

Regional sustainability : tools for integrated governance

Citation for published version (APA):

Grosskurth, J. (2008). *Regional sustainability : tools for integrated governance*. [Doctoral Thesis, Maastricht University]. Roquefort Ontwerpers. <https://doi.org/10.26481/dis.20081209jg>

Document status and date:

Published: 01/01/2008

DOI:

[10.26481/dis.20081209jg](https://doi.org/10.26481/dis.20081209jg)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

REGIONAL SUSTAINABILITY
Tools for integrated governance

Jasper Grosskurth

The research underlying this dissertation was carried out at and supported by the International Centre for Integrated assessment and Sustainable development, previously known as the International Centre for Integrative Studies at Maastricht University.



ISBN 978 90 813676 1 5

© Jasper Grosskurth, 2008

jasper.grosskurth@gmail.com

All rights reserved. / Alle rechten voorbehouden.

Cover design by Roquefort Ontwerpers, Utrecht

www.roquefort.nl

REGIONAL SUSTAINABILITY
Tools for integrated governance

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit Maastricht,
op gezag van de Rector Magnificus,
Prof. Mr. G.P.M.F. Mols
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op dinsdag 9 december 2008 om 12.00 uur

door

Jasper Alexander Großkurth

Promotores:

Prof. Ir. J. Rotmans (Erasmus Universiteit Rotterdam)

Prof. Dr. P. Martens

Beoordelingscommissie:

Prof. Dr. M.G.W.M. Peeters (voorzitter)

Em. Prof. Dr. T.A.M. Beckers (Universiteit van Tilburg)

Dr. R.P.M. Kemp

Prof. Dr. J.B. Robinson (University of British Columbia, Canada)

CONTENTS

Sustainability in a complex world	9
1.1 A complex world.....	11
1.2 Sustainability	13
1.3 Governance for sustainability.....	15
1.4 Scope and objectives	20
1.5 Outline	21
The methodological context	23
2.1 Defining a mess	24
2.2 Sustainability as a mess.....	26
2.3 A new kind of science.....	28
2.4 Research methods	35
2.5 Research process	39
2.6 Case study.....	42
Sustainability assessment	45
3.1 Framework for tool selection.....	47
3.2 The scale and nature of system interventions	50
3.3 Tools and the scale of system interventions	54

3.4	Tools for fundamental system changes.....	61
3.5	Limits to shifting up the hierarchy	64
3.6	Conclusion	67
The SCENE approach to mapping systems.....		69
4.1	The SCENE structure.....	71
4.2	The SCENE process.....	74
4.3	Applying SCENE	79
4.4	Case study.....	84
4.5	Conclusion	87
enriched Qualitative Systems Analysis (eQSA).....		89
5.1	Methods for Qualitative Systems Analysis.....	90
5.2	Method and application of eQSA	95
5.3	Other case studies	114
5.4	Conclusions.....	120
QSSI: A New Type of Sustainability Indicator		123
6.1	Sustainability indicators.....	125
6.2	QSSI: Proposing a System Indicator	132
6.3	Calculating QSSI	136
6.4	QSSI extensions	138
6.5	Conclusion	141

Simulation models.....	145
7.1 Simulation models: an overview.....	147
7.2 Case study: Quantitative models for strategic policy planning..	165
7.3 Models and joined-up thinking.....	174
7.4 Modelling experiments.....	178
7.5 Conclusion	180
Synthesis & Conclusion	183
References.....	193
List of acronyms.....	213
Samenvatting.....	217
Acknowledgements.....	221
Curriculum Vitae	223

Chapter 1

SUSTAINABILITY IN A COMPLEX
WORLD

We live in a world that is already in fact very different from the one which we have begun to comprehend, and by the time our comprehension has caught up with the new reality, the world is likely to be even more drastically different in ways that today may seem unthinkable. (Zbigniew Brzezinski ¹)

The world around us is becoming increasingly complex, difficult to analyse, and fast paced. The concept of 'sustainability' is born from the awareness that our ability to serve the needs of current and future generations is threatened as a consequence of these changes (WCED, 1987). Sustainability encapsulates the ambition of a balanced development of interacting natural, social, and economic systems over many generations. Over the past two decades, sustainability has been formulated as an important target at all levels of government. However, incidental and fragmented changes to the functioning of our society and the way it interacts with the nature system have little effect. This awareness has given rise to the call for more integrated and fundamental change in the form of a sustainability transition. Such a transition confronts decision makers with unprecedented challenges. Identifying strategic priorities and actions is difficult in an environment of complexity, uncertainty, and urgency. A more integrated conceptualisation of complex problems in policy and research is needed to better understand and address these difficulties. This dissertation contributes to this integrated conceptualisation. It helps decision makers at the regional level to integrate the fragmented information available to them, to comprehend sustainability in the context of their own day-to-day responsibilities, and to efficiently identify policy levers for transformative change. This chapter provides a general introduction to sustainability and governance, outlines the scope and objectives of the research covered in this dissertation, and provides an overview of its structure.

¹ This quote was found on the homepage of the Sustainability Institute (www.sustainer.org) in early 2002. I did not succeed to trace its original place of publication.

1.1 A complex world

We live in the 'anthropocene', an era in which natural and human dimensions of the earth system have become inseparable (Crutzen & Stoermer, 2000). Changes in the human system are key drivers for changes in the nature system and vice versa. The Millennium Ecosystem Assessment (MEA) (2005), conducted under the auspices of the United Nations, is the most thorough inventory to date of the state of the world's ecosystems. It concludes that changes in the nature system over the past five decades have been faster and larger than ever before during human history. The MEA lists a wide range of symptoms of that change. These include the following, among many others: species are estimated to go extinct at a rate 1000 times higher than in the distant past; the genetic diversity among species declines globally; more and more ecosystems are being converted, mostly to agriculture; and the availability of nitrogen and phosphorus in ecosystems increases significantly. For the coming five decades, the MEA states that *'the degradation of ecosystem services could grow significantly worse [...]'*.

Changes in the human system have been at least as fast and large. The social, economic and technological systems are becoming increasingly complex: global interconnectedness is widening, deepening and speeding up in every respect (Held, 1999); processes at the global and international level more and more interfere with processes at the national and local level and vice versa (Wilbanks & Kates, 1999); and power and influence is decentralising and values differ widely among groups. While all of these developments contain positive elements, such as easier communication or increased individual freedom, it becomes increasingly difficult to comprehend, predict, and manage the human system and steering problems arise (Rotmans, 2005). Not only the complexity, but also the scale of changes is breathtaking. Economic development illustrates this. During the years 2000-2005, the average of annual absolute growth of the world economy roughly equalled the total size of the economy at the beginning of the 20th century (IMF (2007); Maddison (1995)).

The impact of the human system on the nature system is frequently illustrated on the basis of the Ecological Footprint and the so-called Human Appropriation of Net Primary Production (HANPP). The Ecological Footprint estimates the land surface required to sustain a population and its lifestyle relative to the land surface available (Wackernagel & Rees, 1996). It was estimated to be 1.25 in 2003 (WWF, 2006). This implies that the

resources of 1.25 planets would be required to sustain the current world population and its activities. HANPP estimates the share of total ecosystem production being put to the use of the human species (Vitousek et al., 1986). Human appropriation, for example through intensive agriculture, changes vital ecosystems, food webs, and energy and resource cycles. HANPP for the year 2000 has been conservatively estimated to be 24% (Haberl et al., 2007). The Ecological Footprint and HANPP should be interpreted with some care. Both concepts are scientifically contested, partly for suggesting an exactness and certainty that doesn't match with what many consider a process of oversimplification and omission. Nevertheless, their key message is widely accepted: the nature and scale of the human impact on the world's nature system is unprecedented and in many cases well beyond sustainable thresholds (McNeill, 2001). The MEA lists human induced habitat change, over-exploitation and pollution as the main drivers of that development.

The impact of the nature system on the human system, is also receiving increasing attention. The MEA (2005) concludes:

The changes that have been made to ecosystems have contributed to substantial net-gains in human well-being and economic development, but these gains have been achieved at growing costs in the form of many ecosystem services, increased risk of nonlinear changes, and the exacerbation of poverty for some groups of people. These problems, unless addressed, will substantially diminish the benefits that future generations obtain from ecosystems.

According to the MEA, the degradation of ecosystems affects humans by compromising provisioning services (such as fisheries and fresh water), regulating services (including air quality regulation, erosion regulation, and hazard regulation), as well as cultural services (in the form of spiritual and aesthetic values). There are indications that the costs of these losses already surpass the gains, especially when non-marketed ecosystem services are taken into account. A previous assessment of the environment, the Global Environmental Outlook 3 (GEO-3) lists a number of specific impacts of environmental change on the human system: deteriorating renewable and non-renewable natural resources threaten livelihoods in many parts of the world and trigger migration; loss of fertile soil and desertification threaten food security; insufficient environmental quality is an important cause of poor health; the changing environment brings with it high economic losses (UNEP, 2002). In the Netherlands, environmental costs are estimated to amount to 5% of GDP with an upward trend (RIVM, 2001).

GEO-3 also comments that the interdependencies between the human system and the nature system differ between regions, as do the coping capacity, vulnerability and resilience of societies. Generally, politically stable, affluent and knowledgeable societies find it easier to cope with changes than ones characterised by fragility, poverty and lack of education (UNEP, 2002). However, even wealthy western democracies find it increasingly difficult to avoid the negative side effects of a complex lifestyle dependent on economic growth. Increasing interconnectedness, scale-enlargement, and time acceleration are just as relevant on a regional and local level, as they are globally. Many of these processes are now well documented. The increasing quality and quantity of data, and innovative methods of analysis, as well as other processes of learning lead to an increase in knowledge. More often than not this new knowledge gives rise to new questions and the awareness of yet other uncertainties. The metaphorical wing of a butterfly triggering a hurricane becomes an increasingly frequent occurrence and a challenge for decision-makers everywhere. Many of the symptoms of the pressures on the human – nature system and their possible consequences have been discussed at length for the past half-century. However, the awareness concerning these symptoms and the relationships between them is growing as we approach or irreversibly transgress thresholds of degradation, and as we are more and more faced with surprises and discontinuities (Rotmans, 2005).

1.2 Sustainability

The term ‘sustainability’ and the closely related ‘sustainable development’ refer to the ambition of a balanced development of the combined human-nature system.² The Brundtland Commission has defined the term in its report ‘Our Common Future’ (WCED, 1987):

Sustainable development is development that meets the needs of the present generation without compromising the ability of future generations to meet their own needs.

² I use the term ‘sustainability’ throughout this dissertation, unless direct quotes or names refer to ‘sustainable development’. For more details on the difference between the concepts, see section 2.2 and Robinson (2004).

Generally, sustainability is perceived as having three domains: the environment, the society and the economy. There are many interactions between the three domains and within each domain many more interconnected facets exist. Sustainability is a concept that forces us to think about the long-term. Sustainability integrates aspects of daily life in the present with causes and impacts taking place decennia, centuries and millennia before and after the present (Costanza et al., 2007).

It is a popular sport at scientific conferences to open a speech with the oldest or most surprising reference to sustainability in the general literature, the Bible and Karl Marx's *Das Kapital* being strong contenders. In its current meaning, the concept has grown out of the increasing awareness of the ecological consequences of human activity and of the exponential character of growth of the economy and the population. Arguably, Rachel Carson's *Silent Spring* (1962) and Paul Ehrlich's *The Population Bomb* (1968) first spread this awareness among the general public. The fragmented insights of the 1960's were integrated in the 1970's. The publication of *The Limits to Growth* (D. H. Meadows et al., 1972) was a milestone in that respect and triggered analogous lines of research in many fields. In the 1980's and 1990's, especially after the report of the Brundtland group, the terms 'sustainability' and 'sustainable development' became widely used in policy and research.

Securing sustainability affects almost every aspect of society, for example, technology (McDonough & Braungart (2002); von Weizsäcker et al. (1995)); natural resource use (McNeill, 2001); business (Hawken et al., 2000); poverty, conflicts and international relations (OECD, 2001); and institutional organisation (UNEP, 2002). The resulting system is extremely complex, evolves in a non-linear fashion, is subject to a high degree of uncertainty, and normative judgements are inevitable when assessing sustainability.

Virtually all actors of governance have formulated sustainability as an important target. It ranks high on political agendas. Two world summits in Rio (1992) and Johannesburg (2002) have addressed the issue on a global scale. Sustainability has become a mantra in non-governmental circles, and even permeated the financial and business world. Hawken (2007) counts more than 1 million initiatives addressing sustainability or relevant aspects of it. These include NGOs, citizen groups, government programmes and the like. The concept of sustainability has become almost synonymous with any sound strategy for a better world. In chapter 2, we will take a closer look at the characteristics of the concept and its implications.

Despite the spreading awareness, we are far away from achieving sustainability, independent of our perception of the concept. Relevant studies and cases frequently conclude that to achieve sustainability, fundamental change is necessary, new structures are required, a paradigm shift needs to occur (National Research Council (1999); Rotmans, Kemp et al. (2001b); Martens & Rotmans (2003); Parris & Kates (2003); Edwards (2005); Millennium Ecosystem Assessment (2005)). This transition needs to take place at all levels of society.

1.3 Governance for sustainability

Achieving a transition towards sustainability in a complex world poses unprecedented challenges for decision makers (Rotmans et al. (2001b); Martens & Rotmans (2003)). Lack of information, uncertainty, interdependencies, discontinuities, and urgency characterise the system in which they are operating. Identifying strategic priorities and actions becomes very difficult in this environment (van Asselt, 2000). Decisions have to be taken even though we are not yet ready for them and priorities have to be set in the mist of uncertainty. The difficulties and dangers of trying to manage, control or manipulate complex systems have been widely documented: complex systems often behave in unexpected ways and counter-intuitively (Forrester, 1971); leverage points suited to influence system developments in a controlled manner are difficult to find and even more difficult to use (D. H. Meadows, 1999); people do not perform well when it comes to estimating the behaviour of systems, even in very simple cases when the full structure of the system is known (Sterman (1994); Bossel (2004)); a lack of vision and commitment often compromise actions born out of good intentions (Loorbach, 2007); and people learn very little from system (management) failures and mistakes (Dörner, 2003). Most of these difficulties appear even if the authority for intervening in the system is unquestioned. Once the authority, and thus also the meaning of the concept of sustainability, are questioned by different stakeholders, massive steering problems arise.

Traditional forms of governance seem ill suited to address the changing circumstances. Rotmans (2005) considers top-down governance and free market liberalism as *'outmoded as effective management mechanisms to create sustainable solutions at the societal level'*. These outmoded management

mechanisms have been relatively successful at addressing isolated issues and at triggering gradual change. However, at the societal level, problems seem to exacerbate as a consequence of applying traditional modes of governance in a changed world, leading to government and market failures, as well as so-called persistent problems (Rotmans, Kemp et al., 2001a). We will take a more detailed look at these persistent problems in section 2.1.

Addressing the difficulties of governing a modern world requires the amalgamation of knowledge from many different disciplines and perspectives (Sterman, 2000). The complex dynamics force us to think and act in a more integrative manner (Rotmans, 1998a). The different values of stakeholders combined with increasing uncertainty and high stakes force us to do so in a manner allowing for meaningful societal participation and accountability (Ravetz (1999a); Gibbons (1994)). For policy makers and researchers, who in the past held quasi-monopolies of power and knowledge, respectively, the changes require substantial adjustments in their self-perception and their work. On a higher level, these adjustments require and result in alternative modes of governance. Rotmans (2005) lists four characteristics distinguishing such new modes: the network approach, the interactive approach, the pluralistic approach and the multi-level approach. All of these exhibit an increased level of interaction among stakeholders, compared to traditional top-down approaches (Dirven et al., 2002). The increase in interaction goes together with a process of social learning (Leeuwis (2003); Social Learning Group (2001)), taking into account the social context (Loeber, 2004), as well as an ongoing process of reflection (Voss et al., 2006). The new forms of governance do not only have an impact on the ways in which decisions are being taken, but also on the process of gathering and processing information.

A symptom of a changing role of governance in the context of (environmental) sustainability is the adoption of the UNECE Convention on Access to Information, Public Participation in Decision-making and Access to Justice in Environmental Matters (UNECE, 1998). The Convention explicitly acknowledges that sustainable development can only be achieved through the involvement of all stakeholders. It links governance and environment, focusing on the interaction between the public and public authorities on environmental issues. The Aarhus Convention is a milestone in public participation since it explicitly grants the public rights concerning participation and imposes on public authorities obligations regarding access to information and access to justice.

Four dimensions of integration summarise key issues in the required transformation of science-policy interaction: integration among policy domains or 'joined-up thinking'; integration among scientific disciplines; integration of relevant research in policy processes; and normative integration (derived from Boulanger (unpublished) and Scrase & Sheate (2002)).

1. Integration among policy domains

Integration among policy domains calls for a horizontal and vertical integration in and among policy-making institutions (Rotmans, 1998a). A lack of such integration manifests itself in one department undermining efforts of another, with the consequence of wasted resources and disappointing results. The British New Labour government coined the term 'joined-up' policy as a label for such integrative governance (Lyall & Tait, 2005). Joined-up policies avoid the contradictions, inconsistencies, and inefficiencies caused by sectoral policies, and engage stakeholders as a check of desk studies. The term has now somewhat disappeared from policy agendas, arguably due to a lack of successful implementation, rather than the undesirability of joined-up policies (Caulkin, 2006).

2. Integration among scientific disciplines

In the same way that policy makers need to integrate their different sectoral approaches, researchers dealing with complex problems need to integrate relevant concepts from different scientific disciplines. This contrasts with the compartmented and departmentalised structure of research in the past. Together with the scientific disciplines, the tools and methods used have to be integrated. Morrison-Saunders & Therivel (2006) provide a 7-point scale of levels of integration with respect to sustainability assessments. Environmental assessments including notes on economic and social impacts are at the bottom of this scale. Assessments guided by clear integrated principles for sustainability and decision-making trade-off rules are at its top. Economic, environmental and social pillars are treated equally and in full interdependence. It is this interdependence that requires a systemic understanding. The relationships between different aspects of sustainability take precedence over the states of the individual aspects. Only the ability to amalgamate relevant strands of knowledge from different disciplines will enable researchers to address the complex issues described above (Rotmans &

van Asselt, 1996). Crosscutting questions, such as the quest for sustainability require and facilitate this integration.

3. *Integration of relevant research in policy processes*

Finally, the knowledge resulting from processes of inter-disciplinary and participatory knowledge creation must find their way into the decision-making processes. This may seem the most obvious type of integration among the four, but in reality proves to be hard. Researchers are often frustrated by ‘politicians not listening to their advice’ and policy makers are frustrated by researchers ‘not answering their questions meaningfully’. This type of integration calls for formats that are relevant for integrated knowledge creation and policy decision making alike.

4. *Normative integration*

Researchers and policy makers alike have to open up to processes of participation and social accountability. Meaningful participation of stakeholders inserts different normative perspectives and values into the process (van Asselt & Rijkens, 2002). This is inevitable since decision with high stakes have to be taken in a context of high uncertainty (Ravetz, 1999a). Fairly doing so must be a societal process with the possibility of broad participation. For the researcher this also implies that the traditional accountability to peers is now extended and becomes accountability to society; for the policy maker, the democratic process is extended beyond the voting cycle.

These four types of integration are epistemologically driven. They constitute necessary, but not sufficient conditions for the sustainability transition. Realizing the sustainability transition and the intended contribution of integration to it requires a lot more. For example, political and societal inertia, power relationships, and the quality of governance need to be addressed. This dissertation does not explicitly do so. Another field of research not explored in this dissertation is the legal context of joined-up thinking for the sustainability transition. Since the adoption of the Aarhus Convention in 1998 (UNECE, 1998), several Dutch national laws have been adapted or created, explicitly granting the public the right to participation in environmental decisions and assessments. Unfortunately, the interaction between the collaborative research typical for Post Normal Science and legally prescribed processes of participation is often neglected and this

dissertation is no exception. At the time, the legal domain seemed of little relevance to the project, which probably reflects our ignorance more than the actual importance of the issue. The inclusion of such aspects remains a challenge for future research.

Nevertheless, we are convinced that the focus on the epistemological aspects is valuable for Integrated Assessment of sustainability. Progress in that direction would improve our ability to integrate different types of information and strengthen the understanding of complex sustainability related issues by different stakeholders. If research produces any progress along that path, the lessons learned from the underlying research might well be applicable in different regimes of power and governance. Indirectly, through helping to structure unwieldy complex problems and supporting the quest for truly sustainable pathways of development, the epistemological quest, together with other strains of research, might motivate political and societal players to overcome their inertia and act.

A hierarchy is implicit in the four kinds of integration. The integration among policy domains is what we want to support. At the science policy interface, the integration among policy domains would seem to benefit from an analogous integration on the side of the scientists. This does not imply that all science must be integrated, but rather that there must be relevant integrated science. The relevance of that science facilitates the integration of research in policy processes and thus links the previous two. The normative integration affects all three other types of integration. In chapter 2, this aspect will be treated in more detail.

To achieve an operationalisation of sustainability in governance academics developed a large set of tools, methods and approaches for sustainability assessment to facilitate the sustainability transition (see e.g.: Kates et al. (2001); Robèrt et al. (2002); Scholz & Tietje (2002); de Ridder (2006)). ‘Sustainability assessment’ concerns a subset of these tools, methods, and approaches explicitly aiming for the integration of sustainability concepts into decision-making (Pope, 2006). These approaches employ a broad and heterogeneous range of methods that we will take a closer look at in chapters 2 and 3. However, sustainability assessment is still in its early stages. This dissertation is about helping sustainability assessment mature.

1.4 Scope and objectives

1.4.1 SCOPE

This dissertation contributes to the four-fold integration described in the previous section. Global changes in the combined human-nature system affect governance processes at all scale levels. Appropriate responses require the combined efforts of all agents of governance. In our work, we focus on the role of regional governments in the Netherlands. The capacity to act of these governments is heavily constrained by international and national rules and regulations, and relatively limited resources. Their main policy levers are sectoral and largely disconnected. As a consequence, the administrations of these regional governments greatly benefit from efficient ways to identify robust policy strategies, effective policy levers, and possible side effects resulting from their use. A culture and capacity of joined-up thinking within the organisation ensure the consistency of different policies, the efficient use of resources, and the effectiveness of political action.

In our work, we have developed a set of tools to support these regional decision makers. The tools are meant to help structuring the complexities of regional sustainability in a manner that is accessible to the target audience, to aid the identification of systemic problems and solutions, and to stimulate joined-up thinking. We have co-developed and applied these tools in a specific setting and in co-operation with the Province of Limburg. We elaborate on the possibility to generalize or transfer the lessons learned from our work to other regions and other sustainability related issues at the end of each methodological chapter (chapters 4-7) and in the conclusion (chapter 8).

1.4.2 OBJECTIVES

This dissertation advances an integrated systems perspective of sustainability in regional decision- support. It is the objective to substantiate joined-up thinking at the science-policy interface. Consequently, the core question in this dissertation is:

How can tools for sustainability assessment stimulate joined-up thinking in political decision processes?

Four sub-questions concern the qualities of existing tools, possible extensions, the combination of different tools, and the use of assessment tools in policy processes:

1. How appropriate are existing assessment tools to stimulate joined-up thinking?
2. How can existing tools be extended to improve their capacity to stimulate joined-up thinking?
3. How can the integration among different sustainability assessment tools be improved?
4. How can the integration of assessment tools in political decision-making be improved?

The research is situated at the science-consultancy-policy interface and aims to mobilise the strengths of these three domains in order to improve decisions with long-term implications.

1.5 Outline

Chapter 2 provides the scientific context and structure of the dissertation research. It discusses the kind of science required for sustainability and decision support in a complex world. Key concepts in this context are Post-Normal Science, Integrated Assessment, and Sustainability Science. The research approach follows the basic principles of these concepts.

Chapter 3 thoroughly reviews existing tools for decision-support used in the context of sustainability. The evaluation of these methods results in a set of methodological ambitions and a set of research challenges. The methodological ambitions provide the benchmarks against which each of the tools described in the subsequent chapters will be evaluated. The research challenges concern the capacity of all four tools combined to support joined-up thinking.

Chapters 4, 5, 6 and 7 explore four interdependent tools at the interface of sustainability and decision-support. Chapter 4 introduces the SCENE approach to mapping complex systems for sustainability assessments. The approach is participatory and qualitative, resulting in rich system diagrams.

These have an added value in their own right, as they facilitate and stimulate joined-up thinking. But they are also used as input for the tools described in the subsequent three chapters.

Chapter 5 introduces the eQSA approach for the structural qualitative analysis of complex systems. eQSA analyses the role of different elements in a system, identifies policy levers and possible pitfalls, and explores uncertainties. The analysis results in guidelines for the implementation of policy strategies and provides a framework of integration for assessment tools. The approach is a key element in our ambition to improve integration among assessment tools and the integration of assessment tools in policy-making.

Chapter 6 explores the subject of sustainability indicators. A short review of existing indicators is followed by the introduction of a new kind of qualitative indicator, the QSSI. This experimental method of monitoring and stimulating progress towards sustainability is also based on the SCENE diagram. The approach is unique in its focus on the relationships between different aspects of sustainability and results in an agenda for a sustainability transition.

The first part of chapter 7 contains a review and evaluation of relevant quantitative models. The second part provides a detailed analysis of a set of models used in regional strategic planning, their coverage and their appropriateness for joined-up policies. The examination results facilitate the embedding of the quantitative models in qualitative processes, as well as guidelines for the efficient model improvement.

Chapter 8 concludes this dissertation and evaluates the methods presented in chapters 4, 5, 6, and 7 based on the methodological ambitions and research challenges formulated in chapter 3. It evaluates the impact the different tools have had on the political decision process and it explores the further research paths.

A case study and several case experiments illustrate all approaches described. The case study runs through all chapters except chapter 6 (on indicators) and concerns the Dutch Province of Limburg. The regional administration was actively looking for ways to achieve a culture of joined-up thinking within the organisation. Case experiments are a preliminary test of concepts. These experiment with the application of innovative tools or methods without achieving the depth of the case study.

Chapter 2

THE METHODOLOGICAL CONTEXT

English does not contain a suitable word for 'system of problems'. Therefore I have had to coin one. I choose to call such a system a 'mess'. The solution to a mess can seldom be obtained by independently solving each of the problems of which it is composed. (Russell L. Ackoff (1974))

Uncertain, complex, and, as we will see in section 2.2, normative problems are a challenge for scientific processes. Only quite recently were so-called unstructured, wicked, and persistent problems added to scientific research agendas. Traditional scientific methods struggle with the meaningful analysis of such problems. A new kind of science is needed, especially designed for these characteristics. Mode2 science, post-normal science, integrated assessment and sustainability science are the corner posts of this new science in the context of sustainability. This dissertation aims to advance the methods used within these fields and applies them for the benefit of regional decision-making.

This chapter provides the broader scientific context of the dissertation research. Section 2.1 provides an overview of complex societal problems in the scientific literature. Section 2.2 takes a closer look at sustainability as a complex problem. In section 2.3 we explore the new kind of science needed to address sustainability with scientific methods. The research methods applied in this dissertation are discussed in section 2.4. Section 2.5 outlines the research process.

2.1 Defining a mess

Striving for sustainability in a complex world brings with it the necessity to scientifically address a class of problems that has long been considered unfit for scientific analysis: the so-called unstructured, wicked, or persistent problems. Those who defined such problems seem to have made a deliberate effort to exclude any aspect facilitating a scientific process and to include aspects underlining the urgency of conducting such a process.

1. *Unstructured problems*

Hisschemöller & Hoppe (1996) define unstructured problems based on the criteria of 'certainty on knowledge' and 'consensus on values'. A

problem qualifies as unstructured, when both criteria score a low value. In other words, we have little certainty about the scope, structure and factual details of the problem. And even if we had more certainty on this account, there would likely still be widespread societal dissent on what a good strategy or a good outcome would be.

2. *Wicked problems*

Rittel & Webber (1973) have coined the term ‘wicked problem’. Such problems are defined by the lack of a generally accepted scientific definition, complex dynamics, ambiguous analysis, and subjective and normative choices. Wicked problems have a connotation of being wicked in terms of governance. The culture of governance has to change to solve wicked problems.

3. *Persistent problems*

Rotmans et al. (2001a) and Dirven et al. (2002) extend the list of characteristics of wicked problems to define persistent problems and can be seen as the superlative of wicked problems. They add ‘steering difficulties’ due to the interaction of stakeholders. Current symptoms of un-sustainability have multiple interacting causes, often deeply engrained in societal structures, covering many fields. A structural change in governance will not be sufficient to solve persistent problems, unless it goes hand in hand with fundamental shifts in the structure of society. Nevertheless, solving wicked problems contributes to solving persistent problems.

The three types of problems and their definitions overlap to a large extent. Also, it holds for all three that ways to solve them are far from obvious. Partial solutions are often ineffective, and, worse still, frequently contribute to the exacerbation of a problem. Especially the concept of persistent problems emphasizes how partial solutions contribute to the development of permanent societal lock-ins. For the traditional sciences, such problems were considered out of bounds for two reasons: uncertainty and normativity. In his unfinished *Rules for the Direction of the Mind* (about 1619), Descartes writes:

Those who seek the direct road to truth should not bother with any object of which they cannot have a certainty equal to the demonstrations of arithmetic and geometry.

And for those who ignore his advice and encounter persistent problems, he continues:

If in the matters to be examined we come to a step in the series of which our understanding is not sufficiently well able to have an intuitive cognition, we must stop short here. We must make no attempt to examine what follows; thus we shall spare ourselves superfluous labour. (translation by Haldane & Ross, 1971)

If scientists are ‘*those who seek the direct road to truth*’, then Descartes leaves no doubt that in his opinion any effort spent on persistent problems is wasted. Yet, as laid out in the previous chapter, there is an urgent need to ignore this advice. We must study persistent problems. And scientists cannot opt out of the societal process of finding strategies to solve persistent problems.

2.2 Sustainability as a mess

Sustainability is primarily a political concept. It cannot be derived from first principles, nor can it be defined objectively. Sustainability is complex, ambiguous, normative and subjective (Rotmans & van Asselt (1999); Weaver & Rotmans (2006)). In other words, sustainability well deserves the overlapping labels unstructured, wicked, and persistent.

1. Complexity

All the elements of sustainability are connected and form a system that exhibits a behaviour that can often not be explained from the behaviour of its elements. The complexity is a result of the dense interconnectedness of processes at different temporal and geographical scale levels. Economic, social, environmental and institutional aspects of the human-nature system, form intractable feedback cycles and result in trade-offs between conflicting targets. These change over time, and the speed of that change is increasing (see chapter 1).

2. Ambiguity

Sustainability is a multi-faceted concept. It lacks a clear guidance on resolving trade-offs between conflicting targets concerning different facets of sustainability. The set of targets itself is also indistinct, lending itself to multiple equally compelling interpretations.

3. *Subjectivity*

Even though there is incomplete information on current implicit values and even less on future values, sustainability cannot be meaningfully defined without these values. Subjective individual values, as well as societal values (possibly conflicting with individual values) determine sustainability as much as physical characteristics do.

4. *Normativity*

There currently seems to be a wide consensus about the importance of inter-generational equity, the normative aspect of sustainability. However, the equal valuation of the needs of current and future generations remains an arbitrary norm.

The characteristics of the concept of sustainability are also reflected in ongoing conflicts about its interpretation. A general conflict concerns the distinction of strong and weak sustainability. Weak sustainability assumes that it is theoretically possible to substitute natural capital with human capital; strong sustainability rejects that option (Pearce et al., 1989). As a consequence, technological solutions, and the growth of the economic capital are seen much more positively within a framework of weak sustainability compared to one of strong sustainability. Related is the discussion about the distinction between 'sustainability' and 'sustainable development'. The general thrust of the two concepts is the same. But especially academics and NGOs dislike the emphasis of development, often interpreted as economic growth by private sector and government actors (Robinson, 2004). This might seem an academic discussion at first, but the implications of the different perceptions cause substantially different views on strategies for sustainability. The role of GDP in different sustainability indicators illustrates this. All widespread sustainability indicators have a strong correlation with GDP. Strikingly, for some the correlation is a positive one (for example the Genuine Progress Indicator GPI), for others the correlation is a negative one (for example the Ecological Footprint). While the first perspective would encourage 'good' economic growth as a pathway to achieving sustainability, the second perspective would judge it counter-productive.

The analysis of sustainability is not possible from a single discipline, a single perspective or even a single definition. In order to evaluate the sustainability of a system, we would optimally take into account time, scale and domain. A measure of sustainability should represent system changes that are relevant

in the long-term of 25 to 50 years. It should reflect developments within the system and trade offs to systems on other scale levels. It should cover the economic, ecological and social aspects of sustainability, and incorporate normative, and subjective aspects.

2.3 A new kind of science

Unstructured, wicked, and persistent problems are not easily compatible with the traditional scientific process. Two partially overlapping concepts have prepared the ground for the scientific analysis of such problems. Gibbons et al. (1994) proposed the concept of Mode2 science, Funtowicz & Ravetz (1993) the concept of Post-normal science. Both concepts aim to extend the notion of traditional science by accommodating normativity and uncertainty within the scientific process through, among others, the meaningful engagement of stakeholders. Key notions in Mode2 science are trans-disciplinarity, heterogeneity, a societal context and societal accountability. This contrasts with the homogenous disciplinarity, academic context and academic accountability of traditional science. Post-normal science emphasizes high decision stakes and high systems uncertainties. This contrasts with so-called 'normal science', where stakes would be irrelevant and uncertainties low (Kuhn, 1970). Overall, the proposed new science is issue-driven and problem-focused in a context of disputed values, high stakes and urgent decisions. The relevant peer-community is extended beyond academia and includes all those affected who are prepared to enter into a dialogue about it. In the words of Robinson (2008), the practitioners *'do not find themselves at the margins between disciplines, but in the sometimes uncomfortable borderlands between the academy and the larger world.'*

Both concepts are contested. Caswill & Shove (2000) summarize some of this critique in their introduction to the field of interactive social science, which Gibbons, the main author of Mode2 science, himself considers to be one branch of Mode2 science. In their view, three points of critique are especially relevant:

[T]here is a case for distance from the subject of study; too much close interaction can be a threat to proper development of theory; and the need for quality control of research and results.

In other words, the interactivity of Mode2 and Post Normal Science requires giving up the critical distance from the object of study. The continuous involvement with actors outside the academic domain is seen by some to compromise the ability to develop scientific theories. And even though some traditional measures of quality control, for example empirical falsification, are not always appropriate for this new kind of science, there is little agreement on what other quality control measures should take their place. If one emphasizes this 'relaxation' of scientific rigor and ignores attempts to establish suited substitutes, then one would have to conclude that Mode2 and Post-normal Science are neither new, nor different, but rather no science. Taking into account the relative youth of both disciplines one could still lament the lack of scientific methods to make explicit all hidden values, handle uncertainties and deduct scientific proof from outside the domain of science. A non-academic point of critique is that the social accountability of both concepts can also interfere with choices that should really be taken in the democratic domain.

Despite this criticism, Mode2 science and Post-normal science are now a practical reality in many fields of research, especially in the social scientific study of long-term issues, such as climate change and sustainability. And there are good reasons for it. Persistent problems cannot be captured with traditional scientific methods. However, our understanding of these problems can be improved and this understanding benefits from the qualities of the scientific process. As is the case in many social sciences, such an understanding is largely qualitative, case-based, reflexive and context-sensitive (see e.g. Voss et al. (2006) and Biermann (2007)).

The term context sensitivity describes the far-reaching consequences of processes of co-production with users and stakeholders. Gibbons (2000) sees context sensitivity as an inevitable symptom of an ever more interactive societal system, where institutions, such as universities and governments, have become more and more permeable. According to Gibbons, the boundaries between these institutions are weakening or even disappearing, while society '*can (and will with increasing frequency) speak back to science*'. Consequently, context sensitivity influences what research is being done (goal), how it is being done (process), and with whom (interaction). In addition, the role of the researcher changes. Where traditional sciences take the researcher to be a problem solver, Mode2 and Post Normal Science are calling on the researcher to become a problem recogniser or problem mediator (Hisschemöller & Hoppe, 1996). All this makes context sensitive

research a complex undertaking. However, the rewards for this approach lie in the generation of knowledge that is not only scientifically reliable, but also socially robust.

While Mode2 and Post Normal Science are necessary to address persistent problems, they are not sufficient. Integrated Assessment (IA) and Sustainability Science are two fields of research that were crucial in translating the general characteristics of the new sciences into tools and methods for the analysis of sustainability and other persistent problems.

Rotmans (2006) defines IA in the following words:

Integrated Assessment is the science that deals with an integrated systems approach to complex societal problems in a process-based context.

It is a relatively young field of research striving for a sound understanding of complex issues for better decision-making. Critics warn that Integrated Assessment (IA) is superficial as a consequence of its ambition to capture hugely complex systems in relatively simple concepts. Proponents of IA, however, argue that the complexity of the issues demands an integrated approach to ensure that key interactions, feedbacks and effects are not inadvertently omitted from the analysis. And, despite the youth of the field, IA practitioners do have a track record in support of their argument for more integration.

Initially, IA focused on the topic of climate change, where IA furthered the integration of the social dimension in simulation models dominated by the natural sciences. Models, such as RAINS (Alcamo et al., 1990) and IMAGE (Rotmans, 1990) are characteristic for that early generation of IA projects. Over the years, the field quickly broadened its horizon to other wicked problems. In the process, stakeholder involvement gained importance. Examples for these trends are the ULYSSES project on urban lifestyles and sustainability (De Marchi et al., 1998) or the VISIONS project on integrated scenarios (visions) for a sustainable Europe (van Asselt et al., 2005). More recently, IA projects are more and more embedded in societal processes, extending the boundaries of systems conceptualisations beyond the interacting three pillars of sustainability. Integrated Sustainability Assessment (ISA) is an example for this (Weaver & Rotmans (2006); Weaver et al. (2008)). ISA focuses on reframing existing paradigms of (unsustainable) development, rather than developing strategies to achieve sustainability within them. In other words, ISA aims to solve persistent problems by

redesigning the paradigmatic context causing their persistency. ISA does this in a cyclical, participatory process of scoping, envisioning, experimenting and learning. Today, Integrated Assessment is explicitly normative. Its results must aim to be relevant for decision-making.

Since the turn of the century, Sustainability Science has emerged as a related field. Kates et al. (2001) list four key methodological challenges for sustainability science: (i) spanning the range of spatial scales; (ii) accounting for temporal inertia and urgency; (iii) dealing with functional complexity and multiple stresses on human and environmental systems; and (iv) recognizing the wide range of outlooks. IA practitioners have criticised these ambitions as too deterministic, missing many of the challenges that are characteristic for wicked problems. Swart et al. (2002) proposed the following extension of the list: (v) linking themes and issues (e.g., poverty, ecosystem functions, and climate); (vi) understanding and reflecting deep uncertainty; (vii) accounting for human choice and behaviour; (viii) incorporating surprise, critical thresholds, and abrupt change; (ix) effectively combining qualitative and quantitative analysis and (x) linking with policy development and action through stakeholder participation.

Yet, the ambitions of both fields seem to overlap to a large extent, albeit with different emphasis. Kates et al. (2001) call for

fundamental advances in our ability to address such issues as the behaviour of complex self-organizing systems as well as the responses, some irreversible, of the nature-society system to multiple and interacting stresses.

Swart et al. (2002) call for

methods that can adequately and rigorously capture uncertainty, the capacity for system discontinuity, and the normative content of sustainability problems [...].

IA has grown out of the combination of the fields of risk assessment, technology assessment and policy analysis. The field's theoretical toolbox has the broad field of systems science at its core. To this core IA research attaches appropriate theories from a wide range of disciplines, such as the life sciences, geography, economics, sociology, the political sciences and psychology.

The methodological toolbox of IA initially consisted mostly of methods developed in other fields of research that are adapted for applications in IA.

Rotmans (1998b) proposed a typology of IA methods consisting of analytical and participatory methods. Analytical methods are generally adopted from the natural sciences and include model analysis, scenario analysis and risk analysis. These methods represent and structure scientific knowledge in an integrated manner. Participatory methods are much more diverse and include among others expert panels, Delphi methods, gaming, policy exercises and focus groups. These methods aim for the participation of non-scientists and stakeholders in the process. The theoretical and methodological toolboxes are not clearly defined and still developing and expanding. They are complementary, but not strictly modular with the consequence that only very few methods or theories can be combined with each other in a straightforward manner. Recent IA projects always combine different methods to approach an issue from different perspectives. These combinations frequently span across the participatory versus analytical divide.

The relative youth of IA as a discipline is reflected in the fact that many of the tools applied are still in a stage of experiment and development. This also holds for tools co-produced with decision makers and tools taking into account the specific characteristics of sustainability. Traditional techniques are geared towards the optimisation or simple issues. What is required, are methods and tools for efficiently finding good, but not necessarily optimal, strategies in a complex context: a comprehensive systems approach (see e.g. Clayton & Radcliffe (1996), Schellnhuber & Wenzel (1998), Sterman (2000), Allen et al. (2001), Grosskurth & Rotmans (2005), and Fiksel (2006)). This call for new systemic methods for the analysis of sustainability has led to the development of a large set of tools, methods and approaches for sustainability assessment to facilitate the sustainability transition (see e.g.: Kates et al. (2001), Robèrt et al. (2002), Scholz & Tietje (2002), and de Ridder (2006)).

There are many good reasons to apply IA at the interface of sustainability assessment and policy-making. The methodological issues raised by the concept of sustainability overlap with the ambitions of IA. The interdisciplinarity called for when analysing sustainability lies at the core of IA research and the improvement of the process of integrating knowledge from different fields is a continuous ambition of IA practitioners. The systemic character of sustainability and the complexities arising from this character are of a type that IA practitioners love to address. The IA attention for participation and learning, stakeholder relevance, uncertainty, and

perspectives perfectly overlap with the characteristics of sustainability. The development of Integrated Sustainability Assessment underlines this (Weaver & Rotmans (2006); Weaver et al. (2008)).

In the early years of IA, analytical methods dominated the field. According to Rotmans (1998b), these include models, scenarios and risk assessment. Early IA models were mostly system dynamic models following the tradition of the World3 model underlying the Limits to Growth (D. H. Meadows et al., 1972). Current IA models are methodologically more diversified. Recently, there is a tendency towards explorative modelling techniques, such as cellular automata and agent based modelling. Chapter 7 describes that evolution and some of its results. Scenarios have undergone an evolution similar to that of the models. Scenario studies used to be mostly sector-specific extrapolations of past and current developments distinguishing three alternative routes: high, low and medium. These became subsequently more integrated while also exploring more possible futures. A good example for that process of integration in the context of sustainability are two scenario studies conducted by the Stockholm Environment Institute and the Global Scenario Group, *Branch Points* (1997) and *Bending the Curve* (1998). Scenarios were never meant to be forecasting the future. Yet, their public communication and discussion frequently tended in that direction. Current IA scenario studies try to emphasize the explorative and reflective (learning) capacities of scenarios. Examples are the VISIONS scenarios (van Asselt et al., 2005) and the back casting scenarios employed in the COOL project (Tuinstra et al., 2002). A useful discussion of scenario analysis methods in sustainability science can be found in Swart et al. (2004). Risk assessment has also followed that broadening evolution. Early IA related risk assessments mostly focused on the assessment of the impact of technologies (O'Riordan, 1985). This has subsequently broadened considerably to include other, non-technological aspects of integrated decision-making (Kämper, 1998). Currently, IA studies commonly refer to risk and uncertainty, rather than risk alone, emphasising the unpredictability of the future, while at the same time accommodating that indeterminability in the scientific process (van Asselt, 2000).

Participatory methods now play a crucial role in IA as they are primary tools to address the inherent subjectivity and normativity of sustainability and to structure the communication processes. The input of non-scientific and practical knowledge and expertise, valuation and preferences through the involvement of actors by means of participatory methods enriches

sustainability exercises. Participatory methods can be applied to map out diversity or to reach consensus. When the goal of the research process is the description of differences in subjective and normative perspectives between different actors, it is most helpful to map out the diversity of their (world)views. Exploring the different choices and the implicit trade-offs makes the ambiguities visible. If one strives for concerted action of all stakeholders, it is most helpful to apply consensus-building methods. In an optimal case, both methods are combined: First, the diversity of opinions and possible strategies are mapped out. Differences in view are explained and discussed. Subsequently consensus-building methods are applied to arrive at an efficient and effective strategy for action.

With respect to policymaking and sustainability van Asselt (2000) distinguishes five categories of potential participants related to public policy issues, the so-called actors: government, citizens, interest groups, business, and scientific experts. It is neither feasible nor necessary to employ all actors for every sustainability study. The strategic selection and the recruitment of committed stakeholders is often one of the more challenging steps in the process. The use of participatory methods in policy-related research is now widely accepted (Kasemir et al., 2003). However, it seems that in some cases participation has become a goal in itself. Without a sharp vision on what the ultimate goal of a participatory process is, it is unlikely to contribute to any project. A frequently noted disadvantage of using participatory methods is the fact that current topics tend to be over-represented. An example would be a dominance of water-related issues after a flood has taken place. However, not only stakeholders, but also scientists and more analytical methods are prone to this tendency. The balancing of participatory input and material derived from desk-studies would level this problem somewhat. Nevertheless, it is evident that it is currently not possible to develop a model for sustainability, that is deemed to be evenly useful and applicable at all times.

For a detailed discussion of participatory methods in Integrated Assessment, see van Asselt et al. (2001), van Asselt & Rijkens (2002), and Kasemir et al. (2003). Combinations of analytical and participatory methods are now commonly experimented with and implemented in IA. For example, VISIONS (van Asselt et al., 2005), COOL (Tuinstra et al., 2002), International Futures (Hughes, 1999) and Quest (Tansey et al. (2002); Carmichael et al. (2004) each apply unique methodological combinations. This dissertation contributes to the further development of such combinations.

2.4 Research methods

The most important research methods applied in the context of this dissertation can be grouped under three headings: participatory methods, systems analysis, and case study research.

Participatory methods

In the previous section, we introduce participatory methods in the context of Integrated Assessment. In this dissertation, they were used throughout the research process. The major part of the research has been executed in close co-operation with professional consultants, staff members of the government of the Province of Limburg, fellow academics and other external experts. Participatory methods were used to facilitate this transdisciplinary interplay and to maximize learning from the case study.

The most important methods were different kinds of structured group dialogues and individual questionnaires (some of which again served as input for structured group dialogues). The dialogues mainly took place among a group of six to eight staff members of the province's government, a facilitating consultant, and the author. This set-up is atypical for participatory methods in Integrated Assessment as the group does not reflect the stakeholders of the process. However, as we mentioned above, the research methods and process in Mode 2 science is not pre-determined by the researcher, but rather the outcome of a process of negotiation. In this case explicit preliminary negotiations with the province revealed a trade-off between access to internal information, staff-time, and funding on the one hand and full stakeholder representation on the other.

The guiding questions and formats structuring the dialogues and the individual questionnaires will be explained in detail in the methodological chapters describing each stage of the project. It holds for all stages that the interaction between all participants was close enough to speak of a process of co-development. The choice of participatory methods is driven by two main arguments: we want to explore the systemic issues of sustainability in as much breadth as possible and build the capacity of the target organisation to skilfully continue doing so after the process. The first argument is epistemological. Much of what we want to know is best learned from those whose daily work it is to deal with the issues in question. The second argument concerns the legitimacy of results within the target organisation, as well as processes of learning by insight and learning by doing. A third

argument would be societal: actors (including among others users and stakeholders) are becoming increasingly aware and reflexive (Kemp & Martens, 2007). It is therefore necessary to embed iterative steps of reflection in the research.

Systems analysis methods

Results from the application of the participatory methods are an important source of information to feed into methods of systems analysis. The field of systems analysis is broad and vaguely defined. We focus on methods of representing a region as a system and methods of analysing such systems. The choice to employ methods of systems analysis has far reaching consequences for this dissertation as these methods produced much of the structure for the participatory processes and had decisive influence on our perspective on the region. Other possible approaches would have been histories, policy analysis, or institutional economics, to name a few. Each of these would have provided us with insights into a specific segment of the dynamics driving the ongoing development of the region. We opted for systems analysis methods as these seemed best suited to integrate the various types of data and information we expected to gain access to during the process, for example monitoring and modelling data, reports on sectoral and integrated planning, and individual (sometimes latent) expertise on specific sectors. Also the results we expected from the development and application of methods of systems analysis targeted those issues, where the province in negotiation preceding the project had foreseen the largest impact within the organisation and outside. In addition, the staff had previously experimented with systems methods, increasing the desire to deepen and experiment with their use. Different sustainability related methods of systems analysis are extensively evaluated in chapter 3. Section 5.1 provides an overview of qualitative methods for systems analysis.

Case study method

A case study with the Dutch Province of Limburg forms a crucial element of the research presented in this dissertation. The participatory and systems analytical methods discussed above are embedded within the case study. As is typical for context sensitive science, the research approach and methods are the result of a negotiation of the participating actors. The close co-operation of political administrators, consultants and academics, on which the research is based, reflects this. Case studies are a widely applied, yet contested, method in the social sciences. In short, a case study comprises the

in-depth analysis of a single unique example out of a set of many similar ones. Scholz & Tietje (2002) describe it in the following way:

A case could be a university department, a railway company, a city, or even a child. A case is considered from a specified perspective and with a special interest. It is unique among others (Stake, 1995, p. 2), and always related to something in general. Cases are empirical units, theoretical constructs (Ragin, 1992), and subject to evaluation, because scientific and practical interests are tied to them, They are used for purposes of demonstration and learning, both in education and research.

Critics of case studies as a research method in the social sciences point out several weaknesses. These are summarized by Yin (2003), whose 'Case Study Research: Design and Methods' has been a standard work on the subject since its first publication in 1984. Yin expands on the stereotypical perception of case studies having insufficient precision, lacking objectivity, being biased towards the confirmation of prejudices, suffering from the absence of scientific rigor, as well as being of little value in the development of theories. While Yin counters these arguments spread throughout his book, his main argument in favour of case studies as a research methods is their track record, which he shows to falsify the stereotype, whenever case studies are carefully executed.

A denser five point defence of the case study research method can be found in Flyvbjerg (2004) in which he argues that for the development of meaningful and context sensitive knowledge, the method of case study is indispensable. First, he asserts context-dependent knowledge to be at least as valuable as general theories in the social sciences. Since hard proof is a rarity, the possibility of learning and insight resulting from the in-depth study of a case is a promising path to knowledge. Second, a single case study or a powerful example can have an impact on the development of general theory. A single case, if carefully selected, can form a significant contribution. Consequently, and thirdly, the appropriateness of case studies is not limited to the generation and testing of hypotheses. Flyvbjerg further expands on this in his fourth point in which he states that contrary to widely held beliefs, *'the case study contains a greater bias towards falsification of preconceived notions than towards verification'*. Finally, he admits that the value of summarizing case studies in a meaningful way is difficult. However, he attributes this to the real-life object of study rather than case study research. In other words, summaries and generalisations compromise learning.

Yin, Flyvbjerg and examples from the practice of Mode2 science, Post Normal Science, Sustainability Science and Integrated Assessment leave little doubt, that a carefully executed case study can be useful. Therefore, in the context of this dissertation, the main point is not whether a single case is sufficient, but whether the case chosen and the way in which it has been implemented contributes to the wider aim of the research.

Other methods

Additional research methods are embedded in the research process in order to inform or to interpret the research and the case study. Most importantly, several literature reviews have been executed, covering among other existing and experimental tools for participatory systems analysis, as well as quantitative models for regional decision-making. A group interview is applied as part of the evaluation of the project.

Relevance

We formulate the main question of the research in chapter 1: 'How can tools for sustainability assessment stimulate joined-up thinking in political decision processes?' The four sub-questions focus on the qualities of existing tools, possible extensions, the combination of different tools, and the use of assessment tools in policy processes. The research methods reflect the research questions. Sustainability Assessment and joined-up thinking require participatory methods. The first because of the normativity and subjectivity of sustainability, the second because a process of group learning must take place in order to achieve joined-up thinking. During the research process, specific participatory methods were selected through from groups of methods considered useful for reflexivity and/or learning, depending on the stage of the process. The process of selection was based on insights from experts in participation, as well as on negotiations with participants in the process. A detailed argument for the choices made is given in the different methodological chapters (4-7). The complexity of sustainability and the interconnectedness inherent in joined-up thinking also require systems methods (though not necessarily the qualitative systems analysis approach, we opt for). Sustainability cannot be meaningfully analysed without taking into account the trade-offs and influences between its different aspects and the concept of joined-up thinking is born out of the awareness, that much efficiency and effectiveness can be gained when hitherto fragmented plans and strategies are integrated in a holistic framework. Participatory research at the science policy interface requires a committed group of participants. In

our case, this commitment was secured by embedding the research in a case study in which the participants had a strong self-interest to advance their understanding of the complexities related to sustainability and joined-up thinking. The selection of the specific case and the formulation of our research took place in an interdependent process. The context of the case study is described in further detail in section 2.6. Before that, section 2.5 describes the application of the research methods in the research process.

2.5 Research process

Research processes in Post Normal Science (PNS) projects substantially differ from research processes in normal science. Because PNS is normative, issue- and demand-driven, the exact focus of the research, as well as the methods and tools applied to get there, are emergent properties of the process and therefore not entirely predictable. The research process underlying this dissertation reflects this.

Typical for PNS, the research process underlying this dissertation begins with a rather broad normative research interest: How can a region become more sustainable in a systemic rather than symptomatic sense? This broad interest raises a large number of questions, for example: When is a region sustainable? How can we measure systemic sustainability? How can we account for uncertainty, surprises and discontinuities? What is the relevance of our research interest for other stakeholder, including political decision-makers? The research questions in chapter 1 are also part of this list.

Addressing all these questions simultaneously is beyond the scope of any individual research project. The criteria for narrowing the scope of the research are rooted in PNS. For example, a relevant research project has to be judged as interesting, and efficient by at least some stakeholders. Only then will they be willing and able to commit time and resources to the research process. This stakeholder interest is often connected to a case providing the framing and context in which the research question is to be answered. In addition, the capabilities and skills of the stakeholders and researchers taking part in the process or who could realistically be included in the process are a steering factor in the selection and implementation stages of the process.

The set of potential research questions combined with the opportunities to explore, with the help of committed stakeholders, different facets of our research interest. Due to the nature of our previous work, all of these opportunities were related to governance actors. Not all of these opportunities were taken up for this dissertation. As a result of the negotiations with the stakeholders on what we could expect and contribute, we rejected several research opportunities. Others were taken up by colleagues with interests and skills more appropriate to the specific case. Several regional cases turned out to fit our interest and capabilities. We describe the most important of these cases, a question for advice from the administration of the Dutch province of Limburg, in section 2.6. All other cases are introduced where appropriate.

The process of discussing and selecting the different cases in itself provided a rich ground for learning, since arguments often took the shape of insights into the daily work, perceptions and needs of potential project participants.

The adopted cases shared a number of stages reflected in the research process. During the early phase of the co-production process, the main question for researchers and policy makers was to find a common understanding on *how the different facets of sustainability are connected*. We proposed SCENE as a framework for a shared systems understanding of sustainability (chapter 4). The methodological choice for this qualitative and participatory method was based on the assumption that a common understanding could best be achieved through a collaborative and constructive process of learning. The SCENE projects provided a testing ground for all participating parties.

Once a common systems understanding was achieved, ways to manage and influence the system were explored. Doing so without losing the richness of the initial participatory process required a more structured approach than in the first stage. That process would still have to be open for participation, triggering the common learning process. The more widespread IA methods were either too quantitative or lacked the structure needed to further process the results from the first stage of the process. The balance was finally achieved with the development of the eQSA method (chapter 5). The basic elements of the method were imported from the field of futures studies, where structural analysis and 'la prospective' had been applied since the 1970's. But these elements were only suitable to answer the first question of how current policies affected sustainability in a 'what-if' manner. It was not

fit to identify policy levers with the required level of specificity. We therefore extended the method to suit our needs.

Once a tool is in place to identify policy levers, the question arises of *how to measure progress towards or away from sustainability?* In the relevant cases, traditional indicators would have provided a highly unsatisfactory answer to researchers and policy makers alike. These would not have captured the systemic character that lies at the core of our original research interest and the requirement for quantification would have implied a significant loss of information. The type of indicator proposed, the QSSI (chapter 6) directly resulted from the type and structure of the results of the first two phases and the methodological insights of the eQSA methods. The QSSI has never been fully implemented for a region, due to organisational constraints. Yet, it is an important element of the bundle of methods presented in this dissertation.

SCENE, eQSA, and QSSI all share the same methodological philosophy. They test the limits of systems understanding without quantitative measures. Avoiding quantification as far as possible meant that the breadth of the process was preserved. The restriction to qualitative methods in answering questions that were traditionally answered with the help of quantitative methods forced us to be creative with respect to the methods we used and developed.

Each case contributed in many different ways to the research process, including inspiration, framing of research questions, methodological developments, learning and testing. But cases are not laboratory experiments where circumstances and the environment can be controlled. Therefore separating between the different roles and assigning them to specific cases or even elements of cases or the associated processes is difficult and could only be done ex-post.

In PNS, the people involved are an important element of the research and all participating actors do shape the research and its outcomes. It is therefore crucial to be as conscious as possible about the different roles, participants, including researchers, play. Since these roles differ in different phases and in the different cases, we describe the roles played in the detailed descriptions of the processes throughout this dissertation.

To sum up, the research process began with a broad research interest in regional sustainability. A process of exploration and negotiation resulted in a number of potentially relevant cases, which in turn served as a framework

to select the specific research focus. During the participatory work on the cases, methodological needs were identified, in response to which we selected and developed the methods described in this dissertation. Other stakeholders in the co-production processes contributed, for example, missing elements of factual knowledge or insights into practices of decision-making. The cases and related experiments form the basis for a better understanding of integrated decision-making and for the evaluation of the newly developed methods.

2.6 Case study

One case study has been crucial to the developments described in this dissertation. In this section we describe the context of this main case study and its relevance for the research question.

ICIS, the institutional home of our research, was set up in Maastricht in 1998. During the founding period, formal and informal contacts were established between ICIS and the administration of the Province of Limburg, of which Maastricht is the capital. These contacts were characterized by a shared interest in integration. ICIS wanted to advance the development of Integrated Assessment (IA) and was actively looking for cases in which the young science could test its relevance for stakeholders and realise its trans-disciplinary ambitions. At the same time, the administration of the Province of Limburg wanted to develop a more holistic vision by integrating its fragmented forward looking activities into an integrated strategic plan, the so-called *Provinciaal Omgevingsplan Limburg* (POL). The POL project was to be pioneering work. In the Netherlands, legal mandates require forward-looking research studies at regular intervals for some sectors in order to support the development of robust policy strategies. With POL, Limburg was the first of the country's twelve provinces to combine its required sector plans and a number of optional plans into a single, integrated strategic plan. Activities related to POL and its results have been at the core of the relationship between ICIS and the administration ever since.

The informal contacts between ICIS and the Province are not documented. Formally, ICIS's involvement in POL related activities begins with ICIS preparing two working papers for the administration. Rotmans, Rijkens et al. (1998) provided an inventory of recently executed studies on the state and future development of environmental, socio-cultural, economic, institutional

and spatial factors in Limburg. Rotmans, van Asselt et al. (1998) proposed a framework to coherently explore and structure these diverse issues. The framework was based on the triangle of environmental, economic and socio-cultural aspects and their interactions. The framework and the participatory process applied to operationalise it were the first seed of what would later become the structure of the SCENE approach (chapter 4).

In order to gain experience in the application of the framework and the associated process, the province contracted ICIS to execute a series of training workshops and to implement the approach for selected issues and cases, as well as for the province as a whole. The latter was termed the *denkmodel* (think-model) for POL and was first documented in Rotmans et al. (1999b). In our description of the SCENE approach in chapter 4 we will return to this process and model in more detail. Van Asselt et al. (2000) describe the training workshops and the resulting triangular frameworks, including one for the Parkstad, a sub-region in the Eastern part of Limburg, one on energy and several ones on different aspects of water. It also includes a summary of the think-model. The work done until then was a major building block for the influential first fully documented version of the strategic POL (Provincie Limburg, 2001). Other building blocks were additional studies by external academic and commercial advisors and the results of an extensive process of consultation and political negotiation within the Province's administration and government.

Even though POL was a milestone in the evolution of strategic planning for the Province, the administration and ICIS both identified shortcomings that needed to be resolved in order to live up to the ambition of joined-up thinking. While every effort had been made to comprehensively cover a wide range of different environmental, economic and socio-cultural aspects and while in the graphical representation of the think-model and in the running text references to their interactions were plentiful and common, there was a clear bias to further elaborate interactions that would facilitate policy making and lead to win-win situations. Interactions causing dilemmas were relatively neglected.

More visibly, the numbers of the underlying forecasts and projections did not add up. This was most evident with respect to spatial projections. Many sectors, including among others housing, recreation and transport infrastructure, were planning to claim more space in the future without any reference to where that additional space should come from. But also in terms of demographic factors, inconsistencies were evident. For example, Limburg

was to become a 'knowledge region' while being home to an ageing and shrinking population with one of the lowest average education levels in the country.

Since many decisions on investments and regulations were based on the inconsistent forecasts and projections, the administration felt the need for a follow-up project aimed at the harmonisation of the prognoses. According to the tender for that project, the administration was looking for an improvement and extension of its tools for foresight. The project would have to improve the manner in which national forecasts were translated to the regional level and the way in which uncertainties were explored and represented in the process. The results were to facilitate exploring the causes and effects of surprises and qualitative changes that were outside the scope of existing models. The resulting tools would have to be compatible with existing models and monitoring activities. Process wise, the tender specifically asked for a process of co-development. ICIS won the tender for this project with a proposal for a qualitative systems analysis approach. The implementation of the project resulted in the development of the methods described in chapters 5 (eQSA) and 6 (QSSI). The implications of this process for the quantitative models concerned can be found in section 7.2 and the evaluation of our efforts is presented in chapter 8.

Chapter 3

SUSTAINABILITY ASSESSMENT

“Right idea,” said Mark with a sigh. “Wrong method.” (Douglas Adams & Mark Carwardine, 1990)

Stimulating joined-up thinking for sustainability requires sustainability assessment tools. Sustainability assessments explicitly aim for ‘*the integration of sustainability concepts into decision-making*’ (Pope, 2006). Variations of the concept are sustainability appraisal, sustainability impact assessment, integrated sustainability assessment, and others. They all employ a broad and heterogeneous range of methods that is the subject of this chapter.³

Recent years have seen an increasing number of inventories and reviews of tools for sustainability assessment. These generally include an evaluation of a tool’s appropriateness for a particular purpose, a phase of a project, or a typology of questions. This chapter adds to the ongoing discussion the dimension of systemic leverage (the scale of interventions required for transformative change in complex human-nature systems and the appropriateness of sustainability tools to explore and manage them). The results show a relative lack of tools for complex interventions and a shifting research focus to fill this gap.

In section one we provide an overview of tools for sustainability assessment. This is followed by an introduction to a hierarchy of leverage points to intervene in systems in section two. Section three links the set of tools to the hierarchy and evaluates which tools are most likely to provide useful support for complex system interventions. Section four explores, how existing tools can be improved to cover those leverage points that are up to now weakly supported by sustainability tools. In section five, we take a closer look at some limitations to the potential of improving the performance of sustainability assessments.

³ This chapter is based on Grosskurth et al. (in preparation). The author of this dissertation was responsible for the idea of introducing the Meadows typology, as well as the evaluation of tools based on the typology.

3.1 Framework for tool selection

Recent years have seen an increasing number of inventories and reviews of tools for sustainability assessment. These aim to provide some guidance through the increasing number of available tools, each of which has particular strengths and weaknesses and each of which highlights a different perspective of sustainability. Most overviews focus on particular aspects or tool groups. For example: Kasemir et al. (2003) survey the field of participatory methods; Hak et al. (2007) provide a state of the art overview of sustainability indicators; Robert et al. (2002) compare a series of tools for the sustainability assessment of production systems, including Life Cycle Assessment, Factor 4 and Factor 10, The Natural Step, and Natural Capitalism; and in the IQ tools project, six European research institutes co-ordinated by the Centre for European Economic Research (ZEW) have compiled an inventory of quantitative modelling methods for impact assessment (see <http://gloster.iwr.uni-heidelberg.de/>). An effort to provide an inventory of the latest generation of relevant tools and to advance their development and application is the EC-funded Methods and Tools for Integrated Sustainability Assessment (MATISSE) project. Not all results of the project scheduled to finish in April 2008 are available, but a preliminary methodological overview can be found in Weaver et al. (2008).

The most extensive inventory of assessment tools used to assess a policy's contribution to sustainability has been developed in the SustainabilityA-Test (SustA-Test) project (de Ridder, 2006) and is available at www.SustainabilityA-Test.net. 18 research institutes co-operated on the 3-year EU funded project to provide practitioners, policy-makers and interested laymen with an overview of state-of-the-art tools in sustainability assessment. Susta-Test grouped 50 tools in seven clusters: i) Assessment frameworks; ii) Participatory tools; iii) Scenario tools; iv) Multi-criteria analysis tools; v) Cost-benefit and cost-effectiveness analysis tools; vi) Accounting tools, physical analysis tools and indicator sets; and vii) Models. Each of these very generic clusters contains a broad range of methods. Most relevant in the context of this dissertation are dialogue related methods in the participatory methods cluster, the trends and cross-impact analysis methods in the scenario cluster and the method of qualitative systems analysis in the models cluster.

The focus of the overview lies explicitly on tools for policy support. As the Susta-Test is very recent and at the same time the most comprehensive

overview in terms of types of sustainability assessment tools, we explain it in more depth and use it as a point of reference for the remainder of this chapter.

A framework for tool selection has been developed to provide some guidance for choosing appropriate tools from the inventory. The framework relates the tasks and types of questions associated with generic four phases of a policy process to the tool groups that can support these tasks (de Ridder et al., 2007). The framework (see Table 3.1) will be briefly described in what follows.

The row labels list six of the seven tool groups. Assessment frameworks do not appear in the table. They can be considered ‘procedural tools’ that are designed to outline and guide an assessment process and the combination of tools. Within such a process, a broad range of tools can be applied (Finnveden et al., 2003). The assumption underlying the SustainabilityA-Test framework is that the role of individual tools will not fundamentally change between different types of assessment frameworks. The column labels indicate four generic phases of a policy assessment cycle. The phases are derived from an analysis of various types of assessments, which share the same origin in policy analysis (Dunn, 2003) and system analysis (Miser & Quade (1985); Quade (1983)). They were found to be the lowest common denominator of different types of assessments studied.

Each cell describes a task that is to be done in a particular phase and what a given tool can contribute to it. The shaded cells in the table represent task/tool combinations that are ‘in the lead’ in a particular phase. According to de Ridder (2006) the purpose of phase I, for example

is to understand the problem and to frame it, whilst accounting for different views on it. Useful tools in the problem analysis phase are tools that can steer the process of mobilising knowledge and articulating values, by means of stakeholder participation (experts, policy makers, laymen).

Tasks to be done include the development, mobilisation and discussion of knowledge, with the aim to reduce or understand uncertainty and to elicit views and if possible build consensus on the problem to be addressed. Participatory tools are recommended for this phase. Other tools could be added to structure (scenarios) or inform the debate (all others). In this way, the tool framework facilitates the choice of a tool by placing it in the context of the level of our understanding.

	<i>Phase I</i> <i>Problem analysis</i>	<i>Phase II</i> <i>Finding options</i>	<i>Phase III</i> <i>Analysis</i>	<i>Phase IV</i> <i>Follow-up</i>
<i>Participatory tools</i>	Problem framing (mobilising and integrating knowledge and values)	Supporting scenario building	Providing the context for and improve robustness of MCA, CBA and CEA	Evaluating the assessment process
<i>Scenario tools</i>	Providing the future perspectives to problem framing	Visioning futures, finding options and setting objectives	Providing references for the application of analytical tools	–
<i>Multi-criteria analysis tools (MCA)</i>	–	Definition of criteria	Comparing different alternatives	–
<i>Cost–benefit analysis (CBA) and cost-effectiveness analysis (CEA) tools</i>	Providing the analytical basis for problem-framing	Supporting objective setting	Full analytical characterisation of options to enable comparison	Ex-post assessment
<i>Accounting tools, physical analysis tools and indicator sets</i>				
<i>Model tools</i>				

Table 3.1: The role of tools in assessments according to de Ridder et al. (2007)

Different tools are not only applicable to different phases of an assessment cycle, but also to different types of problems. Remember the four problem types defined by Hisschemöller & Hoppe (1996): structured problems;

moderately structured problems; badly structured problems; and unstructured problems set out along the axes of consensus on values and the certainty on knowledge (see section 2.1). The cost-benefit approach, for example, is best used when all relevant impacts expected from a policy intervention are known and can be valued in monetary terms (certainty on knowledge and values). If there is uncertainty with respect to knowledge (e.g. unknown impacts or impacts that cannot be monetarised) and/or if there is no shared perception on values (e.g. what discount rate to use), a cost-benefit analysis becomes less useful. In this way the typology of problems is used to refine the four phases, the purpose of each phase and the tasks to be done. Analogously, tools vary in their suitability to address wicked or persistent problems. Some are better suited to support decisions about structural changes, others serve better in a context of incremental changes. We will return to this point in more detail in section 3.3.

In conclusion, it is of great importance to choose the most appropriate tools for any given question. This choice is difficult, as there is no direct relationship between tools and questions. In some cases, several tools can answer a given question, or one tool can address several questions. Unfortunately, many questions cannot be answered with current tools. Nevertheless, inventories of existing tools provide some guidance. The SustainabilityA-Test primarily uses policy phases and a problem typology to help selecting appropriate tools. In addition to these process criteria, practical criteria are often mentioned as important factors for choosing one tool over another. Examples are the availability of data, time, and resources. Less attention has been paid to the appropriateness of methods with respect to the scale of system changes one wants to explore or assess.

3.2 The scale and nature of system interventions

The scale of a policy intervention in the human-nature system is explicitly not covered by the tool evaluation in the previous section. The tool framework is suggested to be applicable to different types of assessments and within different assessment frameworks, from small-scale project appraisals to wide transition processes. The assessment cycle and the problem typology say little about how much will have to change in a system to solve a problem. Other tool inventories also do not cover this aspect. The question raised here is whether some tools are more adequate to support

incremental changes, while others are better to support transformative change in human-nature systems?

Answering this question requires a hierarchy of system interventions. A suitable hierarchy has been proposed by Donella Meadows. In 1997, Meadows published a paper titled 'Places to intervene in a system' (D. H. Meadows, 1997). It lists 10 levels of system intervention ranging from fiddling with parameters to the power to transcend paradigms. This hierarchy triggered a lot of discussion and in 1999 a revised version was published as a working paper: 'Leverage Points: Places to Intervene in a System', now distinguishing 12 leverage points in systems listed in table 3.2 on the right hand side (D. H. Meadows, 1999).

Meadows describes leverage points '*as places within a complex system [...] where a small shift in one thing can produce big changes in everything*'. This concept of a leverage point is very much related to systems and models. Policymakers, for example, are likely to perceive leverage points to be more related to the practice of decision-making and management rather than the epistemology of an abstract systems perception. The leverage points are listed from top to bottom in the increasing order of their effectiveness. Generally, a large change in parameters is needed to achieve even a small effect in terms of system behaviour. On the other hand, a relatively small shift in the goal(s) of a system is likely to have huge consequences. This implies that if we want to achieve an incremental change, lower level leverage points are often sufficient, while transformative change, such as a sustainability transition, would require higher level leverage points.

The hierarchy is based on decades of practical experience, as well as discussions with peers and practitioners. It describes the leverage points in a language that is adequate for qualitative and quantitative systems analysis. Its range transcends that of purely technical aspects of system interventions, making it especially useful for the classification of sustainability tools in terms of their systemic depth and broader context.

There is a lot to be said about the hierarchy. Meadows herself describes the list as *tentative* and the order as *slithery*. For the purpose of this paper, it is sufficient to acknowledge that Meadows got the order about right. Also, the wording of the list clearly reflects Meadows' background in system dynamic simulation. Less systemically grounded and mathematically rigid tools, such as participatory tools, are difficult to assess in these terms. When it comes to that, we will resort to an analogous interpretation of the list. How we draw

these analogies, we will explain when we link the relevant tool groups to the system intervention levels in section three.

<p>Static leverage points</p>	<p>12. Constants, parameters, numbers (such as subsidies, taxes, standards)</p> <p>11. The size of buffers and other stabilizing stocks, relative to their flows</p> <p>10. The structure of material stocks and flows (such as transport network, population age structures)</p>
<p>Dynamic leverage points</p>	<p>9. The length of delays, relative to the rate of system changes</p> <p>8. The strength of negative feedback loops, relative to the effect they are trying to correct against</p> <p>7. The gain around driving positive feedback loops</p> <p>6. The structure of information flows (who does and does not have access to what kinds of information)</p>
<p>Contextual leverage points</p>	<p>5. The rules of the system (such as incentives, punishment, constraints)</p> <p>4. The power to add, change, evolve, or self-organize system structure</p> <p>3. The goal of the system</p> <p>2. The mindset or paradigm that the system — its goals, structure, rules, delays, parameters — arises out of</p>
<p>Paradigmatic leverage points</p>	<p>1. The power to transcend paradigms</p>

Table 3.2: Leverage points to intervene in a system in increasing order of effectiveness according to Donella Meadows (1999) (right) and a clustering of these (left).

To not be overly distracted by minor differences between the different leverage points, we cluster the list into four groups. These groups are based on the methodological requirements it takes to learn more about each of them: (i) static, (ii) dynamic, (iii) contextual, and (iv) paradigmatic leverage points (table 2 on the left).

Static leverage points

These (12-10) are found among the elements of a system, the values of stocks and flows. These can be understood using a reductionist approach, as the knowledge about them is generally accumulative.

Dynamic leverage points

These (9-6) intervene in the interplay of relationships and forces between these elements. An analysis of these relationships requires an integrated approach, as the behaviour of the system at this level is often an emergent property of the underlying forces.

Contextual leverage points

These (5-2) must be searched for among the roots of the system. This concerns the exogenous drivers, the context prescribing the possibilities of system evolution, the framing of an issue, and shifts in the societal mindsets and perspectives. The context is generally normative. Scientific assessments can contribute to the understanding of contextual leverage points if they integrate other (non-scientific) perspectives in a dialectic process. Even then, contextual leverage points are not easily identified and managed. An example for such a shift in perspective would be the following: in our globalised world, a locally grown apple and one shipped around half the planet are not perceived as being substantially different from each other when sold in a shop. A more sustainable mindset or perspective would automatically recognize the fundamental difference between the two and thus change many of the lower level leverage points.

Paradigmatic leverage points

These (1) focus on the paradigms in the traditional Kuhnian sense, and shape all the other leverage points. Of all the leverage points, these are certainly the hardest to trace and influence. They are deeply rooted, often unquestioned, societal convictions. Meadows herself provides the example of the concept of land ownership. For many cultures in history

and some in the present, the notion of being able to own land was beyond conception. Vice versa it is virtually impossible for us to not perceive land as something that can be owned.

Meadows commented that 90-99% of our attention is focused on the lowest level leverage points. These require few changes elsewhere, few systemic considerations and certainly no reframing or paradigm shift. Those levers rarely change a system's behaviour. They merely change symptoms of it and are inadequate tools to drive changes in system goals, mindsets and paradigms. When people are calling for fundamental change, new structures or a paradigm shift, they are asking to work with higher-level leverage points. In the next section we link the tool groups distinguished in the SustA-Test project to the hierarchy of leverage points. We propose that most sustainability tools focus on the lowest level leverage points. At the same time we see, how progress is made to higher levels of leverage through the evolution of existing tools, development of new tools and new combinations of tools.

3.3 Tools and the scale of system interventions

Linking the overview of tools to the hierarchy of leverage points provides insight into the adequacy of the tools to inform and support decision-making concerning each intervention level. Does a tool merely describe the system surface (and the corresponding static leverage points) or does it delve into the deeper system layers of dynamic, contextual and paradigmatic leverage levels? A tool can address a system level in two ways. It can describe and document the system level (e.g. the individual system elements). In this case, any change is exogenous. The tool raises the question why the change occurs, but does not necessarily answer it. Or a tool can address change endogenously and thus explain the occurrence of change. This enables the user to conduct experiments at a given system level on different ways to intervene at a leverage point. Generally, the highest leverage level a tool addresses is treated exogenously. Lower ones can be either exogenous, endogenous, or not be addressed at all. Table 3.3 provides an overview of the relationship between the typical application of types of tools and the leverage levels.

LEVERAGE POINTS	static	dynamic	contextual	paradigmatic	Participatory tools (PT)	Scenario tools (ST)	Multi-criteria analysis (MCA)	Cost-benefit and cost- analysis (CBA)	Accounting and physical analysis tools (APA) and indicator sets (IS)	Models (MO)
	++	++	+		++	++	+	+	+	++
	++	+							(+)	+(++)
	+(++)	(++)								(++)

Table 3.3: Adequacy of sustainability assessment tools for the analysis of different levels of leverage (simplified)

- ‘++’ corresponds to a tool group that endogenously analyses a given level of leverage. The respective tools are routinely used to evaluate the state and/or development of the relevant types of leverage points, their changes and possible strategies for intervention.
- ‘+’ denotes that a tool exogenously documents a given system level. These tools often unearth information that is outside the scope of other tools, but relevant to the interpretation of the results of these other tools.
- (Brackets) represent cases, where experimental applications have been executed that are relevant to the given leverage levels. These are usually atypical ways to use a tool or unusual adaptations of a

tool. (++) denotes experimental uses treating a system level endogenously, (+) an exogenous treatment.

- A blank cell denotes the absence of such experiments. This does not exclude the possibility that such experiments are conceivable (or unknown to the authors).

The difference between endogenous and exogenous handling can be made relatively clearly for the predominately quantitative tools. Their application to different leverage levels can be deduced from their mathematical set up. For more qualitative tools this distinction is more difficult: it is problematic to distinguish between exogenous and endogenous approaches in a participatory process. Are the results of a group discussion (e.g. in a storyline workshop) an endogenous part of the approach, as the participants themselves are a part of an approach (much computer software is part of a modelling approach), or is the participatory input exogenous to the approach? We have chosen to denote those aspects, of which the causes of change are discussed or analysed in a participatory process as endogenous, and those where the method is more descriptive, as exogenous.

Tools for static leverage points

The more analytical tools are concentrated at the lower levels of intervention. The identification, measurement and representation of system elements is the single purpose of indicator sets and a prerequisite for the application of multi-criteria analysis, cost-benefit analysis, accounting and physical analysis tools, and models. But not all these types of tools provide the same level of insight into changes. In IS, MCA, CBA and APA, any change in the data is exogenous. Many concrete applications of these tools efficiently address the three domains of sustainability, and identify and raise awareness about symptoms of un-sustainability in each of these domains (Moldan & Dahl, 2007). This sometimes triggers action or further research. But the tools do generally not provide detailed insight into the interdependencies between the three pillars of sustainability (Dahl, 2007). Side effects are signalled if they fall within the scope of the analysis but remain unconnected to their cause.

The leverage point of information structure (6) can be triggered in some cases if the results reach those who take the relevant decisions. An example is the Ecological Footprint (Wackernagel & Rees, 1996). It has reached a broad audience and the tool has been applied in hundreds of case studies.

The results have changed the information structure by making tangible the elusive impact of humans on their planet.⁴ If the audience is not effectively reached, these methods are restricted to documenting constants and parameters. However, if the audience is reached, changes, such as environmentally conscious behaviour or appropriate laws and regulations may arise.

Many participatory methods, such as focus groups or Delphi surveys (Lindstone & Turoff, 1975) also focus on static leverage points (in some cases including normative or subjective perceptions of these). At the same time, there are a number of participatory tools, for example back casting (Tansey et al., 2002) or group model building (Vennix et al., 1996), exploring the dynamics of a system explaining the static leverage points endogenously. It is not surprising that participatory methods endogenously covering static leverage points are often related to models and scenario tools. MO analyse changes in the static leverage points in their context and endogenously represent the impact of one element on another (Gilbert & Troitzsch, 1999). Among others, the most widespread types of models, general equilibrium models and system dynamic simulation models, assume a set of rules or relationships explaining the development of the individual static system elements (Greeuw et al., 2000). The same holds for most quantitative and qualitative scenario tools. Quantitative scenario tools are a sub-group of models and share their properties. In semi-quantitative approaches, such as cross-impact analysis or formative scenario analysis these rules take the form of influence matrices (Scholz & Tietje, 2002).

Dynamic leverage points

At the dynamic level, scenario tools and models are the most widely used and often most appropriate tools, explicitly addressing the system structure. Models do so most rigorously, scenario tools in many cases most creatively. Most sustainability related modelling techniques in some way or another represent the underlying elements of the system and their mutual relationships. From this, future trajectories of static leverage points are calculated. The delays, negative and positive feedback cycles and

⁴ Whether actual changes in behaviour can be attributed to the Ecological Footprint is disputed. Rees is currently collecting cases, indicating that change has indeed occurred.

information flows are part of the underlying assumption and exogenous to most modelling techniques. Models, especially system dynamic models, explicitly address the structure of the system (11), delays (9), negative (8) and positive (7) feedback loops (see e.g. Bertalanffy (1968), D. H. Meadows et al. (1972), and Bossel (2004)). These models hit a ceiling in current applications when it comes to dynamically evolving system structures (4). Due to their heritage in engineering and the natural sciences, current system dynamic methods are not capable to endogenously model changes in the very relationships that form the underlying assumptions of the model. Such evolution of a system structure is quite common in human systems.

Many sophisticated scenario tools use models as an integral element and benefit from their advantages (e.g. Hughes (1999), and Tansey et al. (2002)). But even qualitative scenario tools can pay explicit attention to systemic interactions, feedbacks and delays (e.g. van Asselt et al. (2005)). In these methods, the use of models and quantitative tools is generally restricted to informing the process. This is the price to pay for a broader coverage, richer images and the benefit of being able to let system structures evolve (4) and discussing the goal of a system (3) or even the mindset of systems (2) (Grosskurth, 2007). Semi-quantitative scenario tools represent the structure of delays, feedbacks and information in impact matrices. These matrices and the conclusions drawn from their structural analysis trigger some reflection on the shape of the system, but, as in models, changes in the structure are exogenous. Qualitative and explorative scenario and participatory methods are often more flexible. They explore many different possible system structures, albeit often without a detailed analysis of the impacts on the static level. Rich narrative scenarios, story lines and creative participatory processes trigger this reflection – in some cases even on the drivers of system change.

Contextual leverage points

Towards the higher levels of intervention, the tools seem to peter out. However, experiments within the participatory, scenario and model tool groups do reach into the higher leverage ranks. Experimental modelling techniques address contextual leverage points. Especially agent-based models endogenously represent the evolution and self-organisation of system structures. These are now being tested in the context of sustainability (see e.g. van der Veen & Rotmans (2001), Moss & Pahl-Wostl (2001), Rotmans & Rothman (2003), Valkering et al. (2004)). Given the goal of a system, these models are able to deduce better alternatives for the set up of

dynamic leverage points. Some participatory methods specifically designed for a context of normativity, uncertainty and trans-disciplinarity also provide relevant insights into the highest leverage levels. Transition management and interactive back casting deserve a mention here and we will cover them in more detail in the next section. Other tool groups generally only reach up to this level when they are combined with participatory approaches. Some applications of participatory methods are difficult to place without the context of their specific application. Participatory methods can range from a pseudo-deliberative legitimisation of pre-defined policy options concerning the lowest leverage points to paradigm shifting stakeholder discourses (Arnstein, 1969).

Paradigmatic leverage points

Acknowledging the existence of paradigmatic leverage points is a necessity. The mindset, perspective and paradigms that shape the world around us have a tremendous influence. Changing them requires the boldness to address them combined with the humbleness of being aware that it is by definition impossible to escape one's own paradigm and that changes in mindsets are notoriously difficult to detect and to manage. In Meadows' example of land ownership, we might be able to analyse the possible consequences of not allowing land ownership anymore, we are not able to return to a state, where landownership is inconceivable. These changes go beyond a change in perspective or mindset.

Assessment frameworks

The previous evaluation clearly shows complementarities in the tools used. At the same time, there is enough overlap to look at some leverage points from different perspectives, especially at the lower intervention levels. While one tool might be best suited to provide information on the state of a leverage point (and thus potentially identify the need for using it), another tool might be more suited to explore the consequences of using it. At the higher leverage levels, tools become scarcer. These higher levels might be achieved by applying a toolbox of methods (Rotmans, 1998b) and by combining different tools, even if they are not linked to each other (de Ridder, 2006).

Assessment frameworks provide guidance to the combination and integrated application of different tools. Applications of assessment frameworks range from highly specific inventories of the lowest leverage

points to broad integrated assessments that often combine a wide range of tools, especially model and scenario tools. At the lower end, concepts like the EU Environmental Impact Assessment (EIA) can be found. The EIA process analyses the environmental effects, documents the effects, undertakes a public consultation exercise, reflects on the comments, and takes a final decision about which the public is being informed (see <http://ec.europa.eu/environment/eia/home.htm>). It combines typical low leverage methods, such as cost-benefit-analysis and indicators with participation. This participation is limited to reflection on a report drafted without participatory involvement.

At the high end of the leverage scale, more experimental concepts are found, such as Transition Management (Rotmans et al. (2001a), Loorbach & Rotmans (2006), Loorbach (2007)). It illustrates recent efforts to move up the hierarchy, to meet the need for assessment tools addressing persistent issues. Transitions are long-term fundamental social transformation processes in which (a complex subsystem of) society changes structurally. Influencing the occurrence, speed and direction of transitions is referred to as transition management. According to the SustA-Test web book (2006):

Transition management is a multi-actor governance approach aiming at long-term sustainability through a process of learning by doing and doing by learning. The operational model consists of the creation of a joined problem perception and long-term visions in a so-called transition arena, a shared transition agenda, innovation networks and experimental playgrounds. The basic underlying rationale is that transitions imply structural and institutional changes that influence the market-, innovation- and political- administrative systems, which can and need to be governed into a sustainable direction.

The explicit focus of transition management is on the higher leverage points, especially on the evolution and self-organisation of systems (4). The development of long-term visions supports changes in the goals of a system (3). Several studies document past and current transitions (e.g. on energy (Rotmans, Kemp, van Asselt et al., 2001), waste (Loorbach, 2005), and water (van der Brugge et al., 2005)). Some case studies on managing transitions have been performed or are under way (e.g. Watson et al. (2004), and Loorbach (2007)). Nevertheless, it is too early to judge the effects of transition management cases.

The urge to move up the hierarchy can also be seen in individual tool groups, especially scenarios and models. Chapter 6 of this dissertation also provides an example of applying indicators to higher leverage levels.

3.4 Tools for fundamental system changes

Let's take a closer look at some experiments on how current methods have been applied in specific cases to become relevant for higher-level leverage points. We focus on those tool groups with relatively high scores: assessment frameworks, participatory tools, scenario tools and models. These examples illustrate that innovative adjustments to the tools or a creative synthesis of them can help to make tools more relevant at higher levels leverage points.

Of the participatory tools, we already mentioned the interactive back casting approach as one addressing high level leverage points. A well documented example at the interface of participatory and scenario tools is the Georgia Basin Futures project (Tansey et al., 2002). The Georgia Basin back casting process benefits from an integration of participatory methods and a model simulation (GB-Quest). The purpose of Quest lies more in reflection on and education about the complexity of a sustainable future than in the development of scenarios in their own rights. Through a very user-friendly interface, the user, who is usually a local stakeholder, explores different strategies for regional development. Users define their own image of sustainability and a desirable future, based on choice offered by the interface. Embedded in a participatory process of narrative vision and scenario development, this approach to back casting is an example of a successful combination of participatory and simulation tools for scenario development (Robinson & Tansey, 2006).

Another shift towards higher leverage levels can be observed in the use of scenarios. Futures studies have evolved from single-issue predictions, the exploration of alternative futures (Kahn & Wiener, 1967), the exploration of discontinuities (van Notten, 2005), to institutional capacity building and a learning society (Inayatullah, 2002). The related scenarios are seeking to portray system changes at higher and higher levels moving towards the ideal of macro thinking and explaining the big picture. This macro perspective would stand above the discontinuities and noise of current scenarios and explore the patterns shaping the past and the future

(Sondejker et al., 2006). The four phases of a transition (pre-development; take-off; acceleration; stabilization) can be understood as one element of such a pattern (Loorbach, 2007).

In modelling, one example of stretching to a higher leverage point is the TARGETS model (Rotmans & de Vries, 1997). TARGETS stands for Tool to Assess Regional and Global Environmental and Health Targets for Sustainability. TARGETS puts its emphasis on the assessment of long-run environmental developments. The model addresses uncertainty and differences in mindsets. It uses a generic description of the users worldview and management style and combines these into so-called perspectives based on cultural theory. Changes in perspective lead to adjustments of the underlying assumptions (relevant parameters, delays and to some extent the structure of the model itself). The choices made influence the way in which adjustment in behaviour, learning and other aspects takes its effect during the run of the model. Users can test their own perspective, take on the roles of other stakeholders and discover their perspective and thus discover the consequences of different world views and management styles.

The methods described in the subsequent chapters of this dissertation also focus on higher-level leverage points. The SCENE approach provides a framework for a better understanding of dynamic leverage points and triggers discussions among the process participants about contextual leverage points (chapter 4). The eQSA approach is a heuristic for efficient interventions in complex systems at the dynamic and contextual level (chapter 5). An indicator based on systemic insight is the QSSI (Qualitative Systems Sustainability Indicator). It is based on the direction and strength of the causal relationships between elements of sustainability instead of measurements of elements. Using basic methods of qualitative systems analysis, targets are tested for their potential of simultaneous achievement and aggregated into a single value. The value of that indicator can actually change depending on ones perspective based on a participatory process about the goals, rules and mindsets (chapter 6). In the next section we will take a look at the more general characteristics of tools fit for higher-level leverage.

The characteristics of the tools change as we move up the levels of system intervention. At higher leverage levels, we encounter what one might call the comprehensiveness trap. Comprehensiveness refers to the degree to which the different elements or characteristics of a system can be represented and/or analysed using a given method. The need for

comprehensiveness increases at higher system levels, as seemingly irrelevant side-aspects often become determining factors of a future. However, with the increasing complexity encountered at higher-level leverage points, comprehensiveness becomes more difficult to achieve, if not an impossibility. It has been noted – often with reference to Godet (1983), that the restriction to quantitative measures severely limits the comprehensiveness of a given tool. Chapter 7 documents this for a specific case concerning models for the support of policy strategies for regional sustainability. Many determining factors are soft factors or emergent properties of several factors. The focus in quantification is a classical case of looking where the light is. The reach of the light can be extended. The Sustainability Geoscope (<http://www.sustainability-geoscope.net/>) project aims to increase the reach of the light with respect to sustainability indicators. It is planned to be an observation instrument for the anthropocene. As such, the Geoscope should collect an unprecedented range of data related to sustainability and human-nature systems. Another approach to make the light shine further is to quantify previously non-quantifiable variables to enlarge the comprehensiveness of tools. Examples are attempts to monetarise eco-system services (Costanza et al., 1997) or a national index for happiness (Diener, 2000).

Another way to increase comprehensiveness would be to look at deeper system characteristics. Instead of adding more individual elements to a system, integrate them in one approach. Instead of looking at the system structure and its dynamics, look at the powers that drive its structural evolution. Systemic depth refers to improvements in the tools with respect to analysing fundamental system properties. This can range from ensuring that all aspects of a given system are covered in a participatory process to improved statistical methods for diluting information from incomplete data sets. Qualitative methods generally lack a certain rigor to test comprehensiveness and consistency in assumptions.

Higher leverage levels also pose higher demands on deliberative qualities. Normative elements and uncertainties must be integrated in the process. The methods must therefore be better and better suited for participatory applications. Deliberative quality refers to the ability of methods to cope with uncertainty, ambiguity, normativity and dissent in a context of high stakeholder involvement. It is on this aspect that progress is possibly most difficult to achieve. A highly profiled and well-documented project to gain insights into deliberative quality has been the COOL project (Tuinstra et al.,

2002). The project contained dialogues at three scale levels: Global, European and National Dialogue. The assessments explored the possibilities for 80% reduction of GHG emissions by 2050. The National and European Dialogue used interactive back casting as their main method. While COOL was explicitly designed to address dissent, the process turned out not to be apt to fulfill that ambition. One could conclude from this and other examples that sustainability assessment tools are not yet up to the challenges posed by truly deliberative processes beyond informing such processes.

With increasing leverage, the process becomes more complex. Levels of incremental change at the bottom end of the hierarchy can often be adequately addressed using relatively quick and technical mono-disciplinary methods with a narrow focus. Higher levels require longer lasting processes with a broad focus involving deliberative multi-disciplinary methods and a high degree of participation. As a consequence, applications at the static level are much more frequent and widespread than on the dynamic level or even higher. With increasing scale levels, these methods become much more demanding in terms of time, resources, process management and stakeholder engagement.

3.5 Limits to shifting up the hierarchy

When moving up the hierarchy, the role of science changes. The most fundamental point from the perspective of the scientist is that a shift towards the peak of the hierarchy of leverage points might imply that we leave the realm of science. The answer depends on your perception of science. A shift up the hierarchy certainly goes together with a shift from Mode1 science to Mode2 science (Gibbons et al., 1994), from traditional science to Post-normal science (Ravetz, 1999b). This new science is issue-driven, problem-focused and trans-disciplinary. The context is generally one of uncertainty, disputed values, high stakes and urgent decisions (see section 2.3). Research undertaken in the fields of Integrated Assessment and Sustainability Science implements these principles with respect to sustainability (Rotmans (2005); Martens (2005))

Placing such endeavours outside the realm of science carries a danger. There exists an interaction between the different levels of leverage points. Out of the system paradigm arise the goals, structure and parameters. In turn, the consequences of changes in the lower levels might challenge the current

paradigm. Scientists who restrict the definition to the purely objective and norm-free values of traditional science are hiding their head in the sand. Even the most traditional science is fully embedded in the normative and often ideological cycle of knowledge generation. Not to acknowledge would mean to ignore a major variable in the model of social evolution. It is therefore not surprising that society increasingly expects from academia to participate in and to organize trans-disciplinary processes and that researchers more and more respond to that demand (Rotmans, 2005).

One cannot develop an objective and falsifiable system goal or paradigm and then inform the larger public about it. For the knowledge creation required for the use of these leverage points, the dialogue of people is inevitable: citizens, policy-makers, scientists, and so on. For this dialogue to be successful, at least a part of the scientific community must open up to a fundamental interaction with other societal players, learn carefully from them and contribute scientific understanding and their interpretation of the meaning of it for the requirements concerning system goals and paradigm to it. Integrated Assessment, Mode2 science and Post-normal science and Sustainability Science aim to be part of that open part of the scientific community. The tools for that kind of science are in a very early stage of development. But there are promising pathways, there is commitment and enthusiasm, and there definitely is a need.

Equally fundamental is the gap between the supply and demand of information at the science policy interface. Institutional, conceptual and paradigmatic limitations and differences are highly persistent. The common ground required for successful processes of co-development and social learning is still relatively small. The different cultures with respect to agenda setting, discourse and decision-making are one among many barriers to effective collaboration. The perception of concepts, such as sustainability, transitions or governance differs significantly between policy-makers academia and other societal actors. For example, Turnpenny's exploration of Integrated Sustainability Assessment in the European Union shows, how sustainability assessments were reframed to fit the dominant aims of economic growth and de-regulation, leading to a neglect of sustainability principles despite increased interest (Turnpenny, 2008).

In addition, a large number of concepts are not part of even tentative common ground and therefore generally lost in trans-disciplinary processes. Paradigms differ widely within and between societal actors. Major differences run, for example, along the lines of the degree to which society is

malleable, the predictability of consequences, or whether top-down or bottom-up approaches are more effective.

There are also several technical and practical limits to our ability to improve current tools in terms of their relevance to higher levels of system intervention. In this section, we address some of them.

The gap between the ambition and political realities

There is a gap between the theory of integrated assessments and the reality of policymaking (Lee, 2006). Trans-disciplinary, deliberative integrated assessment frameworks and tools that aim for higher levels of intervention, often promoted by the scientific community to address complex problems related to sustainability (e.g. Weaver & Rotmans (2006)), are rarely used in practice. One possible explanation is that the reality of policy making constrains the possibilities of integrated assessments, in particular as there is often no room for open processes that might undermine the goals set and options proposed by politicians. In addition to that, time and budget constraints could limit such open processes.

The energy cost of problem solving

Tainter (1996) proposed that problem solving requires resources and time. The more complex a problem is, the more resources are required to solve it. The marginal returns to these resources are decreasing. That implies a point beyond which the extra resources needed to solve a problem are larger than the cost of the problem itself. In this sense, processes to solve sustainability related issues do only make sense if the resources required to execute them are smaller than the benefits derived from the expected results. The urgency of some of these issues means that the time it takes to find solutions through complex research processes is possibly too long to adequately address them.

Increasing complexity and urgency

Increasingly complex issues with high stakes that have to be addressed with respect to sustainability only exacerbate this problem as more stakeholders need to be involved and more complex systems understood (Rotmans, 2005). Significant progress in research heuristics would help to relieve some of this pressure. This implies a shift from the time-consuming quest for optimal solutions to faster processes for finding better solutions.

The limited life-span of information and knowledge

In the social sciences, indicators have a limited life span. Bossel (Bossel, 1999) illustrates this effectively when he observes:

As systems change and develop in a changing environment, individual indicators may lose their relevance and may have to be replaced by others that are more relevant under current conditions. Where once coal consumption per capita may have been a useful indicator of living standard, the number of computer chips in use per person may be a better indicator at another time.

Time-series analysis obviously becomes problematic under such circumstances. Also, system structures change and evolve. Estimates for relationships between different parameters change over time and their roles as analysed by tools should change accordingly. Only tools that endogenously represent such shifts, only indicators that represent fundamental systemic properties could circumvent this problem.

3.6 Conclusion

We have seen a concentration of tools focused on the least effective leverage points when it comes to interventions in complex societal systems. Tools and processes addressing higher levels are scarce and in many cases experimental. The SustA-Test team proposes as its main conclusion, that combinations of tools are a promising trajectory. Unfortunately, this only holds for the system intervention levels to which at least one of the tools used apply. This limits the relevance of these tools for planners, politicians and other decision makers as far as higher order system interventions are concerned. In addition to combining tools, using existing tools more effectively is a promising way forward. Gains in this respect are most likely where tools are applied with a greater awareness for the (need of) common ground between the participating actors. Adjusting the type of questions addressed with the tools to the specific context of a science policy-interface is one way to achieve this.

Nevertheless, the urgent need for new and innovative tools for relevant sustainability assessment is evident, as is the need for applying them in the policy process. The path to meeting that need will first and foremost require more sophisticated participatory methods and a better understanding of the effects of participatory processes. In the second place, better methods will

require more sophisticated tools from the other tool groups that can be used in a participatory manner. As methods used in sustainability assessments often originate from other disciplines, this is partly a question of a faster transfer of adequate methods and partly one of developing more tools specifically tailored to the type of issues at hand. And like the science positions itself at the interface of science, politics and society, the upcoming methods will be shaped by that interface. Rotmans (2006) summarizes these challenges in a call for Triple-I tools: Integrated, Interactive and Innovative. The very first seeds for adequate tools and science have been planted. Back casting, transition management, and agent-based models are but a few symptoms of this development. Nevertheless, there is a long way of growing and maturing ahead. The next four chapters are steps in that direction.

Chapter 4

THE SCENE APPROACH TO MAPPING SYSTEMS

Methods become mentalities. (Karl Stern (1986))

Sustainability is widely accepted as a political leitmotiv. But, as explained in section 2.2, scientifically, its complexity, ambiguity, normativity, and subjectivity make it a 'wicked problem' to address. While there is an abundance of tools focusing on the monitoring and extrapolation of quantifiable elements of sustainability, there is a lack of tools with the capacity to stimulate joined-up thinking for the sustainability transition. Such tools would include qualitative elements in a structured manner, take a comprehensive systems view of sustainability, and enable the meaningful engagement of stakeholders (see chapter 3).

In this chapter, we present the SCENE approach as the basis for a set of tools contributing to closing that gap in available methods of sustainability assessment.⁵ The acronym SCENE stands for the three domains of sustainability: SoCial, ENvironmental, and Economic. We developed the approach to foster a better understanding of sustainability's underlying dynamics. SCENE is based on the participative and qualitative representation of stocks and flows in the format of an influence diagram. In the context of our aim to achieve four-fold integration at the administration of the Dutch province of Limburg, the SCENE approach was the seed that planted awareness for interconnectedness within the organisation, while at the same time creating tangibility and structure.

Section one describes the structure of SCENE. Section two details the process of the SCENE approach. In section three, we provide a general overview of applications. Chapters 5, 6, and 7 describe a set of tools for which the SCENE approach serves as a basis or provides valuable input.

⁵ This chapter is based on Grosskurth & Rotmans (2005). The author of this dissertation had a key role in the refinement of the SCENE approach, as well as the design and implementation of the case study process.

4.1 The SCENE structure

4.1.1 CAPITAL DOMAINS

SCENE is based on the widespread differentiation of three capital domains of sustainability: SoCial, ENvironmental and Economic. Together, these domains form the so-called sustainability triangle. The triangle and its variations have served as a framework for the communication and analysis of sustainability in politics and science since the 1990's. In several studies, lists of indicators were structured along the lines of forms of capital. For example, UN-DPCSD (1999) distinguish social, environmental, economic and institutional capital. The World Bank applies the slightly different structure of social, human, man-made and natural capital (Munasinghe (1993), World Bank (1997)). Spangenberg & Bonnoit (1998) extended this approach with policy targets for key indicators and interaction between the different domains in order to make 'proactive policy steering' possible.

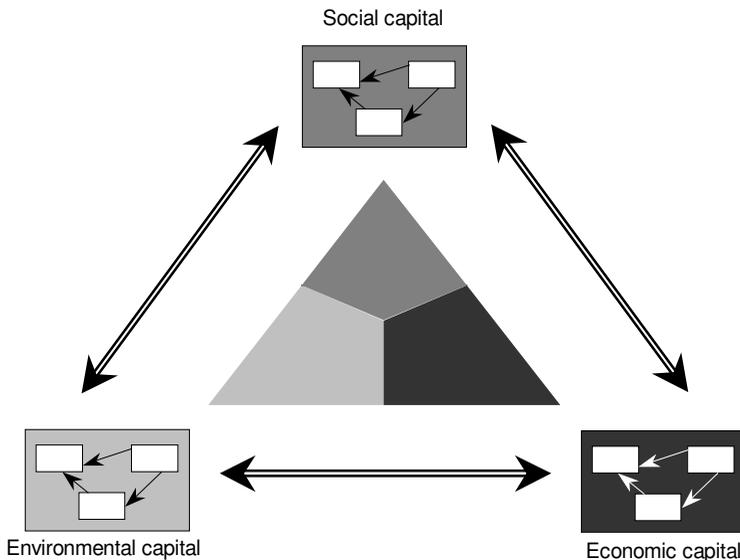


Figure 4.1: The SCENE model (Rotmans, 1998)

With SCENE, we extended the concept of interaction between the capital-domains and the way in which capital stocks are described helping to

portray and analyse context and dynamics more accurately. With these additions, the triangle is transformed from a concept for the structuring of sets of indicators for sustainability to a model, that allows for the analysis of underlying dynamics and integrated strategy analysis and decision support. Figure 4.1 is a schematic representation of the 'naked' SCENE model.

4.1.2 STOCKS

Rotmans et al. (1998) first developed the structure of the SCENE model as an experimental support tool for an analysis of regional sustainability in the Dutch Province of Limburg. They emphasize the importance of stocks and the interactions between them in long-term sustainability studies. Each capital domain contains a number of stocks. These can be quite generic terms, such as 'quality of life' (social capital), 'environmental quality' (ecological capital), or 'economic vitality' (economic capital). The genericity of the stocks decreases the tendency to favour stocks in the selection, for which quantitative data are readily available. The main criterion for the inclusion of a specific stock must be its relevance for the issue under research. From the legacy of system dynamic modelling and indicator studies, there is a tendency to describe a stock using a single dimension, i.e. quantity. This approach does not do justice to the frequently apparent interaction of different aspects of a single stock. We therefore represent four aspects of a stock, namely its quantity, its quality, its function and its spatial component.

Taking 'education' as an example of a stock in a research study on a national level, its quantitative aspects could be described in terms of the distribution of education over the population and the number of years in education, the quality could contain the efficiency of the educational system, the function of education could contain how well education prepares for earning a living and taking responsibility, but also the potential for innovation, the spatial aspect could contain the geographic distribution of education and the land surface needed for educational purposes.

A stock can be described using multiple descriptions of the same characteristic. In other words, a stock can be described with an arbitrary number of quantitative, qualitative, functional and spatial terms. These can be compiled in a fact sheet as well as a narrative abstract. In this way it is possible to capture the wide range of effects that a stock can have in a system. The clustering of these different characteristics of a stock enhances the intuitive understanding of the interaction between the characteristics.

The detailed description of the stocks provides the groundwork to conceptually test policy strategies for sustainability in an integrated way.

Describing the characteristics of a stock is not the same as finding measures or indicators for a stock. It is only in a subsequent step that our focus turns on the selection of indicators for each of the characteristics.

Indicators describe complex phenomena in a (quasi-) quantitative way by simplifying them in such a way that communication is possible with specific user groups. The term 'quasi' indicates that, although indicators are mostly quantitative in nature, in principle they can also be qualitative. (Rotmans, 1997)

Qualitative indicators may be preferable to quantitative indicators where the underlying quantitative information is not available, or the subject of interest is not inherently quantifiable (see e.g. Gallopin (1996), and Bossel (1999)).

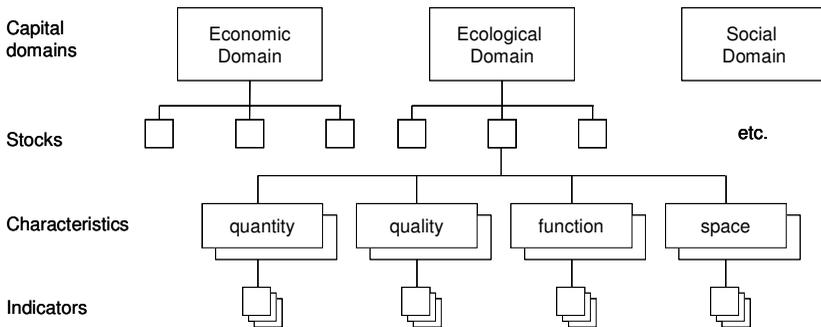


Figure 4.2: The layers of the SCENE model

Each characteristic is assigned one or more indicators or indices from available information-resources. We are aware of the fact that the selection of a measurable indicator for a characteristic with the goal of quantifying or monitoring the development of a certain characteristic often goes together with a loss in descriptive quality. It is for this reason, that we put so much emphasis on the selection of generic terms to describe a system. In this way, the consciousness of what an indicator is meant to measure (and any

deviation from that) is best preserved. However, selecting indicators is not always necessary, depending on the type of research question.

The selection of indicators completes the variable tree implicit in the SCENE structure. The stem of the tree are the three capital domains, the branches are the stocks within each capital domain, twigs branch off in the form of characteristics, connecting to the leaves, the indicators (see Figure 4.2.).

For an integrated description of an issue it is crucial that all three capital domains are 'filled' with stocks with the same scrutiny. It is of little use to take the three-capital model as a basis for analysis and subsequently neglect one or even two of these capital domains in the analysis.

4.1.3 FLOWS

To complete the structure of the SCENE model it is necessary to define the relationships between the different stocks. We call these relationships 'flows'. Flows are visualised in the triangular SCENE model as arrows. We distinguish 'intra-flows' and 'inter-flows' (Rotmans, van Asselt et al., 1998). An intra-flow connects two stocks of the same capital-domain. Analogously, an inter-flow connects two stocks of different capital-domains. Only inter-flows can be carriers of substitution processes, where one form of capital is substituted for another. Depending on the research question, the focus on placing the links can be on links between the generic stocks, between their characteristics or between the selected indicators.

4.2 The SCENE process

In the previous section, we presented the structure of the SCENE model with its four layers ranging from the three capital-domains to the selected indicators. In order to be of any use in a practical application, this 'naked' structure needs to be filled in with information. In general, there are several ways in which this can be done. One can use the insight of experts in a given field, one can rely on participatory methods, or combine both approaches. Which method to use depends on the research question at hand and on the available resources.

4.2.1 EXPERT INSIGHTS

The least integrative form of completing the SCENE model using expert insight is a desk-study, where experts select the relevant stocks, characteristics and flows based on her knowledge of the research question at hand. The SCENE model in this case functions as a structuring framework for complex research issues. For contested research questions, a researcher can develop different plausible sets of stocks, characteristics and flows. In several case studies, we use 'perspectives' to arrive at a small set of consistent, but inherently different SCENE models. These different models can be used to describe, and in a later stage, to bridge different views on an issue. The concept of perspectives is derived from Cultural Theory (Thompson et al. (1990); Thompson (1997)).

Perspectives may be considered as aggregations of the different points of view humans have, and can be defined as consistent, hybrid descriptions of how the world functions and how decision-makers should act. (Rotmans & de Vries, 1997)

For a more extensive description of the use of perspectives in the field of Integrated Assessment, see van Asselt & Rotmans (1997) and van Asselt (2000).

Usually, the knowledge of experts alone is too limited and/or too entrenched to arrive at a consistent and generally accepted set of stocks, characteristics and flows. Therefore, it is advisable to always allow for feedback from other stakeholders. For contested issues, a development process involving other experts and stakeholders is advisable.

4.2.2 PARTICIPATORY METHODS

Participatory methods play a crucial role in the application of the conceptual SCENE model as they help to address the inherent subjectivity and normativity of the concept and at the same time structure the communication process between the modellers, other scientists, stakeholders, and users. The input of non-scientific and practical knowledge and expertise, valuation and preferences through the involvement of actors by means of participatory methods enriches assessment exercises (as outlined in section 2.3). Also, the ambiguity of the concept of sustainability is inherent and cannot be reduced. However, by mapping out the differences between perspectives and by analysing trade-offs between different

strategies and their effects, it is possible to make the inherent ambiguities visible and thus offering an opportunity for a transparent choice.

The participatory methods we use to generate input depend on the characteristics of the problem for which the model is developed and on the type of participants we can expect. In general, different forms of focus groups are applied together with interviews and open feedback sessions. This iterative process of individual work and group work has worked well in several cases. Comparing individually prepared work in a group setting highlights differences in world-views and serves as a structure for a group discussion on agreements and disagreements. At the same time, individual gaps of knowledge and information are identified and often filled during the group sessions.

Developing a policy relevant model for the sustainability of a given region, such as the province of Limburg, is hardly possible without a structural interaction with stakeholders. They map out the relevant factors and their context. In addition, they provide the lines along which complexity must be reduced in order to keep the policy relevant information of the model. In this way, the structure of the model and the dynamics it describes are tailor-made for the users of the model, making its policy relevance more likely. This support in the simplification of the model is vital to the later acceptance of the model. The acceptance is further enhanced by the active participation of those, who will at a later stage work with the outcomes of the model. Developing the model together with the end-users also means that the process of communicating the results will be facilitated. Other than acceptance, structure, policy relevance and communication, there is an inherent benefit of the development in the form of a learning process. In our experience, users who were trained to think in a more integrated fashion were more likely to expand their frame of thinking beyond day-to-day pressures. This, combined with external developments and new insights implies that a SCENE model is never finalised, but rather a momentary snap-shot. Therefore, the SCENE approach is an ongoing process of social learning, and while the SCENE models evolve, awareness for that change is growing.

4.2.3 COMBINING EXPERT KNOWLEDGE AND PARTICIPATORY METHODS

In our practical applications of the SCENE model, we combined expert knowledge and participatory methods in order to satisfy the multiple goals of these studies. The participatory research process is then usually limited to

the stock level of the SCENE model. On this level, it is possible to derive consistent sets of variables (stocks) and the flows between them in relatively little time. The process from first draft to completed SCENE model consists of an iterative cycle of reflection from stakeholders and researchers. The ongoing interaction is needed in order to avoid dissociation, where the relevant stakeholders do not recognize their own input back in the final draft of the model. The input from the researchers helps to reduce selectiveness, bias and the effects of 'group think'. The case study in section 4.4 describes a practical implementation of this process.

4.2.4 NORMATIVITY, SUBJECTIVITY, AND COMPLEXITY

The clear structure of the triangular model makes it possible to represent the complexity of sustainability in a way that does justice to the needs of communication and analysis alike. For communication purposes one can choose an appropriate level of detail to provide insights into the developments and context of sustainability related issues. For analysis the derivation of key-issues and stocks using qualitative participatory methods reduces the complexity without reducing their semantic richness. By postponing the translation of the conceptual model into the black box of a quantitative model, we enable participants to support maximally the process of defining the details of the stocks and flows. The effect of this is, that in the process of reducing the complexity, policy relevant aspects are preserved as much as possible.

4.2.5 ROLE DISTRIBUTION

One of the characteristics of Post Normal Science (PNS) is the progressive integration of knowledge from science, policy and other societal domains. In applied PNS this sometimes leads to confusion about the different roles of in a project. While this cannot be totally avoided, awareness about the different roles helps significantly in the process (for example in the management of expectations) and the interpretation of the results (for example in the awareness of whose perspective the result is a reflection of).

In a case study, there often is a problem owner who has a specific interest in the conduction of a case study and sometimes, as in the case of the province of Limburg case study, also commissions the research. The role of the problem owner is a matter of preparation and negotiation. A case study is of little use to all parties when the influence of the problem owner limits the scope of the project or the application of any results from it. Only when the

independence and the transparency of the project can be asserted, should a SCENE process be considered. On the positive side, a pro-active problem-owner is more likely to integrate results from the process in regular decision-making processes. This makes a significant portion of the after-care the responsibility of the problem-owner.

Before the beginning of the project, the problem owner and researchers decide the degree to which the results of the stakeholder-based processes are binding for any subsequent decisions. This can range from an informing role to one of co-decision-maker. The envisioned role also influences the (self-) selection of the stakeholders. People are much more likely to commit to the process if they can expect that their input will be taken seriously and likely lead to appropriate changes. We have therefore always encouraged problem-owners to in advance commit to an influential role of stakeholder input. On the stakeholder side, it greatly facilitates and enriches a process if the stakeholders are taking part on personal title, rather than as representatives of a group. Representatives will find it much harder to neglect political interests and other (hidden) agendas for the time of the development process.

The researchers involved in a SCENE project have several different roles. The roles closest to the traditional role are that of process designer, analyst and reflector. The researcher usually is the initial developer of the overall structure of the process, which is then refined in negotiation with the problem owner and later in the process also with the stakeholders. Chapters 5, 6 and 7 document several methods of analysis that are based on the SCENE model. The relevant calculations and interpretations are in the domain of the researchers. As a reflector the scientist continuously provides feedback based on selected criteria. These concern broad systems properties (a lack of systemic feedbacks is a common observation) or specific aspects (for example pointing out scientific research that supports or contradicts statements documented during the stakeholder process).

In normal science, the researchers would also have the role of solving the issue in the way that he would process the input and come to a transparent and objective conclusion on a good strategy. In SCENE this role as problem solver shifts to one of problem recognizer. The researcher distills relevant questions from the input and uses these to structure the group learning process without a predetermined result. This also helps to distinguish the scientific process from the societal process. In the case of contested issues with little scientific proof available, taking a decision on the path to choose

between the ambiguous options should not be a role for researchers, but for democratic decision-making.

Finally, an independent person, who is well aware of the methods applied, the objectives and the different roles, best fulfils the role of process facilitator or problem mediator. It is increasingly common in PNS to hire professional external facilitators or mediators.

4.3 Applying SCENE

The SCENE approach is a basic tool for a wide range of sustainability related research issues and research goals. The SCENE model can serve as a framework for integrated and structured thinking about complex issues, for monitoring sustainability, for the evaluation of complex sustainability related issues and for strategy planning. SCENE can also be used as a qualitative modelling-framework for quantitative modelling. In addition, the model provides a tool for the communication of these issues.

Integrated thinking

A straightforward, yet highly effective and efficient application of the SCENE model is as a tool for mapping out and structuring complex interrelations. In this way, the integrated context of stocks can be represented. From this conceptual approach, points of attention and points of discussion can be derived. One example of the SCENE model used as a structuring tool is the NMP4 biodiversity project (Rotmans, van Asselt, de Groot et al., 2000). This project was an integrated analysis of biodiversity as a support study for the 4th Dutch National Environmental Outlook (NMP4). The SCENE model was used to describe the current state of the stock of biodiversity, in order to establish the driving forces for its development, place biodiversity in its context with developments in other domains, sketch future developments and define policy levers for an improvement of biodiversity. The added value of the SCENE model lies in the clear picture it provides of driving forces, effects and policy options. The structural presentation facilitated the quest for win-win strategies in preserving biodiversity.

When applying the SCENE model as a structuring tool, one can choose to describe the model either at the stock-level or at the level of the

characteristics, depending on the level of detail requested. With a process of full participation it takes a set of highly dedicated stakeholders to go beyond the stock level. At the same time, a process of co-production between researchers and stakeholders adds significantly to the content of the model and simultaneously raises the awareness for interactions and the context of their own field among the participants. The importance of this effect should not be underestimated. It has been the most important outcome in many of our governance consulting projects.

Evaluation framework

Choices concerning the inherent trade-offs related to the concept of sustainability can be explicitly weighed against each other, based on the SCENE model. The inherent ambiguity of the concept can thus be addressed. This application is especially useful in the evaluation of projects where trade-offs between the three domains of sustainability have to be negotiated, for example large infrastructural projects. In order to make the choices between different ambiguous options meaningful, it is helpful to focus on the layer of characteristics. In our experience, this layer best represents the consequences of choices that are made to improve certain key-indicators, like economic growth and maps out the hidden consequences in terms of quality, function and space.

In an expert based process, the well-structured overview of economic, ecological and social stocks and characteristics provides a framework for an integrated and balanced evaluation. Any bias towards one or two corners of the triangle becomes explicit. Generally, the content gains from a participatory approach, where implicit choices made by stakeholders become explicit. The awareness for the existing sets of trade-offs in combination with the perspectives of the stakeholders opens up the space for a well-funded discussion on ambiguous options.

The practical implementation of the SCENE model at the province of Limburg provides many examples for this type of application. The consequences of different policy strategies can be mapped out. Including the positive and negative side effects. Based on these options, trade-offs between economic, ecological and social aspects are made explicit, allowing for transparent choices. As a consequence, systemic relationships and trade-offs are now much more visible in the Province's projects and documents, as we will see in chapter 8.

Another relevant example in this context is the TOK project. In this project a set of scenarios for the year 2030 was developed for the city of Maastricht (Kockelkorn et al., 1999). The version of the SCENE model used in this project represents the driving forces and effects of policy strategies and offers to the decision maker well structured insights into possible future developments.

Surface

The surface of the triangle can be used as an additional way of storing and communicating information. The size of each corner of the triangle is a visual indicator for the strength (or capital value) of that domain. We distinguish three possible developments of the surface of the triangle as a whole: weakening, substitution and strengthening. These are visualized in Figure 4.3.

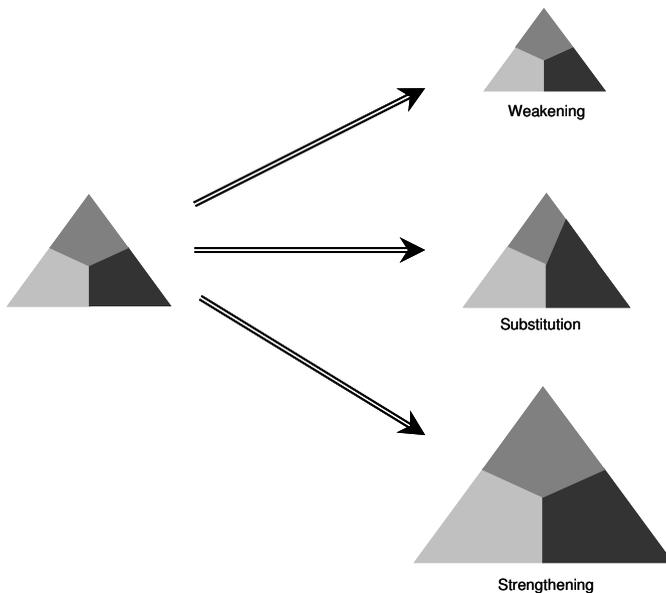


Figure 4.3: Possible capital developments of the SCENE model

Weakening implies that there are losses in the capital of all three domains. Strategies that have the effect of weakening can be considered as the least

sustainable option as on all three domains the ability of future generation to satisfy their needs is compromised. The notion of substitution can be used to describe the concept of weak sustainability with all its advantages and disadvantages. With substitution, one domain grows at the expense of another. A widely discussed example is the growth of economic capital at the expense of natural capital in the western world, especially in the 1960's and 70's. But it is important to note, that comparable links can be described between the economic and social domains, and the ecological and social domains as well. Also, these relationships are not necessarily uni-directional. The economic costs for protecting the environment and for repairing damage done in the past are an example where ecological capital grows at the expense of the economic capital. Strengthening is the third general form of development of the triangle. This process is related to the notion of strong sustainability. By the improvement of all three capital domains, the future ability to satisfy needs is improved. An example would be the successful creation of a natural reserve that is also exploited for recreation. The ecological domain benefits from the extension of natural area, in the economic domain employment and income are created and in the social domain new recreational facilities are added.

Strategy development

In addition to evaluating strategies, the SCENE model can also serve as a tool for the development of policy options and strategies (combinations of options). A useful starting point for the development process is the identification of policy-levers, i.e. stock, characteristics and flows that governance has a direct impact on. Evaluating the consequences of influencing one or more of these policy levers results in strategies that are not influenced by entrenched goals. For example, it is not possible for provincial governance in the Netherlands to directly influence the population size. But, population size being a crucial variable in sustainability related issues, strategies are often focused on achieving the impossible. Letting that focus go and taking a closer look at feasible interventions has delivered more feasible approaches to sustainability, some of which have the side-effect of supporting the preferred development of the population size.

In a process of continued interaction with stakeholders SCENE stimulates all parties to continuously check the underlying assumptions and uncertainties. If done properly, the questioning of assumptions leads to more robust

strategies. Chapters 5 and 8 address the capacities of SCENE to support strategy development in more detail.

Monitoring

The list of indicators derived during the model development process is a useful basis for the monitoring of sustainability. The advantage of the list is the fact, that the indicators are not loosely selected to represent a range of disciplines or topics, but that they also form a consistent and interrelated set for further analysis and interpretation. It is obvious, that in order to select indicators, it is necessary to complete the SCENE structure to that level of detail for any meaningful results.

Based on an expert process, the Telos Institute has built a monitoring tool for the Dutch province of North-Brabant based on the SCENE model (Telos, 2001). The method has later also been applied in the provinces of Limburg, Zeeland and Flevoland (Knippenberg et al., 2007). The state of indicators and their development during the past few years are set against benchmarks. In this way, sustainability related developments are monitored. The 'State of Limburg' collection of monitoring indicators is fully based on the SCENE model (Provincie Limburg, 2007a). This publication is based on an extensive systems analysis of the attributes of the stocks. In contrast to earlier editions, the changes in indicators are highlighted, rather than their absolute values. This emphasizes the dynamic nature of sustainability. Chapter 6 contains a detailed proposal on how the SCENE model can not only help in the selection of indicators, but how an indicator for sustainability can be derived from the system structure itself.

Quantitative modelling

The structure of the SCENE model, including the derived set of indicators is a transparent framework for the development of quantitative models in the tradition of system dynamics. Generally, the development of each given indicator in the model can be expressed as a differential equation. The set of explanatory variables consists of all indicators that have a link towards the given indicator, i.e. influence that indicator.

In some cases the complexity of the conceptual model is preventing the straightforward implementation in system dynamic form while still providing any useful results. In these cases, the conceptual model is simplified by selecting a set of key-indicators that are especially important for the development of a system. The criteria for this selection procedure

depend on the question at hand. During the selection process it is crucial to check, whether the main dynamics of the systems as the expert and/or participants understand them are still represented within the system.

Even in a simplified system, it would be exceptional if all links between indicators in that system were thoroughly researched with widely accepted and reasonably certain results. It is important to explicitly address the less researched links either based on a participatory process or based on a set of scenarios, that represent the range of scientific dissent. Chapter 7 addresses the quantification of a SCENE diagram in detail.

Communication

A major motivation for the development of SCENE has been to communicate complex sustainability related with process participants or third parties. It was our intention to structure the complexity before reducing it. This increases the acceptance of the conceptual model as well as spin-off products, such as quantitative models model. Decision-makers are not confronted with a black-box tool that provides them with ready-made strategies, but their own input is digested and structured. As we have chosen for political decisions not to be taken within the model itself, but left this to the policy circuit, decision makers are more likely to accept the tool and use it intelligently. This turned out well in the POL project that will be introduced in detail in the next section.

4.4 Case study

The SCENE approach was tentatively tested and applied in the POL project (Rotmans et al. (1999a) and (2000)). Section 2.6. describes the context of that project. The development of SCENE was still work-in-progress at that time, the common ground between researchers and problem owners was limited and scientific analysis of complex problems in trans-disciplinary settings was scarce. This is reflected in the methods applied, the relative shortage of criteria, checks and balances, and in the documentation, that paid relatively little attention as to the roles of the participating parties and intermediate products of the process.

Province administrators from a range of different departments participated in the project. In a series of workshops facilitated by ICIS-based researchers

the group filled the SCENE triangle. For each of the three corners, the participants freely mentioned possibly relevant aspects, which were documented on post-it notes. The ICIS researchers proposed a grouping of the notes in about 50 clusters, on which the participants reflected: Did the clusters cover all initial notes? Did the clustering reveal blind spots? Did the clustering trigger associations leading to more notes? The iterative process of clustering, adjusting and adding new notes lead to a preliminary set of stocks.

The flows were added in a process of group discussion in which participants described the existence of a flow between two stocks and what he or she perceived that flow to be. The group would decide together, whether or not to record a flow. ICIS researchers/facilitators took part in these discussions when they noticed that flows did not match empirical data or other scientific or public research. The selection of flows offered was undoubtedly influenced by the background of the participants and by related activities. For example, the attention to details in water related stocks might well stem from sector specific training exercises, mostly on water related issues, that ICIS had conducted to show the value of an integrative approach. In the process of adding flows, minor re-arrangements in the stocks and clusters were made.

The expertise of province administrators and ICIS-based researchers, as well as the contents of an inventory of forward-looking studies with relation to the province of Limburg served as main check for the comprehensiveness of the model. The POL model (in Dutch) of 52 stocks and 244 flows resulting from that process is shown in Figure 4.4.⁶

With hindsight, this model lacks comprehension and balance. Despite this, the model seems to have fulfilled some of our ambitions. It was printed on large posters displayed in the rooms of those members of the administration involved in the development of the strategic plan. In this manner it served as a constant reminder and basic tool for the horizontal integration of policy domains. The process itself also revealed areas about which relatively little was documented, such as the qualitative aspects of the social domain.

⁶ The model was developed to be printed on posters. Unfortunately, scaling the model down to the format of a book renders it undecipherable. It is included here for illustrative purposes. A full-size electronic version is available from the author.

Arguably the most important result was self-consciousness among the process participants for the system they were trying to manage. The project stimulated the desire to continue along this road, further develop the structure, and operationalise a systemic perspective in the organisations, the tools used within the administration, the reporting, and the decision-making process. The POL project laid the foundation for a cultural change in the province administration. The following chapters concern the further development of that foundation.

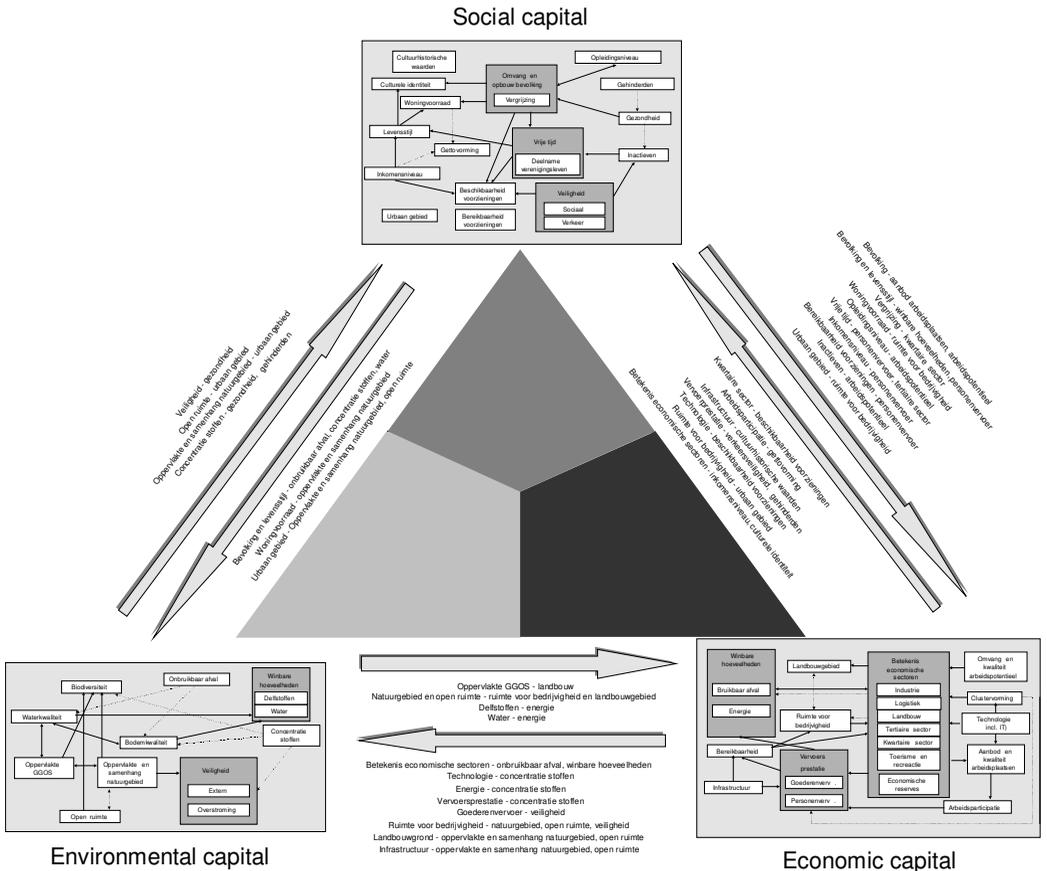


Figure 4.4: The 1st generation SCENE model for the Province of Limburg

4.5 Conclusion

Any approach should as an end result not only highlight the driving forces of sustainability, but also the levers available to influence the system in cooperation with those who handle these levers. Otherwise such an approach will remain of academic interest only and not do justice to the concept of sustainability at the interface of science and politics. In the context of, among others, regional policy-making, an interactive process of exploring and mapping the systemic structure of a region and its society seems a necessary condition for a better understanding. This need inspired the development of the SCENE approach.

Interestingly, the first (and, as described above, rather clumsy) implementation of the SCENE model has had a number of positive effects. In an interview with some of the group participants in the summer of 2007, the initial work with the triangle was generally seen as an important stepping-stone towards more integrative analysis and more horizontal co-operation. The use of the exact structure of the model has been limited to a series of informal thought experiments and discussions about what should be evaluated together or which externalities should be taken into account in specific sectors. However, the discussions about the structure, specific stocks and flows, as well as arguments on whether or not one could derive policy relevant information from the model also had a motivating effect. It triggered a desire for more support in integrative and complex decision-making. All parties, including the researchers, felt that it was feasible and useful to follow the initial lines and explore qualitatively the sustainability related complexities in Limburg. Nevertheless, significant improvements in the process, its implementation and the interpretation and representation of results were necessary. As we will see in chapter 5, a number of these were soon to be implemented in follow-up projects.

In chapter 1, we have formulated the ambition of a four-fold integration that is the quest of this dissertation. The SCENE approach is a basic tool to facilitate the integration among policy domains through raising awareness for the interconnectedness. For researchers, it provides a list of issues to be integrated and a benchmark concerning the level of knowledge within the organisation. This includes not only the items themselves, but also the policy maker's perception of the wording, structure, and emphases. In turn, this facilitates the integration of the scientific findings in the framework of the SCENE model, which has through the process already partially permeated

the target organisation. The active involvement of representatives from many different backgrounds helped to address some of the normative elements and legitimise the product within the province.

In terms of the research priorities drafted in chapter 3, the SCENE approach initially focuses on the dynamic leverage points and serves as an integrative framework for the static leverage points. Its qualitative, yet structured process supports a broad and comprehensive coverage of the system. Its specific design for participatory input enables dialectic qualities. The SCENE process has been successfully applied in several different settings in the Netherlands, providing us with an indication that process is transferable to other settings. The transferability of the context specific results is, of course, limited to illustrative purposes and rough comparisons between regions. The SCENE approach is a useful tool to support joined-up thinking that will be expanded in chapter 5.

Chapter 5

ENRICHED QUALITATIVE
SYSTEMS ANALYSIS (eQSA)

There is something new under the sun. (John McNeill (2001))

The previous chapter described the participatory SCENE approach, a method to improve the systemic understanding of an issue by mapping it in the form of stocks and flows. The value of the SCENE approach lies mostly in the process of drafting the systems map. This chapter focuses on reaping the information contained in such a map by way of structural and logical analysis. The enriched Qualitative Systems Analysis (eQSA) approach is a novel extension of existing participatory structural analysis methods.⁷ We developed the method aiming to support systemic understanding and to facilitate the intuitive identification, evaluation and communication of policy levers and strategies. In this way we hope to clarify the consequences of strategic options.

The first section of this chapter provides an overview of methods for qualitative systems analysis. The second section introduces the eQSA approach, illustrated with an extensive case study on the Dutch Province of Limburg. An overview of the case study and a detailed description of the four project phases and their results introduce the eQSA approach step-by-step. In section three, we take a closer look at a number of other case studies testing some of the assumptions underlying the approach and illustrating the strengths and weaknesses of the eQSA method. An evaluation of the approach, including thoughts on further research round off the chapter.

5.1 Methods for Qualitative Systems Analysis

In previous chapters we echoed the widespread call for improved research methods with respect to persistent problems and sustainability assessments. Quality criteria for such methods are:

- the ability to comprehensively analyse complex systems without a limitation to quantifiable aspects.

⁷ This paper is based on Grosskurth (in preparation). The author of this dissertation has developed the eQSA method and was responsible for its implementation, especially as process designer and as one of three process facilitators.

- the capacity for meaningful stakeholder involvement.
- the possibility to support fundamental systemic changes.

The results yielded from such methods need not necessarily be objective, falsifiable and reproducible. In this context, problem recognition and exploration take precedence over problem solving. The most popular and widespread method for the analysis of complex systems is quantitative modelling, such as system dynamics. We take a closer look at the potential of quantitative models in chapter 7. At this point it suffices to say that the major strength of such models lies in their ability to simulate a part of the dynamics of highly complex systems. These simulations provide valuable insights into the underlying dynamics, as well as the role and interplay of different variables in a system. These advantages come at a price. The need for quantification leads to a over-representation of economic and technological factors and an under-representation of soft factors including socio-cultural and institutional. In addition, the development of such models is very time- and resource intensive, the capacity for stakeholder engagement is generally limited, their structure is rather rigid, and the need for quantification is a cause of shortcomings with respect to the breadth, depth, and balance.

In response to the limitations of quantitative simulation models, a number of qualitative reasoning methods for systems analysis have been developed. The roots of some of these methods reach back to military research around the end of World War II, when complex strategies had to be evaluated in an uncertain environment (operations research). In the 1970's, about the time when Meadows et al. published the results of one of the most influential system dynamic studies in *The Limits to Growth* (1972), qualitative methods for systems analysis were further developed and first applied in regional assessments and sustainability related studies. The methods are known under many different names, depending, among others, on their disciplinary and geographical origin. In this dissertation, we use the term Qualitative Systems Analysis as an umbrella. The following overview of methods is far from complete. Similar methods are used in a number of fields with little reference to each other. Examples of disciplines of origin are strategic planning and management, operations research, artificial intelligence, cybernetics, and psychology.

The closest qualitative methodological relative to system dynamics is the method of Qualitative Simulation (Kuipers, 1994). The approach explores all

logically possible developments of a system consistent with a rigidly pre-defined set of qualitative rules and assumptions. Qualitative differential equations are being used to identify all future states of a system that are consistent with its current state. The result of the method is a pattern of system states revealing, for example, possible lock-in situations or the periodical recurrence of system states, as well as bifurcation points, where the system might evolve in several directions.

A method with a focus on analysing the information requirements for managing a system is the Soft Systems Methodology (Checkland (1981); Checkland & Scholes (1999); Wilson (2001)). Like SCENE and eQSA, this methodology is based on system diagrams. The method forces its users to be as specific as possible about the structure of the system and the information required to evaluate possible interventions. Most applications of the method are found in the fields of business strategy and change management, where it is used to distil concrete definitions, and actions starting from a vague problem conception.

A whole set of QSA methods have quite another methodological heritage: cross-impact analysis methods. The methods are generally grouped under the header of 'structural analysis'. An early predecessor of these methods is Trend-Impact Analysis (TIA). This method requires a curve fitted to time-series data to estimate a future trend in the absence of unprecedented events. Experts then identify possible future trend changing events and judge the probability and impact of each event. The method was soon extended to improve the way of taking cross-impacts into account: events that made other events in the list more or less likely. Detailed descriptions of both methods can be found in Glenn & Gordon (2003), chapters 5 & 6.

The KSIM method, developed by Kane (1972), is related, but less demanding with respect to quantification. A qualitative impact matrix is constructed from a list of variables and their interactions. The KSIM method is based on a few simple rules:

- Variables have a pre-defined upper and lower bound, rescaled to be 1 and 0.
- A variable increases when the net impact of the other variables is positive and decreases when the net impact is negative.
- When the upper and lower bounds are approached, the variable's response to impacts decreases.

- Large variables have large impacts and small variables have small impacts, *ceteris paribus*.
- Looped networks of binary interactions describe complex interactions.

The impact matrix and assumptions on the current value of each variable serve as a hypothesis, from which logical projections are derived within the bounds of the set rules. The method has since been applied to a wide range of cases in the fields of industry, public services and policy.

Another approach based on impact matrices is the Sensitivity Model (Vester, 2002). The Sensitivity Model (SM) is a computerized planning and mediation tool for complex systems. Before computers became standard tools for analysis Vester called a simplified version of the approach 'Papiercomputer' (paper computer). The term is fitting as Vester uses simple matrix calculations, which can be done on a sheet of paper without mathematical skills, to support the structural interpretation of a system. The most important calculations are the active and passive sum, representing the degree of influence and dependency of each variable relative to all other variables in the impact matrix. Vester embeds this and other analyses in a participatory process that is mainly applied in contexts of regional and local planning processes.

The MICMAC method developed by Duperrin and Godet (1973) is quite similar to the SM approach in its initial stages, but extends the method of mapping active and passive sums to include indirect effects. The acronym stands for *Matrice d'Impacts Croisés-Multiplication Appliquée à un Classement* (Cross impact matrix-multiplication applied to classification) and already describes the core characteristic of this approach. When multiplying the impact matrix by itself, the resulting new matrix reveals the impacts of indirect effects. Detailed descriptions of the method, including an overview of strengths and weaknesses, can be found in Godet (1986), Scholz & Tietje (2002), and Arcade et al. (2003).

The eQSA approach presented in this chapter is an extension of the SM method and the MICMAC method. In addition to the theoretical input, two similar tools have inspired the work on eQSA: RAP and the Sensitivity Model Prof. Vester.

RAP is a computer-based tool for the analysis of the effects of different policy strategies for influencing complex societal issues (Arcadis, 2004). Its step-wise procedure allows the user to enter variables and relations between

them in a square influence matrix. The user sets relations between variables in the range from + + + (strongly positive) to - - - (strongly negative). The resulting system can be influenced using policy options that have user defined first-order effects on selected variables, where the influences are defined in the same way as the relations. The software deduces the multiple-order effects to an arbitrary degree, providing some insights into the indirect effects of selected policy strategies applied individually or in combination under a set of assumptions.

The Sensitivity-Model Prof. Vester is comparable to RAP (Vester, 2002). Relations between a set of user selected variables are described in an influence matrix with integer values ranging from -3 (strongly negative) to $+3$ (strongly positive). From the resulting matrix, a series of descriptive values for each of the variables is deduced, using simple matrix algebra. These include the active and passive sum (the weighted sum of outgoing and incoming relationships per variable), the multiplication of the active and passive sum as an indicator for the importance of a variable in the network (force) and the division of the active sum by the passive sum as an indicator of the character of a variable (dominating or listening). Smaller subsets of the system can be simulated in a user-friendly manner using the basis of fuzzy logic.

We opted for a the combination of SM and MICMAC, because the methods are readily compatible with system maps, such as those resulting from a SCENE approach, and well suited for intense participatory engagement. They all share the format of conceptual system maps with system dynamics. Such maps can alternatively be denoted as matrices where each cell denotes a potential uni-directional arrow between two vertices. Even rudimentary matrix algebra is sufficient to visualise some characteristics of a system and its elements that are otherwise hidden to the observer. SM and KSIM are treated in more detail in section 2 where we describe the eQSA approach in detail based on a case study. However, applying these methods is not a panacea to understanding complex systems. In the evaluation in section 5.4, we will discuss the merits and problems of the approach, partially based on the experience of users.

5.2 Method and application of eQSA

In order to introduce the eQSA method, we describe its application in our work with the Dutch province of Limburg on the Provinciaal Omgevingsplan Limburg, or POL (Provincie Limburg, 2001). Introducing the method with the case reflects the process of co-development between researchers and stakeholders leading to the refinement of the method.

The Limburg case is outlined in sections 2.6. Section 4.4 describes the first steps along the lines of conceptual systemic integration based on SCENE. The SCENE model in the previous chapter preceded the draft of the first POL plan. During the development of this plan and during the discussions following its publication it became evident that little attention had been paid to the interactions between the different sectors and to uncertainties concerning the forecasts underlying the plan (Rotmans, van Asselt, Grevers et al., 2000). In order to address these issues, the provincial government commissioned a year-long participatory project to review and redesign the tools for exploratory studies.

The aim of the project was to develop and operationalise within the Province Administration a systemic understanding of the Province reaching beyond the information provided through indicators, mathematical models, and sector specific studies. External expertise and internal latent knowledge about the province should become an integral element of the systemic understanding in order to identify and evaluate issues of concern and potential policy interventions.

The project was to result in a transparent, integrative, comprehensive and flexible set of tools representing the region as a system and highlighting system elements of specific relevance for governance, especially policy levers. The set should be compatible with existing national and regional forecasting models and results should help to improve and integrate the forecasting models already used by the administration. The set of tools should also enable groups of members of the administration to execute system scans of upcoming issues without continuous external support. The project provided us with an opportunity of an extensive case study for the eQSA method, that had until then not been tested on a large scale.

To fulfil the project criteria, we proposed a participatory approach based on structural analysis, paying special attention to existing quantitative forecasting models. The project proposal consisted of four work packages: 1.

a participatory conceptual inventory and structural analysis of the region from a systems perspective (Qualitative Systems Analysis – QSA); 2. an inventory and evaluation of quantitative indicators and forecasting models used by the provincial administration for strategic planning; and 3. the development of a support tool for strategic planning based on the integration and enrichment of the two inventories (enriched Qualitative Systems Analysis –eQSA). 4. a phase for documentation, implementation and training concluded the project.

Six senior civil servants with different backgrounds formed the core group for the participatory process. The participants had all previously expressed an active interest in co-operating horizontally with other departments, taking into account external uncertainties, surprises and qualitative developments. From the process they expected a better understanding of how their domain of expertise interacted with others. They also were looking for tools to embed this understanding in their daily processes of strategic planning and sustainability.

Especially early in the process, additional members joined this group to provide specific practical expertise, to increase internal transparency, and to satisfy general interest. During the year, the group worked with our team during 17 workshop sessions of approximately three hours each.

5.2.1 SYSTEM INVENTORY

To facilitate horizontal integration in the administration and among the different forecasting models, we dedicated the first phase of the project to the participatory development of a conceptual integrated system representation of the province. The purpose of this process was to strengthen mutual understanding, to sharpen awareness for the interdependencies between different characteristics and regional developments, and to structure the available information and knowledge in an integrated manner. We drafted the conceptual integrated system using the SCENE approach described in the previous chapter.

Some of the participants of the POL working group were already familiar with the SCENE method, as they had already taken part in drafting the first SCENE model described in the previous chapter. For the POL project, the group felt the need for a new system map as the earlier model was not intended for a structural analysis, nor had it been structured for an analysis of relevant quantitative models. In addition, the group expected to be able to

improve on the lack of comprehension and balance found in the first SCENE project.

The process had been improved on several points in order to alleviate some of the previous shortcomings. First and foremost, we took more care in defining and distributing the different roles in the process. There were two main researchers involved in the project and a professional consultant/facilitator, separating the roles of facilitator and researcher. Two of the participants acted as project leaders on the side of the Province, representing the project within the organisation and taking care of input from outside the group of participants, when needed. Next to the roles, we also took more care to define, together with the participants, what stocks and flows were and at which level of aggregation the model would be most useful. Also, criteria were added and/or developed to judge the comprehensiveness and degree of balance in the system. A lot of attention was paid to related processes (for example statistical forecasting) and sources of information (for example national reports on related issues). Finally, the documentation of the process was more detailed and focused on the process as well as the results.

As in the previous chapter, drafting the influence diagram for the POL project was an iterative two step process: 1) selecting and describing the stocks that comprehensively represent the province of Limburg and 2) mapping the influences between these stocks. These two steps were again repeated several times. Each iteration and the discussions that went with it, helped to clarify outstanding points, clear up misunderstandings, helped to embed newly arrived external information and increased the familiarity with the system. The external criteria and input from outside the group helped to avoid so-called group-think.

The participants were advised to develop the diagram of the province in the present and to try not to anticipate possible future changes. The concept of sustainability was implicit in the process as the balanced development of the economic, socio-cultural, and environmental domains. During the first several sessions, the group proposed a set of issues and topics in the same way, they had during the first SCENE project (section 4.4): the participants offered aspects that were noted on post-its that were then clustered. The clusters were proposed by the researchers involved with as only guidance the intuitive degree of 'togetherness' of the aspects in any one cluster. In an interactive and iterative process of screening, revising, reformulating, and restructuring, this initial list was brought together into a set of 18 stocks.

The level of aggregation was set by the group in such a way that we would end up with 15 to 25 stocks. There are several reasons for this. Experimenting with stocks at different scale levels, we found that the chosen level of aggregation corresponded to the level of abstraction of many strategic policy documents as well as to clusters of monitored data. Also, higher aggregation of stocks would hide a lot of the relevant flows within the stocks, lower aggregation (more stocks) would make the system unwieldy and any quantitative analysis a black-box operation. Choosing the level of aggregation alone took the time of 4 sessions, since a lot of experiments with different scale levels were required, as well as external information.

We spent several hours individual work and group discussion time to find the right title for each stock. To this end, each stock was described in writing by a researcher or group participant. These descriptions were then put to the discussion. The stock description was also checked against the initially mentioned aspects that the stock was meant to cluster. Titles were then proposed and tested outside the group of participants within the organisation. We asked the interviewees to describe their associations with a given title word. The titles would have to invoke associations in line with the description. In some cases, this testing has also led to the addition of new aspects and extensions of stocks. The structure and final wording of the stocks corresponded to the administration's internal organization and communication habits.

The researchers and facilitators were aware of the fact that they might unintentionally influence the wording, clustering and selection of the stocks during the participatory sessions. Therefore a number of checks and balances were introduced. Initial checks on comprehensiveness and balance were executed by comparing the stocks and their descriptions with the set of indicators regularly monitored by the administration documents on strategic planning published by the Province. Some of the comparisons were based on non-public documents and executed by participants from the administration. The final version of the stocks covered a wider range than any single one document we compared it to.

In addition to a reflection on the set from Province administrators and fellow researchers not part of the drafting process, the set of stocks was also screened for comprehensiveness based on Vester's (2002) system criteria. Vester derived these criteria based on his experience and system understanding to guarantee the balanced representation of all relevant aspects of the system. These system criteria require at least the implicit

inclusion of seven specific areas of life (e.g., economy, population, and infrastructure), three physical categories (material, energy, and information), four dynamic categories (e.g., flow or time dynamic), and four system-relational categories (that open the system to outside influence or can be influenced from within). The 18 stocks selected satisfied the areas of life and the physical categories of Vester's criteria. The dynamic and system-relational categories were not satisfied at this stage as they can only be met after the relationships among the stocks have been represented.

Yet, despite these checks and balances, the selection and structuring of the stocks is to some degree arbitrary. Later in this chapter, we will take a closer look at how different choices influence the results from a systems analysis by comparing the first SCENE model to the eQSA SCENE model described in this chapter.

The 18 stocks (and their labels) are: population (Popu), consumption (Cons), social structure (Soci), public amenities (Amnt), housing (Hous), security (Secu), space (Spac), air (Air), groundwater (G-wat), surface water (S-wat), soil (Soil), nature (Natu), entrepreneurship (Entr), production (Prod), knowledge and innovation (Know), work (Work), mobility (Mobi), and infrastructure (Infr). For each stock, the group formulated an abstract with a description of the stock, its characteristics, and its scope. For example, the characteristics of the stock 'knowledge and innovation' include knowledge institutions, public and private research and development, applications of knowledge and innovation, as well as the spatial distribution of these factors.

To draft these relationships, the participants individually filled in an empty matrix with the 18 stocks as row and column titles. Each cell represented a potential unidirectional flow from one stock to another (though not to itself), resulting in 306 possible flows. Working independently, the participants marked the cells where they perceived a direct influence. They then added a strength estimate (weak, medium, or strong) and a short description of the influence they had in mind. We then combined the individual matrices into a single document, checked the results with experts and, in instances of doubt or dissent, presented it to the group. Among themselves, the participants noted their agreements, discussed and clarified their disagreements, and identified uncertainties. We added comments from the expert scan. These remarks were mostly concerned with additional information about previously identified relationships, thus enriching the group's discussions.

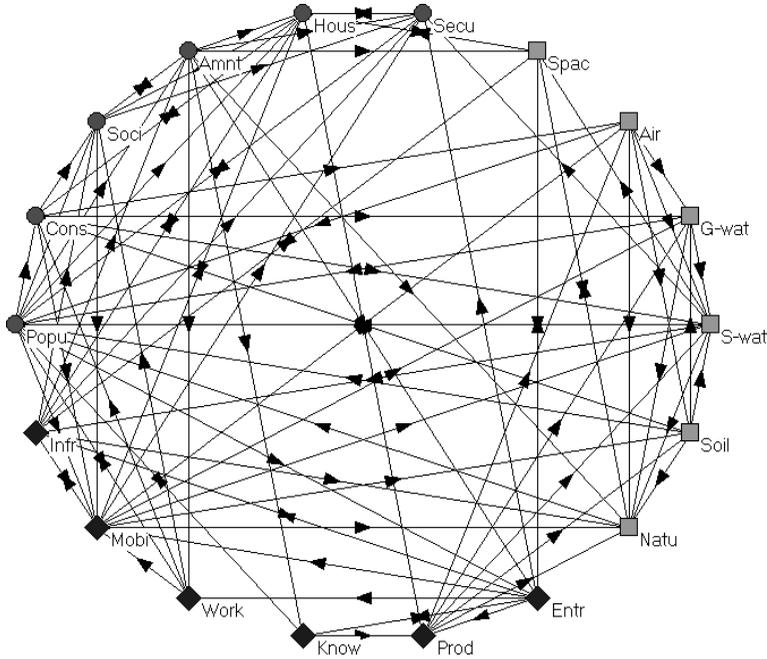


Figure 5.1(a): The conceptual SCENE model for the Province of Limburg as a system graph.

In total, 95 flows were selected and documented—15 strong, 46 medium, and 34 weak. For example, both participants and experts agreed that the stocks ‘population’, ‘public amenities’, and ‘entrepreneurship’ all influence the stock ‘knowledge and innovation’. This latter stock, in turn, influences ‘entrepreneurship’, ‘population’ (both are bi-directional relationships), and ‘production’.

Figure 5.1(a) shows the complete influence diagram (without the flow strength). The stocks are coded by colour and shape to make clear the distinction between stocks in the various domains: sociocultural (red circle), environmental (green box), and economic (blue diamond). Figure 5.1(b) shows the corresponding influence matrix. In the matrix, each cell corresponds to a possible unidirectional flow in a system, where direct flows from a variable to itself are excluded. In the matrix, a flow from ‘population’ (Popu) to ‘consumption’ (Cons) is noted as ‘3’ in the second cell of the first row. The reverse relationship can be found in the first cell of the second row

	Popu	Cons	Soci	Amnt	Hous	Secu	Spac	Air	G-wat	S-wat	Soil	Natu	Entr	Prod	Know	Work	Mobi	Infr
Popu		3	3	3	3	1							2		2	1	2	
Cons			2		2			2	2	1	1						2	
Soci				2		2											2	
Amnt			2		2	1	2					1	1		1	2		
Hous		1	2			1	3										2	1
Secu	1				2													
Spac					2							2	1					1
Air	2								1	1	2	2						
G-wat	1									1		2						
S-wat	1					3	2					2						1
Soil	1								2	1		2						
Natu	1						3											
Entr				1		2	3							3	2	3	3	1
Prod					2			3	2	2	2	1						
Know	1												2	2				
Work	2	2	1														3	
Mobi					2	2		3	1	1	1	1						2
Infr			2	1	1	2	2					2	3				2	

Figure 5.1(b): The conceptual SCENE model for the Province of Limburg as an influence diagram.

(0). The values correspond to weights of the relationships. We use the values 1=weak, 2=medium and 3=strong to describe these weights. The value 0 or an empty cell stands for the absence of a direct relationship between two variables.

Up to this point, the process has delivered a well-documented, transparent systems overview of the Province. It structured the group discussion, triggered reflection, made gaps of knowledge visible, and prompted the consultation of internal and external literature and expertise as input for the group's discussion about the different flows.

5.2.2 STRUCTURAL ANALYSIS

Beyond being a tool for process and information management, the influence matrix contains a large amount of information about the system. Much of this information can help to better understand and improve the systems representation. Foremost, the general density of the flows is an indication of the complexity of the system. In this case 95 out of 306 possible flows have been selected, corresponding to a density of 31%. Variables in rows with many links filled in, influence many other variables, variables with columns with many variables filled in, are dependent from many variables. In terms of influence, 'popu' stands out, in terms of dependency 'secu' and 'natu' seem significant at a glance.

Most information contained in the influence matrix is not visible at first glance. The MICMAC method is one approach to elicit and visualise properties of variables and the system as a whole from such a matrix. In MICMAC, the character of a variable is mainly determined by its passive and active potential. The direct passive potential of a variable is the sum of the absolute values of the incoming relationships influencing that object. The direct active potential is the sum of the absolute values of the outgoing relationships from an object that exert an influence on other objects. In other words, the passive potential describes the degree to which an object is influenced by other objects within the system. The active potential describes the degree to which an object influences other objects. The influence matrix allows an easy calculation of the potentials: the passive potential of a variable equals the vertical sum over the column of that variable in the matrix; the active potential equals the horizontal sum over the row of that variable.

The passive potential of Popu = $1+2+1+1+1+1+2 = 10$; the active potential of Popu = $3+3+3+3+1+2+2+1+2 = 20$. As the absolute values of active potential and passive potential yield little information, the results are generally represented in the form of a ranking as in Table 5.1. For the active ranking, variables with identical active potential have been sorted by their passive potential as a secondary criterion and vice versa for the passive ranking.

Another way to represent the same information is in the form of a system grid. In figure 5.2 the passive potential is shown on the y-axis and the active potential on the x-axis. For the decision maker, the relative position of each variable in that figure provides a rough guide to the character of different system variables.

Active ranking	Passive ranking
1. Popu	1. Mobi
2. Entr	2. Hous
3. Infr	3. Spac
4. Mobi	4. Natu
5. Amnt	5. Secu
6. Cons	6. Soci
7. Prod	7. Popu
8. Hous	8. Entr
9. S-wat	9. Air
10. Air	10. G-wat
11. Work	11. Amnt
12. Spac	12. S-wat
13. Soci	13. Infr
14. Soil	14. Cons
15. Know	15. Work
16. Natu	16. Soil
17. G-wat	17. Prod
18. Secu	18. Know

Table 5.1: Active versus passive ranking of the POL stocks.

Variables in the bottom left sector are hardly embedded in the system. Interventions in these variables can be implemented with little attention for their systemic context. In terms of sustainability, the so-called low hanging fruit is generally found among these variables. The top left sector contains dependent variables with little influence. If a decision maker wants to change such a variable, the policy levers to do so are generally found in the two remaining sectors towards the right hand side of the chart. A direct intervention in a dependent variable is generally not effective, as other determinants within the system tend to overrule it. Active variables are found in the bottom right sector of the chart. From a systemic perspective, appropriate levers for intervention are often found in this sector as these variables influence other variables within the systems without being subject to strong feedback cycles that might dampen or exacerbate the effects of the original intervention. Variables in the top right sector of the system are critical. They are part of a maximum number of feedback cycles in the system. An intervention in such a critical variable sets a large part of the system in motion. Surprises and counter-intuitive developments are most

likely, when these variables are affected. Variables close to the centre of the graph sometimes function as regulating variables. However, the effects of interventions in this area are difficult to determine.



Figure 5.2: The active and passive potential of the POL variables

The analysis in figure 5.2 is restricted to the evaluation of direct influences. Repeatedly multiplying the influence matrix by itself and thus raising it to the power of successively increasing values enables us to take indirect relationships into account. Raising the matrix to its second power yields active and passive potentials that take into account direct and indirect influences and feedbacks up to a path length of two steps. Successively increasing the power implies that longer and longer paths are taken into account. After a number of iterations, usually between three and eight, the

corresponding ranks in active and passive potential stabilise. In our case study, the ranking of the active potential stabilises after three iterations, the corresponding passive rank after four iterations. The corresponding passive and active potentials are mapped in figure 5.3.

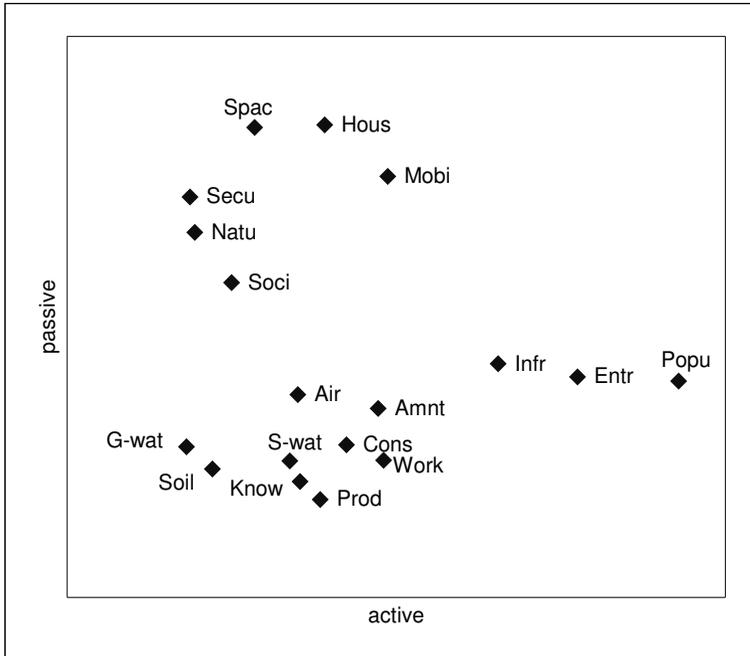


Figure 5.3: The active and passive potential based on indirect relationships calculated using the MICMAC method.

A comparison between the direct and indirect ranks yields further information about the character of the different variables as we can see in table 5.2. The variables 'population', 'entrepreneurship', and 'infrastructure' are clearly the most influential variables in the system. 'Space' and 'housing' score high in terms of dependency. 'Mobility' is the only variable in the critical sector, albeit leaning towards dependency.

The rank changes of 'work', 'know', and 's-wat' are the most significant as far as the active potential is concerned. Especially interventions in 'work'

and 'know' may trigger many unexpected developments that are easily overlooked, as they are the source of relatively few direct flows. With respect to the passive potential 'infr', 'g-wat', and 's-wat' are subject to the biggest changes in rank. Especially 'infr' is a variable to watch in this respect as the many indirect influences on this variable point towards a limited effectiveness of interventions from the side of the administration.

rank active direct		rank change	rank active indirect	
1	Popu		Popu	1
2	Entr		Entr	2
3	Infr		Infr	3
4	Mobi		Mobi	4
5	Amnt	-1	Work	5
6	Cons	-1	Amnt	6
7	Prod	-2	Cons	7
8	Hous		Hous	8
9	S-wat	-3	Prod	9
10	Air	-1	Know	10
11	Work	+6	Air	11
12	Spac	-1	S-wat	12
13	Soci	-1	Spac	13
14	Soil	-1	Soci	14
15	Know	+5	Soil	15
16	Natu		Natu	16
17	G-wat	-1	Secu	17
18	Secu	+1	G-wat	18

rank passive direct		rank change	rank passive indirect	
1	Mobi	-2	Hous	1
2	Hous	+1	Spac	2
3	Spac	+1	Mobi	3
4	Natu	-1	Secu	4
5	Secu	+1	Natu	5
6	Soci		Soci	6
7	Popu	-2	Infr	7
8	Entr		Entr	8
9	Air	-1	Popu	9
10	G-wat	-3	Air	10
11	Amnt		Amnt	11
12	S-wat	-3	Cons	12
13	Infr	+6	G-wat	13
14	Cons	+2	Work	14
15	Work	+1	S-wat	15
16	Soil		Soil	16
17	Prod	-1	Know	17
18	Know	+1	Prod	18

Table 5.2: Comparing direct and indirect influences. The centre column states the number of ranks a variable has moved up or down in the process of the MICMAC analysis.

The distribution pattern of the variables as a whole also gives an indication of the stability of a system. If the variables are mostly located in an L- shape

near the axes, the system is likely to be rather stable as variables are either distinctly influential, dependent, or not embedded in the system at all. If, on the other hand, the variables are predominately clustered around the 45-degree axis, the system tends to be less stable due to the high number of variables in the undetermined and critical sector. In our case the variable cloud seems to suggest a relative stability of the system.

While a statistical validation of the conclusions drawn from a structural analysis is not possible, several tests can be performed in order to test the robustness of the conclusions. The results from these tests often trigger improvements in the system in the way that participants notice inconsistencies or clarify perceptions of a stock.

Segregating the matrix into 9 fields along the dividing lines between the three dimensions of sustainability serves as a rough check on the balance of the model. In our case the distribution of the number of flows over the 9 segments is as shown in Figure 5.4.

	social	environmental	economic
social	18	7	11
environmental	7	13	3
economic	12	13	11

Figure 5.4: The distribution of flows per combination of domains.

The first draft of this model exhibited a severe imbalance, with some sectors of the matrix being almost empty. The segregated matrix served as check on the comprehensiveness and balance of the system. The model in its final form emphasizes links within the social domain. Environmental impacts on the economic sector are least prevalent, in contrast to the reverse flows from economic variables to environmental variables. Impacts of environmental variables are mostly domain specific.

Explaining and discussing each variable in relation to its position within this analysis also sharpens the definitions of the variables and their relationships. This is of significant value in heterogeneous groups that have to find their way to a common language and understanding. Documenting these changes and discussions is also valuable input for an uncertainty analysis.

As the active and passive potential are key elements of the analysis, their robustness should be thoroughly tested. A first test is concerned with the dependency of the conclusions on the weights assigned to the relationships. In order to identify those conclusions, for which these weights are crucial, we reproduce the first and higher order overviews based on a matrix where every relationship is assigned a weight of one. As weak flows are often seen by the group as being somewhat tentative, a second test for robustness excludes these flows and only analyses the system formed by the medium and strong flows.

If any conclusion concerning a variable significantly differs from the original analysis, the causes of the change can generally be traced to a single flow or set of flows. The research on these flows deserves special attention from scientist and decision-maker, as the system seems to be sensitive to changes in its value.

For the administration, detailed results of the structural analysis were documented in an electronic database and a hard copy handbook for each variable and the connected flows. The character of each variable, its potential as a policy lever from the perspective of structural analysis and important feedback cycles including the variable were taken into account. Together with the abstract defining the variables, this information formed the foundation for the set of system tools. This set was considerably extended in the next two work packages of the project.

Up to this point, the process has been a more thorough and more carefully executed replication of the first SCENE process with a different group of stakeholders and with a more concrete idea of what the model could and should be used for. The participants, who were all aware of the first SCENE model, confirmed the improvement in terms of structure and communicability. According to the participants, the process refreshed and strengthened the lessons from developing and applying the first SCENE model. It was in the interpretation and operationalisation of the system that they were hoping for new insights.

5.2.3 MODEL INVENTORY

An inventory and evaluation of the quantitative indicators and forecasting models used by the provincial administration for strategic planning was the core element of the second work package. To compare the completely updated SCENE model described in the previous section with the coverage of the available quantitative models, we conducted an inventory of 13 quantitative models that the administration had previously used for strategic planning. We translated each model into an influence diagram using the stocks described in the previous section. It was obvious from a comparison between the SCENE model derived from the participatory process and that derived from the inventory of quantitative models that the loss of information moving from qualitative to quantitative analysis is dramatic. The quantitative models used for strategic planning cover 30 of the 95 flows in the conceptual influence diagram. The information concerning the quantitative coverage of each stock and each flow was added to the electronic database in order to maximise the synergy between the qualitative and quantitative potential of the tool aimed for. A detailed account of this aspect of the project is provided in chapter 7 and in Grosskurth (2007).

5.2.4 INTEGRATION AND ENRICHMENT

eQSA stands for enriched Qualitative Systems Analysis. It is enriched by characteristics given to each vertex (variable) and arrow (relationship). Up to now we have only processed pure system information that could be implemented in the form of a simulation model if the relevant quantitative information were available. But a qualitative systems view allows for the addition of additional information. Theoretically, any information that can be related to either an object or a relationship and that can be represented in the form of a scale or a typology can be added and processed.

Two sets of information added that proved particularly useful in our work are the speed of change of a variable and the degree of influence on a variable. We asked participants to individually assign values to the speed of significant change of each stock. The following answers were possible: 0 = never changes, 1 = within a generation or slower, 3 = within 5 years or slower, 5 = within a year or slower. At the same time we asked them again individually to assign a value to how much control they thought, the province government had over the development of that variable: 0 = none at all, 1 = little, 2 = reasonable, 3 = almost full control. The results of the individual judgments were compiled and compared. Those items where

outliers or significant differences in judgments were apparent were discussed in the group. This again led to an important exchange of information within the group and provided insights for policy interventions within the system. Items with a large agreement between the participants were not further discussed.

Judging the speed of change provided valuable insights into the time frame one had to consider when addressing an issue or when using a stock as a policy lever. The answers are summarised in table 5.3.

Popu	Cons	Soci	Amnt	Hous	Secu	Spac	Air	G-wat	S-wat	Soil	Natu	Entr	Prod	Know	Work	Mobi	Infr
1	4	2	4	4	4	2	3	2	2	1	2	2	4	2	2	3	3

Table 5.3: The estimated speed of change of the POL stocks.

In general, the social stocks change fastest, the environmental stocks slowest. While this might not be surprising, the assessment concerning the speed of change was introduced to prevent users from running their finger along the flows and imagine, how an ingenious intervention in one point would change other parts of the system and eventually solve many issues at once. It forces the user of the tool set to think about time, an obvious point from the analytical point of view, but often neglected in political processes.

The presentation of table 5.4 surprised the group. The table represents the degree of direct influence the Province and other stakeholders have on the development of a stock. The higher the value, the more influence a stakeholder is perceived to have. The numbers for the Province show the aggregate values after group discussion on a scale from 0 to 5. For the other stakeholders, each number represents the estimate of one group member on a scale from 0 to 3, illustrating the differences in perception. Including other actors improves the understanding of possible conflicts and alliances, as well as the stakes of other actors.

	Province	Communities	National government	International government	Business	Societal organisations	Individual citizen
Social capital							
Popu	1	12001	22111	21011	11001	10010	10?01
Cons	2	22110	12222	11101	12222	11111	21222
Soci	1	22111	11112	00001	21012	21222	32122
Amnt	5	22333	22332	11021	10112	21223	10112
Hous	4	33333	21222	00011	00001	10112	00102
Secu	3	22222	22222	11?11	20212	10112	20112

Environmental capital

Spac	4	22333	21213	11111	10212	21112	00101
Air	2	12?12	22?22	22?21	22?12	01?01	11?01
G-Wat	2	21?12	22?12	12?01	22?12	01?01	01?01
S-Wat	2	11?12	22?22	22?01	12?12	01?01	01?01
Soil	2	21?12	12?12	02?01	22?12	11?01	01?01
Natu	4	21233	22223	12221	01102	21212	10001

Ecological capital

Entr	3	21222	12222	01212	22333	00011	00000
Prod	3	21221	22222	22212	22333	11001	000002
Know	2	10111	21212	11111	22233	00101	00000
Work	2	11112	01112	01102	32323	00102	10010
Mobi	2	22112	22212	11102	11122	21102	30?22
Infr	5	22333	22333	11112	11111	10111	00000

Table 5.4: The degree of influence on the POL stocks by different stakeholders ranging from 0 (no influence) to 3 (very high influence). In the case of the Province, the scale runs from 0 to 5.

The surprise was due to the fact that there seems to be an inverse relationship between the internal importance attributed to a stock by the administration and the degree of influence. In other words, time and resources were concentrated on such issues, where direct influence was rather low. On a side note, the individual participants judged the degree of influence higher in the sectors, to which they were affiliated in their day-to-day work. Subsequent group discussions often corrected that perception. Being explicitly aware of where ones primary influences are located strongly facilitates the search for effective policy levers. The table above serves as a rough indication as to where that influence is located. To confirm the hypothesis, one would have to measure the true levers of influence or, at least, find other proxy indicators to compare this subjective perception to.

A third enriching criterion was attributed not to the stocks, but to the flows. In a non-participatory experiment, we assigned a degree of uncertainty to each flow ranging from 1 to 3. The larger the uncertainty value, the less certain a flow was. Uncertain could either mean, that the flow was likely to appear or disappear, little was known about it or we were not quite sure whether there really was a flow or not. We based this judgement on the preceding group discussions and the uncertainties documented in the process.

We then conducted a Monte Carlo experiment, deriving 12 new matrices from the original influence matrix by letting the strength of highly uncertain flows fluctuate randomly in an interval the size of which increased with the level of uncertainty. An uncertainty value of 1 implied that the flow would always keep the same strength value, an uncertainty value of 2 allowed for a change of +/-1 and an uncertainty value of 3 allowed for the full strength range from 0 to 3 to be covered.

An analysis of the influence matrices resulting from this process delivered several flows that were crucial to the conclusions concerning the systems behaviour. One such flow was the flow between employment and population. Data from the past decades show that the relationship between vacancies on the labour market and incoming migration from other provinces to Limburg has been weak. However, at the time of the project, another Dutch Province, Zeeland, was running an advertising campaign to attract immigration with, among others, the argument of good employment opportunities. Zeeland was trying to activate a relationship between employment and immigration (it seems, without obvious success). In the case of Limburg, the uncertainty analysis shows us that a rising popularity

of Limburg as a destination of Dutch internal migration would have significant influence on a large part of the system.

The Monte Carlo experiment also functions as a sensitivity analysis of the model. In combination with the structural analysis, the experiment helps to separate those flows, for which a detailed understanding is vital for the understanding of the system behaviour from those, where changes might be interesting locally, but in terms of system behaviour, their discussion would be purely academic.

The addition of appropriate enriching criteria allows the use of the system tool as a filtering heuristic for effective and efficient policy strategies. For example, the structural analysis provides an overview of promising policy levers, which is further narrowed down to those levers, on which the province has some influence and that change within the time period envisaged for a given policy programme. It should be noted, however, that what are policy levers in the rather abstract logic of the eQSA process is only a starting point to discover what policy makers would consider policy levers. This translation from one to the other is far from evident, but the participants considered eQSA a useful heuristic to help them in this.

Other enriching criteria are possible for stocks and flows. For stocks, an indication of the current state, for flows, the manifestation (physical, information, monetary, etc. ...) would be worthwhile enrichments in the study at hand. For stocks, the addition of enrichments is relatively easily embedded in the process. For flows, this is more difficult in terms of process, analysis, and visualisation.

5.2.5 OPERATIONALISATION

The information from the previous three work packages was compiled in a handbook, an electronic database, and an Excel based project tool. The handbook and database detailed the system graph and its analysis for the system as a whole. For each stock, the QSA and eQSA interpretations, the underlying indicators monitored by the administration, the flows related to that stock, the representation of these flows in the relevant quantitative models, and excerpts from the group discussions in the different phases were included. The electronic database provided the background for the project tool. This tool supports the embedding of upcoming issues in a systems context, and facilitates the development and evaluation of strategic options. For a specific project, users detail a problem or an intervention,

distributing the different aspects over the relevant stocks. When a stock is filled in, the flows connected to that stock come up for evaluation. These can be switched off, described in more detail, or left unchanged. The different options, together with the information on each stock sourced from the electronic database, result in a rough guide to the impacts, the intervention might have from a systemic perspective, which quantitative models are relevant, as well as whether and how these might be used to gain further information. The group was trained in the use of the project tool during two work sessions.

The tool itself was initially well received, but later evaluations showed that it was not used and perceived to be too complex. However, the content of the database underlying the tool was used extensively. In practical terms, the crude interface of the tool failed to deliver added value above hard copies of the catalogues of stocks, flows, definitions, characteristics and related models. Based on the hard copies, the participants used the information in the same way that it was intended to be used within the tool. The time spent on training and experimenting with the tool might have had an effect in this, though that cannot be measured.

5.3 Other case studies

The POL project has been the most extensive application of the eQSA approach to date. In addition to the POL project, the methods have been applied to a series of other cases, sometimes purely experimental applications. A selection of these smaller applications allow testing some of the assumptions underlying the POL process, as well as illustrating the value of specific aspects of the method. These include

- the 1st generation SCENE project executed in co-operation with the Province's administration allows a comparison between two different systems on the same region.
- a regional analysis for the Region of Flanders, Belgium, commissioned by the ministry of the Flemish region allows the comparison of similar systems for different regions.

- an experimental desk study on the sustainability of the pork meat sector in the Netherlands illustrates strategic insight resulting from the MICMAC process.

Applications not included in this overview are: a study concerning the infrastructure and traffic sector in Limburg, later providing the foundation for the provinces 2007 traffic and mobility plan (Provincie Limburg, 2007b); an experimental study of the Dutch river basin management, that will be introduced in chapter 6; and an unpublished study on the sustainability of a large Dutch utilities company.

1st generation SCENE Limburg

Section 4.4 describes the development of the first SCENE model for the Province, while the majority of this chapter describes the process and analysis of the second generation SCENE / eQSA model. The group drafting the first system only partially overlapped with the group that followed the eQSA trajectory. The first influence diagram was never meant to undergo structural analysis. Nevertheless, the SCENE 1 diagram provides us with the opportunity to test for the robustness of some of the conclusions we have drawn from the eQSA process by qualitatively comparing three systems about the same region drafted at different times under different circumstances by different people. Figure 5.5 shows the active and passive potential of the stocks in the 1st generation SCENE model (in Dutch).

The most active stocks in this model are population in the right hand bottom corner and entrepreneurship (space for business in the top right corner and a dense cluster of different economic sectors to the right of the centre of the map). The transport of people scores highest in terms of passive potential. Other high scorers are energy, the transport of goods, and the space for business. We assigned each of these 50 stocks to one of the 18 stocks of the diagram drafted during the eQSA trajectory and compared their locations within the system. We found the eQSA version more balanced, more comprehensive and more structured. System patterns were difficult to detect in the SCENE 1 diagram. The intensive work with the eQSA diagram has paid off in that respect. But looking only at the passive and active potential, we found no fundamental differences. The conclusions documented in the electronic database fit the stocks of the SCENE 1 model surprisingly well. On the basis of our qualitative comparison, we were positively surprised about how similar the two systems were. We therefore have some indication for the robustness of our conclusions.

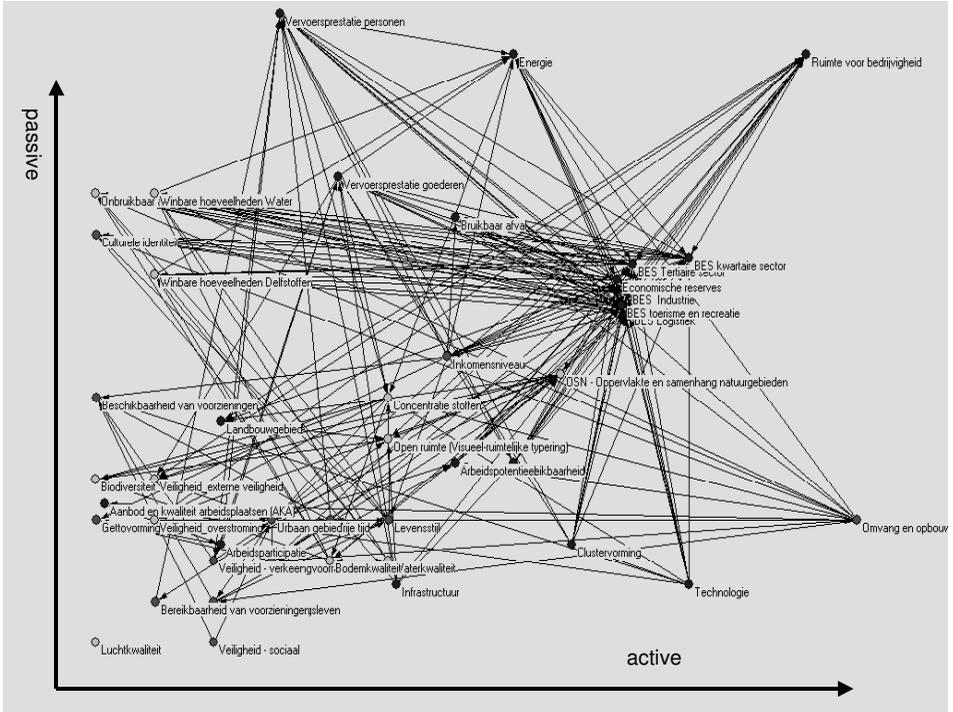


Figure 5.5: Active and passive potential of each stock in the 1st generation SCENE model.

Flanders case

We also had the opportunity to test some of the transferability of the results. The department for planning and statistics of the ministry of the Flemish region in Belgium invited ICIS for a project on regional futures. During this project, we followed a short-tracked SCENE trajectory and mapped out the Flemish region as an influence diagram. The Flanders case shows similarities and differences when different groups about similar types of regions in different countries draft a system. The participating group was much larger and broader in this case. About 50 people from all departments of the ministry participated. An active / passive potential map resulting from this process is shown in figure 5.6

The system consisted of 16 stocks: quality of life (Levenskwaliteit), lifestyle (Levensstijl), housing (Woningvoorraad), public amenities (Voorzieningen), social cohesion (Burgerzin en sociale cohesie), cultural identity (Culturele identiteit), population (Bevolking), economic structure (Economische structuur), work (Arbeid), financial strength (Financieel draagvlak), infrastructure (Infrastructuur), knowledge (Kennis), natural resources (Natuurlijke rijkdommen), soil-air-water (Bodem, lucht, water), nature (Natuur), spatial structure and landscape (Ruimtelijke structuur en landschap). 100 out of 240 potential flows were filled in.

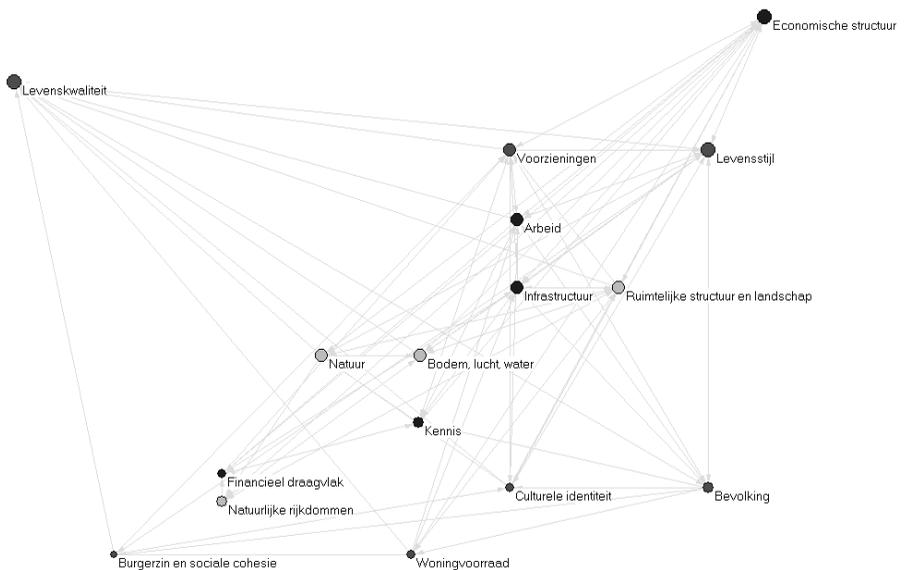


Figure 5.6: Active (horizontal axis) and passive (vertical axis) potential in the Flanders case study.

The main differences between this system and the eQSA Limburg case were found in the structuring of the stocks and aspects that needed to be emphasized. The different structure in the stocks and differences in wording were needed to fit the different organisational structure of the Flemish

ministry. Otherwise, the general results from this case study turned out to be very similar to the eQSA trajectory in Limburg.

Pork meat case

The following study of the pork meat production sector in the Netherlands was an experimental test of concept, which we executed with an internal group of researchers in order to familiarise ourselves with the sector and prepare a modelling study on the sustainability of the industry.

	rank active direct	rank change	rank active indirect			rank passive direct	rank change	rank passive indirect		
1	qt-anml	-1	catast	1		1	env-press	-4	hlth-a	1
2	catast	1	qt-anml	2		2	trsp	-8	catast	2
3	hlth-a	-1	im-anml	3		3	qt-anml		qt-anml	3
4	pop	-5	hlth-a	4		4	catast	2	nl-regul	4
5	im-anml	2	eu-regul	5		5	hlth-a	4	env-press	5
6	nl-regul		nl-regul	6		6	nl-regul	2	m-cons	6
7	ex-anml		ex-anml	7		7	m-cons	1	eu-regul	7
8	eu-regul	3	spac-anml	8		8	ex-anml		ex-anml	8
9	hlth	-3	pop	9		9	eu-regul	2	fod-qual	9
10	im-fod	-6	nl-sub	10		10	hlth	-4	trsp	10
11	inc	-6	frag	11		11	fod-qual	2	im-anml	11
12	spac-anml	4	hlth	12		12	gdp	-1	lfsty	12
13	frag	2	m-cons	13		13	im-anml	2	gdp	13
14	eu-sub		eu-sub	14		14	inc	-4	hlth	14
15	nl-sub	5	trsp	15		15	spac-anml		spac-anml	15
16	env-press	-3	im-fod	16		16	frag	-4	eu-sub	16
17	trsp	2	inc	17		17	lfsty	5	im-fod	17
18	m-cons	5	fod-qual	18		18	pop	-1	inc	18
19	fod-qual	1	env-press	19		19	im-fod	2	pop	19
20	gdp		gdp	20		20	eu-sub	4	frag	20
21	ex-fod		ex-fod	21		21	nl-sub		nl-sub	21
22	lfsty		lfsty	22		22	ex-fod		ex-fod	22

Table 5.6: Active / passive ranks before and after a MICMAC analysis in the pork case

The draft model consisted of 22 stocks: population (pop), human health (hlth), income (inc), meat consumption (m-cons), number of animals (qt-anml), animal health (hlth-a), fodder quality (fod-qual), space for animals (spac-anml), fragmentation (frag), environmental pressure (env-press), transport (trsp), import animals (im-anml), import fodder (im-fod), export animals (ex-anml), export fodder (ex-fod), european regulation (eu-regul), Dutch regulation (nl-regul), European subsidies (eu-subs), Dutch subsidies (nl-subs), gross domestic product (gdp), lifestyle (lfsty), risk of catastrophes (catast). The density of the model was unusually low with a value of 18% (462 possible flows, 84 filled). An outside expert briefly scanned the model. Nevertheless it should be emphasized, that the model is a rough draft and has not undergone the usual checks and balances. Table 5.6 shows the active and passive ranks and rank changes.

The active ranking stabilises after four iterations, passive ranking after 3 iterations. Jumps in the rankings are relatively large, as, due to the low density, many stocks have very similar active and passive potentials. At first sight, several moves are significant, for example, the drop in the passive potential of 'trsp' – transport by 8 ranks. What was most revealing in this case study at closer inspection was the gaining of the first spot in passive potential by 'hlth-a' – the health of the animals, a stock on which government regulation is quite possible (high degree of influence). A detailed analysis of the embedding of this stock in the system suggested, that one could essentially stabilise almost the whole system through this variable. Finding a single, most promising policy lever is unusual, but not impossible. On the passive side, a comprehensive regulation on the minimal health of the animals would force farmers to optimally balance all inputs, such as space for the animals, transport, quality of the fodder and the like. On the active side, risks to trade, human health and catastrophic events would be reduced, and lifestyles would be more likely to include meat consumption. This made the health of the animals the single most promising policy lever.

The pork meat case study illustrates the value of the MICMAC method as many of the factors included in our conceptual model could not be accommodated in a quantitative model, the analysis raised the right question (look for the causes of the changing role of 'hlth-a'), and a promising point of departure for further studies could be found.

5.4 Conclusions

The chapter provides an overview of different Qualitative Systems Analysis methods. We apply the newly developed eQSA approach in the main case study of this dissertation. In the following conclusions, we evaluate the approach in two ways: first, we explore its value for the process participants and their organisation and, second, we consider eQSA's contribution to the four types of integration formulated in chapter 1.

The effects of a single project within a complex institution are not easy to measure. In our evaluation we rely on three types of sources: comments from the participants during the project, a group interview with the participants in the summer of 2007, and publications by the Province since the completion of the project.

From the perspective of the participants, the eQSA approach has been an overall success. Despite this, there are a number of points that should not go unnoticed and we list these first.

eQSA is not the Swiss-Army Knife of regional governance for sustainability. According to the participants, eQSA is too complex and unwieldy to directly support change on a concrete level. A trade-off between complexity and usability is reflected in the fact that the enrichments of the eQSA were generally perceived as complicating when exploring an issue. Keeping speeds of change, degrees of influence, uncertainty and other enrichments in mind requires extraordinary discipline and costs time. This was not always practical in the contexts in which eQSA had been applied in the administration. Should an extra lane be added to a highway? Should a chemical plant be subjected to additional security regulations? How should cultural subsidies best be distributed? Especially with respect to the flows, eQSA is too abstract to support decisions on such questions. (In chapter 7, we will see how eQSA can support in concrete decisions through its ability to improve the execution and interpretation of quantitative prognoses.)

One of the participants signalled the problem of the discrepancy between noticing something and thoroughly comprehending it. eQSA supports the noticing of side-effects, policy levers and strategies. But to thoroughly comprehend what has been noticed, other catalysts outside eQSA are needed, for example a public discussion.

eQSA is best used in discussions on large, abstract and persistent issues. The transition from growth to reduction in Limburg was given as an example. The region's population is ageing rapidly, as is the labour force. Governing these developments and their side effects requires new ways of thinking about development. With eQSA, the participants felt, they could structure and advance such a discussion and identify innovative strategies. Doing so indirectly influences the more concrete choices by providing a future framework against which specific actions can be evaluated.

The eQSA approach entrenched systems thinking in the work of a strategic group at the province of Limburg. The influence diagram combined well with the organisational structure, the different elements of the system provided links to other reporting and strategic planning tools and tasks within the organisation. Therefore, the participants have become ambassadors for the integrated approaches within their organisation. Three participants agreed that: *'When people ask me, what sustainability is about, I show them the SCENE/eQSA triangle.'* This has led to the spreading of integrated work. One of the outcomes of this spreading has been that the updated SCENE framework including the terminology was used to structure the Limburgmonitor, the key publication on the province's monitoring data (Provincie Limburg, 2007a).

In terms of the four types of integration we are aiming for, eQSA provides an amalgamating tool to bridge different scientific disciplines and their respective tools. The results support integration in policy processes and policy domains. In terms of normative integration, the eQSA tool is rooted in a participatory process. However, normative integration was limited by the fact that the only stakeholders actively participating were the employees of the province administration. Methodologically, it is easily conceivable to include a broader range of stakeholders in the process to strengthen this dimension, given the time and resources. The overview of the influence of different stakeholders in table 5.4 provides a glimpse of this.

With respect to the three research priorities formulated in chapter 3, the focus of the analysis lies on dynamic leverage points and on integrating static leverage points. The questions raised when working with the results of the eQSA process trigger insights on contextual leverage points. This is due to the transparent and visual way in which the results are processed and communicated. However, complexity and level of abstraction are limiting factors in this respect. Nevertheless, the participants can trace back every single conclusion to its initial assumptions and question these qualitatively.

In chapter 7, we explain the consequences of the eQSA case in Limburg for the use and interpretation of quantitative forecasting models. In chapter 8, we describe the organisational impact in more detail.

Chapter 6

QSSI: A NEW TYPE OF
SUSTAINABILITY INDICATOR

Long and detailed lists of sustainability indicators are good indicators for a lack of integration. (Bas Amelung (2006))

Sustainability indicators are a popular tool in sustainability assessments. They are easier to implement than quantitative models, suggest more objectivity than participatory methods, and their communication seems comparatively straightforward. However, sustainability and indicators are both complex concepts in their own rights. Combining the two in one tool inevitably creates room for problems. Hak et al. (2007) summarize these: Conceptual problems, for example, are caused by the normative characteristic of sustainability that is confronted with the objectifying ambition of indicators. Methodological problems include issues of legitimacy, stability and soundness of calculation methods. Examples of application problems are dealing with different scale levels and translating indicators to policy targets. Importantly, in the context of this dissertation, existing indicators and indices generally do not take the interconnectedness among different aspects of sustainability into account. Yet, these relationships are critical for a meaningful interpretation of any information resulting from indicator based assessments, especially with respect to the dynamics of a system. In this chapter, we resort to the other extreme, by proposing an indicator exclusively based on the relationships.

Section one defines and reviews existing sustainability indicators. Section two lists some challenges related to the ambition of a four fold integration: integration among policy domains or 'joined-up thinking'; integration among scientific disciplines; integration of relevant research in policy processes; and normative integration. The Qualitative Systems Sustainability Indicator QSSI is introduced in section three and proposed as an alternative type of indicator better suited for assessing the sustainability of systems. Section four contains a concluding synthesis of the overview and the QSSI.⁸

⁸ This chapter is based on Grosskurth & Rotmans (2007). The author of this dissertation has developed the idea and method. Subsequent refinements were done with the help of Jan Rotmans. The case of river basin management was conducted in collaboration with Rutger van der Brugge.

6.1 Sustainability indicators

This section of the chapter defines the concept of indicators, relates this definition to the concept of sustainability and introduces existing sustainability indicators. Sustainability is complex, ambiguous, normative, and subjective (see chapter 2). The characteristics of sustainability bring with them special requirements when it comes to sustainability indicators. *In its most general sense, an indicator is a sign* (Bossel, 1996). These signs are generally represented as numbers, but other representations are possible, such as symbols, graphics, or colours (Moldan & Dahl, 2007). Independent of its symbolic representation, an indicator *is a one-dimensional systems description, which may consist of a single variable or a set of variables* (Rotmans, 1997). Indicators summarise a property of a complex system. The summary facilitates communication at the price of a decrease in the depth of what is being communicated.

The systemic properties of sustainability have repercussions for the development of sustainability indicators. Dahl (2007) challenges the scientific community to develop sustainability indicators that *measure those characteristics or processes of the system that ensure its continuity and functionality far into the future*". According to Dahl, these indicators should *reflect the whole and not just the parts*. Indicators should highlight problems rather than symptoms. An indicator of sustainability must therefore also be based on a systemic perspective. This is widely accepted, yet has had little resonance in the development of sustainability indicators. An indicator should signal, whether the interplay of different aspects of sustainability supports or compromises the ability of the system to sustain itself in the long run. One implication of the time aspect is that future system states (similar to future images) are relevant. Another implication is that sustainability cannot be a static state, but is subject to dynamics in time (similar to movies showing the path to future images).

Beyond these comparatively technical signals, a sustainability indicator should also signal aspects of the normative and subjective characteristics of sustainability. This further delineates the concept of sustaining referring to a desired state or trajectory. We choose the word 'desired' in order to make the subjectivity of anything except pure physical sustainability explicit. The fact that an Earth without a human population would not be sustainable is due to the anthropocentric subjectivity of the concept. Even though the system without humans would quite probably be perfectly able to sustain

itself in the long run, it would be on an undesired trajectory. Any indicator based on the system structure must use a representation of the real world in a system format. The representation and evaluation of a system is intrinsically subjective, normative, ambiguous, uncertain and incomplete (Rotmans & de Vries, 1997). In response, several authors have proposed public participation (Spangenberg & Bonniot, 1998) or stakeholder participation (Jaeger et al., 2001). For our work this implies the challenge to develop methods or interfaces for integrated sustainability assessments that are suitable for stakeholder participation.

Public participation should not be confused with broad public application. In other words, there is a difference between something that is developed in a process of collaboration and something that is developed without collaborative processes, but used by many. The Ecological Footprint (Wackernagel & Rees, 1996) has been widely applied in different forms. Many in the general public have calculated their personal Ecological Footprint. Schools, communities, regions and countries have done the same.⁹ But we are not aware that public participation has played a role in the continuous development of the index beyond the general feedback that is part of many scientific activities with a target group outside the scientific community.

More and more sustainability indicator studies now include elements of public participation. These are often executed within the context of local Agenda 21 initiatives. Scientists or consultants develop lists of indicators based on the input and reflections from stakeholder groups. This sometimes feeds into modelling exercises, where participants interact with a simulated system and thus provide input for defining sustainability for a given region or issue, selecting indicators and communicating results. SDRI Quest is an example for such an approach.

In addition to participation, the quality of a sustainability assessment also depends on the quality of the system description. Several frameworks have been developed in order to facilitate the quality of a system description with

⁹ I am not aware of a comprehensive inventory of Ecological Footprint studies. However, a Google search on 'ecological footprint, calculator' delivers 45,700 hits. 'Ecological footprint, school' delivers 233,000 hits and 'ecological footprint, community' 362,000.

reference to underlying definitions of sustainability. We mention these frameworks at this point as each framework implicitly or explicitly relates to a definition of sustainability and significantly influences the structure of an indicator. Examples for such frameworks are the PSR (Pressure-State-Response) framework applied during the development of the OECD indicator list (OECD, 1993), the sustainomics triangle applied by the UN-DPCSD (1999) in their approach, the 'system conditions' used by The Natural Step (K-H. Robèrt & Anderson, 2002) and the orientor framework developed by Bossel (1996). These frameworks are also tools to help choose what issues should be included in the analysis of the system and in what structure they should be represented. While a PSR approach emphasizes the cause-and-effect structure of issues, the sustainomics triangle emphasizes the balance of social, economic and environmental issues. In our work, we have used elements of PSR and sustainomics to develop the SCENE framework introduced in chapter 4.

In general, existing indicators for sustainability come in two distinct forms: lists of indicators and indices (aggregated indicators). Lists of indicators consist of a series of indicators that are not explicitly put in relation to each other. In indices the fragmented information of lists of indicators is aggregated in order to represent the essence of a list of indicators. A series of more or less comprehensive lists of indicators of sustainability have been developed at different geographical scale levels. Examples for frequently cited lists of regional sustainability indicators are the OECD Environmental Indicators (OECD, 1993), the Sustainability Institute's Indicators and Information Systems for Sustainable Development (D. L. Meadows, 1998), and the UN-DPCSD Indicators of Sustainable Development (UN-DPCSD, 1999). On a national level, the most widespread list of sustainability indicators is currently prepared by the Netherlands Environmental Assessment Agency (MNP, 2006). On a regional level a list including participatory elements in its development is used by the province of Brabant (Telos, 2006). The Province of Limburg applies its broader monitoring system to derive sustainability indicators (Provincie Limburg, 2007a).

Indicators for a given list are selected on the basis of content relevance and data availability. The content relevance is in some cases derived from the frameworks that relate to the underlying definition of sustainability as described in the previous section. From an integrated assessment perspective, these lists of indicators were a first step to include different disciplines in sustainability assessments. Instead of looking at the parts

independently of each other, we are now looking at the parts simultaneously. This makes it possible to conceptually recognize some emergent patterns. Lists of indicators primarily monitor the constituents of sustainability by assigning indicators to these constituents. The awareness that this does not capture some of the relevant dynamics has resulted in the monitoring of determinants as well.

Indicators are sometimes aggregated to form indices. In indices, different indicators are assigned a weight and their developments therefore suggest a direction of the overall sustainability. However, the process of weighing and aggregation is widely accepted to involve several problems that are not yet satisfactorily solved. In order to aggregate different indicators, they must be assigned a common denominator. In existing studies this denominator is money (ISEW, GPI), a dimensionless index between 0 and 1 (HDI) or physical land-use (Ecological Footprint). The selection of a denominator influences the set of indicators that can be part of the system. It is for example a direct result of the denominator of physical land-use that the Ecological Footprint has to exclude social aspects from its analysis. Generally the weights carry a connotation of importance. A higher weight implies a higher importance with respect to sustainability. Beyond that, weights are fine tuned to make sure that the resulting index is neither too volatile nor too rigid. A volatile index changes strongly with relatively small changes in underlying indicators and therefore exaggerates developments of less importance. A rigid index has the tendency not to change significantly even with big changes in the underlying indicators. Important developments are therefore not detected by the index. It is common that an index is rigid with respect to some indicators and volatile with respect to others. Which indicators should be dominant in an index should be determined by the underlying system. The most dominant indicators should be those that provide a reliable early warning of arising problems. But that perspective is not widely applied.

The advantage of indices above indicators lies in the simplification of otherwise more complex information. A good index highlights the essential information from a list of indicators. This goes together with some loss of information if the underlying indicators are not communicated. Examples for indices of sustainability are the Index of Sustainable Economic Welfare (ISEW - Daly & Cobb, 1994), the closely related Genuine Progress Indicator (GPI - Redefining Progress, 1995), the Human Development Index (HDI - UNDP, 2003) and the Ecological Footprint (Wackernagel & Rees, 1996). From

an Integrated Assessment perspective, these indices are a step forward with respect to the communication of complex information. By communicating a single piece of information the essence of a complex phenomenon is understood.

Lists of indicators and indices aggregated from these lists are therefore excellent tools for the communication of the urgency of sustainability. The sign sent through a quickly deteriorating indicator, such as biodiversity or GDP is clear and well understood. The message is: be aware and act! An Ecological Footprint larger than 1, a decreasing ISEW or a low ranking HDI communicate the same well-understood message. Simulation models can support this message by showing not only past and current developments of indicators and indices, but also projecting these developments into the future. This projection often reveals that one deteriorating indicator is followed by others with a delay in time.

However, with respect to the general direction of developments, lists of indicators add little information. The fact that some indicators from a list move into a more sustainable direction and others move into an unsustainable direction does not convey any information on the development of the overall degree of sustainability. Only if all indicators simultaneously move into an either desired or undesired direction can we make a statement on the direction without considering the weights of different indicators. What is noteworthy and disturbing is that none of these classes of indicators are based on the system interrelations. The primary information of all these indicators is the past, current and future state of the elements of the system rather than the system structure. In other words, we receive information concerning symptoms rather than problems.

It is in the underlying system that the unsustainability with its causes and possible solutions can be found. It has, for example, been widely noted that the economy has (within certain limits) a detrimental effect on the environment (for example: Commoner (1972); Chertow (2001); IPCC (2001)). In a list of indicators we would see that the indicator economy goes up and the indicator environment goes down. An index might not change at all as the two movements might weigh each other off. The message taken from the indicator is to act in order to turn around the development of the indicator environment. Even though the trade-off between the economy and the environment is in this case the major sustainability problem, it takes an analysis of the context of an indicator to understand this. This context does not play an explicit role in lists of indicators and indices of sustainability.

Any view that does not take into account this system can only identify symptoms of sustainability, but never find solutions without more extensive context analysis. In this respect it is important to notice that all three types of indicators are based on the past, current and possibly future states of variables and do in their calculations not take into account the way these variables are embedded in their neighbourhood. As a consequence, policy action, for example, is more likely undertaken not on the basis of a systemic understanding, but on the basis of scattered symptoms of un-sustainability, such as air-pollution, crime rates or government debt, which can lead to inefficient solutions and counter-intuitive surprises. (Of course, sustainability indicators and the attempted rationality and epistemology embedded in them, are at this stage not prime drivers of policy action.)

This logic is comparable to that of industrial agriculture. If a pest occurs (a symptom of un-sustainability according to the concept of sustainability prevalent in that sector), one should act and remove the pest and apply pesticides. Other symptoms will appear, the same symptoms will reappear, often stronger than before, and all will be handled in the same way. The underlying problem, that causes the symptom to appear is neither identified nor solved, as this would require a systemic understanding of the relationships between the different indicators.

When we provided advice on regional sustainability to civil servants in strategic departments of sub-national governance bodies, we found that they were quite capable of identifying economic, environmental and social states and trends in their region, such as a low education profile in the population or a rapid loss of biodiversity. The target audience could not map these fragmented parts as a system of stocks and flows. Consequently, they had difficulties to understand the complex dynamics resulting from these interactions. The target audience was even less able to design and test systemic policy actions for their effects on the sustainability of their region. For our work, this meant that a selected list of sustainability indicators or an aggregated indicator of sustainability would not address the most urgent questions and would not help to make better decisions for a sustainable future. Instead of focusing on the states of the parts of the system, we needed to focus on the system structure itself.

Quantifiability is still widely considered a defining criterion for an indicator. Only information, that can be expressed and processed in quantitative format can be included. Issues for which such information is not available because of a lack of possible indicators, a lack of measurements or

a lack of measurability are excluded from the analysis. There is a lot to be said for quantitative indicators. Well-applied quantitative methods can in many cases help to achieve objectivity, accuracy, comparability and reproducibility, all characteristics that are desirable in indicators.

However, the exclusion of qualitative information is independent of the content relevance of the excluded issue. In other words, we are looking exclusively where the light is, not necessarily where the problem lies. This exclusion also causes an imbalance between the three dimensions of sustainability. Traditionally, economic information is conveyed in quantitative form. This partly holds for ecological information as well. Information from the social and institutional dimension frequently takes a qualitative form, especially when reputedly 'soft' issues like social cohesion or cultural identity are analysed. In addition to that, indicators for social issues change rather quickly from suitable to unsuitable. If for example, the memberships in political parties and or the number of churchgoers were reasonable indicators for public participation, these are now replaced by memberships in NGOs or volunteer work.

In our description of the SCENE framework we have argued that even though quantitative information is in some cases useful, it is helpful to postpone quantification of systems for as long as possible in order not to reduce the available information on the wrong assumptions.

Where available, quantitative indicators proved useful to communicate the urgency of key issues in participatory settings. However, the exact same information caused confusion and frustration when the goal of the information was extended beyond raising awareness to include the development of policy options for improvement. The reason for this was the lack of information on the systemic causes of indicator values and thus a lack of information on what one could do to change the value of a given indicator without negatively affecting other indicators. The availability of quantitative information caused a neglect of issues for which quantitative indicators were not available. In response to this we found it necessary to develop methods to better integrate qualitative and quantitative information.

6.2 QSSI: Proposing a System Indicator

With existing sustainability indicators and the challenges in mind it is possible to develop a new type of indicator. If we assume sustainability to be a system state, then two aspects would characterise that state. First, the values of the stock variables should not show unsustainable levels. Secondly, the associated flows should not drive the stock variables into an unsustainable direction. In other words the system structure should be such that stock variables are not driven towards unsustainable values.

To analyse the first aspect, lists of indicators and indices are useful tools. As far as the second aspect is concerned, only tedious data mining will reveal this information, if at all. We therefore base our indicator on the same type of system diagrams described in the previous chapters. If sustainability is interpreted as the balanced long-term development of the three domains of sustainability, then the development of one part of the system towards a desirable state should not structurally go at the cost of developments elsewhere in the system, as this would compromise its continuity and functionality. In this section we propose the QSSI as an indicator for the degree to which the system structure causes such compromises.

We illustrate the QSSI on the basis of a conceptual model that we drafted for a SCENE case study on the transition in Dutch river basin management between 1970 and 2000 (van der Brugge et al., 2005). The model in Figure 6.1 describes the system for the year 1970. We chose this example, as it is the smallest model we have drafted in terms of stock (10) and flow (23) numbers. The system properties in terms of connectedness, the relative number of feedback cycles and cluster formation are comparable with other systems that we have drafted. We therefore consider this model to be representative for illustration purposes.

In order to introduce the reader to the causal thinking represented in this diagram we will follow some of the causal chains. For example, the size of the retention capacity influences the risk of floods. The risk of floods is an important factor when the number and size of dams to be built are determined. These dams take up land that could be used for agriculture and buildings if there were no dams. Buildings and agriculture produce economic benefits that reduce (opportunity) costs. Thus the flows connect stocks that are directly related. 'Direct' meaning that there is no other stock

within the system via which the flow described would take effect. In other words, the arrows only present first order flows.

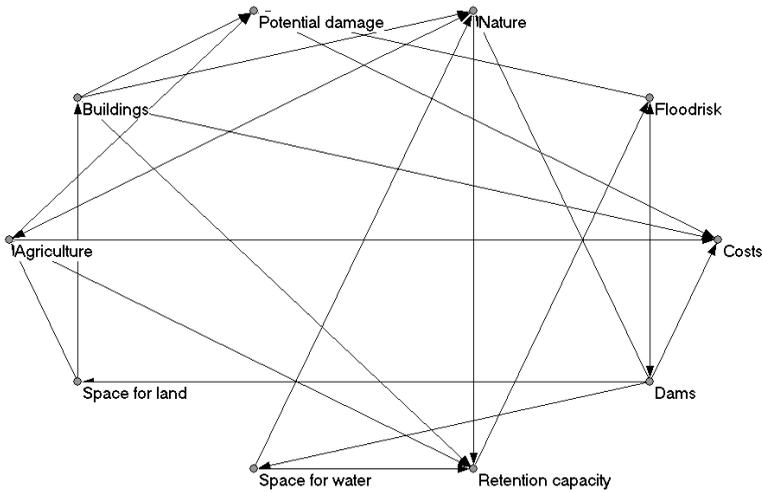


Figure 6.1: Influence diagram for a river water management system

In order to better represent the system, the flows can be described in terms of positive and negative flows. As in the notation of system dynamics positive flows stand for relationships with similar direction of the originating and receiving variables (an increase in flood risk stimulates the building of Dams, a decrease of flood risk has the opposite effect), negative values stand for relationships with opposite direction of the originating and receiving variables (a decrease in the number of dams increases the risk of floods, an increase in the number of dams reduces the flood risk). As this description gives rise to different possible interpretations, the background material of each case study contains extensive documentation on the arguments and discussions raised during the drafting process of the system, the interpretations considered and the relevant scientific literature.

If we want to evaluate ‘the sustainability’ of the system described in our model, we need to define sustainability by making the inherently normative choice of what is desirable. We proposed to let these judgments be formulated in a stakeholder process. This way the normativity is made explicit rather than being buried implicitly in the framework underlying an indicator or index.

For the purpose of the QSSI this choice is made for every stock separately, independently of the effects of that stock on other stocks. The choice to attribute a low desired state to agriculture can therefore not be based on the argument that agriculture reduces retention capacity or has a detrimental impact on nature. Due to this criterion of independence of other stocks, the discussion will not be concerned with the choices that should be made between mutually exclusive stocks. A ‘+’ indicates that a high value for that stock is desired, a ‘-’ indicates that a low level is desired. For example, a low flood risk and low costs are desired, but a high retention capacity and much agricultural activity are desired. This information is summarized in Figure 6.2.

The matrix is read from left to right. Each cell in the matrix stands for a potential flow from the stock in the row towards stock in the column of that cell. A flow from Flood risk to Dams (if there is a threat of floods, more dams are being built) can be found in the third cell of the first row and is represented by the value 1. The reverse flow from Dams to Flood risk (if more dams are built, the flood risk decreases) is represented by the value -1 in the first cell of the third row. As in the notation of system dynamics positive flows (1) reinforce the original signal and negative flows (-1) dampen the original signal.

In order to evaluate the long-term continuity and functionality of the system we tested the consistency of the desired directions and flows connecting the stocks. The consistency check for a flow is passed if the value of the flow (‘+1’ or ‘-1’) is equal to the desired direction of the originating stock multiplied by the desired direction of the receiving stock of that flow. For example: -1 (Flood risk) * -1 (Dams) = 1 (the value of the flow from flood risk to dams found in the third cell of the first row. If these signs are not consistent, then the desired direction of a stock is not consistent with the flow driving it. For example: -1 (Dams) * -1 (Flood risk) \neq -1 (the value of the flow from dams to flood risk).

		Desired direction									
		-	-	-	+	+	+	+	+	-	+
Desired direction		Flood risk	Costs	Dams	Retention capacity	Space for water	Space for land	Agriculture	Buildings	Potential damage	Nature
	-	Flood risk			1						1
-	Costs										
-	Dams	-1	1			-1	1				-1
+	Retention capacity	-1									
+	Space for water				1						1
+	Space for land							1	1		
+	Agriculture		-1		-1					1	-1
+	Buildings		-1		-1					1	-1
-	Potential damage		1								
+	Nature				1			-1			

Figure 6.2: The water management influence diagram and desired states.

An inconsistency implies that if a desired development for one stock is realized, another stock comes under some pressure to develop into an undesired direction. Thus we can pinpoint those system elements the development of which undermines the system elsewhere. The more of these inconsistencies we encounter, the more difficult it will be to steer the system onto what is normatively chosen to be a sustainable trajectory. In our example, 9 out of 23 flows are inconsistent with our sustainability goals. The inconsistencies are highlighted in grey in figure 6.3.

These inconsistencies are not inconsistencies in the policy interventions. They are structural patterns within the system that make a consistent policy strategy for sustainability impossible. Weakening or removing these inconsistencies is a necessary, though not sufficient, condition for an integrated sustainability strategy. Each inconsistency in itself seems trivial: if more room is provided for nature, this goes at the cost of agriculture. The building of more dams goes at the cost of land for other purposes. The construction of buildings within the rivers flood bed increases the potential

damage. Summed up, they determine the difficulties for a sustainable management of the river basin.

		Desired direction										
		-	-	-	+	+	+	+	+	-	+	
Desired direction		Flood risk	Costs	Dams	Retention capacity	Space for water	Space for land	Agriculture	Buildings	Potential damage	Nature	
	-	Flood risk			1						1	
	-	Costs										
	-	Dams	-1	1			-1	1				-1
	+	Retention capacity	-1									
	+	Space for water				1						1
	+	Space for land							1	1		
	+	Agriculture		-1		-1					1	-1
	+	Buildings		-1		-1					1	-1
	-	Potential damage		1								
	+	Nature				1			-1			

Figure 6.3: Inconsistencies between system relationships and desired developments.

6.3 Calculating QSSI

The simplest version of a sustainability indicator related to the model in the previous section is to divide the number of inconsistencies by the total number of flows, resulting in an index between 0 and 1, in our example $9/23= 0.39$. A lower number implies a higher level of sustainability and vice versa. We call this the Qualitative System Sustainability Index (QSSI). This indicator is independent of the current state of the stocks or related indicators. Similar to existing sustainability indicators, the QSSI consists of two layers of information: an index, i.e. a single number that summarizes the information contained in the underlying system, and the body of information that in this case consists of a model in matrix form and in more

common approaches of lists of indicators. In both cases, a sound interpretation of the index is impossible without an insight into the underlying body of information. Following the logic of the QSSI the number of inconsistencies in the system can initially be decreased in three ways.

- We can aim to make an inconsistent flow disappear (decoupling of stocks). In our example there is an inconsistency as buildings take up retention capacity. Both are desired and if there was a possibility to build houses above the potential flood water level (e.g., on poles), then the negative flow between the two stocks would disappear. This would result in a QSSI of $8/22=0.36$.
- The QSSI is reduced even more if we actually manage to turn the sign of the flow. If we would be able to stimulate such agriculture that can be applied in retention areas, and actually profits from occasional flooding of these areas, then the QSSI in our example would decrease from 0.39 to $8/23=0.35$.
- A third way to decrease the QSSI is to add consistent flows. If we can find a way to give space to water in such a way, that agricultural activity is stimulated at no cost for nature development (admittedly a far fetched example), then we would add a consistent flow. The QSSI would then decrease from 0.39 to $9/24=0.37$.

These interventions for increasing the ability of the system to sustain itself on a desired trajectory in the long run are all concerned with the structure of the system. By changing the system structure (adding, removing or changing flows) in a desired manner we make the system more sustainable. Considering the enormous resources needed to actually change existing flows in the real world, the QSSI is only likely to change very slowly over time. This does right to the fact that the indicator is concerned with the slowly changing system structure and its direction, rather than short-term symptoms in the form of flows.

But not all inconsistencies can be solved this way. At some point no further system improvements are possible. We will have to make choices between different sets of inconsistencies. In our example one of these choices could be to weigh off the importance of agriculture and the economic benefits of it against the importance of retention capacity, the potential damage done to the agriculture by floods and nature (all the stocks, that agriculture is influencing). It is at this point that the desired direction of a stock dependent

on the desired direction of other stocks is discussed. In the Dutch case, policy-makers have made a very explicit choice and decided that agriculture should recede in order to provide room for nature (RIVM, 1997).

At the cost of the foregone economic gain from agriculture we resolved our conflicts with the retention capacity, the potential agricultural damage caused by floods and nature. We thus add one inconsistency and remove three others. This results in a QSSI of $7/23=0.30$. The QSSI is reduced through a change in our goals. This choice is not a choice to be taken in a sustainability assessment, but one that should be delegated to democratic processes.

The ultimate consequence of this is that we can achieve a higher degree of sustainability if we are willing to give up some of our goals and thus redefine what is desired. Traditional lists of indicators can help to make an informed choice concerning desired states and the relative urgency of different inconsistencies.

6.4 QSSI extensions

We extended the QSSI to accommodate flows of different strength by assigning a value to the flows ranging from 1=weak to 3=strong. The weights of the flows can be deduced from stakeholder participation, expert judgments and empirical evidence. Figure 6.4 illustrates this.

With weights assigned, the QSSI is calculated as the total sum of the weights of the inconsistencies divided by the total sum of weights of all flows. In our example, the QSSI is $18/47=0.38$. Solving a strong inconsistent flow has a larger impact on the QSSI than removing a weak flow. The QSSI is 0.33 when a strong inconsistency is removed and 0.36 when a weak inconsistency is removed. With weights added, possible interventions can also be aimed at the strength of flows, giving room for issues like eco-efficiency and technological progress that push the frontiers of sustainability without fundamentally changing the system.

In addition, weighted flows and inconsistencies provide arguments to reject inconsistencies based on their significance. If a stock is strongly consistent with one and weakly inconsistent with another stock, then the weak inconsistency is of low priority when it comes to action for sustainability

unless the inconsistency itself is highly undesirable. The weighted QSSI implicitly takes this into account. It requires a high consistency (thus a high divisor) to compensate for a weak inconsistency, resulting in a low QSSI value. With the addition of a set of standard stocks and a standard procedure to define flows for certain classes of systems (e.g., industrial sectors or regions) the QSSI could actually be extended in such a way that it allows for comparisons between different systems. However, developing such a set of standard components and ensuring its broad acceptance raises many problems. A process to do so for the Netherlands alone would require the extensive co-operation of a broad range of societal actors. In the end questions of legitimacy, neutrality and acceptance would still be the subject of dispute.

		Desired direction									
		-	-	-	+	+	+	+	+	-	+
Desired direction		Flood risk	Costs	Dams	Retention capacity	Space for water	Space for land	Agriculture	Buildings	Potential damage	Nature
	-	Flood risk			1						3
-	Costs										
-	Dams	-3	3			-3	2				-1
+	Retention capacity	-1									
+	Space for water				2						2
+	Space for land							2	2		
+	Agriculture		-1		-1					1	-2
+	Buildings		-2		-3					2	-2
-	Potential damage		2								
+	Nature				2			-2			

Figure 6.4: The QSSI matrix with weighted flows.

Also, with methods of general network analysis (for example eQSA) it is possible to derive more system information, such as the character of stocks

(active or reactive) and process additional information, such as the speed of change of stocks or the influence of different stakeholders. Any type of relevant information that can be translated into a range (e.g., from low to high), a vector of ranges or into an ordinary scale can theoretically be processed. This leaves plenty of room for refinements of QSSI. However, applying these extensions also exacerbates the disadvantages of QSSI, for example in terms of complexity and communication.

QSSI is based on the SCENE approach. However, SCENE is only one aspect of many methods for the qualitative understanding of systems and thus integrate quantitative and qualitative information (see chapter 4). Any other approach that comprehensively maps system components and the relationships between them can be used as a base. Therefore, the concept is not restricted to the three pillars of sustainability, but can be adjusted to accommodate additional domains (e.g., Institutional) or entirely different structures (e.g., the DPSIR approach).

The quality of the QSSI hinges on the quality the system description. As with any other type of model, the principle of garbage in = garbage out holds also in our case. In contrast to empirical work, the potential quality of qualitative system descriptions is not widely accepted. However, it can be said that the quality of such systems is not necessarily lower than the quality of the structures of quantitative models (and at least any system dynamic model can be translated directly into a qualitative influence matrix with comparatively little loss of information). We will come back to this point in detail in chapter 7 where we compare the SCENE system developed for Limburg with the structures of the quantitative models used for the same region.

The top information layer of our qualitative influence matrices is actually concerned with the system structure and not with the state of the variables as is the case with system dynamic models. In addition, the system is formulated in such a way that any interested layman can understand and contest it. The structure of the system is therefore open to discussion. How the results of such a discussion could feed back into the QSSI is technically straightforward, but the problems associated with participatory methods in general are particularly difficult to circumvent in such a process. Especially the overlap of the scientific process and the democratic process raises methodological and conceptual problems that are hard to resolve.

Several versions of the system can be developed in parallel. These different representations of the same system make explicit differences of opinion and some of the underlying uncertainties. Nevertheless it must be said, that much research still rests to be done in order to determine the suitability of qualitative systems for rigid sustainability studies. These studies would have to give extra attention to issues including guidelines for formulating and testing suitable system descriptions, issues of scale, robustness of results and normativity versus objectivity.

6.5 Conclusion

We propose the QSSI as a first example of an indicator that is not primarily based on the measurement of flows. The QSSI combines methods from soft and hard systems thinking. The QSSI roots in the SCENE (chapter 4) and eQSA (chapter 5) approaches.

The question is whether the QSSI is a useful addition to the existing set of indicators. The QSSI emphasizes the system properties and is able to process more qualitative information than other sustainability indicators. While from the perspective of sustainability its dependence on stakeholder participation can be seen as an advantage, from the perspective of indicators, especially the objectivity and consistency of indicators, it is a disadvantage.

QSSI integrates the three domains of sustainability through its roots in the SCENE approach. In decision-making, the strength of the QSSI lies in its indication of interventions for more than cosmetic sustainability. QSSI will only register changes, which fundamentally change the ability of the system to sustain itself in the long run. The models are formulated in such a way that an interested layman can understand and contest them.

The structure of the system is therefore open to discussion and several versions of the system can be developed in parallel. These different representations of the same system make differences of opinion and some of the underlying uncertainties explicit. Most importantly, the resulting priorities for sustainability action are easily understood and communicated and quite different from priorities derived from traditional indicator sets. Where a traditional set might point to a loss in biodiversity, the QSSI might point to the trade-off between the building of industrial areas and reserving

space for nature (which in turn supports biodiversity as well). In other words, the QSSI points out the problem rather than the symptom.

These advantages come at a cost. A major drawback of the QSSI is its lack of comparability. The QSSI is highly context dependent. The way in which we choose system boundaries and system representations are key factors influencing the way in which we see and interpret that context. In addition, the process of deriving the QSSI is costly and requires the commitment of stakeholders since the normative judgements of the stakeholders are a crucial element of deriving the QSSI. If that commitment is present, the issue of communication is lightened, but the QSSI itself remains difficult to communicate to a broad audience.

The question is not whether the QSSI is suitable to replace the existing classes of indicators, but rather if it is a useful addition. In general, the QSSI scores high on aspects on which the existing classes of indicators score relatively low and vice versa. In our approach, we have chosen to address the systemic challenges in indicators at almost all costs. These costs are high, as the QSSI does not communicate the urgency of sustainability. More importantly, the QSSI cannot be compared between different systems or scales. It remains a value for a single system that can be either increased or decreased until a standard set of stocks is developed. Unfortunately, the challenges in developing such a standard set of stocks and ensuring its legitimacy and acceptance for regional applications are enormous.

By developing the QSSI indicator we have shifted our attention from the development of quantitative variables over time with little knowledge of the underlying system and even less room for change in that system towards a system approach that deserves the name. Doing so it has become apparent that the value of developing this indicator with participatory methods is at least as important as the resulting value. Considering the enormous resources needed to actually change existing flows in the real world, the QSSI changes very slowly over time. One might be tempted to judge that therefore QSSI is rather insensitive to changes, but the slow change in the indicator reflects that the indicator is concerned with sustainability itself, rather than its symptoms. And achieving fundamental sustainability takes more time and more radical changes, than changing symptoms does.

Currently, the QSSI is a little tested concept for a new type of sustainability indicators. Much research is still needed in order to develop a sound, solid and robust sustainability indicator. Future research questions must include

its relation to existing indicators, the normativity of the index, its results' robustness, its susceptibility to different perspectives, as well as questions of scale and system boundaries, the role of stakeholders, the suitability of different frameworks to develop systems from which the QSSI can be calculated and the potential use in the policy cycle. While this list might seem overwhelming, the majority of these issues are not specific to the QSSI, but need to be researched in the broader context of sustainability as a scientific concept. Nevertheless, we are confident that the road we have taken will stimulate a fruitful discussion on alternative types of sustainability indicators and that we are progressing towards an indicator that resolves the challenge of reflecting the whole instead of the parts.

Chapter 7

SIMULATION MODELS

One restriction entails another: We attain objectivity, but we fail to attain knowledge of the *object as a whole*. (E.F. Schumacher (1977))

Quantitative models are among the most widespread methods in the Integrated Assessment toolbox. They are applied as stand-alone products, or as building blocks of systems approaches integrating different kinds of assessment tools (Rotmans (1998b); Rotmans & Dowlatabadi (1998); Robinson (2004); de Ridder (2006)). This chapter takes a two-tier approach to finding out how to improve the capacity of quantitative models to support joined-up thinking in decision making.

Section one provides a broad overview of integrated simulation models using well tested modelling techniques. All are either milestones in the development of such models and/or are still in active use. The models are not only described with respect to their methodological characteristics, but also with respect to the underlying ideas. The descriptions are based on an active scan of the literature and the input of experts in scenario modelling. A direct comparison of the different approaches that are used for different purposes is done. This provides valuable insights into relevant issues, such as possibilities and problems with further inter-disciplinary integration, ways to model cause-effect chains and time-horizons as well as more general insights into the policy relevance of simulation models.¹⁰

The lessons learned in this inventory provide the structure for a detailed analysis of the quantitative models used by the Province of Limburg to support strategic planning, presented in section two. We explore the coverage of the available models by comparing it with the SCENE diagram developed in chapters 4 and 5. A structural analysis reveals the misrepresentation of dynamics in the quantitative models.¹¹ Section three

¹⁰ A previous version of this overview has been published in Greeuw et al. (2000) and Bakkes et al. (2000). The author of this dissertation was responsible for the initial overview and evaluation of models (chapter 4 in Greeuw et al.). The chapter significantly benefitted from comments from Marjolein van Asselt, Dale Rothman and Jan Rotmans.

¹¹ A less detailed version of this analysis has been published in Grosskurth (2007).

provides policy makers dealing with complex strategic problems with practical advice on the interpretation of quantitative model results in strategic planning and encourages scientists to improve the models by developing new techniques for the integration of quantitative and qualitative analyses.

Section four explores experimental modelling techniques and what they can offer to decision-making and joined-up thinking. Especially agent-based models are increasingly popular in integrated assessment and allow for the exploration of new pathways of analysis.

7.1 Simulation models: an overview

In the following, we provide a survey of past and present integrated assessment models. The models are clustered in three groups. The first group contains those models that have been built with the explicit objective of providing an integrated insight into a broad range of environmental, economic and socio-cultural aspects of sustainability. This group of models includes the legendary World 3 model and the more recent International Futures, TARGETS and Threshold 21 models, as well as the accounting based models PoleStar and Quest. The second group contains models that put an emphasis on the link between the energy sector and the environment. These models usually concentrate on the issues of emissions, climate change or acid rain and depletion of resources. This study includes as examples RAINS and the set of six models used in the recent IPCC Emission Scenarios study, namely AIM, ASF, MESSAGE-MACRO, MARIA and MiniCAM. The third group of models focuses on the link between the economy and the environment. It includes JOBS, GTAP, GEM-E3, E3ME and WorldScan. Most models in this group exhibit sophisticated and detailed scenarios of economic developments and translate these into rough sketches of possible environmental impacts.

None of the models covered were developed as sustainability models, though all of them are relevant to one or more aspects of sustainability. In fact, finding a unifying integrated 'sustainability model' might well be impossible due to the characteristics of sustainability. Up to now, all we could identify were proxies like the ones listed below. Nevertheless, we evaluate these models in terms of their suitability as tools for integrated sustainability assessments. In other words, we are testing the models on

terms that are not their own. The descriptions of the models focus on the breadth and depth of models as well as the modelling techniques used. At the end of this section, some preliminary conclusions are drawn concerning the general usefulness of existing models for a thorough and well-balanced analysis and projection of developments related to sustainability.

The criteria for our evaluation are selected from Rotmans & Dowlatabadi (1998). These criteria were formulated for integrated assessment models, not for sustainability models. As a consequence the need for covering the breadth of sustainability is not reflected. However, they do appropriately cover the characteristics of dynamic, trans-disciplinary and integrated models.

- *The exploration of interactions and feedbacks:* We separate the evaluation of interactions and feedbacks into the levels of horizontal and vertical integration. Horizontal integration refers to the integration between different aspects of a models' domain. Issues of integration between the environmental, the economic and the socio-cultural domain as well as integration within each of these domains are important in that respect. Vertical integration refers to integration between different stages of cause-effect chains in the sense of the pressure-state-response cycle. A causal chain combining several domains would be an example of both horizontal and vertical integration and score high in terms of exploring interactions and feedbacks.
- *Flexibility and ease/speed of use:* Good sustainability models can handle changing system structures and rapidly explore the consequences of different system settings. In the best case, these explorations can be executed by or in co-operation with stakeholders. In the policy context, this would allow the prototyping of a large number of options before taking decisions.
- *Communication:* The ideal sustainability model facilitates the communication among scientists of different disciplines, policy makers from different domains and among stakeholders in terms of process and results.
- *Complexity:* Overly complex models compromise the flexibility and communication aspects and are therefore less useful for highly integrated processes and stakeholder involvement. Yet, the increasing

coverage of interactions and feedbacks drives models towards more complexity.

- *Level of aggregation:* There is no perfect level of aggregation for a sustainability model. But there is an appropriate level for different kinds of sustainability assessments.
- *Treatment of uncertainty and random behaviour:* Dynamic complex systems in the real world are full of uncertainty and surprises. Part of the uncertainty lies in the assumptions made. Being able to test these flexibly and produce according results is valuable in a sustainability model. Other uncertainty lies in non-deterministic behaviour. A good sustainability model would therefore allow for the exploration or analysis of discontinuities, shocks, and surprises.

7.1.1 INTEGRATED MODELS

World 3

World 3 was the first comprehensive integrated global simulation model. It became known to a wide audience through the publication of *The Limits to Growth* (D. H. Meadows et al., 1972) and the revised *Beyond the Limits* (D. H. Meadows et al., 1991). The model is a milestone with respect to dynamically integrating different policy domains and provoking joined-up thinking.

World 3 is a system dynamics model covering a wide range of population, food, energy, environmental and economic issues. System dynamics models are based on sets of difference equations. These equations are used to express levels of stock variables and rates as a measure of change in the stock variables. Auxiliary variables, mostly parameters, quantitatively describe the relationships between the different stocks and flows. This modelling technique enables modellers to model complex interactions without having to tackle massive sets of time-series data. Scenarios can be forecasted and backcasted from a base year, which is the only year for which data is needed for a scenario run (for validation and calibration other data sets might be needed). In the case of the revised World 3 model, the base year is 1990. The forms of the equations, the interactions and the values for the auxiliaries and the constants are based on statistical analysis and scientific judgement without participative input. Scenario simulations run from 1900 until 2100. The so-called back casting is used in order to validate

and calibrate the model to actual history. A comprehensive technical description of the original World 3 model can be found in *Dynamics of Growth in a Finite World* (D. L. Meadows et al., 1972).

World 3 is the first comprehensive example of horizontal integration. There is a plethora of feedbacks between the different domains of the models. However, the socio-cultural domain and the institutional domain are not represented in the model at all. Representation of the environmental domain is relatively poor. The level of vertical integration is limited. It is not possible to trace clear cause-effect chains, as the descriptions of the different processes are too crude to allow for a detailed analysis. A major drawback of the model is the fact that it can only provide scenarios on a global level without a regional disaggregation. The scenarios generated are of limited use for policy makers as they allow only very general conclusions about the possible effects of different policy strategies. As each new scenario run requires adjustments in the code of the program, the ease of use for people other than the developers is very limited.

International Futures

International Futures (IF) is a global system dynamics model based on a similar methodology as World 3 (Hughes, 1999). The model simulates population, food, energy, environmental, and economic developments for the period from 1992 until 2050 at a global level and for 14 world regions. IF adds the domains of domestic and global social and political systems to the coverage of World 3. However, these aspects are covered rather implicitly in the model. An important feature of IF is the interactive way of developing scenarios, which allows users to flexibly adjust virtually all values of variables, parameters and constants in order to explore variations of a given base-line. The user cannot adjust the form of the underlying equations. The IF base scenario draws on the extrapolation of current trends. The greater breadth of the model increases horizontal integration. However, the environmental domain is limited and poorly represented. With respect to vertical integration the model suffers from similar problems as World 3. The possibility to quickly create a set of partially custom-made scenarios is valuable for policy analysis. Users are more likely to identify with the results as they are based on their own assumptions.

TARGETS.

The Tool to Assess Regional and Global Environmental and Health Targets for Sustainability (TARGETS) is another example of a global system dynamics model (Rotmans & de Vries, 1997). As the title suggests, TARGETS puts its emphasis on the assessment of long-run environmental developments, taking into account biophysical, social and economic processes. Estimates of impacts on ecosystems and humans are the central results of the models. Like World 3, the model is restricted to the global level as data sets are fully aggregated at this level. Scenarios run from the year 1900 to 2100. The explicit introduction of cause-effect chains differentiates the model from previous system dynamics models described in this overview. For this purpose the pressure-state-response approach developed by the OECD (1993) has been refined. The state part is split into state and impact to allow for a more detailed distinction between changes in the various functions of (sub) systems and the changes in the state of the system. TARGETS consists of five sub-models and an economic scenario generator, which are integrated and interlinked with each other. The five highly aggregated sub-models are meta-models of expert models in the respective fields. The structures of the expert models have been simplified in order to make integration between the models possible. The result is a set of meta-models capturing the behaviour of their expert-parents quite well, without describing the field concerned in infeasible detail. The sub-models cover the aspects of population and health, land and food, energy, water, and global bio-geo-chemical cycles. The economic scenario generator is a relatively primitive tool to feed exogenous GWP (Gross World Product) trajectories into the model. TARGETS has a unique way of dealing with uncertainty. The model uses a generic description of the users worldview and management style to adjust the underlying assumptions. The options offered for both choices are derived from cultural theory (Thompson et al., 1990): egalitarian, hierarchist and individualist. They influence the way in which adjustment in behaviour, learning and other aspects takes its effect during the run of the model.

TARGETS focuses on the assessment of environmental and health effects of sustainability. Within these fields there is interaction between the different sub-models. There is no further horizontal integration with other sectors. Within the coverage of the model, horizontal integration through the consequent application of the pressure-state-impact-response approach (PSIR), vertical integration within the domain of environmental assessment is rather advanced. The model has little use for concrete policy analysis or

advice other than to quantify the urge for a reduction in resource use and emissions. The model's strength lies in its ability to offer insights with regard to uncertainties and risk and with regard to interconnectedness, rather than answering the questions concerning policy options arising from these insights.

Threshold 21

Threshold 21 (T21) is a system dynamics model that covers the domains of the three previous models.¹² It combines the social, economic and environmental domains in an extremely transparent model. Within the structure of the model, market and government behaviour plays a key role. This allows users to easily adjust political strategies and view its impacts. These adjustments can be introduced in the form of changes in the otherwise static parameters that quantify the relationship between the different stock and flow variables. The model is generally adjusted for use on the national level and has already been used in a number of developing countries as well as in Italy and the US. Global influences in issues such as GHG emissions are also included in T21. The level of detail in the output variables is impressive. However, this level of detail is not applied in the description of the cause-effect chains.

Horizontal integration in T21 is very high. All the sectors directly influence each other through a plethora of feedback cycles and other linkages. Vertical integration is also implemented, but less detailed. The high accessibility of the model and the ease with which every link can be adjusted allows the fast analysis of a large number of different policy options under changing assumptions. One should, however take care that the model is calibrated only for a given reference data set. Playing with the model in itself does not provide substantive input for operational policy analysis, but communicates a feeling for the complexity of sustainability and provides insights into strategic options for governance.

POLESTAR.

PoleStar uses a very different technique from the models above. PoleStar is an integrated accounting framework developed by the SEI Boston Center

¹² Details and papers on applications can be found at <http://www.threshold21.com/resources/elibrary>

and applied by the Global Scenario Group (Gallopín et al., 1997). The backbone of the model is an extensive data set containing a wide range of social, economic and environmental variables. Some political and cultural variables are also included. In a stepwise procedure the user can introduce assumptions from other studies and models and check these against the values of indicators of sustainability for future years. The first step of the scenario building process contains the external development of scenarios for population change and economic development. Based on these scenarios, the model calculates the consequences for the environment and for world resource availability based on standard parameters. Societal responses to certain developments can be introduced externally to explore a set of different pathways. The geographical focus of the model is rather flexible. For the year 2050, two reference scenarios serve as default options for all interrelations not defined otherwise by the user. Also, new variables can be introduced and existing ones ignored. These options make PoleStar a very flexible accounting tool (SEI Boston Center, 1999).

PoleStar exhibits limited horizontal and vertical integration. Horizontal integration is restricted to strictly necessary interactions, vertical integration is lacking on the whole. This limits the possibility to derive integrated and balanced answers for issues related to political decision making from the model. The user-friendly model provides a useful first insight with respect to very specific rather short-term questions. But PoleStar's extrapolative nature prevents insights into the underlying dynamics of sustainability.

Quest

Quest is a framework that combines the properties of an accounting framework with elements of input-output, GIS, and other modelling techniques for a computer simulation of regional developments (SDRI, 1999). Quest simulates the development of a region in an integrated way. The purpose of Quest lies more in the education about the complexity of a sustainable future than in the development of scenarios in their own rights. Through a very user-friendly interface, the user, who is usually a local stakeholder, explores different strategies for local development. The user also defines his view of what sustainability is in choosing a few benchmark variables that are important in his/her opinion. These benchmarks mainly affect the visualisation, but not the calculation of the scenarios. The computer program runs in steps of ten years after which the development progress is being reviewed and the strategy can be adjusted up to 2030 (Carmichael et al., 2004). In this way, Quest is quite similar to the popular

computer game SimCity. Playing with different scenarios teaches the user a feeling for the complexities involved in the sustainability of a region. These features make the model a suitable tool at the interface of scientific advice, political decision-making and communication with the general public (Tansey et al., 2002).

Quest explores issues related to sustainability on a sub-national level with good horizontal integration. However, this integration is too rigid in its interactions to allow for reasonable long-term evaluations as most parameters for interaction are constants or develop in a linear way. Vertical integration in Quest lacks the final feedback loop in the cause effect chain. Quest is very useful for policy advice on the local level as it is detailed, flexible and allows stakeholder engagement. The strength of the model turns into a weakness if one wants to apply such a model for nations, as it is currently not possible to cover the many co-existing contradicting trends in different regions within Quest. In addition, the model is highly region-specific. Transferring it to other regions requires a substantial re-development.

7.1.2 ENVIRONMENT-ENERGY MODELS

RAINS

RAINS (Regional Air pollution INformation and Simulation) is a first generation integrated assessment model developed by IIASA. The model focuses on the analysis of SO₂, NO_x and NH₃ emissions and their effects on sensitive ecosystems through acidification and eutrophication (Alcamo et al., 1990). The projections are based on the technique of linear programming, although non-linear components have been included in later versions. RAINS consists of a large number of sub-models that are clustered in three modules: the emission-cost module (EMCO), the acid deposition and ecosystem impact module (DEP) and the optimisation module (OPT). Estimates of emission costs are based on information about economic activity, energy consumption, fuel characteristics and agricultural activities derived from national statistics and projections. The DEP module allows the manipulation of geographical data and functions as an interface between the other two modules. The optimisation module finds the cost-minimal allocation of measures to reduce emissions for a set of given target emissions.

Horizontal integration of RAINS is limited by its specific focus from the perspective of comprehensive environmental assessment. Within the domain of the model, vertical interactions are modelled in great detail. RAINS provides useful output within its focus. The model has been widely and actively used in policy analysis and negotiations on an international and national level (van der Sluijs, 1997).

Multi-Model Approach for Developing Emissions Scenarios

The IPCC uses a multi-model approach for developing GHG emission scenarios (IPCC, 2000). The IPCC considers the selected models as representative. For their previous emissions scenarios six models were used: The Asian Pacific Integrated Model AIM (Morita et al., 1994), the Atmospheric Stabilization Framework Model ASF (Lashof & Tirpak, 1989), the Integrated Model to Assess the Greenhouse Effect IMAGE (Rotmans (1990); Alcamo (1994)) used in connection with the WorldScan model described in the next section, the Multiregional Approach for Resource and Industry Allocation MARIA (Mori & Takahashi, 1999), the Model for Energy Supply Strategy Alternatives and their General Environmental Impact MESSAGE (Messner & Strubegger, 1995) and the Mini-Climate Assessment Model MiniCAM (Edmonds et al., 1996).

AIM is a general equilibrium model developed by the National Institute of Environmental Studies in Japan. The model focuses on the assessment of greenhouse gas emissions. It calculates the level and type of energy use on the basis of external socio-economic scenarios. The model consists of six sub-models: The Socio-Economic Scenarios sub-model is an input module for the external scenarios. It requires the exogenous input of GDP, population, resource base and lifestyle developments. The Bottom-up Model provides information about sectoral energy and resource efficiency. The Energy-Economic Model calculates a market equilibrium for energy markets. It does so for 17 world regions. Among them is the OECD-West. The Land Equilibrium Model calculates the land use based on biomass energy demand, food consumption patterns and technological change. From the latter two sub-models GHG emissions are calculated which are feeding into the AIM Climate Model and subsequently the AIM Impact Model. Outputs from the latter model can be used to provide feedback concerning the socio-economic scenarios.

Horizontal integration in AIM is limited by the extremely narrow focus of the model. The only interaction between the socio-economic module and the

environmental module concerns the climate related land-use patterns. AIM lacks horizontal integration with other aspects of environmental sustainability. Vertically, there is only a limited feedback mechanism between the environmental impact calculation and the socio-economic scenarios, which interrupts the pressure-state-response cycle. The exogenous treatment of socio-economic developments limits the level of vertical integration. The geographical focus of the model is primarily meant to cover global issues rather than issues relevant on a regional scale. A very detailed version of the model exists for the Asian Pacific region where the model distinguishes 3200 regions.

ASF is an accounting framework covering the fields of energy, agriculture, deforestation, GHG emissions and an atmospheric model. Emission estimates are provided for nine world regions, among them the OECD-West. ASF consists of a cluster of five sub-models. The energy model contains four sectors: residential, commercial, industrial and transportation. An equilibrium for the energy market is based on prices, which differ by region, type of energy and conversion costs. An agricultural model is linked to a deforestation model. Interactively, these two models calculate the agricultural production and the area of deforestation based on population and GNP developments. The emissions model uses the output of the three previous models to estimate GHG emissions. The atmospheric model uses this as an input to calculate the effects with respect to temperature and CO₂ concentration.

Integration in the ASF model is limited horizontally by the narrow focus of the model. Within this focus the level of integration between the different modules is actually quite high. The model does not represent any socio-economic factors. The horizontal integration of economic and environmental issues is very limited. Vertically, there are some gaps in the PSIR causal chain, especially concerning responses. The possibility to derive substantial input for policy analysis from the model is limited by these factors.

IMAGE was developed at RIVM in order to assess the impact of anthropocentric climate change (Rotmans, 1990). IMAGE 2.1 was used in the assessment. It consisted of three sub-models. The Energy-Industry System (EIS) calculates greenhouse gas emissions in thirteen world regions. Emissions from the energy sector are modelled with the TIMER simulation model, a system dynamics model for the prediction of energy related information in 5 sectors of the economy. The Terrestrial Environment System (TES) simulates land-use and land-cover change and their consequences for

climate related biophysical processes. The Atmosphere-Ocean System (AOS) calculates the behaviour of greenhouse gases in the atmosphere and its effects on temperature and precipitation patterns. Sub-models of the AOS cover the atmospheric composition, zonal atmospheric climate, oceanic climate, oceanic biosphere and oceanic chemistry. Economic data, technological change, demographic developments and control policies are exogenous input into the model. Simulations run up to the year 2100. The geographical focus of the model ranges from a 0.5 degree \times 0.5 degree latitude-longitude grid to world regional level (CIESIN, 1995). An overview of the history and further developments of IMAGE can be found in Bouwman et al. (2006).

IMAGE 2.1 exhibits a lack of horizontal integration from the perspective of general sustainability analysis, but within its narrow focus horizontal integration is exemplary. Horizontal integration is introduced at all computing levels of the model and not restricted to higher domains. There are a multitude of important feedbacks between models in the sub-systems, and between sub-systems. Vertical integration is more limited as there is a lack of feedback from the outcome of the model and the input. Only the output of land-use and climate data is used as input for the next round of calculations without affecting other exogenous inputs. In perspective of a more general model of sustainability in Europe, it is questionable whether the level of detail and integration could be achieved in a model that covers a wider domain.

MESSAGE-MACRO: MESSAGE is a dynamic linear programming model. It is a sub-model of the IASA integrated modelling framework and is generally used in a tandem with the MACRO macro-economic model. The two models are a combination of top-down and bottom-up modelling techniques. MACRO (top-down) calculates the maximal utility of a representative consumer for each region and its relations with macro-economic development and energy use. MESSAGE (bottom-up) calculates energy demand, supply and emission patterns on the basis of economic input. The model requires the exogenous input of population and GNP scenarios on a regional level. From this, the Scenario Generator (SG) derives scenarios for future energy demand that are consistent with empirical results derived from historical data stored within the SG. The model divides the world into eleven regions. In the IASA integrated modelling framework MESSAGE is also used in combination with RAINS, MAGICC and other specific models.

The MESSAGE-MACRO tandem in its own right lacks vertical as well as horizontal integration. Horizontal integration can be improved through the interaction with other models, but is then limited to the highest scale level and does not directly affect sub-sections of the model. Vertical integration is also a problem because effect chains have to be limited to rather general descriptions. As such the model is more a tool for the analysis of the consistency of given scenarios rather than a relevant basis for political decision-making.

MARIA is an inter-temporal non-linear optimisation model focusing on the assessment of technology and policy options available to address global warming. The model is based on the DICE model (Nordhaus, 1994) and covers aspects of economics (consumption and trade), land use, natural resources, and energy. MARIA divides the world into eleven regions. The model requires the exogenous input of population and potential per capita GDP growth rates. Economic activities are then modelled assuming constant substitution elasticities. The analysis of emissions is limited to global carbon emissions.

The level of horizontal integration of the model is limited by its narrow focus. Vertical integration is limited by the exogeneity of key economic indicators. With respect to policy relevance there is a lack of detail concerning the policy options available. The effects of specific policy choices are hard to evaluate. The strength of the model lies in the general macro-level evaluation of various consistent options. The regional division of the model makes it unsuitable for the application to EU specific problems.

MiniCAM is an integrated model developed at Battelle Pacific Northwest Laboratories. The model consists of a combination of the Edmonds-Reilly-Barnes (ERB) energy model and the Model for the Assessment of Greenhouse-Gas Induced Climate Change (MAGICC) atmospheric global climate model. The ERB model describes long-term trends in economic output, energy use, and greenhouse gas emissions for nine world regions through detailed sub-modules representing energy resources, primary energy supply and demand, energy markets including world trade and electricity conversion, and fuel-specific emissions factors. An exogenous economic scenario is fed into the ERB model, which calculates the energy supply, demand, balance and resulting GHG emissions. This output is used as input into MAGICC where the gas-cycle module feeds into the climate module and the sea-level module. The atmospheric composition, radiative forcing, mean global temperature rise, and sea level rise feed back into the

ERB model. Some applications use the MERGE model to translate the output from MAGICC into input for the adjustment of the economic scenarios in the form of market and non-market damages using simple damage functions. A simple graphical spreadsheet-based interface allows the fast exploration of different scenario runs. The geographical focus of the model distinguishes 14 world regions.

MiniCAM exhibits low levels of horizontal and vertical integration. Horizontally, the economic and environmental issues addressed are very limited. There is a complete lack of socio-cultural aspects. Vertically, there are no feedback mechanisms other than the final link of output to input, which is one of the weakest links of the model. A full-feedback version including the link between climate change and the economy is under development. The purpose of MiniCAM lies more with the quick exploration of a set of possibilities rather than a detailed description of the domains involved.

7.1.3 ECONOMIC-ENVIRONMENTAL MODELS

The third group of models focuses on the assessment of the link between economic developments and its environmental consequences.

JOBS

JOBS is a rather simple simulation model solving a sequence of static equilibriums. It does so under the assumptions of national income accounting where aggregate investment equals aggregate savings for each time step. Substitution elasticities are held constant throughout a model run. JOBS is a global model that divides the world into ten regions. The model uses the GTAP data base (see below) as a source. The model has been developed at the OECD (Bakkes et al., 2000).

The level of horizontal integration of the model is limited to a selected few ecological and economic indicators. The level of vertical integration is limited by the non-dynamic modelling technique. The breadth and width of the model are more limited than with the other economic-environmental models. The lack of accessibility to the model is a major drawback. It limits the possibilities of analysis to the simulation of isolated events and their effects in a very limited domain.

GTAP

The Global Trade Analysis Project (GTAP) is a general equilibrium model (Hertel & Tsigas, 1997). The model represents the developments and interactions between 50 sectors in 45 regions on the basis of the assumptions of perfect competition and constant returns to scale. The model is restricted to the economic analysis of consumption and trade. The model does not address environmental and socio-cultural aspects. So there is a lack of horizontal integration. Vertical integration is much more advanced. Similar to WorldScan (see below), the model uses an Armington trade specification in its representation of inter-regional trade.

GTAP is a sophisticated tool for the simulation of global economic and trade patterns. Within that domain, the model exhibits a good degree of vertical integration. Horizontal integration is limited to the trade sector. Other economic perspectives of sustainability as well as environmental or socio-cultural aspects are lacking. However, GTAP might become relevant as a provider of economic scenarios, an aspect that has been rather neglected in many of the other models described here. In its own right the domain of the model is insufficient to address questions of sustainability. An interesting part of the model is the GTAP database. It covers a wide range of variables and is constantly extended according to the needs expressed by its users. Not all of the variables in the database are used in GTAP as the database is open to all users and has been used in other models as well.

GEM-E3

GEM-E3 (General Equilibrium Model for Energy – Economics – Environment) is another general equilibrium model (Capros et al., 1996). GEM-E3 focuses on the macro-economy and its interaction with the environment and the energy system. The model covers the EU with an explicit national subdivision. Each member state economy is divided into 11 sectors. Trade in goods and services as well as dynamics of capital accumulation and technological progress are endogenous. Public consumption is exogenous. GEM-E3 calculates the equilibrium for all represented markets annually. A fully operative sub-model represents the monetary side of the economy. A speciality of the model is the close attention it pays to distributional effects between countries and industrial sectors. In order to be able to cover the environmental and energy effects of consumption, the representation of the consumption choice allows for substitution at a highly disaggregated level. Also, the model discriminates

between durables, which are treated as stocks and non-durables, which are treated as flows. Utility functions dominate the choice between work and leisure as well as present and future consumption. A consequence of this is that the strength of the model lies in the assessment of those environmental impacts that are related to energy and resource use.

GEM-E3 lacks feedback mechanisms between the environmental sector and the economy. There is no further horizontal integration with other sectors. The level of detail within the environmental sector is not sufficient to horizontally integrate several different aspects of sustainability. Institutional and socio-cultural aspects are missing from the model. There is a good degree of vertical integration. The results of the model are relevant for policy decisions, but the analysis of possible political strategies is difficult. The model is best suited for a qualitative analysis of trade-offs between conflicting policy objectives. Of course, this analysis is restricted to the domain of the model. The modelling technique restricts the inclusion of dynamic effects and surprises. These can only be introduced exogenously.

E3ME

E3ME (Energy-Economy-Environment Model for Europe) was developed under the co-ordination of Cambridge Econometrics (CEC, 1995). It is an econometric input-output model in a general equilibrium framework representing sectoral and regional developments. The model divides each economy into 30 industries, 27 consumer categories, 17 fuel users and 11 fuels. CO₂ and SO₂ emissions are represented with a great level of detail. The economic part of the model is based on neo-Keynesian assumptions, i.e. there is no market clearing assumption for labour and goods markets. The major advantage of the model over existing national models is the common base of definitions and classifications so that results can easily be aggregated to the European level (Barker & Koehler, 1998). Values for output, prices, employment, fuel demand and supply are calculated for annual intervals. The data for the model are derived from EUROSTAT statistics.

Horizontal integration of the model is achieved within a restricted domain. The model limits its attention on environmental sustainability to emissions. Socio-cultural sustainability is not addressed, although an extension of the model represents 14 socio-economic groups for classifying households. The level of vertical integration is difficult to judge on the basis of the information available. The rather short time horizon limits the policy relevance of the model for long-term sustainability planning as it misses the

representation of developments in capital stocks. The regional focus on Europe with the opportunity for national disaggregation is a great advantage of the model.

WorldScan.

WorldScan (WORLD model for SCenario ANalysis) is a global applied general equilibrium (AGE) model with a strong emphasis on economic aspects (CPB, 1999). The underlying assumptions of the model are based on neo-classical economic theory. The model is suitable for a wide range of applications in the fields of economics, energy, transport, trade and environmental policy.

WorldScan divides the world into twelve regions. The economy is divided into 11 sectors with differing factor requirements. These are rather crudely defined to allow major shifts in production within sectors. Primary inputs in all sectors are low-skilled labour, high-skilled labour, capital and a fixed factor. Economic growth is predicted in line with neo-classical growth models on the basis of physical capital, labour and technology. Technology is allowed to differ between regions and can easily be adopted by developing countries. Labour is divided into low-skilled labour and high-skilled labour and in developing countries a fraction of the labour force works in low-productivity sectors, i.e. the subsistence sectors. Trade is modelled in such a way as to avoid abrupt specialisation patterns. Consumption is allocated over time, categories and regions. Consumption patterns generally tend to converge with OECD preferences and are assumed to change with changes in GDP per capita. Richer societies spend relatively less on agricultural products and relatively more on services than poorer societies. These allocation decisions influence the amount of transport required and thus the level of emissions.

An important application of WorldScan has been to analyse policy strategies for the implementation of the Kyoto protocol on climate change. For this purpose, the trade module of the model has been adjusted to incorporate the trade of different kinds of permits and the Clean Development Mechanism (CDM). With respect to the environmental effects of globalisation and the related increase of transport, four scenarios for the transition towards the year 2020 and in some respects towards 2050 have been developed using WorldScan. The assumptions underlying these four quantitative scenarios are derived from the qualitative scenarios developed by Van Veen-Groot and Nijkamp for the GITAGE project. Future work with WorldScan is

planned to be more explicit with respect to the effects of trade, consumption and production on international transport and the related environmental costs.

WorldScan is useful with respect to the analysis of global economic trends. Horizontal integration with environmental, institutional and socio-cultural factors is lacking. Vertical integration is insufficient due to a lack of feedback mechanisms. The lack of integration makes it difficult to derive balanced policy advice from WorldScan. For the economic domain, specific policy advice would be improved through a better access to different scenario runs, which would allow the analysis of the effects of different policy strategies concerning a specific question.

7.1.4 LESSONS FROM THE INVENTORY

It would hardly be fair to conclude from the overview above that the available models do not fulfil the objectives of full horizontal and vertical integration and policy relevance. A model fulfilling all these criteria, a Swiss Army Knife of models is an impossibility as should be evident from chapter 3. Each of the models was designed for a specific context and purpose and many models have been highly effective within their scope. The value of the criteria lies in helping to show differences between models that are often perceived as being similar.

Some patterns in the descriptions of the models are striking and provide insights into problems and possibilities with respect to the development of models for sustainability assessment. These insights concern the trade-offs between the depth and the breadth of models as well as between the transparency of a model and its policy relevance.

Those models that have been developed in order to assess sustainability in an integrative way, i.e. World 3, IF and TARGETS are the least relevant for concrete political decision making as they do not offer a regional analysis of variables that policy makers could directly influence. The main advantage of these models lies in their role as a building block of long-term policy advice and policy support processes. The models efficiently communicate possible outcomes of societal (lack of) action with respect to highly complex issues. The World 3 model, for example, has had a pivotal role in bringing environmental issues from the science domain to the public domain. By doing so, the model has triggered a societal discourse and many policy decisions. It is interesting that all four models are based on system

dynamics. The TARGETS model combines system dynamics with other modelling techniques that providing sector specific input scenarios. The strength of system dynamics seems to lie in a more general description of interactions between a wide variety of fields. So there is a trade-off in favour of breadth at the cost of depth.

Another interesting fact is the use of general equilibrium modelling techniques by all economy-environment models described in section four of this chapter excluding JOBS. JOBS uses the closely related concept of static equilibriums. One should note that most other models require the exogenous input of economic and demographic data. None of the seven environment-energy models included features endogenous treatment of economic factors. The six models of the first group also lack a reasonably detailed representation of the economy. So there is a definite need for the integration of more sophisticated economic modules in existing models with an environmental focus. These models provide the user with considerable depth at the cost of crucial breadth. However, the integration of existing models with an environmental focus with models that put an emphasis on economics encounters the difficulty of conflicting modelling techniques.

It could be argued that the more communication and user-friendly models such as PoleStar and Quest have externalised the process of political feedback to scenario developments by letting the user provide input at given intervals during the scenario run. This accommodates normative choices and enables researchers to gain insight into links that have previously been neglected due to lack of quantification. This also allows for rapid prototyping and the exploration of uncertainties and surprises in those cases, where the models allow changes in the system structure during a model run.

The socio-cultural domain has not been explored in satisfying depth in any of the above models. The knowledge, data and methods to do so do not (yet) exist. The indicators available to represent this domain are still largely demographic or economic and only marginally correlated to the underlying issues. Building an integrated model that fulfils the objectives of breadth and depth combined with policy relevance would require a concerted effort of several innovative research institutes. Based on the existing models combined with some new elements, such a project could result in a model that can provide a base for horizontally balanced and vertically integrated policy analysis with respect to concrete policy targets. According to van der Sluijs (1997)

[a] perfect IAM [integrated assessment model] would model the complete so-called causal chain, including all the feedbacks within this chain. The causal chain starts with socio-economic drivers leading to economic activity and other practices, leading to emissions and other pressure on the environment leading to environmental changes, leading to physical impacts on societies and ecosystems, leading to socio-economic impacts, eventually returning to cause changes in the socio-economic drivers.

Such a model would nest a detailed regional analysis in a less detailed world model. Furthermore it would allow for an analysis of sustainability at the national and sub-national level. This would require a detailed representation of the different issues related to sustainability including economic, environmental and socio-cultural aspects.

This ideal model is a utopian ambition. However, outlining utopia sketches different directions of improvement. Modellers and researchers working with models should be aware of these ambitions and carefully select their priorities among them according to the specific context in which they want to apply their tools. In the past of integrative research, many innovative models have been a driving force of societal innovation with large volumes of qualitative work following the lead. This has especially been the case when the models were developed and interpreted with a high degree of attention to the characteristics of the public or political discourse. The World3 model is the most exponent example of this. A less known example is the DICE model (Nordhaus, 1992) that was widely used to support the negotiations on international greenhouse gas regimes. In the following section, we explore how the models used at the province of Limburg can be improved in that sense.

7.2 Case study: Quantitative models for strategic policy planning

7.2.1 INTRODUCTION

The previous section has shown the difficulty of developing a policy relevant sustainability model. In this section, we revisit the POL-prognoses case study (see chapter 5 for details) in order to explore the application of quantitative models on the regional level in the Netherlands. For this analysis, we examine those models in use at the administration of the

province, rather than models that might be used. The case study allows us to study in detail the discrepancies between ambition and reality in models for public strategic planning and regional sustainability. This insight facilitates the integration of existing models in the policy process. The analysis is based on an inventory and evaluation of quantitative forecasting models used by the provincial administration to support its strategic planning. We explore the differences between the relationships covered in the conceptual POL-SCENE model (see chapter 5) and in the structure of the forecasting models. As a reminder, the POL-SCENE model consisted of the following 18 stocks: population (Popu), consumption (Cons), social structure (Soci), public amenities (Amnt), housing (Hous), security (Secu), space (Spac), air (Air), groundwater (G-wat), surface water (S-wat), soil (Soil), nature (Natu), entrepreneurship (Entr), production (Prod), knowledge and innovation (Know), work (Work), mobility (Mobi), and infrastructure (Infr). These were connected by 95 flows.

7.2.2 MODEL INVENTORY

To compare the completed SCENE model analysed in chapter 5 with the coverage of the available quantitative models, we conducted an extensive inventory and represented the underlying structures in the same format as previously. In total, we identified 13 quantitative models that the administration had previously used for strategic planning:

- The local research institute ETIL contributed separate models on three issues: population, labour markets, and housing. These models translate national forecasts on these issues down to the regional level using a shift-share approach (ETIL, 2000).
- The PRIMOS model produces forecasts on overall population size, number of households, and housing requirements. The acronym stands for PRognose-, Informatie-, en MONitoring Systeem (Prognosis, Information, and MONitoring system) (ABFresearch, 2006b).
- The RDP model (a Dutch acronym for Spatial Demographic Forecasting model—Regionaal Demografisch Prognosemodel) produces region-specific population projections based on national estimates and is quite similar to the ETIL models, though the underlying assumptions differ (ABFresearch, 2006c).

- The InterProvinciaal woningBehoeftte-model (IPB—Interprovincial Housing Needs Model) is a highly simplified shift-share module that produces population forecasts on the community level (ABFresearch, 2006a).
- The RAIL (Regionale Arbeidsmarkt Informatie Limburg—Regional Labor Market Information Limburg) model produces a regional sector-specific forecast on labor market supply and demand (Teunis, 1996).
- The COMBI (een COMBINatie van modellen—a COMBINATION of models) model forecasts regional migration patterns for the country as a whole based on population, labor market, housing, and higher education modules (Heida & Poulus, 1993).
- The Dutch Central Planning Bureau (CPB) produces a set of three quantitative scenarios with respect to environmental quality, mobility, space, and energy. These scenarios are documented in CPB (1996).
- The OGM model (Overdraagbaar Groeimodel—Transposable Growth Model) forecasts developments in regional mobility (4cast, 2006).
- The MIOW-PROV model forecasts the socioeconomic consequences of environmental regulation. The acronym stands for Marktsituatie, Internationale concurrentie, Omvang en Weerstandsvermogen PROVinciaal milieu-economisch model—Market Position, International Competition, Volume and Resilience Provincial Environmental-Economic Model (van der Woerd et al., 1997).
- PROVEST (PROgnosemodel VESTigingslocaties—Forecasting Model on Locations of Commercial Settlements) forecasts the physical space required for the development of business and industry in a shift-share manner (BCI, 1998).
- The LeefOmgevings Verkenner (LOV—Environment Explorer) is a geographic information system (GIS)-based model for exploring changes in land-use patterns. The evolving maps cover a large set of land-use types, and so include social aspects such as housing, environmental aspects such as open landscape, and economic aspects such as infrastructure (Engelen et al., 2003).

Overall, four models had a demographic focus, five had a socio-economic orientation, two pertained to the labour market, and two concentrated on

land use. Some of these tools were simple shift-share models that translated national extrapolations down to the regional level, while others were sophisticated stand-alone constructions based on statistical regression, general equilibrium, or GIS approaches. Approximately half were developed specifically for the region and the other half are nationally run models that are disaggregated and optimised over all Dutch provinces. These two model types often cover similar topics.

Due to divergent assumptions, regional and national models are often at odds in their projections. The differences are a consequence of statistical estimation biases to which outliers in panel-data sets are subject. For example, labour is highly mobile between most Dutch provinces when one location has vacancies in employment and another has unemployment. However, this mobility does not extend to the peripheral provinces, including Limburg. As Limburg historically had many vacancies unfilled, the national forecasts assumed high immigration into the region and thus projected a higher population than region-specific models that neglected this relationship. In the past, the lower population forecasts have turned out to be the better estimates for Limburg.

As many of the evaluated models are either run at the national level or shared among the provinces, it is safe to assume that the set is representative for other provinces in the Netherlands. Members of the Limburg administration, who frequently took part in inter-provincial meetings on monitoring and strategic planning, supported this view. At smaller-scale levels, the coverage of the models seems to get even thinner; at higher-scale levels, the coverage is slightly improved as more general and global models become relevant.

We translated each model into an influence diagram using the stocks described in the previous section. To accomplish this, we screened the available primary and secondary model documentation and conducted extensive interviews with the model developers and/or owners where documentation was insufficient. We accepted a flow as being covered by a model when its framework incorporated any aspect of the flow (either exogenously or endogenously). For example, the flow from 'knowledge and innovation' to 'entrepreneurship' is represented in the relevant models as an exogenous constant by which productivity increases each year. This facet of the model would obviously be insufficient to explore the effects of changes in the education system and it certainly does not answer potential questions of public officials about increasing entrepreneurship through stimulating

knowledge and innovation. More generally, the description provided in the conceptual influence diagram is much richer than in the models themselves. Figure 7.1 combines the separate influence diagrams for each model into a single overview.

Quantitative models used for strategic planning cover 30 of the 95 flows in the conceptual influence diagram. It is important that these modelled flows are not necessarily comprehensively addressed and the different models are not at all compatible in their underlying assumptions or the methods used to make the ultimate calculations. The models are certainly not calibrated for use within a single system of models. This fragmentation of coverage is lost in the illustration above.

The stocks 'social structure' and 'groundwater' are not covered at all. A third of the remaining stocks are only connected to the system via a single remaining flow. Some flows are covered by several models, most of which concern population and labour-market dynamics. We did not detect a single flow covered in any of the models that group participants had not previously identified. Well-covered stocks include the extent and condition of 'housing', the effects (but not the causes) of changes in the 'population', the economic effects of 'entrepreneurship', changes in land use ('space'), and the situation regarding 'mobility'. Other economic—and most social and environmental—stocks are not covered at all or, at best, their treatment is highly superficial. A more favourable picture emerges when the weights of the flows are taken into account. There is coverage of 12 (out of 15) strong flows, 16 (out of 46) medium flows, and only two (out of 34) weak flows.

7.2.3 DISCREPANCIES

It is immediately obvious that the loss of information moving from qualitative to quantitative analysis is significant. The discrepancies below the surface of figure 7.2 are likely even more significant. While the models covered well the dynamics between the stocks labelled 'population', 'housing', 'entrepreneurship', and 'mobility', the majority of other flows were reduced to exogenous parameters. The structural analysis tools described in chapter 5 make it possible to gain further insight into the changing role of variables.

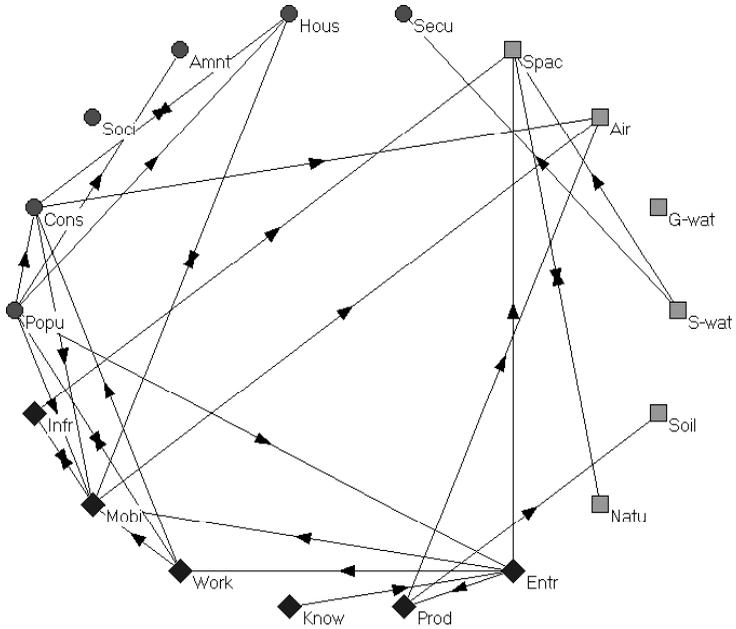


Figure 7.1: Flows covered by the quantitative forecasting models.

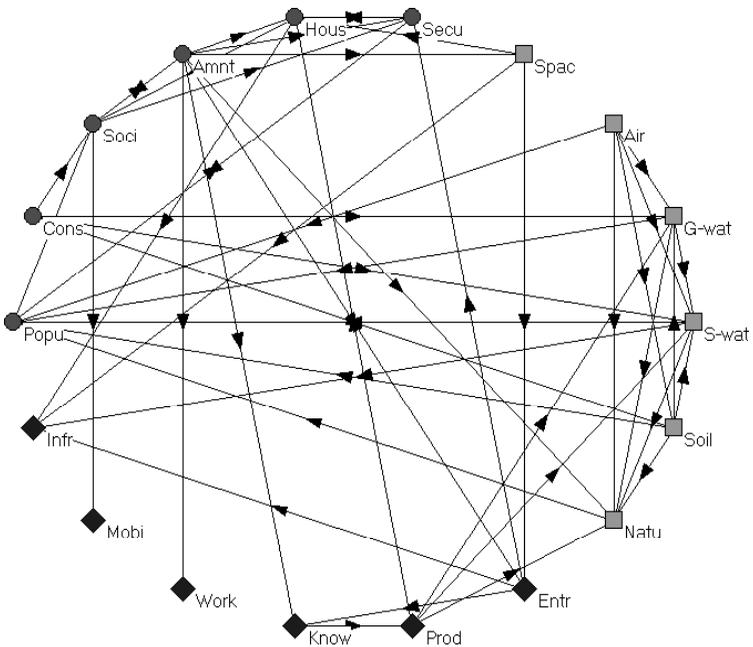


Figure 7.2: Flows not covered by the quantitative forecasting models.

Figure 7.2 shows the 65 flows that are not covered in the models, but that were identified during the project's first phase. It shows the discrepancy between the ambition of van der Sluijs's ideal model and the reality of models that typically inform public strategic planning.

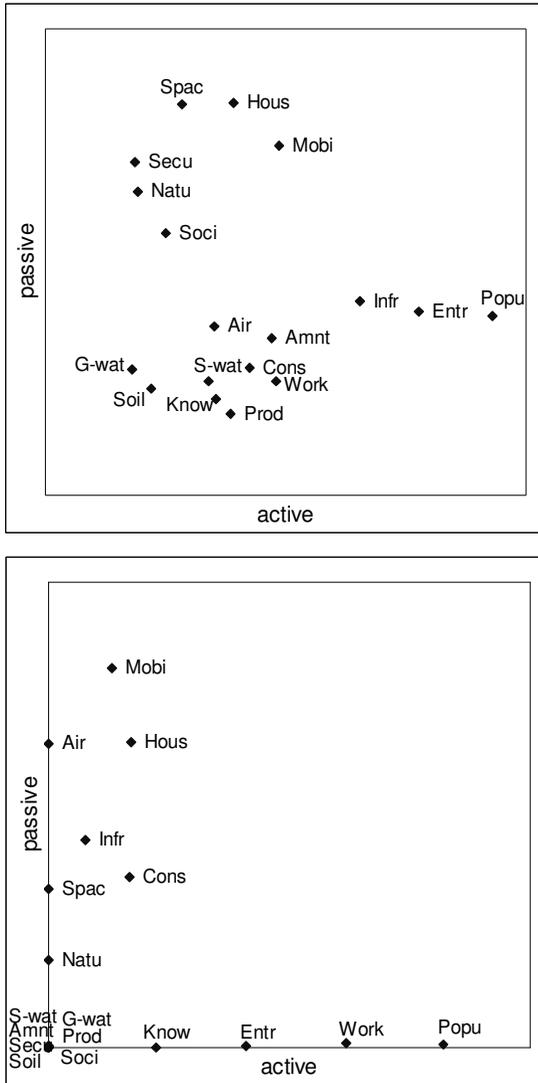


Figure 7.3: Results of a MICMAC analysis of the SCENE model for the POL case (above) and of the combined quantitative forecasting models (below).

Figure 7.3 compares the role of each of the 18 stocks in the dynamics of the POL-SCENE model with those in the quantitative models for strategic planning used by administration of the Province of Limburg.

The graphs show the results of the MICMAC analysis as described in chapter 5. As a reminder, with MICMAC, the character of a variable is mainly determined by its passive and active potential. The passive potential describes the degree to which a stock is influenced by other stocks within the system. The active potential describes the degree to which a stock influences other stock. Variables in the bottom left sector are hardly embedded in the system. The top left sector contains dependent variables with little influence. Active variables are found in the bottom right sector of the chart. Variables in the top right sector of the system are critical. They are part of a maximum number of feedback cycles in the system.

The MICMAC analysis clearly illustrates the misrepresentation of variables in the system and loss of dynamic information. In the conceptual model, all variables are to some degree influential and dependent. This important characteristic is lost when we look at the dynamics of the quantitative models: seven stocks are excluded from the dynamic interplay and three more have an active potential of zero. The stocks 'know', 'entr', 'work', and 'popu' lose virtually all their passive potential. A potential of zero, be it active or passive, interrupts all feedback cycles running through a stock. Stocks become either active, or passive, or excluded. The resulting clear-cut L shape of the variable cloud strongly suggests a misleadingly stable system. Surprising developments and neglected side effects will remain unnoticed based on the given set of models.

One could summarise these findings by saying that the system represented in the models fundamentally differs from the relevant conceptual model in terms of coverage, dynamics and behaviour. The absent passive potentials for the majority of variables imply that the dynamics of the system rely fully either on a statistical analysis of the past or on a rough estimate of parameter value. Modellers are not to blame for this relative poverty in systems information. It is a reflection of how little we know about many of the relationships involved.

Nevertheless, technical improvements are possible and important. In many cases, the exact parameter values were not documented. Inaccessible parameters mean that flows are putatively covered, but not documented in an accessible way. Because of this lack of transparency, even simple scenario

studies cannot be executed in an informed manner. The richness of the qualitative influence diagram is almost completely lost.

Not surprisingly, there is a clear relationship between the ease with which a stock can be quantified and its inclusion in the models. We were able to deduce this straightforward relationship from the inventory of provincial indicators. On the one hand, 'security' and 'social structure' are lost from the inventory and from the models. Their core characteristics are 'soft' and notoriously difficult to quantify. 'Population', on the other hand, is a stock with very high data availability and, as such, is prevalent in the models.

We also witnessed a clear relationship between the coverage of different stocks and the modelling methods used to forecast them. More dynamic modelling techniques, such as general equilibrium of system dynamics, allow for scenario-type, longer-term explorations and often provide insights into counterintuitive developments (Forrester, 1971). Integrated models of causal chains, as described above, require such dynamic techniques (Gilbert & Troitzsch, 1999). Parameter-based shift-share modules largely ignore these connections, give little insight into possible surprises, and only very rarely produce counterintuitive insights.

In our case, 'housing' and 'mobility' were modelled dynamically and included a large number of feedbacks and interactions. 'Population', 'work', and 'entrepreneurship' were modelled as exogenous drivers. General equilibrium modelling is the method of choice for the more sophisticated of these models. 'Consumption' and 'production', two stocks related to lifestyles and type of economic activity, are mostly covered as parameters in shift-share modules (e.g., translating economic activity into environmental impacts by assuming emissions per unit of production based on statistical analysis). All other stocks are covered in the form of calculated results of these simple shift-share procedures. This pattern is illustrated in Figure 7.4. A better integration of models would require coverage of a larger number of aspects using dynamic modelling techniques.

One might respond by noting that the modellers apply Occam's razor quite skilfully. This principle requires that any explanation of a phenomenon should rely on as few assumptions as possible. The modellers minimize the number of assumed relationships by 'shaving off' those with little statistical significance and thus reduce the system's complexity without significantly changing the dynamic properties. For the short run that might be true, but in the long run discontinuities in trends and other surprises are likely to result

from causes located virtually anywhere in the system. Strategic planning must not, by definition, exclude these causes. Qualitative aspects of the stocks are not considered and influences currently deemed weak are omitted, severely limiting the models' adequacy for exploring the consequences of novel events and developments. Occam's razor is not compatible with rich, exploratory, forward-looking studies.

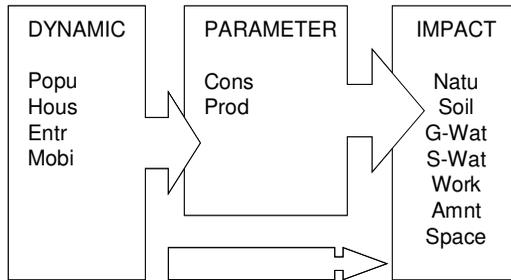


Figure 7.4: Stocks and predominant modelling approaches

However, the fact that flows with high strength estimates are generally incorporated into the models indicates that the modellers are judicious in their application of Occam's razor. For pragmatic business-as-usual scenarios they do appear to set the right priorities. Just as the participants tended to give a higher strength estimate to flows subject to little uncertainty, volatility, and discussion, modellers focus on easy-to-measure and relatively stable flows.

7.3 Models and joined-up thinking

Based on our analysis of the conceptual model and the corresponding model inventory, we formulated three recommendations for the provincial administration on how to better use its available models and how to efficiently improve the overall set.

- First, a better understanding of the structure of the models is imperative for better use of their results. Our primary advice was to require disclosure of the parameter values whenever a model was

commissioned. This small amount of extra insight would enable public officials to place a particular model in its proper context and to improve interpretation of its output in terms of validity and robustness. Also, simple scenario exercises that introduced different change rates—or even discontinuities—into the development of the parameters over time would create opportunities for transparency. Different futures can thus be explored in a rule-of-thumb manner. The results could also illustrate narrative scenarios of various futures.

- Second, extension of the individual models would be an inefficient way to improve the integrated dynamics of the whole set. Harmonization of the underlying assumptions (and thus the possibility to vary them with full intention) would be a much more important step.
- Finally, a fully functional integrated model with a sufficient level of detail is indeed possible. A model of this kind would, however, require considerable resources to implement while its comparability with the model runs of other provinces would be lost.

The first step in bridging the discrepancies between qualitative and quantitative analyses of complex systems must be a better understanding of the potential and the limits of both approaches.

Quantitative models have often been critically appraised, not in the least by people with no practical modelling experience of their own. Godet (1983) criticized dependency on inaccurate data, coupled with unstable models and the explanation of the future in terms of the past, as an inherent property of statistical models. In response, he called for a global, qualitative approach. The ensuing two decades have seen no fundamental progress in this respect. *‘Mathematical modelling is only possible if one is willing to except important parts of the problem and to limit an originally comprehensive understanding of the problem to the computable part of the problem’* (Weimer-Jehle, 2006). Also, uncertainties are notoriously difficult to represent in quantitative studies that suggest precision and objectivity (Jaeger et al. (1998); (2001)). Modelling results have little value once deprived of their underlying assumptions and context.

During a group interview in August 2007, we explored the consequences of the combined SCSNE/QSA approach for the way in which the process participants used models and model results. We found that the insights gained by provincial administration officials during the project helped them

to make more appropriate use of the model results and better embed the models in the decision making process. An understanding, of which factors were accounted for in the different studies, as well as an appreciation of the implications of different methods, helped the participants to use model results as arguments rather than indisputable facts. For new quantitative prognoses models are now selected and applied with much more care. One of the consequences of this is that the differences between regional and national expectations have increased, pointing towards more awareness for region-specific factors.

And even without an exact understanding of the dynamics, mapping the province in the form of an influence diagram facilitated horizontal integration in organizational decision-making. Civil servants from different provincial departments are now more likely to contact colleagues from other administrative units when their decisions might generate ripple effects and their mutual comprehension has improved.

Vester (2002) provides anecdotal evidence of similar effects for a series of case studies: a working group on systems thinking set up in Bad Aibling (which continued to be active after his intervention), an altered structure for a regional planning group in Frankfurt, more integrated project perceptions at NERIS (Netzwerk in der Risiko-Sensitivitätsanalyse—Network in Risk and Sensitivity Analysis).

This experience confirms observations in similar projects. In reflecting on the ULYSSES project on urban lifestyles and sustainability, De Marchi et al. (1998) describe indications of mutual learning where the participatory processes included a somewhat odd 'Mr. Computer'. The perception of the computer as a group member indicates the experience of mutual learning among human participants, the computer (model) and the modellers. Siebenhüner & Barth (2005) evaluate three major integrated assessment projects that embedded models in their participatory approaches: the ULYSSES project described above, the COOL project (focused on long-term climate options), and the VISIONS project (focused on developing integrated visions for a sustainable Europe). Even though these processes were mainly concerned with learning from models rather than about models, the authors indicate that some models *served as a trigger for debates about the implied uncertainties*. Unfortunately, empirical studies on the effects of participation, the drafting of system diagrams, and the implications of learning about factual information—including model structures—are few and far between

and quality criteria for participatory processes are disputed and generally hard to test (van de Kerkhof, 2004).

The direct path towards more comprehensive coverage in the sense of the cause-effect cycle described above would be to extend existing models to capture a larger number of dynamics. Such an effort would shift some stocks in Figure 7.4 from the right side to the left side, from impact and parameter status toward becoming integrated elements of the dynamic model. However, in many cases this would be mere window dressing as these models have been developed with a more focused target in mind. Toth (2003) warns modellers that they

should not extend their tools far beyond their original objectives because this might jeopardize the internal consistency and integrity of the [model].

Every flow added to a model exponentially multiplies the number of indirect influences and feedback cycles. This increase in complexity limits the dynamics that can be modelled without losing an overall understanding of the system's behaviour.

Many model extensions would also require more data. The most far-reaching project in this respect is probably the Sustainability Geoscope that attempts to be an observation instrument for the anthropocene, the era in which natural and human dimensions of the earth system have become inseparable (see <http://www.sustainability-geoscope.net>). But it remains to be seen whether these additional data are actually useful for modelling purposes when there are breaks in the trends and the meaning of indicators changes. Time-series analysis is problematic under such circumstances. The most ambitious option is to develop better ways to integrate qualitative and quantitative analyses. Such methods would combine some of the rigor of quantitative analysis with some of the richness of qualitative analysis. Numerous scholars have been calling for years for the development of such tools (Rotmans (1998b); Kates et al. (2001); Toth (2003); Robinson (2004)). We believe that the QSA approach and its application in highly practical cases is a step in that direction. It increases the mutual understanding between modellers and model-users in the widest sense as well as providing a methodological stepping-stone between quantitative and qualitative ways of analysing systems.

7.4 Modelling experiments

So far, our evaluation of existing models for integrated sustainability assessment is sobering. Models can only answer very specific questions when it comes to sustainability and even the answers to these are not always satisfying. This might be due to the fact that all models listed above are embedded in the current regime. Such models are unlikely to provide us with the right insights to manage what we know must be a radical change of regime. The relevant questions concerning the dynamically and rapidly evolving system structures of a sustainability transition lie outside the scope of the models used during the past decades. In response to this dissatisfaction, a series of modelling experiments are currently underway, many of them with a link to transition management.¹³

Geels & Schot (2005) and (2007) provide a taxonomy of the type of regime change, relevant in this context. They distinguish transformations, technological substitutions, de-alignment and re-alignment, and reconfiguration. These differ in the way different levels interact in the process of the transition and in terms of the timing of the interactions.

Models of radical change or regime change developed in this context are not predictive, but explorative. The models become metaphors rather than truth machines (Ravetz, 2003). Such a change in the role of models also changes the criteria for their suitability, or at least the relative importance of the different criteria. They emphasize the capacity to explore interactions and feedbacks that are beyond the scope of traditional models. Once developed, these models allow flexible and fast runs, often in combination with a Monte Carlo technique for attributing probabilities to different outcomes. The results they produce can usually be narratively communicated, describing patterns and dynamics, rather than numerical values. While the outcomes might sometimes seem complex, the rules underlying these models are often surprisingly simple and transparent. The focus of these models often lies on the individual stakeholder in interaction with other stakeholders rather than an assumed system structure. Surprises, uncertainties and random behaviour are frequently the target of exploration, rather than disturbing factors.

¹³ Pieter Valkering and Alex Haxeltine provided valuable input for this section.

Most of these experiments involve elements of agent based modelling. This is not surprising as being able to capture the dynamics between different stakeholders is a crucial prerequisite for the ability to model regime changes. In chapter 3, we shortly mentioned agent-based models as tools to endogenously represent stakeholder behaviour and the evolution and organisation of system structures driven by stakeholders. Agent based models can be rather freely combined with other methods.

Pahl-Wostl (2002) emphasizes the benefits of combining agent based models with participatory methods. Stakeholders would supply relevant issues, as well as subjective and normative perceptions. Also, agent based models can help to prepare and structure stakeholder processes. Valkering et al. (2004) have used agent based models in this way in a case of river management. Van der Veen & Rotmans (2001) have proposed the use of agent based models in the context of land-use in a way that is analogous to the implementation by Valkering et al.

Bergman et al. (in prep.) combine agent based models with system dynamics and propose applications to the transport, hydrogen fuel, and housing sectors. The model prototype is comparatively complex, representing the macro-level as a landscape, different agents and their interactions with a complex internal and external structure, as well as external signals that are designed to represent the different triggers for different types of transitions.

The attraction of agent-based models lies in their capacity to represent actors and their behaviour in the context of sustainability. Nevertheless, as all modelling techniques, agent-based models have their shortcomings. They often deal with homogenous, simple, individual agents rather than with heterogeneous, complex and collective agents. Also, they need to be integrated with physical models in order to develop agent-physical models with relevance for sustainability.

Very early experiments are also conducted with evolutionary algorithms. In these cases, fitness landscapes represent degrees of sustainability in the form of a multi-criteria evaluation and different strategies 'travel' through them. We are not aware of any of these experiments having transcended the early experimental stage or having been published.

The experiments described and the ambitions underlying them are the symptoms of a significant bifurcation in the development of models. While the majority of researchers tries to twist and tweak tested modelling

techniques to achieve progress, another group considers these attempts as a dead end for the reasons described above and focuses on the utilisation of experimental modelling techniques in order to answer new types of questions. Letting go of the ambition to be predictive opens up a wide playing field for the development of explorative tools. However, these are unlikely to fit traditional policy processes, which are set in the current regime. Proponents of these models therefore often call for much wider societal change, including new types of governance. The change in worldview can be described as a change from plan, control and management to one of self-evolution, where the role of the government is restricted to supplying a fertile soil for that self-evolution without controlling or steering it.

7.5 Conclusion

The all-encompassing sustainability model is a utopian idea. Combining comprehensive depth, breadth and policy relevance in quantitative sustainability models seems impossible. However, models focusing on one or two of these dimensions have in several cases been of great value within the scientific community and beyond. This has especially been the case when the context and the scope of a model was well defined and modellers and model users were aware of what the underlying assumptions implied for the interpretation of modelling results.

The eQSA method can be used as a tool to explore this context and scope in great detail, directly compare it with the structure of quantitative models, and to stimulate a learning process. In our case study, we have, not surprisingly, found a significant loss of information when moving from qualitative analysis to quantitative modelling. There was a strong bias with respect to which types of information were lost: Social aspects, very fast and very slowly moving factors and uncertain or instable relationships. The resulting quantitative system turned out to suggest much higher systems stability than a full mental map of the same region.

A highly detailed and structured analysis of which information was lost and how the quantitative models treated the remaining items was applied as an initial step to map and bridge the gap between conceptual and quantitative analysis. While this does at this stage not encompass any technical improvement in the models, there is evidence that this analysis has

significantly altered the way in which model runs and model results are executed and used by the administration of the province of Limburg. There was a tangible learning process ranging from a better understanding of models to more attention dedicated to the interpretation of model results in the context of a complex environment.

The case has also provided a list of advice on how to improve current models and approach the development of future models. At this stage, this advice has not been implemented in any significant manner. There are several possible reasons for this. The development of models is resource intensive and a shift from one type of models to another is subject to high transition costs for the whole organisation. The time lag between the actual study and the evaluation might have been too short to show an ongoing shift. The improvements in understanding and interpretation combined with the qualitative tools supplied in the shape of QSA might have reduced the pressure to redevelop the set of models used. Finally, our results might not be sufficiently specific to trigger such a change.

Generally, whether a loss of information outweighs a gain in analytical possibilities must be determined for each case individually. With the methods presented in this and previous chapters, we have tried to ease that trade-off by increasing the capacity to analyse qualitative information in a structured manner. Instead of having to choose between knowing a little of the whole or a whole lot of a little, we have tried to embed the detailed and highly specific information in its broader context, while at the same time keeping it highly accessible to the users. There is evidence that our efforts have been successful.

The structure applied has its main advantage in its high compatibility with the different types of qualitative information produced in engaged participatory processes and with quantitative models. However, this structure also reinforces a perspective of sustainability in which actors are peripheral and the future is relatively malleable. This and many other aspects in combining qualitative and quantitative approaches will remain a challenge for the future. Neither methodological approach can be expected to meet the other party on its own terms. Policymakers are caught between the two as numbers and graphs provide them with authority and suggest objectivity while narratives and the simplification of complex issues are the butter and bread of politics.

Experiments to resolve these problems are being conducted. Models are increasingly explorative and open to new types of dynamics. It remains to be seen how fast this branch will progress from clumsy first steps to a brisk walk.

Chapter 8

SYNTHESIS & CONCLUSION

In chapter 1 we formulate the following research question:

How can tools for sustainability assessment stimulate joined-up thinking in political decision processes?

The different chapters each provide elements of an answer to that question.

Chapter 3 would respond that joined-up thinking could be stimulated by taking a closer look at higher level leverage points, by increasing the dialectic capacities of tools for sustainability assessment and by increasing their comprehensiveness. Chapter 4 emphasizes the learning process involved. By making intangible knowledge and perceptions explicit in a structured manner, the systemic perspective can be shared, discussed and further developed. The tools described in chapter 5 facilitate the transitions from awareness and discussion to implementation. The ability to structurally analyse the complex societal systems in a transparent way and to answer policy relevant questions on leverage points, possible side-effects, and the like is reinforced by the seamless integration of the tool with existing organisational structures, reporting, and strategic activities. Chapter 6 triggers creative interventions in a system in order to make it more sustainable and thus deepens the perception of a system. Chapter 7 offers little advice on how simulation models can stimulate joined-up thinking, but rather helps to integrate existing models in an organisational transition towards more integration.

This fragmented answer to the research question does justice to the fact that the all-round and context independent sustainability assessment is an impossibility. No single sustainability assessment tool can fundamentally stimulate joined-up thinking; neither can a fragmented set of standardised tools. This dissertation therefore responds to the question through an exploration and extension of promising pathways. Each of the answers above highlights a different assessment perspective and its value for joined-up thinking. The range of perspectives addressed is not comprehensive and neither is our exploration of pathways to explore the perspectives. In addition, it is difficult to judge how well the approach taken can be transferred to other regions, to other scale levels, to other contexts.

Each chapter also provides an answer to at least one of the four sub-questions concerning the qualities of existing tools, possible extensions, the combination of different tools, and the use of assessment tools in policy processes:

1. How appropriate are existing assessment tools to stimulate joined-up thinking?

Chapter 3 elaborates extensively on this question. It is not only a tool as such, that is appropriate, but also the way in which it is being employed. All types of tools described in this dissertation are currently experimented with in the context of more joined-up analysis or thinking. Traditional applications of the same tools mostly fail the test for the ability to stimulate joined-up thinking. Relatively minor technical modifications combined with great attention for the appropriateness of a tool within a specific context often make a significant difference in terms of how appropriate a tool is. Chapter 6 explicitly illustrates this for indicators and indices, chapter 7 for simulation models.

2. How can existing tools be extended to improve their capacity to stimulate joined-up thinking?

The SCENE approach presented in chapter 4, the eQSA approach at the core of chapter 5, and the QSSI in chapter 6 are all extensions of or variations of existing tools. This reinterpretation of tools goes together with the change in the way tools are applied. A process of reframing is taking place, where hitherto rigid assumptions are being relaxed in order to open up pathways for innovative developments in new directions. Chapter 7 illustrates, that sometimes the tools will not need to change, but our interpretation of their results. This dissertation does not complete this process of reframing. The methods presented are still very much rooted in a traditional perspective on sustainability. This has the advantage of high compatibility with different types of existing information and tools. At the same time it limits the range of perspectives covered by the tools. For example, actors, power relationships, and the evolution of a system as time passes are only superficially looked at.

3. How can the integration among different sustainability assessment tools be improved?

Chapter 5 provides a technical example as an answer to this question. The eQSA approach forms a bridge between tools. The latent knowledge present in the organisation and the traditional processes taking place in that organisation are being reframed through the systemic amalgamation of qualitative information, indicators and models. The development of such amalgamating tools that are able to integrate a wide range of different types

of information in any manner is an under-explored field where in our perception much can still be gained. Chapter 3 also hints at the possibilities of combining different tools to that end. The SCENE approach (chapter 4) or similar other approaches that trigger learning of systemic understanding are a prerequisite to the successful implementation of such amalgamating tools in an organisation. However, there might be a more fundamental problem that this dissertation does not extensively deal with: co-operation on a broad range of tools among diverse groups is a substantial challenge. Many tools have their own scientific community with more competition than exchange between the communities and many participants and users prefer one tool to another, often for personal, political or intuitive reasons. This power play between different tools hinders their integration as much as technical boundaries do.

4. How can the integration of assessment tools in political decision-making be improved?

SCENE, eQSA, QSSI and our exploration of models currently used by the province administration all closely explicitly allow to follow existing organisational cultures in terms of structure and wording (while provoking radical systemic changes in actions). This is meant to stimulate the acceptance of these tools within a organisation. For different stakeholders, these tools provide either valuable output or valuable input to their own routines. The tools help to reframe their traditional work in a systemic context without having to change dramatically the ways of their work or the organisational structure.

If this process is successful, the change resulting from the reframing process would come from within the organisation by triggering the ambition and capacity to address higher-level leverage points. We have improved the dialectic capacities of the tools used in the process, or at least opened opportunities to easily do so. And we have improved the coverage of the tools, especially concerning soft and seemingly less important factors.

The answers to the research questions imply that our ambition to achieve a four-fold integration has been at achieved to some degree:

- Integration among policy domains
- Integration among scientific disciplines
- Integration of relevant research in policy processes

- Normative integration

The true test of this would be an analysis of whether the mindsets of the participants and the wider organisation have evolved towards more integration and how deeply embedded that evolution is. Unfortunately, such a study is beyond the scope of this dissertation and relevant evaluation methods are scarce.

We therefore rely on an overview of public documents published by the Province since the trajectory finished in 2004 and on a group interview evaluating the impact of the projects within the administration taken in July 2007. In the documents we find ample evidence for change:

- The sustainability strategy clearly reflects the long-term impact of the SCENE approach (Provincie Limburg, 2005). The SCENE triangle is prominent in the structure and the content of the document. The path towards sustainability is seen in the integration of the concept in the regular decision-process on all issues. The document indicates broad support for joined-up thinking within the organisation. In practical terms, the province installed one of the participants as co-ordinator for sustainability. His responsibility lies in the stimulation and facilitation of sustainability issues in all relevant processes. It is the position of a SCENE ambassador.
- The new integrated strategic plan (Provincie Limburg, 2006) draws extensively on the SCENE model and the eQSA analysis. The chosen policy priorities are consistent with the results of the analysis. The Province focuses on areas, where it has a strong influence. For other issues, co-operation with partners is aimed for. The document translates the broad support into strategic guidelines. The strategic plan is one of the most important documents published by the province. It serves as a guideline for all strategic decisions that are not explicitly treated elsewhere. It also is published and read widely, making it a benchmark against which stakeholders judge the decision taken. Inconsistencies between the strategic plan and actual decisions are hard to be taken without raising questions.
- The Limburgmonitor ('State of Limburg') collection of indicators (Provincie Limburg, 2007a) is fully based on the SCENE model. Where we executed the eQSA analysis on the level of stocks, this publication is based on an extensive system analysis of the attributes of the stocks (see

chapter 4). In contrast to earlier editions of the publication, the changes in indicators are highlighted, rather than their absolute values. This emphasizes the dynamic nature of sustainability. This way of representing the state of the province is unique in the Netherlands. The document proves the transposability of the analysis results to more detailed scale levels.

- The traffic and mobility plan (Provincie Limburg, 2007b) draws heavily on the SCENE model and the eQSA analysis, placing the issue in a detailed socio-cultural, economic, and ecological context. Specific flows from these models are extensively evaluated in the process of strategy development. The document indicates the success of vertical integration by translating the results to a specific topic. It also proves the relevance of the full integrative framework to topic- or sector-specific issues.

The combination of these documents points towards a successful integration of qualitative and quantitative, static and dynamic assessment tools, and their subsequent integration in the governance process. However, this selection of documents is subject to a positive bias. It could easily be extended with a list in which few or no traces of an evolution towards joined-up thinking are visible. The documents provide insights into the embedding of an integrated paradigm within the administration.

The group interview conducted with the process participants provides deeper insights into hidden changes within the organisation and personal learning experiences. During the interview, the different tools were discussed (with the exception of the QSSI, which was never applied for Limburg) and the personal experiences of the participants were shared.

A striking result is the degree to which the SCENE triangle is being used by the participants. For any new project they consider the contribution that an integrated approach might have, and stimulate such an approach, where they consider it appropriate. For themselves the participants would often execute little systems exercises, for example when delineating a problem. Unlearning the integrated perspective is hard to imagine for the participants. One participant states that as an analytical and structural tool SCENE has been *'a useful aid in the quest for the real problems, making it explicit, making choices and making these concrete.'*

This is also reflected in how the participants describe their personal use of SCENE:

In all my work I have always tried in one way or another to use the concepts and the [SCENE] model is the foundation for how we approach sustainability, especially in the last 9 months, where sustainability has become much more important.

If given a number of targets, I always try to find the relationships between them [with the SCENE model].

It is very much dependent on the type of project. [...] The triangle is not the bible, but it is important.

The SCENE triangle has also become a much used tool for communication:

If people ask me, what sustainability is about, I show them the triangle.

You can use it communicate the integrality to others, who are not familiar with the reports and that works very well. It makes people think.

With respect to the wider organisation, the participants largely confirmed the findings of the document overview. The results had been applied in a number of projects, often in a simplified or less stringent manner. In several internal group projects, in which 'our' participants and other staff of the administration took part, integrated approaches were selected. In this way, the knowledge reached a wider circle within the organisation. One such project is the Limburgmonitor mentioned above where, according to one participant 'the 60% who knew [the SCENE / eQSA approach] convinced the 40% who didn't know to use it as a basis.'

An indication for the quality of the model is the fact that, according to the participants, the structure of the stocks, flows, and attributes selected during the process is still fully intact. More recent projects have adapted to the SCENE model rather than the SCENE model adapting to new projects. On the basis of these statements, we have many reason to believe that the SCENE process and results have stimulated joined-up thinking. The impact of the other tools is more difficult to trace.

The participants statements concerning the eQSA tool imply that the process of developing the eQSA tool has had more impact than the tool itself. The most concrete application of the eQSA tool was the development of the traffic and mobility plan.

We have taken the eQSA framework and explicitly used it to look for relevant domains, stocks and attributes, that should be taken into account

in our strategic environmental assessment. And that assessment was the extended beyond the environmental domain.

A shortcoming of the tool is its perceived complexity:

The tools seems terribly complicated and is not being used. [...] There are so many dimensions (speed, strength, direction, influence). That can't be handled.

Once you use it, all the uncertainties come up with the consequence that the tool is quickly contested.

The complexity of the issues, which is reflected in the complexity of the tools provided, does hinder their frequent thorough application.

Translating the outcomes of the analysis to the level of detail required form most governance documents remains a challenge for the participants.

The same statements also indicate, that the tool succeeds in provoking and structuring discussions on relevant questions. In addition, concrete applications of the tool might be realised in the future. One participant mentions that *'we have more than we think and we could use that more often'*. During the discussion several ideas for new applications were brought up:

We could put the coalition programme into the tool.

We could analyse larger, more strategic questions, for example the consequences of a shrinking economy.

We could integrate the tool and a cost-benefit analysis to better evaluate investments.

With respect to quantitative modelling, a learning process has taken place, though few steps have been taken to implement changes in the models. One participant described the impact of the in depth model analysis in the following words:

The combination of quantitative models and the triangular thinking has nuanced the discussion on the level of detail included in models. We have recently put more thought into the positioning of models [...]. We do not simply use any model, just because it is available.

This discussion has also provided input for the discussion on the significant differences between the future expectations calculated by the Central Planning Bureau and expectations at the provincial level.

Even though, some of the participants' statements might make it seem otherwise, it is not possible to directly attribute the impacts and changes to the projects described in this dissertation. Many other developments, including a rising national awareness for sustainability worked in favour of integration. Despite this and based on the evidence provided we consider the project to have been quite successful in stimulating joined-up thinking. This process does not yet seem irreversible, especially at the organisational level, but with additional opportunities to experience the added value of the tools a lasting effect seems likely. On a personal level, the participants seem to be irreversibly infected with a joined-up perspective, though that might not always be tangible in their daily routines.

It should be noted that the approaches and processes described in this dissertation are still in the early stages of development, as is the general development of tools for integrated sustainability assessment. More applications and cases will be required to test assumptions and judge the transferability of processes and results. As in any young field, the need for future research is extensive. Some of these needs are listed below and conclude this dissertation.

(Chapter 3) Generally accepted quality criteria for integrated sustainability tools need to be developed and studies assessing tools and processes according to these criteria must take place.

(Chapter 4) Processes of participation rely on many assumptions, that have until now not been empirically tested in a satisfactory manner. The influence of, among others, group facilitators, participants, settings, and processes need to be studied in detail in order to consolidate the field to some extent. System diagrams are at the basis of many integrative tools. The appropriateness of these diagrams and the processes they result from tools for learning about systems has not been studied extensively. On the contrary, Sterman (1994) even documents a lack of the ability to learn about systems without dynamic simulation in specific cases.

(Chapter 5) Even though related techniques and processes have been toyed with since the 1970's, the eQSA approach and similar techniques still offer many obvious extensions, especially when combined with other methods by

design. Especially the integration with more recent modelling techniques might provide ways in which the analysis of the underlying systemic structure could be reflected upon more broadly and effectively.

(Chapter 6) The QSSI still awaits its true proof of concept. An application of the technique in a fully developed case study would be the first step towards the extended development of systemic indicators along the lines of the QSSI. Other methods for integrating systemic information in indicator studies must be explored as well.

(Chapter 7) Traditional models seem to be in a phase of small incremental improvements. While researching these might be valuable in the very short term, alternative techniques will certainly be needed in the future. These alternatives will follow a transformative, rather than incremental paradigm. The earlier these are implemented in many different cases, the earlier will it be possible to weed out what will remain at the experimental stage and develop the future standards of explorative models.

(This dissertation) The process of integration of different domains of knowledge is far from complete. One example for a frequently omitted domain is the absence of legal expertise in most integrated sustainability assessments. Integrating the discourse in the legal domain on the rights to participation with respect to sustainability with the experience with participation in integrated sustainability assessment bears much potential for better decision making and improved methods to support the sustainability transition.¹⁴

An important problem with the intensely collaborative and explorative work in this dissertation and similar projects elsewhere is the inability to concretely document the concrete impacts. The field would greatly benefit from efforts to develop societal impact assessments documenting the actual effects of a given project as concisely as possible.¹⁵

¹⁴ This point was brought to my attention by Marjan Peeters in reaction to an earlier version of this dissertation.

¹⁵ This point and the term 'societal impact assessment' were drawn from comments made by John Robinson on an earlier version of this dissertation.

REFERENCES

- 4cast. (2006). OGM. Retrieved 30 October 2006 from <http://www.4-cast.nl/default.php?id=1028>
- ABFresearch. (2006a). IPB. Retrieved 30 October 2006, from <http://www.abfresearch.nl/default.asp?r=58>
- ABFresearch. (2006b). PRIMOS. Retrieved 30 October 2006, from <http://www.abfresearch.nl/default.asp?p=33>
- ABFresearch. (2006c). RDP. Retrieved 30 October 2006, from <http://www.abfresearch.nl/default.asp?r=64>
- Ackoff, R. L. (1974). *Redesigning the Future: A Systems Approach to Societal Problems*. New York: John Wiley & Sons.
- Alcamo, J. (Ed.). (1994). *IMAGE 2.0: Integrated Modeling of Global Climate Change*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Alcamo, J., Shaw, R., & Hordijk, L. (Eds.). (1990). *The RAINS Model of Acidification: Science and Strategies in Europe*. Dordrecht: Kluwer Academic Publishers.
- Allen, P., Malchow, H., Matthies, M., & Kriz, J. (2001). *Integrative Systems Approaches to Natural and Social Dynamics*: Springer.
- Amelung, B. (2006). *Global (Environmental) Change and Tourism - Issues of scale and distribution*. Maastricht: Amelung Publishers.
- Arcade, J., Godet, M., Meunier, F., & Roubelat, F. (2003). Structural analysis with the MICMAC method & Actors' strategy with MACTOR method. In J. C. Glenn & T. J. Gordon (Eds.), *Futures Research Methodology - Version 2.0* The American Council for the United Nations University - The Millennium Project.
- Arcadis. (2004). Rapid Assessment Program (RAP). version 1.9.
- Arnstein, S. (1969). A Ladder of Citizen Participation. *JAIP*, 35(4), 216-224.

- Bakkes, J. A., Grosskurth, J., Idenburg, A. M., Rothman, D. S., & van Vuuren, D. P. (2000). *Descriptions of selected global models for scenario studies on environmentally sustainable development : Global Dynamics and Sustainable Development Programme*. Bilthoven: RIVM.
- Barker, T., & Koehler, J. (1998). Equity and Ecotax Reform in the EU: Achieving a 10 per cent Reduction in CO₂ Emissions Using Excise Duties. *Fiscal Studies*, 19(4), 375-402.
- BCI. (1998). *Marktanalyse Bedrijventerreinen Zuid-Limburg*. Nijmegen: Buck Consultants International.
- Bergman, N., Haxeltine, A., Whitmarsh, L., Köhler, J., Schilperoord, M., & Rotmans, J. (in prep.). Modelling Transition Patterns and Pathways.
- Bertalanffy, L. v. (1968). *General Systems Theory: Foundations, Development, Applications*. New York, USA: Braziller.
- Biermann, F. (2007). 'Earth system governance' as a crosscutting theme of global change research. *Global Environmental Change* (doi: 10.1016/j.gloenvcha.2006.11.010).
- Bossel, H. (1996). Deriving indicators of sustainable development. *Environmental Modeling & Assessment*, 1(4), 193-218.
- Bossel, H. (1999). *Indicators for Sustainable Development: Theory, Method, Applications*. Winnipeg: International Institute for Sustainable Development.
- Bossel, H. (2004). *Systeme, Dynamik, Simulation - Modellbildung, Analyse und Simulation komplexer Systeme*. Norderstedt: Books on Demand.
- Boulanger, P.-M. (unpublished). *Integration in Sustainability Impact Assessment: Meaning, Patterns and Tools*. Ottignies: Institut pour un Développement Durable.
- Bouwman, A. F., Kram, T., & Klein Goldewijk, K. (Eds.). (2006). *Integrated modelling of global environmental change: an overview of IMAGE 2.4*. Bilthoven: Netherlands Environmental Assessment Agency (MNP).
- Capros, P., Georgapopoulos, P., & Zografakis, S. (1996). Double Dividend Analysis: First results of a General Equilibrium Model (GEM-E3). In C. Carraro & S. Siniscalco (Eds.), *Environmental Fiscal Reform and Unemployment*. Dordrecht: Kluwer.

- Carmichael, J., Tansey, J., & Robinson, J. (2004). An integrated assessment modelling tool. *Global Environmental Change*, 14, 171-183.
- Carson, R. (1962). *Silent Spring*. London: Penguin.
- Caswill, C., & Shove, E. (2000). Introducing interactive social science. *Science and Public Policy*, 27(3), 154-157.
- Caulkin, S. (2006). Why things fell apart for joined-up thinking. Retrieved 10. July 2007, from <http://observer.guardian.co.uk/business/story/0,,1717898,00.html>
- CEC. (1995). *E3ME: An Energy-Environment-Economy Model for Europe* (No. EUR 16715 EN). Luxemburg: Office for Official Publications of the European Communities.
- Checkland, P. (1981). *Systems Practice, Systems Thinking*. New York, USA: John Wiley and Sons.
- Checkland, P., & Scholes, J. (1999). *Soft Systems Methodology in Action*. Chichester: John Wiley & Sons.
- Chertow, R. M. (2001). The IPAT Equation and Its Variants; Changing Views of Technology and Environmental Impact. *Journal of Industrial Ecology*, 4(4), 13-29.
- CIESIN. (1995). IMAGE 2.0 Model Guide [online]. Retrieved 25-7-2000, 2000, from <http://sedac.ciesin.org/mva/image-2.0/image-2.0-toc.html>
- Clayton, A., & Radcliffe, N. (1996). *Sustainability: A Systems Approach: Earthscan*.
- Commoner, B. (1972). The Environmental Cost of Economic Growth. In *Population, Resources and the Environment* (pp. 339-363). Washington, DC: Government Printing Office.
- Costanza, R., d' Arge, R., de Groot, R. S., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., & O'Neill, R. V. (1997). The Value of the World's Ecosystem Services and Natural Capital. *Nature*, 387(6630), 253-260.
- Costanza, R., Graumlich, L. J., & Steffen, W. (Eds.). (2007). *Sustainability or collapse? An integrated history and future of people on earth*. Cambridge, MA: Massachusetts Institute of Technology.

- CPB. (1996). *Surroundings scenarios: Long term exploration 1995-2020* (No. 89). The Hague.
- CPB. (1999). *WorldScan: the Core version*. The Hague: CPB.
- Crutzen, P. J., & Stoermer, E. F. (2000). The "Anthropocene". *Global Change Newsletter*(41), 17-18.
- Dahl, A. L. (2007). Integrated Assessment and Indicators. In T. Hák, B. Moldan & A. L. Dahl (Eds.), *Sustainability Indicators - A Scientific Assessment*. Washington: Island Press.
- Daly, H. E., & Cobb, J. B. (1994). *For the Common Good* (2nd ed.). Boston: Beacon Press.
- De Marchi, B., Funtowicz, S. O., Gough, C., Guimarães Pereira, Ã., & Rota, E. (1998). *The ULYSSES Voyage: The ULYSSES Project at the JRC*. (No. EUR 17760EN). Ispra: Joint Research Centre, European Commission.
- de Ridder, W. (2006). *Tool use in integrated assessments - Integration and synthesis report for the SustainabilityA-Test project*. Bilthoven: Netherlands Environmental Assessment Agency (MNP).
- de Ridder, W., Turnpenney, J., Nilsson, M., & von Raggamby, A. (2007). A framework for tool selection and use in integrated assessment for sustainable development. *Journal of Environmental Assessment, Policy and Management*, 9(4), 423-441.
- Descartes, R. (about 1619). *Rules for the Direction of the Mind* (E. S. Haldane & G. R. T. Ross, Trans.). Chicago: Encyclopaedia Britannica.
- Diener, E. (2000). Subjective well-being. The science of happiness and a proposal for a national index. *The American Psychologist*, 55(1), 34-43.
- Dirven, J., Rotmans, J., & Verkaik, A. (2002). *Samenleving in transitie: Een vernieuwend gezichtspunt*. Den Haag: Innovatienetwerk Agrocluster en Groene Ruimte.
- Dörner, D. (2003). *Die Logik des Misslingens: Strategisches Denken in komplexen Situationen*. Hamburg: rororo.
- Dunn, W. N. (2003). *Public Policy Analysis: an introduction*. (3rd ed.): Pearson-Prentice Hall.

- Duperrin, J. C., & Godet, M. (1973). *Méthode de hiérarchisation des éléments d'un système*. Paris: Rapport Economique du CEA No. R-45-41.
- Edmonds, J. M., Wise, M., Pitcher, H., Richels, R., Wigley, T., & MacCracken, C. (1996). An integrated assessment of climate change and the accelerated introduction of advanced energy technologies: An application of MiniCAM 1.0. *Mitigation and Adaptation Strategies for Global Change*(1(4)), 311-339.
- Edwards, A. R. (2005). *The Sustainability Revolution: Portrait of a Paradigm Shift*: New Society Publishers.
- Ehrlich, P. R. (1968). *The Population Bomb*. New York: Ballantine.
- Engelen, G., White, R., & de Nijs, T. (2003). Environment Explorer: a Spatial Support System for the Integrated Assessment of Socio-Economic and Environmental Policies in the Netherlands. *Integrated Assessment*, 4(2), 97-105.
- ETIL. (2000). *Harmonisatie Prognoses POL, Onderdeel: beoordeling gebruikte prognoses en voorstellen nieuwe prognoses*. Maastricht: ETIL.
- Fiksel, J. (2006). Sustainability and resilience: toward a systems approach. *Sustainability: Science, Practice & Policy*, 2(2), 14-21.
- Finnveden, G., Nilsson, M., Johansson, J., Persson, Å., Moberg, Å., & Carlsson, T. (2003). Strategic Environmental Assessment Methodologies - Applications within the Energy Sector. *Environmental Impact Assessment Review*, 23(1), 91-123.
- Flyvbjerg, B. (2004). Five misunderstandings about case-study research. In C. Seale, G. Gobo, J. F. Gubrium & D. Silverman (Eds.), *Qualitative Research Practice* (pp. 420-434). London and Thousand Oaks, CA: Sage.
- Forrester, J. W. (1971). Counterintuitive Behavior of Social Systems. *Theory and Decision*, 2(2), 109-140.
- Funtowicz, S. O., & Ravetz, J. R. (1993). Science for the Post-Normal Age. *Futures*, 25(7), 739-755.
- Gallopin, G. (1996). Environmental and sustainability indicators and the concept of situational indicators as a cost-effective approach. *Environmental Modelling and Assessment*, 1(3), 101-117.
- Gallopin, G., Hammond, A., Raskin, P., & Swart, R. (1997). *Branch Points: Global Scenarios and Human Choice - A Resource Paper*

- of the Global Scenario Group (No. 7): Stockholm Environment Institute, Sweden.
- Geels, F., & Schot, J. (2005). *Taxonomy of transitions pathways in socio-technical systems*. Paper presented at the Paper presented at workshop by the ESRC Sustainable Technologies Program, May 12, 2005, London.
- Geels, F., & Schot, J. (2007). Typology of sociotechnical transition pathways: Refinements and elaborations of the multi-level perspective. *Research Policy*, 36(3), 399-417.
- Gibbons, M. (2000). Mode 2 society and the emergence of context-sensitive science. *Science and Public Policy*, 27(3), 159-163.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. New Delhi: SAGE.
- Gilbert, N., & Troitzsch, K. G. (1999). *Simulation for the Social Scientist*. Buckingham: Open University Press.
- Glenn, J. C., & Gordon, T. J. (Eds.). (2003). *Futures Research Methodology - Version 2.0: The American Council for the United Nations University*.
- Godet, M. (1983). Reducing the blunders in forecasting. *Futures*, 15(3), 181-192.
- Godet, M. (1986). Introduction to La Prospective. *Futures*, 18(2), 134-157.
- Greeuw, S. C. H., van Asselt, M. B. A., Grosskurth, J., Storms, C. A. M. H., Rijkens-Klomp, N., Rothman, D. S., & Rotmans, J. (2000). *Cloudy crystal balls : An assessment of recent European and global scenario studies and models*. Copenhagen, Denmark : European Environmental Agency (EEA) Experts' corner report no. 4 : Prospects and scenarios.
- Grosskurth, J. (2007). Ambition and Reality in Modeling: a case study on public planning for regional sustainability. *Sustainability: Science, Practice, & Policy*, 3(1), 3-11.
- Grosskurth, J. (in preparation). enriched Qualitative Systems Analysis.
- Grosskurth, J., de Ridder, W., Martens, P., & Rotmans, J. (in preparation). Scale in selecting sustainability assessment tools.

- Grosskurth, J., & Rotmans, J. (2005). The SCENE model: getting a grip on sustainable development in policy making. *Environment, Development and Sustainability*, 7(1), 135-151.
- Grosskurth, J., & Rotmans, J. (2007). QSSI: A New Type of Sustainability Indicator. In T. Hák, B. Moldan & A. L. Dahl (Eds.), *Sustainability Indicators - A Scientific Assessment* (Vol. 67, pp. 193-210): Island Press.
- Haberl, H., Erb, K. H., Krausmann, F., Gaube, V., Bondeau, A., Plutzer, C., Gingrich, S., Lucht, W., & Fischer-Kowalski, M. (2007). Quantifying and mapping the human appropriation of net primary production in earth's terrestrial ecosystems. *PNAS*, *Early Edition* (www.pnas.org/cgi/doi/10.1073/pnas.0704243104), 1-6.
- Hák, T., Moldan, B., & Dahl, A. L. (Eds.). (2007). *Sustainability Indicators - A scientific assessment* (Vol. 67). Washington, DC: Island Press.
- Hawken, P. (2007). *Blessed Unrest, How the Largest Movement in the World Came into Being and Why No One Saw It Coming*. New York: Viking Press.
- Hawken, P., Lovins, A., & Lovins, L. (2000). *Natural Capitalism: The Next Industrial Revolution*: Earthscan.
- Heida, H., & Poulus, C. (1993). *COMBI: Een combinatie van modellen*. Delft: Focus Onderzoek & Advies.
- Held, D. (1999). *Global Transformations*. Stanford, CA: Stanford University Press.
- Hertel, T. W., & Tsigas, M. E. (1997). Structure of GTAP. In T. W. Hertel (Ed.), *Global Trade Analysis: Modeling and Applications*. Cambridge: Cambridge University Press.
- Hisschemöller, M., & Hoppe, R. (1996). Coping with Intractable Controversies: the case for problem structuring in policy design and analysis. *Knowledge and Policy*, 8(4), 40-60.
- Hughes, B. B. (1999). *International Futures: Choices in the face of uncertainty* (3rd ed.). Oxford, UK: Westview Press.
- IMF. (2007). *World Economic Outlook: Spillovers and Cycles in the Global Economy*. Washington, D.C.: International Monetary Fund.
- Inayatullah, S. (2002). Reductionism or layered complexity? Futures of futures studies. *Futures*, 34(3-4), 295-302.

- IPCC. (2000). *Emissions Scenarios*. Cambridge: Cambridge University Press.
- IPCC. (2001). *Special Report on Emissions Scenarios: a special report of Working Group III of the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge University Press.
- Jaeger, C. C., Renn, O., Rosa, E. A., & Webler, T. (1998). Decision analysis and rational action. In S. Rayner & E. Malone (Eds.), *Human Choice and Climate Change* (Vol. 3, pp. 141-215). Columbus, OH: Battelle Press.
- Jaeger, C. C., Renn, O., Rosa, E. A., & Webler, T. (2001). *Risk, Uncertainty, and Rational Action*. London: James & James, Earthscan.
- Kahn, H., & Wiener, A. (1967). *The Year 2000*. New York: MacMillan.
- Kämper, E. (1998). *Sociology of risk: Implications for the analysis of environmental policy* (<http://www.iue.it/WGES/ISS17/kaemper.h>). Florence: European University Institute.
- Kane, J. (1972). A Primer for a New Cross-Impact Language - KSIM. *Technological Forecasting and Social Change*, 4.
- Kasemir, B., Jäger, J., Jaeger, C. C., & Gardner, M. T. (Eds.). (2003). *Public Participation in Sustainability Science*. Cambridge: Cambridge University Press.
- Kates, R. W., Clark, W. C., Corell, R., Hall, J. M., Jaeger, C. C., Lowe, I., McCarthy, J. J., Schellnhuber, H. J., Bolin, B., Dickson, N. M., Faucheux, S., Gallopin, G. C., Grubler, A., Huntley, B., Jager, J., Jodha, N. S., Kasperson, R. E., Mabogunje, A., Matson, P., Mooney, H., Moore, B. I., O'Riordan, T., & Svedlin, U. (2001). Sustainability Science. *Science*, 292(5517), 641-642.
- Kemp, R., & Martens, P. (2007). Sustainable development: how to manage something that is subjective and never can be achieved? *Sustainability: Science, Practice & Policy*, 3(2), 1-10.
- Knippenberg, L., Beckers, T., Haarmann, W., Hermans, F., Dagevos, J., & Overeem, I. (2007). Developing Tools for the Assessment of Sustainable Development in the Province of Brabant, the Netherlands. In T. Hak, B. Moldan & A. L. Dahl (Eds.), *Sustainability Indicators: A Scientific Assessment* (pp. 309-328). Washington, DC: Island Press.

- Kockelkorn, G., van Asselt, M. B. A., & Rotmans, J. (1999). *Toekomst op koers : Stadsvisie en planningsinstrument voor Maastricht op weg naar 2030* (Plan van aanpak). Maastricht: ICIS BV.
- Kuhn, T. S. (1970). *The Structure of Scientific Revolutions*. Chicago, USA: University of Chicago Press.
- Kuipers, B. (1994). *Qualitative Reasoning: Modeling and Simulation with incomplete knowledge*. Cambridge, MA: MIT Press.
- Lashof, D. A., & Tirpak, D. A. (1989). *Policy Options for Stabilising Global Climate*. Washington, USA: US Environmental Protection Agency.
- Lee, N. (2006). Bridging the gap between theory and practice in integrated assessment. *Environmental Impact Assessment Review*, 26, 57-78.
- Leeuwis, C. (2003). *Van strijdttonelen en luchtkastelen*. Twente: Communicatie en Innovatie Studies.
- Lindstone, H., & Turoff, M. (Eds.). (1975). *The Delphi Method*. Boston: Addison Wesley
- Loeber, A. (2004). *Practical wisdom in the risk society*. Amsterdam: Universiteit van Amsterdam.
- Loorbach, D. (2005). Transitions in Waste Policies in the Netherlands and Flanders. In P. Hjerp, C. Monkhouse & A. Farmer (Eds.), *EFIEA Workshop on Improving Waste Policy Through Integrated Environmental Assessment, 6 - 7 December 2004, Scotland House*. Brussels.
- Loorbach, D. (2007). *Transition Management - new mode of governance for sustainable development*. Utrecht: International Books.
- Loorbach, D., & Rotmans, J. (2006). Managing transitions for sustainable development. In X. Olshoorn & A. J. Wiczorek (Eds.), *Understanding Industrial Transformation. Views from different disciplines*. Dordrecht: Springer.
- Lyall, C., & Tait, J. (2005). Shifting Policy Debates and the Implications for Governance: Developing an Integrated Policy Approach to Science, Technology, Risk and the Environment. In C. Lyall & J. Tait (Eds.), *New Modes of Governance* (pp. 1-18). Hampshire: Ashgate.
- Maddison, A. (1995). *Monitoring the world economy 1820-1992*. Paris: Development Centre of the Organization for Economic Co-operation and Development.

- Martens, P. (2005). Sustainability: Science or Fiction? (inaugural speech). Maastricht: Universiteit Maastricht.
- Martens, P., & Rotmans, J. (Eds.). (2003). *Transitions in a Globalising World*. Maastricht: ICIS.
- May, K. (1892). *Durch die Wüste*.
- McDonough, W., & Braungart, M. (2002). *Cradle to Cradle*. North Point Press.
- McNeill, J. (2001). *Something new under the sun - an environmental history of the twentieth century*. London: Penguin.
- Meadows, D. H. (1997). Places to Intervene in a System (in increasing order of effectiveness). *Whole Earth* (Winter), 78-84.
- Meadows, D. H. (1999). *Leverage Points - places to intervene in a system*. Hartland, VT: The Sustainability Institute.
- Meadows, D. H., Meadows, D. L., & Randers, J. (1991). *Beyond the Limits*. London: Earthscan Publications Ltd.
- Meadows, D. H., Meadows, D. L., Randers, J., & Behrens, W. W. (1972). *The Limits to Growth*. New York: Universe Books.
- Meadows, D. L. (1998). *Indicators and Information Systems for Sustainable Development*. Hartland, VT: The Sustainability Institute.
- Meadows, D. L., Meadows, D. H., & Randers, J. (1972). *Dynamics of Growth in a Finite World*. Cambridge, UK: Wright-Allen Press.
- Messner, S., & Strubegger, M. (1995). *User's Guide for MESSAGE III* (No. WP-95-69). Laxenburg, Austria: IIASA.
- Millennium Ecosystem Assessment. (2005). *Ecosystems and Human Well-being: Synthesis*. Washington, DC: Island Press.
- Miser, H., & Quade, E. (1985). *Handbook of Systems Analysis: overview of uses, procedures, applications and practices*. Chichester: John Wiley & Sons.
- MNP. (2006). *Indicatoren en duurzaamheidsindex: Verantwoording van het werk rond indicatoren voor de Duurzaamheidsverkenning 'Kwaliteit en Toekomst'*. Bilthoven: Milieu en Natuur Planbureau (MNP).
- Moldan, B., & Dahl, A. L. (2007). Challenges to Sustainability Indicators. In T. Hák, B. Moldan & A. L. Dahl (Eds.),

Sustainability Indicators - A Scientific Assessment. Washington: Island Press.

- Mori, S., & Takahashi, M. (1999). An integrated assessment model for the evaluation of new energy technologies and food productivity. *International Journal of Global Energy Issues*, 11(1-4), 1-18.
- Morita, T., Kaihuma, Y., Harasawa, H., Kai, K., Dong-Kun, L., & Matsuoka, Y. (1994). *Asian-Pacific Integrated Model for Evaluating Policy Options to Reduce GHG Emissions and Global Warming Impacts*. Tskuba, Japan: National Institute for Environmental Studies.
- Morrison-Saunders, A., & Therivel, R. (2006). Sustainability integration and assessment. *Journal of Environmental Assessment, Policy and Management*, 8(3), 281-298.
- Moss, S., & Pahl-Wostl, C. (2001). Agent-based integrated assessment modelling: the example of climate change. *Integrated Assessment*, 2(1), 17-30.
- Munasinghe, M. (1993). *Environmental Economics and Sustainable Development* (World Bank Environment Paper No. 3). Washington: The World Bank.
- National Research Council. (1999). *Our common journey: a transition toward sustainability*. Washington: National Academy Press.
- Nordhaus, W. D. (1992). *The DICE Model: Background and Structure of a Dynamic Integrated Climate Economy*. USA: Yale University.
- Nordhaus, W. D. (1994). *Managing the Global Commons: The Economics of Climate Change*. Cambridge, USA: MIT Press.
- O'Riordan, T. (1985). The impact of EIA on decision-making. In V. T. Covello, J. L. Mumpower, P. J. M. Stallen & V. R. R. Uppuluri (Eds.), *Environmental Impact Assessment, Technology Assessment and Risk Analysis* (Vol. G 4). Berlin: Springer.
- OECD. (1993). *Environmental Indicators: Basic Concepts and Terminology*. Paper presented at the Indicators for use in environmental performance reviews, Paris.
- OECD. (2001). *Helping prevent violent conflict*. Paris: OECD.
- Pahl-Wostl, C. (2002). Participative and Stakeholder-Based Policy Design, Evaluation and Modeling Processes. *Integrated Assessment*, 3(1), 3-14.

- Parris, T., & Kates, R. W. (2003). Characterizing a sustainability transition: Goals, targets, trends, and driving forces. *PNAS*, 100(14), 8068-8073.
- Pearce, D., Markandya, A., & Barbier, E. A. (Eds.). (1989). *Blueprint for a Green Economy*. London: Earthscan.
- Pope, J. (2006). Editorial - What's so special about sustainability assessment. *Journal of Environmental Assessment, Policy and Management*, 8(3), v-x.
- Provincie Limburg. (2001). *Liefde voor Limburg - Provinciaal Omgevingsplan Limburg (POL)*. Maastricht: Provincie Limburg.
- Provincie Limburg. (2005). *Duurzaamheid duurt het langst: Programma Duurzaam Limburg 2005-2007*. Maastricht: Provincie Limburg.
- Provincie Limburg. (2006). *Provinciaal Omgevingsplan Limburg 2006*. Maastricht: Provincie Limburg.
- Provincie Limburg. (2007a). *Limburgmonitor 2007*. Maastricht: Provincie Limburg.
- Provincie Limburg. (2007b). *Provinciaal Verkeers- en Vervoersplan (PVVP)*. Maastricht: Provincie Limburg.
- Quade, E. S. (1983). *Analysis for public decisions* (3rd edition ed.). New York: Elsevier.
- Ravetz, J. (1999a). Citizen participation for integrated assessment: new pathways in complex systems. *International Journal of Environment and Pollution*, 11(3), 331-350.
- Ravetz, J. (1999b). What is post-normal science? *Futures*, 31, 647-653.
- Ravetz, J. (2003). Models as Metaphors. In B. Kasemir, J. Jäger, C. C. Jaeger & M. T. Gardner (Eds.), *Public Participation in Sustainability Science* (pp. 62-77). Cambridge: Cambridge University Press.
- Redefining Progress. (1995). *Genuine Progress Indicator*. San Francisco: Redefining Progress.
- Rittel, H. W., & Webber, M. M. (1973). Dilemmas in a General Theory of Planning. *Policy Sciences*, 4, 155-169.
- RIVM. (1997). *National Environmental Outlook 4 1997-2020 (in Dutch)*. Bilthoven: Samsom H.D. Tjeenk Willink.

- RIVM. (2001). *Milieubalans 2001*. Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu.
- Robèrt, K.-H., & Anderson, R. M. (2002). *The Natural Step Story: Seeding a quiet revolution*. Babriola Island, B.C.: New Society Publishers.
- Robèrt, K.-H., Schmidt-Bleek, B., Aloisi de Lardereel, J., Basile, G., Jansen, J. L., Kuehr, R., Price Thomas, P., Suzuki, M., Hawken, P., & Wackernagel, M. (2002). Strategic Sustainable development - selection, design and synergies of applied tools. *Journal of Cleaner Production*, 10, 197-214.
- Robinson, J. (2004). Squaring the circle? Some thoughts on the idea of sustainable development. *Ecological Economics*, 48(4), 369-384.
- Robinson, J. (2008). Being undisciplined: Transgressions and intersections in academia and beyond. *Futures*, 40, 70-86.
- Robinson, J., & Tansey, J. (2006). Co-production, emergent properties and strong interactive social research: the Georgia Basin Futures Project. *Science and Public Policy*, 33(2), 151-160.
- Rotmans, J. (1990). *IMAGE: An Integrated Model to Assess the Greenhouse Effect*. Dordrecht: Kluwer Academics.
- Rotmans, J. (1997). Indicators for Sustainable Development. In J. Rotmans & B. de Vries (Eds.), *Perspectives on Global Change: The TARGETS approach*. Cambridge, UK: Cambridge University Press.
- Rotmans, J. (1998a). *Integrated assessment: a necessary good (in Dutch)* (inaugural speech). Maastricht: ICIS.
- Rotmans, J. (1998b). Methods for IA : The challenges and opportunities ahead. *Environmental Modelling and Assessment*, 3(3), 155-179.
- Rotmans, J. (2005). *Societal Innovation: between dream and reality lies complexity* (inaugural speech). Rotterdam: Erasmus University Rotterdam.
- Rotmans, J. (2006). Tools for Integrated Sustainability Assessment: A two-track approach. *The Integrated Assessment Journal*, 6(4), 35-57.
- Rotmans, J., & de Vries, B. (1997). *Perspectives on Global Change: The TARGETS approach*. Cambridge: Cambridge University Press.

- Rotmans, J., & Dowlatabadi, H. (1998). Integrated Assessment of Climate Change: Evaluation of Methods and Strategies. In S. Rayner & E. Malone (Eds.), *Human Choice and Climate Change: An International Social Science Assessment*. Washington, USA: Battelle Press.
- Rotmans, J., Kemp, R., & van Asselt, M. B. A. (2001a). More evolution than revolution: transition management in public policy. *Foresight*, 3(1), 15-31.
- Rotmans, J., Kemp, R., & van Asselt, M. B. A. (2001b). Transition management: A promising perspective. In *Interdisciplinarity in Technology Assessment. Implementation and its chances and limits*: Europese Akademie, Germany.
- Rotmans, J., Kemp, R., van Asselt, M. B. A., Geels, F., Verbong, G., Molendijk, K. G. P., & van Notten, P. (2001). *Transitions & Transition management : The case for a low emission energy supply (in Dutch)*. Maastricht: ICIS BV.
- Rotmans, J., Rijkens, N., & van Asselt, M. B. A. (1998). *Limburg in perspectief : Een eerste aanzet*. Maastricht: ICIS institute & ICIS BV.
- Rotmans, J., & Rothman, D. (2003). *Scaling in Integrated Assessment*: Lisse, Swets & Zeitlinger Publishers.
- Rotmans, J., & van Asselt, M. B. A. (1996). Integrated Assessment: A growing child on its way to maturity. *Climatic Change*, 34(3-4), 327-336.
- Rotmans, J., & van Asselt, M. B. A. (1999). Perspectives on a sustainable future. *International journal for sustainable development*, 2(2), 201-230.
- Rotmans, J., van Asselt, M. B. A., de Groot, R. S., van Vliet, A., & Molendijk, K. G. P. (2000). *Biodiversiteit in geïntegreerd perspectief*. Maastricht: ICIS BV.
- Rotmans, J., van Asselt, M. B. A., Grevers, W., Grosskurth, J., & Molendijk, K. G. P. (2000). *Harmonisatie Prognoses POL : Vooronderzoek verbetering POL-prognoses*. Maastricht: ICIS BV.
- Rotmans, J., van Asselt, M. B. A., & Rijkens-Klomp, N. (1998). *Een denkmodel van kapitaalsvormen, voorraden en stromen*. Maastricht: ICIS institute & ICIS BV.
- Rotmans, J., van Asselt, M. B. A., & Rijkens, N. (1999a). *The Conceptual Model for POL (in Dutch)*. Maastricht: ICIS.

- Rotmans, J., van Asselt, M. B. A., & Rijkens, N. (1999b). *Het denkmodel voor POL*. Maastricht: ICIS institute & ICIS BV.
- Schellnhuber, H.-J., & Wenzel, V. (1998). *Earth System Analysis: Integrating Science for Sustainability*. New York: Springer-Verlag.
- Scholz, R. W., & Tietje, O. (2002). *Embedded case study methods*: SAGE.
- Schumacher, E. F. (1977). *A Guide for the Perplexed*. New York: Harper Collins.
- Scrase, J. I., & Sheate, W. R. (2002). Integration and Integrated Approaches to Assessment: What do they mean for the environment? *Journal of Environmental Policy & Planning*, 4(4), 275-294.
- SDRI. (1999). QUEST Overview: envision: sustainability tools Inc, Sustainable Development Reserach Institute - University of British Columbia.
- SEI Boston Center. (1999). PoleStar 2000: SEI Boston Center, Tellus.
- Siebenhüner, B., & Barth, V. (2005). The role of computer modelling in participatory integrated assessments. *Environmental Impact Assessment Review*, 25(4), 367-389.
- Social Learning Group (Ed.). (2001). *Learning to manage global environmental risks*. Cambridge, MA: MIT Press.
- Sondeijker, S., Geurts, J., Rotmans, J., & Tukker, A. (2006). Imagining sustainability: the added value of transition scenarios in transition management. *foresight*, 8(5), 15-30.
- Spangenberg, J. H., & Bonniot, O. (1998). *Sustainability Indicators - A Compass on the Road Towards Sustainability* (No. 81). Wuppertal: Wuppertal Institute.
- Sterman, J. D. (1994). Learning in and about complex systems. *System Dynamics Review*, 10(2-3), 291-330.
- Sterman, J. D. (2000). *Business dynamics - systems thinking and modeling for a complex world*. Boston: Irwin McGraw-Hill.
- Stern, K. (1986). *The Flight from Woman*. Paragon House publishers.
- Stockholm Environment Institute, & Global Scenario Group. (1997). *Branch Points: Global scenarios and human choice* (No. Polestar Series Report No.7). Stockholm: SEI.

- Stockholm Environment Institute, & Global Scenario Group. (1998). *Bending the curve: Towards global sustainability*. Stockholm: SEI.
- SustA-Test. (2006). Retrieved 15 October 2006, from www.SustainabilityA-Test.net
- Swart, R., Raskin, P., & Robinson, J. (2002). Critical Challenges for Sustainability Science. *Science*, 297, 1994.
- Swart, R., Raskin, P., & Robinson, J. (2004). The problem of the future: sustainability science and scenario analysis. *Global Environmental Change*, 14, 137-146.
- Tainter, J. (1996). Complexity, Problem Solving and Sustainable Societies. In J. Olson, R. Constanza, O. Segura & J. Martinez-Alier (Eds.), *Getting Down to Earth: Practical Applications of Ecological Economics*: Island Press.
- Tansey, J., Carmichael, J., VanWynsberghe, R., & Robinson, J. (2002). The future is not what it used to be: participatory integrated assessment in the Georgia Basin. *Global Environmental Change*, 12(2), 97-104.
- Telos. (2001). *De Duurzaamheidsbalans van Noord-Brabant 2001*. Tilburg: TELOS.
- Telos. (2006). *De Duurzaamheidsbalans van Brabant 2006*. Tilburg: Telos.
- Teunis, U. (1996). *De ontwikkeling van een Euregionaal informatiesysteem onderwijs-arbeidsmarkt voor de Euregio Maas-Rijn* (No. ROA-W-1996/7). Maastricht, the Netherlands: ROA - Research Centre for Education and the Labour Market.
- Thompson, M. (1997). Cultural Theory and Integrated Assessment. *Environmental modelling and assessment*, 2(3), 139-150.
- Thompson, M., Ellis, R., & Wildavsky, A. (1990). *Cultural Theory*. Boulder, CO: Westview Press.
- Toth, F. L. (2003). State of the Art and Future Challenges for Integrated Environmental Assessment. *Integrated Assessment*, 4(4), 250-264.
- Tuinstra, W., Berk, M., Hisschemoeller, M., Hordijk, L., Metz, B., & Mol, A. (2002). *Climate OptiOns for the Long-term (COOL) - Synthesis Report*: Wageningen University.
- Turnpenny, J. (2008). "Are we nearly there yet?" Lessons for integrated sustainability assessment from EU environmental

- policy-making. *International Journal of Innovation and Sustainable Development*, 3(1), 33-47.
- UN-DPCSD. (1999). *Indicators of Sustainable Development: framework and methodologies*. New York: Department for Policy Coordination and Sustainable Development - United Nations.
- UNDP. (2003). *Human Development Report 2003*. New York: Oxford University Press.
- UNECE. (1998). *UNECE Convention on Access to Information, Public Participation in Decision-making and Access to Justice in Environmental Matters*: United Nations.
- UNEP. (2002). *Global Environmental Outlook 3*. London: Earthscan.
- Valkering, P., Krywkow, J., Rotmans, J., & van der Veen, A. (2004). *Simulating stakeholder support for river management*. Paper presented at the iEMSs 2004 International Congress: Complexity and Integrated Resources Management, Osnabrück, International Environmental Modelling and Software Society.
- van Asselt, M. B. A. (2000). *Perspectives on Uncertainty and Risk : The PRIMA approach to decision support*. Dordrecht: Kluwer Academic Publishers.
- van Asselt, M. B. A., Mellors, J., Rijkens-Klomp, N., Greeuw, S. C. H., Molendijk, K. G. P., Beers, P. J., & van Notten, P. (2001). *Building Blocks for Participation in Integrated Assessment. An assessment of participatory methods*. Maastricht: ICIS.
- van Asselt, M. B. A., & Rijkens, N. (2002). A Look in the Mirror: Reflection on participation in Integrated Assessment from a methodological perspective. *Global Environmental Change*, vol. 12, 167-184.
- van Asselt, M. B. A., Rijkens, N., & Molendijk, K. (2000). *Werken met het denkmodel*. Maastricht: ICIS institute & ICIS BV.
- van Asselt, M. B. A., & Rotmans, J. (1997). Uncertainties in perspective. In J. Rotmans & B. de Vries (Eds.), *Perspectives on Global Change: the TARGETS approach*. Cambridge, UK: Cambridge University Press.
- van Asselt, M. B. A., Rotmans, J., & Rothman, D. (2005). *Scenario Innovation: Experiences from a European Experimental Garden*: Taylor & Francis.

- van de Kerkhof, M. (2004). *Debating Climate Change*. Amsterdam: Lemma.
- van der Brugge, R., Rotmans, J., & Loorbach, D. (2005). The transition in Dutch water management. *Regional Environmental Change*, 5(4), 164-176.
- van der Sluijs, J., & Schulte Fishedick, K. (1997). *Dealing with Uncertainty in Science for Environmental Policy: An inventory of theories and approaches (in Dutch)* (No. 97017). Utrecht: University of Utrecht and the Rathenau Institute.
- van der Sluijs, J. (1997). *Anchoring amid Uncertainty: On the management of uncertainties in risk assessment of anthropogenic climate change*. Utrecht: Universiteit Utrecht.
- van der Veen, A., & Rotmans, J. (2001). Dutch perspectives on "Agents, Regions and Land Use Change". *Environmental Modelling and Assessment*, 6(2), 83-86.
- van der Woerd, K. F., Dellink, R. B., Boelens, J. J. M., Masurel, E., & Verberne, T. G. (1997). *MIOW-PROV-rekenmodel voor de sociaal-economische gevolgen van provinciaal milieubeleid* (No. R-97/11). Amsterdam: Instituut voor Milieuvraagstukken, Vrije Universiteit Amsterdam.
- van Notten, P. (2005). *Writing on the Wall: Scenario Development in Times of Discontinuity*. Boca Raton, FL: Dissertation.com.
- Vennix, J., Andersen, D. F., Richardson, G. P., & Rohrbaugh, J. (1996). Model building for group decision support: issues and alternatives in knowledge elicitation. In J. D. W. Morecroft & J. D. Sterman (Eds.), *Modeling for learning organizations*. Portland, Oregon: Productivity Press, Inc.
- Vester, F. (2002). *Die Kunst vernetzt zu denken*. Munich: dtv.
- Vitousek, P. M., Ehrlich, P. R., Ehrlich, A. H., & Matson, P. A. (1986). Human appropriation of the products of photosynthesis. *BioScience*, 36(6), 368-373.
- von Weizsäcker, E. U., Lovins, A. B., & Lovins, L. H. (1995). *Faktor Vier: Doppelter Wohlstand - halbiertes Naturverbrauch*. Munich: Droemersch Verlaganstalt Th. Knaur Nachf.
- Voss, J.-P., Bauknecht, D., & Kemp, R. (Eds.). (2006). *Reflexive governance for sustainable development*. Cheltenham: Edward Elgar.

- Wackernagel, M., & Rees, W. (1996). *Our ecological footprint: reducing human impact on the earth*: New Society.
- Watson, J., Tetteh, A., Dutton, G., Bristow, A., Kelly, C., & Page, M. (2004). *UK Hydrogen Futures to 2050*: Tyndall Centre for Climate Change Research.
- WCED. (1987). *Our Common Future*. Oxford, UK: Oxford University Press.
- Weaver, P., Jäger, J., & Rotmans, J. (2008). Integrated Sustainability Assessment: Concept, Process and Tools. *International Journal of Innovation and Sustainable Development (special issue)*, 3(1).
- Weaver, P., & Rotmans, J. (2006). Integrated Sustainability Assessment: What? Why? How? *International Journal of Innovation and Sustainable Development*, 1(4), 284-303.
- Weimer-Jehle, W. (2006). Cross-impact balances: A system-theoretical approach to cross-impact analysis. *Technological Forecasting and Social Change*, 73(4), 334-361.
- Wilbanks, T. J., & Kates, R. W. (1999). Global Change in Local Places: How Scale Matters. *Climatic Change*, 43, 601-628.
- Wilson, B. (2001). *Soft Systems Methodology: Conceptual Model Building and its Contribution*. Chichester: John Wiley & Sons.
- World Bank. (1997). *Expanding the Measure of Wealth: Indicators of Environmentally Sustainable Development*. Washington D.C.: World Bank.
- WWF. (2006). *Living Planet Report 2006*. Gland: WWF International.
- Yin, R. K. (2003). *Case Study Research: Design and methods*. London and Thousand Oaks, California: Sage.

LIST OF ACRONYMS

AGE	Applied General Equilibrium
AIM	Asian Pacific Integrated Model
AOS	Atmosphere-Ocean System
APA	Accounting and Physical Analysis Tools
ASF	Atmospheric Stabilization Framework Model
CBA	Cost Benefit Analysis
CDM	Clean Development Mechanism
CEA	Cost Effectiveness Analysis
CEC	Cambridge Econometrics
CIESIN	Center for International Earth Science Information Network
COMBI	COMBINatie van modellen
COOL	Climate OptiOns for the Long term
CPB	Centraal Planbureau
DEP	Deposition and Critical loads Assessment
DICE	Dynamic Integrated Climate Economy
DPSIR	Driver-Pressure-State-Impact-Response
E3ME	Energy-Environment-Economy Model for Europe
EC	European Commission
EF	Ecological Footprint
EIA	Environmental Impact Assessment
EIS	Energy-Industry System
EMCO	Emission-Cost Module
eQSA	enriched Qualitative Systems Analysis
ERB	Edmonds-Reilly-Barns
EU	European Union
GB-Quest	Georgia Basin Quest for sustainability

GDP	Gross Domestic Product
GEM-E3	General Equilibrium Model for Energy-Economics-Environment
GEO	Global Environment Outlook
GHG	GreenHouse Gas
GIS	Geographic Information System
GITAGE	Globalisation, International Transport and the Global Environment
GNP	Gross National Product
GPI	Genuine Progress Indicator
GTAP	Global Trade Analysis Project general equilibrium model
GWP	Gross World Product
HANPP	Human Appropriation of Net Primary Production
HDI	Human Development Index
IA	Integrated Assessment
IAM	Integrated Assessment Model
ICIS	International Centre for Integrated assessment and Sustainable development (formerly known as the International Centre for Integrated Assessment)
IF	International Futures
IIASA	International Institute for Applied Systems Analysis
IMAGE	Integrated Model to Assess the Global Environment
IMF	International Monetary Fund
IPB	InterProvinciaal woningBehoeftte-model
IPCC	Intergovernmental Panel on Climate Change
IQ tools	Indicators and Quantitative Tools for Improving the Process of Sustainability Impact Assessment
IS	Indicator Sets
ISA	Integrated Sustainability Assessment
ISEW	Index of Sustainable Economic Welfare
JOBS	<i>unknown</i>
KSIM	Kane's Simulation Method
LOV	LeefOmgevings Verkenner

MAGICC	Model for the Assessment of Greenhouse gas Induced Climate Change
MARIA	Multiregional Approach for Resource and Industry Allocation
MATISSE	Methods And Tools for Integrated SuStainability assEssment
MCA	Multi Criteria Analysis
MEA	Millennium Ecosystem Assessment
MERGE	Model for Evaluating the Regional and Global Effects of GHG Reduction Policies
MESSAGE	Model for Energy Supply Strategy Alternatives and their General Environmental Impact - Macro economic model
MICMAC	Matrice d'Impacts Croisés-Multiplication Appliquée à un Classement
MiniCAM	Mini Climate Assessment Model
MIOW-PROV	Marktsituatie, Internationale concurrentie, Omvang en Weerstandsvermogen PROVinciaal milieu-economisch model
MNP	Milieu- en Natuurplanbureau
MO	Models
NERIS	Netzwerk in der Risiko-Sensitivitätsanalyse
NGO	Non Governmental Organization
NMP	Nationaal Milieubeleidsplan
OECD	Organisation for Economic Co-operation and Development
OGM	Overdraagbaar Groeimodel
OPT	Optimization module
PNS	Post Normal Science
POL	Provinciaal Omgevingsplan Limburg
PRIMOS	PRognose-, Informatie-, en MOnitoring Systeem
PROVEST	PROgnosemodel VESTigingslocaties
PSIR	Pressure-State-Impact-Response
PSR	Pressure-State-Response
PT	Participatory Tools
QSA	Qualitative Systems Analysis

QSSI	Qualitative Systems Sustainability Indicator
RAIL	Regionale Arbeidsmarkt Informatie Limburg
RAINS	Regional Air pollution Information and Simulation
RAP	Rapid Assessment Protocol
RDP	Regionaal Demografisch Prognosemodel
RIVM	Rijksinstituut voor Volksgezondheid en Milieu
SCENE	SoCial Environmental Economic
SDRI	Sustainable Development Research Institute
SEI	Stockholm Environment Institute
SG	Scenario Generator
SM	Sensitivity Model
ST	Scenario Tools
SustA-test	SustainabilityA-Test
T21	Threshold 21
TARGETS	Tool to Assess Regional and Global Environmental and health Targets for Sustainability
TES	Terrestrial Environment System
TIA	Trend-Impact Analysis
TIMER	Targets IMage Energy Regional
TOK	Toekomst op Koers
ULYSSES	Urban Lifestyles, Sustainability and Integrated Environmental Assessment
UN-DPCSD	United Nations Department for Policy Coordination and Sustainable Development
UNECE	United Nations Economic Commission for Europe
UNEP	United Nations Environment Programme
US	United States
USA	United States of America
VISIONS	Visions for a sustainable Europe
WCED	World Commission on Environment and Development
WorldScan	World model for Scenario Analysis
WWF	World Wildlife Fund
ZEW	Zentrum für Europäische Wirtschaftsforschung

SAMENVATTING

(Hoofdstuk 1) Duurzaamheid omvat de ambitie, natuurlijke, sociale en economische systemen in balans te laten ontwikkelen. De zoektocht naar duurzaamheid komt voort uit het bewustzijn, dat tal van ontwikkelingen op onze planeet de capaciteit bedreigen, om aan de behoeften van huidige en toekomstige generaties te voldoen. Voor de realisatie van duurzaamheid is een veelomvattende transitie noodzakelijk. Een dergelijke transitie brengt voor beleidsmakers ongekende uitdagingen met zich mee. Met hoge urgentie en in een uitermate complexe en onzekere context moeten zij strategische prioriteiten identificeren en integraal doordachte plannen formuleren. Dit proefschrift maakt deel uit van de zoektocht naar geschikte hulpmiddelen voor geïntegreerde beschrijving, analyse en strategische planning van duurzaamheid in wetenschap en beleid. Daarbij richt zich het proefschrift vooral op regionale beleidsmedewerkers.

(Hoofdstuk 2) Duurzaamheid is in eerste instantie een politiek en geen wetenschappelijk concept. Complexiteit, ambiguïteit, subjectiviteit en normativiteit zijn belangrijke kenmerken van duurzaamheid. Traditionele wetenschappelijke methoden zijn slechts in beperkte mate geschikt om dergelijke vraagstukken op te lossen. In het recente verleden zijn daarom een reeks nieuwe onderzoeksfilosofieën en –terreinen ontwikkeld, met name Mode2 science, post-normal science, integrated assessment en sustainability science. In tegenstelling tot traditionele wetenschappen, zijn de nieuwe wetenschappen transdisciplinair en heteroog. De onderzoeksinhoud is aan maatschappelijke in plaats van academische verantwoording onderhevig. Het voorliggende onderzoek gebruikt integrated assessment methoden, waaronder participatieve methoden, verschillende manieren van systeem analyse en case study onderzoek. Deze worden toegepast in de context van regionaal beleid voor duurzaamheid met het doel, een geïntegreerde besluitvorming op meervoudige wijze te faciliteren.

(Hoofdstuk 3) De wetenschappelijke zoektocht voor geschikte methodes heeft een snel groeiend aantal mogelijke aanpakken opgeleverd. Deze zijn in

enkele recente onderzoeksprojecten geïnventariseerd en op hun geschiktheid voor relevante vraagstellingen geëvalueerd. Aan deze evaluaties voegt dit hoofdstuk een overzicht toe, voor welke schaal van interventie bestaande methoden geschikt zijn. De interventieschaal beschrijft, hoe ingrijpend een systeem zou moeten veranderen om een doelstelling te bereiken. Daaruit blijkt, dat er een (snel afnemend) tekort is aan methoden voor het aanpakken van diepgaande, complexe veranderingsprocessen.

(Hoofdstuk 4) Een mogelijke eerste stap op weg naar het omgaan met complexiteit is het in kaart brengen ervan. De SCENE methode richt zich op het beschrijven van complexiteit bij duurzaamheidsvraagstukken. Het daarvoor gebruikte conceptuele SCENE model omvat de economische, sociale en natuurlijke domeinen, evenals de interactie binnen en tussen deze domeinen. Het conceptuele model wordt in een participatief proces ingevuld, waarbij er kennisuitwisseling en –ontwikkeling tussen de deelnemers plaats vindt. Het tastbare resultaat van het proces is een systeemkaart ofwel invloeddiagram uit duurzaamheidsperspectief omtrent een thema of regio.

(Hoofdstuk 5) De systeemkaarten, die het resultaat van een SCENE proces of vergelijkbare methoden zijn, kunnen met behulp van de eQSA methode diepgaand geanalyseerd worden. Deze methode bouwt voort op een lange traditie van structurele, participatieve analysemethoden voor complexe systemen. Relatief eenvoudige berekeningen op basis van participatieve inbreng volstaan om veelbelovende aanknopingspunten voor duurzaamheidsstrategieën te identificeren. De gevolgen van strategische keuzes worden in kaart gebracht en de interventieschaal wordt duidelijk. Het hoofdstuk beschrijft naast de methode ook de toepassing ervan samen met beleidsmedewerkers van de Provincie Limburg.

(Hoofdstuk 6) Voor een duurzaamheidstransitie moeten systeemveranderingen niet alleen in gang gezet worden, maar ook gemonitord. Het meest gebruikelijke middel hiervoor zijn lijsten van indicatoren of daarop opbouwende geaggregeerde indices. Bestaande duurzaamheidsindicatoren en –indices richten zich daarbij vooral op de huidige toestand van relevante variabelen en negeren de interacties tussen de variabelen. Voor lange termijn veranderingen vormt dit een probleem, omdat deze interacties het krachtenveld vormen, dat de toekomstige ontwikkeling van de variabelen bepaalt. De in het kader van dit proefschrift ontwikkelde QSSI vormt een eerste aanzet om het krachtenveld integraal te

monitoren, ten koste van het monitoren van de huidige toestanden. Door dit unieke perspectief geeft de QSSI inzicht in kansrijke structurele verbeteringen en documenteert hij systeemveranderingen op de weg naar duurzaamheid.

(Hoofdstuk 7) Rekenmodellen zijn een veel toegepaste en op het eerste gezicht zeer aantrekkelijke methode om complexe systemen en hun gedrag te beschrijven en te analyseren. Echter, de karakteristieken van het concept duurzaamheid staan een veelomvattende analyse gebaseerd op kwantitatieve gegevens in de weg. De kracht van de meeste rekenmodellen ligt vooral in het inzichtelijk maken van directe consequenties van zeer concrete systeemveranderingen ofwel in het beschrijven van systeemdynamieken. Het laatste wordt onder meer gebruikt als leermiddel in participatieve processen. Nieuwe modelleermethoden, zoals agent gebaseerde of evolutionaire modellen, beloven verdere vooruitgang op dit gebied. Desalniettemin blijkt uit een vergelijking van de kwalitatieve analyse uit de hoofdstukken 4 en 5 met de modelstructuur van kwantitatieve modellen die door beleidsmedewerkers bij de Provincie Limburg gebruikt worden, dat door kwantificering niet alleen een groot verlies van informatie optreedt, maar ook, dat de systeemeigenschappen van de kwantitatief onderzochte systemen fundamenteel verschillen van de systeemeigenschappen van de kwalitatieve systemen. Dit heeft ingrijpende gevolgen voor de zorgvuldigheid, waarmee modelresultaten geïnterpreteerd en toegepast moeten worden.

(Hoofdstuk 8) De verschillende hoofdstukken in dit proefschrift trachten elk op een andere manier bij te dragen aan het faciliteren van geïntegreerd beleid voor duurzaamheid. Hoofdstuk 3 richt zich vooral op het belang van de keuzes in onderzoeksmethodes en biedt handvaten om deze keuze voor een gegeven vraagstelling te maken. Hoofdstuk 4 beschrijft een manier om in een participatief onderzoek- en leerproces tot een systeemkaart voor een gegeven thema of regio te komen. Hoofdstuk 5 gebruikt de systeemkaarten uit het SCENE proces en analyseert deze met eQSA. Voor de beleidsmedewerkers levert deze methode inzicht op in mogelijke beleidsstrategieën. De effectiviteit van daaruit voortvloeiende interventies kan vervolgens met de in hoofdstuk 6 beschreven duurzaamheidsindicator QSSI gevolgd worden. Hoofdstuk 7 onderzoekt de mogelijkheden om bijvoorbeeld duurzaamheidsbeleid met behulp van rekenmodellen te onderzoeken of te plannen. Een beter begrip van de specifieke waarden van rekenmodellen door beleidsmakers en dus ook een betere interpretatie ervan

zou de waarde van de modellen ten goede komen. Alle beschreven aanpakken zijn slechts kleine stappen op de weg naar methoden voor een geïntegreerd duurzaamheidsbeleid. De daadwerkelijke meerwaarde van deze methoden zal nog moeten worden bewezen en de methoden zelf bieden nog ruimte voor ingrijpende verbeteringen en uitbreidingen. De evaluatie van de vragen of de methoden tot meer geïntegreerd overwogen besluiten hebben geleid en of de besluiten beter waren dankzij de toepassing van de methoden is zeer moeilijk. Er zijn echter indicaties, dat de het kader van dit proefschrift beschreven en ontwikkelde methoden in proces, inhoud en resultaat ingang hebben gevonden in de besluitvormingsprocedures bij de Provincie Limburg.

ACKNOWLEDGEMENTS

Ich bekam festen Boden, er wich aber augenblicklich unter mir. Halb schon im Versinken, fußte ich wieder und raffte mich auf. Ich sank und erhob mich, ich strauchelte, ich trat fehl, ich fand dennoch Grund. Ich wurde hinabgerissen und kam dennoch vorwärts und ging dennoch nicht unter. Ich hörte nichts mehr, ich fühlte nichts mehr, ich sah nichts mehr als nur die drei Männer dort am Hügel, von denen mich zwei mit angeschlagenem Gewehr erwarteten.

Karl May, *Durch die Wüste* (1892)

It would require an additional volume to name those, who have in one way or another aided this dissertation and the underlying research. ‘It takes a village to raise a child’ and it takes many to write a dissertation. From my heart, I thank all of you.

Most of the time I felt like a child in a playground with ideas and concepts as toys, conferences as games of tag and a group of peers sharing the pleasures of exploration, adventure, and meaning. I want to thank those, who provided this academic playground and joined many of the games: the university, my supervisors Jan Rotmans and Pim Martens, many generous colleagues and staff at ICIS and elsewhere, the enthusiastic participants in the case studies, and those who spent long hours of discussions with me on new games, adjusted rules, and the type of games it would require to change the world.

If a joyful game of curiosity and creativity would be sufficient to complete a dissertation, this book would have been in your hands a while ago. I want to thank all those, who were patient and willing to let me find my own way of transforming the games into an academic product. I also want to thank those, who have aided me in my search. Some were not so patient and pushed me on or even took work out of my hand, for which I am even more grateful.

On a private note, I want to thank my parents, especially for striking the difficult balance between letting me find my own way and helping me to take the next step. I want to thank Birthe for being there in times of frustration and motivation.

CURRICULUM VITAE

Jasper Grosskurth was born on 5 May 1975 in Munich, Germany. In Munich, he attended primary and secondary school, graduating in 1994. He studied International Economic Studies at Maastricht University (formerly known as Rijksuniversiteit Limburg) and Trinity College Dublin. Jasper obtained his MSc. degree in 2000 with a thesis titled 'Quantifying qualitative scenarios'. The thesis was written as part of the VISIONS project on European scenarios during an internship at the International Centre for Integrative Studies (ICIS, now known as the International Centre for Integrated assessment and Sustainable development). Until 2005 he stayed at ICIS as a PhD researcher on regional sustainability. In late 2004 he joined forces with three colleagues to found a commercial future studies consultancy, Pantopicon, in Antwerp. Since 2006, he worked free-lance, travelled and finished this dissertation. Currently, Jasper works as a project manager for the STT Netherlands Study Centre for Technology Trends (Stichting Toekomstbeeld der Techniek) coordinating a foresight project on the future of technology in Africa.