Dangerous Goods by Road (ADR), and provided that there are no national or Community restrictions on market access.

In other terms the principle of subsidiary applies to those areas where the Member States do not have exclusive competence, i.e. delineating those areas where the Member States should and should not act. Consequently, in Europe an effective risk governance approach related to the transportation of dangerous goods will built and developed considering the synergies and influences of the global governance level (i.e. European and international) and the National one.

2.7.1 Public and Private Actors and Stakeholders

A part from the European Union, there are other international bodies and Authorities that work on the harmonisation of the transportation of dangerous goods. At the United Nations level, all work related to the transport of dangerous goods is coordinated by the Economic and Social Council Committee of Experts on the Transport of Dangerous Goods and on the Globally Harmonized System of Classification and Labelling [25], which produces the "Recommendations on the Transport of Dangerous Goods" [26], also called the "Orange Book" and the Globally Harmonized System of Classification and Labelling of Chemicals.

In the special case of radioactive material, work is coordinated by the International Atomic Energy Agency (IAEA), which produces the Regulations for the Safe Transport of Radioactive Material. The Recommendations on the Transport of Dangerous Goods and the IAEA Regulations for the Safe Transport of Radioactive Material are addressed not only to all Governments for the development of their national requirements for the domestic transport of dangerous goods, but also to international organizations such as:

- The International Maritime Organization (IMO),
- The International Civil Aviation Organization (ICAO) and
- Regional commissions such as the Economic Commission for Europe – for regulations and international/regional agreements or conventions governing the international transport of dangerous goods by sea, air, road, rail and inland waterways.

If these international authorities define at international level the framework and the constraints for the national governments, it could argue that the implementation power of such provisions remains in the hands of national and regional government, at least within western political systems. Because the approaches suggested by international bodies can not be exhaustive and there areas of concerns that can be tackled only by national authorities, the governance of the dangerous good it seems to have two conflicting objectives. From one side it tries to accomplish with international rules and standards, on the other side it reinforce regional policies in order to accomplish with local stakeholder needs.

For instance the review of the best practice about the applied methodologies for risk analysis of dangerous goods or the different policy approaches to support the land use planning process in order to mitigate of risk of dangerous goods along the main routes of transportation, show clearly that harmonisation of approaches and criteria still hardly exist in Europe and in other countries. This is manly due to the different interests at regional level or to compromises with other peaces of legislation or the consequence of historical development of national technical disciplines and regulations.

This is only an apparent conflict. One of the factors, which these two apparently divergent developments have in common, is that they illustrate the erosion of the assumption that the nation state is the primary and most appropriate locus of authority for political decision-making. The increasing interdependence and interconnectedness of nations and markets of dangerous goods, both as a result of voluntarily undertaken international obligations and the establishment of international institutions of varying degrees of power on the one hand, and as a result of the transnational activities of economic actors and the unplanned trans-boundary effects of many social, political and market activities on the other, pose a powerful challenge to the claim that national governments are best placed to decide upon and to regulate most matters of political importance. Secondly, it can be recognised a more general fragmentation of political power, and the dissolution of boundaries between private and public spheres of regulation and control. It concerns the evolution, in other words, of a concept of governance, which transcends the more traditionally conceived private/public divide and which challenges previous assumptions about the locus of political and economic authority. In other terms stakeholders and public actors are more interested to collaborate in order to develop regional implementations that pay more attention to local interests and to take advantage of the international harmonisation of the rules.

The grey area where regional interests meet the needs for the improvement of international standardization is exactly the space where new challenges for an enhanced risk governance of dangerous good can be defined. This area needs to be exploited with contribution of all stakeholder concerned at different level in order to develop a risk governance approach that go beyond the private and public separation of interests and objectives.

2.7.2 Challenges for the a Common Governance Definition

Risk management of transport of hazardous substances is generally known to suffer from the lack of availability of pertinent data or of information integration [27], as a result of, or resulting in, a generalized analysis approach. There is the added issue that the hazardous material transportation, viewed as a system, is highly heterogeneous and complex [28]. These issues and others like damage minimization and information availability [29] become critical when dealing with an emergency situation. A good risk governance approach in the domain of transportation of dangerous goods is an ideal process, which is difficult to achieve in its totality. In this chapter we will like to present some of available drive forces that may help in the future to improve the risk governance of dangerous goods. For this purpose we can assume that governance in general has, among others, three dimensions:

- The political dimension – processes by which advanced and effective policy are defined, monitored and replaced.
- The economic dimension – process by which public resources are effectively managed, sound policies implemented, and private business activities can have advantages of to contribute to such development.
- The institutional dimension – processes by which citizens and the state itself define the own role for the implementation of policies.

In addition to these dimensions, a crucial role is played by the availability of technology that can support to governance processes.

2.7.3 EU Scale Approach Supported by a Political Moment

Among the main and fundamental objectives of the European Union there is the aim at responding to globalisation by making the European economy more competitive by supporting the reform the programmes of member countries on growth and competitiveness. The goal is to enforce the social cohesion and sustainable development of the Member States and their citizens.

By the year 2000, EU leaders were well aware that the European economy needed thorough modernisation in order to compete with the United States and other major world players. Meeting in Lisbon in March that year, the European Council set the EU a new and ambitious goal: to become, by 2010, 'the most competitive and dynamic knowledge-based economy in the world, capable of sustainable growth with more and better jobs and greater social cohesion'.

This could consider as the ground where to build and to improve the common risk governance of dangerous goods. Maybe dangerous good management can be seen as not priority issue on the political agenda of the EU but there are some relevant ongoing policies under definition and implementation that will have positive effects in the near future. This chapter will illustrate the momentum generated by few of them that will contribute to improve the risk governance.

One of the relevant aspects covered by the 'Lisbon strategy' is the accelerated liberalisation of the telecommunications. The ongoing European transportation policy (ITS – Intelligent Transport Systems) is stimulating communication technology innovation and it will apply more and more information and communication technologies to transport. Computers, electronics, satellites and sensors are playing an increasingly important role in our transport systems. The main innovation is the integration of existing technologies to create new services. ITS as such are instruments that can be used for different purposes under different conditions. ITS can be applied in every transport mode (road, rail, air, water) and services can be used by both passenger and freight transport.

The ITS innovation will provide more information which can contribute to regulate and control the traffic. It will make transport more sustainable, which means efficient, clean and safe, but also it will improve to capability of monitoring freights in real time. The available information will support emergency planning activities and the rescue services.

Other element that will contribute in this direction is the digital tachograph. The Regulation (EEC) 3821/85, as amended in 2006, requires that digital tachographs are fitted into goods vehicles and buses that come into scope of the Drivers' Hours rules. The digital tachograph is obligatory equipment on new trucks and coaches from May 2006 in the European Union and from 2010 (after a 4-year transition) for other ECE member countries, which are Contracting Parties to the UNECE-administered

European Agreement concerning the Work of Crews of Vehicles Engaged in International Road Transport (AETR).

To strength the safety issues related to the transportation of dangerous goods, the new programme of European Commission (2011–2020) could be considered as well. This program defines Seven Strategic Objectives:

Improved Safety Measures for Vehicles: Between 2011 and 2020 a range of new "active safety" measures will come into force for safety equipment including, mandatory electronic stability control (for cars, buses and trucks to reduce the risk of destabilising or rolling), Mandatory lane departure warning systems (for trucks and buses), Mandatory automatic emergency braking systems (for trucks and buses).

Building safer road infrastructure: European funds will only be granted to infrastructure compliant with road safety and tunnel safety Directives. This is already the case for TEN-T funding, the Commission wants to extend it as a general principle for any EU funding, for example cohesion funding.

Boost Smart Technology: as it has illustrate above, the Commission will propose new technical specifications, under the ITS Directive (Intelligent Transport Systems Directive) so that data and information can be easily exchanged between vehicles and between vehicles and infrastructure (for example to enable real time information on speed limits, on traffic flows, congestion, pedestrian recognition).

Strengthening education and training for road users: The road user is the first link in the safety chain and the one most prone to error. Whatever the technical measures in place, the effectiveness of road safety policy depends ultimately on users' behaviour. Education, training, and enforcement are essential.

Better enforcement: Effectiveness of road safety policies largely depends on the intensity of controls and compliance with safety requirements. Enforcement remains a key factor in creating the conditions for a considerable reduction in the number of deaths and injuries. Speeding, drink driving and failure to wear a seatbelt are still considered the three main killers on the road.

Establishing a Road Injuries Target: reducing the number of injuries will be a key priority action for Europe for the next decade. The Commission will develop the elements of a comprehensive strategy of action concerning road injuries and first aid.

A new focus on Motorcyclists: The Commission will focus particularly on motorcycles and other "powered two wheelers" (PTW). While other vehicle transport modes have shown significant decreases in fatalities and serious injuries over time, those for PTW riders have exhibited much lower decreases or remained even static.

Finally it should mentioned the European Earth observation programme GMES (Global Monitoring for Environment and Security). GMES consists in a complex set of systems which collects data from multiple sources (earth observation satellites and in situ sensors such as ground stations, airborne and sea-borne sensors), processes these data and provides users with reliable and up-to-date information through the advanced services. The main goals of this program is providing policy-makers and public authorities, the major users of GMES, with sound information to prepare environmental legislation and policies with a particular focus on Climate Change, monitor their implementation and assess their effects. GMES also supports the critical decisions that need to be made quickly during emergencies, such as when natural or man-made catastrophes and humanitarian crises occur.

Users will be provided with information through services dedicated to a systematic monitoring and forecasting of the state of the Earth's subsystems. Some of these systems and data sources already exist today, as well as prototype services but many developments are still required in all domains. Therefore the risk management of transportation of dangerous good will profit of GMES program.

All these initiative are not strictly dedicated to regulate the transportation of dangerous goods but they will have a positive effect. The question switch then how to take advantage of such programmes and to define an agenda for applied research in order to exploit the available technologies and how to define a business model capable to profit of advanced technological services.

2.8 Sustainable Business Model and Decision – Making Process

Governance of the transport of hazardous materials requires complex decision-making regarding regulation, routing, land-use, and resources for emergency response. Decision makers need to balance the needs of various actors, such as industry and the public, they must take a multi-hazard approach, they must take into account the spatial dimension of the problem, and they must know the vulnerability of their territory. Therefore, they must have appropriate decision support tools that facilitate the required integration of information.

Decision-support systems have demonstrated to be valuable tool in the field of management of dangerous good and prevention of crisis. They rely more and more on complex information technology architectures and powerful communication technologies. Such technologies are mature but there is operational constraint related to the cost the communication. Therefore, the challenge for the coming years is not related to improvement of the available technology but more on the definition of a feasible and sustainable business model able to reduce the operative costs of public controllers and decision-makers and to stimulate the competitiveness among the private operators. Private sectors have already put a lot of effort in order to improve the safety and security of the transportation of dangerous goods but improvements are still required. Beside this, public authorities need to improve the control and the monitoring activities, and this requires large investment of public money.

Therefore, the following questions remain open and need to be systematically addressed in the coming years:

- How much of safety/security is still required?
- How much does it cost to implement more efficient monitoring and management systems? Who is going to pay?
- What are the benefits? Who can benefit of it?

- How to harmonise the administrative and technical procedures?
- What kind of adoption? (on voluntary basis; compulsory; governmental subsidies)
- Who will manage the collected information?

Considering the nature of such questions, it could argue that the answers need to be searched more on the political agenda. Decision makers need to balance the needs of various actors, such as industry and the public, they must take a multi-hazard approach, they must take into account the spatial dimension of the problem, and they must know the vulnerability of their territory. The available technology is mature for supporting them.

The coming challenge is relate how to define a business model that can exploit the available technologies and resources in order to develop a proactive business model based on an ethical ground. From one site the model will need to minimize and to equally share the risk related to the dangerous goods supply chains. From the other one it has to contribute to develop a more environmentally sustainable society (e.g. low carbon society). Setting new standards for supply and production activities based on such principles may open new market and supply opportunities, which will reflect more closely the social and environmental sensitivity of future societies. The driving forces will be the exploitation of the available technology and the increasing demand of the citizens for a more sustainable society.

Addressing systemic risks and vulnerabilities in Europe requires a cross-sectorial approach for identification, prevention and mitigation of systemic risks. This can be achieved by building multidisciplinary partnerships amongst authorities, experts and institutions on different policy implementations to coordinate the cooperation between different sectors at the appropriate level (local, regional, national and international). It also helps to overcome possible conflicts of interest between different stakeholders and service providers.

2.9 Conclusions

The governance of the transport of dangerous substances is not only a technological challenge, but also, and perhaps more importantly, a participatory challenge. Technical standard and rules are already available. Anyway, the future development of ICT technology will contribute enormously to improve the safety and management associated with the transportation of dangerous goods because it will improve the availability of information for authorities but also for all stakeholders. Advanced information and communication technologies can now be incorporated into onboard "Intelligent vehicle systems", offering new solutions to today's transport problems.

These high-tech systems have great potential to:

- Help drivers prevent or avoid traffic accidents;
- Mitigate the consequences of accidents that do occur;
- Provide drivers with real time information about traffic on road networks, thereby avoiding congestion;

- Find the most efficient routes for any journey;
- Improve real time monitoring and emergency planning;
- Provide more efficient rescue service.

The adoption on board of electronic telecommunication units will accelerate the deployment of intelligent vehicle systems on European and international markets, using a mix of policy, research and communications instruments to:

- Ensure interoperability across different EU countries and harmonize technical solutions through a comprehensive European approach;
- Support ICT-based research and development in the area of transport and facilitate the take-up and use of research results;
- Raise awareness among consumers and decision-makers of the potential benefits of ICT-based solutions.

It could be argued that goal of refinancing a more efficient governance of the transport of dangerous goods will be achieve only if along the technological development, a sustainable model will be defined, which will be able to take care of the internal market needs for the transport of dangerous goods, to set common rules among Member States and to consider to the ethical questions related to the environmental sustainability and the safety of citizens.

There are still difficulties among Member States in agreeing on the principles in the field, because each country has its own priorities, local communities, central authorities and different kinds of legislation. Variations arising in definitions, concepts, as well as methodologies and practices are the result of a wide range of different factors, including the different perceptions, attitudes and values regarding risks in different socio–economic-political contexts; and the different needs and specifications of diverse industrial sectors and risks specifications in various regions and countries [30].

It could be concluded that in order to develop an advanced risk governance model for the transportation of dangerous goods, it is of utmost importance to create effective domestic environments that ensure administrative feasibility and operational capability. Institutions are needed to link national and local activities with the international framework in many different sectors. The risk governance of transportation of dangerous goods has to reflect a strategic and integrated approach by incorporating transport effects from other sectors (e.g. organization of industrial production processes, design of cities) and all effort made in the field of the development of a low carbon society. A systemic approach should integrate and go beyond individual projects and support policy packages to achieve more safety of transportation. Policies should aim to achieve the wider sustainable development benefits of transport and remove financial and non-financial barriers. This will be a win-win strategy that will show extensive co-benefits regarding environmental protection (e.g. air quality and noise), equity between social groups (e.g. access and safety) and economic effects (e.g. fair competitiveness, job creation and wealth creation).

References

1. EUROSTAT Statistics in focus 66/2008 Road freight transport by type of goods – 2006 at http://ec.europa.eu/eurostat. Accessed 13 Aug 2010
2. Giannopoulos GA (2001) An intelligent tracking and tracing solution for intermodal transport. In: Proceedings ITS World conference, Sydney, Sept 2001
3. Giannopoulos GA (2004) The application of information and communication technologies in transport. Eur J Oper Res 152:302–320
4. World Nuclear Association (2010) http://www.world-nuclear.org. Accessed 13 Aug 2010
5. MENTORE Web site (2009) www.gnsstracking.eu. Accessed 10 July 2009
6. Globally Harmonized System of classification and labelling of chemicals (GHS) (2010) http://www.unece.org/trans/danger/publi/ghs/ghs_rev03/03files_e.html. Accessed 13 Aug 2010
7. ADR (2010) http://www.unece.org/trans/danger/publi/adr/adr2009/09ContentsE.html. Accessed 13 Aug 2010
8. UN recommendations on the transport of dangerous goods – model regulations, 14th rev. edn. http://www.unece.org/trans/danger/publi/unrec/rev14/14files_e.html. Accessed 13 Aug 2010
9. MENTORE final event ñMENTORE Manifesto (2010) EGNOS/Galileo services benefiting to regulated applicationsî, Brussels, 10 June 2009. Web site: www.gnsstracking.eu/en/about-mentore/mentore-final-event. Accessed 13 Aug 2010
10. Communication on a Community strategy and framework for the deployment of road transport telematics in Europeî (COM (97) 223 final of 20.05.97)
11. Minister of the Interior, and Ministry of Equipment, Transportation and Housing (1999) Task Force for Technical Investigation of the 24 March 1999 fire in the Mont Blanc vehicular tunnel – report of 30 June 1999 (English translation)
12. Nordvik JP (2005) ñTraceability and risk assessment for the transport of dangerous substances: results from an Italian case-studyî. 29th ESReDA Seminar, Ispra, Italy, October 25–26th, 2005
13. Mintsis G, Basbas S, Papaioannou P, Taxiltaris C, Tziavos IN (2004) Applications of GPS technology in the land transportation system. Eur J Oper Res 152:399–409
14. PROMIT Web site (2009) http://www.promit-project.net. Accessed 10 July 2009
15. Planas E, Pastor E, Presutto F, Tixier J (2008) Results of the MITRA project: monitoring and intervention for the transportation of dangerous goods. J Hazard Mater 152:516–526
16. Batarliene N (2008) Mobile decisions of transportation of dangerous goods. In: The 8th international conference ñRELIABILITY and STATISTICS in TRANSPORTATION and COMMUNICATION - 2008î, 15–18 October 2008, Riga, Latvia
17. Intelligent Access Program Australia-A GPS and map-based project for the monitoring of Authorized Transport. IAP Delivering the benefice. A report prepared by Rapp TRANS AG consultants (Basel) for the Transport Certification Australia, 2008–2009
18. Young OR (1992) The effectiveness of international institutions: hard cases and critical variables. In: Rosenau JN, Czempiel E-O (eds) Governance without government: order and change in world politics. Cambridge University Press, Cambridge, pp 160–194
19. Fischhoff B (1984) ñDefining Riskî. Policy Sci 17:123–139
20. Kaplan S, Garrick J (1981) On the quantitative definition of risk. Risk Anal 1(1):11–27
21. Gunningham N, Grabosky P, Sinclair D (1998) Smart regulation: designing environmental policy. Oxford University Press, Oxford
22. Slovic (2000) The perception of risk, The earthscan risk in society series, ISBN: 9781853835285
23. IRGC (2005) Risk governance – towards an integrative approach, white paper No 1, International Risk Governance Council, Geneva, Sept 2005
24. Renn O (2008) Risk governance – coping with uncertainty in a complex world, the earthscan risk in society series, ISBN: 9781844072927

25. United Nations (2009) European agreement concerning the international carriage of dangerous goods by road – ADR, United Nation, Economic Commission for Europe, Geneva, January 2009, ISBN: 13: 978 9211562255
26. United Nations (2009) Recommendations on the transport of dangerous goods. United Nation, Economic Commission for Europe, Geneva. ISBN 13: 978 9211391367
27. Saccomanno F, Living, F (2001) GIS-based integrated model for road accident analysis and prediction. Transportation research record 1768, Transportation Research Board, National Research Council, Washington, DC, pp 193–202
28. ICF (2000) Risk management framework for hazardous materials transportation. U.S Department of Transport Research and Special Programs Administration, Washington, DC, November 1 2000
29. Fabiano B, Curro F, Reverberi AP, Pastorino R (2005) ñDangerous good transportation by road: from risk analysis to emergency planningî. J Loss Prev Process Ind 18:403
30. DCDEP 2000, Risk Assessment in Europe. A summary report from the EU workshop on risk assessment, part 2, Directorate for Civil Defence and Emergency Planning, Oslo 25–26 November 1999, ISBN: 82-993462-8-2

Chapter 3
Technical and Functional Standards to Provide a High Quality Service for Dangerous Good Transport on Road

Mauro Benza, Chiara Bersani, Massimo D'Incà, Claudio Roncoli, and Roberto Sacile

Abstract In the last years, the European Commission has increased its effort to improve safety on road infrastructures by adopting new directives to guarantee users a high, uniform and continuous level of service. In this paper, the authors describe the current normative state of the art on Dangerous Good Transport (DGT) by road and introduce innovative initiatives at European level to prevent risk in this context. Then, the authors introduce the architecture of the DGT Information System developed in Italy exploring technological and methodological approaches applied to monitor and detect in real time DG vehicles. The DGT information system can be applied to implement additional decision support systems to manage the fleet routing, the customer orders, and other support tools for drivers including training, resource management and advanced data mining.

3.1 DG Transport by Road

The transport by road represents two thirds of the transport of goods in Europe (Fig. 3.1).

Dangerous Goods (DG) accounted for an estimated 4.1% of total tonne-kilometre performance of road goods transport in 2007 in the EU-27. Fuel made up the two largest categories of DG transported by road: Flammable liquids (58%) and Gases (12%). DG are substances that due to their physical and chemical properties or the nature of the reactions they provoke, represent a grave danger to man, property or the environment. Transported goods are classified as being dangerous according to the main

M. Benza (✉) • C. Bersani • M. D'Incà • C. Roncoli • R. Sacile
University of Genoa, DELAB – (DIST-ENI Joint Lab on Safety and Security in Logistics and Transport), Via Opera Pia 13, 16145 Genoa, Italy
e-mail: mauro.benza@dist.unige.it; chiara.bersani@unige.it; massimo.dinca@unige.it; claudio.roncoli@unige.it; roberto.sacile@unige.it

E. Garbolino et al. (eds.), *Transport of Dangerous Goods*, NATO Science for Peace and Security Series C: Environmental Security, DOI 10.1007/978-94-007-2684-0_3,

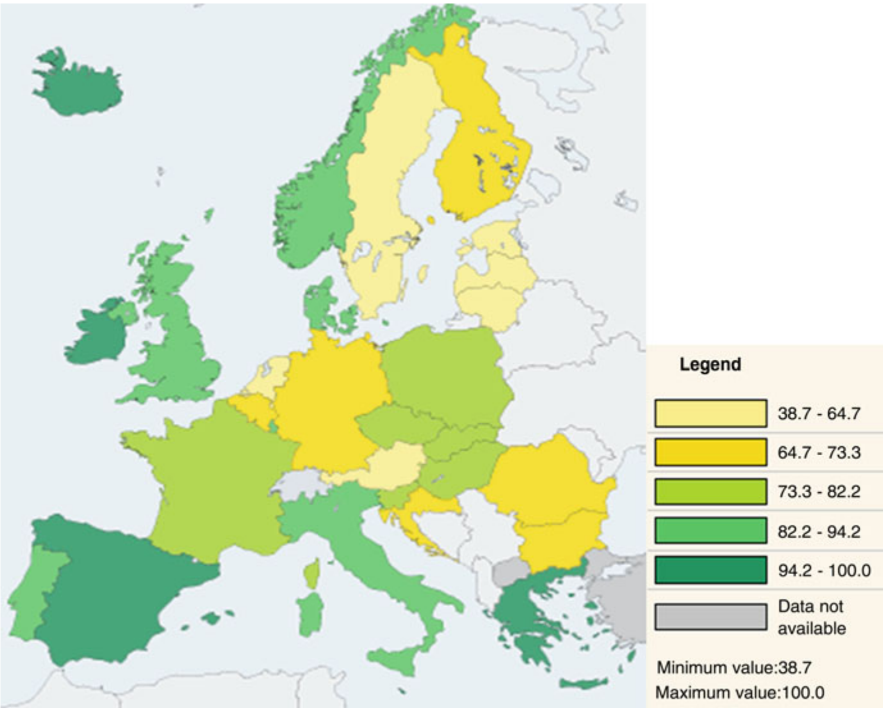


Fig. 3.1 Modal split of freight transport – % in total inland freight tonne-km transported by road in 2008 (*Source*: Eurostat)

categories of Directive 94/55/EC. These goods are classified by the ADR (European agreement regarding the international transport of DG by road, recently revised in 2005 and 2007) into nine classes: goods are included in a class according to their physical-chemical characteristics, the principle type of risk they represent. The transport of DG can be done in tankers, as unpacked goods, or packaged in cylinders or bags. There is also special packaging for radioactive and biological goods. The transport of packages, at times grouped with other types of goods, is carried out in accordance with regulation on quantity, type of package etc. (Fig. 3.2).

In 2007, the five main contributors to EU-27 total were the major economies Germany, France, Spain, the United Kingdom and Italy (2006). In terms of tonnes, they together made up two thirds (66%) of the weight of goods loaded nationally in the EU-27. From 2000 to 2007, in the 13 EU-15 Member States for which data are available, the weight in tonnes of goods transported by road grew at an average yearly rate of 3.0%.

The risk attached to the transport of DG by road is a risk that is complex to understand as it is connected to all the road network and depends on multiple factors such as traffic density, weather conditions, the necessities of undesired events (road accidents, natural phenomenon etc.). This risk is also strongly linked to the nature

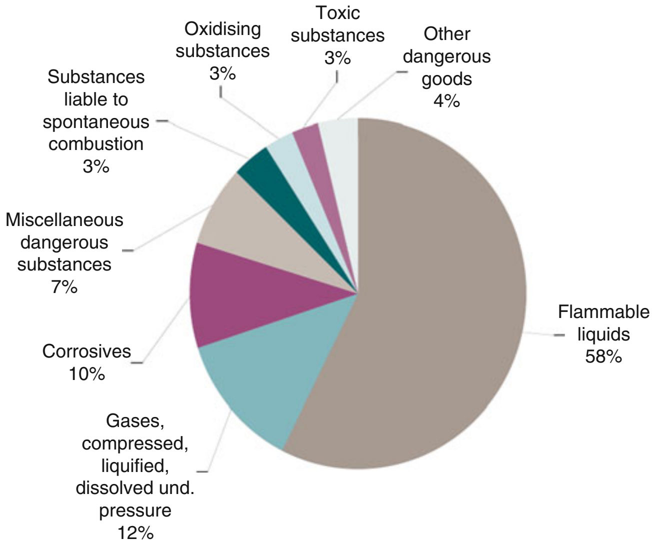


Fig. 3.2 DG transported by road, EU-27, 2007 (%tkm) (*Source*: EUROSTAT Panorama of Transport ISSN 1831–3280 Edition 2009)

of the transported goods and to the presence of exposed humans and materials in proximity to the place of incident [1]. For example, the transport of fuel such as petrol products can cause considerable fires or the explosion of the tankers in which it is transported, with heat, excess pressure and missile effects. Other substances have toxic properties and can be the origin of toxic gas clouds in the case of leakage due to the accidental puncturing of the tanker [2] (Fig. 3.3).

On a national scale it is shown that DG accidents on the roads make up no more than 0.1% of total accidents. But, even though this risk is minimal, the consequences are important when dangerous substances are involved. In France on the 8th September 1997, the collision between a vehicle transporting hydrocarbon and a lorry caused the deaths of 13 people and injured a further 43. In Italy on 9th February 1997 in a motorway accident a collision between a lorry transporting kerosene and a tanker caused a fire and a pile-up which resulted in 1 death and 40 injured. As with any mode of transport, accidents involving can take on far greater proportions than for example the crash itself, opening the door to hazards such as fire, explosion, chemical leak and environmental pollution, as well as possibly affecting areas beyond the actual scene of the accident.

Despite the high risks the public authorities and the management of the motorway and road infrastructure do not precisely know the nature, number and route of the DG transported on the territory. The only statistics produced are the result of manual surveys carried out in various periods of the year, furthermore, carried out by different bodies such as motorway companies, territory institutions or some state services (for example the Ministry of Transport and Infrastructure).

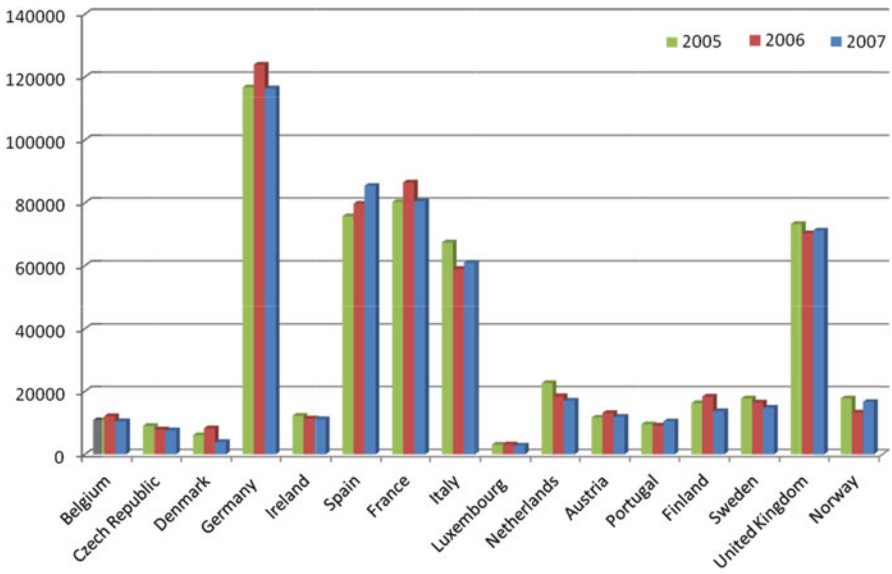


Fig. 3.3 DG Annual road freight transport of petroleum product (1,000 t)

3.2 Regulatory Placement and Background

The transport of DG is an activity which is increasingly international and multi method; the regulation involved can therefore not disregard connecting itself to international level to sustain a future integrated logistics system with multi method efficiency.

The ONU Recommendations for the DG transport, published for the first time in 1957 and periodically updated, are the point of reference for all the laws specific to the different methods of transport (sea, air, road, railway, rivers/canals) at international, community and national level.

The ADR law, acronym of European Agreement concerning the International Carriage of Dangerous Goods by Road, is intended to increase the safety of international transport of DG by road. The Agreement has been regularly amended and updated since its entry into force. This version has been prepared on the basis of amendments applicable as from 1 January 2009. Its Annexes A and B contain the technical requirements for road transport, i.e. the conditions under which dangerous goods, when authorized for transport, may be carried internationally, as well as uniform provisions concerning the construction and operation of vehicles carrying dangerous goods. They also establish international requirements and procedures for training and safety obligations of participants.

Unfortunately, these instructions concern a single shipments while a global framework to identify an overall view of risk for this kind of transport is missing. In particular, the simultaneous presence of more shipments on the same infrastructure or on a particular vulnerable area is not considered. This kind of approach should be tackled by the road infrastructure managers or, in general, by the public authority who supervises the transport system planning.

In the last years, the consciousness of the problem related to the safety on the road is increasing and in its efforts to limit the number of road fatalities by 2010, the European Commission has adopted a new directive. The Directive 2008/96/EC of the European Parliament and of the Council of 19 November 2008 on Road Infrastructure Safety Management aims to bring road safety management to higher standards throughout the EU major roads, i.e. the trans-European transport network (TEN-T) through infrastructure measures and better engineering. It defines guidelines and best practices for all stages of infrastructure management, including road safety impact assessments, road safety audits, network safety management and safety inspections. The directive does not impose technical standards or procedures but invites Member States to make better use of existing procedures and practices. Besides, the accidents in the Mont Blanc (France/Italy) and Tauern (Austria) tunnels in 1999 and in the Gotthard (Switzerland) tunnel in 2001 have highlighted the potential risk for the DG transport in the tunnels. In this context, the new Directive 2004/54/EC of the European Parliament and of the Council of 29 April 2004 on minimum safety requirements for tunnels in the Trans-European Road Network was emanated. The aim of the Directive is to ensure that all tunnels longer than 500 m, whether in operation, under construction or at the design stage, and forming part of the Trans-European Road Network, comply with the new harmonised safety requirements.

Finally, the Directive 2010/40/EU of the European Parliament on the framework for the deployment of Intelligent Transport Systems in the field of road transport and for interfaces with other modes of transport has entered into force on August 28, 2010. The EU Commission have recognize that innovation will have a major role to play in finding appropriate solutions for the Union and, in particular, Intelligent Transport Systems (ITS) represent advanced applications which aim to provide innovative services relating to different modes of transport and traffic management and enable various users to be better informed and make safer, more coordinated and "smarter" use of transport networks. According to this directive which cited "ITS integrate telecommunications, electronics and information technologies with transport engineering in order to plan, design, operate, maintain and manage transport systems. The application of information and communication technologies to the road transport sector and its interfaces with other modes of transport will make a significant contribution to improving environmental performance, efficiency, including energy efficiency, safety and security of road transport, including the transport of dangerous goods, public security and passenger and freight mobility, whilst at the same time ensuring the functioning of the internal market as well as increased levels of competitiveness and employment". Since 2002, Eni and DIST have been testing and developing innovative technological and methodological approaches to monitor DG vehicles in order to manage and control DG flow and prevent risk for people and environmental.

3.3 Innovative Initiative to Prevent Risk in DG Transport in Italy

The definition of the typology characteristics and the vulnerability of the infrastructure network and the relative flows of traffic assume great importance in the analysis of risk in the transport of dangerous goods [3].

Among the first applications in Italy in the monitoring of DG vehicles, SIMAGE and the ReLaMP projects should be mentioned. SIMAGE started in 1999 and lasted 5 years. SIMAGE project aimed at establishing an integrated system in Italy for environmental monitoring, management and reduction of risks, and management of related emergencies. The project has been carried out by the Joint Research Centre on behalf of the Italian Ministry of the Environment. The primary objectives are: a) to support local authorities in coping with pollution and risk related problems; b) to support strategic decision making at political level through improved information assessment and communication; c) to support communication with the public, as prescribed by existing and foreseen European directives. These objectives are being achieved through the integration of different systems and technologies for data persistence, its federation and its analysis [4]. Relamp started in 2006 and was supported by Regione Liguria and set up by Filse (Finanziaria Ligure per lo Sviluppo Economico) in partnership with Elsag and Set Italia. Relamp is based on a system that plans the itinerary for the vehicles' journeys thanks to a combination of territory data and data from the monitored vehicles during the transport. The aim of the project was to supply an efficient product to support the Fire Brigade, Traffic Police and other authorities in case of emergency, supplying useful data for giving aid. On this subject, the work finished in 2006, by Maja, Studer, Rainoldi et al. of the Politecnico di Milano in collaboration with the Ministry of Infrastructure and Transport, is also interesting. It regards the definition of the origin – destination itineraries, of great interest for the analysis of risk applied to the region of Lombardia [5]. The objective of the study was to individuate, among all available itineraries, that with the most interesting telly-control of the vehicles used for the transport of refinery products, aimed at minimizing the risk connected to that transport. After the first phase of geographic aggregation of data, carried out for the analysed origins (Rho-Pero and Sannazzaro de Burgundi, Italy) and for each goods category transported (petrol/oil and GPL), the social risk for each route was calculated, defined as the relationship between the frequency of an accidental event and the number of deaths which followed, represented graphically with curve F-N.

Subsequently, other international research projects were supported by a work group including DIST, Department of Information Department of Computer Communication and System Sciences of the University of Genova, Regione liguria, Department of Infrastructure and Transport, and Armines - Ecole des Mines de Paris (Association pour la Recherche et le Développement des Méthodes et Processus Industriels). In 2006, the project DAGOT, DAngerous GOods Transport (TSA6-CT-2006-044658) was financed by the European Commission in the VI FP. The partnership consisted of Regione Liguria (IT), FILSE (IT), DIST - Università degli Studi di Genova (IT), REC - The Regional Environmental Center for Central and Eastern Europe (Hungary) and ARMINES (FR). The goal of the project has two-folded perspective: to stimulate international cooperation as regards sustainable development and risk management in DGT (Dangerous Goods Trasnportation), by the exchange of experiences and by the production of a comprehensive European Reference Framework (DGT-ERF) to support decisions and to identify shared strategies, based on the integration of common knowledge, coming from different actors related to DGT to control, manage and plan DGT.

In 2005–2007, DIST and ARMINES have been financed by Regione Liguria and European Commission to realize an Interreg Project, a European initiatives of cooperation between France and Italy in the cross-border territories, related to the development of a distributed information system prototype to detect and monitor the hazardous material transport on the road on the territory Nice-Imperia-Ventimiglia [6]. Recently, a new project, to continue and improve the research studies and the methodological and technological developments in DG transport risk and prevention was financed by Interreg Programme in 2010–2012. In this case, also a uniform risk assessment procedure for the transport of DG through road tunnels and by multimodal transport will be developed. Those previous project have been made possible thanks to the collaboration of ENI S.p.A, the biggest Italian oil Company, which provided its know-how in the distribution of DG and who allowed the experimentation of different technologies for the monitoring of the vehicles.

In order to improve the research activities and to guarantee continuity in the technological and methodological approaches in the monitoring and detection of DG vehicle, DIST and ENI S.p.A., R & M Secondary Logistics, in the 2009 created a joint laboratory called DELAB. More in detail, DELAB research focuses on DG Transportation (DGT) integrating intelligent systems in order to prevent accidents to people or infrastructures and damages to the environment and its main activities are related to

- Real time monitoring architecture for vehicles carrying DG
- Support tools for staff and drivers including training, resource management and advanced data mining
- Decision Support Systems (e.g. orders and fleet management)

3.4 DGT Information System

In the last years, Eni, R & M Secondary Logistics, and DIST, have designed and developed innovative initiatives in transport safety and security and adopted stringent technical and functional standards in order to provide a high quality service, respecting the laws and rules of safety and environmental protection and for the protection of the health of the workers and the public.

Since transport services are almost completely outsourced, it is necessary to develop instruments for continuous monitoring of processes and performance, and to ensure an adequate level of control of this important aspect of the supply chain.

Fig. 3.4 Distribution box on the trailer

The general architecture of the developed DGT Information System (DGT IS) consists of

- On board architecture
- Transmission system
- Database
- GIS-based Applications
- Decision Support System

Those components are integrated in the TIP (Transport Integrated Platform), the tool used by Eni R&M Logsec to share data with suppliers of transport and, at the same time, implement an ongoing monitoring of performance. TIP, developed by DELAB, is a web portal, with secure and selective access, accessible via the Internet by both internal users (Eni) and by external parties (e.g. suppliers of transportation).

3.4.1 On Board Architecture

On board, the system consists of two sub-systems located in different positions:

– on the trailer there are various type of device: the electronic counter, a collection of analog/digital sensors for the detection of information critical to monitor the single components of the vehicle and the satellite device called " concentrator" to which it is connected.
– on the tractor there is an on-board computer (PDA - Personal Digital Assistant) in the driver's cabin equipped with a user interface for information and possible manual input by the driver. Furthermore in the back of the tractor there is a distribution box installed, allowing the connection to the CAN (Controller Area Network) transmitting through the protocol FMS Standard, to the odometer, to the power supply, to an emergency switch in the cabin and in the distribution box in the trailer. In case that vehicles operates in explosive atmosphere the box must satisfy some specific requisites ruled by ATEX directive (94/9/CE) (Fig. 3.4).

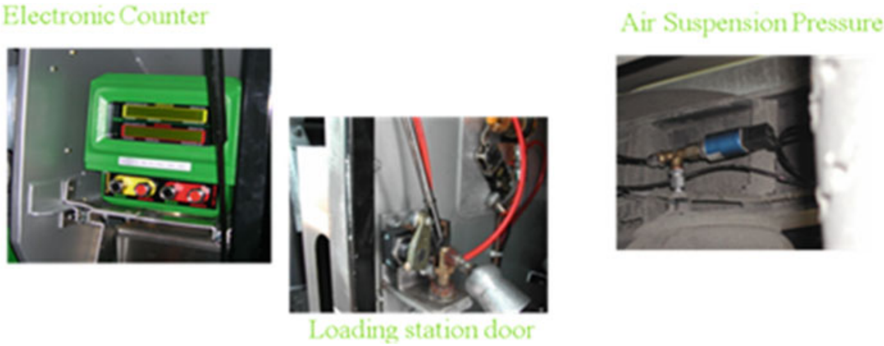


Fig. 3.5 On board sensors

The concentrator is equipped with a GPS antenna and a GPRS transmitter/receiver in order to detect in real time the position of the vehicle and all the parameters relative to the vehicle's operation and the state of the load. All the equipment and the wiring used must conform to the CEI law and be useable in areas with a possibility of explosion.

The concentrator performs the following activities:

- data collection from all the on-board sources
- data processing
- addition of geo-referenced information (by a GPS receiver)
- packing and sending messages on multiple queues (Fig. 3.5)

Such a kind of information must be transmitted at predefined fixed time intervals (e.g. 5 min). The state of the truck is described by the following information: the position of the vehicle, speed and direction (by GPS and odometer), inclination with respect to the two axes, air suspension pressure, CANBUS information, amount, temperature and pressure of product, state of the load (data comes from the electronic counter).

In the transport of high risk DG, the ADR 2005 introduced the creation of safety plans, that had to include the detection of the dangerous goods and the evaluation of the critical operations and the risks to safety also derived from the rest breaks of the vehicles and/or pauses of the goods in the vehicles. With this aim, other information related to particular events detected with the device installed on board the vehicle is:

– Start/stop: raised when electronic meter has turned on/off
– Loading/unloading: these events contains relevant information about operation in depots and service stations
– Loading door opening/closing: raised at the start/end of loading and unloading operations
– Overturning Risk Alarm: raised contextually with threshold crossing for inclinometer sensor

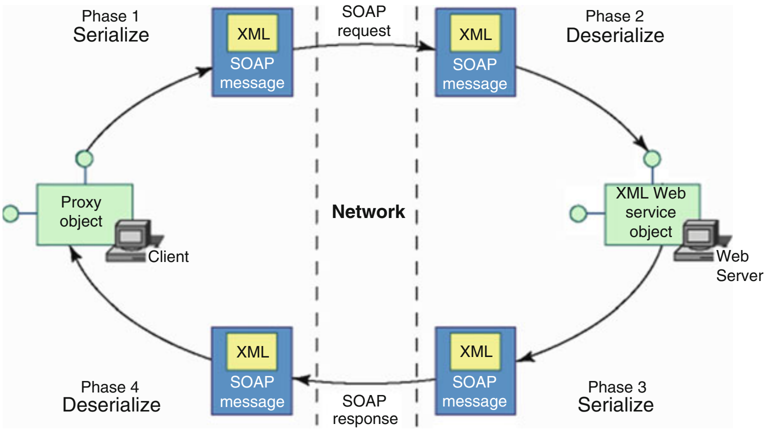


Fig. 3.6 Web Service communication

– GPS signal loss: raised when the concentrator looses the GPS signal due to low coverage or HW anomalies
– Emergency buttons: raised when driver push the proper button in case of emergency

3.4.2 Transmission System

Messages should be sent by a GPRS module equipped with one or more SIM-cards. SIM redundancy copes with the frequent issue of the signal loss due to the GSM uncovering of some areas. Italian Department of Infrastructure and Transport is evaluating the realisation of a SIM card dedicated to DGT. It should work with every mobile operator and it should have transmission priority. Another experiment regards also the introduction of satellite transmission, that can increase the territory coverage. Different technologies can be used to perform transmission between concentrators and a remote server over the net: socket connection over UDP, Socket connection over TCP or a Web Services.

The Web Service is a software system that "exposes" a public interface (webAPI) to be called remotely to exchange data over the net. Messages are "encapsulated" inside a XML envelope (coded with SOAP standard) and transmitted using HTTP over TCP. As shown in this Fig. 3.6, the communication is bidirectional: the server can answer to the remote clients with a simple acknowledgment or more complex commands. Web Service grants a solid infrastructure and can work on interoperable platform but the high overhead increase the costs considerably.

One of the most relevant aspects in the client/server data transmission is the definition of a common data format for the messages packed and sent by the concentrators. Such kind of standardisation make the higher layers of the

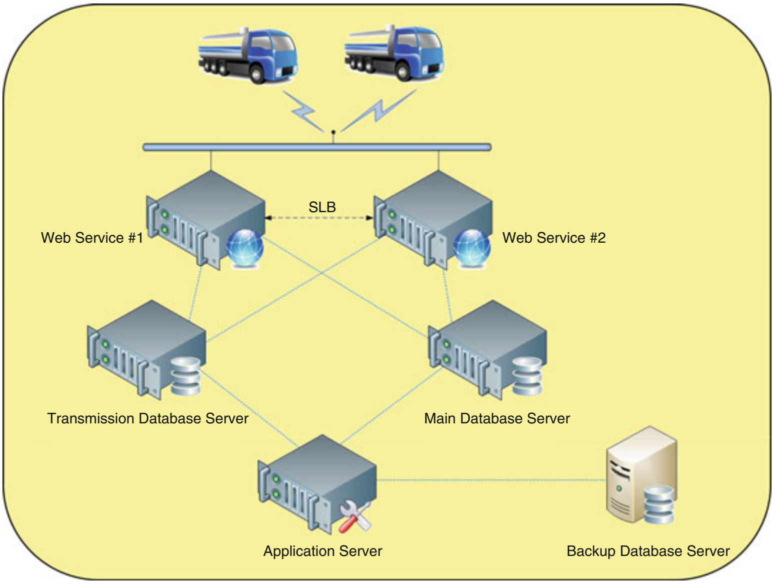


Fig. 3.7 DGT Database Infrastructure

information system independent of different HW infrastructures sited on the trucks. A good data format must be scalable and consistent with older versions. It must code both the request and the response with an efficient diagnostics management. It must clearly define the rate of transmission distinguishing between standard-frequency and on-event messages. The first ones periodically give information about the truck state (telemetry data), the second ones notify a particular on-board event. The approach used in the DGT IS consists of a “fixed part” which does not depend on the class of data followed by a “variable part” whose length and format depend on the kind of data. The “fixed part” includes information about creation and transmission date, vehicle identifier, coordinates and kind of transmitted data.

3.4.3 Database

The approach to the database layer must include:

- Efficient data storage for real-time messages received from the remote vehicles
- Good management to front-end application requests (e.g. GIS–based application)
- Backup and recovery strategy (Fig. 3.7)

This database receives all the raw messages from the Web Service and stores data for the diagnostic activities. The high rate of transmission can lead to a danger of table

Fig. 3.8 Web GIS Architecture

locking. In order to avoid this issue it is suggested that messages are not immediately parsed: an external application could provide this operation at regular intervals (e.g. 1 min), moving the unpacked data to the main DB server (M-DB). The M-DB server receives data from the T-DB via application server, provides data to the web applications, collects generic data (e.g. Trucks Registry, Users authentication) integrating them with the real time information and finally provides efficient mechanism to historicize daily data. The M-DB backups all data on a local storage device. The application server sends the backup files to the Backups-DB (B-DB). It can quickly substitute the M-DB in event of failure with minimum data loss and, periodically, the application server provides to delete oldest backup files (e.g. older than 1 month).

3.4.4 Geografical Information System (GIS)

A geographic information system (GIS) integrates hardware, software, and data for capturing, managing, analyzing, and displaying all forms of geographically referenced information. Considering DGT a GIS can assist to:

- Data presentation
- Risk analysis
- Decision support, i.e. in the definition of the vehicle routing optimization.

The proposed GIS-based application allows to use effective graphic interface, high scalability, a method to retrieve information from the interface and the ability to perform geographic calculation. More in details, available information includes maps, geocoding/reverse geocoding, routing, proximity researches (Fig. 3.8).

The Web server operates in order to provide the information requested by the user and in particular it uses the following procedures:

– Receive the client GIS request
– Call the Web map server to obtain the desired information
– If necessary process the received data and send the data back to the client

3.4.5 Decision Support System

A Decision Support System (DSS) is a class of information systems (including but not limited to computerized systems) that support business and organizational decision-making activities. In the optic of DGT, the final goal is to support decision maker in transportation planning, not only minimizing the cost but also the risk. The proposed approach on DGT DSS focus on the following thematic priorities:

- Risk definition on the routes covered by trucks
- Optimisation of deliveries planning
- Detection and evaluation of the dangerous goods flows on the specific road infrastructure

The first DSS topic is based on the elaboration of the data contained in the database for the definition of the risk ratio associated to the transport infrastructure. The ratio calculated takes into consideration not only the static parameters regarding the road but also the dynamic parameters, acquired in real time, regarding transport and weather conditions. The second allowed the definition of the best routing of the vehicles according to the logic of risk aversion and exposure minimisation. The third one implies the collection of data related to the routing of vehicles in order to analyse the paths of the main flows taking into consideration the exposure and the presence of critical road infrastructures [7].

3.4.6 Transport Integrated Platform (TIP)

The sections currently present in TIP are:

- Control Room
- Remote Monitoring
- Trip Planning
- Technical Management
- Service Stations
- Training
- Accident Reporting
- Document Management
- Quality
- Compliance/audit
- Operational control
- Performance Index (Fig. 3.9)

Each section has specific features and is dedicated, but the data are shared in order to keep the control of each aspect of the transport services.

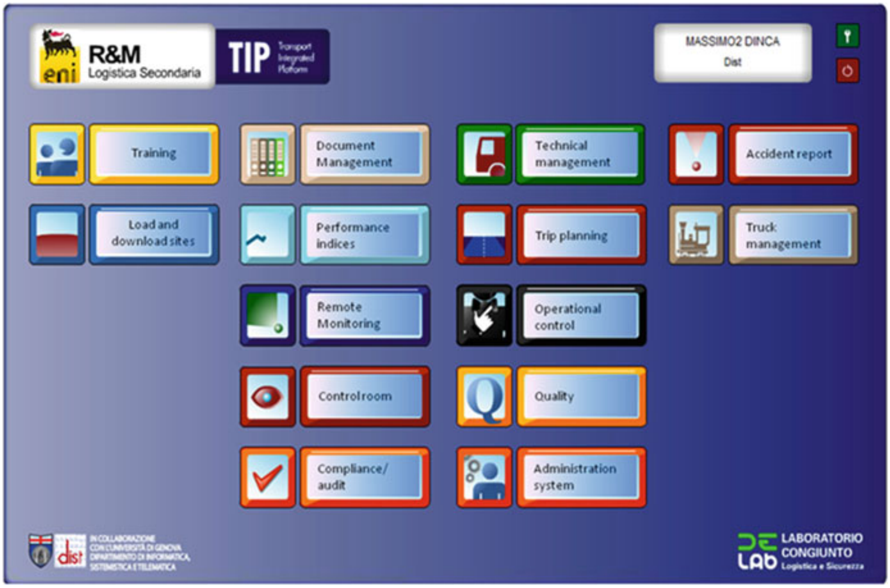


Fig. 3.9 TIP tool

3.4.6.1 Control Room

The control room ensures a continuous, real-time monitoring of vehicles equipped with remote control, both during the day services both where the service is done at night. The control room is equipped with advanced tools for managing alarms and anomalies detected during transport.

3.4.6.2 Remote Monitoring

This module is equipped with real-time monitoring of transport through the representation of data in tabular form and on geo-referenced maps and interactive maps. This allows the users to constantly monitor the operations and to extract and export data about the position of the trucks and the quality and quantity of the products loaded and unloaded (Fig. 3.10).

3.4.6.3 Trip Planning

The planning section is used to provide comparison in real time between the service planned and the service implemented. It's possible to compare the trip programmed with the data coming from the remote monitoring system installed on board (Fig. 3.11).

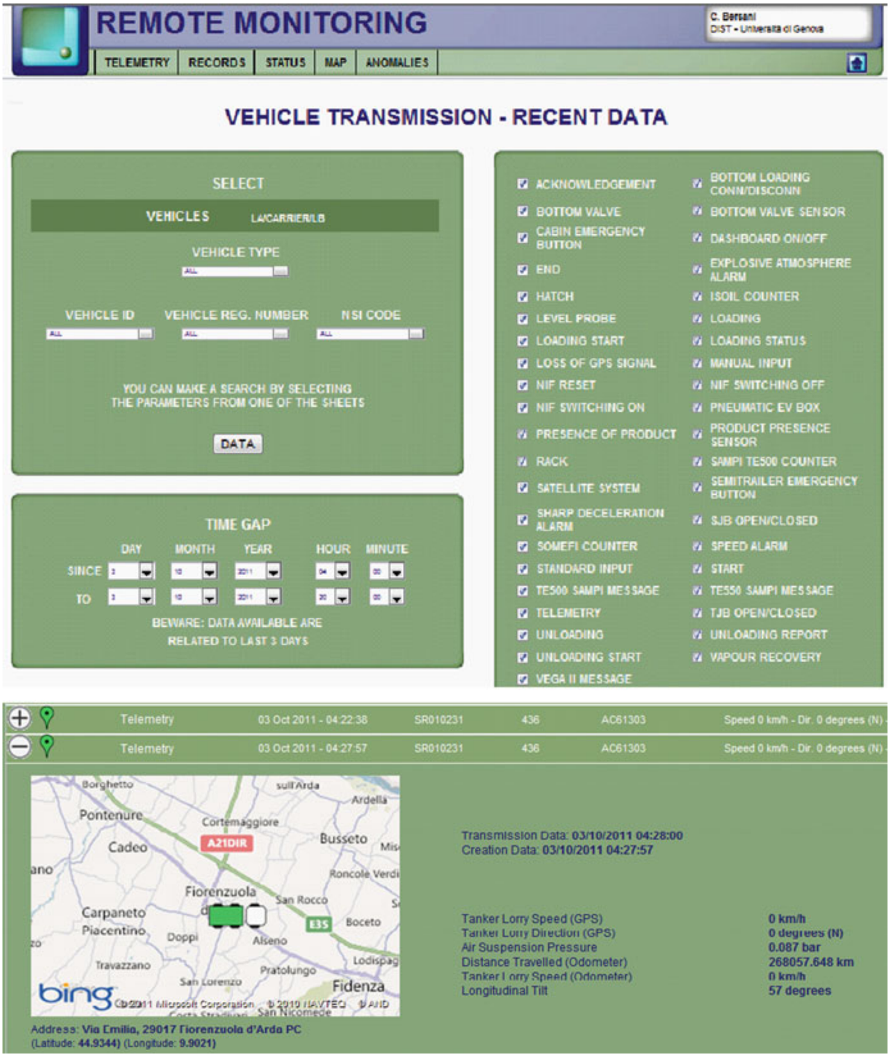


Fig. 3.10 Remote monitoring module

3.4.6.4 Technical Management

All the technical controls of the trucks are included in this section and are constantly updated during each inspection. Are available, in addition to the controls, all the expiring dates (audits, insurance, fire extinguishers, certificates, etc…) In this section are also available the technical data of rail tank wagons.

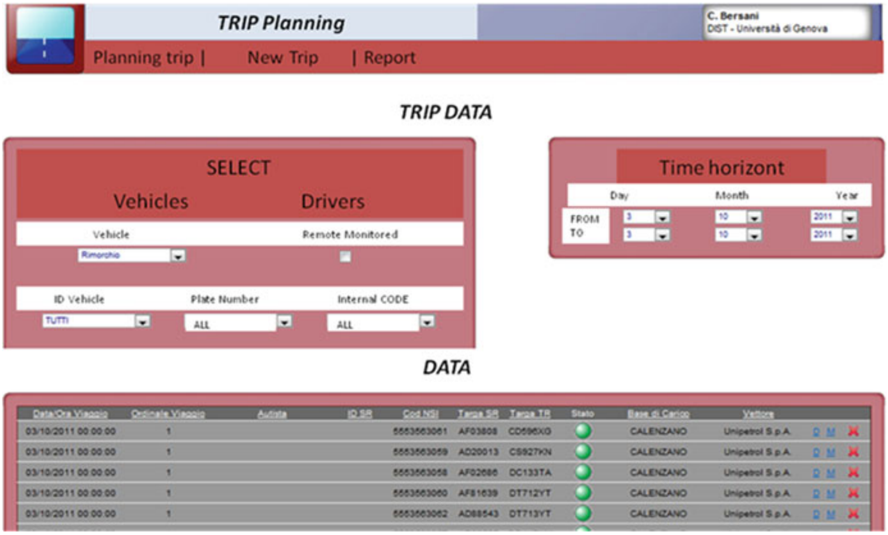


Fig. 3.11 Trip planning module

3.4.6.5 Service Stations

This feature provides the data of the level of products that is in underground tanks at service stations. This makes possible to know the volume of product in the service sations in order to check the correctness of the unloading of the vtrucks and automatically generate orders for delivery.

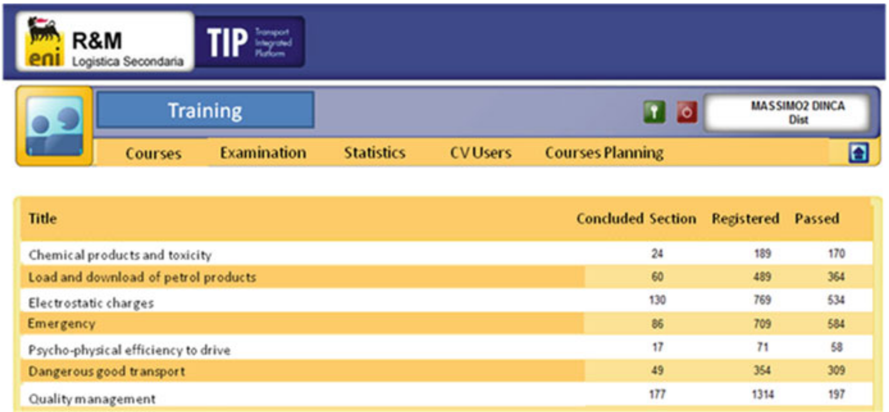
3.4.6.6 Training

With the Training module is possible to create training sessions. After each course is concluded, the tests are automatically generated by the system and the results are stored in the data base (Fig. 3.12).

The training section is integrated with the service Safety Game, an innovative feature that permits to send to the users a daily question about the courses that he attended.

3.4.6.7 Accident Reporting

The section of accident reporting is designed to collect information on any anomalous event that has caused damage or injury. All event information is placed in the database and classified by type of service and product transported, in order to be analyzed, exported and used for statistical purposes.



Title	Concluded Section	Registered	Passed
Chemical products and toxicity	24	189	170
Load and download of petrol products	60	489	364
Electrostatic charges	130	769	534
Emergency	86	709	584
Psycho-physical efficiency to drive	17	71	58
Dangerous good transport	49	354	309
Quality management	177	1314	197

Fig. 3.12 Training module

3.4.6.8 Document Management

All documents relating laws, regulations, standards and business processes are classified and available in this module to ensure a continuous updating and better sharing of information with suppliers.

3.4.6.9 Quality

The management of information flows, procedures, definition of roles and responsibilities and verification of operation of business processes are managed in a quality system. This module provides all the tools (non-compliance, improvement programs, complaints, corrective actions, etc…) to keep under control the Quality Management System in Secondary Logistics.

3.4.6.10 Compliance/Audit

Continuous assessment of compliance with laws and procedures is the best way to detect abnormalities in business processes and ensure continuous improvement. This module is available to suppliers in order to make self-assessments and audits concerning: Environment, Fire Prevention, Safety (D. Lgs. 81/08), Relevant accidents (D. Lgs. 334/99), ADR, RID, IMDG, LPG Storage, Fuel Storage, Quality (Fig. 3.13).

3.4.6.11 Operational Control

The section of operational control is used in combination with technical checks to verify the proper implementation of the operation of the service, procedures and standards contained in the contracts of transport. The correct application of

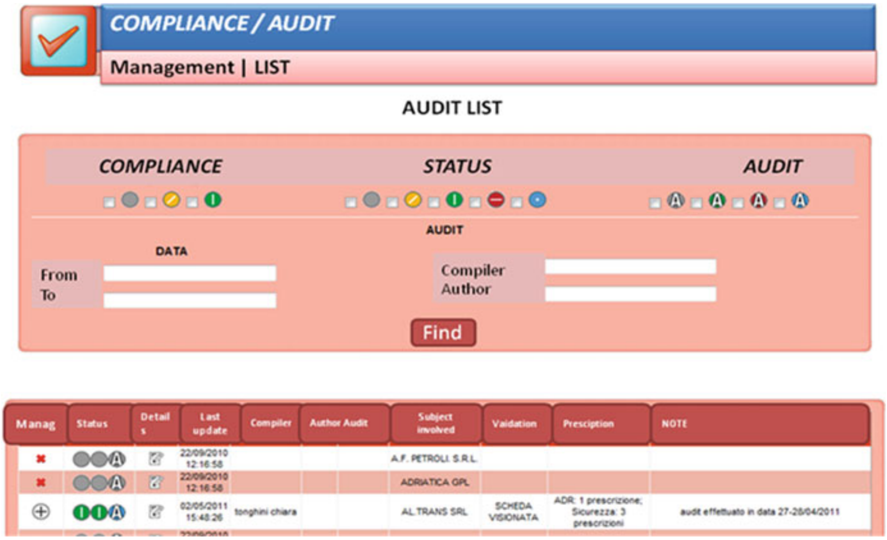


Fig. 3.13 Compliance/audit module

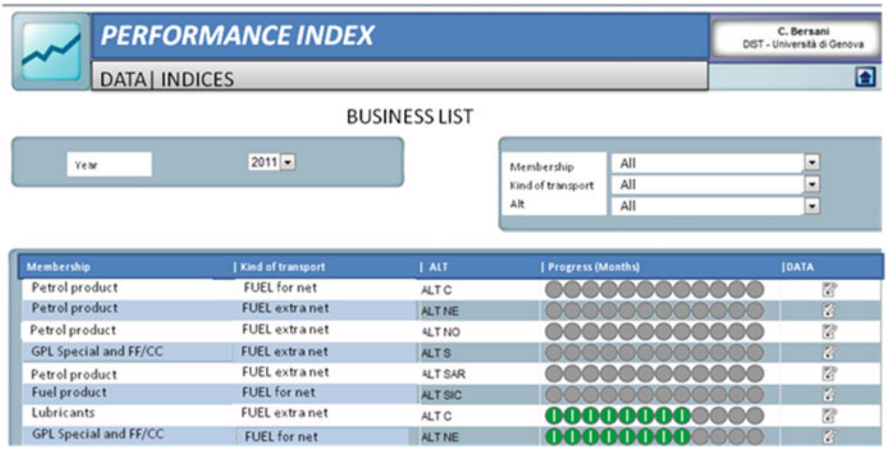


Fig. 3.14 Performance index module

procedures for loading, transport and unloading of goods is the basis for good customer service and to prevent accidents and injuries.

3.4.6.12 Performance Index

All modules above make it possible to produce "dynamic" performance indexes, which are generated in real time. In this way there is a continuous monitoring of performance and it's possible to make comparisons between different suppliers on the basis of objective parameters (Fig. 3.14).

3.5 Conclusions and Future Applications

The proposed Transport Integrated Platform (TIP) is based on dynamic, real time, as well as static data, and takes advantages combining the latest advanced telematics technologies, quantitative risk analysis and methodologies, optimization theories, and dynamic data collection systems, all incorporated into a integrated system.

Starting from TIP development, the main challenge in the next future on the dangerous good transport framework is to create different Regional Control Centres in order to monitor and manage a regional area by an integrated platform which gathers and processes in real time vehicle, traffic and environmental data and other important information provided and used by the different actors which are involved in the dangerous goods transport. The Control Centre should oversee the routing and monitoring of all dangerous goods vehicles within a certain geographical area, providing the necessary traffic and environmental data to them and inform in real time their logistic fleet managers for any unscheduled re-routing required. On the other hand, the Control Centre should also guarantee the divulgation of appropriate and reliable information and/or warning related to traffic or emergency situation to emergency services and public authority, to DG vehicle drivers and to common infrastructures users.

Significant efforts should be dedicated to accurately outline the Control Centre system architecture, taking into account the network of users (DG vehicles drivers, fleet and road infrastructure managers, common users, public authorities, etc.) but, above all, to define a standard communication protocols and data formats to be utilized by the cooperating entities.

Recently, SCUTUM project, (SeCUring the EU GNSS adopTion in the dangeroUs Material transport), promoted by Italian Ministry of Infrastructure and Transports and Telespazio S.p.A (Italy and France), started on February 2010, has launched and pursued a concrete path for the use of EGNOS-based services for the dangerous goods transport market in Europe, implementing large scale trials to support standardization and harmonization at European level. SCUTUM performs the necessary R&D activities enabling the evolution of technological elements of the service provision from prototype to standardized products. The project enhances existing solutions (on-board units and service/application platforms) to use EGNOS CS. In parallel with the implementation and trials activities, SCUTUM initiates a technical standardization related to EGNOS based services for the transport of dangerous material, by launching a CEN (European Committee for Standardization) Workshop aimed at drafting and publishing a CEN Workshop Agreement (CWA). The project will implement an EGNOS best-practice programme involving key players, combined with a standardisation effort and regulatory policy planning, at national level and in parallel with a cross-border basis as a first step, which will be further extended to other Member States and other freight types. SCUTUM is conceived to be in line with European policies and Action Plans in support of the telematics applied to freight transport and the use of EGNOS/Galileo, such as the Directive for the Implementation of the ITS Action Plan, the Freight Logistics Action Plan and related eFreight principles.

References

1. Bersani C, Minciardi R, Sacile R (2010) Economic and risk implications in the distribution of petrol products to service stations under retailer managed and vendor managed inventories. Int J Sustain Transport 4(3):129–153, 1556–8334
2. Emmanuel G, Roberto S, Samuel O, Chiara B, Nicolas A, Eva T, Mauro B, Davide G (2007) A spatial decision support system prototype for assessing road HAZMAT accident impacts on the population in a dense urban area: a case study of the city of Nice, French Riviera. In: The eight international conference on chemical & process engineering
3. Bersani C, Minciardi R, Sacile S, Tomasoni AM, Trasforini E (2008) An integrated system for the hazardous materials transport in a sub-regional scale area. In: Bersani C, Boulmakoul A, Garbolino E, Sacile R (eds) Advanced technologies and methodologies for risk management in the global transport of dangerous goods, NATO Science for Peace and Security Series - E: Human and Societal Dynamics (ISSN 1874–6276). Ios Press, Washington, DC, p 261, ISBN 978-1-58603-899-1
4. Peckham R, Atkinson M, Di Mauro C, Nordvik JP (2003) The SIMAGE project: an information infrastructure for environmental monitoring, risk reduction and emergency management in Italy. In: 9th EC GI & GIS Workshop, ESDI Serving the User, A Coruña, Spain, pp 25–27
5. Bruglieri M, Maja R, Marchionni G, Rainoldi G (2008) Safety in hazardous material road transportation: state of the art and emerging problems, 45th edn, NATO Science for Peace and Security Series - E: Human and Societal Dynamics (ISSN 1874–6276). Ios Press, Washington, DC, p 88, ISBN 978-1-58603-899-1
6. Benza M, Bersani C, Garbolino E, Giglio D, Olampi S, Sacile R, Tomasoni A, Trasforini E (2007) A distributed information system prototype to define and monitor the hazardous material transport on the road on the territory nice-imperia-ventimiglia. In: Second international conference on safety and security engineering, safe, Malta
7. Casazza R, Garbolino E, Olampi S, Bersani C, Trasforini E, Giglio D, Sacile R (2006) Detection and monitoring of hazardous material transportation on road between France and Italy: objectives, methodology and first results. In: Guedes Soares C, Zio E (eds) Safety and reliability for managing risk. ESREL 2006, Estoril, Portugal, vol 3, pp 2659–2666

Chapter 4
Dangerous Goods Transportation and Risk Management: A Cost-Benefit Analysis Approach

Mohamed Tkiouat

Abstract Cost-Benefit Analysis (CBA) is a decision-support tool which can help in discussions with stakeholders. It provides a structured framework for representing and weighting all the elements of a decision, thereby improving the transparency of the decision-making process. The tool is widely used in many countries for assessing environmental regulation, industrial safety and risk reduction in terrorist actions.

This chapter aims to provide guidance for the conduct of a Cost-Benefit Analysis related to Dangerous Goods Transportation (DGT) and the development of regulations in the area of industrial safety and security. It also aims to improve understanding of the underlying economic principles and assess the quality of the analysis of a DGT project. It presents an overview of the economic principles on which a CBA of Dangerous Goods Transportation is based, and describes the steps involved in conducting a study. It suggests data sources which may help to monetize the various nonmarket consequences of industrial activity (impact on people's health, safety and security, environmental pollution, noise pollution).

This chapter formed the foundation for the NATO Advanced Training Course 'Risk Prevention for Environment and Human Society against Dangerous Goods Transport Accidents and Malicious Intents: Methods and Tools' held in Paris, France, 5th–9th July 2010.

Abbreviations

CBA	Cost-Benefit Analysis
DGT	Dangerous Goods Transportation
DG	Dangerous Goods

M. Tkiouat (✉)
LERMA, Ecole Mohammadia d'Ingénieurs,
Université Mohamed V Agdal, Rabat, Morocco
e-mail: tkiouat@emi.ac.ma

E. Garbolino et al. (eds.), *Transport of Dangerous Goods*, NATO Science for Peace and Security Series C: Environmental Security, DOI 10.1007/978-94-007-2684-0_4,

LPHC Low Probability High Consequence
NPV Net Present Value
PIS Project Impact Statement

4.1 Introduction

Transportation is a very important element in the supply chain. When it involves the transportation of dangerous goods (DGT), decision-makers place a high priority on any potential risk and may have to mobilise substantial resources to reduce it. DGT therefore implies a substantial investment; with time, risk and uncertainty as the main components. Investment analysis concerns the evaluation and comparison of investment options. It considers technical, financial, economic and external outcomes in the long-term. A detailed examination of the Cost-Based Analysis approach can be found in the *Best Practice Regulation Handbook* produced by the Australian Government [1] and a bibliographical overview of hazardous materials transportation can be found in the *Handbook in OR & MS* [2].

Conventional investment analysis methods include: present worth, annual equivalent worth, rate of return and benefit/cost ratio amongst others. This chapter proposes an investment analysis based on a cost-benefit approach. The purpose of this is to make a systematic evaluation of the impact of the project through the use of a set of tools which aim to help improve the transparency of the decision-making process.

4.2 Dangerous Goods Transportation

There are thousands of different dangerous goods in use today. Dependence on hazardous materials is a fact of life in industrialized societies. For example, in the civil and military spheres there are:

– more than 50,000 nuclear munitions;
– about 80,000 t of chemical weapons of mass destruction;
– hundreds of thousands of tons of dangerous explosives;
– highly poisonous substances; and
– tens of thousands of highly reactive substances.

The United States Department of Transportation defines a dangerous good (DG) or hazardous material as ‘any substance or material capable of causing harm to people, property, and the environment’ [3]. The United Nations divides them into nine classes according to their physical, chemical, and nuclear properties [4]:

1. explosives and pyrotechnics;
2. gasses;
3. flammable and combustible liquids;
4. flammable, combustible, and dangerous-when-wet solids;

5. oxidizers and organic peroxides;
6. poisonous and infectious materials;
7. radioactive materials;
8. corrosive materials (acidic or basic); and
9. miscellaneous dangerous goods, such as hazardous wastes

Almost all dangerous goods are produced at a location other than their destination. Transportation therefore plays a significant role. The extent of this role depends on the size of the country and its level of industrialization. The transportation of dangerous goods can be classified according to the mode of transport, namely: road, rail, water, air and pipeline. Some shipments are intermodal; they switch from one mode to another during transit.

An accident/incident analysis of the transportation of dangerous good by class [5] shows significant differences in the use of these modes. While truck transportation accounts for approximately 94% of all individual shipments in the United States, it only accounts for 43% of tonnage (as the volume that can be shipped by one truck is limited compared to other modes of transport). In contrast, rail, water, and pipelines carry 57% of tonnage while accounting for less than 1% of all individual shipments. It is clear that these modes carry huge quantities of dangerous goods, although it is difficult to count individual shipments (How many shipments are there in a pipeline? Does a train consisting of multiple containers count as a single shipment?), they can clearly carry much larger quantities per shipment than trucks. The balance of shipments (5% numerically and 0.05% by weight) is made by air [6].

DGT incidents can occur at the origin or destination (loading, unloading or storage), or en-route. Incidents involving dangerous goods can lead to severe consequences characterized by:

- Fatalities;
- Injuries;
- Evacuation;
- Property damage;
- Environmental degradation; and
- Traffic disruption.

All of which must be evaluated in order to facilitate the comparison of DGT options.

The reasons for failures and accidents, according to developers, manufacturers and consumers are:

- Poor quality projects;
- Poor quality testing in development;
- Poor quality operational tests;
- Poor quality monitoring of operations;
- Deterioration and aging of operational equipment;
- Poor quality work carried out by personnel (due to social factors);
- Human error; and
- Terrorist actions;

4.3 DGT Accident Statistics

About 90% of incidents occur on highways. Human error seems to be the single greatest causal factor (for both minor and serious incidents). In the United States in 2003, there were 488 serious incidents (from a total of 15,178 incidents) resulting in 15 deaths, 17 major and 18 minor injuries, and total property damage of US$37.75 million [7]. In Morocco in 2003, there were 18 accidents involving dangerous goods resulting in 15 injuries and 4 deaths; and in 2004, 9 accidents resulting in 9 injuries and 1 death (see [8]).

The annual number of non-DGT accidents in the United States is estimated to be 126,880, which compares to approximately 15,000 DGT accidents and incidents (Federal Motor Carrier Safety Administration 2001 see [9]). Although DGT is only involved in a small minority of all transport accidents, this type of accident can have catastrophic consequences. For example in 2003, 22 containers derailed at Tamaroa (USA), resulting in the release of various types and quantities of dangerous substances from 7 of the containers. Over a 1,000 residents within a three-mile radius were forced to evacuate their homes and a major highway was closed followed the derailment.

The cost of an incident resulting in fire or explosion is significantly higher than the cost of a non-DGT incident. DGT accidents are commonly perceived as low probability–high consequence (LPHC) events and the data seems to support this idea. Most of the transport accidents that impact a large number of people and result in significant economic loss involve the transportation of dangerous cargo.

4.4 The Problem of Decision Making

The seriousness of a DGT accident depends on the probability of its occurrence and the consequences. Decision-makers must act to minimize these two aspects. Their job is made more difficult by the fact that DGT involves multiple players, such as shippers, carriers, packaging manufacturers, freight forwarders, consignees, insurers and emergency response teams. Each has a different role in safely moving dangerous goods from their origin to their destination. There are often multiple handovers of material from one party to another during transportation. The various parties, ranging from individuals or small firms to large multinational organizations, may have overlapping and unclear responsibilities for managing the risks (see [10]). Furthermore, each party may have different priorities and viewpoints.

Although a government department or local authority may be responsible for designating routes that reduce risk, carrier companies generally try to identify the route that minimizes fuel costs and travel times for each shipment. This problem can be modelled using games theory. Some routes are short but pass through heavily populated areas; some routes avoid heavily populated areas but are longer (resulting in higher transport costs and accident probabilities); while other routes use major motorways to minimize travel time, but may be incur higher accident rates.

Therefore, DGT is a typical multi-objective problem with multiple stakeholders. These types of problems are complicated to solve [11, 12]. To overcome this difficulty, the decision-maker must decide priorities. The answers to the following questions may help:

– Does the project support other community or economic goals?
– Does the project fit with the community's long-term goals?
– Does the project correspond to a community need or want?
– If the project is a want, what is compelling about it?
– Will the project put any individual or group at a disadvantage?
– Can the community afford to maintain and operate the project after it is created?
– Is there strong support for the project in the community, across factions?
– Has the project been defined in local plans or studies?
– If this project is undertaken, will it be a duplication of other efforts?
– Does the community fully understand the project and its parameters?
– Does the project resolve any compliance issues?
– How will the project be funded?
– Will there be staff requirements if the project is implemented and operated?

DGT issues are further complicated by public sensitivity. This is rooted not only in risk perceptions, but also in equity concerns. Those individuals benefiting from DGT shipments are usually those who live near the production facility (the origin) or the delivery points (the destination). However, the population living alongside a major highway connecting these origins and destinations is exposed to risks regardless of whether or not they benefit. This imbalance in costs and benefits is a source of public opposition to DGT shipments. The shipment of spent nuclear fuel rods from nuclear power plants to a proposed repository at Yucca Mountain in Nevada, USA, provides a good example of equity-based public opposition. Shipping the rods reduces risk at the power plants. However, this creates risk for the population living along the major east–west highways and railways, who are asked to assume the risk with no clear benefits to themselves. Furthermore, if the same route were selected for shipments originating from other sites, objections from people living along the route would increase considerably. These people are likely to prefer alternate routings that would spread the risk.

Public opposition to DGT shipments has increased in recent years, due to fears of terrorist attacks on DGT vehicles. The Research and Special Projects Administration (RSPA) of the United States Department of Transport accepts that dangerous goods pose a significant threat during transportation. They are particularly vulnerable to sabotage or misuse by terrorists as weapons of mass destruction or as weapons of convenience - particularly given how easy the current placard system makes it to identify a DGT vehicle (and its cargo). As a result some jurisdictions are trying to force the re-routing of DGT vehicles away from populated areas using local legislation.

Many of these points apply equally to the location of dangerous facilities, where the level of risk and public opposition is even higher.

4.5 Cost-Benefit Analysis (CBA) of a DGT Proposal

A CBA is a systematic evaluation of the impact of a DGT proposal. It takes into account all the potential effects on the community and economy i.e. not just the immediate, direct or financial effects, or the effects on just one group. It focuses to the greatest extent possible, on evaluating in monetary terms the gains and losses from the proposal. In principle, CBA measures the efficiency, or the effects of resource allocation. It sums the monetary value of the gains and losses for all of those affected. If the sum is positive, the benefits exceed the costs; the proposal would increase efficiency and would be chosen as a candidate solution.

CBA is useful because it:

– provides decision-makers with quantitative information about the likely effects of a proposal;
– encourages decision-makers to take account of all the positive and negative effects of a proposal, and discourages them from making decisions based only on the impact on a single group within the community;
– quantifies the impact of DGT proposals in a standard manner, which promotes comparability, assists in the assessment of relative priorities and encourages consistent decision making;
– captures the various linkages between the DGT proposal and other sectors of the economy, helping decision-makers to maximise the net benefits to society; and
– helps to discover cost-effective solutions to policy problems by identifying and measuring all costs.

Even when it is difficult to precisely estimate costs or benefits, CBA clarifies and makes transparent assumptions and judgements. Furthermore, attempting to quantify costs and benefits encourages closer examination of these factors. Issues involved in preparing a CBA include: how to deal with costs and benefits that are difficult to measure, how to take equity effects into consideration, how to determine the social discount rate and other common CBA pitfalls.

4.6 Principal Stages of a Full CBA

The CBA is structured using the following 12 steps:

1. Definition of the DGT problem.
2. Establishing the objectives of the intervention.
3. Identification of DGT policy options.
4. Calculation of costs and benefits.
5. Selection of impact and measurement indicators.
6. Assessment of the whole life impact of the proposal.

7. Monetization of the impact.
8. Calculation of future discounted costs and benefits.
9. Calculation of Net Present Value.
10. Sensitivity analysis.
11. Rank policy options.
12. Draft and publish a report.

Conducting a well-executed CBA requires decision-makers to follow a logical sequence that matches the steps involved in the preparation of the Project Impact Statement (PIS). The following provides an overview of these main steps using a costed example.

4.6.1 Definition of the DGT Problem

This stage of the project outlines the fundamental problems that decision-maker propose to address. One key component is to forecast the potential future outcome under a ‘do nothing’ scenario (i.e. with no intervention). Most of the work involves the risk assessment which includes identifying and calculating:

– Any adverse outcomes that result from a lack of intervention;
– The probability that the outcome will occur;
– How widespread the outcome is likely to be;
– Who is likely to be affected by the adverse outcome; and
– What harm or injury is likely to occur?

Risk comes in a variety of forms, and the type of risk affects the forecast techniques used. The following two types of risk can be distinguished:

– Actuarial risks. These can usually be estimated using historical data, and arise when an event occurs frequently. For example, the likelihood of a traffic accident at a particular location can often be forecast from historical crash data.
– Latent risks. This is where there is a long time delay between an adverse event and any negative consequences. These are more difficult to forecast, and analysis might need to be based on scientific research of the risk.

The level of uncertainty surrounding any forecast should be made clear in the Project Impact Statement. A section describing the problem should distinguish between real (or actual) risks and perceived risks. This is important as decision-makers frequently initiate a DGT project in response to the exigencies of international financial institutions and/or public calls for increased protection, particularly following a significant adverse incident. It is also important to note that the problem definition will never provide conclusive scientific evidence of the risk faced by society, but rather, an estimate of future uncertainties based on different scenarios.

4.6.2 Objectives of the Intervention

The PIS should contain an objectives section which clearly outlines the aims of the intervention. In an uncertain situation, objectives might include:

– reducing the probability of an adverse event occurring, or
– reducing the potential cost of an adverse event if one does occur, or
– reducing the probability and the potential cost of an adverse event if one does occur.

However, achieving any level of risk reduction entails costs, and in reality individuals make day-to-day decisions about the level of risk (compared to the cost) they are prepared to accept. The achievement of zero risk is neither an appropriate, nor technically feasible intervention goal. The aim of the PIS is to identify how much risk is acceptable to society, and the cost that society is prepared to pay to achieve it. Transparency and consultation are the best ways to identify this trade-off.

4.6.3 Identification of DGT Policy Options

Decision-makers must specify the set of policy options available to solve the DGT problem. One of the alternatives should always be 'maintain current arrangements', which provides a baseline from which the incremental costs and benefits of other alternatives can be determined [13, 14]. Different options can affect the probability of an adverse event or reduce its consequences in different ways and in different scenarios:

– Risk avoidance: prohibits behaviour that gives rise to risks;
– Risk transfer: causing another party to bear the consequences of risk;
– Risk retention: ensuring parties that face a risk bear the consequences of adverse events; or
– Risk reduction: activities to reduce the probability or consequence of risk events.

Some of the options available to departments and agencies that handle risk include: compulsory insurance schemes, licensing requirements, enforcement of codes and standards, self-insurance and industry-insurance schemes, proscription of particular activities or an outright ban. Some factors that a DGT project has to take into account are:

– the technical capacities of vehicles;
– the infrastructure;
– assessment methods;
– staff training; and
– regulations (safety and security).

4.6.4 Calculation of Costs and Benefits

For most DGT projects, it is more appropriate to measure costs and benefits at a national, rather than an international level. Effectively this means that the costs and benefits to all residents of the country should be counted, as far as practical. Welfare economics is the heart of public policy and at the core of CBA; therefore DGT costs and benefits have the following properties. They are:

- Wide scope and long-term;
- Private and social, direct and indirect, tangible and intangible;
- Benefits are based on standard welfare principles:
- Benefits are based on the consumer's willingness to pay; and
- Costs are based on what the loser is prepared to accept as compensation.

It is necessary to make clear that market prices are not always good indices in DGT projects. Some techniques will be outlined to overcome these difficulties.

4.6.5 Selection of Impact and Measurement Indicators

The project should identify the full range of potential impacts. Identifying the costs and benefits involves comparing the outcome of the proposed change with the outcome if it is not made. For each policy option, it is important to identify the incremental costs and benefits relative to the base case (maintaining current practice). Changes that would have occurred anyway should not be included. Benefits can be thought of as any and all effects considered desirable by those affected, while costs are all undesirable effects. CBA requires the explicit identification of ways in which the proposed changes make individuals better or worse off.

The choice of indicators to measure the impact of policy changes depends on the data available and how easily it can be monetized. For example, a proposal may reduce the risk posed by a hazard. The positive impact might be measured in terms of fewer accidents. The benefit of accidents that don't happen can also be given a monetary value [15].

4.6.6 Assessment of the Whole Life Impact of the Proposal

Impacts should be quantified for each time period in the duration of the project. The total duration of the project needs to be long enough to capture all of its potential costs and benefits. One particular difficulty of long-term costs and benefits forecasting is that environmental regulation can have long-term consequences.

The outcome of a project will always be uncertain. An assessment of these uncertainties should be a standard component of the evaluation of any project. The expected values and variability in costs and benefits should be assessed, and all risks should be taken into account.

A CBA should make a best-estimate of expected costs and benefits, together with a description of the major uncertainties and how they have been taken into account. It should set out how costs and benefits are likely to vary according to the general economic conditions or other influences. For example, would a large relative price change (such as a rise in energy prices or real wages) significantly change net benefits? If so, how would this affect the price path? More generally, the CBA should not make the simple assumption that net benefits for 1 year will be repeated every year [16].

Although it can be difficult to predict the long-term effects of a project, or in some cases, to attach objective probabilities to various scenarios, decisions will require some assumptions to be made. The CBA should make these assumptions explicit and provide a justification. This improves the planning of implementation and identifies areas where forecasting needs to be improved. It is a first step towards dealing with the uncertainties the project may create.

Consultation may help to understand the probabilities or consequences of particular risks. Key experts or interest groups should be contacted during the consultation period. The aim is to collect as much feedback as possible about the probability of an adverse event and its consequence, for each of the available options.

4.6.7 Monetization of the Impact

The net monetary value measures the efficiency effects of change. It represents the gains and losses all of those affected by a particular initiative. The amount an individual would (if necessary) be prepared to pay to obtain (or avoid) a particular change, is a measure of how much it is worth to them. The amount can be positive or negative, depending on whether the change makes them better or worse off. Summing these amounts across all those affected places a monetary value on the community's willingness to pay for the change. If the sum is positive, the change improves efficiency.

Overall costs and benefits to the community are calculated without taking into account individual gains or losses; one person's gain cancels another's loss. This assumption enables resource allocation to be separated from distribution effects, or efficiency from equity effects. However, it does not mean that distribution considerations are unimportant or should be neglected. It does mean that they should be taken into account as a separate part of the overall analysis. It may be more important than the resource allocation assessment, and should be distinct from it.

Monetary values can be estimated from observed behaviour. The value people place on something can be found by observing how much they are willing to pay for it. Market behaviour often reveals people's valuations; it is at least a guide. However, quantification can still be difficult as some impacts are uncertain, others are difficult to value in monetary terms, and some are both uncertain and difficult to value. Outputs and inputs are easier to value when they are bought and sold in competitive markets, even when markets are distorted by taxes. But competitive markets for the

Table 4.1 Sample framework for estimating expected annual costs and benefits

	Public and private costs (Ci)	Benefits (Bi)	Probability of occurrence of Ci and Bi	Expected (E) net benefit E(Bi)–E(Ci)
Injuries prevented and lives saved				
Clean-up costs avoided				
Product loss avoided				
Evacuations avoided				
Property damage avoided				
Environmental damage avoided				
Traffic incidents avoided				

relevant outputs may not exist. For example, a project may result in non-market benefits (such as an increase in safety, security or a reduction in pollution emissions). The following data will help to achieve an estimate of annual costs and benefits:

1. The amount individuals would, if necessary, pay to obtain (or avoid) a change, measures how much it is worth to them.
2. Summing these annual amounts across all those affected gives the community's total willingness to pay for the change (Table 4.1).
3. The annual costs (C) and annual benefits (B) for all those affected are summed without taking into account individual gains or losses.
4. The annual efficiency effect of the proposal = B – C.
5. Any additional resources indicate the size of the additional economic contribution.
6. Costs and benefits can be discounted (this is explained in the next section).

4.6.8 Calculation of Future Discounted Costs and Benefits

Most of the costs and benefits of a project are distributed over a long time period, and their value depends on when they are received. Positive market interest rates mean that the currency will be valued less highly in the future than now. Inflation is another reason for this effect. Future costs and benefits can be valued in nominal or real currency. The two are related by the formula:

$$(1+R)(1+P)=(1+N)$$

Where:

- R is the real interest rate;
- P is the expected rate of inflation (the percentage increase in the price index over a defined period); and
- N is the nominal interest rate;

This gives:

$$R = (N - P) / (1 + P)$$

In the nominal approach, the impact of predicted inflation is explicitly reflected in projections (cost and benefit streams increase more rapidly as expected inflation increases). This real (or constant) price approach expresses all variables in terms of the price level of a given year, usually the current year.

CBA measures the value people place on various policies, ideally using their willingness to pay (as shown by their market behaviour). Consequently, the preferred approach is to base the discount rate on market-determined interest rates, which indicate the value (to the current population) of future net benefits. It can be difficult to determine an appropriate discount rate. Market interest rates indicate the opportunity cost of any capital used in the project, i.e. what it would have produced if used for something else [17]. However, it can be difficult to identify what the alternative uses of capital might have been and what the capital would have produced in those situations. In this case, the social discount rate is used to attempt to fully represent the preferences of all individuals (including future generations) directly or indirectly affected.

4.6.9 Calculation of Net Present Value

To calculate the Net Present Value (NPV) of a policy option, costs and benefits need to be quantified for the expected duration of the policy. NPV is then calculated as:

$$\text{NPV} = \sum (B_t - C_t) / (1 + r)^t$$

Where:

B_t = the benefit at time t
C_t = the cost at time t
r = the discount rate
t = the year
T = number of years over which the future costs or benefits are expected to occur (the current year being year 0)

A worked example

An industrial site must install new equipment to limit air pollution. The equipment costs €5 million to install, and will operate for the following 4 years. On-going (annual maintenance) costs are €1 million/year (assuming constant prices). Benefits are estimated at €3 million/year (constant prices). Discount rates are 3% and 5% (see Table 4.2).

As with any uncertain variable, a sensitivity analysis should be carried out. Net present values should be calculated using the real discount rates of 3% and 5%.

Table 4.2 NPV—a worked example with 3% and 5% (in millions of €)

	Expected cost	Expected benefit	Expected net annual benefit	3% discount rate	5% discount rate
Year 0	5	0	−5	−5	−5
Year 1	1	3	2	1.94	1.9
Year 2	1	3	2	1.89	1.81
Year 3	1	3	2	1.83	1.73
Year 4	1	3	2	1.78	165
NPV				**2.44**	**2.09**

If the NPV changes from positive to negative (or vice-versa) the sensitivity analysis reveals that the choice of discount rate is significant. More thought needs to be given to the choice of an appropriate rate.

The NPV of a policy option is equal to the present value (PV) of benefits (B) minus the present value (PV) of costs (C):

$$NPV = PV(B) - PV(C)$$

If the NPV is positive, the policy improves efficiency. If the NPV is negative, the policy is inefficient. If all costs and benefits cannot be given a monetary value, the report should explain why any non-monetised costs or benefits are large or small relative to monetised elements.

The cost/benefit ratio is calculated as:

$$\mathrm{PV(B)\,/\,PV(C)}$$

If this ratio is greater than 1, the policy improves efficiency.

4.6.10 Sensitivity Analysis

There may be considerable uncertainty about predicted impacts and their associated monetary valuation. Sensitivity analysis provides information about how changes in different variables affects overall costs and benefits [18]. It shows the sensitivity of predicted net benefits to changes in the values of uncertain variables and to changes in assumptions. It tests whether uncertainty over the value of certain variables in fact matters, and identifies critical assumptions. It helps to assess uncertainties and determine reasonable expected values for costs and benefits. The process of thinking about, and trying to quantify uncertainties is in itself valuable. It identifies the factors critical for policy success, and enables decision-makers to focus more attention on estimating and managing them in order to reduce uncertainty.

If sensitivity analysis is to be useful to decision-makers, it needs to be systematic and clearly presented. Common approaches to sensitivity analysis include the following:

4.6.10.1 Worst/Best Case Analysis

The base case assigns the most plausible values to variables to produce an estimate of net benefits that is considered most representative. The worst, or most pessimistic, scenario assigns the least favourable of the plausible range of values to variables. The best, or most optimistic, scenario assigns the most favourable of the plausible range of values to the variables. If the pessimistic scenario results in a NPV below zero, it is then necessary to investigate the critical elements driving the value of the proposal, using the following two techniques.

4.6.10.2 Partial Sensitivity Analysis

This examines how net benefits change as the value of a variable changes through a plausible range (holding other variables constant). It should be used for the most important or uncertain variables, such as estimates of compliance costs, forecasts of benefits and the discount rate. It may also be important to vary the values assigned to 'intangibles', especially when the assumed values are controversial. Partial sensitivity analysis clarifies how CBA results are affected by uncertainty about the level or value of a variable. If varying the value of a parameter has a large effect on net benefits, uncertainty about its value becomes important.

4.6.10.3 Monte Carlo Sensitivity Analysis

This creates a distribution of net benefits drawing key assumptions or parameter values from a probability distribution. If net benefits do not change sign (i.e. from positive to negative or vice-versa) after examining the range of scenarios, there can be confidence in the efficiency effects.

4.6.11 Rank Policy Options

The PIS should specify which of the available options is the most efficient. Generally, this will be the one with the largest NPV. Given that NPVs are predicted (average) values, the sensitivity analysis might suggest that the alternative with the largest NPV is not necessarily the best alternative under all circumstances. For example, the recommended option might be the one with lower expected net benefits, but with a smaller chance of imposing a significant net cost on the community. The recommended option should be clearly explained and the analysis published in a report (explained in the following step).

4.6.12 *Draft and Publish the Report*

The project report should comprise the following elements:

(a) A summary or executive summary presenting the critical hypotheses, the main results and recommendations;
(b) The context of the study, explaining why it was undertaken;
(c) The objectives of the project(s), program(s) or activities studied;
(d) The main risks associated with the various options;
(e) A description of the various scenarios;
(f) The temporal profile of costs, benefits and net benefits;
(g) The main hypotheses underlying the analysis;
(h) The discount rate used;
(i) The calculation of Net Present Value;
(j) An uncertainty analysis;
(k) Other important information such as distributional effects, other quantified costs and benefits, and any relevant unquantified factors;
(l) A comparison of the preferred option with the alternatives.

The report should be disseminated to all stakeholders, and their feedback should be incorporated.

4.7 Costs and Benefits That Are Difficult to Value

When a policy proposal concerns marketable goods, estimating costs and benefits is in most cases conceptually straightforward and is explained in a number of existing CBA guides. However, it is often difficult to identify and measure the effects of a project, especially when there are policy impacts on un-tradable goods, such as pollution levels, access to scenic views and safety improvements [19].

Costs and benefits can be difficult to give a monetary value to, either because their magnitude may be unknown or uncertain, or because, even if their impact is known, they are difficult to express in monetary terms. Examples include environmental, social and cultural considerations, regional impacts, health and safety, and national defence and security. When valuations are uncertain, a sensitivity analysis should be used to test how variations in an assigned value affect the overall viability of the policy option. If the impact cannot be valued, it should still be quantified in non-monetary terms. For example, a regulation to reduce pollution could quantify the expected reduction in emissions. The quantification should aim to identify any assumptions made in determining the effects, the impact on the community (such as the number of people affected and in what way) and the probability of the full impact of the factor.

The process of trying to describe and measure costs and benefits is valuable in itself. By examining what determines the costs and benefits and how they are likely to vary, policy-makers are encouraged to consider different approaches and determine the best way to achieve intangible objectives. Is the current policy optimal, or could a better outcome be produced by another alternative? Even qualitative descriptions of the pros and cons associated with a proposed action can be helpful.

A wide range of tools have been developed to estimate the value of costs and benefits when market information is not directly available. These include: revealed preference techniques, stated preference techniques and others. The following will be described in greater detail:

1. Revealed preference techniques:
   (a) Market analogy method
   (b) Trade-off method
   (c) Hedonic pricing method
   (d) Travel cost method
   (e) Defensive expenditure method
2. Stated preference techniques
3. Impact valuations from secondary sources
4. Costs and benefits that cannot be valued in monetary terms;
5. Cost-effectiveness analysis;
6. Equity accounting;
7. Social discount rate;
8. Accounting for uncertainty;
9. Accounting for future generations;
10. Common cost-benefit analysis pitfalls.

4.7.1 Revealed Preference Techniques

Revealed preference techniques infer value from observed behaviour. A non-marketed good's value may be inferred indirectly from markets for related goods.

4.7.1.1 Market Analogy Method

The public sector supplies many goods and services that are also provided by the private sector. When provided by the public sector these goods and services are often free or at below-market prices. It may be possible to estimate the value of a publicly provided service by using data on a similar good that is produced by the private sector and sold in a 'normal' market.

4.7.1.2 Trade-Off Method

Trade-offs made in other markets may provide information on people's willingness to pay for the non-marketed good. The amount people are willing to pay for increased safety (e.g. air bags, smoke detectors or other risk-reducing goods) can be used to infer the value of an incremental change in the probability of death, or the statistical value of a life (assuming consumers are fully aware of the risks they are avoiding) by using consumer or job market studies' techniques.

4.7.1.3 Hedonic Pricing Method

Some market goods comprise bundles of characteristics, some of which are intangible. For example, the attributes of a house may include a scenic view and its environment, which may be noisy or polluted. The value of these intangible characteristics can be inferred from house prices. The difference in price between a house under a flight path and a comparable house in a quiet street provides an estimate of how much people are willing to pay to avoid aircraft noise. By trading market goods (like houses) consumers express the value they place on intangible goods, and these values can be extracted through the use of statistical techniques (such as regression analysis).

4.7.1.4 Travel Cost Method

Market and non-market goods can be complementary: purchase of market goods and services may be required to access an intangible good or service. This method has been used to value recreational sites. Visitors from different locations incur different travel costs depending on their proximity to the site. The resulting difference in visit rates provides a basis for estimating a demand curve for the site. The situation is complicated, however, by the fact that travel can have value in its own right, and that the same costs might be incurred to access more than one site [20].

4.7.1.5 Defensive Expenditure Method

This method measures the amount people spend to mitigate an unmarketed problem, such as pollution or a terrorist attack. If smog worsens, people may spend more on window cleaning. People might buy double-glazing to reduce traffic noise. This mitigation expenditure provides evidence of how individuals value a reduction in the problem. If defensive expenditure eliminates the problem, the total expenditure provides an estimate of the cost of the unwanted effect.

4.7.2 Stated Preference Techniques

Stated preference techniques rely on surveys to obtain information on how people value costs and benefits. In this case, people are simply asked about their willingness to pay. These surveys are known as contingent valuation surveys. A survey may be the only way to collect information on 'non-use value', which is the value in an individual places on a resource or activity, even though they may not directly use it or participate, either now or in the future. For example, people might be willing to preserve a wilderness area because they place value on knowing that a natural habitat exists for rare animal species. Boardman et al. [21] describe how to conduct contingent valuation surveys and outline some problems with the technique.

Choice modelling is another survey method, used when the benefits from a policy proposal have many attributes and the policy options involve different combinations of those attributes. As a rule, estimates resulting from observed behaviour in the marketplace tend to more credible that those from surveys. Observing purchasing decisions directly reveals preferences, whereas surveys elicit statements about preferences. Survey respondents may have little incentive to take the questionnaire seriously, invest the time necessary to answer it accurately, or be truthful. Participants incur little cost for inaccurate or ill-considered answers and may have an incentive to exaggerate.

4.7.3 Impact Valuations from Secondary Sources

The methods discussed above provide a set of tools for the practical valuation of impacts, but may be difficult to implement. When there are insufficient resources or expertise to conduct an original survey, 'plug-in' values may be drawn from previous studies. Frequently used plug-ins include the value of a statistical life or life year, the value of travel time savings, the cost of noise pollution and the cost of air pollution. While information from secondary sources can provide a quick, low-cost approach for obtaining the desired monetary values, it should be treated cautiously and not used without clear justification. Judgement is required to determine whether the results from a previous study are appropriate for a particular regulatory impact analysis. Estimates gleaned from secondary sources may need to be adjusted, depending on the specifics of the particular application. The accuracy and quality of the original study should be carefully scrutinised. Any technical weaknesses should be identified and taken into account. Clearly, if a study has major weaknesses, it should not be used.

4.7.4 Costs and Benefits That Cannot Be Valued in Monetary Terms

Some costs and benefits resist the allocation of monetary values. A CBA should nevertheless take into account all relevant information and should make explicit allowance for costs and benefits that cannot be valued. Estimated costs and benefits should be divided into the following three categories:

– monetised;
– quantified, but not monetised; and
– qualitative, but not quantified or monetised.

The challenge is to pay adequate attention to non-monetised impacts, but not to overplay them. For example, if a particular policy option is advocated despite the monetised benefits falling significantly short of monetised costs, the study should explain clearly why non-monetised benefits would tip the balance. CBA encourages decision-makers to make explicit the limits they place on non-monetised benefits. For example, the monetised costs of a project may far exceed the monetised benefits. Is the non-monetised benefit valuable enough to outweigh the net monetised costs? It may be reasonable to assume that the value residents place on the non-monetised benefits makes it worthwhile. However, this might not be such a reasonable assumption should those costs rise. If quantification is not possible, intangibles should, at a minimum, be described qualitatively and an evaluation made of the strengths and weaknesses of the relevant arguments for taking them into account.

4.7.5 Cost-Effectiveness Analysis

Cost-effectiveness analysis is a widely used alternative to CBA in circumstances where it is not possible to monetise the most important policy impact. It compares alternatives on the basis of the ratio of their costs and a single quantified, but not monetised, effectiveness measure, such as lives saved. It may be reasonable to use cost-effectiveness analysis if the effectiveness measure captures most of the policy's benefits.

Cost-utility analysis is a form of cost-effectiveness analysis that employs a more complex effectiveness measure, reflecting both quantity and quality. It is generally used in health care. For example, the benefit measure may be quality-adjusted life-years (QALYs), which combine the number of additional years of life and the quality of life during those years (usually measured on a scale where 1 = perfect health and 0 = death). In cost-utility analysis, the incremental costs of various policies are compared to changes in health measured by the QALYs that they produce. A similar cost-effectiveness measure is disability-adjusted-life-years (DALYs).

4.7.6 Accounting for Equity

CBA aggregates costs and benefits across individuals, and does not take into account how equally those costs and benefits are distributed. A CBA implicitly counts a monetary gain to one person as cancelling an equivalent loss to another. It therefore assumes that the unit of currency is worth the same to everyone. In other words, CBA aims to determine whether there is a net monetary gain to society as a whole, rather than explore the details of who receives the benefits or who pays the costs. This assumption separates a policy's efficiency or resource allocation effects from its equity or distributional effects (Commonwealth of Australia 2001; [22, 23]). This separation is useful, as there is no consensus on the weight attributed to equity effects. Ultimately, it is up to decision-makers to decide the trade-off between equity and efficiency. A CBA can only help to inform the decision.

CBA provides a summary of the efficiency effects of a policy. A policy can be said to improve efficiency if its NPV is positive. While efficiency is important and should be given due weight, it is only one consideration. The way in which costs and benefits are distributed across various groups, and over time, can also be important to decision-makers. While CBA cannot resolve equity issues, it can draw attention to them by quantifying the impacts of proposed policies on different groups. If the information is available, a CBA can identify potential winners and losers and the magnitude of their gains and losses. It is then up to decision-makers to decide whether distributional impacts or equity issues are important and need addressing.

A CBA clarifies the trade-offs made when comparing alternative policy proposals, such as how much income may need to be sacrificed to achieve other objectives. For example, a policy option with the largest NPV may be rejected if it has significant adverse equity impacts. The reasons should be made explicit.

4.7.7 The Social Discount Rate

When a project imposes current costs on firms and individuals, it increases the demand for capital. For example, firms may need to borrow, or forgo other investment, to cover increased compliance costs. The capital required by a project must be sourced from displaced investment, newly stimulated savings (i.e. decreased consumption) or extra foreign capital inflow. For the project to increase efficiency, the benefits must exceed those to be had from alternative uses of the capital.

The appropriate discount rate to use in a CBA should reflect the social opportunity cost of the capital needs of a regulatory measure. This is the weighted average of the return on different sources of capital. The weights are the proportion of funds drawn from each source. What is the opportunity cost of the various sources? For displaced investment, it is generally its pre-tax marginal return. When consumption is decreased, it is generally the post-tax return on savings. When foreign capital inflow is increased, it is the marginal cost to the national economy. The social discount rate is the weighted average of the rates attached to each of these sources.

Returns vary across sources because of taxes and market imperfections. For example, income tax drives a wedge between the pre-tax (investment) interest rate (the marginal return of investment in capital) and the post-tax (consumption) interest rate (the return consumers receive on their savings). The difference can be quite substantial. The appropriate social discount rate therefore depends on:

- the weighting of each source of capital; and
- anticipated rates of return for the displaced uses of the capital.

Compiling these figures is generally difficult. Nevertheless, the use of any particular social discount rate implicitly makes an assumption about these variables, and these assumptions should be made explicit.

4.7.8 Accounting for Uncertainty

There will generally be some uncertainty about the costs and benefits of any project. The degree of uncertainty will differ across proposals and affects how the costs and benefits are valued by the individuals involved [11, 24]. Once again, the appropriate approach is to use the market price of risk to determine the value individuals place on risk. A project's net benefits should be discounted by the return on private assets with similar risk characteristics.

A 7% social discount rate reflects the overall average return to private capital in the economy. It is appropriate when the project has an average level of risk (i.e. the same as the market). In general, it is reasonable to assume that the net benefits from most efficient projects contain average levels of market risk and the relevant discount rate is the weighted average market return (i.e. 7% in real terms). However, that is not the case for all projects. Proposals that are more sensitive to market returns should have a higher discount rate, while less sensitive proposals (such as DGT projects) should have a lower rate. Given the uncertainty surrounding the discount rate, it is useful to use sensitivity testing with a range of rates, using risk premiums above and below the market risk premium. The rates used in sensitivity testing represent the weighted average of riskless, market and high-risk returns.

4.7.9 Accounting for Future Generations

Using the weighted average social discount rate to discount future costs and benefits determines the efficiency effects of the policy, i.e. the current generation's aggregate willingness to pay. However, an issue arises when regulatory impacts cross generational lines (for example, when costs are borne by today's generation but the benefits are shared with or received by future generations). Some argue that a lower discount rate should be used for intergenerational discounting. As yet, there is no consensus on how to value impacts on future generations. Rather than use an arbitrarily lower discount rate, some authors suggest that the effects on future generations be considered explicitly (see [25]). One way this could be done is to supplement the CBA with a discussion of how future generations could be affected by the project.

4.7.10 Common Cost-Benefit Analysis Pitfalls

Some common pitfalls, particularly when analysing DGT projects, include the following:

4.7.10.1 Downplaying or Ignoring Non-financial Social Costs and Benefits

DGT projects differ considerably in the ease and accuracy with which the prospective costs and benefits can be quantified. Although a CBA places emphasis on valuing costs and benefits in monetary terms, it is important that the project evaluation process is not biased in favour of those proposals with impacts that are relatively easy to value. Care should be taken to ensure that monetised impacts do not overshadow other important factors in decision- making.

4.7.10.2 Double-Counting Benefits

If the costs and benefits of a DGT project have been estimated from the impact in a primary market, they should not be counted a second time as a result of consequent changes in secondary markets. For example, if a change to transport regulation results in savings in travel time to a particular group of homeowners, it would be inappropriate to add the resulting increase in their house prices to the benefits from the regulatory change (which is merely the capitalised equivalent).

4.7.10.3 Before/After Rather Than With/Without

The costs and benefits of a DGT project relate to what would have happened in the absence of the project. That is, it is necessary to compare the world without the change to the world with the change. It is inappropriate to merely calculate incremental DGT costs and DGT benefits compared to the status quo, unless no further changes would have happened in the absence of the proposal.

4.7.10.4 Using the Riskless Rate of Interest to Discount Net Benefits That Contain Market Risk

The riskless rate of interest should only be used to discount net DGT benefits that are uncorrelated with market returns or that have been converted to certainty equivalents. The certainty equivalent of a risky payment is the smallest certain payment the individual would accept in exchange. When a risky payment increases overall consumption risk for a risk-averse individual, its certainty equivalent is less than the expected value of the payment. It would overstate the NPV of a proposal to discount expected values involving market risk using a riskless rate.

4.8 Conclusion

The management of DGT risks raises numerous questions which cannot be answered with a simple 'yes' or 'no' [26]. Some of these questions include:

- What criteria should be used to decide that risk has been reduced to acceptable level?
- How to arbitrate between conflicting priorities. For example, potential casualties (in the event of an accident), the potential impact on the environment, increased employment, the expropriation of property, etc. which have an impact on multiple stakeholders?

Given these conflicting objectives, decision-making should use the most objective method possible and encourage dialogue between stakeholders. The CBA method meets these needs in the following ways:

- It makes explicit and comparable the various policy determinants and stakes in the decision, facilitating the transparency of decision-making;
- It offers a structured framework for responding to the question of economically acceptable spending limits required by legislation; and
- It allows various scenarios to be compared, and the trade-off between risk management spending and asset protection to be examined.

References

1. Australian Government (2007) Best practice regulation handbook. Australian Government, Canberra
2. Erkut E, Tjandra SA, Verter V (2007) Hazardous materials transportation. In: Barnhart C, Laporte G (eds) Handbook in OR & MS, vol 14. Chapter 9. Elsevier
3. United States Department of Transportation (http://www.dot.gov/)
4. UN (2001) UN recommendation on the transport of dangerous goods, model regulations. United Nations Economic and Social Council's Committee of Experts on the Transport of Dangerous Goods (http://www.unece.org/trans/danger/publi/unrec/12_e.html)
5. US DOT (2004a) 2002 and 2003 summary of hazardous materials transportation incidents. Research and special programs administration, Office of hazardous materials safety, 2 p
6. US DOT (1998) Hazardous materials shipments. Office of hazardous materials safety, Research and special programs administration, US Department of Transportation, Washington, DC, 12 p
7. US DOT (2004c) Hazmat summary by mode/cause: calendar year 2003. Serious incidents, printed on January 4, 2005. The office of hazardous materials safety, US Department of Transportation, Washington, DC. Available at http://hazmat.dot.gov/pubs/inc/data/2003/2003scause.pdf, 1 p
8. Responsible Care (2006) Introduction au 8ème Workshop, Transport des Produits Chimiques & Sécurité Routière, Ministère de l'Equipement et du Transport, Maroc
9. US Department of Transportation (2001) Hazardous materials (HM) Shipper Check 2001. FMCSA Federal Motor Carrier Safety Administration. March 2001. Available at http://www.fmcsa.dot.gov/safety-security/hazmat/shippercheck2001.htm
10. ICF Consulting (2003) Costs-benefit analysis of road safety improvements - Final report. Available at http://ec.europa.eu/transport/roadsafety_library/publications/icf_final_report.pdf

11. Baird BF (1989) Managerial decisions under uncertainty. Wiley, New York
12. Clemen RT (2000) Making hard decisions: an introduction to decision analysis, 2nd edn. Duxbury Thomson Learning, Belmont
13. Transports Canada (1994) Guide de l'analyse coûts-avantages à Transports Canada. Ch. 4, Détermination des options and Ch. 5, Une base de référence commune, Ottawa, Canada
14. Boardman EA, Vining A, Waters WG (1993) Cost and benefits through Bureaucratic Lenses: example of a highway project. J Policy Anal Manage 13(3):532–555
15. Williams T (1995) A classified bibliography of recent research relating to project risk management. Eur J Oper Res 85:18–35
16. Hanley N, Splash CL (1993) Cost-benefit analysis and the environment. Edward Elgar Publishing Ltd, Brokfield
17. Burgess DF (1981) The social discount rate for Canada: theory and evidence. Analyse de Politiques 7(3):383–394
18. Watson K (1995) Sensitivity analysis in outcome evaluations: a research and practice note. Revue canadienne d'évaluation de programme 10(2):113–122
19. Cropper M, Evans W, Berardi S, Ducla-Soares M, Portney P (1992) The determinants of pesticide regulation: a statistical analysis of EPA decision making. J Pol Econ 100:175–197
20. Culley EK, Donkor F (1993) Évaluation des économies de temps de déplacement des passagers. Transports Canada, Ottawa
21. Boardman EA, Greenberg DH, Vining AR, Weimer DL (2006) Cost-benefit analysis: concepts and practice, 3rd edn. Pearson Prentice Hall, Upper Saddle River
22. Commonwealth of Australia (2006) Handbook of cost benefit analysis, January. Financial management reference material no. 6
23. Nwaneri VC (1970) Equity in cost-benefit analysis: a case study of the third London airport. J Transp Econ Policy 4(3):235–254
24. Morgan MG (1990) Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, Cambridge
25. Australian Government, the Department of Finance and Deregulation, Best Practice Regulation Handbook, Appendix E. Cost-benefit analysis. Available at http://www.finance.gov.au/obpr/proposal/handbook/appendix-E-cost-benefit-analysis.html
26. Carter EE, (1972) What are the risks in risk analysis? Harvard Business Review, July–August, pp 72–82

Chapter 5
Mobile Object Framework and Fuzzy Graph Modelling to Boost HazMat Telegeomonitoring

Azedine Boulmakoul and Adil El Bouziri

Abstract This chapter describes a real-time mobile information system which responds to the vulnerabilities of the transportation of dangerous goods. It describes the integration of various software components to boost HazMat (hazardous material) telegeomonitoring and presents a real-time, object-oriented model. It also explains how to integrate a spatial decision support system and how the result can be exploited by computer science. It focuses specifically on two particularly helpful contributions to improvements in the safe transport of HazMat. The first is the problem of finding the fuzzy shortest path. It discusses an original solution using an algebraic dioïd structure to solve the k-best fuzzy shortest paths problem. The second deals with mobile object modelling. A mobile object data model and a mobile query language with a powerful set of spatiotemporal predicates are outlined. This approach is based on the comprehensive framework of abstract data types. It represents the first steps in constructing innovative moving object databases.

5.1 Introduction

Thousands of tons of petroleum, toxic, chemical, corrosive, flammable, and radioactive materials are transported every day. These substances may pose a threat to public safety or the environment during transportation. Previously, such problems could not be studied in real-time. But now, thanks to telecommunication technologies and intelligent real-time sensors, risk and environmental monitoring can be carried out in real-time, although it requires a combination of various distributed data sources and heterogeneous technologies. The first challenge is to find a

A. Boulmakoul (✉) • A. El Bouziri
Computer Sciences Department, Mohammedia Faculty of Sciences and Technology (FSTM), Mohammedia, Morocco
e-mail: a_elbouziri@yahoo.fr; boulmakoul2000@yahoo.fr

E. Garbolino et al. (eds.), *Transport of Dangerous Goods*, NATO Science for Peace and Security Series C: Environmental Security, DOI 10.1007/978-94-007-2684-0_5,

comprehensive framework for an open system architecture with wide interoperability designated to provide risk and environmental monitoring and offer specific services to mobile users.

In this chapter, we describe a mobile information system architecture with real-time capabilities for hazardous materials transportation and environmental monitoring. This system has grown out of previous work by the authors and others [2, 3, 6–9]. It illustrates the integration of various software components and proposes a real-time object-oriented modelling technique using the modelling languages UML 2 [44] and RT-UML [43].

Advances in computer science have dramatically enhanced this type of system. In particular, innovative moving objects' databases have been integrated into information systems for the management of risk, and routing of HazMat transportation. This has been successfully combined with positioning and geographic information systems to increase overall performance and to achieve the secure transportation of hazardous materials [2].

This chapter illustrates how to integrate a Spatial Decision Support System (SDSS) into an overall system for the routing and monitoring of hazardous materials transportation. The design of the SDSS was the result of a project led by Boulmakoul et al. [4–9]. The SDSS was initially based on both geographical information system (GIS) and decision support system technologies. It has since integrated GPS capabilities and fuzzy routing algorithms in fuzzy graphs that capture the concept of risk. The SDSS implements algorithms and algebraic structures developed by [4, 5]. An original solution is found for the k-best fuzzy shortest paths problem. A fuzzy approach is proposed. This uses fuzzy data to model risk en route. The approach is based on some basic concepts of multicriteria analysis and fuzzy set theories. The technique produces a fuzzy risk graph that models both the system of transportation and the concept of the risk of accident. Furthermore this fuzzification of risk enables the results from fuzzy shortest path-finding problems in fuzzy graphs to be applied to lesser risks in the transport of HazMat.

This chapter specifically focuses on an object-oriented modelling of the proposed system, the problem of fuzzy shortest path-finding, and a moving object data model. These are the first steps in the construction of innovative moving object databases. These concepts contribute to achieving useful advances in the safe transport of HazMat. This introduction gives a preliminary overview.

5.1.1 Fuzzy Shortest Path-Finding Problem

The fuzzy shortest path-finding problem (finding the shortest path from a specified source node to other nodes) appears in several different areas. With respect to transportation systems, the corresponding networks use fuzzy information on the arcs, which is assumed to represent transportation time, economic cost, traffic flow, etc. This information is soft and can be represented by fuzzy numbers or a fuzzy set based on fuzzy set theory [49]. The first work on solving the fuzzy shortest path problem came from Dubois and Prade [13]. However, if the search for the shortest path length in a fuzzy graph is feasible, this path does not generally correspond to a real path in the fuzzy graph in question. This exception is explained by the particular behaviour of the generalised min and max operators for the fuzzy numbers. Dubois and Prade [13] comment that the solution to the classical fuzzy shortest path problem is through the use of extended sum, and min and max operators. Floyd's algorithm and Ford's algorithm are applied to solve the problem.

Unfortunately, this approach, even though it can determine the length of a fuzzy shortest path, cannot find a fuzzy path which corresponds to this length in the fuzzy graph. This failure is a consequence of the classical extension principle of the min and max operators. This principle says that the extended min or max of fuzzy numbers may not be one of those numbers. Some approaches based on the concept of α-cut [10, 11, 30, 36] and other models based on the parametric orders [18] or relation order [42], allow the reuse the classic methods in various fuzzy graphs applications in the field of operations research.

A formulation of the fuzzy shortest path problem which does not make reference to the concept of α-cut or parametric orders has been proposed by Klein [28]. Klein's algorithm is based on multi-criteria dynamic programming, and can find a path or paths for a level of membership set by the decision-maker. However this algorithm assumes that the fuzzy graphs in question are acyclic graphs. To apply Klein's algorithm to other graphs, Klein proposed a transformation based on the following observation attributed to Lawler [35]: each graph that has no cycles of negative weight can easily be converted to a directed acyclic graph. Nevertheless the transformation procedure is NP-Hard. Hence for computations, Klein's algorithm is restricted to acyclic graphs. This current work proposes the construction of appropriate and new dioïd structures (path algebra) to solve the fuzzy shortest path problem in a fuzzy graph [20].

The dioïd concept was initially proposed by Kuntzmann [28]. Dioïds have already been used to formulate the path finding problem in traditional (not fuzzy) graphs. Solving the (classical) operations research problem consists of determining an algebraic structure based on dioïds and applying generalised algorithms such as Bellman, Ford, or A* [19]. On this subject the work of Gondran and Minoux [23] provides an excellent presentation.

In this chapter, the proposed dioïd algebraic structure is adapted to solve the k-best fuzzy shortest path problem. The result generalises Klein's work. This work starts the extension of Gondran path algebra [22, 23, 37, 38] given for the crisp case to valued fuzzy graphs.

5.1.2 Towards Mobile Object Databases

Currently, with internet-enabled mobile devices and mobile positioning we talk more and more of location based services (LBS). This refers to wireless services provided to the subscriber based on their current location [27, 33]. The location can be discovered by receiving data from the mobile phone network, or from another positioning service, such as a global positioning system (GPS). Telegeomonitoring can be defined as a new discipline characterised by positioning systems, cartography, the exchange of information between different sites and real-time spatial decision making. The development of telegeomonitoring systems combines two heterogeneous technologies: geographical information systems (GIS) and telecommunications technology [4, 34]. HazMat risk and environmental monitoring systems can be considered as examples of telegeomonitoring.

All such systems are based on mobile objects that change location either discretely or continuously through time. Therefore, LBS applications require database and application support to model and manage mobile objects in both the database and application domains and for querying the motion properties of objects. Supporting this type of spatiotemporal object (the so-called moving object) is one of the challenges addressed in this chapter. Neither spatial nor temporal databases can deal with the moving object. The composition of temporal and spatial properties of real-world objects in a unified data framework results in so-called moving object databases (MOD). The latter is able to process, manage and analyse changing spatiotemporal data. It has to deal with both moving objects and all kinds of spatiotemporal queries [24, 25, 46].

This work describes a mobile object data model using UML 2 [44] and the main elements of an extended SQL query language for representing and querying mobile objects, in particular those with point geometry moving on a transport network. This data model, based on the widely accepted OGC (Open Geospatial Consortium) specification [39] constitutes a framework that provides MOD functionality to relational object DBMS (database management systems) compatible with OpenGIS. It is principally employed in the proposed component-based modelling of the overall system. This framework constitutes the first steps in the construction of new moving object databases.

This chapter is structured into three major sections. The first proposes a generic architecture for a real-time mobile information system for HazMat telegeomonitoring. It describes the real-time object-oriented modelling of the system using UML 2 and RT-UML. It also describes components-based modelling with real-time capabilities. The second section presents a data model of a mobile object in a transportation network and outlines an extension to the SQL query language with a set of spatiotemporal operations. It represents the first steps in the construction of the moving object database. The third section focuses on the multicriteria fuzzy routing component. It presents original work that globally revises the fuzzy shortest paths problem and gives a new solution (using dioïds) to the k-best fuzzy shortest paths problem in fuzzy graphs.

5.2 Architecture and Object-Oriented Modelling

5.2.1 Architecture

Mobile users of the proposed system can, in real-time, access information related to location, such as HazMat and risk data, or the fuzzy shortest path to get to a specific destination. Figure 5.1 shows the main architectural components of a mobile

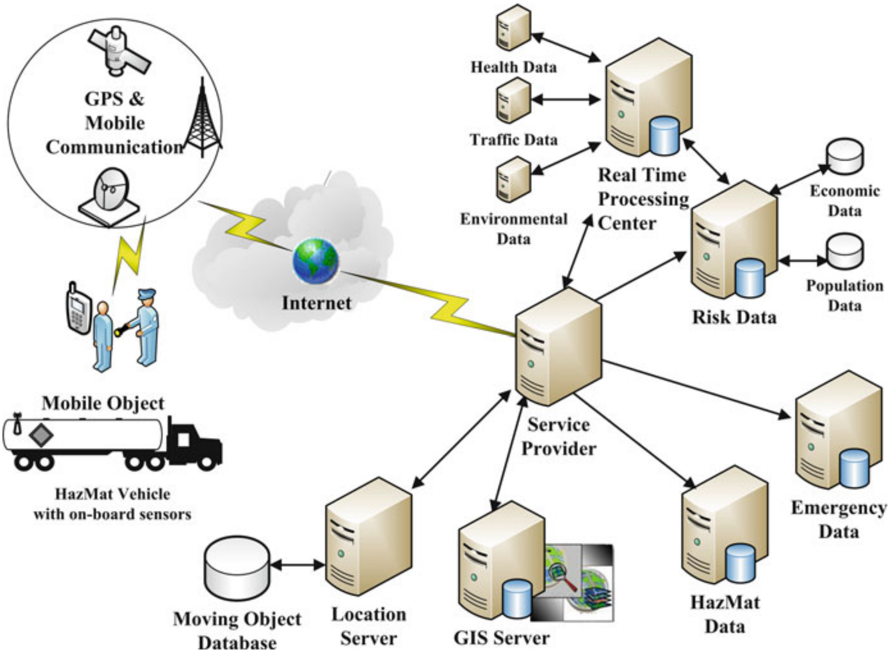


Fig. 5.1 Overall architecture

information system for HazMat telegeomonitoring. It illustrates the seamless integration of the various components and disparate technologies into one system. A global positioning system (GPS) receiver determines the current position of the mobile object (i.e. client truck or mobile sensors) and sends periodic updates of the mobile user's position to the location server. Other sensors, distributed all over the city monitor conditions at different locations, such as temperature, sound, vibration, pressure, motion or pollutants.

The Mobile Object represents the moving object (such as a truck) equipped with an embedded device such as a location detection mechanism. It periodically sends its coordinates to the location system. The Service Provider coordinates the various system components. It can provide anywhere, anytime, real-time data, maps or other services to the mobile object related to its spatial position. It uses web services [29] that can interact with any type of mobile device to deliver the requested services. In order to determine the position of mobile object, the Location Server has to communicate with database servers storing the user's location. These new moving object databases have to deal with both moving objects and all kind of spatiotemporal queries [25]. The GIS Server has a suite of tools to perform spatial operations which include geo-coding, reverse geo-coding, routing and several other services. It needs access to the GIS database to perform its functions. The RealTime Processing Centre is accessed concurrently and receives real-time data from various sources including for example, environmental or traffic data. It analyses up-to-date information and stores the processed data in the RTData Storage database.

Other components enhance the system. A short description is given below:

HazMat Data: this server enables the retrieval of descriptions of hazardous material stored in the HazMat multimedia database (designed to be user-friendly and to allow remote access). HazMat information concerns product identification, the nature of the danger, the physical and chemical properties, risk and safety instructions, etc.

Risk Data: this server provides risk information on the three main targets: population, environment and economy.

Emergency services: in case of emergency, this component facilitates decision-making. It makes it possible to estimate the radius of the affected area with soft and hard consequences. It gives the optimal deployment for emergency response units and minimises the evacuation time in the affected area, by reducing traffic flow.

5.2.2 Web Services-Based Open Architecture

The classical solution, which provides access to specific Internet services through an architecture based on a client browser and a web server has limitations: data accessibility depends on the mobile user's device, interoperability issues, and the need for remote and mobile access control. Web services technology addresses these limitations. The main advantage it offers is an open architecture, accessible by any type of client in a simple way. The server provider shown in Fig. 5.1 represents a set of components and servers. In order to offer the desired services to diverse mobile users, an open architecture based on web services and an n-tiers model is proposed. Any mobile device that supports the SOAP (Simple Object Access Protocol) protocol can access HazMat information independently of its platform, language, and above all device. On the server side, a number of web services are defined to handle HazMat data, to access real-time information and to calculate the fuzzy shortest path to a specific destination.

Figure 5.2 shows the web services-based open architecture adopted to deliver SOAP services. At the presentation level, heterogeneous mobile devices like SMS, WAP, J2ME or Windows CE clients can access services via the same web services and application server.

The logical level is responsible for all processing operations and coordination between distributed components in the overall information system.

5.2.3 UML 2 and RT-UML Specification

UML 2 represents a major revision of the Object Management Group's (OMG) Unified Modelling Language (UML). It is composed of two specifications: Infrastructure and Superstructure. The first defines the basic language constructs

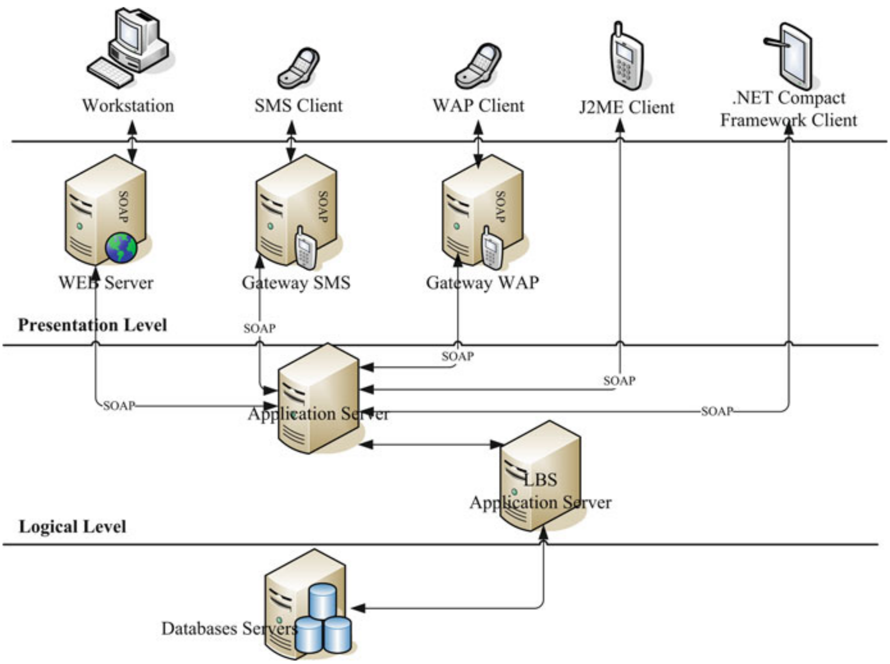


Fig. 5.2 Web services-based open architecture

required. The Superstructure specification uses the architectural foundation provided by the Infrastructure specification, and defines the user level constructs required [44]. The major improvements to UML 2.0 include support for component-based development via composite structures, and integration of action semantics with behavioural constructs [29]. In addition, UML 2 offers some features that support real-time operations: concurrency modelling, timing constraints, etc.

The RT-UML specification is the UML profile for real-time modelling, formally called Schedulability, Performance and Time (UML/SPT) [43] was adopted by the OMG in 2002. It has increased interest in the use of object-oriented technology and UML in particular to model and build real-time systems. This profile is designed to add standard real-time extensions to UML in order to facilitate development with real-time modelling notation. RT-UML is a framework for modelling quality of service, resource, time and concurrency concepts. It provides the user (modeller) with a set of stereotypes and tagged values in order to annotate the UML model. Quantitative analysis (schedulability and performance analysis) can then be applied to these (predictive) UML models.

The structure of the profile is modularised to allow users to only choose the elements they need. The profile defines a basic framework which is a set of subprofiles that represent the general resource modelling framework. There is a sub-profile used for schedulability analysis of systems derived from the basic framework. This model is more interesting in real-time systems where the question

Table 5.1 SPT common stereotypes for schedulability analysis

Stereotype	Real-time Concept	UML Model Element
SAsituation	Real-time situation	Collaboration, Sequence diagrams
SAtrigger	Event	Message, Stimulus
SAresponse	Response	Method, Action
SAaction	Action	Method, Stimulus, Action
SAschedRes	Task, Thread	Instance, Object, Node
SAresource	Resource	Instance, Class, Node
SAengine	CPU, Processor	Object, Class, Node

of when a response to an event occurs is very important for the correct behaviour of the system. The schedulability sub-profile focuses on how to annotate the model in ways that allow a wide variety of schedulability techniques to be applied. The metamodel defined in this sub-profile includes the main concepts involved in the schedulability analysis: the execution engine, threads (task or process), shared resources, external events and the response of the system to external events. To represent these concepts in UML, a set of stereotypes and their associated tagged values are defined in a schedulability sub-profile. Table 5.1 presents a sample of stereotypes. The application of these stereotypes will be illustrated in the object-oriented model described in next section.

5.2.4 An Object-Oriented Model of the Overall System

Figure 5.3 proposes an object-oriented model using UML 2 which follows, as far as possible, the definition found in the open GIS specification [40, 41]. It represents a communication diagram that emphasises the structural organisation of the objects that exchange messages.

The following steps describe the normal scenario that provides the requested service: the mobile object requests, via a wireless network, real-time services from the service provider (SP). After identification of the subscriber and the requested service, the SP parses, evaluates and interprets this spatiotemporal request and calls various components in order to provide mobile geographic services to the mobile client. The SP formats the request data sent to the location server for the positions of all mobile objects mentioned in the request from the mobile object. The location server retrieves the relevant positioning data from the moving object database and sends its response to the SP. The latter opens a connection to a GIS server to send a map request or to search for spatial objects in whose area of influence the mobile object is found. The SP then interacts with RealTimeData, EmergencyData, HazMatData or RiskData object to retrieve the appropriate information. Finally, the SP sends its response to the mobile object describing the service. A mobile application installed on the mobile terminal parses the response. For example, it allows the subscriber to view the processed map with the position of services plotted, and to interact with other functions.

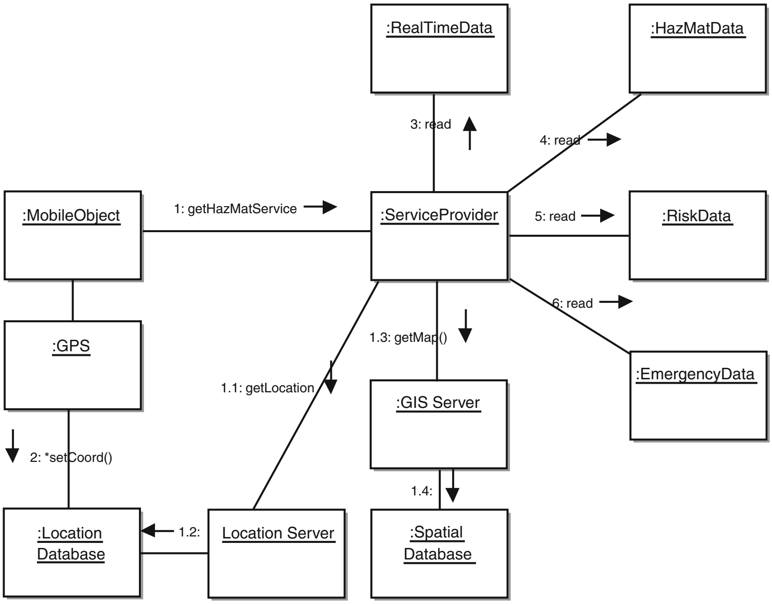


Fig. 5.3 Communication diagram of the proposed system

The communication diagram can be annotated with several stereotypes from the schedulability sub-profile to represent real time considerations of different scenarios within the system. For example, ServiceProvider, HazMatData, EmergencyData, RiskData and TrafficDataGatherer classes in Fig. 5.4 are associated with the "SAschedRes" stereotype. Instances of these classes execute concurrently in the application context. The execution flow of the "SAschedRes" stereotype is identified as a scenario which is started after an activation message stereotyped with "SAtrigger". During this execution, many actions stereotyped by "SAaction" with a specified priority (SAaction.SApriority) may be executed in response, for example, to a method call. The RTduration tag indicates the total duration of action.

The basic structure of the class scenario is characterised by the "SAresponse" stereotype and executes periodically following the event trigger associated with the RTat tag. In addition, the RealTimeData class stereotyped with "SAresource" is a protected resource that is accessed concurrently using mutual exclusion mechanisms. The SAcapacity = 1 means that one element can simultaneously access an instance of this class.

5.2.5 Modelling the Overall System with UML 2

Since the proposed system is complex and necessitates the integration of various components, it is preferable to model it using UML 2 [1, 29, 44]. UML 2 is an

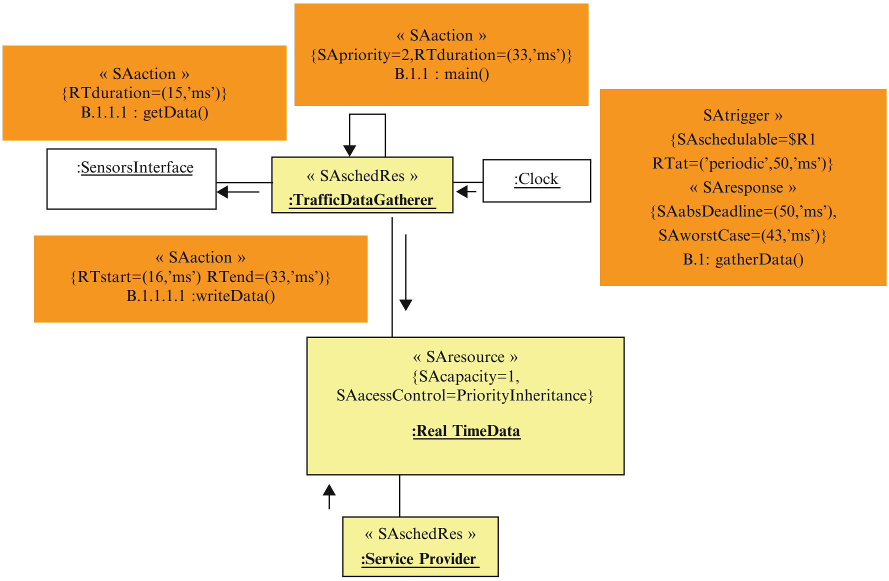


Fig. 5.4 Real-time annotations with RT-UML

improvement over UML 1.x, as it now addresses system structures. It is one of the main improvements in UML 2 which supports component-based development via composite structures. A component is a modular unit with well-defined interfaces. Component interfaces are classified as either 'provided' or 'required'. Provided interfaces have formal service contract which defines what the component provides to other components, while required interfaces are those services it requires from other components in its environment to operate properly.

Figure 5.5 proposes an object-oriented model with UML 2 [44] following the definition found in the open GIS specification [40, 41]. It shows the compositional structure of components. The wiring between components is represented by assembly connectors between provided and required interfaces. This component-based structure aims to hierarchically decompose the complex system into smaller subsystems and then connect these sub-systems together. Any part of the modelling system can be reused in many other contexts. In this component-based modelling, founded on the data model of the mobile object, two components are created relative to the mobile object: one server-side module and another for the client. The client component has some additional classifiers (classes or components) and interfaces in order to deal with location capture and to calculate uncertainty.

The SDSS component plays an important role in this system. It represents a decision-making unit that enables risk analysis through the simulation of scenarios. At this level the decision system uses fuzzy routing algorithms in fuzzy graphs which capture the concept of risk. It provides simulations by analysing the accident scenario impact on the tree main targets: population, environment, and the economy.

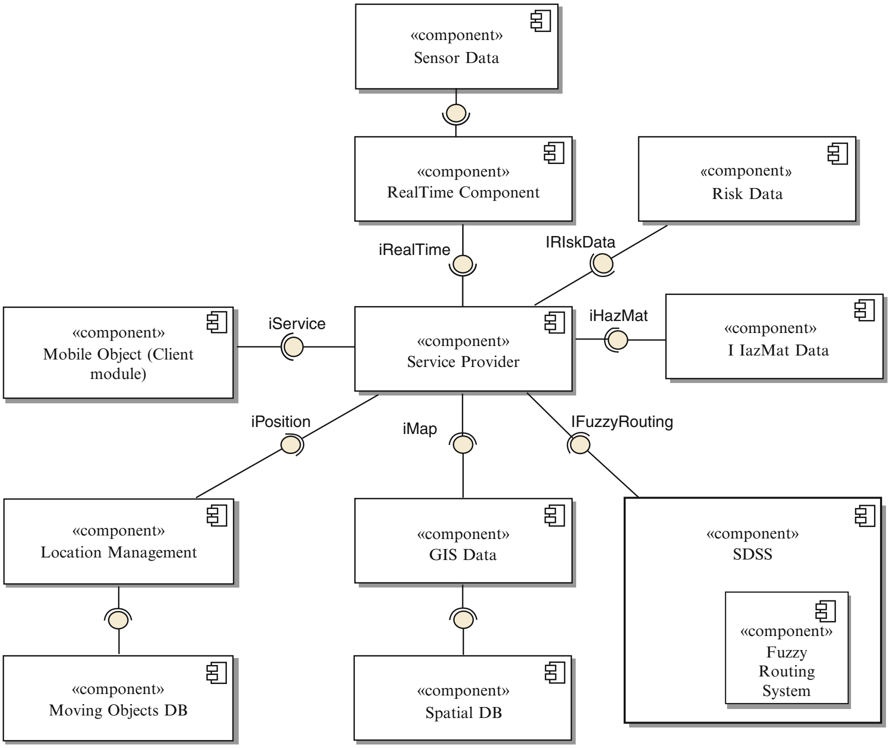


Fig. 5.5 Compositional structure of the overall system

5.3 Toward Moving Object Databases

This section proposes a data model for mobile objects and constitutes a framework for the development of location-based services and applications. This object-oriented model is extensible using UML 2 formalisms. It represents some first steps in the construction of new moving object databases.

Modelling is based on several specifications [26, 39, 47] which relate to the spatial model, the temporal model and that of the transportation network. In the same way, the model is based on research on spatiotemporal data, especially that related to mobile objects, moving objects databases, and spatiotemporal queries [16, 17, 24, 25, 31, 48]. This model defines a set of types and operations able to deal with mobile objects and for which is necessary to define a new query language for mobile data. This new query language is an extension to SQL 3; it enables the formulation of spatiotemporal queries and the interrogation of mobile data in an object-relational database.

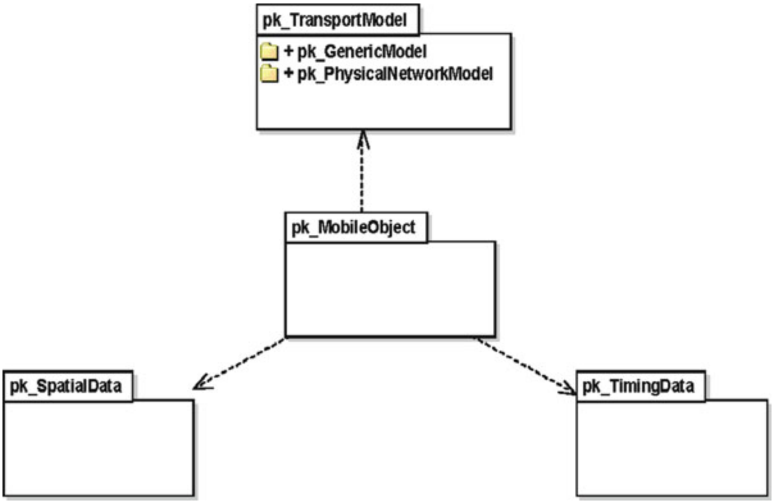


Fig. 5.6 Package diagram of the mobile object data model

5.3.1 Mobile Object Modelling

5.3.1.1 The Global Model

The global object-oriented modelling of mobile objects is presented in Fig. 5.6. The model is represented by spatiotemporal classes with mobility aspects. It is a result of our work on mobile object modelling and location-based services (LBS). The LBS are concerned with mobile point objects, i.e. objects with zero extent that change their location over a predefined network infrastructure. Emphasis is placed on modelling the mobile point object and its relationships with the main classes representing the multimodal transportation network. The trajectory of the mobile object is a polyline in three-dimensional space (two-dimensional space and time). The mobile object does not move in straight lines at constant speed. As an approximation of its motion, its trajectory is represented as a sequence of points (xi,yi,ti). The number of points along the trajectory is proportional to the accuracy of the approximation. Additional parameters (e.g. type of motion) have to be added for every point.

The package diagram will be discussed first. The proposed model contains several components that enable the definition of a set of classes and associations necessary to represent the aspects essential to mobile objects models. This diagram is composed of several packages indicated by: pk_MobileObject, pk_SpatialData, pk_TransportModel and pk_TimingData.

The *pk_MobileObject* package is the core of the mobile object data model. It contains classes representing mobility data and the main operations essential for the interrogation of mobile data. The *pk_TransportModel* package models the

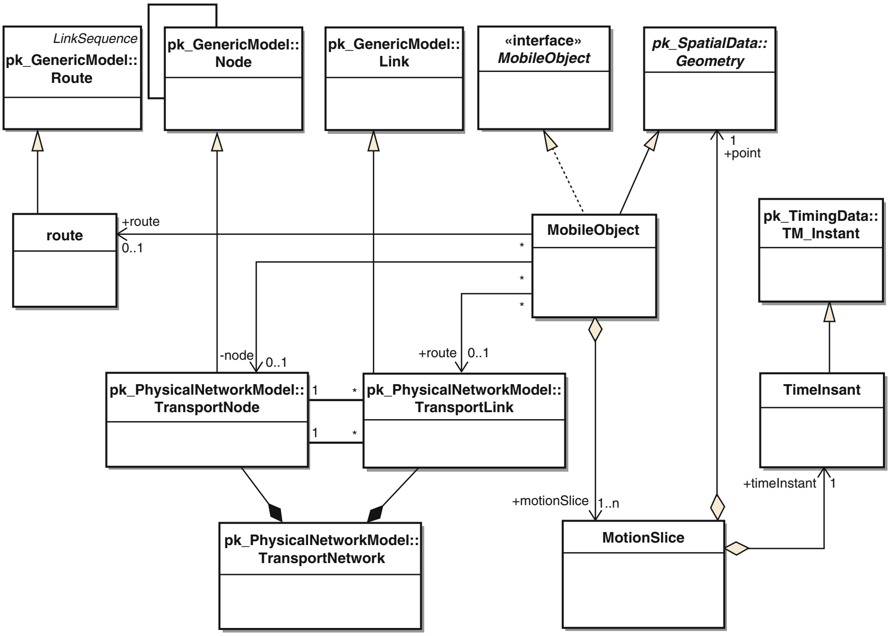


Fig. 5.7 Package diagram of the mobile object data model

transportation network. It consists of two sub-packages: *pk_GenericModel* and *pk_PhysicalNetworkModel*. The first sub-package describes an abstract model of the transport network; it can be instantiated by any real-life structure. The second sub-package models the physical transportation network by identifying the principal classes which represent the structure of the physical transportation network.

The *MobileObject* class (see Fig. 5.7) of the *pk_MobileObject* package represents the structure of the mobile object and its mobility characteristics. This class represents all the attributes and operations which enable interactions with the mobile object. As, in this context, the mobile object continuously changes location over a predefined network infrastructure, it is necessary to represent the associations which connect the *MobileObject* class with the entities that constitute the transportation network. The *TransportLink* class, which can represent for example at the functional level, a section of a road, is added into the model. The *TransportNode* class represents intersections such as crossroads. Any *TransportLink* must start with, and terminate in a *TransportNode*. The orientation of *TransportNodes* does not necessarily refer to the direction of the traffic flow, but is interpreted as an arbitrary orientation. An optional attribute 'driving direction' may be used to specify a direction.

Another important entity is clarified in this package. It is represented by the class *route*. This depends on the physical infrastructure and represents the route taken by a mobile object in a transport network. On the other hand, the class *Route* of the *pk_GenericModel* sub-package is an abstract concept independent of any type of

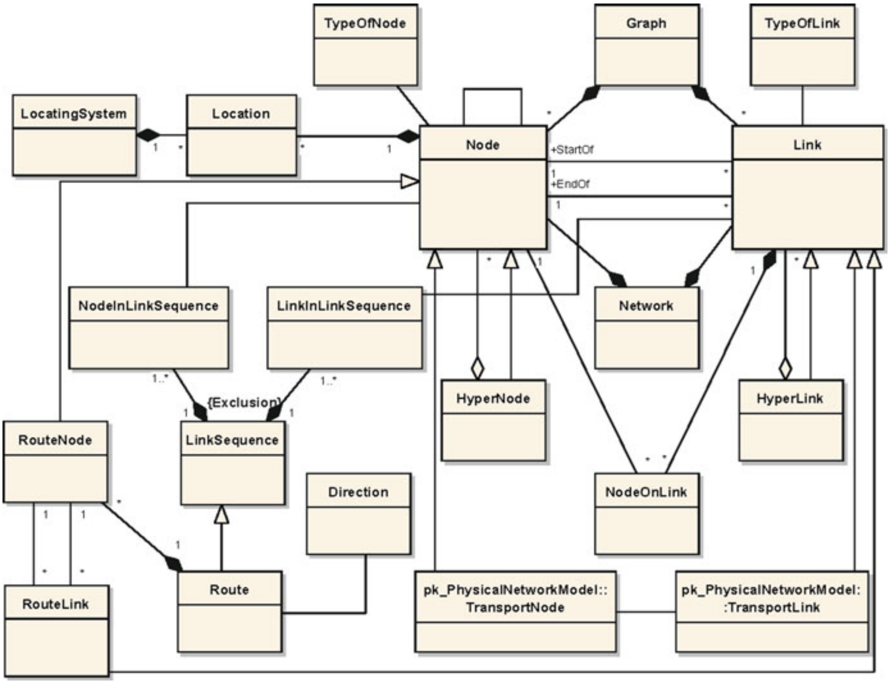


Fig. 5.8 Generic model of a multimodal transportation network

infrastructure. A more detailed overview of the transportation network modelling is given in next section.

The *MobileObject* class inherits the properties of the abstract class *Geometry* defined in the specification *simple OGC features* [39]. The class containing data on the movement of mobile objects is named *MotionSilce.*

The expression of time (representation of a valid time) is carried out via the classes *TM_Instant* and *TM_Period* defined in the ISO 19108 Temporal Schema Standard [26].

5.3.1.2 Multimodal Transportation Network Model

One of the most important aspects of information systems is the representation of the transport networks in which the services operate. This representation describes transport network objects using simplified and conventional topological entities: nodes and links. Specific roles are assigned to these simple elements according to the functional purpose of the description.

The multimodal transportation network is modelled as an oriented graph, whose fundamental elements are nodes and links (see Fig. 5.8). The model is in accordance with the Transmodel specification [47]. The spatial network model begins by focusing specifically on the definitions and semantics of these two entity types and the relationships between them. A *node* (for generic topology) is the smallest identified location in space. It represents a zero-dimensional entity in the network. It can play many different roles in the transportation network (i.e. it is not just a location in space although it can mark the location of bus stops, or parking places). Between two nodes of any type, a *link* may be defined. This represents a 1-dimensional connection between nodes. All links must have a limiting node at each end. Moreover, as the network structures used by different functions may be subject to different conditions and constraints, in some structures, the ordered connection between two nodes may have to be unique. Each node is functionally classified as being of one or more types. A particular type of node is defined as an entity to describe the common role played by a number of nodes. The link entity similarly expresses the various functional roles of a link. It is often necessary to specifically define nodes located on a link of a certain type. Each *node on link* is identified by the *link* it is located on, and by its order on that *link*.

The model also includes *hypernode* and *hyperlink* entities. A *hypernode* is a node composed of one or more nodes. For example, if a node is a station for a single transportation mode, a hypernode is an intermodal station, i.e. a place where people can enter or leave the transportation network or change their mode of transport. A *hyperlink* is a link connecting two *hypernodes*, and it is composed of one or more links.

Figure 5.8 outlines a logical view of the transportation network. The node, the link and the relationships between them are considered as a generic structural pattern which specifies many the specific structures in a transportation network.

Certain types of nodes are important enough to be additionally represented by a separate entity, such as the *route node* entity which represents a route entity. The *Route* class represents an abstract concept. Its purpose is to describe a path independent of the infrastructure pattern. This class represents a conventional way of describing a path through the network. A *route* is composed of nodes and links specifically defined for that purpose. This sequence of nodes and links must be built in such a way that identifies a path without any ambiguity. In most cases, such sequences should be simple, in order to be recognised by the data system and users. The definition of a *route* uses *route nodes*, which are nodes dedicated to the definition of regular service paths. A *route node* may be a point along the route, or an end point. The *route nodes* should identify a route without ambiguity. The definition of a route also involves *route links*; the links defined between two *route nodes*.

The physical transportation network model is represented by the *pk_PhysicalTransportModel* sub-package. The *physical network* describes the infrastructure of the transport network. It is composed of two types of network: a road network and a rail network, in which transport services run. The basic generic entities of a physical network are the *transport link* and *transport node*. All physical nodes (i.e. *road node* and *rail node*) of the physical network are sub-types of the *transport node*; and all the physical arcs (*road link, rail road* etc.) are sub-types of the *transport link*.

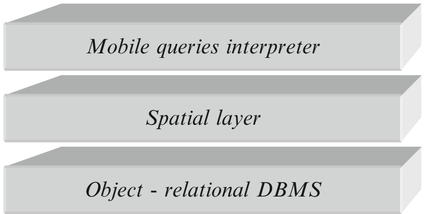


Fig. 5.9 Suggested architecture

5.3.2 Extending the SQL Query Language with Spatiotemporal Predicates and Operations

This sub-section describes our work on the mobile objects query language. We propose an extension to the SQL 3 query language by adding new operations specific to mobility. These new language elements are rich in predicates and describe the specific operations necessary for the presentation and interrogation of mobile data. A set of these operations is defined in the work of Güting et al. [25] whose approach is based on abstract data types in order to define new temporal types. The formulation of this language takes into account criteria related to the position of the subscriber (e.g. the GPS position) and the characteristics of its movement. The proposed architecture for implementation in an object-relational DBMS is shown in Fig. 5.9.

5.3.2.1 New Abstract Data Types

Some data model operations, in particular those of the *MobileObject* class use the concept of spatiotemporal data types introduced by the following studies [16, 17]. These authors introduce the notion of the temporal object by observing that any entity which evolves over time can be expressed as a function of time. The spatiotemporal objects are seen as particular instances of temporal objects. The system is based on a set of types including the basic data types (int, real, string, bool). Instances of these data types have static values which do not change over time. A constructor named *moving* is introduced to create data types whose values are dynamic, *moving* (α). α is a basic type or a spatial data type like *point* and *region*.

The value of the new type (α) varies over time. The notion of time used here is a valid time. For example, the distance between two airplanes can be expressed by $moving(real)$. The location of the moving point in the euclidean plane over time is called *moving(point)* and is noted as *mpoint*.

Figure 5.10 represents a moving point in three-dimensional space (x, y, t). Similarly, a region which evolves over time (e.g. the surface of fire) is a region which can move. It is termed an evolving or moving region, and is noted as *mregion*.

A mobile object moving in a transport network is regarded as a mobile point. Based on the 9-intersection model of Egenhofer [15] and the definition of new spatiotemporal data types, new time-dependent topological relations can be defined.

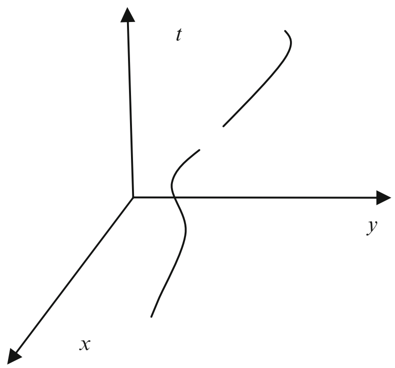


Fig. 5.10 Representation of mobile point

For example, two objects can be "disjoint" at a certain time as they can be in intersection in another. Therefore, a new set of relations were defined for spatiotemporal data types (a temporal version of spatial predicates). This mechanism is called *temporal lifting*. The same transformation can be applied to obtain other operations which deal with mobile objects. In relation to a mobile object these new operations return numbers which change over time, or a boolean. These moving data types will be named respectively *mreal* and *mbool*. They are essential when defining operations on mobile objects.

For example, the spatial predicate *inside* is applied to spatial types {point, region} and it returns a boolean. By applying the *temporal lifting* transformation, the predicate will be applicable to the spatiotemporal data types *mpoint* and *mregion* and return a time-changing boolean (*mbool*). The new operation corresponding to the predicate *inside* is noted as *Inside*. Its signature is given below:

$$\begin{aligned} &mpoint \times point &&\rightarrow mbool \\ &mpoint \times mpoint &&\rightarrow mbool \\ &mpoint \times region &&\rightarrow mbool \end{aligned}$$

Some examples of other operations include:

Trajectory: $mpoint \rightarrow line$; this operation gives the trajectory of the mobile object,

Distance: $mpoint \times mpoint \rightarrow mreal$; this operation gives the distance between two mobile objects.

5.3.2.2 Proposed Operations of the Data Model

The data model can now be completed by adding the spatiotemporal operations necessary to handle mobile objects. The new moving abstract data types which are essential for the data model are: *mreal*, *mbool* and *mpoint*.

We define two additional operations. The *MObjectAt* operation returns the *point* object at a specific time instant. The *MObjectPeriod* operation has a period parameter

Table 5.2 Operations of the mobile object class

Operations	atInstant(): real Lifespan(): Timestamp GetRoute(): LineString Distance(): mnumber
Spatial operations	disjoint(): bool touch(): bool within(): bool
Spatiotemporal operations	Disjoint(OtherGeo: Geometry): mbool Touch(OtherGeo: Geometry): mbool Within(OtherGeo: Geometry):mbool
Spatiotemporal predicates	DISJOINT(OtherGeo: Geometry): bool TOUCH(OtherGeo: Geometry): bool WITHIN(OtherGeo: Geometry):bool
Specific spatiotemporal predicates	Enter(OtherGeo: Geometry): bool Leave(OtherGeo: Geometry): bool Cross(OtherGeo: Geometry): bool

and restricts the sequence of motion of the mobile object according to the specific time object.

Table 5.2 presents a classification of some operations used in the data model. In fact, this model redefines a number of operations proposed in the simple feature specification of OGC by using the formalism presented above. For example, the operation *distance* defined in this specification has only non-temporal arguments. Therefore, we define in the mobile object data model a new function, *Distance*. This corresponding lift version works on mobile objects and returns a temporal value as result (*mreal*). By using the operation *atInstant ()*, the value at a specific instant can be given.

Furthermore, spatiotemporal relations can be defined in ways that correspond to those already well-known for spatial relations. These new relations deal with spatiotemporal data types. Spatiotemporal predicates are added according to the definition given by Erwig [17]; they have to return a boolean. These specific relations are to establish whether a mobile object remains (throughout its lifespan) in the defined area. These simple spatiotemporal predicates are noted here in capital letters (e.g. INSIDE). The trajectory of a mobile object can interact with other spatial objects belonging to the urban infrastructure and other specific relations can be added to the model (e.g. *Enter*, *Cross*, *Leave* and *Bypass*).

In fact, these relations represent more complex spatiotemporal predicates. In the work of Erwig [17] they correspond to sequences of simple spatiotemporal predicates. For example, the predicate *Enter* is a sequence of predicates: DISJOINT, TOUCH and INSIDE. These complex spatiotemporal relations are very useful in the data model as they enable the formulation of more complex mobile queries. They make it possible to determine whether a mobile object did or did not cross a specific region. For example, a mobile object enters a polygonal area during a given time period. If it was outside the polygon at the beginning of the period (Disjoint),

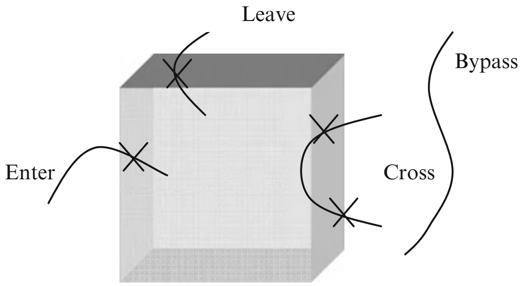


Fig. 5.11 Visual representation of specific spatiotemporal predicates

at a certain instant it would be at the border of the polygon (Touches) and within the region for the rest of the time period (Within). Figure 5.11 gives a visual representation of these specific spatiotemporal predicates.

5.3.2.3 Implementation of the Model Within an Object-Relational DBMS

The implementation of the data model in an object-relational database is based on the use of abstract data types and relations within the DBMS. Spatiotemporal data types can be described using CREATE TYPE. SQL 3 introduces such object extensions to standard SQL. These extensions are supported by DBMS such as Oracle and PostgreSQL.

CREATE TYPE MotionSlice_type AS OBJECT
(MS_ID number, position Point, instant DATE);
CREATE TYPE Motion_type AS OBJECT OF MotionSlice_type;
CREATE TYPE MobilePoint_type AS OBJECT
(MP_ID number, motion Motion_type, etc.);

Example 5.1

Find all taxicabs that are now in "Peace" street.
Select t.id, t.location.MPointAt(Now)
From Taxicab t, Street s
Where s.name = "Peace" and (t.location.MPointAt(now).within(s.type_geo))

Example 5.2

Find people who entered the "iris" region between t1 and t2.
Select p.id, p.name, p.location.MPointAt(Now)
From People p, Region r
Where r.name = "iris" and (p.location.MpointPeriod(TimePeriod(t1,t2)).Enters(r.type_geo))
These examples use the following set of object-relational tables:
Taxicab (id number, name varchar2(35), color varchar2(10), location MobilePoint_type);
Street (id number, name varchar2(30), type_geo LineString);
People (id number, name varchar2(35), location MobilePoint_type)
Region (id number, name varchar2(35), type_geo Polygon)

5.4 Multicriteria Fuzzy Routing Component

This section presents the work that globally revises the fuzzy shortest paths problem and describes an original solution to the fuzzy path problem using dioïds. The main contribution of this work is the construction of adequate and new dioïd structure to solve a fuzzy graph path-finding problem. This algebraic structure is specifically adapted to solve the problem of the k-best fuzzy shortest paths. This work outlines a method for extending Gondran's [21] and Minoux's paths algebra results [37] to fuzzy graphs.

5.4.1 Dioïds and the Shortest Path Problem

The dioïd concept was initially proposed by Kuntzmann [32] to designate an algebraic structure composed of a set *S* endowed with two internal laws denoted $\oplus$ and $\otimes$. The dioïd structure was transported to matrix algebra to generalise the theory's results. The definition of the concept is as follows:

Dioïd Definition [22, 32]

A dioïd is a triplet (S, $\oplus$, $\otimes$) made up of the following elements:
S is a set which has two elements ε and e,
$\oplus$ is an associative and commutative internal law of composition,
$\otimes$ is an associative internal law of composition,

Such that:

$\otimes$ is distributive compared to $\oplus$ on the right and on the left,
ε is the neutral element for $\oplus$ and absorbing for $\otimes$,
e is the neutral element for $\otimes$.

This dioïd is known as commutative if the $\otimes$ law is commutative. In addition, in a dioïd there is an order relation brought about by the $\oplus$ law, if the semiring is not taken into account.

Generalised Algorithms for the Shortest Path Problem

Let us consider a directed graph G = (X, A), where nodes of the set X are numbered *1, 2, … n*, and in which each arc (i, j) of the set A is given a value $a_{ij} \in S$, where *S* is a dioïd structure set. Consider a node $1 \in$ X as an origin.

We search the lengths π(j) ($j = 1 \ldots n$) of the shortest paths between node *1* and the other nodes *j* of the graph. In the case of a graph without p-absorbing cycles, the general algorithm (Algorithm 5.1) is as follows:

Γ is the function successor of the graph.
Γ is the function successor of the graph.

$$(\alpha)\, \pi(1) = e, \pi(i) = a_{1i} \text{ for } i \geq 2$$

$$(\beta) \text{ at step k, do (for } i = 1 \text{ to n)}:$$

$$\pi(1) \leftarrow \underset{j \in \Gamma^{-1}(1)}{\oplus} \left(\pi(j) \otimes a_{j1}\right) \oplus e \qquad \text{Algorithm (5.1)}$$

$$\pi(i) \leftarrow \underset{j \in \Gamma^{-1}(i)}{\oplus} \left(\pi(j) \otimes a_{ji}\right) \text{pour } i \geq 2$$

(χ) Repeat (β) until stabilisation of $\pi(i)$.

In the case of the classical shortest path ($S = R^{+} \cup \{+\infty\}$, $\oplus$ = min, $\otimes$ = +), this generalised algorithm corresponds to Ford's algorithm. For a graph without cycles, the generalised algorithm becomes simple:

$$\begin{cases} \pi(n) = e \\ \pi(i) = \underset{j \in \Gamma(i)}{\oplus} \left(\pi(j) \otimes a_{ij}\right) \end{cases} \qquad \text{Algorithm (5.2)}$$

In the case of the shortest path problem in a graph without cycles, the algorithm (Algorithm 5.2) corresponds to the optimality equation of dynamic programming (or the generalised algorithm of Bellman).

5.4.2 *General Concept of Fuzzy Sets*

Fuzzy Set

Let Ω be a classical set, called the universe. We call fuzzy set Ω the set of pairs $\{(x, \mu(x)), x \in \Omega\}$ where μ is an application in [0,1]. The fuzzy set is a generalization of the classical set concept, for which the values of μ are in {0,1}. We denote the characteristic function of fuzzy set A as follows:

$$A = \mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + \ldots + \mu_A(x_n)/x_n \text{ for } \Omega = \{x_1, \ldots., x_n\} \text{ where}$$

$$A = \int_{\Omega} \mu(x)/x \text{ when } \Omega \text{ is not finite.}$$

The Extension Principle [14]

Let $\Omega = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n$ be the Cartesian product of n universes. Let $\Omega = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n$ be the Cartesian product of n universes. Let $A_1, A_2, \ldots, A_n$ be the fuzzy sets in $\Omega_1, \Omega_2, \ldots, \Omega_n$ respectively. Given that φ is a mapping from Ω to a universe Ξ, where $y = \varphi(x_1, x_2, \ldots x_n)$, the extension principle allows us to define a fuzzy set B in Ξ by $B = \{(y, \mu_B(y) \mid y = \varphi(x_1, x_2, \ldots, x_n), (x_1, x_2, \ldots, x_n) \in \Omega\}$

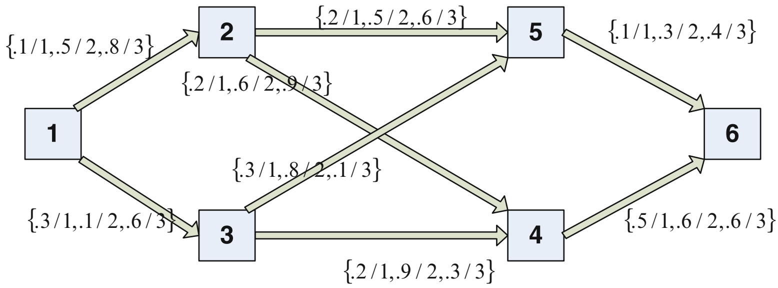


Fig. 5.12 Fuzzy directed graph without cycles

$$\mu_B(y) = \begin{cases} \sup\limits_{(x_1, x_2, \ldots, x_n) \in \varphi^{-1}(y)} \min\left(\mu_{A_1}(x_1), \mu_{A_2}(x_2), \ldots, \mu_{A_n}(x_n)\right) \text{ si } \varphi^{-1}(y) \neq \emptyset \\ 0 \text{ if not} \end{cases}$$

The extension principle is used to generalise the different classical operators to the fuzzy context. Other extensions are described in [13, 14].

The Fuzzy Graph

Fuzzy graph modelling has been applied to various problems [10, 11, 36, 45]. The fuzziness is generally introduced into a graph through the capacity of the arcs, the length of the arcs or node restrictions. We limit this work to the shortest path problem for which the graph has valuations defined by fuzzy sets of discrete D* defined as $D^* = N \cup \{+\infty\}$ where N corresponds to set of natural numbers. Elements whose membership of fuzzy sets is null, can be omitted.

We consider the length of each arc and the length of any path as fuzzy paths. In this case, each arc corresponds to a fuzzy set which indicates its valuation. In the example shown by Fig. 5.12, only those elements having a non-null grade of membership are explicitly represented in fuzzy sets.

5.4.3 Dioïds and the Fuzzy Shortest Path Problem

In this sub-section, we propose a new dioïd structure to solve the problem of the k-best fuzzy shortest paths.

Dioïd of K-Best Fuzzy Shortest Paths Problem

We will formulate the problem of k-best fuzzy shortest paths problem using the algebra of the dioïds.

For each arc (i,j) of the fuzzy graph, we associate a valuation $\tilde{\sigma}_{ij}$ defined by a fuzzy sets $D^* = N \cup \{+\infty\}$, where N corresponds to set of natural numbers.

For each fuzzy set $\tilde{\sigma}_{ij}$ of cardinality *m*, is associated a k-tuple of order k:

$$\sigma_{ij}^{k} = \left(\mu_1(i,j)/1, \ldots \mu_q(i,j)/q, \overbrace{1/+\infty, \ldots, 1/+\infty}^{k-q} \right) q \le m$$

The coefficients $\mu_I(i,j) | I \le m$ correspond to ÏkÓ grades of membership of the multicriteria valuation of the arc (i,j) to fuzzy sets $\tilde{\sigma}_{ij}$ (we generally complete by 1/+∞ to constitute a tuple of order k)

We define the set *S* in the following way:

Given a k-tuple $u = (\alpha_1 / u_1, \ldots, \alpha_k / u_k)$, $u \in S$ if and only if:

$$u_1 \le u_2 \le \ldots \le u_k$$

Where $\alpha_i \in [0,1] et\, u_i \in D^* | 1 \le i \le k$

For each arc *(i,j)* is associated the k-tuple σ_{ij}^k. The operations ⊕ and ⊗ are constructed in the following way:

The Operation ⊗

Consider A and B which are respectively, fuzzy sets of D^*. The sum of A and B is the fuzzy set denoted $A \mp B$ whose membership function is given by:

$$\forall z \in D^*, \mu_{A \mp B}(z) = \underset{z=x+y}{Sup} \left(\min\left(\mu_A(x), \mu_B(y) \right) \right)$$

If $u = (\alpha_1 / u_1, \ldots, \alpha_k / u_k)$ and $v = (\beta_1 / v_1, \ldots, \beta_k / v_k)$ are two k-tuples, then, consider Au and Av as two fuzzy sets associated respectively to u and v in the following:

$$A^u = \{\alpha_1 / u_1, \ldots, \alpha_k / u_k\} \text{ and } A^v = \{\beta_1 / v_1, \ldots, \beta_k / v_k\}$$
$$(\alpha_i, \beta_i) \in [0,1] \times [0,1] et\, (u_i, v_i) \in D^* \times D^* | 1 \le i \le k$$

Given $A^u \mp A^v$, let *w* be a tuple composed of the k smaller values of $A^u \mp A^v$. Then, we define $u \otimes v$ as exactly the tuple *w*, *w* = (u ⊗ v) is the tuple of order k.

The Operation ⊕

Consider A and B which are respectively, fuzzy sets of D^*. The union of A and B is the fuzzy set denoted $A \tilde{\cup} B$ whose function of membership is given by: $\forall z \in D^*, \mu_{A \tilde{\cup} B}(z) = \max(\mu_A(z), \mu_B(z))$

If $u = (\alpha_1 / u_1, \ldots, \alpha_k / u_k)$ and $v = (\beta_1 / v_1, \ldots, \beta_k / v_k)$ are two k-tuples, then:

Let *w* be a tuple composed of the smaller of the values of $A^u \tilde{\cup} A^v$. Then, we define $u \oplus v$ as exactly the tuple *w*, *w* = (u ⊕ v) is the tuple of order k.

The construction of ⊕ and ⊗ is made is such a way to build a dio'd structure.

Let $\tilde{A}$ be a fuzzy set of D^*. The support of $\tilde{A}$ denoted $\Theta(\tilde{A})$ is defined by $\Theta(\tilde{A}) = \{\omega \in D^* | \mu_{\tilde{A}}(\omega) \succ 0\}$, which is a classical set of D^*. Let $[\,]_k$ be the selection or sorting operator defined on classical sets of D^*. If A is the classical set of

D^*, then $[A]_k$ corresponds to set composed of the k first elements of A (k positive number), sorted in ascending order according to the order relation defined on fuzzy sets.

Let $\theta_k(\ )$ be the function defined on fuzzy sets of D^* by: $\theta_k(\tilde{A}) = \tilde{A} \tilde{\cap} [\Theta(\tilde{A})]_k$. The symbol $\tilde{\cap}$ corresponds to intersection operator defined on fuzzy sets.

With these definitions, if $u = (\alpha_1 / u_1, \ldots, \alpha_k / u_k)$ and $v = (\beta_1 / v_1, \ldots, \beta_k / v_k)$ are two k-tuples, then:

- $(u \oplus v)$ is the tuple notation of fuzzy set $\theta_k(A^u \tilde{\cup} A^v)$,
- $(u \otimes v)$ is the tuple notation of fuzzy set $\theta_k(A^u \mp A^v)$, $\theta_k(\ )$ satisfies the following properties:

$$\theta_k(A^u \tilde{\cup} A^v) = \theta_k(\theta_k(A^u) \tilde{\cup} \theta_k(A^v))$$

These properties results from those of operator $[\ ]_k$ defined on the set of natural numbers.

Proposition 5.1
The algebraic structure

$$\left(S, \oplus, \otimes, \varepsilon = (+\infty)^k, e = \left(\overbrace{1/0, \ldots, 1/0}^{q}, \overbrace{1/+\infty, \ldots, 1/+\infty}^{k-q}\right)\right) \text{ is a dio‘d}$$

Proof
The structure (S, ⊕, ⊗) verifies the properties of a dioïd. The distributivity of the operation ⊗ relative to ⊕ results from the distributivity of the addition in the fuzzy sets ($\mp$) relative to the union ($\tilde{\cup}$).

Numerical Example and Interpretation of the Operations ⊕ and ⊗
Let us suppose that the fuzzy valuations relating to the arcs indicate the amplitude of the risk associated with each section of the road (see Fig. 5.13). The risk is regarded as a fuzzy quantity binding the vulnerability and the economic costs.

A value of 1 represents a weak risk; a value of 3 corresponds to a high risk, etc. Consequently we can define a "subjective" sorting on the amplitude of the risk $(1 \le 2 \le 3 \le \ldots \le K)$, $K < +\infty$. In the writing of the fuzzy sets, the elements having a null grade of membership are omitted.

The fuzzy valuation $A^u = \{.1/1,.5/2,.8/3\}$ is associated with the arc (1,2), the valuation $A^v = \{.1/1,.5/2,.6/3\}$ is associated with the arc (2,3), and finally the valuation $A^w = \{.3/1,.6/2,.9/3\}$ is associated with the arc (2,4).

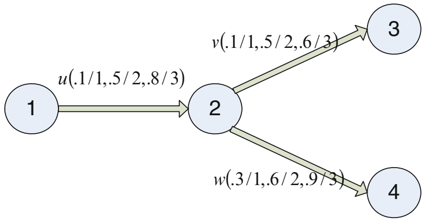


Fig. 5.13 Illustration and interpretation of the operations ⊗ and ⊕

The Operation ⊗

By applying the transformations and the rules of calculations previously developed, we obtain:

$$A^u \mp A^v = \{.1/2,.1/3,.5/4,.5/5,.6/6\},$$
$$A^u \mp A^w = \{.1/2,.3/3,.5/4,.6/5,.8/6\},$$
$$\text{Where } \mu_{A\mp B}(z) = \underset{z=x+y}{Sup}\left(\min\left(\mu_A(x),\mu_B(y)\right)\right).$$

At this stage the operation $\mp$ (fuzzy sets addition) on the fuzzy sets is allowed to give all amplitudes of the risk for the paths $1\to2\to3$ and $1\to2\to4$.

For k = 3, we have the first three fuzzy values according to the sorting defined on the amplitudes of the risk:

$$X = u\otimes v = \{.1/2,.1/3,.5/4\}, \text{ for the path } 1\to2\to3,$$
$$Y = u\otimes w = \{.1/2,.1/3,.5/4\}, \text{ for the path } 1\to2\to4.$$

Operation ⊕

To illustrate the calculation of the operation ⊕, we have:

$$A^u \,\breve{\cup}\, A^v = \{.1/1,.5/2,.8/3\},$$

where $\mu_{A^u \breve{\cup} A^v}(z) = \max\left(\mu_{A^u}(z),\mu_{A^v}(z)\right)$,
For k = 3, we obtain u ⊕ v = (.1/1,.5/2,.8/3).

For the k-best fuzzy path problem, the law ⊕ causes an order relation. In this example, we obtain:

$$(u\otimes v)\oplus(u\otimes w) = (.1/2,.1/3,.5/4)\oplus(.1/2,.3/3,.5/4)$$
$$= (.1/2,.3/3,.5/4).$$

In general, let x_1 and x_2 be two arcs and *u* and *v* their respective fuzzy valuations. We have $\mu_{A^u \breve{\cup} A^v}(z) = \max\left(\mu_{A^u}(z),\mu_{A^v}(z)\right)$, if the maximum is realised with A^u, then it is the arc x_1 which is retained as marker, if not it is the arc x_2. For the other cases, the two arcs are retained as markers (see Fig. 5.14).

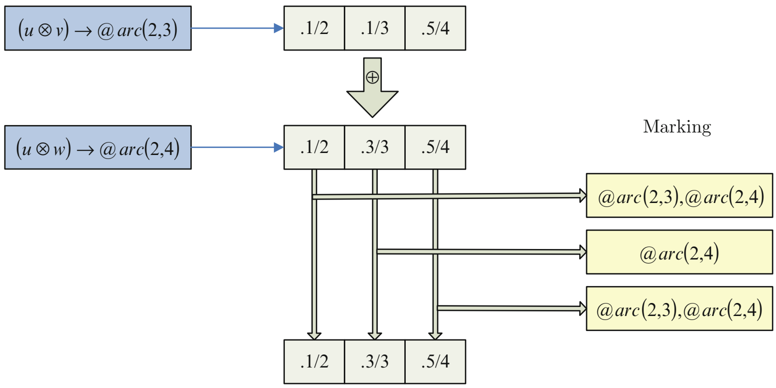


Fig. 5.14 Conservation of marking

5.4.3.1 Implementation of the General Algorithm

On the graph without cycles, we can trace execution of the general algorithm (Algorithm 5.2) to solve the problem of k-best shortest path by applying the dioïd structure given by Proposition 5.1. It is known that any graph without cycles with negative weight can be transformed into a graph without cycles. The use of the generalised algorithm for the graphs without cycles is not restrictive [35]. The process of transformation consists the decomposing of the graph into levels [35]. Klein refers to this observation in [28] which is based on the work of Lawler [35]. This transformation has a combinatory cost, which favours the generalised algorithm of Ford to solve the problem of k-best fuzzy shortest path.

The modelling of the fuzzy shortest path problem, proposed in this chapter is practical. In the case of a graph without cycles, we apply the generalised algorithm of Bellman (Algorithm 5.2). In the case of any graph that has no cycles of negative weight, we exploit the generalised algorithm of Ford (Algorithm 5.1).

5.5 Applying the Generalised Algorithm of Bellman to the Fuzzy Graph with No Cycles

Next we apply the generalised algorithm of Bellman to the fuzzy graph without cycles shown in Fig. 5.12.

Notation: $\pi^3(i)$ represents the marking of node i, for the problem of the k^{th} fuzzy shortest path:

$$\pi^3(6) = (1/0, 1/0, 1/0), \pi^3(5) = (.1/1, .3/2, .4/3), \pi^3(4) = (.5/1, .6/2, .6/3),$$
$$\pi^3(3) = a^3_{34} \otimes \pi^3(4) \oplus a^3_{35} \otimes \pi^3(5)$$
$$= (.2/1, .9/2, .3/3) \otimes (.5/1, .6/2, .6/3) \oplus (.3/1, .8/2, .1/3) \otimes (.1/1, .3/2, .4/3)$$
$$= (.2/2, .5/3, .6/4, .6/5, .3/6) \oplus (.1/2, .3/3, .3/4, .4/5, .1/6)$$
$$= (.2/2, .5/3, .6/4) \oplus (.1/2, .3/3, .3/4)$$
$$= (.2/2, .5/3, .6/4).$$
$$\pi^3(2) = a^3_{24} \otimes \pi^3(4) \oplus a^3_{25} \otimes \pi^3(5)$$
$$= (.2/1, .6/2, .9/3) \otimes (.5/1, .6/2, .6/3) \oplus (.2/1, .5/2, .6/3) \otimes (.1/1, .3/2, .4/3)$$
$$= (.2/2, .5/3, .6/4, .6/5, .6/6) \oplus (.1/2, .2/3, .3/4, .4/5, .4/6)$$
$$= (.2/2, .5/3, .6/4) \oplus (.1/2, .2/3, .3/4)$$
$$= (.2/2, .5/3, .6/4).$$
$$\pi^3(1) = a^3_{12} \otimes \pi^3(2) \oplus a^3_{13} \otimes \pi^3(3)$$
$$= (.1/1, .5/2, .8/3) \otimes (.2/2, .5/3, .6/4) \oplus (.3/1, .1/2, .6/3) \otimes (.2/2, .5/3, .6/4)$$
$$= (.1/3, .2/4, .5/5) \oplus (.2/3, .3/4, .3/5)$$
$$= (.2/3, .3/4, .5/5)$$

$\pi^3(1)$ is a fuzzy subset which gives the weight of the fuzzy shortest path. If the table of nodes (once the marking for node 1 has been carried out) is [3, 4, 4], then the first fuzzy shortest path of length 3 with degree of membership 0.2 is the path:

$$1 \rightarrow 3 \rightarrow 4 \rightarrow 6,$$

The second fuzzy shortest path of length 4 with a degree of membership 0.3. is the path:

$$1 \rightarrow 3 \rightarrow 4 \rightarrow 6,$$

The third fuzzy shortest path of length 5 with a degree of membership 0.5. is the path:

$$1 \rightarrow 2 \rightarrow 4 \rightarrow 6,$$

It should be noted that the paths are obtained by the marking of the predecessor nodes for each value of fuzzy subset $\pi^3(i)$ of a node i.

5.5.1 Overview of Proposed Fuzzy Risk Modelling [3, 4]

The impact of an accident involving hazardous materials transportation can be considerable, affecting the dimensions of the environment, infrastructure, economy, etc. The risk associated with HazMat transportation depends on the type of product transported and the potential damage (taking all dimensions into account).

A fuzzy approach, using fuzzy data, is proposed to model en route risk. This approach is based on some basic concepts of multicriteria analysis and fuzzy set theories. This technique produces a fuzzy risk graph that models both the transportation system and the concept of accident risk. This fuzzification of risk enables the results concerning fuzzy shortest path-finding problems in fuzzy graphs to be applied to lesser risks engendered by TDG.

The approach models the concept of accident risk on each arc of the transportation network. It takes into account the vulnerability of the arc in question and the cost generated in the event of an accident on this arc. This is assessed with respect to the various impacts which are categorised according to the dimension in question. For example, in the dimension of environment, potential impacts include population, protected areas, etc.

In this approach, the concept of the vulnerability of an arc replaces and generalises that of the probability of having an accident on an arc (found in the traditional method). It is evaluated by taking into account not only accident data on a given arc (such data are not always available) but more general information concerning the arc in question and its surroundings, with respect to a given impact. The cost of the consequences generated in the event of accident on an arc is estimated in modules relating to the impacts in question. The proposed method is thus more general than the traditional method [12] as it makes it possible to take more factors into account in risk modelling.

Moreover, the introduction of these parameters is done in a way which simplifies probability calculations used in the classical method, as the level of vulnerability and the eventual accident cost on an arc (with respect to a given impact) are taken as being fuzzy quantities. These fuzzy quantities are obtained by asking network managers (experts and decision makers) to assign, with a degree of plausibility, qualitative evaluations to these various parameters. In this way, complex probability calculations are replaced by human judgment which enables the experience of transport network actors to be integrated into the model. The impact of the risk is then calculated as the product of the vulnerability and the cost components evaluated for each impact on each arc. The overall level of risk on each arc is calculated by the application of an appropriate fuzzy aggregation operator on the fuzzy parameters corresponding to the various impacts that have been taken into account. More details of this approach are given in [3, 4].

5.6 A Prototype System

The city and region of Mohammedia (Morocco) is home to intensive oil and chemical industry activity. The economic life of the region is linked partly, to its geographical position. It is close to both the economic capital, Casablanca and to ports.

Industrial installations receive and dispatch many harmful substances that present, in the case of an accident, risks for people and the environment. For this reason, the city has been chosen as a pilot site for a prototype of a real-time mobile information system for TDG.

The server side uses Oracle 10 g as a spatial database and MapObjects software for mapping and GIS components. MapObjects software is used to create applications that include dynamic live maps and GIS capabilities (e.g. spatial and attribute querying, geo-coding). ArcSDE technology provides database access to spatial data, its associated attributes, and the metadata stored within an object-relational DBMS.

The model integrates multiple layers of data such as accidents, populations, networks and others. It provides multicriteria fuzzy routing of TDG via the selection of origin and destination nodes. In the background, the k-best fuzzy shortest path problem algorithms are applied. The algorithms then collect relevant information concerning accidents and populations from the accident and population layers.

5.7 Conclusion

This chapter describes a mobile information system with real-time considerations for HazMat telegeomonitoring arising from previous work by the authors and others. It illustrates the integration of various software components and describes real-time object-oriented modelling using UML 2 and RT-UML. Components based modelling is also described.

We then explain how advances in computer science can vastly improve this system. We propose a mobile object data model and describe its main relationships with the transportation network. The main elements of a spatiotemporal query language are presented using powerful spatiotemporal predicates. This approach represents the first steps in the construction of innovative moving object databases. Furthermore, the performance of the proposed system is also significantly increased by incorporating a SDSS that analyses the risk posed by hazardous materials and provides routing strategies that minimise the transportation risk. In particular, a dioïd algebraic structure is adapted to give an original solution to the k-best fuzzy shortest paths problem.

Therefore this work illustrates that because of its complexity, this type of system requires a number of considerations to be taken into account to provide satisfactory HazMat services: improvement to communication, the development of more refined web services, real-time and schedulability validation, etc.

References

1. Björkander M, Kobryn C (2003) Architecting systems with UML 2.0. IEEE Software 20(4):57–61
2. Boulmakoul A (2004) Generalized path-finding algorithms on semirings and the fuzzy shortest path problem. J Comput Appl Math 162(1):263–272

3. Boulmakoul A (2006) Fuzzy graphs modeling for HazMat telegeomonitoring. Eur J Oper Res 175(3):1514–1525
4. Boulmakoul A, Bouziri AE (2009) Modélisation d'objets mobiles pour les systèmes d'information géolocalisés itinérants en temps réel. le numéro spécial sur les Systèmes d'information et géolocalisation de la revue Ingénierie des systèmes d'information 14:5
5. Boulmakoul A, Chala M, Bouziri AE, Laurini R (2008) Modeling a real time mobile information system for HazMat telegeomonitoring. In: Bersani C, Boulmakoul A, Garbolino E, Sacile R (eds) Advanced technologies and methodologies for risk management in the global transport of dangerous goods, vol 45, NATO Science for Peace and Security Series: Human and Societal Dynamics. IOPress, Washington, DC, pp 169–193
6. Boulmakoul A, Laurini R (1999) Système d'informations télégéomatiques pour la supervision des transports des matières dangereuses. Revue Internationale de Géomatique 9(3):317–336, Hermès Ed
7. Boulmakoul A, Laurini R, Servignes S, Idrissi MAJ (1999) First specifications of a telegeomonitoring system for the transportation of hazardous materials. Comput Environ Urban Syst 23:259–270, Pergamon
8. Boulmakoul A, Zeitouni Z, Laurini R (1997) Un système d'information environnemental urbain pour la surveillance du transport des matières dangereuses: cas de la ville de Mohammedia-Maroc. Conférence européenne sur les technologies de l'information pour l'environnement, Strasbourg, Metroplolis, pp 187–196, ISBN: 9518-163-3
9. Boulmakoul A, Zeitouni Z, Laurini R, Aufaure MA (1997) Spatial decision support system for hazardous materials transportation planning. In: IFAC transportation systems'97, June 16–18, Chania, Greece, pp 611–616
10. Chanas S, Delgado M, Verdegay JL, Vila MA (1995) Fuzzy optimal flow on imprecise structures. Eur J Oper Res 83:568–580
11. Delgado M, Verdegay JL, Vila MA (1990) On valuation and optimisation problems in fuzzy graphs: a general approach and some particular cases. ORSA J Comput 2(1):75–83
12. DOT (1989) Guidelines for applying criteria to designate routes for transportation hazardous materials, Report No. DOT/RSPA/OHMT, 1989–02, Federal Hwy. Admin, Washington, DC
13. Dubois D, Prade H (1978) Algorithmes de plus courts chemins pour traiter des données floues. RAIRO/Oper Res 2(2):213–227
14. Dubois D, Prade H (1980) Fuzzy sets and systems. Academic, New York
15. Egenhofer MJ (1991) Point-Set topological spatial relations. Int J Geogr Inf Syst 5(2): 161–174
16. Erwig M, Güting RH, Schneider M, Vazirgiannis M (1999) Spatiotemporal data types: an approach to modeling and querying moving objects in databases. Geoinformatica 3(3):269–296
17. Erwig M, Schneider M (2002) Spatio-temporal predicates. IEEE Trans Knowl Data Eng 14(4):881–901
18. Furukawa N (1994) A parametric total order on fuzzy numbers and a fuzzy shortest route problem. Optimization 30:367–377
19. Galperin D (1977) On the optimality of A*. Artif Intell 8(1):69–76
20. Gondran M (1975) Path algebra and algorithms. In: Roy B (ed) Combinatorial programming: methods and applications. D. Reidel Publish Co, Dordrecht, Holland, pp 137–148
21. Gondran M (1975) Algèbre linéaire et cheminement dans un graphe. RAIRO/Oper Res 1:77–99
22. Gondran M, Minoux M (1984) Linear algebra in dioïds: a survey of recent results. Ann Discrete Math 19:147–164
23. Gondran M, Minoux M (1995) Graphes et algorithmes, 3rd edn. Eyrolles, Paris
24. Güting RH, Schneider M (2005) Moving objects databases. Morgan Kaufmann Publishers, London
25. Güting RH et al (2000) A foundation for representing and querying moving objects. Geoinformatica ACM Trans Databases Syst 25(1):1–42

26. ISO/TC211 (2000) Geographic Information/Geomatics: ISO 19108-Temporal Schema
27. Jagoe A (2002) Mobile location services—the definitive guide. Prentice Hall PTR, Denver
28. Klein CM (1991) Fuzzy shortest paths. Fuzzy Sets Syst 39:27–41
29. Kobryn C, Samuelsson E (2003) Driving architectures with UML 2.0. A Telelogic white paper
30. Koczy LT (1992) Fuzzy graphs in the evaluation and optimisation of networks. Fuzzy Sets Syst 46:307–319
31. Koubarakis M, Sellis T (2000) Spatiotemporal databases: the Chorochronos approach. Springer Verlag, New York
32. Kuntzmann J (1972) Théorie des réseaux. Dunod, Paris
33. Küpper A (2005) Location based services. Wiley, Chichester
34. Laurini R (2000) An introduction to TeleGeoMonitoring: problems and potentialities. In: Atkinson P, Martin D (eds) GIS Innovations. Taylor and Francis 1999, London, pp 11–26
35. Lawler E (1976) Combinatorial optimisation; networks and matroids. Holt, Reinehart and Winston, New York
36. Lin KC, Chern MS (1993) The fuzzy shortest path problem and its most vital arcs. Fuzzy Sets Syst 58:343–353
37. Minoux M (1976) Structures algébriques généralisées des problèmes de cheminements dans les graphes: Théorèmes, algorithmes et applications. RAIRO – Rech Opérationnelle 10(6):33–62
38. Minoux M (1977) Generalized path algebras. In: Prekopa A (ed) Surveys of mathematical programming. Publishing House of the Hungarian Academy of Sciences, Budapest, pp 359–364
39. OGC (1999) OpenGIS simple feature specification for SQL. Document 99–049
40. OGC (2003a) OpenGIS location services (OpenLS™): Part 1–5 Core Services". OGC 03-006r1
41. OGC (2003b) Open GIS web services architecture (WSA). OGC 03–025, version 0.3
42. Okada S, Soper T (2000) A shortest path problem on a network with fuzzy arc lengths. Fuzzy Sets Syst 109:129–140
43. OMG (2005a) UML profile for schedulability, performance and time. formal/05-01-02, version 1.1
44. OMG (2005b) Unified modeling language: superstructure. Formal/05-07-04, version 2.0
45. Prade H (1979) Using fuzzy set theory in a scheduling problem: a case study. Fuzzy Sets Syst 2:153–165
46. Stojanovic D, Djordjevic-Kajan S (2003) Modeling and querying mobile objects in location-based services. Sci J Fac Univ 18:59–80, Series Mathematics and Informatics, NIS, Serbia
47. TRANSMODEL (2003) Reference data model for public transport (in UML), version 5.1
48. Vazirgiannis M, Wolfson O (2001) A spatio-temporal model and language for moving objects on road networks. In: Proceedings of 7th SSTD, USA, pp 20–35
49. Zadeh LA (1965) Fuzzy sets. Inform Control 8:338–353

Chapter 6
Multi-path Multi-criteria Routing of Hazardous Materials in Time-Dependent Networks

Solmaz Haji Hosseinloo, Urszula Kanturska, Michael G.H. Bell, and Achille Fonzone

Abstract The transport of hazardous materials carries a risk of releasing dangerous substances. Quantification of risk has two dimensions, probability and consequence. Because accidents involving hazmat vehicles are low-probability-high-consequence events, the probabilities of which are non-quantifiable or unknown, risk-averse planning of hazmat transport needs to be focused on minimising the potential adverse consequences. Such an approach was proposed by Bell (Netw Spatial Econ, 6, 253–265, 2006), who departed from the established practice of finding a single safest route and demonstrated that generally for repeated hazmat shipments the safest strategy is to share these between several routes. This has been shown using a game-theoretic framework where unknown event probabilities are replaced by worst-case disruption/accident probabilities, on the basis of which the optimal route-usage frequencies are determined. In this paper we exploit the equivalence of a game against local demons and the hyperpath generated by the Spiess and Florian (SF) algorithm (Spiess and Florian, Transportation Res Part B Methodol, 23B(2), 83–102, 1989). We extend the SF approach to multi-criteria time-dependent setting to explicitly model the trade-off faced by the hazmat dispatcher between travel costs, potential delays and the potentially negative impacts on people and the environment. A simple case study is used to illustrate the method.

S.H. Hosseinloo (✉) • U. Kanturska • M.G.H. Bell • A. Fonzone
Centre for Transport Studies, Imperial College London, London SW7 2AZ, UK
e-mail: m.g.h.bell@imperial.ac.uk; a.fonzone@imperial.ac.uk; solmaz.haji06@imperial.ac.uk; urszula.kanturska05@imperial.ac.uk

E. Garbolino et al. (eds.), *Transport of Dangerous Goods*, NATO Science for Peace and Security Series C: Environmental Security, DOI 10.1007/978-94-007-2684-0_6,

6.1 Introduction

Dangerous goods are commercially transported substances or materials that can pose a significant risk of fire, explosion or environmental damage to health, safety or property. The term covers a long list (211 pages!) of items such as: explosives; flammable, poisonous or compressed gases; flammable and combustible liquids and solids; oxidisers; poisonous or infectious substances and radioactive materials.

The risk associated with the transport of hazardous materials (hazmat) arises from the possibility of release of dangerous substances and their negative impact on the environment and people. Depending on the type of cargo, the consequences of accidents leading to hazmat release may be considerable (environmental pollution, economic damage, evacuations, injuries and fatalities).

The generic problem of planning routes for the shipment of hazardous materials is to concurrently minimise the transportation cost (by sending the shipments via the shortest and most reliable route) and the potential negative impacts (by avoiding routes leading through densely populated or environmentally sensitive areas). These two aims may often be in conflict, for example when the shortest route leads through a city centre.

Both cost and risk associated with hazmat transport can be reduced by appropriate routing and scheduling. Distributing repetitive shipments between several routes reduces the exposure to loss in the event of an incident; distributing shipments in time has a similar effect. In cases where there are departure time restrictions, the time-varying nature of parameters (population along the routes, link travel times and delays) can also be taken into account, leading to dynamic multi-path routing.

In this paper we briefly review common approaches to hazmat routing and discuss how trade-offs have been considered. We then focus on multi-path planning using a game-theoretic approach, which in one special case is equivalent to the hyperpath generated by the Spiess and Florian (SF) algorithm [16]. We then review some contributions to hazmat scheduling, apply a reversed SF algorithm to a time-dependent network, and show how it can be extended to consider multi-criteria problems, which is the main theoretical contribution of this paper.

6.2 Literature Review

6.2.1 The Nature of Risk

Incidents related to dangerous goods transport may have a wide range of negative consequences; environmental pollution, need for evacuation, economic damage, injuries and fatalities.

Formally, risk is usually measured by the product of the loss and its probability, with the loss usually expressed in terms of the severity of its adverse effects.

The most common measure of risk used in hazmat route planning is the expected number of people affected by a shipment along route R composed of links $i \in R$

$$E = \sum_{i \in R} q_i e_i$$

Where e_i is the number of people affected if an incident occurs on link i and q_i is the probability of an incident occurring on that link.

The value of probability q_i depends on the nature of the threat. The threat can be described by two parameters: intent (neutral or malicious) and cause (delay or road accident). For example, assuming neutral intent, one could derive incident probabilities q_i from historical road safety records; assuming malicious intent, probabilities should take into account security intelligence information.

In all cases, due to their infrequent occurrence, incidents affecting hazardous materials are considered "Low-Probability-High-Consequence" (LPHC) events, probabilities of which are difficult to quantify or unknown.

6.2.2 *Minimisation of Risk Though Route Choice*

Risk associated with transporting hazardous materials can be reduced by appropriate choice of routes. Despite the LPHC property of incidents involving hazmats, many existing decision-support systems for dispatchers attempt to quantify the risk and recommend the best route accordingly.

The recommended route depends critically on the model adopted for the quantification of risk, as shown by Erkut and Verter [9]. Many such models are possible, for example techniques that minimise the total distance, the expected number of incidents (fatal or otherwise), the incident probability, or the residential population within a given distance of a route. The most popular measure of risk remains, however, the product of the incident probability and the number of people affected giving the expected number of people exposed to the incident.

An alternative approach is to acknowledge that when dealing with LPHC events there is often insufficient data to estimate incident probabilities, in particular for individual sections of the route, and that the historical data on accident rates quickly becomes obsolete due to improved vehicle designs and safety measures applied at accident black spots. Consequently, risk-averse planning of hazmat transport tends to focus on minimising the potential adverse consequences, which is sometimes referred to as *catastrophe avoidance*.

For example, Sivakumar, Batta et al. (1993) [15] define the cost of a hazmat release on a link as a loss and suggest that shipments should be directed along the route that has the minimum sum of link losses, until an incident occurs. While such an approach provides a proxy for the maximum possible damage, it can lead to illogical routes, as shown by Erkut (1995) [7]. Routes that minimise the sum of the link losses can also have an unacceptably high risk (the sum of the products of the link losses and their respective probabilities) or route incident probabilities.

Erkut and Ingolfsson (2000) [8] directly introduce risk-aversion into the routing of hazardous materials and propose three ways of doing so; (1) minimise the maximum consequence along a route, (2) incorporate the variance of the losses along a route into route selection, or (3) minimises the expected disutility of the losses with utility expressed by a convex function expressing a risk-averse attitude. It is shown that all three approaches can be solved as shortest-path problems. Even though the first method can be used when incident probabilities are unknown, it leads to the recommendation of a single route only.

Rather than finding a single safest route, Bell [2] demonstrated that for repeated hazmat shipments the safest strategy is to allocate the shipments to several routes on the grounds that this reduces exposure in the event of an incident, defined as the loss in the event of an incident multiplied by the probability of using the route in question. This approach makes use of a game-theoretic framework proposed first by Bell [1].

6.2.3 Game-Theoretic Approach to Route Choice

The game-theoretic framework is applicable in contexts where the outcome depends on the actions of several agents, the interests of which may be different or opposing. With respect to hazmat transport the first player is a dispatcher who wants to move hazmat cheaply and safely. The second player is hypothetical and represents anything that might prevent achievement of the dispatcher's goals.

Different types of games arise depending on the assumption made about the threat, in particular about the intent. If incidents occur at random without malicious intent, the situation is modelled as a *game against nature*. If incidents are caused by an intelligent malicious entity or entities, the situation is known as a *game against one or more demons*. Both types of game are reviewed by Schmöcker, Bell et al. [14]. The solution to a game indicates the best routing strategy. A pure strategy corresponds to the single best route between a given OD pair while a mixed strategy corresponds to using multiple paths with certain frequencies.

6.2.4 Games Against Nature

In games against nature, incident probabilities are unknown, but they are assumed to be independent of the route choice, because nature does not specifically act to harm the dispatcher and its actions do not depend on his choices. In this context the best course of action for a dispatcher can be determined using different solution principles, for example:

- Laplace's criterion of insufficient reason: All the states of nature are considered to be equally likely, and the route with the lowest expected risk is chosen.
- Wald's minmax approach: Actual probabilities are ignored and replaced with worst case probabilities which are then used to select a route with the lowest worst case risk.

- Hurwicz's expected risk minimisation: Expected risk is calculated as a combination of the worst and best case scenario for a chosen coefficient of optimism $0<\alpha<1$ as α*min_risk + (1–α)*max_risk and a route is chosen which has the lowest expected risk.
- Savage's regret minimisation: Regret is experienced after the actual act of nature is known; a regret matrix is constructed from payoff matrices by calculating the difference between each entry and the smallest entry in the column; a route is chosen that guarantees maximum risk minimisation.

Each of the above principles leads to a pure strategy (the selection of one best route).

6.2.5 Games Against a Demon

In games against a demon it is assumed that incidents are caused by a malevolent agent (or demon) who tries to maximise the inflicted damage. Incident probabilities are unknown, but they are assumed to be dependent on the dispatcher's route choice: the demon tries to predict which routes will be used and focuses his attack on these. This situation is usually modelled as a zero-sum game, meaning that the goals of the dispatcher and those of the demon are exactly opposite: the dispatcher's loss is the demon's gain.

The dispatcher is generally assumed to be risk-averse and applies Wald's principle to minimise his maximum expected risk. As for the demon, different assumptions have been proposed in the literature on hazmat transport regarding the demon's capability to impose an incident. The basic idea, first studied by Bell [1], is to replace the unknown event probabilities by worst case probabilities, on the basis of which the optimal routes and their usage frequencies are determined. The usage frequencies and worst case probabilities are related to each other.

6.2.5.1 Global Game Against a Demon

If a demon can cause only one incident in the entire network, path selection occurs only at the origin. This approach is suitable for strategic rather than operational route planning in a context where incidents are rare as planning is based on only one incident. Originally devised to analyse network reliability, it is a game between a traveller, who seeks a least-cost path between his origin and destination, and a demon, who strives to maximise the trip cost by failing one link. Each link is therefore in one of two states - normal or failed - and no more than one link is failed simultaneously. The link cost in the failed state is higher than in the normal state, and signifies the increased traversal time due to reduced link capacity or another link-specific measure of incident consequence.

The equilibrium solution to the game indicates an optimal routing strategy (a set of paths with associated optimal usage frequencies) allowing the traveller to avoid excessive costs irrespective of which link is failed. The optimal strategy for the demon indicates link-failure probabilities that are interpreted as the criticality of links for network performance.

Bell [2] applied the above concept to hazmat transport. He formulated a similar non-cooperative zero-sum game between a dispatcher and a demon, where expected cost has been replaced with expected loss. There are two types of cost associated with all links in the network: traversal costs and loss measured by the population inside a circle of a given impact radius. Population exposure is a product of link loss (it is assumed that the population in the vicinity of a link is distributed evenly so that the population exposed in the event of an incident is independent of the location on a link where the incident occurs) and link use frequency. In this way travel cost and incident consequence are both included in problem formulation. The routing strategy results from the expected loss considering both these costs and the worse-case location of accidents. By employing a mixed routing strategy, the dispatcher can minimise his exposure to loss in the event of an incident. Bell [2] also shows that the overall expected cost at the solution is equal to the expected cost of travel plus the exposure to loss on any link in the event of an incident.

Nagae and Akamatsu (2007) [11] extend Bell's [2] mixed route maxmin problem into a more generalised framework by adding an entropy term in path use probabilities. They reveal that this extended maxmin problem reduces to a single level convex programming problem, which shares a common mathematical structure with the logit stochastic user equilibrium traffic assignment model. Replacing the entropy term in path use probabilities by an entropy term in link-failure probabilities, Bell, Kanturska et al. [4] produced another and arguably more useful single level convex programming problem, which allows the user to vary the level of risk aversion.

6.2.5.2 A Sequence of Local Games Against a Demon

If the demon's actions are considered in a more dynamic setting, path selection occurs as a sequence of on-the-spot decisions at every node as the trip unfolds. It is assumed that at each node there is a 'local demon' that can cause an incident at any of the outgoing links. The solution to such a game indicates the optimal link-use probabilities for links emanating from every node.

Schmöcker, Bell et al. (2009) [14] showed that such repeatedly played games are equivalent to the hyperpath obtained from the SF algorithm [16] when link frequency is replaced by the reciprocal of link loss. This algorithm, first introduced in the context of public transport assignment, returns a routing strategy that minimises exposure to loss upon departure from every node.

Bell (2009) [14] proposed the SF algorithm for use in route guidance, and by so doing provided an algorithm that minimises exposure to maximum loss upon departure from every node. The paper also shows that the A* speed-up can be applied to the SF algorithm, and called the result Hyperstar. This paper does not, however, consider time dependency. In this paper the SF algorithm is reversed so that exposure to loss upon arrival at every node is minimised, time-dependent link travel times and losses are considered and link travel time and loss is evaluated at the (pessimistically) expected time of entering the link. Reversing the algorithm allows the optimal time-dependent hyperpath to be determined for a given dispatch time. If the SF algorithm were not reversed, link travel times and losses would be evaluated at the (pessimistically) expected time of exiting the link, exposure to loss upon departing from every node would be minimised and the optimal time-dependent hyperpath would be determined for a given arrival time. Both variations are potentially interesting.

6.2.6 Time Dimension

Significant improvements in safety and travel time reliability can be achieved not only by the choice of routes, but also by appropriate scheduling of deliveries. This was first pointed out in Nozick, List et al. (1997) [12], who indicated that the probability of an incident and the number of people impacted when an incident occurs are usually dependent on the time of day. Also traffic conditions change in the course of the day. Time-dependence of these input parameters should be accommodated by the algorithms used for finding the best routes and schedules for hazmat shipments.

Bersani, Minciardi et al. (2010) [6] extends Bell (2006) [2] to the case where the consequences depend on when the accident happens. Accident timing determines both the size of population exposed and the frequency of hazmat shipments. It is shown that spreading deliveries over time allows for further reduction in the expected consequences.

Szeto and Sumalee (2009) [18] study simultaneous routing and scheduling of hazardous materials, extending the approach of Bell (2006) [2] to the case with multiple OD pairs, multiple classes of hazmats, and multiple global demons (and hence multiple incidents). The problem is considered over the space-time expanded network (STEN). A STEN, originally proposed by Ford and Fulkerson (1962) [10] as an approach to studying the dynamic maximum flow problem, is an enlarged static representation of the time-dependent network where the traditional (static) shortest-path algorithms can be applied.

6.2.7 Multi-criteria Routing

In the context of hazmat transport, dispatchers may prefer a path that simultaneously minimises travel cost, time, distance, accident likelihood and population exposure or any combination of these. If the cost associated with using each link include: transportation cost (proportional to the length of the link), delay costs (proportional to travel time in excess of the usual travel time) and cost of an incident (adverse consequences to society and the environment, dependent on the hazmat type, and usually proportional to the number of people affected) it is unlikely that a single path exists between an OD pair that is best with respect to all these criteria.

A single-objective approach in which each criterion is considered individually does not reflect the complexity of choice. Several objectives can be considered concurrently in two ways: a single-objective optimisation performed on a combined measure (utility) or a multi-objective optimisation.

The multi-criteria (or multi-objective) shortest-path problem (see [19] for a review of such algorithms) often results in multiple paths (a set of Pareto-optimal or non-dominated paths) [3]. So does the game-theoretic framework proposed of Bell (2006) [2] , where the trade-off between transportation cost and risk of population exposure is implicitly considered.

Opasanon and Miller-Hooks (2006) [13] review methods for generating Pareto-optimal paths in the case of multi-criteria problems in time-invariant networks, both deterministic and stochastic. They propose three algorithms for finding multi-criteria hyperpaths in stochastic time-dependent networks. The first algorithm generates all Pareto-optimal all-to-one hyperpaths with respect to the expected value of multiple criteria. These hyperpaths comprise a set of path strategies that provide the traveller with Pareto-optimal directions at each node en route in response to knowledge of the arrival time at intermediate locations. The authors acknowledge that any technique that generates all Pareto-optimal solutions has exponential worst-case computational complexity and may require enormous computational effort. Thus two variations of the algorithm are proposed that rely on a linear utility function and produce only a single hyperpath that minimises the expected disutility.

6.3 Problem Formulation and Notation

6.3.1 Proposed Method

In this paper we propose an algorithm for multi-criteria hazmat routing in time-dependent networks where the values of input parameters change over time. Since the use of the linear or exponential utility function allows the use of any labelling algorithm for finding the optimal path in static networks without violating Bellman's principle, in this paper we propose extending the reversed SF algorithm to a multi-criteria time-dependent setting.

The reversed SF algorithm proposed has recently been adapted by Bell, Trozzi et al. (in press) [5] to generate hyperpaths in time-dependent networks where link travel time and link delay are a function of time of arrival at the link. Time-dependency has been incorporated into the model following the approach of Sung, Bell et al. (2000) [17] known as the *flow speed model*. The main advantage of this method is that having speed intervals rather than time intervals ensures that all vehicles on a given link travel at the same speed at any point in time, so there is no overtaking and the first-in-first-out principle holds. We extend this approach to a multi-criteria setting, and present an example with accident cost (link loss) as a dimension of choice additional to travel time and delay.

In recognition of the importance of hazmat scheduling, all formulations presented in this paper are directly applicable to a STEN, and as such can also be used to find the optimal split of shipments between departure times. For clarity we omit the presentation of the scheduling aspect.

6.3.2 Notation

Consider a road network consisting of a set of links A and a set of nodes I. The following variables are defined:

A	Set of links
I	Set of nodes
H	Set of links in the hyperpath from origin r to destination s
r	Origin node
s	Destination node
A_i^+	Set of links entering node i
A_i^-	Set of links exiting node i
$c_a(t)$	Undelayed travel time on link a for trips arriving at link a at time t
$d_a(t)$	Maximum delay for link a for trips arriving at link a at time t
$e_a(t)$	Maximum loss for link a for trips arriving at link a at time t
p_a	Probability of using link a
w_i	Exposure at node i arising from expected probability of delay in arriving at node i, based on the frequency of using the links entering that node
z_i	Exposure at node i arising from expected probability of loss for arriving at node i, based on the probability of using the links entering that node
u_i	Pessimistically expected arrival time at node i
f_i	Sum of inverse maximum link delays for attractive links entering node i
y	Probability of using node i

6.3.3 Problem Formulation

The original routing strategy problem formulated by Spiess and Florian has a form of linear program and, as shown by Schmöcker, Bell et al. (2009) [14], is equivalent to a sequence of games against a local demon. The reversed time-dependent version proposed by Bell, Trozzi et al. (in press) [5] is also formulated as a linear program:

$$\min_{p,w} \sum_{a=(i,j)\in A} p_a c_a(u_i) + \sum_{j\in I} w_j(u_i)$$

subject to

$$\sum_{a=(i,j)\in A_j^+} p_a - \sum_{a=(i,j)\in A_j^-} p_a = b_j \text{ for all } _{j\in I}$$

$$w_j(u_i) \geq p_a d_a(u_i) \text{ for all } a=(i,j)\in A_j^+, j\in I$$

$$p_a \geq 0 \text{ for all } _{a=(i,j)\in A}$$

Link travel time is a function of pessimistically expected arrival time at every node, and link usage is calculated to minimise pessimistically expected travel time and delay. If $p_a \geq 0$ then $a = (i,\ j) \in H$ and $u_i + c_a(u_i) \leq u_j$ since a link would not be attractive (and therefore not in the hyperpath) if it did not offer the possibility of reaching node *j* earlier.

The extension of the above problem to a multi-criteria setting is conceptualised as a series of local games against as many demons as there are criteria. Each demon represents a different incident type, e.g. one demon can cause delays, while another can cause an accident leading to loss resulting from population exposure and environmental damage. The increase in link costs will only occur if the link is affected by delay, by accident or both. The formulation with two criteria can be written as follows:

$$\min_a \sum_a p_a c_a(u_i) + \sum_i w_i(u_i) + \sum_i Z_i(u_i)$$

subject to

$w_i(u_i) \geq p_a d_a(u_i)$ where $a \in A_i^+$

$z_i(u_i) \geq p_a e_a(u_i)$ where $a \in A_i^+$

6.3.4 Adaptation of the SF Algorithm for Multi-criteria Applications

The solution algorithm to the above problem is given below. The algorithm is merely an extension to the original SF algorithm with the following modifications:

- The algorithm is reversed which means the search starts from the origin.
- The algorithm is time-dependent with link travel time, link maximum delay and link loss all time-dependent.
- In addition to that, there are two criteria for consideration, maximum delay and maximum loss, denoted by $d_a(u_i)$ and $e_a(u_i)$ respectively.

1. Initialisation

   $u_i \leftarrow \infty$ for $i \in I - \{r\};\ u_r \leftarrow 0$

   $f_i \leftarrow 0$ for all $i \in I$

   $y_i \leftarrow 0$ for $i \in I - \{s\};\ y_s \leftarrow 1$

   $L \leftarrow A$

   $H \leftarrow \varnothing$

2. Select link *a*

   Find $a = (i, j) \in L$ with minimum $u_i + c_a(u_i)$

   $L \leftarrow L - \{a\}$

3. Update node i If $u_j \geq u_i + c_a(u_i)$ then

If $u_j = \infty$ and $f_j = 0$ then $\beta \leftarrow 1$ else $\beta \leftarrow f_j u_j$

$$u_j \leftarrow \frac{\beta + \left(\frac{1}{d_a(u_i) + e_a(u_i)}\right)\left(u_i + c_a(u_i)\right)}{f_i + \frac{1}{d_a(u_i) + e_a(u_i)}}$$

$$f_i \leftarrow f_j + \frac{1}{d_a(u_i) + e_a(u_i)}$$

$$H \leftarrow H + \{a\}$$

If $L = \varnothing$ or $u_i + c_a(u_i) > u_s$ then go to Step 3, else go to Step1

4. Loading For every link $a = (i, j) \in A$ in decreasing order of $u_i + c_a(u_i)$

If $a = (i, j) \in H$ then

$$P_a \leftarrow \frac{1}{f_j}\left(\frac{1}{d_a(u_i) + e_a(u_i)}\right) y_j \quad \text{and} \quad y_j \leftarrow y_j + p_a$$

else $p_a \leftarrow 0$

6.4 Numerical Example

6.4.1 Example Network

Here a numerical example is presented to illustrate the trade-offs faced by the dispatchers and different outcomes for different objective functions. A small network shown in Fig. 6.1 is considered over a 1 h analysis period from 8:30 to 9:30.

In this example, each link is characterised by three parameters: non-delayed travel time (c), maximum expected delay (d) and cost (human, environmental, operational etc.) of a hazmat accident involving release on a link (e). The analysis period is divided into three intervals, each 20 min in duration. Table 6.1 below shows values of these parameters for each interval. The average speed on all links is assumed to be 50 km/h during the first and the last interval, dropping to 30 km/h during the middle interval, affecting the expected travel time (c). The number of people affected in case of an accident (e) is assumed to be constant over time for all links except for link BK that traverses the Central Business District. The number of people present in this area increases as they arrive during the peak period. The maximum expected delay (d) is assumed to be constant over time.

We first show a simplified case when the static algorithm is applied to the average (over the entire time interval) values of link parameters listed in Table 6.2.

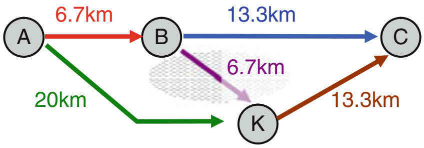


Fig. 6.1 Example network

Table 6.1 Time-dependent network parameters

	c			*d*			*e*		
	8:30–8:50	8:50–9:10	9:10–9:30	8:30–8:50	8:50–9:10	9:10–9:30	8:30–8:50	8:50–9:10	9:10–9:30
A-B	8	13.3	8	5			10		
B-C	16	26.7	16	10			30		
A-K	24	40.0	24	10			5		
B-K	8	13.3	8	0.1			2	6	7
K-C	16	26.7	16	5			20		

Table 6.2 Average (static) network parameters

	c	*d*	*e*
	8:30–9:30	8:30–9:30	8:30–9:30
A-B	10	5	10
B-C	20	10	30
A-K	30	10	5
B-K	10	0.1	5
K-C	20	5	20

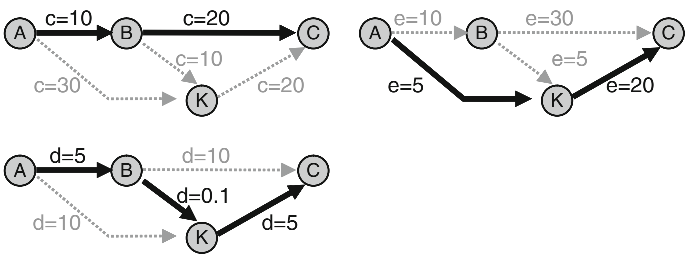


Fig. 6.2 Alternative paths from origin A to destination C

There are three alternative paths from origin A to destination C shown in Fig. 6.2. The shortest time path is 30 min leading through nodes A-B-C. The lowest population exposure path is A-K-C with a total of 250 people at risk. In turn, path A-B-K-C is the most reliable with a maximum delay of just over 10 min. Which should be chosen? (Table 6.3)

Table 6.3 Routing models

	Static	Dynamic
Single-objective	Fig. 6.2	–
Multi-objective	Fig. 6.3	Fig. 6.4

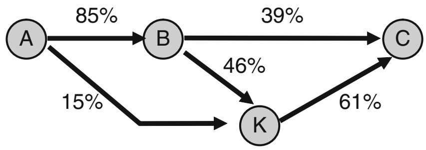


Fig. 6.3 Static multi-objective hyperpath and frequency of use

Obviously, a single objective approach where each criterion is considered individually does not reflect the complexity of choice. Here we use the bi-criterion algorithm developed in the previous section to illustrate the problem. It is first demonstrated in a static case and then applied to a dynamic case with a departure time 8:30.

6.4.2 *Static Multi-objective Hyperpath*

The algorithm presented in Sect. 4 has been applied to the example network static case (input parameters listed in Table 6.2) returning the hyperpath composed of three paths A-B-C, A-K-C and A-B-K-C. The three paths each have different probabilities of use. Users are more inclined to choose link AB when departing the origin. Taking delay into account leads to the total expected travel cost of 30.36 Fig. 6.3.

6.4.3 *Dynamic Multi-objective Hyperpaths*

We first consider two dynamic cases with departure time at 8:30:

(a) Travel time (c) changes over time as indicated in Table 6.1, while d and e assume constant average values as in Table 6.2; and
(b) All parameters (c, e and d) change as indicated in Table 6.1

In case (a) shown in Fig. 6.4a there are two alternative paths included in the hyperpath: Starting at origin node A, the hazmat shipments should be sent via A-K-C in 60% of cases, where the chance of encountering a delay is lower than on the path via A-B-C which should be used for 40% of shipments. Such a strategy leads to the total expected travel cost of 31.2.

In case (b) shown in Fig. 6.4b the number of people at risk in the vicinity of link BK increases with time, leading to a different hyperpath in which link BK is no longer included. At the origin 62% of shipments are directed through link AK, since it has a lower exposure value. The total expected travel cost is 42.60.

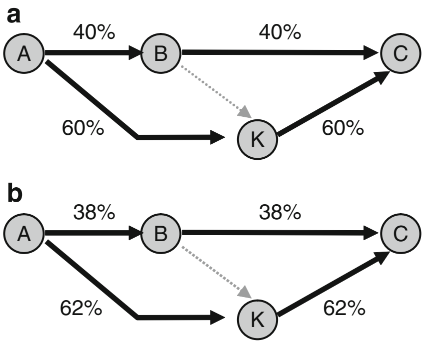


Fig. 6.4 Dynamic multi-objective hyperpaths

6.5 Discussion and Conclusions

In this paper we have presented a new approach to time-dependent vehicle routing for application in repeated transport of hazardous materials. Hazmat dispatchers may prefer a path that simultaneously minimises travel cost, time, distance, accident likelihood and population exposure or any combination of these. This leads to a multi-criteria optimisation problem.

The concept of a mixed-route strategy is introduced in this paper by reference to a game theoretic framework and its equivalence to a linear program proposed by Spiess and Florian [16] is exploited by extending the SF algorithm into a time-dependent multi-criteria setting. Population exposure cost is considered as a route choice criterion concurrently with travel time and maximum delay. As in many multi-criteria analysis frameworks, the objective is not "solving" the problem, but rather providing a reasonable set of alternatives.

The resulting routing strategy (hyperpath) reflects a risk-averse attitude in which a pessimistic view on the possible delay and exposure cost is taken. This implies that the routing strategy of a dispatcher is to minimise the risk of potential maximum delay and potential exposure of the population to any spillage of dangerous goods. Time-dependency is explicitly considered in this paper. Travel times, delays and population exposure change in the course of a day. It is assumed in the model that their values are a function of the time of entry to particular links.

A small numerical example of a network was used to illustrate the implications of adding the time dimension as opposed to using static averaged values on input variables, as well as the implications of introducing an additional criterion to the routing problem.

References

1. Bell MGH (2000) A game theory approach to measuring the performance reliability of transport networks. Transportation Res B: Methodol 34:533–546
2. Bell MGH (2006) Mixed route strategies for the risk-averse shipment of hazardous materials. Netw Spat Economics 6:253–265

3. Bell MGH (2009) Hyperstar: a multi-path Astar algorithm for risk averse vehicle navigation. Transportation Res B: Methodol 43(1):97–107
4. Bell MGH, Kanturska U et al (2008) Attacker-defender models and road network vulnerability. Philos Trans R Soc A: Math Phys Eng Sci 366(1872):1893–1906
5. Bell MGH,Trozzi V, et al. (in press) Time-dependent reverse Hyperstar algorithm for vehicle navigation in road networks
6. Bersani C, Minciardi R et al (2010) Risk averse routing of hazardous materials with scheduled delays. In: Bell MGH, Hosseinloo SH, Kanturska U (eds) Security and environmental sustainability of multimodal transport. Springer Netherlands, Dordrecht, pp 23–36
7. Erkut E (1995) On the credibility of the conditional risk model for routing hazardous materials. Operations Res Lett 18:49–52
8. Erkut E, Ingolfsson A (2000) Catastrophe avoidance models for hazardous materials route planning. Transportation Sci 34:165–179
9. Erkut E, Verter V (1998) Modelling of transport risk for hazardous materials. Oper Res 46:625–642
10. Ford LR, Fulkerson DR (1962) Flows in networks. Princeton University Press, Princeton
11. Nagae T, Akamatsu T (2007) An efficient algorithm for a maximin routing model for hazardous materials. 3 rd International Symposium on Transport Network Reliability (INSTR), Ashgate Publications, Delft, Holland
12. Nozick LK, List GF et al (1997) Integrated routing and scheduling in hazardous materials transportation. Transportation Sci 31(3):200
13. Opasanon S, Miller-Hooks E (2006) Multicriteria adaptive paths in stochastic, time-varying networks. European J Operational Res 173:72–91
14. Schmöcker JD, Bell MGH et al. (2009) A game theoretic approach to the determination of hyperpaths in transportation networks. Selected Proceedings of the 18th International Symposium on Transportation and Traffic Theory (ISTTT), Springer, Hong Kong
15. Sivakumar RA, Batta R et al (1993) A network-based model for transporting extremely hazardous materials. Operations Res Lett 13:85–93
16. Spiess H, Florian M (1989) Optimal strategies: a new assignment model for transit networks. Transportation Res B: Methodol 23B(2):83–102
17. Sung K, Bell MGH et al (2000) Shortest paths in a network with time-dependent flow speeds. European J Operational Res 121(1):32–39
18. Szeto WY, Sumalee A (2009) A game theoretic approach to routing and scheduling hazardous materials in transport networks with multiple origin-destination pairs. 88th transportation research board annual meeting Washinbgton, DC
19. Tarapata Z (2007) Selected multicriteria shortest path problems: an analysis of complexity, models and adaptation of standard algorithms. International J Applied Math Comput Sci 17(2):269–287

Chapter 7
Adaptation of Graph and Game Theories to Reliability Problems

Natalia Yankevich

Abstract This paper is dedicated to application of Graph and Game theories regulations for investigation of reliability of complex technical systems (vehicles). Special efforts are attached to application of system approach, taking into account interaction of details, included in it. Such problem definition supposes consideration of a vehicle as a system, including a number of elements (as elements may be examined subsystems as a whole), connections between them as well as structural functional characteristics, having probabilistic character as a rule. Method, based on game theory, for getting the most reliable distribution low of technical system failures as a linear combination of known ones is developed. Such combination of statistic data and modern mathematic theories can create new possibilities in reliability analysis for creation of modern active safety systems, based on preventive diagnostics.

7.1 Introduction

The objectives of European Union (EU) sustainable transport policy are that its transport system meets society's economic, social and environmental needs. Effective transportation systems, having significant impact on economic growth, social development and the environment, are essential to Europe's prosperity. The transport industry accounts for about 7% of European GDP and for around 5% of the employment in the EU. It is an important industry, which makes a major contribution to the functioning of the European economy as a whole. Only congestion of freight vehicles is causing an estimated loss of 1% of the EU's GDP per year [1, 2].

N. Yankevich (✉)
Joint Institute of Mechanical Engineering of National Academy of Sciences of Belarus, Laboratory of Working Processes and Safety of Technical Systems,
12, Academicheskaya str, 220072 Minsk, Belarus
e-mail: lab_12@tut.by

E. Garbolino et al. (eds.), *Transport of Dangerous Goods*, NATO Science for Peace and Security Series C: Environmental Security, DOI 10.1007/978-94-007-2684-0_7,

The volumes of internal and international trucking industry grow steadily. Even World Crisis does not change this tendency in much degree. It connected with energization of transport market, rising of turnover of goods and passengers and features of economic and geographic location of Europe. The expert prognosis says that the trucking industry in Europe can grow on 1/3 during the period 2004–2015. At the same time the volume growth of Europe – Asia trucking industry may run up to 80–150% per year. Under the aforesaid, it easy to predict the situation, when the transport motion intensity will increase to such level, that any barrier on the way (little damage, repair work, etc.) may be a reason of big traffic jams. All this leads to time and funds losses, in most degree lowers the effectiveness of transport work and makes worse the ecological situation. Recent fact is especially important for big towns, since in the places of traffic jams appearing the concentration of exhaust gases increases in few times more. That is why development and introduction of technologies, allowing realization of real-time reliability analysis of transport means, are actual problem.

One of the most frequently occurring reason of big traffic jams or damage situations is a defective car. So improving of road motion predictability and making transport motion more "self-adaptive" may be very important for road safety increasing. One of possible approaches is based on the following algorithm. If a diagnostic system of a vehicle fixes abnormality in sensors data, the reasons and consequences of this situation must be analyzed by special algorithms and a conclusion about possibility of further motion of a vehicle must be made. If further motion of a car is dangerous, the nearest cars are alarmed about possible dangerous situation by means of radio-channel and then the motion of vehicles (dangerous and nearest one) is corrected in such a way, that dangerous car moves to roadside. At the same time the information about technically defective car can be transmitted to a Service Center for quick help for driver and (if necessary) re-orientation of transport flows.

This charter is devoted to development of foundations for intelligent analysis of technical station of a vehicle, including:

– neural net for analysis of possible failures (causes, consequences, their possible importance for further exploitation);
– method for estimation general distribution of failures of vehicle details during simultaneous action of few lows of loadings.

7.2 Analysis of Vehicle Failures on a Base of Graph Theory

7.2.1 General Regulations of System Analysis in Reliability

General approaches and restrictions, used in reliability analysis of complex technical systems, were determined in the course of numerous investigations [3]. Informative base of developed models is formed on basis of condition of system insularity, which is mean, that behavior and station of a technical system are caused by structure of interaction of its component processes. At the same time availability of closed boundary doesn't exclude influence of some external factors upon it.

Mathematical models in reliability can be divided into four classes:

1. **Determinate models**, which are presented in the form of equations and inequalities, describing a behavior of a complex technical system or its elements. Such models can be named as record one.
2. **Optimization models**, containing expression, which it is necessary maximize or minimize under some restrictions. So optimization prescribes the best policy, such models are called normative.

   All methods of optimization are based on conception of profit preference. As a rule, in decision-making practice all alternatives accessible to a system are considered. Then a variant, giving the most result (efficiency) or having the most utility must be chosen. Another words, it is necessary:

   – to define preferences between different alternatives;
   – to attribute profits to these preferences, indicating how much in numerical expression one preference is bigger then other.

   Mentioned profits define an objective function, which is defined on set of alternatives. An objective function is determined in such a way, that more preferable objects receive more high priority, and objects of equal profit - the same.

   Special difficulties can be connected with definition of a group profit. So if it is firstly defined, that the group profit of three apples and one pear is higher, than group profit of two apples and one pear, we have no reasons for conclusion, that first combination has higher group profit than two apples and two pears. In any case, definition of group profits for complex technical system is directly connected firstly with determination of its individual profits of its subsystem.
3. **Probability models**, which can be represented in a form of equations and inequalities, having probabilistic sense. For example, theory of solutions, being a branch of optimization theory, deals with just maximization of average value of profit.
4. **Complex models**, which include traits of determinate, optimization or (and) probability models.

During development of mathematical models, reflecting real phenomena, the most difficulties appear, when it is necessary to describe existing ideas and conceptions by precise correlations, in other words to formulate a problem correctly. Execution of mathematic investigations of existence and uniqueness of a solution, definition of its properties and construction of a solution are the next steps.

For all evidence of a method itself formal approaches in complex system investigations are not high-developed for this time. At that it is significant, that in many respects purely mathematical modeling is based mainly on experimental data and so can not be a panacea. Therefore, let us mark out those aspects, for which mentioned models can help in selection of reasonable solutions.

It can be stated, that mathematic formulation itself of a problem of investigation of complex system reliability already carries information about selection of possible issues and optimal strategies.

Let us formulate the reasons by which mathematical model construction can be troublesome during analysis of reliability of complex technical systems:

– structure of a system is very complicated;
– structure of a system is clear, but it includes vagueness, so correspondent probabilities can't be estimated.

At the same time it is essential to distinguish sharply necessity and probability. Probability takes place, when values of some parameters are known in probabilistic sense. In this case structure of a model must reflect this peculiarity. Such fickleness, connected with probability, must be considered in received solutions. But vagueness itself means lack of understanding of a problem in question or uncertainty of different factors interaction [4].

Explanation of mechanisms of system behavior and its stations are firstly connected with necessity of clarification of system characteristics (integral features, in what degree, by what conditions). So first of all it is necessary to specify a structure and properties of a technical system with the help of aggregates of models having different organization. This way of concrete definition of abstract system is effective in case of appearing of difficulties during specification of knowledge about system properties with the help of detailed elaboration of special structures, representing these properties. This approach can be realized in hierarchical structural model. However notwithstanding its obviousness, this model doesn't carry sufficient information for describing of reliability of functioning system.

At the second stage the models in the form of graphs of states (nets or failures tree) are developed for solution of such problem. Every graph branch represents failure station of a separate element, which can be a reason of a failure of a system as a whole. For this case parameters often are specified on a base of expert estimates. Grounding on these values the analysis of functioning of a system under consideration is carried out, having as a result conditionally optimal values of some its characteristics (for mechanical systems, for example, the quantity of failures of every subsystem or a whole system). A number of researches call in question a value of such approach, because of it would be reasonable to speak not about conditional optimal solution, but about compromise one, being not optimal, but acceptable in a diapason of solutions. But the compromise theory is not for a while yet developed enough. So the method of expert estimates remains the most usable approach now, in spite of the fact that its application can lead to unjustified overstating of risk estimate and as a consequence - to unjustified costs on prevention of improbable failures or failures what don't lead to serious consequences.

There exist the other approach, based on quantitative models, intended for more precise definition of the average risk as a function of probabilities of system failures and heaviness of their after-effects. A quantitative assessment of contribution (weight) of every failure in forming of general risk valuation may be obtained. Application of such quantitative models is oriented on use of probability conception for description of vagueness of different natures, including engineering.

Definite a priori information as well as results of observations on single situation (object) is used for construction of probability models and selection of their parameters. Probability estimates, relating to the concrete situation (object), are established on this base.

It must be mentioned, that numerous probability problems lead to so-called scheme of independent tests, described by the normal law. Normal distribution is of the great importance in probability theory and mathematical statistics [2]. This is explained by the fact, that in accordance with limiting theorems distributions of those random quantities, which are formed under the influence of numerous independent factors (at that influence every of such factor is low) are approximated by normal distribution. At the same time, one ought not to overstate its field of application. Cited references to satisfactory approximation of a central part of distribution of empiric data by normal low (region of not too large and not too small probabilities) are not well-founded because of namely non-concerned regions are of great importance in the reliability theory.

There is supposition, abundant in scientific papers devoted to reliability problems, that non-failure operating time is a quantity distributed according to exponential law. In this case forthcoming character of failures of a technical system will never change up to the end of its exploitation. This property of invariance (antecedent use of a product does not influence nowise on its capacity for work in the subsequent time) defines the adaptability boundaries of the distribution sufficiently fully. So property of invariance is unacceptable for elements, in exploitation process of which take place irreversible ageing processes. At the same time for a first approximation it is possible to consider a complex technical system, composed of many elements with non-exponential distributions of uptime, as a system with exponential distribution of uptime [5].

It must be underlined, that there is not necessity to overstate a role of mathematician theories for getting of a result. Essence of these approaches is in elaboration of search strategy, able to lead to the result. So, counting of chances in some gambling game does provide neither win nor defeat - it only assists for choice of a behavior strategy in sets. At that a gambler understands situation, but can not change it for his benefit in every concrete set. Information of this kind can be of great significance only in strategy choice.

Therefore further text may be regarded as instruction or more strictly as support in development of decision-making methodology based on results of investigations of complex system reliability.

7.2.2 System-Defined Approach in Analysis of Vehicle Failures

System-defined approach can be a base of analysis of failures of a vehicle, regarded as a complex cascading system [6]. Presentation of a vehicle as a system allows, on the one hand, to divide its projection on steps, on the other hand - to take into account existing ties between its components, determining capacity for work of whole construction. It is evident, that a vehicle is a complex system, analysis of which is possible only on a base of detailed examination of cause-and-effect relations between its components. At the same time, superfluous detailed drafting and concretization may not only make difficulties in much degree, but make this analysis practically impossible. So it can be certified, that during modeling of such system as a vehicle it is inexpediently to create one universal model, which would be able to reproduce functioning of a whole system, as well as functioning of its separate components. With the aim of prognosis of reliability of functioning of a vehicle as a whole it is considered, that it is possible to create a complex of probability models, determining each of its components as well as a model, determining their interaction.

Development of functional model of a vehicle can be realized into three directions, defined by hierarchical level of system elements:

– level of components (an engine, a transmission, etc.);
– level of component subsystems (e.g. for an engine – coal-handling system, starting system, etc.);
– level of details (e.g. crankshaft, connecting-rod, etc.).

Such approaches can be realized on a base of representation of vehicle failures as a **net graph** or **failure tree**, at that every approach has its own advantages and defects. Both of these methods give possibility of wide using of experimental data about vehicle failures, but the second approach is more cumbersome and complicated. Furthermore, net analysis gives integrated presentation about interdependency of vehicle components and general representation about reasons and consequences of vehicle failures. So in what follows we will keep namely to it.

Investigation can be realized in accordance with the following scheme:

- **the first step**
  Vehicle components are considered as graph nodes, interaction between them – as ties, failures, induced by cause-and-effect relations between components -as values of section flow, failures of namely components themselves – as carrying capacities of vertexes, permissible failure number for system – as total flux.
- **the second step**
  Subsystems of vehicle components are considered as graph nodes, interaction between them – as ties, failures, induced by cause-and-effect relations between subsystems within every vehicle component -as values of section flow, failures of namely subsystems themselves – as carrying capacities of vertexes, permissible failure number for every subsystem – as carrying capacity for correspondent component, considered on the previous step.
- **the third step**
  Details of each subsystem of vehicle components are considered as net graph vertexes, interaction between them – as ties, failures, appearing during their interaction within subsystem – as values of section flow, failures of namely details themselves – as carrying capacity of vertexes, permissible failure number for every subsystem – as carrying capacity for correspondent subsystem, considered on the previous step. On this stage reliability estimation for each detail of a subsystem is developed, basing on new or existent approaches. This stage is the

most laborious, because of it demands knowledge of physical phenomena, which are stimulated and accompanied failure appearing.

Analysis failures of a vehicle as a complex system can be regarded on this stage as accomplished. Synthesis is the diverse process to the stated above.

When requirements on vehicle failures are established, it is necessary to take into account complexity of the problem, namely, to separate reliability specifications for all system, its components and subsystems, and single details and mechanisms.

It must be underlined, that usually practiced expert estimation of reliability specifications for a vehicle as well as its components, subsystems and details, based only on engineering practice and exploitation experience is the most wide-distributed and simple approach. At the same time, criteria of improvement of reliability indices for different subsystems, components and vehicle as a whole must be at least coordinated, since in reality increasing of reliability indices of one subsystem (component) can lead to decreasing of analogous indices for other subsystems (components).

Usually used in a number of cases specification of reliability indices of a vehicle, based only on analysis of available statistic information of existing objects with indices, similar to parameters of model under consideration, is in most degree only prognosis, indicating orientation of its technical perfection. However such prognosis at its own heart is based on data, obtained during mentioned above investigations, so both this approaches can be united.

Investigations of vehicle failures, based on system approach as well as net analysis, make it possible to specify optimal level of reliability indexes for a vehicle as well as for its components and subsystems. At that one can say not only about increasing of reliability indexes of all components and subsystems of a vehicle as usually accepted (and so – increasing reliability of a vehicle as a whole), but about reasonable (with relation to field-performance data) redistribution of reliability indexes between components and subsystems.

7.2.3 Application of System Approach to Failure Analysis by the Example of an Engine of a Vehicle

7.2.3.1 Reliability Analysis of Non-recoverable Engine

General Approach

With the aim of simplicity during explanation of stated theoretic foundations the concrete examples of analysis of failures are given only for vehicle engine. Since this approach is universal, it can be extended on failure analysis of a vehicle as whole.

Presentation of theoretical model for estimation of failure stations of an engine by means of a failure tree restricts possibilities of taking into account impact of

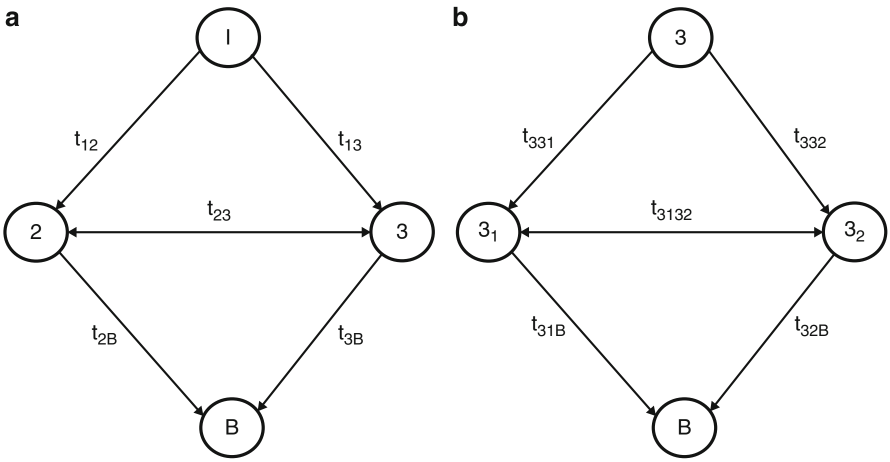


Fig. 7.1 Flow diagram for failure analysis: (**a**) nonrecoverable engine (1 – source, 2 – piston-cylinder-unit, 3 – crank mechanism, B – drain); (**b**) crank mechanism of a nonrecoverable engine (3 – source, 3_1 – a connecting-rod, 3_2 – crankshaft, B – drain); t_{ij} – failure quantities on each tie

Table 7.1 General reliability indices for a party of non-recoverable engines

Parameter	General formula	Specified formula for a case under consideration
Probability density	$f(t)=\lambda \exp(-\lambda t)$	$f(t)=1,11 \exp(-1.11{*}t)$
Distribution function	$F(t)=1-\exp(-\lambda t)$	$F(t)=1-\exp(-1.11{*}t)$
Reliability function	$R(t)=\exp(-\lambda t)$	$R(t)=\exp(-1.11{*}t)$
Failure rate	$\lambda(\tau)=\lambda$	$\lambda(\tau)=1.11$
Expectation value	$E(T)=1/\lambda$	$E(T)=0.90$
Dispersion	$D^2(T)=1/\lambda^2$	$D^2(T)=0.81$

driver and environment. At the same time it was established, that influence of construction and technological differences to reliability indexes of serial engines is insignificant as against influence of operational factors [3].

Thus, reliable analysis of engine failures is possible only on a base of investigation of a system "environment - driver - engine". That can be realized by construction of a net model (Fig. 7.1) by means of introduction of graph node, describing the influence of external impacts. Standard problems, solved with the help of graph theory, can be formulated for this graph. However, we'll dwell on problem definition, allowing take into account probability character of fatigue damage accumulation during operational process of an engine.

Let parameters t_{ij}, representing failure quantities on each tie (Fig. 7.1a), are distributed by normal low (see Sect. 1.1.2). Since an engine can be reconstructed practically after all its failures, so suggested flow diagram for failure analysis of a nonrecoverable engine (Table 7.1) is very simple. As a rule, failures of a such engine are connected with fault of a piston-cylinder-unit or crank mechanism.

At that time it must be marked, that tie between graph nodes 2 and 3 is two-way in common case, so the question about its orientation is solved separately in every concrete case.

If we approximate a distribution of t_{ij}, a standard formulation of a problem, related to graph theory, can be obtained. According context in question it can be formulated in the following way. If an engine is exploited in accordance with existing technical requirements, the following problems can be regarded:

– search of optimal distribution of failures between subsystems (details);
– reduction of general or minimum allowable failure amount for an engine.

Similar net graphs can be constructed for all subsystems of an engine. At that data, obtained during failure analysis of a vehicle as a whole, must be used as input parameters for its engine analysis. As failure analysis of all subsystems is carried out by the same method, let us consider only one of them, namely crank mechanism, which in much degree determines an engine cost (Fig. 7.1b).

Models, describing reliability of every detail, must be regarded after calculations of all subsystems. This step demands other approaches (statistical models, testing, etc.). These approaches are shown in a number of scientific articles, but as a rule, a lot of questions, which can be a base of separate investigations, appear in every concrete case. Some of them will be touched upon below.

Let the reclamation data of separate subsystems and details of a non-recoverable engine are given as a guide. At the same time suggested algorithm is oriented on the ordinary situation, when statistical data about failures are collected for engines during backup service very carefully. Furthermore, suggested approach is common in analysis of failures for any engine, so a character of used data have not an influence upon formulation and solution of a problem.

One suppose, that in 2009, after 1 year after beginning of exploitation of 100 identical engines of one lot, it was be fixed 34 failures of crank mechanisms and 33 one – of piston-cylinder-units. Taking into account, that mean operating time of failures of such complex system as engine is distributed according to exponential low, by simple calculations one can find parameter λ.

It is evident, that the amount of failures of non-recoverable engine is defined by the ratio: $S=34+33=67$ failures by 100 engines. So the probability of failure appearing is equal to: $f=67/100=0.67$. From the other side, substituting $t=1$ in formula for definition of probability density (Table 7.1), we can receive, that $F(1)=1-exp(-\lambda * 1)=1-exp(-\lambda)$.

So the continuous distribution is limiting distribution for statistic data, obtained expressions can be equated. As a result of not complicated calculations we obtain, that $\lambda \sim 1.11$. Corresponding reliability indices are given in Table 7.1. So one can to forecast the failure amount of the engines lot under consideration during further exploitation (Table 7.2).

Let us know, that data about failure reasons are the following (Table 7.3).

Analysing this data with regards of net graph for non-recoverable engine (Fig. 7.1a), one can obtain the net graph (Fig. 7.2). Further investigation is carried out in accordance with graph theory methods.

Table 7.2 Predictable amount of failures in years for a lot of engines under consideration

	Predictable amount of engines failures in years for 100 pcs.				
Subject	2009	2010	2011	2012	2013
Engine failures	67	22	7	2	2

Table 7.3 Reclamation data about subsystems of a non-recoverable engine and failure reasons

No.	Ties and corresponding net graph nodes	Failure reason
1.	1–2	Overheating
2.	1–3	Overheating
3.	2–3	Engine seizure
4.	2–B	Non-recoverable piston-cylinder-unit failure
5.	3–B	Non-recoverable crank mechanism failure

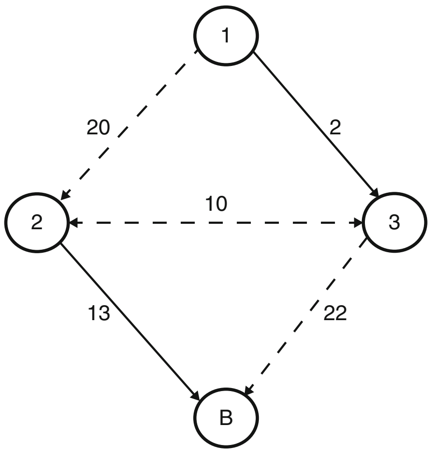


Fig. 7.2 Data flow diagram for investigation of failures of non-recoverable engine (failure data represented according to Table 7.3): 1 – source, 2 – piston-cylinder-unit, 3 – crank mechanism, B – drain; – critical path

It is evident, that in this case total flow along all ties is equal to failure amount of engines lot during the first year of exploitation. With the aim of minimization of failure amount for a engine as a whole system, it is naturally to consider the possibility of decreasing of maximal flow value in obtained net graph. So one can find critical path, computing at the same time maximal and minimal failure amount, as well as free and total floats. It is evident, that path {(1, 2), (2, 3), (3, B)} is critical.

So with the aim of increasing of reliability of engines from a lot under consideration, it is necessary to lower failures along the following paths:

1–2 – piston-cylinder-unit failures, caused by overheating;
2–3 – non-recoverable crank mechanism failure, caused by piston seizure;
3–B – non-recoverable crank mechanism failures.

Table 7.4 Reclamation data of a non-recoverable engine and failure reasons for crank mechanism subsystem

No.	Ties and corresponding net graph nodes	Failure reason	Failure amount, nr.
1.	$3–3_1$	Seizure of a unit "connecting-rod – crank pin"	10
2.	$3–3_2$	Twisting of main bearing cap bolts	2
3.	$3_1–3_2$	Oil starvation	5
4.	3_1–B	Connecting-rod breakage as a consequence of construction imperfection, technology error, etc.	3
5.	3_2–B	Crank mechanism breakage as a consequence of construction imperfection, technology error, etc.	14

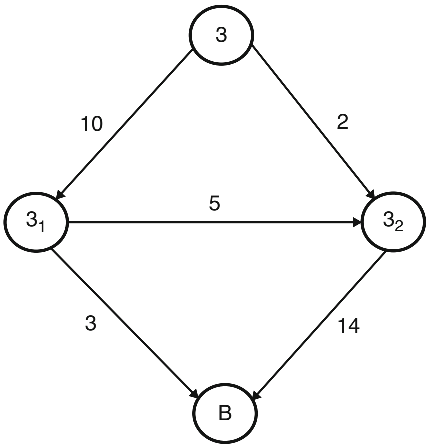


Fig. 7.3 Data flow diagram for investigation of failures of crank mechanism of a non-recoverable engine (failure data represented according to Table 7.6): 3 – source, 3_1 – a connecting-rod, 3_2 – crankshaft, B – drain

After analysis of all components as a whole it is necessary to investigate failure distributions in their subsystems. Let us consider this method, taking a crank mechanism of engine as example.

Failure analysis of a crank mechanism (Fig. 7.1b) will be carried out by analogous scheme. From the data of Table 7.3 it is evident, that for a lot of engines under consideration it was observed 34 events of failures, caused by non-recoverable breakages of crank mechanism, namely: 21 failures of crankshafts and 13 – of connecting-rods. Reclamation data of a non-recoverable engines and failure reasons for crank mechanism are given in Table 7.4. Analysing obtained data with reference to overview graph (Fig.7.1b), we'll obtain net graph as in Fig. 7.3.

Magnitudes of flows along ties $3–3_1$ and $3–3_2$ show the amount of non-recoverable failures of a connecting-rod and crankshaft accordingly. As a such failures can be regarded their fatigue breakings (in particular, abrasion of crankshaft journals is recoverable failure), at that we'll regard, that breakage of a crank-shaft leads to failure of an engine as a whole.

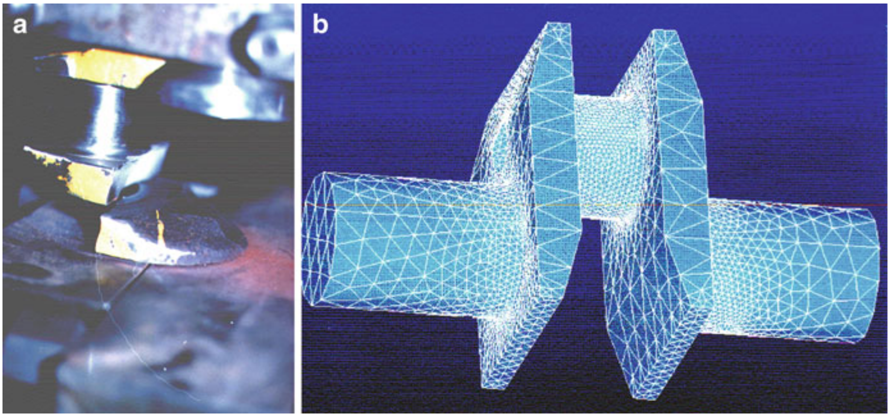


Fig. 7.4 Crankshaft: (**a**) fatigue breaking of crankshaft web; (**b**) finite-element mesh for crank of a crankshaft

Magnitudes of flows along ties 3–3_1 and 3–3_2 correspond amounts of failures of a connecting-rod and crankshaft accordingly, caused by impacts of other subsystem of an engine (total impact of all subsystems). It is evident, that critical path passes over nodes $\{3, 3_1, 3_2, B\}$, at that $T_{critical} = 10 + 5 + 14 = 29$. Consequently, for this lot of engines $N = 29$ is the lower limit of total amount of failures.

It is evident (Fig. 7.3), that from the point of view analysis of non-recoverable failures appearing, the most vulnerability of subsystem "crank mechanism" is a crank-shaft. So investigation methods of fatigue resistance of crank-shafts can be regarded. However described approaches are common for investigations of fatigue resistance of any other details.

Methods of Investigations for Mode of Deformation and Fatigue Resistance of Crankshafts

It is known, that during crankshaft exploitation the following failures appear [7]:

– abrasion of main journal parts and crankpins (recoverable failures);
– fatigue breaking of main journal parts and crankpins (non-recoverable failures);
– fatigue breaking of crankshaft web (non-recoverable failures, Fig. 7.4a).

Calculus of Approximations

Wide application of calculus of approximations for investigation for mode of deformation of details including crankshafts is caused by intense evolution of computer techniques and difficulties appearing during use of analytic methods. It is considered, that in case of adaptation of these methods the problem of estimation of elastic stresses and deformations for details under known loading does not exist more. Different circumstances, reflecting specific of concrete objects, caused appearing of different methods, based on approximate solution of differential equations.

Different types of continuous problem digitization, represented by differential equations, are used.

One of the simplest methods of a digitization is transformation to **finite differences**. In this case a solution of differential equation is changed by solution of a system of algebraic equations. If data array is very big, solution is lighten by algorithm optimization. This approach extends a field of application of finite differences in much degree, but it does not reduce existing difficulties of approximation of curved boundaries and a gradient, prescribed on a boundary. Moreover, obtained matrix of algebraic equations is non-symmetric, that makes solution too intricate.

Variationally-differential method extends possibilities of solution of problems under consideration in much degree. In this case minimization of definite integral is changed by minimization of function of many variables, describing balance of mechanical system. This approach simplify boundary-value problem definition, investigation of equation properties and convergence of solutions. In this case matrix of algebraic equations is always symmetric, that makes process of their solution simpler.

Method of finite elements is one of the most effective approach in view of numerical solutions of problems of elasticity theory (Fig. 7.4b). Its main point is concluded in approximation of a solid under consideration by some discrete model, being totality of elements with finite number of degrees of freedom. This elements are interconnected in reseau points, to which fictitious forces, equivalent to surface stresses, distributed on elements boundaries, are applied [4].

Parameters of such idealized system are defined from correspondent variational solutions. In contrast to variationally-differential approach, interpolation properties of form functions are of great importance in the finite element method. A number of algorithms, demanding in different degree manual labor for input data forming, are developed.

The other methods for investigation of mode of deformation for different objects are developing now. **Boundary equations method** is generally recognized last time. But it is very complicated to define unambiguously its advantages in comparison with finite element method. Owing to sufficiently high developed mathematical and computational base namely this approach is of very wide use in research process.

Application of finite element method during investigation of mode of deformation of different objects (details and model of complex form) leads to acceptable coincidence of obtained results with experimental data in overwhelming majority of cases. At the same time in a number of works it is ascertained supposition, that satisfactory solution for analysis of stress concentration can be obtained only for simple details in elastic problem definition.

As it is known, the main defect of this method is in complexity of receiving of independent estimations. Estimation of reliability of this method can be obtained only by its examination on the strict solutions, as established functions of

Fig. 7.5 Crank model

displacements don't satisfy anytime to demands of continuity between adjacent elements. As a result the conditions of equilibrium can be disturbed at the boundaries of elements, and concentration of equivalent strains in nodes leads to local violation of balance conditions into elements or on their boundaries. All this can be confirmation of thesis about unsatisfactory convergence of design (finite element method) and experimental (polarization-optical method, holographic interferometry, etc.) data. These contradictions have stimulated appearing of combined application of numerical and experimental methods. In this case stresses are defined experimentally on a contour of detailed zones. Further these stresses are used for finite element calculation algorithm as boundary conditions.

Modeling

Guaranteeing of durability of mechanical construction and its elements is defined mainly by actual tests for the time being. However, such investigations are very expensive and in a number of cases – simply impossible. So as a rule researches restrict themselves by results of serial details checkout under development testing. In much degree these questions can be solved with the help of modeling of strain state and fatigue resistance of details with the help of the theory of similarity and dimensions (Fig. 7.5).

Physical as well as mathematical models are used in investigations of different processes. Physical modeling provides for scale change under saving of nature of a phenomenon. Qualitative and quantitative ties of these phenomena are established as criterion limitations, which allow to pick the most essential parameters. Mathematical modeling is based on identity of equations, describing processes of a model and original. Comparison of phenomena, received on such models, with results of corresponding calculations, makes it possible to define more exactly modeling of conditions, provides accuracy rise for reproduction of phenomenon in qualitative as well as quantitative respect. Often analytic solution of such problems is impossible. Moreover analytic solution demands experimental tests on an original or a model. Latest one is achieved simpler and inexpensive by means of investigation of variations of different constructional parameters and other factors, defining character of a process.

Theorems of Newton, Kirpichev-Gukhman, π-theorem and other are usually used for formulation of similarity criteria. But the first and the third theorems, being only necessary conditions, does not indicate methods of showing up a similarity during modeling.

At the same time π-theorem in general form is valid only for processes, which are determined by full functional dependences, but this is often impossible. So the Kirpichev-Gukhman theorem, defining necessary and sufficient conditions for modeling, is of wide use in scientific investigations. However, it is very difficult to apply this theorem to phenomena, for which mechanical saving of a similarity of a model to object does not depend on absolute values of a system parameters. It must be underlined, that the accuracy of similarity criterion affects on character of a process under consideration. So in engineering for establishing of similarity conditions and modeling only part of processes, connected with construction under consideration, is essential, but not all of them. This fact implies, that any simplicity in engineering is approximate.

For the cases, when it is impossible to carry out full-scale experiment for a number of physical (time, cost, etc.) reasons, simulation method, mathematical foundation of which Monte-Carlo method forms, is applied. Solutions, obtained with the help of such models, lead to such big number of different outcomes, so that it is very difficult to interpret received information. Furthermore, analysis of simulation results is based only on mathematical statistics, and thus, demands numerous reiteration of imitating experiments.

It must be marked, that results, obtained with the help of these methods, are very restricted in a number of cases. At the same time, comparison of theoretical results with experimental data can lead to principal new conclusions.

Fatigue resistance modeling of details, based on statistic theory of fatigue failure theory is considered as very perspective. Statistic durability theory of "the most vulnerability", proposed by Weibull, is a base for this theory. According to it reliability of endurance limit of detail is defined by trustworthiness of studying of maximal stresses in stress concentrators of models. At that characteristics of fatigue resistance are defined from the condition of appearance of the first fatigue macro-crack. Further process of a crack development is described by methods of fracture mechanics.

It must be mentioned, that similarity relations, occurring in literature, are taken for restricted number of constructive forms and stress concentrators, as their definitive parameters demands of analytic description of stress fields, what is impossible under complex geometry of a detail.

Brief Review of Existing Experimental Methods for Investigation of Mode of Deformation of Details

For the time being calculus of approximations application does not exclude active using of experimental methods in investigations of detail fatigue resistance. At that, problems, arising during investigation of stressed state in the domain of stress concentration, can be overcome in much degree by combination application of computational and experimental methods. Such approach also supposes partial substitution of an experiment by specified calculation. In this case a selection of experimental methods can not be defined by only one parameter, because in a process of their application it is necessary to find the optimal combination of experimental and computational solutions. So the most known experimental methods of stresses measurement are considered below.

Method of Nets

Mutually transverse lines are marked on detail part in question by photo-method, etching or rolling. These lines (thickness 0.01–0.05 mm and step 0.1–0.5 mm) generate a net with square or rectangle cells. In case of loading sides and angles of cells, marked on elastic material, are changed. This changes are fixed by microscope, and deformation and stresses are calculated on this base with the help of special calibration. Method is widely known, but sufficiently difficult, especially on curvilinear surfaces. All this leads to its comparatively low accuracy.

Method of Moiré Stripes

It is based on determination of displacements and stresses by interference lines (fringes), obtained after superposition of a net, drifted on detail surface before its deformation, and model net. Parameters of material warping are defined on a base of moire fringes picture by special algorithm. Error of this method is defined mainly by possibility of experimental data elaboration.

Method of Optically-Sensitive Coatings

Optically-sensitive material is fixed on a surface of a detail under consideration. It is regarded, that during loading process the deformation of surface and coating is identical. At the same time deformation of a coating is characterized by modification of refraction coefficient. This can be fixed by special apparatus. Method is widely used for investigations of stresses in different constructions, but only for surfaces, accessible for observer. At the same time deformation of optically-sensitive coating can reflect ambiguously a state of a detail.

Method of Marking of Brittle Strain-Sensing Coatings

Specially produced material is coated by thin layer on detail surfaces under consideration. Then detail is loaded by static or dynamic force. At that time cracks appear in a coating, length and density of cracks are defined pro rata to loading growth.

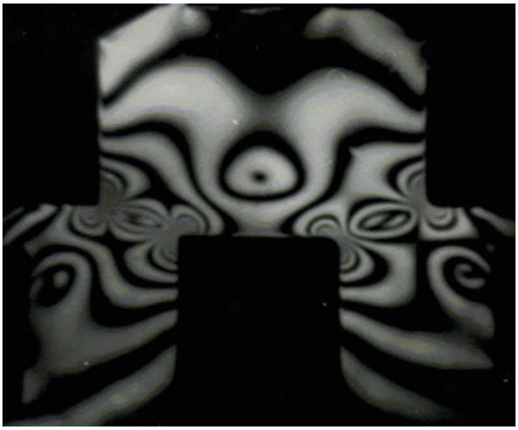


Fig. 7.6 Polarization-optical model of a crank

Method is applied independently as well as jointly with other methods, for example, with the aim of preliminary definition of direction of principal stresses.

Polarization–Optical Method

This method is based on proportional modification of parameters of refraction of light of optically-sensitive material during its loading. Measurements are carry out on models, made mainly from epoxy materials. On a base of measurements with a help of known methods, stressed state of a detail can be defined (Fig. 7.6).

This method provides especially high accuracy for two-dimensional models. However with the help of this method it is possible to determine directly only differences and directions of principal stresses at a plane, which is perpendicular to the beam of polarization light. In a case of three-dimensional stress state method demands additional improvements.

Strain Metering Method

This method is based on change of active resistance during electricity pass through metal under its warping. Method is used for investigation of mechanical behavior of materials under different loading conditions, down to its destruction. Recommendations for application of this method as well as apparatus and resistive-strain sensors are wide known.

Method of Physical Fields of Lüders Lines

This approach is based on determination of electric potential level, appearing on different regions nonuniform deformable steel by means of chemical etching. Method is widely used for definition of parameters of plastic deformation during

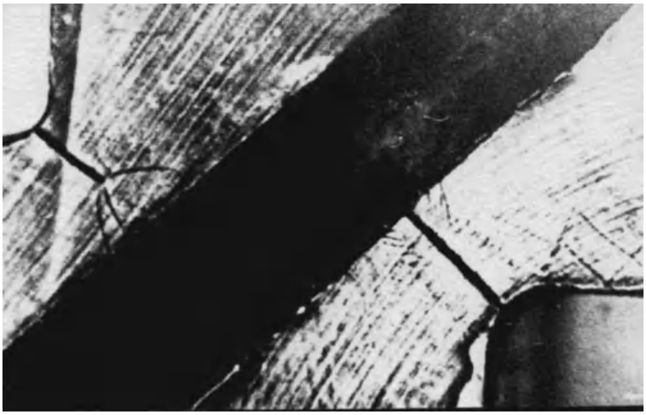


Fig. 7.7 Luder's lines in a crank

static warping mainly on models, which had been made from armco iron. Application of structural and alloyed steels (Cu – 0.15 … 3.0%) extends capability of these investigations in much degree.

Thus, application of this method for investigation of crankshafts capacity for work, produced from mentioned steels, made it possible to estimate kinetics of their destruction in conditions of test benches investigations as well as real exploitation (Fig. 7.7). Parameters of Luder's lines are defined by photos of etching regions. After this parameters of stressed state of a detail in the domain of plasto-elastic deformation can be defined on a base of known methods. Directions and values of maximal tangent stresses in a domain of strengthening are established on disposition of thin Luder's lines (incubation period of fatigue). Crude lines are related to the loosening period, which takes place as a result of intergrain and transcrystalline destruction. State of invariability of Luder's lines parameters in a process of cyclic loading of a detail is defined by malleability of construction, that is plastic deformations, appearing on the first cycles, have no further evolution. Method is very labour-intensive and has not very high accuracy.

Thus, application of this method for investigation of crankshafts capacity for work, produced from mentioned steels, makes it possible to estimate kinetics of their destruction in conditions of test benches investigations as well as real exploitation (Fig. 7.7). Parameters of Luder's lines are defined by photos of etching regions. After this parameters of stressed state of a detail in the domain of plasto-elastic deformation can be defined on a base of known methods. Direction and value of maximal tangent stresses in a domain of strengthening are established on disposition of thin Luder's lines (incubation period of fatigue). Crude lines are related to the loosening period, which takes place as a result of intergrain and transcrystalline destruction. State of invariability of Luder's lines parameters in a process of cyclic loading of a detail is defined by malleability of construction, that is plastic deformations, appearing on the first cycles have no further evolution. Method is very labour-intensive and has not very high accuracy.

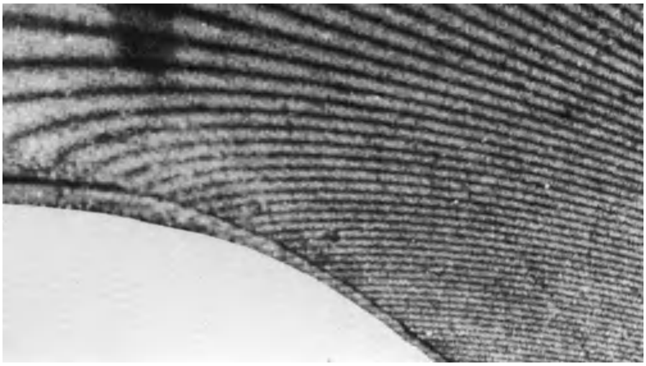


Fig. 7.8 Method of holographic interferometry applied to crank model

Method of golographic interferometry allows to extend interference changes to diffusely reflecting objects. An object is lighted by two optically interfering beams of light. With the aim of receiving of information about displacements of a surface under consideration, obtained speckl patterns, established before and after warping, are compared (Fig. 7.8). High availability of this method for problems of experimental mechanics is marked In a number of papers.

Development Testing

It is usually considered, that the most reliable data are obtained during actual tests in conditions of ordinary exploitation. However if we set oneself the aim of effective rise of construction resource, such data lose their urgency owing to very long duration of process of their obtaining. So development testing with a variable degree of imitation of exploitation conditions are of wide use in such investigations.

Analysis of Results

Let us know, that as a result of carried out investigations, the design solutions, allowing to improve a failure-free operation of a crankshaft, were obtained. At that a number of non-recoverable failures after the first year of exploitation was reduced to 3 (including ties: $3–3_2$—1, $3_1–3_2$—1 and 3_2–B—1 event). Thus net graph will be of the form, showed at Fig. 7.9a.

It is evident, that critical way passes via nodes {3, 3_1, B}, at that its length decreases to 13 units. Reliability of a subsystem "crank mechanism" for this construction of engine as a component of a vehicle is regulated already by reliability of a connecting rod. On necessity of increase of connecting rod fatigue resistance analogues investigations are carried out using methods described earlier.

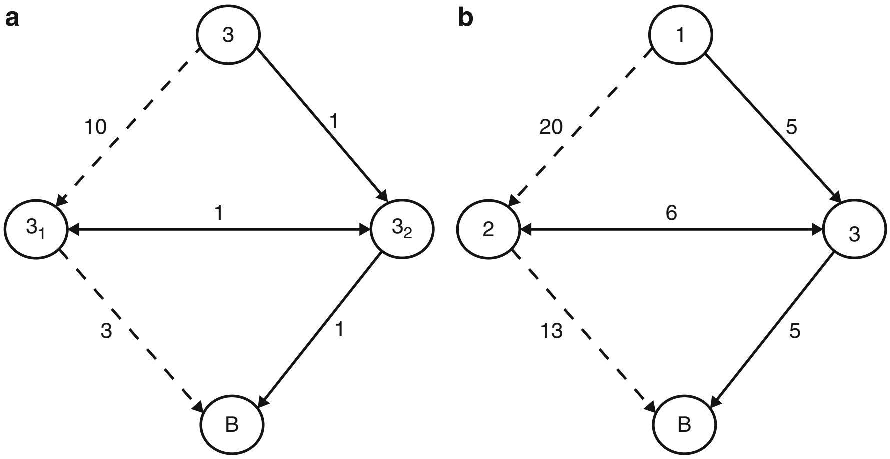


Fig. 7.9 Data flow diagram for investigation of failures of improved non-recoverable engine: (**a**) crank mechanism (3 – source, 3_1 – a connecting-rod, 3_2 – crankshaft, B – drain); (**b**) engine as a component of a vehicle (1 – source, 2 – piston-cylinder-unit, 3 – crank mechanism, – critical path, B – drain)

Table 7.5 General reliability indices for a party of improved non-recoverable engines

Parameter	General formula	Specified formula for a case under consideration
Probability density	$f(t) = \lambda \exp(-\lambda t)$	$f(t) = 0.67 \exp(-0.67 * t)$
Distribution function	$F(t) = 1 - \exp(-\lambda t)$	$F(t) = 1 - \exp(-0.67 * t)$
Reliability function	$R(t) = \exp(-\lambda t)$	$R(t) = \exp(-0.67 * t)$
Failure rate	$\lambda(\tau) = \lambda$	$\lambda(\tau) = 0.67$
Expectation value	$E(T) = 1/\lambda$	$E(T) = 1.49$
Dispersion	$D^2(T) = 1/\lambda^2$	$D^2(T) = 2.23$

Let us consider data flow diagram, describing failures of a component "engine" as a whole. In accordance with Fig. 7.9a it takes a form, indicated on Fig. 7.9b. Critical way passes via nodes {1, 2, B}, at that its length is equal to 33.

General reliability indices of improved engine are shown in the Table 7.5, and comparative graphs for reliability indices – at Fig. 7.10.

7.2.3.2 Analysis of Failures of a Recoverable Engine

General Approach

All mentioned above net models describe failures of non-recoverable systems, to which a vehicle and, in particular, an engine can be related only by convention. A big number of engine failures can be removed by repair. So general model for estimation of reliability of a recoverable engine can be represented as a net graph

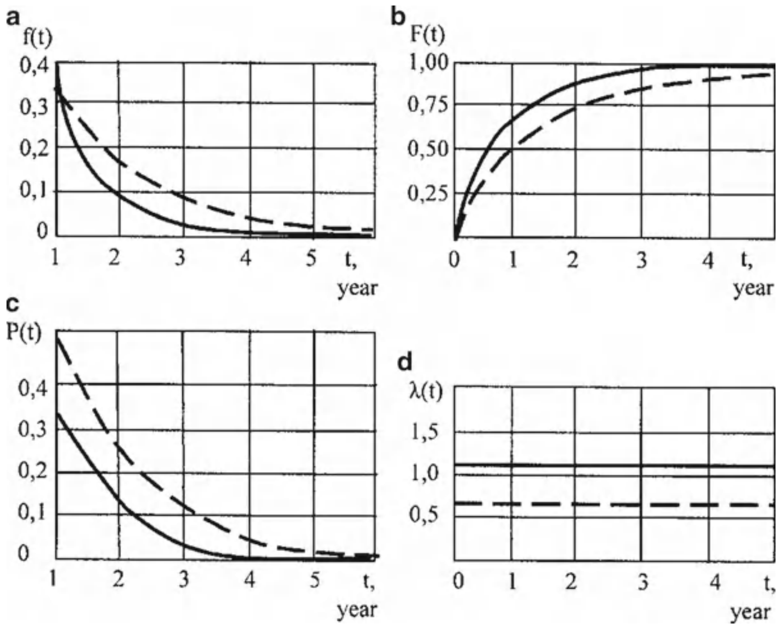


Fig. 7.10 Reliability indices of a lot of non-recoverable engines of initial and improved construction: a – probability density f(t); b – distribution function F(t); c – reliability function R(t); d – failure rate initial construction; – improved one

(Fig. 7.11), where t_{ij} – quantity of failures along every tie, r_{ij} – intensity of failures appearing along every tie. It is evident, that a flow diagram for a non-recoverable engine is a special case of flow diagram for a recoverable engine.

If we can determine parameters t_{ij} and r_{ij} , standard formulation of the theory of planning and management can be obtained (for example, a problem about distribution of limited resources on a net graph).

If an engine is exploited in accordance with technical requirements, than the question about optimality of failure distribution between subsystems and details of an engine appears. Naturally, that concernment in reduction of a total number (or at least minimal possible) of failures exists. Modification of quantity of failures of the separate subsystems (reduction or, maybe, increase) can lead to increase of failures of this component as a whole. Objective consists in determining of such limit failure quantities for every subsystem, for which total quantity of failures does not exceed agreed value R or deviation from this value would be permissible.

Example

Let it is known, that after 1-year exploitation of a lot of engines, including 100 pcs., it was fixed 68 failures of their subsystems:

- control system – 24;
- coal-conveying plant equipment – 14;
- piston-cylinder-unit – 10;
- crank mechanism – 20.

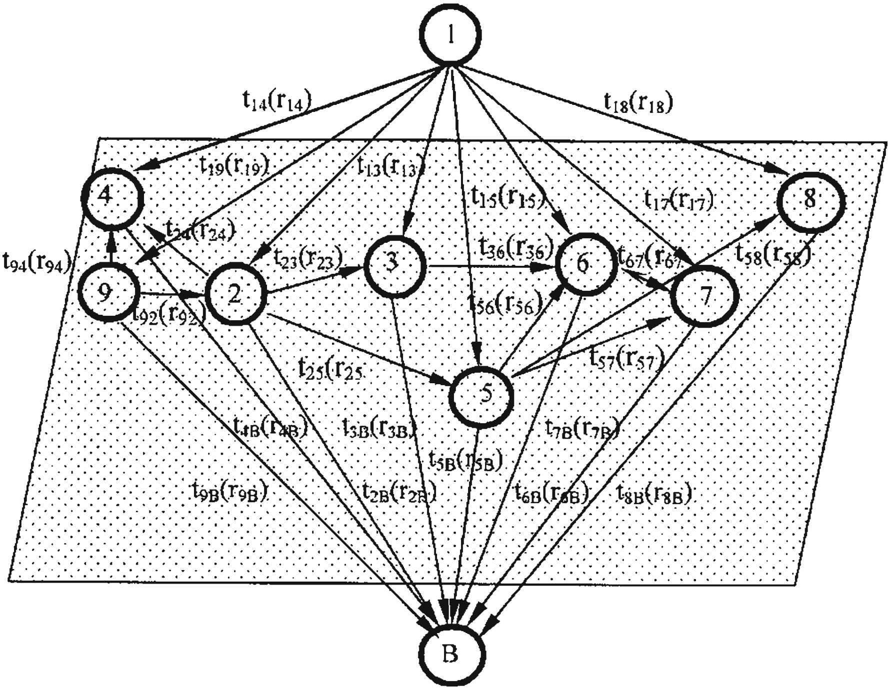


Fig. 7.11 Data flow diagram for failure analysis of a recoverable engine (without turbocharging): 1 – source, 2 – control system, 3 – coal-conveying plant equipment, 4 – starter, 5 – lubrication system, 6 – piston-cylinder-unit, 7 – crank mechanism, 9 – electrical equipment, t_{ij} – failure quantities on each tie, r_{ij} – failure intensity on each tie

Corresponding total time for repair of these failures is equal to 26 days (or 0,071 portion of year):

- control system – 10 days;
- coal-conveying plant equipment – 14 days;
- piston-cylinder-unit – 12 days;
- crank mechanism – 18 days.

It is evident, that calculation of λ is analogous to aforecited (Sect. 1.1.4.1). Coefficient τ will be found in the following way.

By a definition, mean time of an engine repair is equal to a time, expended on reconstruction of a lot of engines, divided on total number of engines failures: $\tau = 0.071/68 = 0.001$.

Let us accept the exponential low for operating time between failures and restoration time of these engines lot. So formulae for host parameters are given in Table 7.6.

From the other side, $\tau = 1/\mu$. Consequently, $\mu = 1/\tau = 1/0.001 = 1.000$.

Taking into consideration, that $\lambda/\mu = 1.1/1{,}000 = 0.0011 << 1$, reliability indices for a lot of engines under consideration are of the form, indicted in Table 7.6. Let us

Table 7.6 Reliability indices for recoverable engine (operating time between failures $F(t)=1-e^{-\lambda t}$, restoration time $G(t)=1-e^{-\mu t}$)

Paramerter, shorthand notation	Strict formula	Approximate formula ($\lambda t_0 << 1$; $\gamma=\lambda/\mu << 1$)
Reliability function in time interval from 0 to t_0, $P(t_0)$	$e^{-\lambda to}$	$1-\lambda t_0$
Probability of failure in time interval from 0 to t_0, $Q(t_0)$	$1-e^{-\lambda to}$	λt_0
Mean operating time between failures, T	$1/\lambda$	–
Mean restoration time of an element, τ	$1/\mu$	–
Stationary availability function of an element, K	$\mu/(\lambda+\mu)$	$1-\gamma$
Stationary stoppage function of an element, k	$\lambda/(\lambda+\mu)$	γ
Intermittent availability function of an element, K(t)	$K+k\,e^{-(\lambda+\mu)t}$	$1-\gamma(1-e^{-\mu t})$
Intermittent stoppage function of an element, k(t)	$k(1-e^{-(\lambda+\mu)t})$	$\gamma(1-e^{-\mu t})$
Stationary operational availability function, $R(t_0)$	$K\,e^{-\lambda\, to}$	$1-\gamma-\lambda t_0$
Intermittent operational availability function, $R(t, t_0)$	$(K+ke^{-(\lambda+\mu)t})^{*}\,e^{-\lambda to}$	$[1-(1-\gamma)e^{-\mu t}]^{*}(1-\lambda t_0)$

know, that data on failure reasons are the following (Table 7.7). Than flow diagram is of the form (Fig. 7.12). This graph is nicely numbered, so further calculations are carried out according to known methods.

It is evident, that ties {(1, 2), (2, 3), (3, 4), (4, 5), (5, B)} belong to critical path, but {(1, 3), (3, B)} – to minimal one. From the very definition of the critical path we can obtain, that the least possible number of failures of a lot of engines under consideration, being subject to repair, is equal to 40 events per 100 engines in 1 year. Let us make corresponding line diagram (Fig. 7.13) and analyze it.

Applying known standard algorithm, one can obtain, that a number of works $\{(P_1–P_3), (P_1–P_4), (P_1–P_5), (P_2–P_B), (P_4–P_B)\}$ can be moved to the right along axis, reflecting failure amounts. This is means, that this construction of an engine allows use of repair elements in a number of subsystems (on condition that a resource of used subsystems will be close to new one).Thus in engines under consideration, having prescribed distribution of failures, it is possible application of coal-conveying plant equipment, piston-cylinder-unit and crank mechanism, having exploitation prehistory (failures, connected with influence of other vehicle components and environmental impact), at that this will not be a reason of subsequent increase of total amount of engine failures.

For example, displacement to the right work $P_2–P_3$ (application of recovered coal-conveying plant equipment, having as exploitation pre-history the failures, connected with control system work), will be a reason of a displacement of the following works:

- $P_3–P_4$ – failure of a piston-cylinder-unit, induced by work of coal-conveying plant equipment; $P_3–P_B$ – internal failure of a subsystem “coal-conveying plant equipment”;

Table 7.7 Reclamation data and failure reasons for subsystems of a lot of recoverable engines

No.	Ties and corresponding net graph nodes	Failure reasons	Failure amount, nr.	Time for repairing, 24 h
1.	1–2	Impact of other components of a vehicle	11	10
2.	1–3	Impact of other components of a vehicle	4	4
3.	1–4	Impact of other components of a vehicle	1	2
4.	1–5	Impact of other components of a vehicle	1	2
5.	2–3	Failure of a coal-conveying plant equipment caused by control system failure	8	8
6.	2–B	Failure of a control system	13	13
7.	3–4	Failure of a piston-cylinder-unit caused by coal-conveying plant equipment failure	2	3
8.	3–B	Failure of a coal-conveying plant equipment	2	2
9.	4–5	Wreck of a crank mechanism caused by piston-cylinder-unit failure	5	6
10.	4–B	Failure of a piston-cylinder-unit	7	7
11.	5–B	Failure of a crank mechanism	14	10

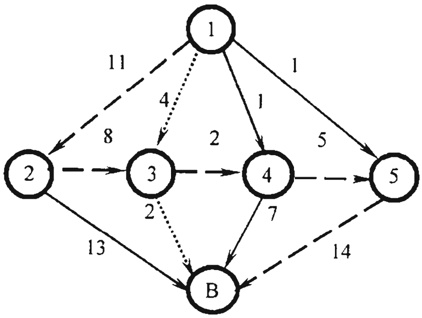


Fig. 7.12 Data flow diagram for investigations of failures of a recoverable engine (without turbocharging): 1 – source, 2 – control system, 3 – coal-conveying plant equipment, 4 – piston-cylinder-unit, 5 – crank mechanism, B – drain, – critical path, – minimal path

- P_4–P_5 – failure of a crank mechanism, caused by wreck of a piston-cylinder-unit;
- P_4–P_B – internal failure of a subsystem “piston-cylinder-unit”;
- P_5–P_B – internal failure of a subsystem “crank mechanism”.

Hence in accordance with accepted assumptions failure appearance in subsystems of an engine is possible. So design study, allowing make competent conclusions about permissibility of redistribution of possible failures of engine, must be carried out after failure analysis of its subsystems.

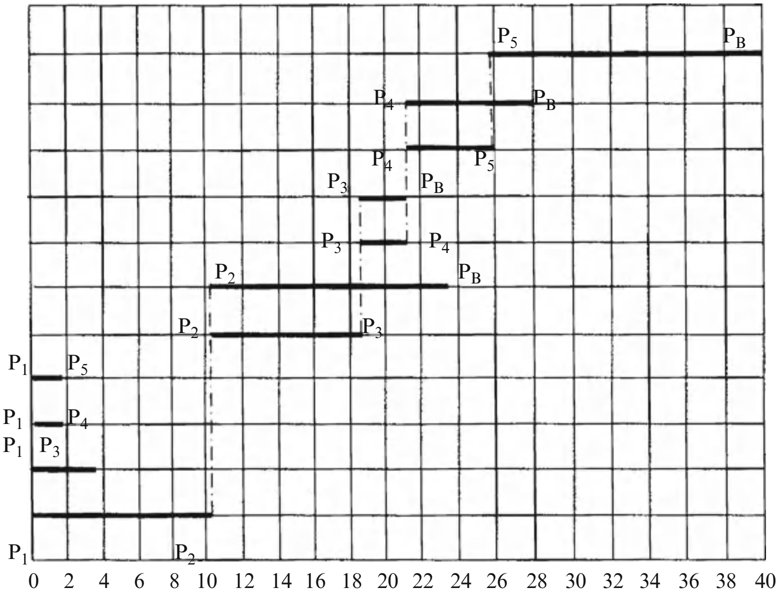


Fig. 7.13 Linear diagram, corresponding for the net graph (Fig. 7.12)

Approaches, considered above, deal with different research aspects of reliability of complex technical systems. It must be underlined, that application of these approaches is not restricted by examined problems. Thus, in direct communication with developed solutions, for example, the problem about search of such distribution of subsystems failures, which is maximal upon exploitation of technical objects during definite time interval. can be regarded. At the same time this principle of analysis of vehicle failure interaction can be called for in a number of applied problems, for example, creation of the modern active safety systems for transport means, based on preventive diagnostics principles.

7.2.4 Application of Game Theory to the Reliability Problems

Last time in a number of reliability investigations significant attention is devoted to development of methods for reliability prediction of details of complex technical systems as a base for reliability analysis of the system as a whole. Usually practiced expert formulation of reliability demands, based only on engineer practice and exploitation experience is not only the most simple, but the most wide-useful approach. As usual, design prognoses, being unconditionally essential, have not so high authenticity (excepting cases, when wide experimental experience had been used). Attempts of some specialists already on designing stage of a complex technical system to predict with high confidence terms of its servicing and overhaul, etc. by calculations are evidence of extreme optimism and must be estimated correspondingly [8]. Suggested experiment-calculated method for estimation of detail reliability is based on a priori information (generalization of statistic data), and the results of supervisions of a unit situation or a single object. Probabilistic estimations, concerning to a concrete situation or a concrete object are calculated on this base.

Let us consider a general formulation of a problem. Suppose, that a factory during k years produce a detail D, which consists of $\boldsymbol{N}$ elements $E_1, E_2, \dots E_N$, that is $\{D\} = \{E_1, E_2, \dots E_N\}$, and failures of $\{D\}$ are described by **n** different lows $\{Z\} = \{Z_1, Z_2, \dots Z_n\}$.

It is evident, that influence of all lows $Z_1, Z_2, \dots Z_n$ can lead to the situation, when Φ(t), representing density of distribution of failures of a detail D, can not be specified as an explicit function.

Let us suppose, that

$$\phi(t) \approx \sum_{i=1}^{n} C_i(t) \cdot \varphi_i(t), \tag{7.1}$$

where $\varphi_i(t)$ – density of distribution of failures, caused by action of a low Z_i, $C_i(t)$ – experiment-calculated coefficients.

Let us consider statistic data (Table 7.8) about exploitation failures of a detail D and its elements ($m_{tl}^{(s,p)}$ – quantity of failures of elements of detail D, t – year of exploitation of a detail, produced in l-th year, s – number of an element, p – number of a low) and mark by $m_t^{(s,p)}$ and M_t – average empiric number of failures for the s^{th} element E_s and detail D correspondingly:

$$\begin{aligned} m_t^{(s,p)} &= \frac{\sum_{l=1}^{k} m_{tl}^{(s,p)}}{k}, \qquad t = \overline{1,j}, \quad s = \overline{1,N}, \quad p = \overline{1,n} \\ M_t^{(p)} &= \frac{\sum_{l=1}^{k}\sum_{s=1}^{N} m_{tl}^{(s,p)}}{k}, \qquad t = \overline{1,j}, \quad p = \overline{1,n}. \end{aligned} \tag{7.2}$$

According to the law of large numbers for sufficiently big j and n the element of fortuity (with which individual characteristics of the detail elements are connected) disappears in summary parameters $m_t^{(s)}$ and M_t. Number sequences $\{m_t^{(s)}\}$, $t = 1\dots j$ and $\{M_t\}$, $t = 1\dots N$ define discrete density of distributions of failures of elements and a detail D as a whole. They can be approximated to continuous with the help of known methods. We will regard fact of appearing of limiting state of a detail in accordance with any low under consideration as failure.

Solution of this problem can be carried out in accordance with game theory – sufficiently new, but thriving part of modern mathematics [9]. In contrast to optimization theory, studying possibilities of the optimal solution construction for all system as a whole, game theory studies optimization methods of individual profit in

Table 7.8 Statistic data on failures of a detail D

Low	Element	Year of exploitation	Year of putting into operation of a detail (technical system 1	2	…	k	Average value
Z_1	E_1	1	$m_{11}^{(1,1)}$	$m_{12}^{(1,1)}$	…	$m_{1k}^{(1,1)}$	$m_1^{(1,1)}$
		…	…	…	…	…	…
		j	$m_{j1}^{(1,1)}$	$m_{j2}^{(1,1)}$	…	$m_{jk}^{(1,1)}$	$m_j^{(1,1)}$
	…	…	…	…	…	…	…
	E_N	1	$m_{11}^{(N,1)}$	$m_{12}^{(N,1)}$	…	$m_{1k}^{(N,1)}$	$m_1^{(N,1)}$
		…	…	…	…	…	…
		j	$m_{j1}^{(N,1)}$	$m_{j2}^{(N,1)}$	…	$m_{jk}^{(N,1)}$	$m_j^{(N,1)}$
…	…	…	…	…	…	…	…
Z_n	E_1	1	$m_{11}^{(1,n)}$	$m_{12}^{(1,n)}$	…	$m_{1k}^{(1,n)}$	$m_1^{(1,n)}$
		…	…	…	…	…	…
		j	$m_{j1}^{(1,n)}$	$m_{j2}^{(1,n)}$	…	$m_{jk}^{(1,n)}$	$m_j^{(1,n)}$
	…	…	…	…	…	…	…
	E_N	1	$m_{11}^{(N,n)}$	$m_{12}^{(N,n)}$	…	$m_{1k}^{(N,n)}$	$m_1^{(N,n)}$
		…	…	…	…	…	…
		j	$m_{j1}^{(N,n)}$	$m_{j2}^{(N,n)}$	…	$m_{jk}^{(N,n)}$	$m_j^{(N,n)}$
Z_1	Detail *D*	1	$\sum_{l=1}^{N} m_{11}^{(l,1)}$	$\sum_{l=1}^{N} m_{12}^{(l,1)}$	…	$\sum_{l=1}^{N} m_{1k}^{(l,1)}$	$M_1^{(1)}$
		…	…	…	…	…	…
		j	$\sum_{l=1}^{N} m_{j1}^{(l,1)}$	$\sum_{l=1}^{N} m_{j2}^{(l,1)}$	…	$\sum_{l=1}^{N} m_{1k}^{(l,1)}$	$M_j^{(1)}$
…		…	…	…	…	…	…
Z_n		1	$\sum_{l=1}^{N} m_{11}^{(l,k)}$	$\sum_{l=1}^{N} m_{12}^{(l,k)}$	…	$\sum_{l=1}^{N} m_{1k}^{(l,k)}$	$M_1^{(n)}$
		…	…	…	…	…	…
		j	$\sum_{l=1}^{N} m_{j1}^{(l,k)}$	$\sum_{l=1}^{N} m_{j2}^{(l,k)}$	…	$\sum_{l=1}^{N} m_{jk}^{(l,k)}$	$M_j^{(n)}$

competition with other persons (events), which rationally aspire to satisfaction of their own profits. Under such consideration the total quantity of failures of detail *D*, defined by influence of lows $\{Z\}$ can be regarded as a win in a game with nonzero sum. Estimation of a win (failure quantity) can be carried out from a point of view of minimax theory (lower limit of estimation – minimal, but guaranteed win), as well as from the point of view of construction of equilibrium solution (such strategy, according to which any attempt of any gamer to change his strategy, when his partner insist on initial choice, will not lead to increase of a win of a gamer, breaking a strategy).

With the aim of the better obviousness of the subsequent text, let us consider case $k=2$ (impact of only 2 lows Z_1 and Z_2). Solution of this problem is analogous to a problem about arbitrary number of acting lows. Let us develop a payoff matrix for this formulation. It must be mentioned, that generally speaking its coefficients depends from a moment of observation, and so it reasonable must speak about a totality of payoff matrixes.

Low Z_2

$$\text{Low } Z_1 \quad \begin{array}{c|cccc} & E_1 & E_2 & & E_n \\ \hline E_1 & \left(m_i^{(1,1)}, \overline{m_i^{(1,2)}}\right) & \left(m_i^{(1,1)}, \overline{m_i^{(2,2)}}\right) & \ldots & \left(m_i^{(1,1)}, \overline{m_i^{(n,2)}}\right) \\ E_2 & \left(m_i^{(2,1)}, \overline{m_i^{(1,1)}}\right) & \left(m_i^{(2,1)}, \overline{m_i^{(2,2)}}\right) & \ldots & \left(m_i^{(2,1)}, \overline{m_i^{(n,2)}}\right) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ E_n & \left(m_i^{(n,1)}, \overline{m_i^{(n,1)}}\right) & \left(m_i^{(n,1)}, \overline{m_i^{(2,1)}}\right) & \ldots & \left(m_i^{(n,1)}, \overline{m_i^{(n,2)}}\right) \end{array} \tag{7.3}$$

Analysis of such payoff matrix is carried out using known methods. Preferences of gamers are usually indicated by arrows (direction corresponds to larger win). Equilibrium point is defined as a point to which vertical arrows (the first gamer preference of the first strategy owing to connected with it larger win) indicate, as well as horizontal one(preferences of the second gamer).

It is evident, that in accordance with such formulation one can find equilibrium point (A, B) by:

$$A = \max_j \{m_i^{(j,1)}\},$$
$$B = \min_j \{m_i^{j,2)}\}, \quad j = \overline{1,n}. \tag{7.4}$$

Then ratio between amount of failures of a detail D, described by lows Z_1 and Z_2 for a concrete time moment will be defined as:

$$C = \frac{A}{B} = \frac{M_t(Z_1)}{M_t(Z_2)}, \tag{7.5}$$

where $M_t(Z_1)$ and $M_t(Z_2)$ are discrete failure densities according to lows Z_1 and Z_2 respectively.

Taking into account that the fact of failure appearance is registered on reaching of limiting state of a detail D according to one of acting lows Z_1 or Z_2 and with respect to (5), we can obtain the formula for density of distribution of failures in the fixed moment of observation:

$$\Omega = \frac{A}{A+B} \cdot M_t(Z_1) + \frac{B}{A+B} \cdot M_t(Z_2). \tag{7.6}$$

In a case of regular observations during a definite time interval (Table 7.9), basing on totality of payoff matrixes we can build piecewise constant function

$$\Phi(t) \approx \frac{A}{A+B} \cdot M_t(Z_1(t)) + \frac{B}{A+B} \cdot M_t(Z_2(t)). \tag{7.7}$$

Table 7.9 Statistic data, describing failures of crankshafts

Reclamation on elements	Failure of element	Failure amount, nr. 2003	2004	2005	2006	2007	2008	2009	Average value
Fatigue breaking of shaft journals in the presence of their wear	Fatigue breaking of shaft journals	100	80	120	110	90	120	80	100
	Wear of shaft journals	500	450	550	480	520	550	450	500
Fatigue breaking of crankshaft webs in the presence of shaft journals wear	Fatigue breaking of crankshaft webs	400	360	440	320	480	420	380	400
	Wear of shaft journals	300	350	250	340	260	310	290	300
Fatigue breaking of crankshaft webs in the presence of crankshaft web wear	Fatigue breaking of crankshaft webs	560	640	600	580	620	610	590	600
	Wear of crankshaft webs	0	0	0	0	0	0	0	0
Fatigue breaking of shaft journals in the presence of crankshaft web wear	Fatigue breaking of shaft journals	430	370	410	390	350	400	450	400
	Wear of crankshaft webs	0	0	0	0	0	0	0	0

Using the function theory, this piecewise constant function can be approximated to continuous.

Guaranteed level and maximin strategies can be defined on lack of knowledge about actions of other gamer. However basing on such method of payoff matrix construction, one can obtain, that the results of maximin theory are identical to the equilibrium point.

The question remains open, in which sense a win and corresponding to it strategy are the best. It is considered, that a gamer guarantees to himself maximin (and possibly, the largest) win, using the equilibrium strategy.

After definition of equilibrium strategy (and so to the maximum guaranteed win, e.g. quantity of failures, determined by action of one or another low), the conclusion about the most significant failures for this construction can be made. In such a way, the most effective direction of investigations for increasing of reliability can be determined.

Example 7.1

A plant produces crankshafts in terms of numbers – 33,000 per year. Control of exploitation failures is carried out during the first year of exploitation in 2003–2009. At that the data shown in Table 7.10 are obtained.

Since during exploitation of crankshafts one can observe fatigue breaking (normal low) as well a uniform wear (uniform distribution), the following failures are the most essential:

- fatigue breaking of shaft journals in the presence of wear;
- fatigue breaking of crankshaft webs in the presence of wear;
- fatigue breaking of crankshaft webs;
- fatigue breaking of shaft journals.

Fact of failure was fixed by sight on reaching of critical value of parameter under consideration (wear value or fatigue crack opening displacement).

Payoff matrix for the first year exploitation of lots of crankshafts taking into account lows Z_1 and Z_2, determining process of appearing of failure by reason of wear and fatigue breaking, is of the form:

$$
\begin{array}{c|l l}
 & \text{Wear} & \\
\hline
 & & \begin{array}{ccc} \text{Shaft journals } (B_1) & & \text{Crankshaft} \\ \text{webs } (B_2) & & \end{array} \\
\text{Fatigue breaking} & \text{Shaft journals } (A_1) & \begin{bmatrix} (100, 500) & \longleftarrow & (400, 0) \\ \downarrow & & \\ (400, 300) & \longleftarrow & (600, 0) \end{bmatrix} \\
 & \text{Crankshaft webs } (B_1) &
\end{array}
$$

Table 7.10 Statistic data, describing failures of crankshafts

Reclamation on elements	Failure of element	Failure amount, nr. 2003	2004	2005	2006	2007	2008	2009	Average value
Fatigue breaking of shaft journals in the presence of their wear	Fatigue breaking of shaft journals	480	520	550	450	420	580	500	500
	Wear of shaft journals	180	220	200	210	190	300	100	200
Fatigue breaking of crank pins in the presence of their wear	Fatigue breaking of crank pins	450	550	500	480	520	430	570	500
	Wear of crank pins	180	20	100	120	80	110	90	100
Fatigue breaking of shaft journals in the presence of crank pins wear	Fatigue breaking of shaft journals	560	220	180	180	220	200	230	170
	Wear of crank pins	350	450	400	320	480	400	400	400
Fatigue breaking of crank pins in the presence of shaft journals wear	Fatigue breaking of crank pins	280	320	300	350	250	220	380	300
	Wear of shaft journals	500	450	550	420	580	500	500	500

It is evident, that the solution can be founded in pure strategies. Saddle point exists for the matrix under consideration – (400, 300). At that win is equal to:

– 400 failures accordingly low Z_1 (m_2^{11});
– 300 failures accordingly low Z_2 (m_2^{12}).

As m_2^{11} and m_2^{12} are guaranteed wins for gamers, it is possible to estimate a ratio of failure amounts according different lows Z_1 ad Z_2, at the first year of exploitation:

$$\frac{p(Z_1(1))}{p(Z_2(1))} = \frac{m_2^{11}}{m_2^{12}} = \frac{4}{3}$$

and to formulate the function, determining a ratio between densities $p(Z_1(1))$ and $p(Z_2(1))$ (or integral functions) of normal and uniform distributions. It is evident, that for $p(Z_1(1)) << p(Z_2(1))$ only $Z_2(1)$ is essential for investigations of reliability of a crankshaft. In the case under consideration we obtain, that density of constructed distribution is defined in the following way:

$$\Omega = \frac{3}{7} \cdot P_{uniform} + \frac{4}{7} \cdot P_{normal},$$

where P_{uniform}, P_{normal} are values of discrete densities accordingly of uniform and normal distributions.

Example 7.2

Let a plant under the same output program of crankshafts per year carried out the reconstruction of its production. As a result a ratio of reclamations on elements of crankshafts is changed. During 1 year after putting into operation of lots of crankshafts, produced in 2003–2009 the following failures were observed:

– wear of main journal parts in the presence of their fatigue destruction;
– wear of crank pins in the presence of their fatigue destruction;
– wear of main journal parts in the presence of crank pins fatigue destruction;
– wear of crank pins in the presence of main journal parts fatigue destruction.

Fact of failure was fixed by sight on reaching of critical value of parameter under consideration (wear value or fatigue crack opening displacement). Statistic data are shown in Table 7.11.

Let us formulate the payoff matrix for the first year exploitation of lots of crankshafts taking into account lows Z_1 and Z_2, determining process of appearing of failure by reason of wear and fatigue breaking.

It is evident, that solution in pure strategies for this matrix does not exist. Let us find this solution in mixed strategies. In order to guarantee equal (or sufficiently close) failure amounts for fatigue destruction (gamer 1) out of dependence of wear (gamer 2), let concentrate our attention on the amount of failures, appearing owing to wear, and so we'll calculate mixed strategies, determined by wear.

Table 7.11 Statistic data, describing failures of crankshafts

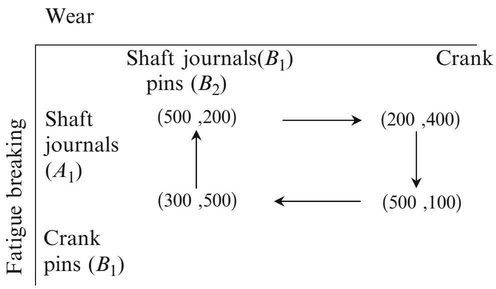
Reclamation on elements	Failure of element	Exploitation year	Failure amount, nr. 2003	2004	2005	2006	2007	2008	2009	Average value
1	2	3	4	5	6	7	8	9	10	11
Fatigue breaking of shaft journals in the presence of their wear	Fatigue breaking of shaft journals	1	100	80	120	110	90	120	80	100
		2	180	220	210	190	170	230	–	200
		3	380	390	370	370	390	–	–	380
		4	1,040	1,000	1,080	1,040	–	–	–	1,040
		5	380	360	400	–	–	–	–	380
		6	220	180	–	–	–	–	–	200
		7	100	–	–	–	–	–	–	100
	Wear of shaft journals	1	500	450	550	480	520	550	450	500
		2	530	510	500	540	520	520	–	520
		3	500	540	460	500	500	–	–	500
		4	500	510	520	510	–	–	–	510
		5	490	490	490	–	–	–	–	490
		6	480	520	–	–	–	–	–	500
		7	480	–	–	–	–	–	–	480
Fatigue breaking of crankshaft webs in the presence of shaft journals wear	Fatigue breaking of crankshaft webs	1	400	360	440	320	480	420	380	400
		2	520	680	550	650	700	500	–	600
		3	860	880	840	860	860	–	–	860
		4	2,080	2,000	2,160	2,080	–	–	–	2,080
		5	860	860	860	–	–	–	–	860
		6	550	650	–	–	–	–	–	650
		7	400	–	–	–	–	–	–	400

(continued)

Table 7.11 (continued)

Reclamation on elements	Failure of element	Exploitation year	Failure amount, nr.							Average value
			2003	2004	2005	2006	2007	2008	2009	
1	2	3	4	5	6	7	8	9	10	11
	Wear of shaft journals	1	300	350	250	340	260	310	290	300
		2	300	280	270	310	270	310	–	290
		3	360	240	300	300	300	–	–	300
		4	270	330	330	310	–	–	–	310
		5	290	290	290	–	–	–	–	290
		6	300	320	–	–	–	–	–	310
		7	300	–	–	–	–	–	–	300
Fatigue breaking of crankshaft webs in the presence of crankshaft web wear	Fatigue breaking of crankshaft webs	1	560	640	600	580	620	610	590	600
		2	890	910	920	880	800	1,000	–	900
		3	1,200	1,320	1,270	1,250	1,260	–	–	1,260
		4	3,000	3,160	3,100	3,060	–	–	–	3,080
		5	1,200	1,320	1,260	–	–	–	–	1,260
		6	1,000	800	–	–	–	–	–	900
		7	600	–	–	–	–	–	–	600
1	2	3	4	5	6	7	8	9	10	11
	Wear of crankshaft webs	1	0	0	0	0	0	0	0	0
		2	0	0	0	0	0	0	–	0
		3	0	0	0	0	0	–	–	0
		4	0	0	0	0	–	–	–	0
		5	0	0	0	–	–	–	–	0
		6	0	0	–	–	–	–	–	0
		7	0	–	–	–	–	–	–	0

Fatigue breaking of shaft journals in the presence of crankshaft web wear	Fatigue breaking of shaft journals	1	430	370	410	390	350	400	450	400
		2	780	820	770	830	750	850	–	800
		3	1,300	1,460	1,200	1,530	1,380	–	–	1,380
		4	4,000	4,200	4,100	4,100	–	–	–	4,100
		5	1,300	1,460	1,380	–	–	–	–	1,380
		6	900	700	–	–	–	–	–	800
		7	400	–	–	–	–	–	–	400
	Wear of crankshaft webs	1	0	0	0	0	0	0	0	0
		2	0	0	0	0	0	0	–	0
		3	0	0	0	0	0	–	–	0
		4	0	0	0	0	–	–	–	0
		5	0	0	0	–	–	–	–	0
		6	0	0	–	–	–	–	–	0
		7	0	–	–	–	–	–	–	0



In compliance with game theory terminology, if gamer 2 choose the first column of a payoff matrix with probability q and the second column – with the probability 1−q, then the expectation value for both rows of payoff matrix of a gamer 1 must be equal, that is: $500*q+200*(1-q)=300*q+500*(1-q)$. Consequently, $q=3/5, 1-q=2/5$.

Thus gamer 1 must choose the first column of the payoff matrix with probability 3/5 and the second column – with probability 2/5. At that anticipated payoff of gamer 1 will be equal to 380. As is easy to see, that this value of anticipated payoff will be saved when gamer 1 uses any mixed strategy (p, 1−p), since

$$500*p*3/5+200*p*2/5+300*(1-p)*3/5+500*(1-p)*2/5=380.$$

Just as gamer 1, desiring to obtain, that the anticipated payoff of a gamer 2 does not depend of choosing by him mixed strategy, with the help of payoff matrix of gamer 2 writes: $200*p+500*(1-p)=400*p+100(1-p)$. Hence, $p=2/3$, $1-p=1/3$.

Then mixed strategy is of the form: $(s_1, s_2)=(2/3*A_1+1/3*A_2)+(3/5*B_1+2/5*B_2)$, and equilibrium gain $(A, B)=(380, 300)$.

Estimating the influences of fatigue destruction and wear, one can obtain:

$$\frac{p(Z_1(1))}{p(Z_2(1))}=\frac{A}{B}=\frac{380}{300}=\frac{19}{15} \text{ or } \Omega=0{,}44\cdot P_{uniform}+0{,}56\cdot P_{normal}.$$

Example 7.3

It is evident, that the problem of finding of distribution density parameters (for example, normal distribution) can be solved only during long-term accumulation of statistic data about failures of different details. So let us suppose, that extending observation to 7 years, one have obtained statistic data about failures of crank-shafts, shown in Table 7.12. Performing elaboration of this statistic data, one can obtain values of discrete densities of distributions (Table 7.13).

Then, after not-complicated calculations in accordance with known methods, general form of densities of normal and uniform distributions can be obtained:

$$\varphi^*_{normal}(t)\approx\frac{1}{\sqrt{2\pi}}e^{\frac{(t-4)^2}{2}}$$

$$\varphi_{uniform}(t)\approx 0{,}14.$$

Table 7.12 Densities of normal and uniform distributions for statistic data (Table 7.11)

		Year of exploitation						
Low	Failed element	2003	2004	2005	2006	2007	2008	2009
Fatigue breaking	Webs	0.05	0.09	0.15	0.42	0.15	0.09	0.05
	Shaft journals	0.06	0.10	0.15	0.38	0.15	0.10	0.06
	Total	0.06	0.09	0.15	0.40	0.15	0.09	0.06
Wear	Shaft journals	0.14	0.14	0.14	0.15	0.14	0.15	0.14
TOTAL		0.07	0.10	0.15	0.36	0.15	0.10	0.07

Table 7.13 Saddle points for different years of exploitation of crankshafts

	Year of exploitation						
Parameter	1st	2nd	3rd	4th	5th	6th	7th
Equilibrium point (A,B)	(400, 300)	(600, 290)	(860, 300)	(2,080, 310)	(860, 300)	(600, 290)	(400, 300)
A/B	1.33	2.07	2.87	6.71	2.87	2.07	1.33
$A/(A+B)$	0.57	0.67	0.74	0.87	0.74	067	0.57
$B/(A+B$	0.43	0.33	0.26	0.13	0.26	0.33	0.43

Table 7.14 Calculation of densities of distribution of failures of crankshaft lot by suggested method ($\varphi(t)$), for adopted for a lot under consideration normal distribution ($\varphi^*(t)$) and experimental data ($\varphi_{exp}(t)$)

t, year	1st	2nd	3rd	4th	5th	6th	7th
$\varphi_{normal}(t)$	0.004	0.050	0.242	0.399	0.242	0.050	0.004
$\varphi_{uniform}(t)$	0.140	0.140	0.140	0.140	0.140	0.140	0.140
$C_1(t)$	0.653	0.718	0.757	0.770	0.757	0.718	0.653
$C_2(t)$	0.347	0.282	0.243	0.230	0.243	0.282	0.347
$\varphi(t)$	0.050	0.080	0.217	0.339	0.217	0.080	0.050
$\varphi^*(t)$	0.004	0.050	0.240	0.399	0.240	0.050	0.004
$\varphi_{exp}(t)$	0.070	0.100	0.150	0.360	0.150	0.100	0.070
$\Delta(\varphi, \varphi_{exp})$,%	28.6	20.0	44.7	5.8	44.7	20.0	26.6
$\Delta(\varphi^*, \varphi_{exp})$,%	94.3	50.0	60.0	10.8	60.0	50.0	94.3

So payoff matrix for every year of exploitation can be formed and saddle points can be calculated foe every year of crank-shafts exploitation (Table 7.14).

We can determine two piecewise constant functions:

$$\frac{A(t)}{A(t)+B(t)} \quad and \quad \frac{B(t)}{A(t)+B(t)},$$

that will define influence of discrete densities of normal and uniform distributions on general function of failure distribution (Table 7.14, Fig. 7.14).

Let

$$C_1 = \psi\left[\frac{A(t)}{A(t)+B(t)}\right] \quad and \quad C_2 = \psi\left[\frac{B(t)}{A(t)+B(t)}\right] \quad -$$

are continuous approximations of mentioned piecewise functions (Fig. 7.14).

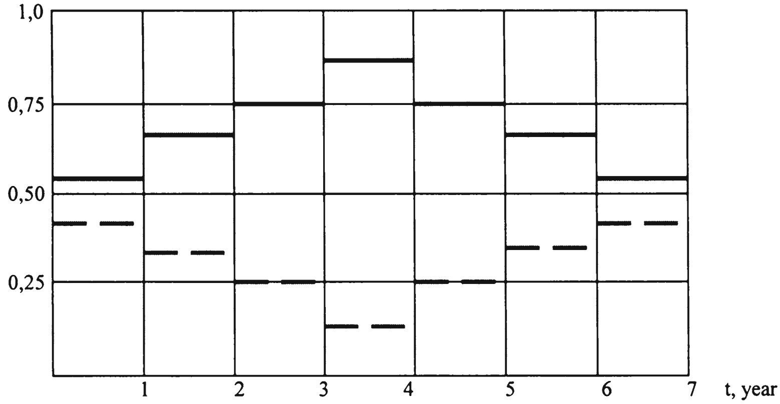
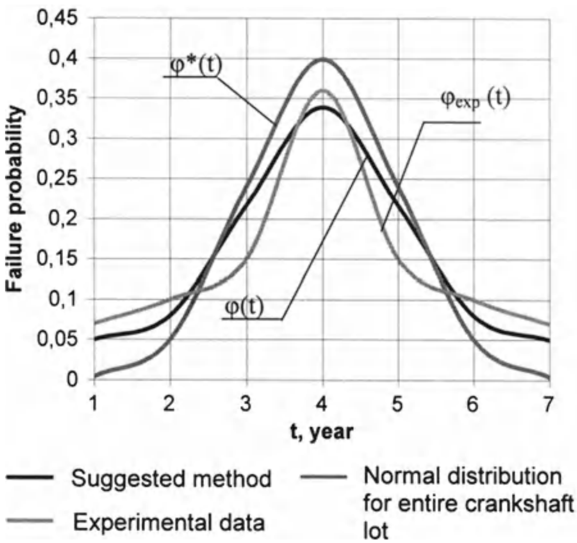


Fig. 7.14 General form of piecewise constant functions: – A(t)/(A(t)+B(t)) and – (B(t)/(A(t)+B(t))

Fig. 7.15 Graphs of densities of distributions of failures of a lot of engines in accordance with: 1 – suggested method (φ(t)), 2 – for adopted for a lot under consideration normal distribution (φ*(t)) and 3 – experimental data (φexp(t))



It is evident, that in this case functions $C_1(t)$ and $C_2(t)$ are close to quadratic function. So they can be represented in the form:

$$C_1(t) \approx 0,77 - 0,013 \cdot (t-4)^2 \quad and \quad C_2(t) \approx 0,23 + 0,013 \cdot (t-4)^2 .$$

Let us construct the function

$$\varphi(t) = C_1(t) \cdot \varphi_{normal}(t) + C_2(t) \cdot \varphi_{uniform}(t).$$

Function $C_1(t)$ and $C_2(t)$ will determine the importance of influence of normal and uniform distribution on general density function. Results of calculations are shown at Fig. 7.15. It is evident, that in the domain of small probabilities normal low, accepted for all lot of engines, has sufficiently high error in comparison with experimental data. At that general form $\phi(t)$ is practically identical to experimental data (Table 7.14, Fig. 7.15).

Suggested approach makes it possible to analyze modification of quantity of failures, caused by different reasons (lows) during definite time limits. Jointly with methods, formulated in Sect. 1, this mathematical formulation can be a basis for development of modern active safety systems for vehicles, based on principles of preventive diagnostics.

7.3 Conclusion

Development of new technologies, allowing optimizing the process of driving, is actual problem for transport motion in Europe. One of the most perspective ways for improving of city logistics is applying of Intelligent Transport Systems (ITS). This is the thrust of the European Commission's "Action plan for the deployment of intelligent transport systems in Europe" – the "ITS action plan" for short – and the accompanying proposal for a directive laying down the framework for the deployment of ITS, both adopted on 16 December 2008. The plan aims to make road transport (including city road transport) more environmentally friendly, more efficient, safer and more secure [1, 2]. So new technologies coming to market in the near future will gradually provide new services to citizens and allow improved real-time management of traffic movements and capacity use for environmental and security purposes.

References

1. European Commission. Directorate-General for Energy and Transport (2006) Road transport policy: open roads across Europe. Office for Official Publications of the European Communities, Luxembourg, 4
2. European Commission (2006) Keep Europe moving – sustainable mobility for our continent – mid-term review of the European Commission's 2001. Office for Official Publications of the European Communities, Luxembourg, 37
3. Bolotin VV, Gusenkov AP, Nefedov SV, Tananov AI (1993) Reliability in engineering. ISTC "Reliability of machines", Moscow
4. Yager RR (ed) (1986) Fuzzy sets and theory of possibilities: the latter advances, Radio i Svyaz, Moscow [in Russian].
5. Bolotin VV, Nefedov SP, Chirkov VP (1993) Reliability in engineering. Methods of the probability theory. ISTC "Reliability of machines", Moscow
6. Bolotin VV, Kovex VM, Nefedov SP (1996) Methods of prognosis of reliability indices. System approach. ISTC "Reliability of machines", Moscow
7. Yankevich NS (2006) Rise of reliability of engines. JIME, Minsk
8. Kugel RV (2003) About balanced life of machines. Machinostroenie, Moscow
9. Saati TL (1977) Mathematical models of conflict situations. Soviet radio, Moscow

Chapter 8
An Expected Risk Model for Rail Transport of Hazardous Materials

Morteza Bagheria, Manish Vermab, and Vedat Verter

Abstract A number of risk measures including those based on incident probability and population exposure have been developed for rail transport of hazardous materials (hazmat). This chapter presents an expected risk model, which incorporates the sequence of events leading to hazmat release from derailed railcars and the resulting consequence, and demonstrates its use on a realistic size problem instance from the United States. It was very interesting to note that although risk models developed for rail are distinct than those for roads, use of different models resulted in different solutions for rail transport of hazmat – much like for road transport.

8.1 Introduction

Hazardous materials (hazmat) are harmful to humans and the environment because of their toxic ingredients, but their transportation is essential to sustain our industrial lifestyle. A significant majority of hazmat shipments are moved via the highway and railroad networks. In the United States, railroad carries approximately 1.8 million carloads of hazmat annually, which translates into 5% of rail freight traffic [1]. On the other hand, in Canada, approximately 500,000 carloads of hazmat – equivalent to 12% of total traffic – are shipped by railroad [2]. The quantity of hazmat traffic

M. Bagheria (✉)
Desautels Faculty of Management, McGill University, Montréal, QC, Canada
e-mail: morteza.bagheri@mail.mcgill.ca

M. Vermab
Faculty of Business Administration, Memorial University, St. John's, NL, Canada
e-mail: mverma@mun.ca

V. Verter
McGill School of Environment (MSE), Desautels Faculty of Management,
Operations Management Unit, 1001 rue Sherbrooke Ouest, Montreal H3A 1G5 Quebec, Canada
e-mail: vedat.verter@mcgill.ca

E. Garbolino et al. (eds.), *Transport of Dangerous Goods*, NATO Science for Peace and Security Series C: Environmental Security, DOI 10.1007/978-94-007-2684-0_8,

on railroad networks is expected to increase significantly over the next decade, given the phenomenal growth of intermodal transportation and the growing use of rail-truck combination to move chemicals. Fortunately a host of industry initiatives, such as the formation of inter-industry task force in the 1970s, and the emphasis on reducing the frequency of tank car accidents, as well as the likelihood of a release, are collectively responsible for making railroad one of the safest modes for transporting hazmat.

In spite of the favourable safety statistic of railroads [3], the possibility of spectacular events resulting from multi railcar incidents, however small, do exist. For example in the United States, between 1995 and 2009, around 120 train accidents resulted in release from multiple tank cars, which translates into an average of eight accidents every year [4]. To that end, we highlight three recent train derailments in North America: first, in January 2005, a freight train carrying a number of hazmat railcars derailed in South Carolina. Three chlorine carrying tankers ruptured and released, resulting in the death of nine individuals, injury to 100 and evacuation of about 6,000 residents. Total damages from this incident exceeded $ 6.9 million [5]. The next two incidents happened in quick succession in western Canada. In August 2005, a freight train derailed in Alberta. A total of 43 railcars derailed, including 27 carrying hazmat. Approximately 800,000 l of different hazmat spilled, into Wabamun Lake, resulting in around $132 million worth of property, environmental and biological damage. Two days later, nine derailed cars released caustic soda into the Cheakamus River in British Columbia [6].

Such episodes had motivated the implementation of a comprehensive safety plan, by the Federal Railroad Administration (FRA) in 2005, for the railroad system. In response to derailments in Alberta and British Columbia, Transport Canada announced a formal review of the Railway Safety Act to consider ways to manage and reduce the risk associated with hazmat transport [7]. Although we provide a detailed literature review in Sect. 8.1, it is important to mention that while the early works dealing with railroad transportation of hazmat extrapolated from highway transportation, most of the later works focused on estimating accident probabilities with little or no consideration to consequence from hazmat release. Note that appropriate safety planning and installation of emergency response system is contingent on level of detail and exactness of parameters used in the a priori risk assessment. Through this research project, we outline a methodology for railroad transport of hazmat that incorporates the characteristics of a railroad accident –namely, accident rate, conditional probabilities of derailment, point of derailment and hazmat release –and the resulting consequence to determine transport risk.

The remainder of the chapter is organized as follows: Sect. 8.1 provides a detailed literature review of relevant works, thereby setting the stage for developing a comprehensive risk assessment methodology in Sect. 8.2; Sect. 8.3 outlines a case example based on realistic-size railroad network, discusses the solution, and then presents some managerial insights into the problem; and, Sect. 8.4 sketches conclusion and directions of future research.

8.2 Literature Review

It was interesting to note that despite the quantity of hazmat moved by rail, an overwhelming majority of research on hazmat transportation focuses on road shipments [8]. The sparse literature on railroad transportation of hazmat mainly deals with analyzing past accident data in an effort to increase railroad safety by improving rail-tracks or railcar tank designs. To the best of our knowledge, no published work incorporates the characteristics of railroad accidents –namely, the sequence of events leading to hazmat release and the resulting consequence –in the development of a risk assessment methodology. Such consideration is important, since a detailed analysis of the freight accident data revealed that rail transport risk depends on: quality of tracks; derailment probabilities for every position in the train-consist; initial point of derailment; number of railcars derailed; and, resulting consequence.

Although railroad transportation has been a popular area of research (see [9] for a comprehensive review), the literature on the use of trains for hazmat shipments is rather sparse. We next trace the two relevant domains of research for our work: hazmat risk analysis and mitigation in railroad transportation; and, application of air dispersion models in assessing transport risk.

The first thread, focused on hazmat risk analysis and mitigation, contains the early works of Glickman [10], and Glickman and Rosenfield [11]. While the former demonstrated that rerouting of trains with/without track upgrades can reduce risk, the latter derived three forms of risk distributions viz., number of fatalities in a single accident, total number of fatalities from all the accidents in a year, and frequency of accidents that result in any given number of fatalities. Barkan et al. [12] conducted a statistical analysis of the railroad accident data to conclude that the speed of derailment and the number of derailed cars are highly correlated with hazmat release, and then proposed estimating direct and conditional probabilities in conducting risk analysis [13]. Most recently, Verma [14] analyzed the FRA accident records to conclude that front of the train is riskier, and that seventh to ninth train-deciles are the safest position to place hazmat railcars for freight-trains of any length. The last 20 years has also witnessed a number of efforts geared towards risk mitigation with a focus on reducing the frequency of tank car accidents and likelihood of releases. While Raj and Pritchard [15] reported that DOT-105 tank car design is safer than the 111 type, Barkan et al. [16] showed that tank cars equipped with surge pressure reduction devices experienced lower release rates than those without. Finally, Saat and Barkan [17] developed a metric to assess the performance of a tank car in an accident, and analyzed the trade-off between increased damage resistance and greater exposure to accidents [18].

A sub-stream within risk mitigation focused on reducing the probability that a hazmat railcar gets involved in a train derailment. To that end, the early work of Fang and Reed [19] suggested that the front of the train is more prone to derailment under loaded conditions, and hence hazmat railcars should be placed in the rear of the train. A later study commissioned by the Federal Railroad Administration (FRA) and the Department of Transportation (DOT) concluded that derailment probabilities are highest in the first and lowest in the fourth quarter of the train [20]. The same report also explored the possibility of implementing commodity-based blocking, as opposed to the destination-based blocking practice of railroad industry, and ruled it out due to time and hence cost considerations. The issue of additional handling and the resulting increase in time and cost was recently visited by Bagheri [21], who then outlined a placement strategy that could be implemented during the railcar blocking process that in turn will minimize hazmat risk on pre-defined transportation corridors [22].

The second thread deals with the application of air dispersion models for assessing hazmat transport risk. The most comprehensive study thus far is carried out by Hwang et al. [23], who used a Lagrangian-integral dispersion model to estimate impact zones for six toxic-by-inhalation materials. In analyzing the chlorine-handling facility in England, Leeming and Saccomanno [24] made use of dense-gas dispersion model to estimate the impact areas stemming from each possible release scenario. The Gaussian dispersion plume model (GPM), however, is by far the most popular dispersion model used by micro-meteorologists, air pollution analysts, and regulatory agencies [25]. Patel and Horowitz [26] were the first to use the GPM, coupled with a geographical information system (GIS), for risk assessment of road shipments under varying wind conditions. In an effort to develop closed-form expressions, they assumed that dispersion parameters are equal to one, and focused on the technical risk by assessing the total expected contaminant concentration due to a potential spill. Later on Zhang et al. [27] modeled the probability of an un-desirable consequence as a function of the concentration level and used the expected consequence representation of transport risk. Most recently, Verma and Verter [28] made use of GPM to develop a methodology for railroad that takes into consideration both the volume and type of hazmat in the determination of transport risk.

As discussed before [21], only focuses only on derailment probabilities, while [28] consider just the consequence. Interestingly, both studies refer to the importance of including the missing component, and to the best of our knowledge this is the first work that incorporates detailed accident probabilities and resulting consequence in the determination of in-transit risk associated with hazmat railcars. To that end, we take into consideration the quality of tracks, varying derailment probabilities by railcar position, initial point of derailment, number of hazmat railcars derailed and releasing, and the resulting consequence in developing a risk assessment methodology for railroad transportation of hazmat.

8.3 Risk Assessment Framework

In this section, we first outline the determinants of railroad transport risk, which are then used to discuss the risk assessment methodology. The sub-section on methodology will highlight the different components of transport risk, and the appropriate evaluation/estimation techniques.

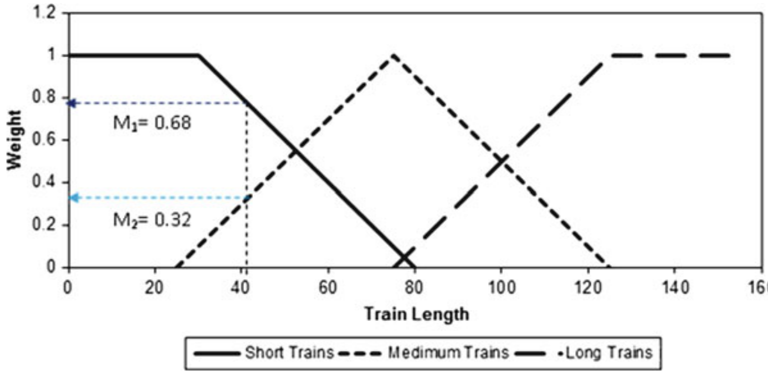


Fig. 8.1 Rail corridor

8.3.1 Determinants of Transport Risk

Based on the work of Bagheri et al. [29], we incorporate rail-track segment and block of railcars to determine transport risk. Consider a railway corridor, composed of six segments, where freight trains originate at the rail-yard (following train make-up) and service demand at the three stations (Fig. 8.1). Railcars move from origin to destination yard in blocks (group of railcars with same intermediate handling points), and hence we assume that a block is set-off at each of the three stations. Although the objective of blocking operations in the railroad industry is to minimize the number of intermediate handling (and hence cost), we are also interested in minimizing the transport-risk (TR) over the given corridor. It should be noted that geometric attributes of rail segment affect the derailment probability profiles and consequently transport-risk. Thus, the rail corridor is divided into segments with homogenous attributes. Hence,

$$TR = R^1 + R^2 + \cdots + R^6 \tag{8.1}$$

where, TR is the transport risk; and R^s is the risk over rail-segments.

It should be clear that transport risk depends on the sequence of blocks, which in turn make-up the train. Based on the existing practice, the sequence of blocks on a given train follows the order in which intermediate destinations will appear on a train route, such that the block designated for the first destination is shunted to the front of the train, followed by the block assigned to the next destination, and so on.

Notice that the first block (destined for station A) traverses over only the first two rail-segments, whereas the second block is moved over the first five segments and the third over all the segments. This implies that

$$R^s = R^{s1} + R^{s2} + R^{s3} \tag{8.2}$$

where, R^{sb} is the risk over rail-segments due to block b. On substituting Eq. 8.2 in Eq. 8.1, we have:

$$TR = \sum_{s}\sum_{b} R^{sb} \tag{8.3}$$

Now, transport-risk of a particular block on a specific rail-segment is jointly determined by all the hazmat railcars in that block. For example, if there are n_b cars in block b, then the corresponding risk over segment s can be determined as:

$$R^{sb} = \sum_{i=1}^{n_b} Risk_i^{sb} \tag{8.4}$$

where, $Risk_i^{sb}$ is the risk of hazmat railcar in position i of block b over rail-segments. The preceding discussion implies that risk from a hazmat railcar varies over different rail-segments, and more importantly across different positions within a block (and hence a train). For exposition purposes, we just focus on railcar position i on a given rail-segment, which simplifies $Risk_i^{sb}$ to $Risk_i$.

8.3.2 *Transport-Risk*

In this section, we outline a comprehensive risk assessment framework that is novel and different than the in-transit risk framework studied in [21]. Our approach takes into consideration not just the sequence of events leading to hazmat release but also the associated consequence, unlike [21] which focused only on derailment probabilities.

For a given rail-segment, risk from a hazmat railcar, $Risk_i$, is defined as the product of derailment probability for position i, P_i, and the expected consequence, C_i.

$$Risk_i = P_i \times C_i \tag{8.5}$$

where, the probability of derailment for position i can be calculated as the product of the probability of train derailment on this segment $\left(P^{TD}\right)$, and the conditional probability that position i is the point where derailment starts (POD) given that the train has derailed P_i^{POD}. Unlike [21], the proposed framework focuses on railcars after the point of derailment in the computation of risk.

$$P_i = P^{TD} \times P_i^{POD} \tag{8.6}$$

and, the consequence of derailment can be calculated as the product of the conditional probabilities that m railcars derail as a result of derailment beginning at

Table 8.1 Derailment rates (Source: [13])

	FRA track class[a]				
Derailments per	1	2	3	4	5 & 6
Million freight *train* miles	48.54	6.06	2.04	0.53	0.32
Billion freight *car* miles	720.10	92.70	31.50	7.80	4.90

[a]The FRA classified tracks based on various quality and speed considerations. Class 1 represents the poorest tracks and wherein speed limit is 16 km/h, whereas Class 6 is the best with permissible speed of 177 km/h

position i, $P(m|i)$, and j hazmat railcars among the m derailed will release, $P(j|m)$, together with the population exposure associated with j railcars, PE_j.

$$C_i = \sum_{m=1}^{n-i+1} P(m|i)\left[\sum_{j=D}^{m} P(j|m) \times PE_j\right] \tag{8.7}$$

Substituting Eqs. 8.6 and 8.7 in Eq. 8.5 results in the complete expression for determining transport-risk on a particular rail-segment (or route):

$$TR = \sum_{i=1}^{n}\left\{P^{TD} \times P_i^{POD} \times \sum_{m=1}^{n-i+1} P(m|i) \times \left[\sum_{j=D}^{m} P(j|m) \times PE_j\right]\right\} \tag{8.8}$$

Next, we will discuss the techniques to calculate and/or estimate various components in Eq. 8.8.

8.3.2.1 Train Derailment

Probability of train derailment is calculated using the model proposed in [13]. Federal Railroad Administration (FRA) accident data from 1992 to 2001 was analyzed to conclude that derailment probability of a train – a function of travel distance, train length, track quality –can be calculated as:

$$\mathrm{P^{TD}} = 1 - \mathrm{e}^{\{-\mathrm{M}[\mathrm{RC(L)} + \mathrm{RT}]\}} \tag{8.9}$$

where, M is the travel distance; L is the train length; RC is the derailment rate per billion freight car-miles; and, RT is the derailment rate per million freight train-miles. The model was based on aggregate data for accident rates for different types of track class in terms of the number of derailments per billion freight car-miles and number of derailments per million freight train-miles (Table 8.1). It is important to note that $\mathrm{P^{TD}}$ will change with rail segment and number of blocks in the train consist.

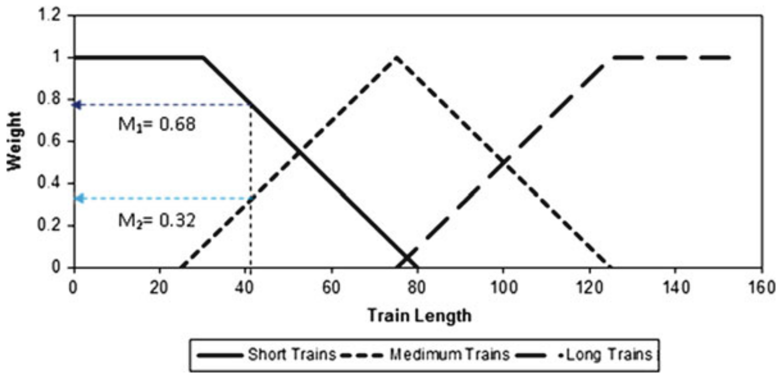


Fig. 8.2 Weights for the train types (*Source*: [22])

8.3.2.2 Point of Derailment (POD)

In a recent work, Bagheri et al. [22] grouped derailment causes into three classes depending on the part of the train likely to derail first. More specifically, causes those affect the front of the train (C_F), the rear of the train (C_R), and the middle of the train (C_M). In addition, trains were categorized into three types: short (up to 40 railcars); medium (between 41 and 120 railcars), and long (more than 120 railcars). Finally, a nonparametric Kruskal-Wallis test was applied to show that train-length and causes provide a statistically significant explanation for median point of derailment (POD). Best fit POD distributions were obtained for all nine length-cause combinations, and different weights were used to account for uncertainty caused by overlap in train-length at the boundaries (i.e., around 40 and 120 railcars). Figure 8.2 can be used to determine the two types of trains under consideration and the appropriate weights.

For example, for a 41-railcar train over a rail-segment subject to cause C_M, the respective weights are $M_1 = 0.68$ for a short-train classification, and $M_2 = 0.32$ for a medium-train classification, and zero for long-train classification. Now the probability of derailment for the tenth position (in a train with 41 railcars) with cause C_M is obtained by applying the following steps: (i) determine the weights for relevant train-lengths (i.e., M_1 and M_2); (ii) compute the normalized point of derailment, which for this example is: 10/41 = 0.244; (iii) obtain the derailment probabilities for the position of interest (for this example, i = 10) the values are $f' = 0.033$ and $f'' = 0.07$ from [22]; and, (iv) substitute values in Eq. 8.10 to obtain P_i^{POD}, which for this example is 0.0448.

$$P_i^{POD} = \frac{f' M_1 + f'' M_2}{M_1 + M_2} \tag{8.10}$$

8.3.2.3 Number of Cars Derailed

In a previous work, Anderson and Barkan [30] suggested that the number of railcars derailing is affected by the dissipation of kinetic energy following a train derailment,

Table 8.2 FRA data causal factor statistical summary (Source: [22])

Parameters	Estimates	Std. error	Z statistics	Lower 95%	Upper 95%
Intercept	–2.013	0.082	–24.46	–1.85	–2.17
Residual length (R_L)	0.001	0.001	1.266	0.002	–0.001
Speed effect (S_E)	–0.032	0.002	–17.07	–0.029	–0.036
Roadbed (R_B)	0.419	0.018	2.367	0.766	0.072
Track geometry (T_G)	0.171	0.089	1.921	0.346	–0.003
Switches (S)	0.715	0.119	6.013	0.949	0.482
General car (G_C)	0.841	0.085	10.132	1.03	0.697
Axles/wheels (A_W)	1.108	0.077	14.404	1.26	0.958
All other (O)	0.444	0.073	6.056	0.587	0.30

which implies that train speed and distance from the point of derailment are important elements. Note that the latter point is relevant because as the distance to the POD increases, forces of instability acting on the remaining railcars decrease. Given that risk is posed by a derailed hazmat railcar, all hazmat railcars placed before the POD or beyond the derailed block do not pose any risk on the given rail-segment. A truncated geometric expression, to estimate the probability that m railcars will derail given that derailment started at position i [31], is proposed.

$$P(m|i) = \frac{p(1-p)^{m-1}}{1-(1-p)^{L_r}} \tag{8.11}$$

where, $m = 1,2,\ldots,L_r$, and L_r is the number of railcars in the train past the point of derailment, and (1-p) is the probability of derailment for a position after the point of derailment.

In an effort to evaluate the significance of other causal factors on the probability of derailment (i.e., p in Eq. 8.11), summary statistics for the FRA database was generated at the 95% significance level. From Table 8.2, it is possible to conclude that number of railcars derailing beyond the POD has a strong association with all the primary causes. In addition, [32] suggested using a logistic regression of the following form to estimate the probability of derailment beyond the POD.

$$P = \frac{e^{\{\beta_0 + \Sigma_k \beta_k X_k\}}}{1 + e^{\{\beta_0 + \Sigma_k \beta_k X_k\}}} \tag{8.12}$$

where, Xk indicates the impact of kth independent factor on the probability of derailment beyond the POD. Table 8.2 depicts the summary statistics for the various causal factors for the 1997–2006 FRA dataset [22].

Table 8.2 shows that at 95% confidence level, each of the eight factors is significant, and that increasing the speed will increase the probability of derailment when using a logistic regression function. In addition, the impact of the factor A_W, representing axles/wheels, on derailment is less compared to that of track geometry.

Table 8.3 Number of derailments resulting in Hazmat release (Source: [4])

Year	Number of accidents	Hazmat involved	Hazmat derailed	Hazmat release
1997	545	152	64	11
1998	620	201	95	23
1999	547	206	102	20
2000	598	202	91	22
2001	627	223	97	20
2002	541	192	95	16
2003	582	196	95	17
2004	594	174	74	14
2005	580	171	69	13
2006	563	204	103	11
Total	5,797	1,921	885	167

8.3.2.4 Hazmat Release

The conditional probability that j hazmat railcars release, given that m has derailed, is estimated from empirical data. For the 10-year period under discussion (1997–2006), around 5,800 freight train accidents of which one-third included hazmat, were recorded (Table 8.3). A total of 885 hazmat railcars derailed resulting in release from 167 (or, 18.8%). It was interesting to note that a number of hazmat release episodes involved multiple railcars, which necessitates taking into consideration potential volume released. For instance, in eight of the eleven accidents in 2006, multi railcar release episodes were recorded. In fact, the worst incident involved release from 20 hazmat railcars.

In an effort to take advantage of the empirical dataset, we propose that the conditional probability of release from a derailed hazmat car (q) be independent of each other, such that:

$$P(j|m) = q \times q \times \cdots .. \times q = q^j \tag{8.13}$$

According to the FRA database (1997–2006), the conditional probability of release a derailed hazmat railcar (q) is 0.0903, obtained at an aggregate level by dividing the number of railcars releasing by the total number of hazmat railcar derailed.

8.3.2.5 Population Exposure

Population Exposure (i.e., total number of people exposed to the possibility of an undesirable consequence) is the last element in Eq. 8.8. We focus on hazmat that become airborne in the event of an accidental release (such as chlorine, propane and ammonia) since they can travel long distances due to wind and expose large areas to health and environmental risks. Spatial distribution of toxic concentration level is estimated using Gaussian plume model, and at any given distance the maximum concentration is observed at the downwind locations [25]. We use immediately dangerous to life and health (IDLH) concentration levels of hazmat being shipped in determining the threshold distances for fatality and injuries (http://www.cdc.gov/niosh). In estimating the population exposure, we adopt the worst-case approach by assuming least favourable weather conditions and focusing on maximum concentrate levels.

As explained in [28], a population center is exposed if the aggregate concentrate level exceeds the critical IDLH level for the hazmat being shipped. To make this explicit, consider that n hazmat railcars are traveling on rail-segments. By making use of GPM and the methodology developed in [28], the aggregate concentrate level at downwind distance y can be determined as:

$$\bar{C}_n(y) = \frac{Q}{\pi u a c x^b x^d} \times n \tag{8.14}$$

where, Q is the release rate; u is the wind speed; a, b, c, and d are atmospheric constants; and, y is the downwind distance from the hazmat median (center of the hazmat railcar block) of the train. At an IDLH level of $\tilde{C}$, the threshold distance can be determined using (8.15):

$$\tilde{X} = \sqrt[b+d]{\frac{n \times Q}{\pi u a c \tilde{C}}} \tag{8.15}$$

The movement of the danger circle, of radius $\tilde{X}$, along the rail-segment will carve out a band, and the number of people within the band is the population exposure because of the release from n hazmat railcars. We invite the reader to consult [28] for methodological details, and [33, 34]34 for application of the risk assessment methodology for solving realistic size problem instances involving hazmat transportation.

8.4 Case Example

8.4.1 Experimental Setting

The assessment methodology developed in the previous section is illustrated through a realistic problem analysis involving rail transportation of hazmat from Barstow (California) to Chicago (Illinois), with two intermediate stations (Albuquerque and Kansas City). Figure 8.3 depicts the three rail-routes recreated, along with population zones, in ArcView GIS [35]. This corridor is interesting since it presents the geographical and geometric track attributes necessary for our analysis, and is one of the primary routes that connect the supply and demand points for hazmat by rail in California and Illinois [36, 37].

Note that each route has a number of segments, including the three common ones (i.e., segments from Barstow to Albuquerue; and, Kansas City to Chicago). Clearly the variance in attributes for the three routes is stemming from the remaining rail segments (Table 8.4). For exposition reasons, we label the three routes as follows: one through Denver and Omaha as the North (N) route; one through El Dorado as the South (S) route; and the third as the Mainline (M) route. The three routes differ just

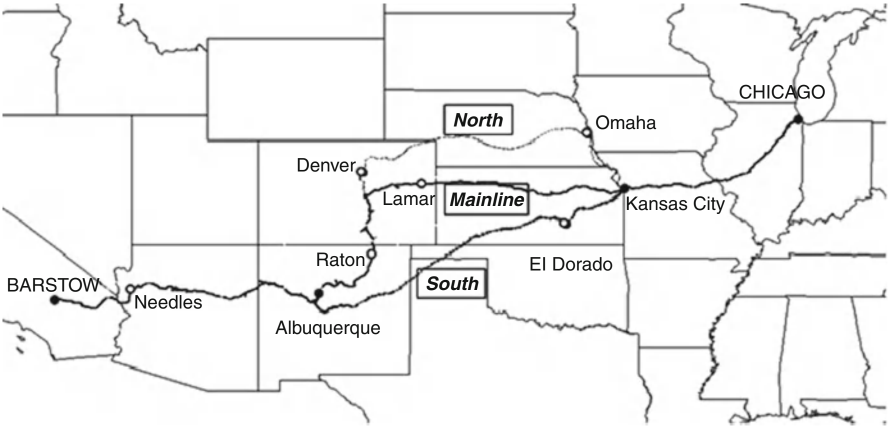


Fig. 8.3 Railroad routes

on the stretch between Albuquerque and Kansas City, which could be a focal point for subsequent analysis. To facilitate appropriate assessment, it is important that route attribute data be complete, and hence Table 8.4 also contains the relevant information. More specifically, statistic on segment length, track class and type, maximum possible speed,[1] elevation difference, and topology is provided for each segment, which will be used to compute derailment probability. Arc Map (in ArcView GIS) was used to separate the rail network into two levels: sea and mountain. The former covers mostly the central and eastern parts of the country and has an elevation difference of less than 1,500 m, while the latter has a higher elevation difference.

For illustration purposes, we focus on 110 railcars including 10 with hazmat to be shipped from Barstow to the three demand points in a single freight-train. The first demand point (i.e., Albuquerque) needs 25 railcars including 2 hazmat ones, Kansas City requires 30 railcars including 3 with hazmat, while Chicago has demanded 55 railcars (including five with dangerous goods). The objective is to determine the best way to move the ten hazmat railcars such that the transport-risk is minimized. Note that this implies not just a decision about the route, but also the railcar placement strategy at the rail-yard in Barstow.

8.4.2 Analysis

In an effort to replicate the practice of railroad industry, it is presumed that railcars will move from origin to destination yard in blocks (i.e., group of railcars with same

[1]To clarify the track class of each segment, timetable speeds for passenger trains have been used. For example, the average speed of Southwest Chief passenger train, operating between Los Angeles and Chicago, was calculated and this value was then calibrated against the FRA track classification data to determine the track class.

Table 8.4 Relevant attributes for the three routes

Routes/segments	1	2	3	4	5	6	7
Mainline							
Distance (miles)	169	599	243	157	495	437	X[a]
Track class	4	5	4	5	5	5	
Track type[b]	1	2	1	2	2	2	
Max speed (mph)	60	80	60	80	80	80	
Elevation difference (mts)	4,507	3,830	3,147	3,391	1,024	470,291	
North							
Distance (miles)				182	530	188	
Track class				5	4	5	
Track type				2	1	2	
Max speed (mph)				80	60	80	
Elevation difference (mts)				3,391	1,397	1,024	
South							
Distance (miles)			630	175		X	X
Track class			4	5			
Track type			1	2			
Max speed (mph)			60	80			
Elevation difference (mts)			3,147	1,024			

[a]X: Implies this segment is not a part of the given route
[b]Track type: 1 = Poor; 2 = Good

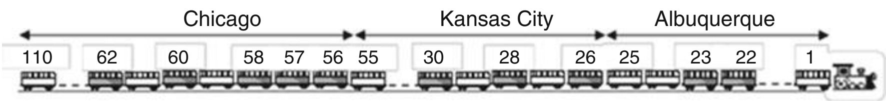


Fig. 8.4 Destination-based blocks and position of Hazmat railcars

handling point). Furthermore, it is assumed that three blocks – one for each railway station – is being built, and that incoming railcars at Barstow are assigned to one of the blocks on a first-come first-served basis. To facilitate exposition and later discussion, Fig. 8.4 depicts the resulting position of hazmat railcars (i.e., ones in grey) within each of the three blocks.

Probability of train derailment, PTD, for a specific segment (i.e., Barstow to Needles in this instance) of a particular route is calculated using Eq. 8.9, the derailment rates from Table 8.1, and assuming that approximately 25% of all derailments can be classified as train-mile caused and the remaining as car-mile caused (Anderson and Barkan 2005). For example, for first rail segment, between Barstow and Needles, where distance is 169 miles and track class 4, $RC = 0.75\times\left(7.8\times10^{-9}\right) = 5.85\times10^{-9}$, and $RT = 0.25\times\left(0.53\times10^{-6}\right) = 0.132\times10^{-6}$. Thus, the probability of derailment for a 110-railcar train on this segment is approximately 0.00013. Similarly, train derailment probabilities for other segments and routes can be obtained.

Table 8.5 Weights and train-length

Train length	M_1	M_2
110	0.3	0.7
85	0.8	0.2
55	0.4	0.6

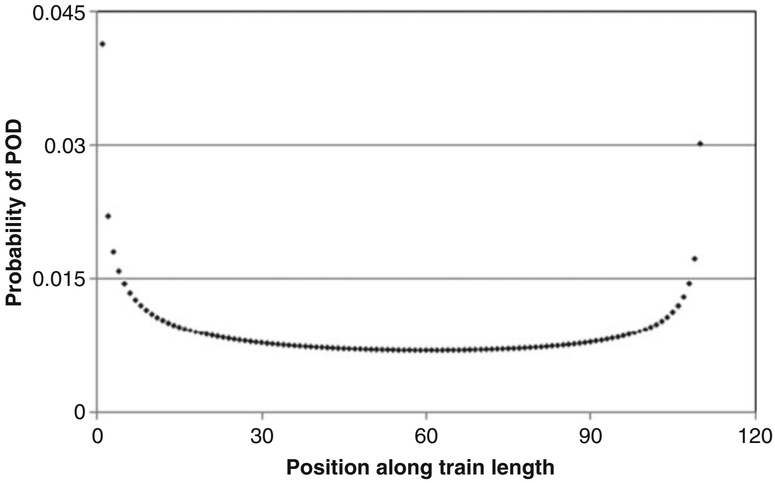


Fig. 8.5 Conditional probability for POD for each position on the first rail segment

Probability of point of derailment starting at position i (i.e., P_i^{POD}) is estimated using the best fit distributions from [22]. Figure 8.2 is used to estimate weights, M_1 and M_2, for a given train-length. Table 8.5 depicts the membership values based on three train-lengths, since the lengths will not change over segments between two demand points. P_i^{POD} for each of the 110-positions in the train is calculated using the four steps outlined in Sect. 8.3.2. Figure 8.5 depicts the probability of derailment (i.e., product of train derailment and point of derailment) for each of the 110 positions in the train for the rail segment from Barstow to Needles. It should be noted that based on the probability estimates for this segment, it is advisable to place hazmat railcars at the rear of the first two blocks (destined for Albuquerque and Kansas City), while the front of the third block appears safer.

Now, consequence associated with derailment at position i would require information on: conditional probability of derailment for other hazmat railcars $P(m|i)$; conditional probability of derailed hazmat railcars releasing $P(j|m)$; and, population exposure associated with releasing cars (PE_j).

$P(m|i)$ is estimated, using Eq. 8.11, together with the summary statistic for different causal factors in Table 8.2. For example, the freight train will carry 110 railcars on the first rail segment between Barstow and Needles. Assume that the derailment is initiated at the very first position in the train (i.e., $i=1$). Now, the derailment probabilities for the remaining 109 railcars on this segment are between 0.0279 and 0.00125.

$P(j|m)$: It is important to note that conditional probability that a certain number of derailed hazmat railcars will release depends on two factors: the order in which hazmat railcars arrived at the marshalling yards; and, the number of hazmat railcars that could potentially release given the POD. Since we are aware of the sequence of arrival of railcars at the Barstow yard, we know not just the exact position of each railcar within a block, but can also determine the number of hazmat railcars likely to release given a particular point of derailment (Fig. 8.4). For example, if the derailment was initiated at $i = 1$ and 22 railcars derailed (i.e., $m = 22$), the maximum number of hazmat railcars that can release is one (i.e., $j = 1$, for the hazmat railcar in the 22nd slot). However, the maximum release number increases to two, if $m = 23$. For the latter, we will have two risky scenarios with $j = 1$ and $j = 2$ (since $j = 0$ does not result in any hazmat risk). Similarly, if $i = 1$ and $m = 62$, then j can assume any element from the sample space: {1, 2, 3,…, 9, 10}.

Two observations are important in this regard: derailment starting in one block can impact railcars in other blocks; and, derailment starting after the last hazmat railcar poses no risk (i.e., beyond the 62nd slot in Fig. 8.4). In the absence of any peer-reviewed work demonstrating the interaction between various types of hazmat transported on railroad, we assume that all railcars are carrying gasoline.

(PE_j): Two pieces of information –the number of hazmat railcars releasing and the quantity released from each –would be required for the a priori determination of population exposure. Note that we are addressing the first concern by building scenarios based on every possible value j can assume. Based on simulation exercises, we are assuming loss of entire lading within 10 min, which is less than the anticipated response time for emergency response providers [14].

For each possible value of j, and for every rail-segment on the three routes, Eq. 8.15 was used to determine the threshold distance. For each j value and for each rail-segment, the threshold distance was used in ArcView GIS, together with Avenue Programming, to generate the corresponding exposure zone, which was then overlaid on the population centers to compute the population exposure (see [28] for details). It is important to note that this kind of processing is very time intensive, since it has to be done individually for each j –rail segment pair. For our case example, a total of 110 computations had to be done, and they are listed in Table 8.6.

Table 8.7 depicts the transport risk for each segment, and for the three routes. Transport Risk for the segment between Barstow to Needles is 2109×10^{-6}, which represents the sum of risk (i.e., Riski) over all the 110 railcar positions in the train consist. For example, Risk_i for $i = 62$ is the product of P_i (i.e., $P^{TD} \times P_i^{POD} = 0.00013 \times 0.006867$) and C_i (i.e., probability that one railcar will release as a result of derailment from m railcars x population exposure from one railcar from Table 8.6 = 0.0903×316). Similar calculation is done for all other positions in the train, which will add to $2{,}109 \times 10^{-6}$. Although the South route is the shortest, it is 38% riskier than the longest route (i.e., North), but the Mainline route has the lowest risk and should be preferred over the other two. In the next subsection we comment on some related insights.

Table 8.6 Number of hazmat railcars releasing and population exposure

Release – route/ segments	1	2	3	4	5	6	7
Mainline							**X**
j =*1*	316	1,118	43	28	87	419	
j =*2*	523	1,855	71	46	145	695	
j =*3*	674	2,388	92	59	186	894	
j =*4*	856	3,033	116	75	237	1,136	
j =*5*	1,023	3,627	139	90	283	1,359	
j =*6*	1,135	4,023	154	100	314	1,507	
j =*7*	1,215	4,306	165	107	336	1,613	
j =*8*	1,348	4,776	183	118	373	1,789	
j =*9*	1,457	5,163	198	128	403	1,934	
j =*10*	1,544	5,473	210	136	427	2,050	
North							
j =*1*				235	174	963	
j =*2*				390	289	1,598	
j =*3*				502	371	2,056	
j =*4*				638	472	2,611	
j =*5*				762	564	3,123	
j =*6*				846	626	3,464	
j =*7*				905	670	3,707	
j =*8*				1,004	743	4,112	
j =*9*				1,085	803	4,446	
j =*10*				1,150	851	4,712	
South						X	X
j =*1*			1,412	147			
j =*2*			2,342	244			
j =*3*			3,014	314			
j =*4*			3,829	399			
j =*5*			4,579	477			
j =*6*			5,079	529			
j =*7*			5,436	566			
j =*8*			6,030	628			
j =*9*			6,518	679			
j =*10*			6,910	719			

Table 8.7 Transport risk for the three-routes ($\times 10^{-6}$)

Routes/Segments	1	2	3	4	5	6	7	Total
Mainline	2,109	18,268	278	78	758	905		22,397
North				756	2,448	3,185		27,951
South			23,714	453				45,449

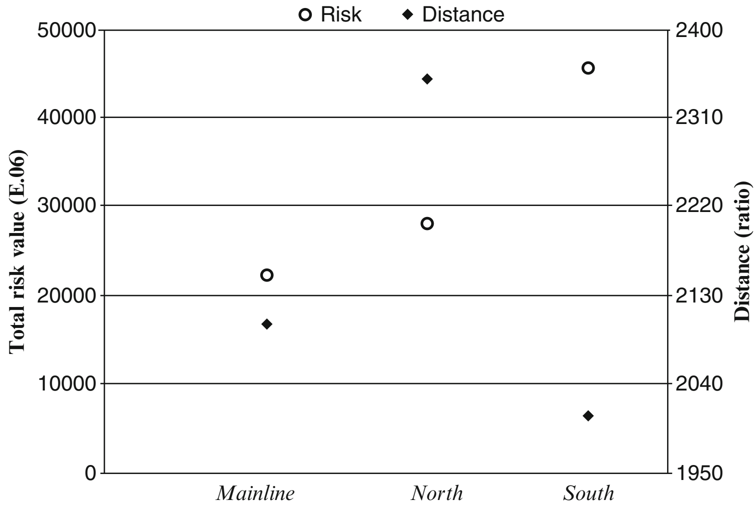


Fig. 8.6 Snapshot of transport-risk v/s distance

8.4.3 Managerial Insights

Given the objective of railroad companies to minimize cost, it is conceivable that shortest path is important –although number of intermediate handling also plays a crucial role. Because each of the three routes has two intermediate destinations, we can deduce that although South route is the shortest at 2,010 miles, railroad companies should select the Mainline route wherein the transport risk is less than half. This is because the rail segments between Albuquerque to Kansas City –through El Dorado –are traversing through more densely populated areas, and hence the population exposure as a result of an accident is higher than that for the other routes. Figure 8.6 provides the length and transport risk for each of the three routes.

Since the three routes differ just on the stretch between Albuquerque and Kansas City, our remaining analysis focuses only on those segments. Drawing inspiration from the methodology developed in [38] subsequently corrected for errors in [39], we made use of a route based approach to compute population exposure for the rail segments in question. Resulting population exposure numbers were plotted against the combined probability profile (i.e., probability of train derailment, probability of point of derailment, probability of a given number of railcars derailing, and the probability of a given number of hazmat railcars releasing). Figure 8.7 depicts the combined probabilities and corresponding population exposure numbers for the rail segments, for each route, for hazmat release from one to eight railcars (since two hazmat railcars are dropped-off in Albuquerque).

It is clear from Fig. 8.7 that, for each of the eight release scenarios, Mainline route dominates the other two, and should be taken as far as possible. But it is very interesting to note that although the North route exposes fewer people for

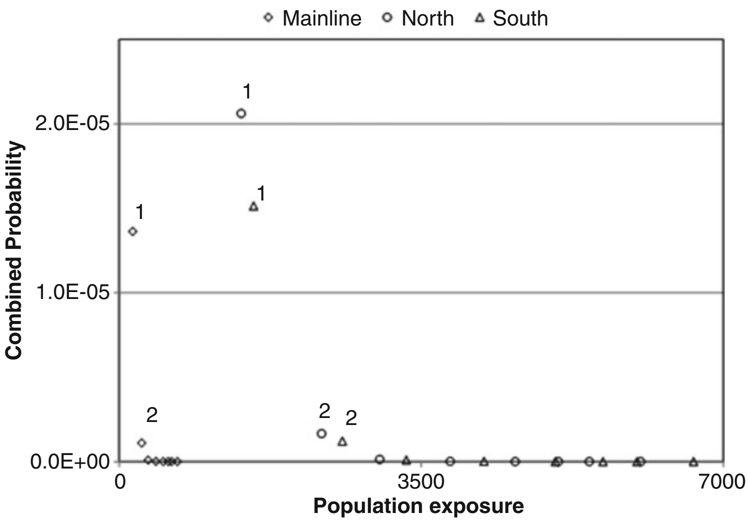


Fig. 8.7 Combined probability v/s population exposure

any of the eight release scenario than the South route, the former is worse off in terms of combined probabilities of release. This implies that if exposure is not a consideration, then South route would be preferred by railroads, but North route would be chosen if incident probability is paramount. This is why it is important to use the proposed risk model since it incorporates all elements of transport risk, and not look a singular factor (such as, derailment probability or consequence).

8.5 Conclusion and Future Research

This chapter outlines an expected risk model that is distinct from the other models proposed for rail transport of hazmat. The model considers the sequence of events leading to hazmat release from railcars and the resulting consequence from various release sources, and is applied to analyze a realistic-size problem instance based in the US. To that end, an innovative risk assessment framework – that considers train derailment, point of derailment, number of railcars derailing and releasing hazmat, and the number of people exposes – is outlined. It is expected that the current work will culminate in a tool that can assist railroad companies to determine not just the best route but also the more preferred position to carry hazmat railcars. Through the case study it was noticed that while a single dominant route is possible, different risk models can result in different solutions, not unlike road transportation of hazmat.

There are a number of research projects arising from the expected risk framework outlined in this chapter. First, deals with the placement of hazmat railcars within a block and the train consist. To that end, the sequence of railcars within a block and the order of block to form a train, and how they result in increased yard-risk are being investigated. Second, the insights gained from these works would be used to develop an analytical approach to determine railcar routing, hazmat railcar position in a train consist, and yard activities in a network level setting.

References

1. AAR (2006) Association of American Railroads, Current rail hazmat conditions called "Untenable", AAR News Press, June. http://www.aar.org/View Content.asp?Content_ID=3763
2. TSB (2004) Transportation Safety Board of Canada, Statistical summary railway occurrences 2004. http://www.tsb.gc.ca/en/stats/rail/2004/statssummaryrail04.pdf
3. Oggero A, Darbra RM, Munoz M, Planas E, Casal J (2006) A survey of accidents occurring during the transport of hazardous substances by road and rail. J Hazard Mater 133A:1–7
4. FRA (2010) Federal Railroad Administration, Accident data on demand, Federal Railroad Administration Office of Safety Analysis. http://safetydata.fra.dot.gov/officeofsafety
5. NTSB (2005) National Transportation Safety Board, Railroad accident report. NTSB Report Number: RAR-05-04. http://www.ntsb.gov/publictn/2005/RAR0504.pdf
6. TSB (2005) Transportation Safety Board of Canada, Railway investigation report. Number: R05E0059. http://www.tsb.gc.ca/eng/rapports-reports/rail/2005/r05e0059/r05e0059.pdf
7. TDG (2007) Transport Canada Newsletter. http://www.tc.gc.ca/tdg/newsletter/spring2007.pdf
8. Erkut E, Tjandra S, Verter V (2007) Hazardous materials transportation. In: Barnhart C, Laporte G (eds) Handbooks in operations research and management science: transportation. Elsevier, Amsterdam
9. Cordeau J-F, Toth P, Vigo D (1998) A survey of optimization models for train routing and scheduling. Transp Sci 32(4):380–404
10. Glickman TS (1983) Rerouting railroad shipments of hazardous materials to avoid populated areas. Accid Anal Prev 15(5):329–335
11. Glickman TS, Rosenfield DB (1984) Risks of catastrophic derailments involving the release of hazardous materials. Manag Sci 30(4):503–511
12. Barkan CPL, Tyler Dick C, Anderson R (2003) Railroad derailment factors affecting hazardous materials transportation risk. Transp Res Rec J Transp Res Board 1825:64–74
13. Anderson RT, Barkan CPL (2004) Railroad accident rates for use in transportation risk analysis. Transp Res Rec J Transp Res Board 1863:88–98
14. Verma M (2010) Railroad transportation of dangerous goods: a conditional exposure approach to minimize transport risk. Transp Res Part C: Emerg Technol 19:790–802
15. Raj PK, Pritchard EW (2000) Hazardous materials transportation on US railroads. Transp Res Rec: J Transp Res Board 1707:22–26
16. Barkan CPL, Treichel TT, Widell GW (2000) Reducing hazardous materials releases from railroad tank car safety vents. Transp Res Rec J Transp Res Board 1707:27–34
17. Saat MR, Barkan CPL (2005) Release risk and optimization of railroad tank car safety design. Transp Res Rec J Transp Res Board 1916:78–87
18. Barkan CPL, Ukkusuri S, Waller ST (2007) Optimizing railroad tank cars for safety: the tradeoff between damage resistance and probability of accident involvement. Comput Oper Res 34:1266–1286
19. Fang P, Reed HD (1979) Strategic positioning of railroad cars to reduce their risk of derailment. Technical report, Volpe Transportation Systems Center (DOTITSC), 7, p 67
20. Thompson RE, Zamejc ER, Ahlbeck DR (1992) Hazardous materials car placementin a train consist. Technical report, Battelle Columbus Division
21. Bagheri M (2009) Risk-based model for effective marshaling of dangerous goods railways cars. PhD dissertation, University of Waterloo, Waterloo

22. Bagheri M, Saccomanno FF, Chenouri S, Liping F (2010a) Reducing the threat of in-transit derailments involving dangerous goods through effective placement along the train consist. J Accid Anal Prev. doi:10.1016/j.aap.2010_09_008
23. Hwang ST, Brown DF, O'Steen JK, Policastro AJ, Dunn W (2001) Risk assessment for national transportation of selected hazardous materials. Transp Res Rec J Transp Res Board 1763:114–124
24. Leeming DG, Saccomanno FF (1994) Use of quantified risk assessment in evaluating the risks of transporting chlorine by road and rail. Transp Res Rec J Transp Res Board 1430:27–35
25. Arya SP (1999) Air pollution meteorology and dispersion. Oxford University Press, Cambridge
26. Patel MH, Horowitz AJ (1990) Optimal routing of hazardous materials considering risk of spill. Transp Res Part A 28(2):119–132
27. Zhang J, Hodgson J, Erkut E (2000) Using GIS to assess the risks of hazardous materials transport in networks. Eur J Oper Res 121:316–329
28. Verma M, Verter V (2007) Railroad transportation of dangerous goods: population exposure to airborne toxins. Comput Oper Res 34:1287–1303
29. Bagheri M, Saccomanno FF, Fu L (2010) Effective placement of dangerous goods cars in rail yard marshaling operation. Can J Civil Eng 37(5):753–762
30. Anderson RT, Barkan CPL (2005) Derailment probability analysis and modeling of mainline freight trains. In: Proceedings of the 8th international Heavy Haul conference, Rio de Janiero, pp 491–497
31. Saccomanno FF, El-Hage S (1989) Minimizing derailments of railcars carrying dangerous commodities through effective marshalling strategies. Transp Res Rec J Transp Res Board 1245:34–51
32. Saccomanno FF, El-Hage S (1992) Establishing derailment profiles by position for corridor shipments of dangerous goods. Can J Civil Eng 19(1)
33. Verma M, Verter V (2010) A lead-time based approach for planning rail-truck intermodal transportation of dangerous goods. Eur J Oper Res 202(3):696–706
34. Verma M, Verter V, Gendreau M (2010) Tactical planning model for railroad transportation of dangerous goods. Transp Sci 45:163–174
35. ESRI (2007) ArcView GIS 9.1. ESRI Inc., 380 New York Street, Redlands, CA, USA
36. US CFS (2002) Commodity flow survey, research and innovative technology administration: US DOT. http://www.bts.gov/programs/commodity_flow_survey
37. FAF (2002) Freight analysis framework, FHWA office of freight management and operations. http://ops.fhwa.dot.gov/freight/freight_analysis/faf/faf2_com.htm
38. Erkut E, Verter V (1998) Modeling of transport risk for hazardous materials. Oper Res 46:625–642
39. Kara BY, Erkut E, Verter V (2003) Accurate calculation of hazardous materials transport risks. Oper Res Lett 31:285–292

Chapter 9
Transportation and Storage of Spent Nuclear Fuel: Security and Theory

Barseghyan Artak and Martoyan Gagik

Abstract The problem of spent nuclear fuel is a major issue. The safety and efficiency of nuclear power plants depends on finding a solution to this problem. The transportation and storage of radioactive materials is already strictly regulated to ensure a high degree of safety. Nevertheless, recommendations to ensure the safe transportation and storage of many types of radioactive materials are still being developed.

This chapter describes the processes occurring in radioactive waste from the viewpoint of radiochemistry. It gives an overview of transport security requirements and describes the development and implementation of recommendations for safe storage. The areas of security are defined.

9.1 Introduction

9.1.1 Radiation Chemistry in the Nuclear Fuel Cycle

The nuclear fuel used in nuclear power plant exploitation is obtained, enriched, and used in nuclear reactors. All of this, and particularly the transportation and storage of spent nuclear fuel (SNF) is impossible without an understanding of radiation chemistry.

B. Artak (✉)
"AREV" Scientific-Industrial CJSC, Ecoatom LLC Adana 1, Yerevan 0082, Armenia, Engineering Academy of Armenia, Teryan 105/17, Yerevan 0009, Republic of Armenia

Engineering Academy of Armenia, Teryan 105, bld. 17, Yerevan 0009, Republic of Armenia
e-mail: artakbarseghyan@yahoo.com

M. Gagik
AREV SI CJSC, Adana 1, Yerevan 0082, Republic of Armenia

E. Garbolino et al. (eds.), *Transport of Dangerous Goods*, NATO Science for Peace and Security Series C: Environmental Security, DOI 10.1007/978-94-007-2684-0_9,

The development of radiation chemistry as a science has been inextricably linked with the emergence and development of the nuclear industry. Since the first reactors produced fissile materials, the study of the products of uranium fission has shown the need to investigate the effects of ionizing radiation on structural materials, fluids, reagents and solutions. At the same time the kinetics of chemical reactions that occur in the technology must be understood, in order to isolate target products.

Improving fissile materials' technology and power production at nuclear plants required the development of databases containing experimental data on the effects of radiation on materials and processes. It also required the development of a body of theory on radiation-chemical processes, ranging from the interaction of radiation with matter, to the formation of the final radiolysis products and changes in the performance properties of materials. Current theory makes it possible to adequately describe almost any problem in nuclear power and the radiochemical industry associated with the effects of exposure to ionizing radiation on materials and processes. In most cases, a description is possible at a level that allows practical solutions to be found.

9.1.2 The Nuclear Fuel Cycle

Radiation-chemical problems manifest themselves at various stages in the nuclear fuel cycle. The most common configuration of the nuclear fuel cycle is shown in Fig. 9.1. In the first two stages of the fuel cycle (mining uranium and the manufacture of fuel assemblies) problems related to the action of ionizing radiation on materials and processes do not arise. This is due to the low doses of natural radioactivity of products and reagents used in the technology in these stages.

Within the nuclear power plant (NPP) the main problem related to radiation chemistry is to identify and minimize the effects of ionizing radiation on coolants, moderators and other materials in the core of the nuclear reactors. The problem is made more difficult by the fact that in these conditions the composition of the radiation that these materials are exposed to is complex (slow and fast neutrons, α-, β-, and γ-radiation with a broad energy spectrum, 'hot' fission fragments and recoil), it happens at very high rates (up to MG/s), at high temperatures (up to 600°C) and pressures (up to 200 atm).

In some cases it is economically preferable to obtain uranium from mines using underground uranium leaching schemes (Fig. 9.2).

In this scenario profitability is determined as the combination of uranium extraction and the risk of environmental pollution. Maintaining the safety of staff and the environment is very important during the extraction and enrichment process. Using this process it is possible to extract separate elements and isotopes from SNF using a hydrometallurgical method. The method is based on differences in the electrochemical parameters of elements and it is able to segregate elements with high selectivity and automation (Fig. 9.3).

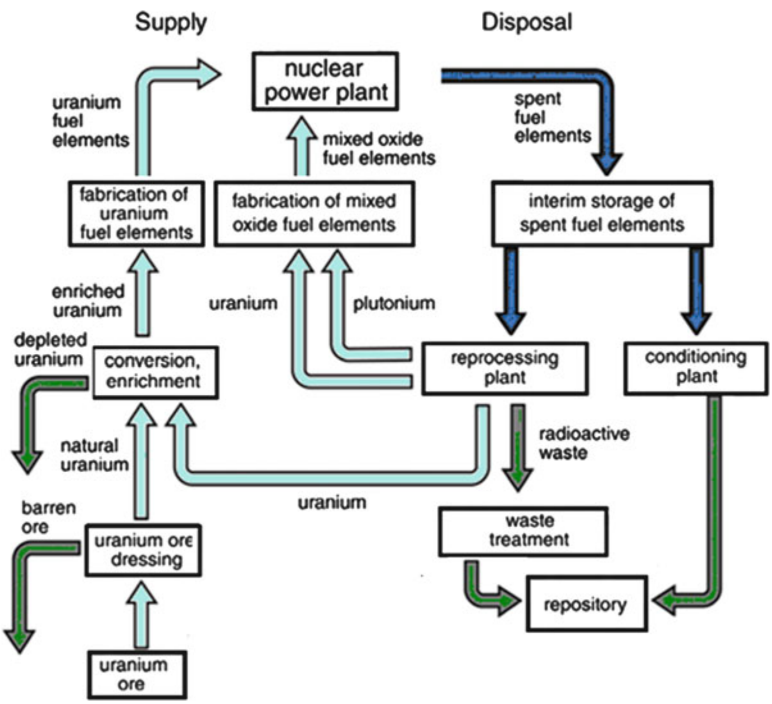


Fig. 9.1 The nuclear fuel cycle

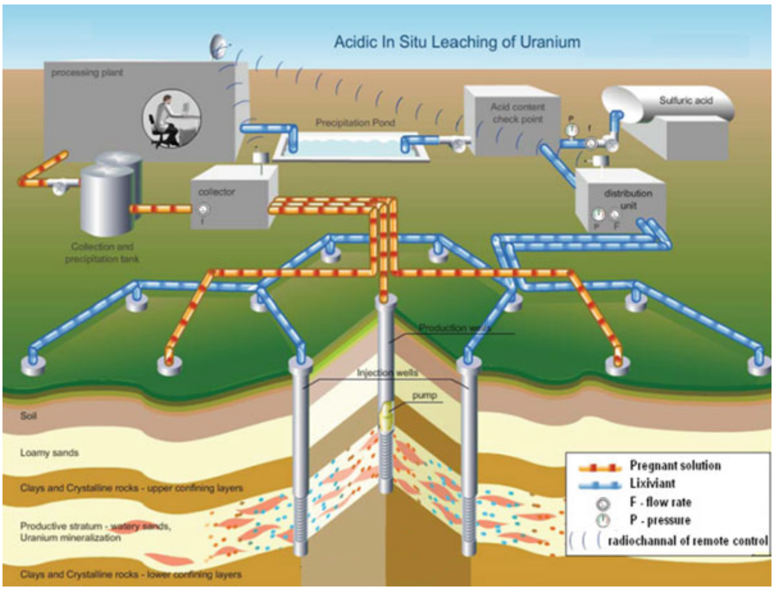


Fig. 9.2 Underground uranium leaching

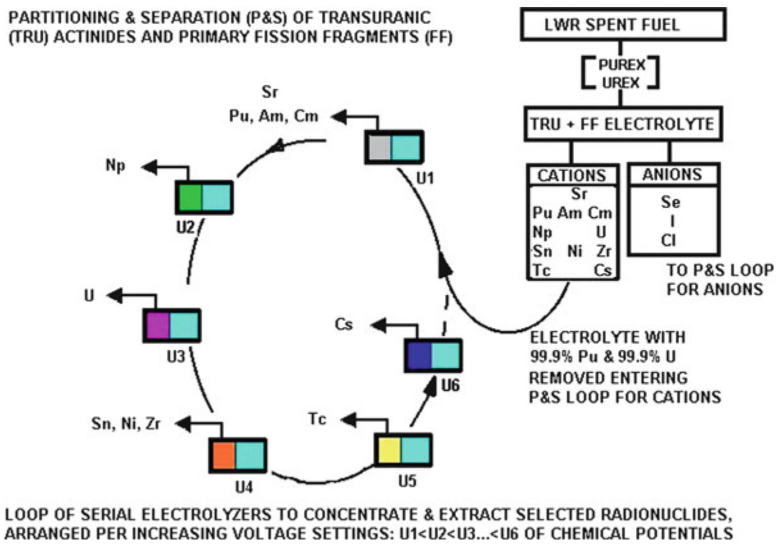


Fig. 9.3 Single stage extraction, partitioning and chemical separation of selected transuranic actinides and primary fission fragments radionuclides from post-PUREX aqueous solution

9.1.3 Radiation-Chemical Aspects of Spent Nuclear Fuel

In accordance with the technology of nuclear power plants, fuel assemblies are periodically unloaded from the reactor core. They are unloaded into special pools of various designs situated near the reactor. The fuel assemblies are placed on racks or in special canisters in these water-filled basins, and stored for 3–8 years (or more).

The water in the storage pools plays three roles: it provides biological protection, neutron shielding and acts as a coolant. However, two problems arise in connection with this type of storage, which are due to the effect of radiation on water:

- The formation of hydrogen and, consequently, the problem of preventing a possible explosion should the hydrogen mix with oxygen;
- Intensive corrosion of the storage basins due to the effects of radiation.

The water in the pools is constantly exposed to special treatment using anion exchange columns. As the water in the storage basins is in constant contact with air, in addition to containing oxygen and nitrogen, it also contains carbon dioxide.

Molecular hydrogen is formed in the storage pools. Radiolysis of the water in the basin is mainly a result of exposure to gamma radiation (beta and alpha radiation is trapped by the fuel assembly shells and its energy is converted into heat). The initial rate of radiolytic hydrogen formation is low and a continuous flow of air above the pool ensures protection from fire or explosion.

Others problems are associated with the effects of ionizing radiation. More serious problems are created in the next phase of the nuclear fuel cycle, transportation of spent fuel from nuclear power plants for reprocessing. Spent fuel is transported in special containers. The spent fuel assemblies are placed into these containers in special covers. The walls of the container are both structurally strong and provide protection from radiation. Heat produced during radioactive decay is released into the environment (air or nitrogen) or into water-cooled containers. The water content of air plays a role both as a fluid and offers protection from fast-neutron fission.

Containers are mounted on railroad cars, trailers, ferries, etc., and transported on public highways. This imposes additional security requirements. Currently, the IAEA imposes requirements to ensure the safe transportation of radioactive materials (including spent fuel from nuclear power plants). These rules require that during transportation, including emergency situations, there should be no leakage of radioactive materials from the container. For example, the container must remain sealed after falling from a height of 9 m onto a concrete base; it should withstand temperatures of up to 800°C for 1 h, or submersion in water at a depth of 10 m. The rules also require that free volume water-filled containers under all conditions of carriage can withstand a hydrogen explosion. This is achieved through various safety mechanisms which liquidate the hydrogen-oxygen mixture. As it is impossible to conduct experiments in such complex situations, simulation methods have been developed to predict hypothetical accident scenarios.

9.2 Methods and Characteristics of Spent Nuclear Fuel Transportation

Nuclear power can come from the fission of an atom of uranium, plutonium or thorium, or the fusion of hydrogen into helium. The primary method currently used is uranium fission. Uranium atom fission creates ten million times the energy produced by the combustion of an atom of carbon from coal.

Natural uranium is almost entirely a mixture of isotopes: 0.7% U-235, 99.3% U-238 and a trace amount of U-234 by mass. Natural uranium may be used as a fuel in nuclear reactors and from point of radioactivity the appropriate percentages are: 2.2% U-235, 48.6% U-238, and 49.2% U-234 [1].

9.2.1 What Is Spent Nuclear Fuel?

Spent fuel is the waste product from nuclear power reactors. Uranium is able to power reactors for a number of years, until its potential to produce electrical power is exhausted. The exhausted uranium is referred to as spent fuel [2].

9.2.2 Spent Nuclear Fuel Storage Methods

Nuclear power plants store spent fuel either in closed cooling ponds or, in some cases, in dry storage casks to await shipment to a temporary storage or permanent disposal facility. Acceptable methods for storing spent fuel after it has been removed from the reactor core are shown in Fig. 9.4.

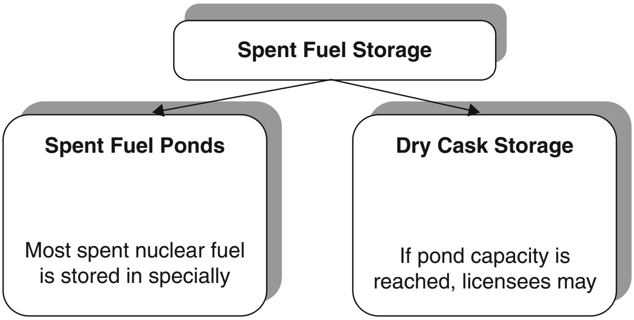


Fig. 9.4 Spent nuclear fuel storage methods

Fig. 9.5 Fuel ponds at the Centrale nucleare di Caorso (Photos by Simone Ramella [4])

9.2.3 Spent Nuclear Fuel Ponds

Spent nuclear fuel ponds are constructed using very thick steel-reinforced concrete walls with stainless steel liners, which are located in special protected areas, designed for spent fuel from nuclear reactors. Many fuel ponds are located below ground level. Many are shielded by other structures, and many have intervening walls to protect the structure from impact with aircraft, for example. Spent fuel ponds contain enormous quantities of water. Spent fuel rods are stored under at least 20 ft (6 m) of water. This provides adequate shielding from radiation. The rods are moved from the reactor into the ponds through water canals, in order to shield workers from the spent fuel.

About a quarter to a third of the total fuel load in the ponds becomes spent and is removed from the reactor every 12–18 months. It is replaced with fresh fuel. Nuclear plants have access to many sources of water which provide backup supplies to the spent fuel ponds [3].

Figure 9.5 shows the fuel ponds at the *Centrale nucleare di Caorso* (Italy) [4].

Fig. 9.6 Dry cask storage

9.2.4 Dry Cask Storage

In the late 1970s an alternative storage solution became necessary as the ponds at many nuclear reactors began to fill up. Dry cask storage (Fig. 9.6) was developed to increase capacity. This type of storage can handle already cooled spent fuel which has been in the spent fuel pond for at least 1 year. The spent fuel is held in an inert gas inside a container called a cask. The casks are typically steel cylinders which provide leak-tight containment of the spent fuel and are either welded or bolted closed. Casks are designed to resist situations such as tornadoes, projectiles, floods and temperature extremes.

There are horizontal and vertical designs (Fig. 9.7) [3].

9.3 Ensuring Security of Spent Nuclear Fuel Transportation

9.3.1 Methods of Radioactive Material Transportation

Spent nuclear fuel is highly radioactive and must be heavily shielded and tightly contained to be transported safely. A robust spent fuel container or cask is essential for safe transportation. The following describes some methods of safe transportation for spent nuclear fuel. Package design is the primary safety measure. Some issues that must be taken into account are: stability following damage, fire protection and structural resistance. In order to demonstrate that containers are safe,

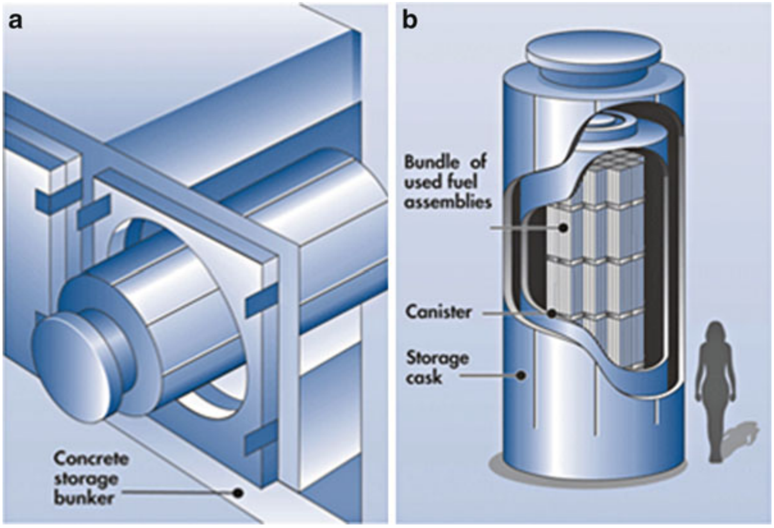


Fig. 9.7 (**a**) Horizontal and (**b**) vertical dry cask storage facilities (images from www.nrc.gov)

designers may use computer analyses, comparisons with component testing, scale-model testing, etc. or a combination of these techniques. In many cases they use a combination of computer analyses and physical testing.

9.3.2 Sea Transportation

The International Maritime Dangerous Goods Code (IMDG Code) was published by International Maritime Organization (IMO) in 1965. It applies to the carriage of dangerous goods of any kind by sea. This international instrument includes matters such as packing and container stowage, and makes particular reference to the segregation of incompatible substances. The IMDG Code offers guidance to those involved in the handling of radioactive material using sea transportation.

In 1993 the International Maritime Organization established the Code for the Safe Carriage of Irradiated Nuclear Fuel, Plutonium and High-Level Radioactive Wastes in Flasks on Board Ships (INF Code). The INF Code was designed to complement the International Atomic Energy Agency (IAEA) Regulations. The Code introduced recommendations for the design of vessels transporting radioactive material. These complementary provisions address such issues as stability after damage, fire protection, and structural resistance [5]. In January 2001, the INF Code was made mandatory and renamed the International Code for the Safe Carriage of Packaged Irradiated Nuclear Fuel, Plutonium and High-Level Radioactive Waste on Board Ships.

Table 9.1 Agreement concerning the International Carriage of Dangerous Goods by Road (ADR): Annexes A and B

Annex A: General provisions and provisions concerning dangerous articles and substances	
Part 1	General provisions
Part 2	Classification
Part 3	Dangerous goods list, special provisions and exemptions related to limited and excepted quantities
Part 4	Packing and tank provisions
Part 5	Consignment procedures
Part 6	Requirements for the construction and testing of packaging, intermediate bulk containers (IBCs), large packaging and tanks
Part 7	Provisions concerning the conditions of carriage, loading, unloading and handling
Annex B: Provisions concerning transport equipment and transport operations	
Part 8	Requirements for vehicle crews, equipment, operation and documentation
Part 9	Requirements concerning the construction and approval of vehicles

9.3.3 Air Transportation

The International Civil Aviation Organization (ICAO) has responsibility for all aspects of international civil aviation [6]. In 1981, the ICAO adopted Annex 18 covering the air transport of dangerous goods and published a set of technical instructions detailing the requirements for these movements. The technical instructions contain a list of dangerous goods, as well as requirements for packing, marking, labelling and documentation in order to be fully consistent with the International Air Transport Association Regulations [7].

9.3.4 Land Transportation

The United Nations Economic Commission for Europe (UN/ECE) has published the European Agreement concerning the International Carriage of Dangerous Goods by Road (the ADR). The Agreement itself is short and simple. The key article is the second, which states that apart from some excessively dangerous goods, other dangerous goods may be carried internationally in road vehicles subject to compliance with:

- the conditions laid down in Annex A (Table 9.1) for the goods in question, in particular with respect to their packaging and labelling; and
- the conditions laid down in Annex B (Table 9.1), in particular with respect to the construction, equipment and operation of the vehicle carrying the goods in question [8].

The IAEA Regulations have been adopted to apply to the transport of radioactive material under the ADR. Currently there are 40 contracting States to this Agreement.

The Intergovernmental Organisation for International Carriage by Rail (OTIF) is responsible for the regulations concerning the International Carriage of Dangerous Goods by Rail (RID).

9.3.5 Other Modes of Transportation

For inland waterways, the UN/ECE has developed a European Agreement.

Transport of radioactive material by post is regulated by the Universal Postal Convention and its detailed regulations, which are published by the Universal Postal Union [9]. The Convention allows the transport of exempted quantities of radioactive material, within the terms of the IAEA Regulations, which must conform to IAEA prescriptions.

9.3.6 Safety Regulations for the Transportation of Radioactive Material

Detailed information concerning safety regulations for the transportation of radioactive material is shown in Table 9.2.

9.4 Radiation-Chemical Processes During the Storage of High-Level Liquid Radioactive Waste

A necessary stage in any radiochemical reprocessing of spent nuclear fuel is the intermediate storage in steel tanks of high-level liquid waste (HLW). HLW consists of solutions of nitrate salts, fission radionuclides and transuranic elements, sodium nitrate, organic acids, etc. These solutions contain process impurities, such as extractants, diluents, silicates, corrosion products, etc. This stage is one of the most dangerous in the workplace. The requirements that must be fulfilled for the safe storage of high-level waste are outlined below:

- The removal of heat produced during the decay of radionuclides should ensure the absence of both general and local heating and, especially, evaporation of waste. The mode of heat removal depends on the waste to be reprocessed and its thermal properties.
- Explosive concentrations of flammable gases and vapours should not be allowed to form in the free storage capacity. They may arise from the radiolysis of aqueous waste hydrogen and methane, or of extractants and diluents, hydrogen and 'light' hydrocarbons, butanol, etc.
- During storage of the waste, hazardous corrosive substances (e.g. chloride ions in nitric acid solutions), or substances that are components of waste compounds should not be allowed to form (this may be caused by thermoradiation processes).

Table 9.2 Safety regulations for the transportation of radioactive material

Mode of transport	International/regional organization	Name of regulation/agreement/code	Scope
All	IAEA	Regulations for the Safe Transport of Radioactive Material, TS-R-1	Worldwide
All	United Nations	Recommendations on the Transport of Dangerous Goods	Worldwide
Sea	International Maritime Organisation	International Maritime Dangerous Goods Code (IMDG Code)	Worldwide
		International Code for the Safe Carriage of Packaged Irradiated Nuclear Fuel, Plutonium and High-Level Radioactive Wastes on Board Ships (INF Code)	Worldwide
Air	International Civil Aviation Organization (ICAO)	Technical Instructions for the Safe Transport of Dangerous Goods by Air (TI)	Worldwide
		Dangerous Goods Regulations (DGR)	Worldwide
Road	United Nations Economic Commission for Europe (UNECE)	European Agreement concerning the International Carriage of Dangerous Goods by Road (ADR)	Regional
Rail	Intergovernmental Organization For International Carriage By Rail (OTIF)	Regulations Concerning the International Carriage of Dangerous Goods by Rail (RID)	Regional
Road and rail	MERCOSUR/MERCOSUL (Southern Common Market)	Agreement of Partial Reach to Facilitate the Transport of Dangerous Goods	Regional
Inland waterways	UNECE	European Agreement concerning the International Carriage of Dangerous Goods by Inland Waterways (ADN)	Regional
	Central Commission for the Navigation of the Rhine (CCNR)	Provisions concerning the Carriage of Dangerous Goods on the Rhine (ADNR)	Rhine navigation
Post	Universal Postal Union (UPU)	Universal Postal Convention and its detailed regulations	Worldwide

These are difficult to remove from storage and complicate the treatment of waste. Moreover these substances are able to concentrate the fissile radionuclides (such as the degradation products of extractants).

- Throughout its duration waste storage should be homogeneous, i.e. it should not form precipitation.
- During normal storage operations the minimum technically achievable discharge of radionuclides should be ensured (in aerosol form in the free volume of the container). This can be achieved through optimization of the blow-off of radiolytic explosive gases generated by radiolysis.

The foregoing shows that environmentally sound technologies for the interim storage of liquid HLW must take into account the radiation-chemical processes in high-level liquid waste which are under the influence of ionizing radionuclides radiation. Radiation-chemical transformations of waste components can cause the appearance of potentially dangerous substances with potentially harmful effects during storage. The mechanisms of radiation-chemical reactions in the waste must be understood. Detailed theoretical analysis and the determination of analytical dependencies of the kinetics of the transformations in the waste are unfortunately impossible with today's level of knowledge because of the complexity and variability of waste composition, even within a single process. For this reason it is useful to determine the critical parameters that control these or other radiation-chemical transformations in the waste, to develop a phenomenological picture of the storage technology and develop a methodology for calculating storage modes using semi-empirical laws.

The final stage in the disposal of radioactive waste in the environment is to bury it in deep underground reservoirs. The waste is safely isolated from external influences (primarily water exposure) in salt mines, special tanks, under rock or in wells. However, the final disposal of any technology also involves the temporary storage of solidified waste in blocks of glass, concrete, asphalt, ceramics, metal, etc. for some time in near-surface repositories, where the waste can cool under constant control. It is impossible to guarantee against water penetration in this case, especially for long storage periods. There are two effects when water comes into contact with the waste: leaching of the waste and water radiolysis. Leaching has been studied very intensively. Less attention has been paid to radiolysis of water through the radiation of radionuclides incorporated in the waste. Nevertheless, this effect may be the cause of many problems.

9.5 Mathematical Modeling of Water-Cooled Systems in Containers During Transport and Disposal of Radioactive Waste

9.5.1 Introduction

The transportation of waste is a necessary part of overall activity related to the nuclear cycle. Spent fuel and high-level waste are transported by road, rail or sea in special containers. The design of these containers takes into account heat dissipation, and protection against radiation. They are capable of withstanding any hypothetical accident without loss of integrity. Heat released into the container environment during radioactive decay is removed by cooling with water through a built-in sprinkler system. Here, water plays not only the role of coolant but also provides protection against fast neutrons. These special containers are transported by rail, truck, or ferry on highways open to the general public. This imposes additional safety requirements on nuclear power facilities.

The reaction between the radiation and the water generates chemically active components such as excited molecules and ions, radicals and ion-radicals. Reactions between these components and other dissolved substances are determined by the macro-changes occurring in the water during irradiation, such as the division into radiolytic gases (oxygen and hydrogen), the formation of hydrogen peroxide, and redox processes involving dissolved particles.

The water-cooled container itself is a closed vessel in which there is a certain amount of water and a free volume filled with gas (air or inert gas). The water in the vessel is subject to irradiation; in particular high doses of 'light' γ-radiation, while α and β-radiation (the debris of uranium nuclei) is trapped by the fuel assembly envelope and converted into heat. The conditions which the water is subject to bring about the inevitable onset of radiation-chemical steady-states of hydrogen, oxygen and hydrogen peroxide. The composition of the gas mixture in the free volume will be determined by the partial pressure equilibrium of hydrogen and oxygen, consistent with their concentration in the liquid phase. The task of preventing the explosion of the gas mixture is therefore reduced to minimizing the steady-state concentrations of the stable products of water radiolysis in a closed system. This is achieved by preventing oxygen from the container's free volume from entering the water. The last action before sealing the container should be to fill it with nitrogen under pressure; this is in addition to the use of devices to combust the hydrogen produced by the radiolysis of water.

The hydrogen-oxygen mixture cannot explode if their concentrations in the mixture in the container's free volume are below the lower explosive limit concentrations. These concentrations can be deduced from the laws of radiation chemistry. However, these limits are not always easily respected. Operational nuclear power plants sometimes encounter problems for which a solution is only possible following the quantitative determination of changes in the composition of the coolant, which have occurred as a result of exposure to ionizing radiation. Experiments that might provide solutions to these problems are impossible as they would violate rules governing the safe operation of the container. The difficulty of performing these experiments makes it necessary to instead model these processes using hypothetical cases in order to predict cases where the system may be become explosive. To carry out such research is necessary to create a system model and develop specialised software applications.

The level of steady-state concentrations of the products of water radiolysis is significantly affected by these technological parameters:

- The linear energy transmission associated with the qualitative composition of the radiation which surrounds the water system.
- The temperature of the system under irradiation and the presence of admixtures in the water.

The above discussion demonstrates that the organization of environmentally sound technologies for the interim storage of high-level liquid waste must take into account the radiation-chemical processes in this waste under the influence of the ionizing radiation of radionuclides.

It is therefore important to understand the mechanisms of these processes. Unfortunately, it is impossible to carry out a detailed theoretical analysis and determination of the analytical dependencies of the kinetics of these transformations in the waste with today's level of knowledge because of the complexity and variability in the composition of the waste, even within a single process. For this reason it is appropriate to determine the critical parameters that control these, or other radiation-chemical transformations in the waste, in order to develop a phenomenological picture of the storage technology and develop a methodology for calculating storage modes using semi-empirical laws. The question therefore arises of how to identify these critical parameters?

9.5.2 The Identification of Critical Parameters

There are very important gains to be made in the development of mathematical models of physical and chemical processes in such systems, and in the forecasting methods that use these models in high-performance computing.

A model usually consists of several modules. The central module is a physical description of the phenomenon (radiolysis of coolant in this case) as a set of equations representing chemical reactions, balance, kinetics, transport, etc. The parameters of these equations provide fundamental constants that are characteristic of the phenomenon, e.g. activation energies, rate constants for chemical reactions, and thermodynamic constants of the coolant, as well as the conditions under which the phenomenon occurs (temperature, hydrodynamic conditions, properties of radiations and particle streams). This set of parameters forms the second module. The third module describes the mathematical apparatus and software applications which make it possible to calculate the dynamics of the system and to forecast its response to a given change in conditions. The main issue during the creation of the model is proof of its adequacy, i.e. its capacity to describe at a given accuracy with known assumptions all the available experimental data related to the phenomenon, obtained both from the reactor and laboratory experiments. Such a proof is obtained using statistical methods through the comparison of calculated values with experimental values. As a result of this proof process, the model can be refined.

To describe the kinetics of the accumulation of radiolysis products in the event of continuous exposure to radiation the so-called 'homogeneous approximation' technique is used. In a numerical study we consider the possibility of minimizing the steady-state concentration of hydrogen peroxide controlled by the induced reaction of water vapor in the presence of molecular oxygen in a water-cooled radioactive waste container. Hydrogen peroxide was chosen in this case as it is totally soluble in water, unlike hydrogen and oxygen, which accumulate in the container's free volume and can be neutralized.

9.5.3 The Calculation Method

The rate of change of concentrations of chemical components (c_i) in a kinetic model can be represented as follows:

$$\frac{dc_i}{dt} = f_i(c_1, \ldots . c_n, U), \tag{9.1}$$

Where c_i is the concentration of the *i*-th chemical component, and U is the control parameter defined in the interval $U_{\min} < U < U_{\max}$.

It is often necessary in the chemical process, to obtain the maximal or minimal yield of a product P. This condition can be written as:

$$c_p^{t_f} = \int_{t_0}^{t_f} r_{c_P} \, dt \to \text{extremum} \tag{9.2}$$

The fulfilment of condition (9.2) at condition (9.1) is written as [10]:

$$\begin{cases} \dfrac{dc_i}{dt} = \dfrac{\partial H}{\partial \psi_i} = f_i \\ \dfrac{d\psi_i}{dt} = -\dfrac{\partial H}{\partial c_i} \end{cases}, \quad i = 1, 2, \ldots n \tag{9.3}$$

Here $\psi_i(t)$ is the function adjacent to the functions $c_i(t)$, and H is the Hamiltonian system.

$$H = \psi_0 r_{c_P} + \sum_{i=1}^{n} \psi_i f_i, \tag{9.4}$$

Where: $\psi_0 = \pm 1$ depending on whether the problem of minimum or maximum is solved.

According to Pontryagin's maximum principle [10], the optimal trajectory corresponding to the solution of Eq. 9.3 and target condition (9.2) is written as:

$$\sup H = 0 \tag{9.5}$$

From Eqs. 9.4 and 9.5 it follows that:

$$r_{c_P} = \pm \sum_{i=1}^{n} f_i(t) \psi_i(t) \tag{9.6}$$

Condition (9.5) is obtained for U equal to U_{min}, U_{max} and U_{cl}, where U_{cl} is a classic type of control determined by the condition $\delta H / \delta U = 0$.

The physical sense of $\psi_i(t)$ is the value of the rate of change of accumulation of the i-th component in the change of rate of accumulation of the target component P [11–14]. Equation 9.6 can be rewritten as:

$$r_{c_P} = \pm\sum_j^L G_j r_j, \tag{9.7}$$

Where L is the number of elemental reactions in the scheme, r_j is the rate of j-th stage, G_j is the value of j-th stage ($\mathrm{m+n \rightarrow k+1}$) in the rate of formation of the target component P and is determined as [14]:

$$G_j = \psi_k + \psi_l - \psi_m - \psi_n \tag{9.8}$$

The magnitude of the value contribution of reaction (h_j) also takes into account the rate of the j-th stage:

$$h_j = G_j r_j. \tag{9.9}$$

In the evaluative approach to control of the chemical reaction system, the rate of elemental reactions is proposed as a control parameter. In this case any other aforementioned control parameters are given directly or indirectly, depending on the rates of elementary steps of reaction. This representation is useful as in this case the Hamiltonian system has a linear relationship to the rates of elementary reaction steps. The latter circumstance is very convenient for the effective application of Pontryagin's maximum principle [10]. On the other hand, the "physical", i.e. kinetic interpretation of values ψ_i, G_j, h_j [11–14] demonstrates in specific cases the key efficiency factors – control parameters which are very important for the simulation and controlling of chemical processes.

Therefore, the evaluative approach offers a choice of the most effective control parameters among great number of others. It is then possible to find the optimal values and trajectories of these parameters. At the same time, it makes it possible to perform targeted numerical studies, which form the basis for a model of a real reaction system obtaining hydrogen and hydrogen peroxide when water vapor and molecular oxygen are mixed.

At high temperatures steady-state concentrations of hydrogen and hydrogen peroxide increase. However, a decrease in temperature may not always be sufficient to minimize these concentrations. For a given system there is a minimal residual temperature which is due to the magnitude of the difference between the rates of heat generation and heat recovery in the system itself. In order to answer the question of whether, in low-temperature conditions, a critical radiolysis phenomenon is possible it is necessary to determine the criticality limits depending on temperature and the concentration of active radicals formed during radiolysis. To confirm this, it is necessary to consider the probability of hydrogen peroxide formation in radiolysis conditions at low temperatures, and define the conditions that favor the formation of high steady-state concentrations of H_2O_2.

The evaluative approach also enables the numerical representation of critical reaction conditions, according to which the critical state of the reaction corresponds to extreme behavior in the total concentration of the components of the reaction system [14].

According to this, the target condition is written in the form:

$$J = \int_{t_0}^{t_f} \sum_{i=1}^{n} c_i \, dt \to \text{extremum} \tag{9.10}$$

Identification of critical reaction conditions is important in determining the range of intense reactions. This indicates the search interval from which optimal control parameters are selected.

9.5.4 The Kinetic Reaction Model

The reaction scheme shown in Table 9.3 was selected for the numerical investigation of the potential to obtain hydrogen peroxide through a controlled induced conversion reaction of water vapor in the presence of molecular oxygen. In this case, the kinetic scheme of a branched chain reaction of hydrogen with oxygen in the gaseous phase is a traditional subject of research [14–18]. Note that the hydrogen-oxygen reaction mechanism, which also takes into account how the reaction proceeds when water molecules and hydrogen peroxide are present, is sufficient to describe the chemical conversion of water vapor in the presence of molecular oxygen. A numerical study of this reaction can be carried out as the rate constants of elementary reactions in the chemical mechanism can be reasonably accurately determined over a wide temperature range [15, 18]. In this case the reaction scheme adequately describes a large set of experimental data on the spread of hydrogen-oxygen flames and air mixtures [18].

Calculations are carried out in the framework of a homogeneous model of the system which does not provide for the removal of the products formed by radiolysis of water. It should be noted that if a reaction occurs in the transitional state of the gas-liquid transformation the input parameters present in reaction scheme must be corrected [19].

9.6 Results and Discussion

The best temperature range for obtaining hydrogen peroxide by thermolysis is around T = 1,500 K. This is due to the intense decomposition of molecular oxygen and the relative stability of hydrogen peroxide which occurs at this temperature. Radiolysis increases the probability of hydrogen peroxide formation at low temperatures because

Table 9.3 Kinetic model of the hydrogen-oxygen reaction

		Rate constants			
1.1.1	Stages	k_0 (cm^3/parts)	n	E (kJ/mole)	ΔH^0 (kJ/mole)
1.	$H_2O_2+M \rightarrow OH+OH+M$	6.75E-08	0	175	208
2.	$OH+OH+M \rightarrow H_2O_2+M$	[a]1.20E-30	0	0	−208
3.	$H_2+OH \rightarrow H_2O+H$	4.00E-11	0	21.6	−62.5
4.	$H+H_2O \rightarrow H_2+OH$	3.20E-11	0	72.9	62.5
5.	$H+O_2 \rightarrow OH+O$	1.30E-10	0	66.5	71
6.	$OH+O \rightarrow H+O_2$	5.00E-11	0	5	−71
7.	$O+H_2 \rightarrow OH+H$	4.20E-12	0	32.1	8.3
8.	$OH+H \rightarrow O+H_2$	1.00E-11	0	0	−8.3
9.	$H+O_2+M \rightarrow M+HO_2$	[a]5.00E-33	0	0	−196
10.	$HO_2+M \rightarrow H+O_2+M$	2.80E-09	0	205	196
11.	$OH+OH \rightarrow H_2O+O$	2.40E-11	0	0	−71
12.	$O+H_2O \rightarrow OH+OH$	2.70E-10	0	75.8	71
13.	$H+H+M \rightarrow H_2+M$	[a]1.00E-32	0	0	−430
14.	$H_2+M \rightarrow H+H+M$	9.15E-08	0	436	430
15.	$O+O+M \rightarrow O_2+M$	[a]5.00E-33	0	0	−492
16.	$O_2+M \rightarrow O+O+M$	2.50E-07	0	500	492
17.	$OH+H+M \rightarrow H_2O+M$	[a]1.00E-31	0	0	−492
18.	$H_2O+M \rightarrow H+OH+M$	1.00E-06	0	503	492
19.	$H+HO_2 \rightarrow OH+OH$	1.00E-11	0	0	−154
20.	$OH+OH \rightarrow H+HO_2$	8.00E-13	0	157.0	154
21.	$H_2O_2+OH \rightarrow HO_2+H_2O$	1.30E-11	0	6.7	−129
22.	$HO_2+H_2O \rightarrow H_2O_2+OH$	1.00E-13	0	129.0	129
23.	$HO_2+HO_2 \rightarrow H_2O_2+O_2$	4.00E-11	0	6.3	−167
24.	$H_2O_2+O_2 \rightarrow HO_2+HO_2$	1.00E-10	0	167.0	167
25.	$O+H_2O_2 \rightarrow H_2O+O_2$	4.70E-11	0	26.9	−178.5
26.	$O_2+H_2O \rightarrow O+H_2O_2$	8.00E-11	0	383.9	178.5
27.	$H+H_2O_2 \rightarrow HO_2+H_2$	1.20E-11	0	17.6	−67.2
28.	$H_2+HO_2 \rightarrow H+H_2O_2$	2.70E-06	0	84.0	67.2
29.	$O+H_2O_2 \rightarrow OH+HO_2$	4.70E-11	0	26.9	−58.8
30.	$OH+HO_2 \rightarrow O+H_2O_2$	4.80E-12	0	85.7	58.8
31.	$H+H_2O_2 \rightarrow H_2O+OH$	3.70E-09	0	49.6	−232.6
32.	$OH+H_2O \rightarrow H+H_2O_2$	4.00E-10	0	340.2	232.6
33.	$O_2+H_2O \rightarrow HO_2+OH$	1.25E-13	0.5	307.4	243.6
34.	$OH+HO_2 \rightarrow O_2+H_2O$	8.70E-14	0.5	89.0	−243.6
35.	$H_2+O_2 \rightarrow OH+OH$	1.38E-10	0	188.5	77.9

[a]For intramolecular reactions k_0 is expressed in units [cm^6/parts^2]; k_0, E, and n are parameters of the Arrhenius equation: $k = k_0 T^n \exp\left(-E/RT\right)$ where ΔH^0 is the standard enthalpy of reaction, and M is the arbitrary third particle

it is even more stable at these temperatures. Atomic oxygen is also obtained by radiolysis. First, we define the basic reactions which are important in obtaining a controlled hydrogen peroxide chain reaction induced by the conversion of water in the presence of molecular oxygen. Second, we look at what are the critical concentrations of atomic hydrogen for specific temperatures.

Table 9.4 Reduced value contributions ($\bar{h}_j$) of the individual stages of the reaction of water with oxygen ($t = 10^{-6}$ s)

1. 6.677D-02	2. −6.677D-02	3. −3.299D-02
4. 3.299D-02	5. 3.299D-02	6. −3.299D-02
7. 3.283D-02	8. −3.283D-02	9. −1.274D-03
10. 1.274D-03	11. −6.582D-02	12. 6.582D-02
13. 4.967D-05	14. −4.967D-05	15. −1.119D-04
16. 1.119D-04	17. −3.294D-02	18. 3.294D-02
19. 6.714D-02	20. −6.714D-02	21. −3.331D-02
22. 3.331D-02	23. 1.638D-03	24. −1.638D-03
25. 8.418D-04	26. −8.418D-04	27. −3.142D-04
28. 3.142D-04	29. 3.251D-02	30. −3.251D-02
31. 3.383D-02	32. −3.383D-02	33. 3.167D-02
34. −3.167D-02	35. 6.581D-02	

To this end, the problem is framed as how to achieve the maximum yield of hydrogen peroxide with minimal energy consumption. This problem is solved by the numerical simulation of reactions using the evaluation method of the kinetics of complex chemical reactions [15–18].

To find the optimal reaction control mode it is necessary to write the objective function. As the target functional we choose the concentration of hydrogen peroxide.

$$J = [\mathrm{H_2O_2}]_t = \int_{t_o}^{t_f} f_{\mathrm{H_2O_2}}\, dt \to \text{extremum} \tag{9.11}$$

The choice of the integrand in the target functional (9.11) determines the initial values $\psi_i(0)$, required for integrating the system of Eq. 9.3. Based on the type of target functional $\psi_{\mathrm{H_2O_2}}(0) = -1$ the original values $\psi_i(0)$ for all other chemical components of the reaction are zero.

Before we address the problem of the maximal functional (9.11) using the Pontryagin maximum principle, we determine the effect of atomic hydrogen and oxygen produced in the elementary chemical reactions, on the formation of hydrogen peroxide. With this aim, we calculated the value of contributions at T = 750 K in isothermal conditions; the results are given in Table 9.4.

Where: $\bar{h}_j = h_j \Big/ \left(\sum_0^L h_j^2 \right)^{\frac{1}{2}}$.

As the data presented in Table 9.4 shows, reactions with positive values are those in which oxygen atoms are formed. Conversely, those reactions leading to the formation of hydrogen atoms have a negative effect on the formation of hydrogen peroxide. From this it follows that the introduction of oxygen atoms into the reaction system has a positive impact on the formation of hydrogen peroxide and, conversely, the introduction of hydrogen atoms into the reaction system will decrease the output of $\mathrm{H_2O_2}$. Therefore, an increase in velocity of the 5th, 12th and 16th elementary stages,

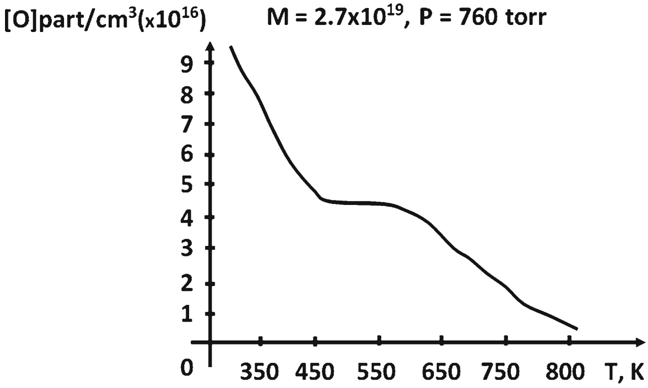


Fig. 9.8 The dependence of critical conditions (continuous shedding) on the initial concentration of atomic oxygen (in particles per cm^3) and temperature of the reaction system at atmospheric pressure

which correspond to positive value contributions, and increase the concentration of oxygen atoms, should greatly increase the yield of hydrogen peroxide. The results of direct calculations prove the validity of the findings.

We now consider the possibility of critical phenomena manifesting in the water vapor-molecular oxygen system, depending on the initial concentration of oxygen and the temperature. As is known, critical phenomena are a characteristic feature of branched-chain reactions [17], and can serve as additional information to determine the initial concentration of oxygen atoms at a given temperature needed for the triggering the reaction that leads to the formation of hydrogen peroxide. As noted above, for an optimal solution, this factor is important in determining the limits of control parameter changes; the selected parameter is the initial concentration of oxygen atoms.

As follows from the Pontryagin maximum principle, in the case of Hamiltonian linear dependence on the control parameter the quantity of oxygen atoms must be as large as possible to ensure the condition $\sup H = 0$. Obviously, the maximum possible amount of atomic oxygen is limited by method of its production and expediency. Therefore it is also important to find the lower limit of this amount (the 'minimum maximum') at which processes in the system both qualitatively and quantitatively will execute at high speed, i.e. when the system will be in a critical condition.

Critical system conditions were found according to the evaluative analysis of the kinetic reaction model [14]. The values obtained for the initial concentration of atomic oxygen shown in Fig. 9.8, correspond to the limits which mark the border between different modes of reaction, defined for the same initial reaction conditions which are: initial concentration of $[H_2O]_0 = 95$ mol%, $[O_2]_0 = 5$ mol% and $P_0 = 1$ atm. The parameters of the reactor are: S/V = 1 where V is the volume, S is the inner surface of the reactor, and the coefficient of thermal conductivity $\alpha = 2.7{\cdot}10^{-6}$ J/(s·m^2·K). The region at the bottom of the curve in Fig. 9.9 represents the almost imperceptible transformation of water, while the region above it represents an intense reaction.

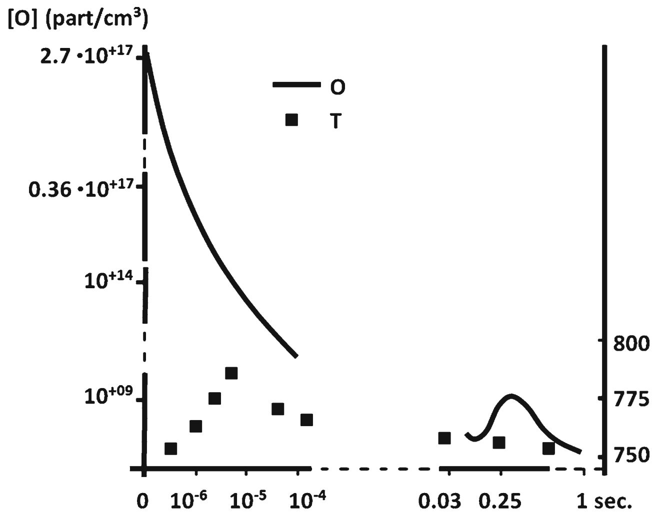


Fig. 9.9 Kinetics of change of temperature T (K) and concentration of atomic oxygen [O] in the supercritical region of the reaction system. [O] o=2.7I1017 part/cm^3. T0=750 K

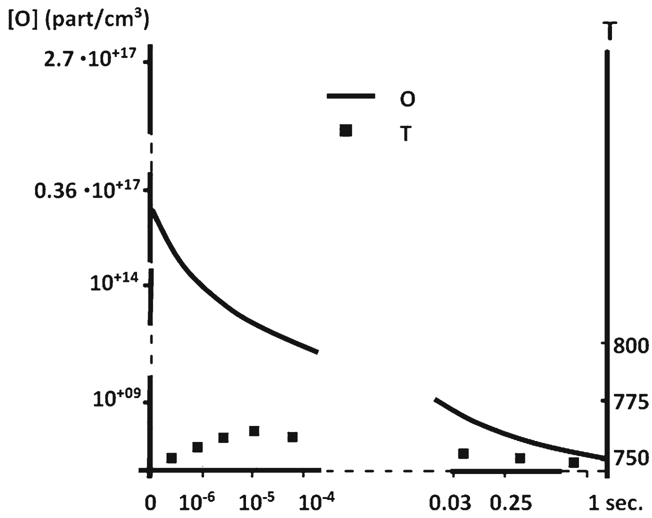
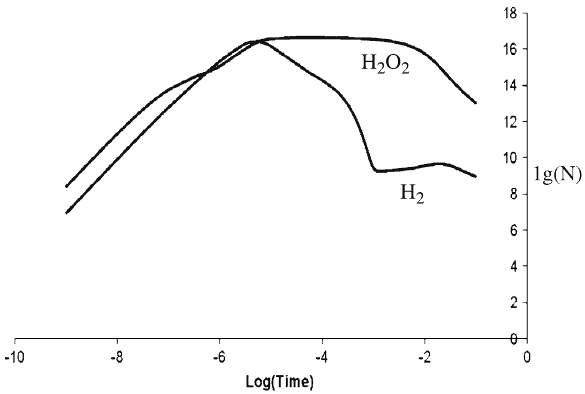


Fig. 9.10 Kinetics of temperature change T (K) and concentration of atomic oxygen [O] in the subcritical region of the reaction system. [O]0=1016part/cm^3. T0=750 K

It follows from the results shown in Figs. 9.9 and 9.10 that the consumption of the initial reaction components in the supercritical region of the reaction system differs from than in the subcritical region of the reaction. The concentration of atomic oxygen does not decrease continuously, but at certain points in time it rises and a concentration peak is observed which is attributable to a branching-chain reaction.

Fig. 9.11 Change of concentrations of hydrogen [H_2] (part/cm^3) and hydrogen peroxide [H_2O_2] (part/cm^3) depending on time in seconds. T0=750 K, [O] o=2.7·1017(part/cm^3)



It follows from the curve describing the critical conditions of the reaction (Fig. 9.10), that by ensuring the availability of oxygen atoms in the initial reaction at low temperatures the system can reach supercritical conditions. Under these conditions, (a high rate of reactions) hydrogen peroxide should be produced in large amounts.

Figure 9.11 below shows the results of calculations from which it is possible to predict that even with 1% of atomic oxygen in the total number of particles in the initial reaction system at T=750 K the amount of hydrogen peroxide produced does not exceed the corresponding level at T=1,500 K, but without the presence of atomic oxygen in the initial reaction system.

Therefore we have identified those reactions which make a major contribution to the production of hydrogen peroxide from water in the presence of molecular oxygen, and demonstrated that they can be attributed to the formation and use of atomic oxygen. For the same system we have obtained the critical dependencies of atomic oxygen concentrations on temperature.

Now let us return to the general question of whether the creation of critical concentrations of atomic oxygen in water-cooled containers is possible at low temperatures. Estimates were made for radiolysis conditions for given system and the role of admixtures is very important. Therefore a more detailed description of the kinetic model is needed. Performing calculations using the kinetic model shows that based on the potentially wide range of energies and intensities of radiation doses, the contribution of the reaction $e+O_2 \rightarrow O+O+e$ ($k_0=1.434\times10^{-9}$) alone can cause the atomic oxygen concentration to rise up to a level of $8\cdot10^{16}$ [O]/cm^3 which is a critical quantity at low temperatures (see Fig. 9.10).

9.7 Conclusions

1. The evaluative theory outlines the main reactions leading to the formation of hydrogen peroxide and ways to inhibit these reactions.
2. The evaluative theory opens great possibilities for the study of more complex kinetic systems (in terms of predictive capacity).

3. Under radiolysis conditions at low temperatures the formation of atomic oxygen at critical concentrations is possible.
4. The transportation and storage of radioactive waste must ensure subcritical conditions for oxygen depending on temperature.

References

1. http://www-formal.stanford.edu/jmc/progress/nuclear-faq.html
2. http://www.nrc.gov
3. http://www.nrc.gov/waste/spent-fuel-storage/pools.html
4. http://www.flickr.com/photos/12557829@N00/272376923
5. http://www.imo.org
6. http://www.icao.int
7. http://www.iata.org
8. http://www.unece.org/trans/danger/publi/adr/adr_e.html
9. http://www.upu.int
10. Pontryagin LS, Boltyanski VG, Gamkrelidze RV, Mischenko EF (1961) Mathematical theory of optimal processes. PH “Physmatgiz”, Moscow (in Russian)
11. Martoyan GA, Tavadyan LA (2004) Lecture notes in computer science, vol 3044. Springer, Berlin, p 309
12. Martoyan GA, Tavadyan LA (1992) Selective influence on the multicentre chain reactions. Kinet Catal 33(3):851–860. (in Russian)
13. Tavadyan LA, Martoyan GA (1994) Value principal in the study of kinetics of complex chemical reactions. Chem Phys 13(5):261–267 (in Russian)
14. Martoyan GA, Karamyan GG, Barseghyan AR (2005) The Influence of Monatomic Hydrogen on Critical Conditions of Branching Chain Reactions in Water Steam, Proceedings of NAS and SEUA of Republic of Armenia, Technical Series, LVIII(3):523–528 (in Russian)
15. Basevich VYA, Belyayev AA, Posvyanskiy VS (1982) Distribution of laminar flames. The role of H_2O_2 in the H_2-O_2 combustion systems. Chem Phys 6:842–847 (in Russian)
16. Ivanova AN, Andrianova ZS, Azatyan VV (1998) The Contribution from Quadratic Reactions to the Thermal Regime of Branched Chain Combustion. Chem Phys Rep 17:1511 pp 164–169 (in Russian)
17. Semyonov NN (1986) Chain reactions. PH “Nauka”, Moscow (in Russian)
18. Pikaev AK, Ershov BG (1967) Primary products of the radiolysis of water and their reactivity. Russ Chem Rev 36(8):602–620
19. Martoyan GA, Tavadyan LA (1998) Invariance principle in studying the effect of medium on chemical reaction. Chem Phys Rep 17(5):851–860