

Kaare Gether

Transition to Large Scale Use of Hydrogen
and Sustainable Energy Services

Choices of technology and infrastructure under path
dependence, feedback and nonlinearity

Doctoral thesis
for the degree of Philosophiae Doctor

Trondheim, March 2004

Norwegian University of Science and Technology
Department of Energy and Process Engineering

ISBN 82-471-6355-1
ISBN 82-471-6354-3(electronic ver.)

Preface

Chance favours the prepared mind

Louis Pasteur

The present thesis originates within the environment of the Department of Energy and Process Engineering at the Norwegian University of Science and Technology (NTNU). The department has extensive competence in the field of energy technologies, while a challenge exists to develop knowledge required to change our current unsustainable energy practice. On one hand is the production of offshore oil and gas, increasingly needed to sustain Europe with reliable supplies of energy. This by itself is a daunting technological challenge. On the other hand is the growing awareness of further exponential growth in the already extensive consumption of energy throughout the world. The environmental impact of this, in combination with the time it will take to convert to sustainable practices, may demand drastic measures within the next decades. We may elect to ignore it for a while yet. We may, however, also adopt the position that a shift towards sustainable use of energy embodies creation of value, and look for the possibilities and opportunities this conveys, rather than seeing problems and obstacles. This is the outset of the present work.

Achieving the creation of value inherent in transition to a sustainable energy system is a task at the borderline between technical solutions and human endeavour. Some of the problems to be solved are technological, or will be greatly eased by technological contributions, while all solutions will involve organisations and human relationships. The challenge is well described by Simon in his speech at the Nobel Price Banquet in 1978:

"...May I express my deep gratitude for the honor bestowed on me, and through me, on the colleagues with whom I have collaborated in studying economic behavior and the human mind.

Ten years ago, economics was added to the list of disciplines recognized in these ceremonies. About ten years ago, also, in the United States, the social and behavioral sciences were given a full seat in our National Academy of Sciences. The decisions to take these steps were acknowledgements that all the great problems that face our world today have both technical and human content - the one intermingled inseparably with the other.

To deal with these problems - of world population and hunger, of peace, of energy and mineral resources, of environmental pollution, of poverty - we must broaden and deepen our knowledge of Nature's laws, and we must broaden and deepen our understanding of the laws of human behavior. And we must do this in the spirit of deep concern for human values that is symbolized by the presence here also of the domain of literature."

Herbert A. Simon's speech at the Nobel Banquet,
December 10, 1978

The problems we focus on, the questions we raise and the answers we find, are related to our deeper perception of reality in the environment in which we live and work. This goes for the present work as well. I believe that if the world continues as it has done since the industrial revolution, it is a matter of time before the world's population will hit a constraint (e.g. available resources or environmental concerns). In such a setting it seems likely that humanity will be both tried and compromised. There is little agreement over when this 'time' will

occur, but considering the many exponential growth regimes we observe today, the term 'fire-fighting' may well come to have a new meaning halfway through the 21st century.

To avoid this scenario I believe sustainable development is the best way to ensure a liveable world for our descendants. In an attempt to reveal 'neutral' facts, a 'devil's advocate' role has been adopted as a counterweight to many 'hallelujah' people in the hydrogen promoting societies. This work investigates the role of hydrogen as an energy carrier in fierce competition with other alternatives. The scope and perception of the work has changed from initially being a rather static 'well-to-wheel' study of hydrogen, to a study of dynamic change and transition in the broader energy system. Hopefully this will put us in a better situation to see how hydrogen might help us towards sustainability.

The practical process of change puts a major focus on the firms that bring forth new products and services. In the present work these firms are understood to function as agents of society to accomplish practical change. This means that the firms operate under a social contract, rather than have as their prime *raison d'être* to earn money for their owners. Under this social contract they receive help in various ways to achieve their objectives. Even so, each individual firm functions in markets and requires a sound economic basis. Each firm therefore requires a *strategy* to guide its efforts so as, in cooperation with other firms, to meet the social contract. The strategy must tell the firm what products and services to deliver, and what knowledge and equipment to invest in. It must also tell the firm how to compete.

As a shift towards sustainability involves radical change in technology, the setting of competition is special: the firm must simultaneously *compete* with other firms supplying the same products/services, and *cooperate* with yet other firms in the supply

chain to secure competitiveness for their common end-products. The latter pay for all efforts leading to them, and therefore pay each individual firm. The essential challenge is to make sustainable end-products competitive against the ones that exist now. Broad knowledge of the technological, organisational, and business issues involved, is required to work out both overall strategy for transition, and strategy for individual firms. Clearly any substantial change will have its basis in the energy structure that is in place now: new development will depend on the competitive picture as the new seeks to replace the old.

Optimisation is used to find “theoretically best” solutions. However, theory is not reality. Decisions in reality are influenced by path dependence and bounded rationality. The present work includes these phenomena. Furthermore it seeks to identify potentials and opportunities and to combine this with understanding of how to achieve change. That is to get from where we are to what’s identified as ‘best’. A major component is to bring knowledge to decision makers in a way they can comprehend and communicate further.

From the ‘how to achieve change’ focus, an important realisation is reached: sequences are crucial. Not only what is decided, but in what order, must be considered. This introduces temporal development as an important factor. Dynamic modelling is a way to better understand the importance of sequences and path dependence. This dynamic aspect requires a type of modelling that will capture reality as it unfolds over time. One needs to obtain an overall understanding of both the forces involved and the circumstances created along the way that influence further decisions.

Perhaps equally important: this understanding is needed *where decisions are made*, i.e. with public authorities and with executives of individual firms, rather than in the towers of scientific research. While the latter is the natural place where the necessary tools and

understanding may be hatched, a means needs to be in place to allow decision makers themselves gain requisite competence, understanding and overview. They should be able to try out "what if...?" from their own, individual perspectives. An important target of the present work has thus been to provide a tool for communication and to develop a common ground for communication.

The subject matter of the present thesis is complex. It is difficult to convey overview and understanding of dynamic behaviour. In the context of string theory, B. Greene expresses it this way:

"I like to say things more than one way. I just think that when it comes to abstract ideas, you need many roads into them. From the scientific point of view, if you stick with one road, I think you really compromise your ability to make breakthroughs. I think that's really what breakthroughs are about. Everybody's looking at a problem one way, and you come at it from the back. The different way of getting there somehow reveals things that the other approach didn't.

Brian Greene, Scientific American, Nov. 2003:50

The best way to understand *dynamic* systems is not through textual, static description. Dynamic modelling is not a spectator sport: you really need to get your hands dirty. This means actually running computer simulations and watching what happens as outsets are changed, and as hypotheses and alternative approaches are tried out. Only then is a good grasp of the dynamics attainable. Such a dynamic approach has become feasible through modern computers and software, although this development is very recent. (The ability to handle extensive hierarchical structures in system dynamic software became available in May 2003).

A metaphor of dynamic models is the game of chess. Chess has a set of pieces that represents a system with a set of fixed rules for the

behaviour of these pieces. These fixed rules are analogous to the established patterns we observe in the real world. The dynamics are created through irreversibility in not taking moves back, corresponding to irreversibility from production of entropy in the real world. It is an often expressed view among chess players that "chess mirrors life". Once the game of chess is available, the dynamics of chess is played out in individual games. With time, the dynamics have been extensively investigated, with well established principles and general rules of strategy as important parts of the overall understanding of the game. These rules, and this understanding, are not available from a single game of chess. Instead it has evolved through thousands of games, with contributions from some very gifted players. In the same way, we seek a "game of change" to develop a broad understanding of all the issues involved in the transition away from our present non-sustainable energy systems.

What's been developed in the present work is intended to be used as a 'learning machine' where decision makers and opinion formers can 'play games'. This will develop understanding for the necessity of including dynamic thinking and to develop understanding of how dynamics play out and the ensuing consequences. The result is the Energy Infrastructure Competition Model (EICOMP), intended as a 'flight simulator'. Due to the inherently multidisciplinary nature of the challenge at hand, an attempt has been made to make the dissertation available to a broad audience.

This thesis is an attempt to try something new.

Acknowledgements

First of all I would like to thank my supervisors Geir A. Owren and Arne M. Bredesen for their encouragement, support and flexibility throughout my PhD study. No one quite knew what would come out of this work when we initiated the project, and many interesting and constructive discussions have resulted. For all the opportunities they have created for me in this work, I am deeply grateful.

During my research visit at the Massachusetts Institute of Technology from September 2002 to March 2003, I had the pleasure of meeting and gaining knowledge from a number of people. John D. Sterman at the Sloan School of Management in particular helped me to a deeper understanding of dynamic issues, for models to reflect all sides relevant to reality, and to the need for being critical. His kind guidance and wide-ranging knowledge have helped the project and have been an inspiration.

Stephen Connors, Mary E. Gallagher and David H. Marks in their different roles at the Laboratory for Energy and the Environment, deserve a special thanks. For integrating me into their networks, they all have their share in this thesis.

Magne Myrtveit, the original architect of Powersim, has played a crucial role. His programming experience and ability to incorporate ideas into programming code has played a vital role in the development of the EICOMP model.

The encouragement by Joan Ogden (Princeton University, Energy Group, now at UC Davis), and by Johan Hustad (NTNU) was much appreciated when on unknown paths and at times when the approach had no guarantee of succeeding. Joan Ogden has also kindly reviewed the manuscript for this thesis.

To my many friends and colleagues, I am grateful for their consistent lively discussions and scientific and social support... As time goes by, happy memories remain!

Norsk Hydro ASA and the Gas Technology Center, NTNU-SINTEF, have funded this research project. Additional funding for the research visit to MIT was given by NTNU's strategic program on Energy & Environment. Without this, there would have been no thesis!

I owe my sincere thanks to my family for their interest and practical support in this doctoral work. Specifically I thank my brother, Harald Gether, for helping out with issues on innovation and strategy, and my father, Jørgen J. Gether, for being a vital sparring partner for ideas, focus and quality of the work. Finally, I would like to thank my Ninja for simply being fantastic.

Summary

We live in a world of becoming. The future is not given, but forms continuously in dynamic processes where path dependence plays a major role. There are many different possible futures. What we actually end up with is determined in part by chance and in part by the decisions we make. To make sound decisions we require models that are flexible enough to identify opportunities and to help us choose options that lead to advantageous alternatives. This way of thinking differs from traditional cost-benefit analysis that employs net present value calculations to choose on purely economic grounds, without regard to future consequences.

Time and dynamic behaviour introduce a separate perspective. There is a focus on *change*, and decisions acquire windows of opportunity: the right decision at the right time may lead to substantial change, while it will have little effect if too early or too late. Modelling needs to reflect this dynamic behaviour. It is the perspective of time and dynamics that leads to a focus on sustainability, and thereby the role hydrogen might play in a future energy system. The present work develops a particular understanding relevant to energy infrastructures. Central elements of this understanding are:

- *Competition*
- *Market preference and choice beyond costs*
- *Bounded rationality*
- *Uncertainty and risk*
- *Irreversibility*
- *Increasing returns*
- *Path dependence*
- *Feedback*
- *Delay*
- *Nonlinear behaviour*

Change towards a “hydrogen economy” will involve far-reaching change away from our existing energy infrastructure. This infrastructure is viewed as a dynamic set of interacting technologies

(value sequences) that provide services to end-users and uphold the required supply of energy for this, all the way from primary energy sources. The individual technologies also develop with time.

Building on this understanding and analysis, an analytical tool has emerged: the Energy Infrastructure Competition (EICOMP) model. In the model each technology is characterised by a *capacity*, an *ordered-*, and an actually *delivered* volume of energy services. It is further characterised through physical description with parameters like efficiency, time required for extending capacity and improvement by learning. Finally, each technology has an *attractiveness*, composed of costs, quality and availability, that determines the outcome of competition.

Change away from our present energy infrastructure into a sustainable one based on renewable energy sources, will entail substantial change in most aspects of technology, organisation and ownership. Central results from the overall work are:

- *Change is dynamic and deeply influenced through situations with reinforcing feedback and path dependence. Due to this, there is a need for long-term perspectives in today's decision making: decisions have windows of opportunity and need to be made at the proper time.*
- *Strategies aimed at achieving change should team up with reinforcing feedback and avoid overwhelming balancing feedback that counteracts change.*
- *The EICOMP model is now available as a tool for further analysis of our existing energy infrastructure and its dynamic development into possible, alternative energy futures. As the model is intended for practical guidance in decisions, a central practical aim has been to allow it to be used close to where decisions are actually made; i.e. decentralised and*

locally in firms and in public institutions. In this respect much effort has been made in an attempt to make it transparent and easy to communicate.

- *The EICOMP model may be used to analyse situations of reinforcing feedback throughout the alternative energy infrastructures that we may come to have in the future.*

Contents

Preface	i
Acknowledgements	vii
Summary	ix
1 Introduction.....	1
2 Energy structures and environmental futures	11
2.1 Background.....	11
2.2 Oil	14
2.3 Coal.....	15
2.4 Natural Gas	17
2.5 Electric power.....	18
2.6 Future trends	19
3 Vision of a hydrogen future.....	23
3.1 Hydrogen as a fuel	25
3.2 Production of hydrogen from fossil fuels	30
3.3 Production of hydrogen by electro- and photolysis	32
3.4 Storage and distribution of hydrogen.....	38
3.5 Fuel cells	42
3.6 The process of transition to a hydrogen economy	47
4 Understanding change.....	51
4.1 The physical nature of change	51
4.2 Feedback, path dependence and lock-in	56
4.3 Barriers to change	59
4.3.1 Human costs.....	59
4.3.2 Defence of existing investments	60
4.3.3 Lack of capital.....	61
4.3.4 Lack of knowledge.....	62
4.3.5 Lock-in and catch 22 situations	62
4.4 How to achieve change	63
4.4.1 Describing change by means of activities.....	64

4.4.2 Motivation for change.....	64
5 Facing uncertainty and risk.....	67
5.1 Concepts and consequences of value.....	68
5.1.1 Value related to price.....	68
5.1.2 Value related to competition.....	70
5.1.3 Value related to effectiveness and efficiency.....	70
5.2 Uncertainty and technological change.....	71
5.2.1 Quantitative uncertainty.....	72
5.2.2 Quantitative uncertainty in the public sphere.....	74
5.2.3 Qualitative uncertainty and robustness.....	74
5.3 Risk versus probability and consequences.....	76
5.4 Scenarios as tools.....	77
5.5 How the EICOMP model may reduce risk.....	80
6 Models and modelling.....	83
6.1 Basic concepts of modelling.....	84
6.2 Approaches to modelling.....	85
6.2.1 Basics of system equations.....	86
6.2.2 Complex systems.....	90
6.3 The relationship between models and reality.....	100
6.4 Specification of information to models.....	103
6.5 Value sequences, end-products and activities.....	103
6.6 Porter's model of industry competition.....	106
6.7 System dynamics with causality, feedback and delay.....	108
6.7.1 Basic mechanism.....	109
6.7.2 Networks, causal loops and hierarchies.....	113
6.7.3 Endogenous and exogenous variables.....	117
6.8 Examples of existing energy models.....	119
6.9 Communicating dynamic behaviour.....	126
7 The Energy Infrastructure Competition Model.....	129
7.1 Modelling tools.....	131
7.2 Model rationale.....	135
7.3 Main structure and parts.....	142
7.3.1 Overview of structure.....	142
7.3.2 Specification of region and data input.....	145

7.4 Substructures with mode of operation	145
7.5 Model boundaries and assumptions	155
7.6 The challenge of documenting the EICOMP model	160
7.7 The challenge of data quality	162
7.8 Model output	164
7.8.1 Procedure of output	165
7.8.2 The prepared mind	165
8 Discussion and conclusions	167
8.1 Approach and relevance to purpose	169
8.1.1 Critique of approach	169
8.1.2 Information related to change	174
8.1.3 Positive feedback and increasing returns	178
8.1.4 Modelling of irreversible behaviour	181
8.1.5 Categories of dynamics	182
8.1.6 A basis for detailed studies	182
8.2 Critique of the EICOMP model	183
8.2.1 Choice of structure	187
8.2.2 Degree of aggregation	188
8.2.3 Further development to the EICOMP model	189
8.3 Strategy and how to achieve desired change	190
8.3.1 Decisions and policy making	191
8.3.2 Value today – value tomorrow	193
8.3.3 Positive feedback and strategic opportunities	194
8.3.4 Intermittent supply of renewable energy	195
8.3.5 Investments already made	196
8.3.6 New opportunities for creation of value	197
Epilogue	198
References	201

Introduction

"I believe it is important to view problems in a dispassionate way, to put aside ideology and to look at evidence before making a decision about what is the best course of action"

Joseph Stiglitz, 2002

The purpose of this work is to better understand the dynamic change inherent in transition towards sustainable use of energy. This involves identifying and studying complex feedback structures as they continuously evolve and interact, as basis for sound strategy and policy making. The ambition is not to model energy systems as such, but to achieve practical change.

Three matters or truths appear beyond discussion with respect to the supply of energy to modern society. The first is that modern society as we know it, is totally dependent on a steady supply of energy. If this fails, then production and provision of essential needs and services cannot be upheld. Secondly, the dominant part of the current supply is based on non-renewable oil, gas, coal and nuclear energy. The third matter has gained less attention in public debate, but appears equally clear: fundamental changes in the prevailing pattern of energy supply will require a long time, perhaps 3-4 decades or more, corresponding to the lives of the next generation of people.

The main concern with the supply of energy has traditionally been that of becoming dependent on other nations. This concern is growing more acute. For example, a "green paper" adopted by the

Chapter 1

European Commission in November 1999 starts its executive summary thus:

"The European Union is consuming more and more energy and importing more and more energy products. Community production is insufficient for the Union's energy requirements. As a result, external dependence for energy is constantly increasing. The dramatic rise in oil prices which could undermine the recovery of the European economy, caused by the fact that the price of crude oil has tripled since March 1999, once again reveals the European Union's structural weakness regarding energy supply, namely Europe's growing dependence on energy, the role of oil as the governing factor in the price of energy and the disappointing results of policies to control consumption."

EU, Green Paper, 2000:2

The concern over dependency is thus a concern over resources that are non-renewable, and their origin. Over the last decades a further concern has come into focus, namely the environmental impact from the use of fossil fuels. This may in turn be divided into the three broad categories of local, regional and global. Local effects are to a large extent associated with transport and with some kinds of stationary use of energy: emission of particulates, of carbon monoxide, of polycyclic aromatic hydrocarbons (PAH), of volatile hydrocarbons (VOC) and of oxides of nitrogen (NO_x). The latter two are also intimately involved with the regional problems of photochemical smog and low-level ozone. A further regional concern is acid rain, induced mostly through sulphur dioxide (SO₂) from burning of oil and coal. The main global concern is that of global warming, generally believed to be induced mainly by carbon dioxide from the burning of fossil fuels, but also from the release of methane and some other chemical compounds into the atmosphere.

An impression of the time required for change is gained from the historical picture of changes in energy carriers and infrastructure in Figure 1.1:

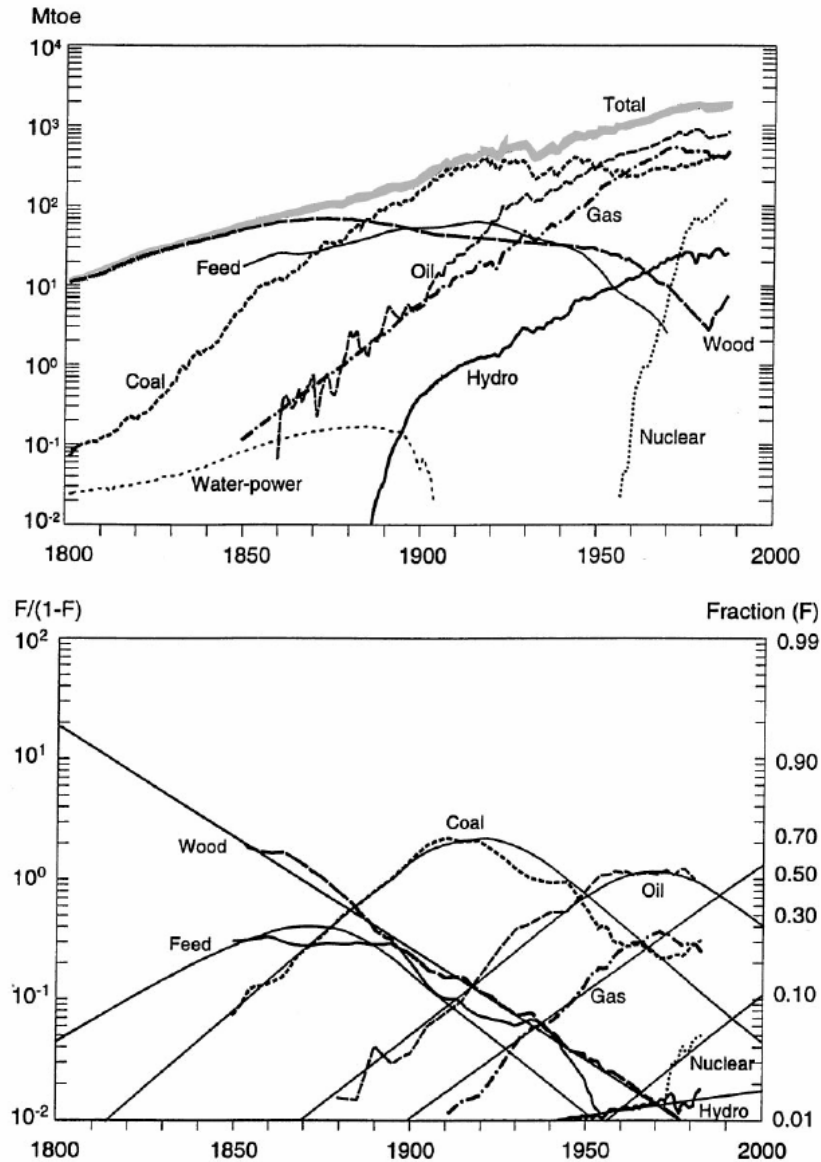


Figure 1.1 *Historical transitions in energy resources in the US.* (The upper graph depicts total use of energy resources. The lower illustrates the fraction covered by each resource)

A. Gröbler et al, 1999:265

Note that the vertical axes are logarithmic: Straight lines correspond to exponential development. The lower figure shows the fractions (F) of total energy supply for each energy carrier.

Another side of this challenge is the wide disagreement over whether or not we will run out of resources or when price will reflect this limit, and when this may happen. Also, there are many possibilities for conversion from one type of fuel to another, albeit at a cost money-wise and in terms of energy efficiency. However, non-renewable resources are just that: non-renewable. No matter how we look at it, as we extract the more accessible fuels, the more inaccessible ones remain. Furthermore there appears to be agreement that unless something is done about it, we will become increasingly dependent on oil from the Middle East. An extensive analysis of development of fuel demand patterns in Europe over the period 1970-1995 is given by Haugland et al (1998), cf. Chapter 2.

A compounding factor is that "business as usual" is accustomed to see steady growth of some few percent every year. This in reality amounts to exponential growth, which is characterised by fixed doubling times. For example, a yearly rate of growth of 2.3% means a *doubling* of consumption in 30 years. In the upper part of Figure 1.1 the "Total" is an almost straight line, demonstrating exponential growth in energy consumption in the reported time frame. In combination with the need for substantial time in order to change into new patterns of energy use, this may have grave consequences unless we heed the need for change in time. Therefore, we need a clear perspective of time as we proceed with these questions. Following Meadows et al (1992) an attempt to illustrate this is given in Figure 1.2:

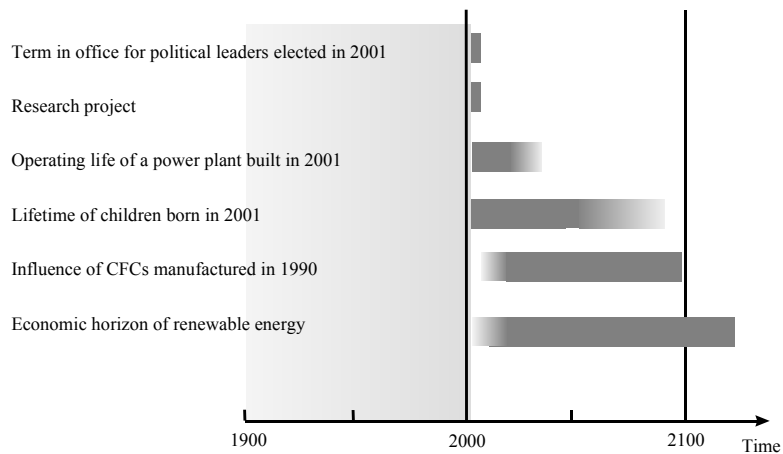


Figure 1.2 *Time perspectives of some administrative and energy-related issues*

Following D. H. Meadows et al, 1992:235

Although there is at present disagreement and uncertainty over when, and to what extent efforts need to be focused towards changing of the current pattern of energy supply, there is broad agreement that we need to know. Also, it is clear that a change-over to more sustainable practices promises great opportunities for new creation of value (cf. discussion in Chapter 2). This is most likely to be reaped by those prepared to take advantage of the new opportunities, which again translates into needs for knowledge. What is the picture ahead of us? What opportunities are present? What should be our goals and our more near-at-hand objectives? What decisions need to be made at the present time? What will be the ramifications of the decisions made? What path dependencies will develop as a result of our decisions, or lack thereof?

This type of questions needs to be analysed both from a societal point of view, and from the point of view of individual firms that will be involved. What new knowledge needs to be built up within the firm? What investments should be made? What markets should be focused on, and with what products and services? At the societal level there

are questions relating both to national interests, and to larger political regions as is the focus of the EU Green Paper (2001) referred to above. Haugland et al (1998:13) underline the multidisciplinary nature and framework of the analysis of these questions. This challenge is deeply dynamic, and development is path dependent. We live in a reality with a single "now". We cannot go back in time and try new approaches or strategies, as what we have already done has set up path dependencies that partly determine the outcome of the next round. In short, what we do and what decisions we make, set the stage for what will happen next, and what window of opportunities we will have. It is this multidisciplinary and dynamic picture that has formed the backdrop and the underlying thinking behind the approach chosen in the present work, listed in Table 1.1:

<i>Local optimisation through market competition</i>
<i>Market preference and choice beyond costs</i>
<i>Bounded rationality</i>
<i>Uncertainty and risk</i>
<i>Irreversibility</i>
<i>Increasing returns</i>
<i>Path dependence</i>
<i>Feedback</i>
<i>Defence of existing investments</i>
<i>Delay</i>
<i>Nonlinearity</i>

Table 1.1 Essential features of reality

With this outset we now turn to how these features of reality may be met. We may create models, which in telling ways imitate the essence of reality, without drowning in detail. The thesis combines the following fundamental modelling tools and concepts to understand *change*:

- *Dissipative, nonlinear and irreversible systems*
- *Value sequences* by means of hierarchies of value-oriented activities*
- *Competition between energy carriers and their technologies*
- *System dynamics to model causality and feedback in networks*
- *Policy instruments and innovation*

(* Value sequences are defined in section 6.5).

An instrument to understand and carry out this process is the Energy Infrastructure Competition Model (EICOMP). It is also intended for communicating results and understanding.

As developed in the following chapters the elements above open for a combined tool to evaluate alternatives including strategy, innovation and economic development. An overview of these chapters is given below.

The first six chapters are devoted to the background (1-3) and tools (4-6) involved in the present work. Chapters 7 and 8 report the outcome, and discuss and critique approach and results.

Chapter 2, *Energy structures and environmental futures*, presents an overall picture of the energy sector that a transition to a sustainable energy structure will have to start from.

Chapter 3, *Vision of a hydrogen future*, has two main objectives: to identify the parameters that appear most important to future development and that are most amenable to change, and secondly to present a picture of sustainability, where hydrogen functions as a substantial energy carrier. A hydrogen future seems to solve a number of problems associated with the current supply of energy, in a single leap:

Chapter 1

- *All carbon-based pollutants are removed at the site of emission*
- *Fuel cells are not essential, but promise both higher efficiency and no NO_x emissions*
- *Hydrogen offers a way of storage of energy from intermittent generation of electric power from renewable sources*
- *In the long run there may be an opportunity for ample sustainable generation of hydrogen from sunlight*

While the potential seems promising, there are still technological and institutional hurdles to overcome, as well as the dynamic challenges associated with overcoming resistance.

Chapter 4, *Understanding change*, seeks to provide a better understanding of the concept of *change* in its own. A first focus is to understand change based in irreversible thermodynamics. We then consider important factors influencing the direction change may take; in particular path dependence and feedback. Finally, we look at factors that form barriers to intended change, and factors that form driving forces for change. Reality is not completely deterministic, nor is there complete randomness.

Chapter 5, *Facing uncertainty and risk*, focuses on three themes that are central when we seek to channel change towards particular outcomes. These themes are respectively the issues of *value*, *uncertainty*, and *risk*. Value defines the direction in which we want change to take place, and allow us to rank alternatives. Furthermore, value strongly depends on *ownership* that defines the framework of stakeholders related to change. In the present work we distinguish in particular between public and private ownership. Uncertainty and risk differs according to ownership. Uncertainty may further be seen with respect to whether or not we will reach our goals quantitatively, or pertains to situations in which value changes qualitatively. We refer

to this as *quantitative* and *qualitative* uncertainty, respectively. Risk is the product of probability and consequences related to uncertainty.

Chapter 6, *Models and modelling*, presents available modelling techniques, limitations, how they should and should not be used, and how models are part of communication. Central themes of this chapter are system equations, the relationship between models and reality, specification of information. This is applied to value sequences and end-products, Porter's model of competition and system dynamics. The chapter also describes the present status of energy models, and how dynamics are visualised.

Chapter 7, *The Energy Infrastructure Competition Model*, describes the model developed in the present work. The purpose is to study how transition will play out due to the system's structure, causality, feedback and the many changes that are taking place. A main focus is the integration of hydrogen related technology into existing energy carriers and associated infrastructures. The goal is to be able to identify and understand reinforcing- and balancing feedback loops. Such loops are central mechanisms in the system, and form the basis for both path dependence and business opportunities. In turn this becomes the basis for policy making. The purpose is thus not to model energy systems as such. The chapter is concerned with selection of tools, rationale, main structure and parts, substructures, boundaries and assumptions. The last part of the chapter describes the challenge posed by data quality and documentation of the model.

Chapter 8, *Discussion and conclusions*, presents a critique of the approach and the outcome of the present work. The main themes are the appropriateness of the approach and methodology with respect to purpose, the perspective of moving modelling closer to where decisions are made, and strategy for achieving sustainable change. The most concrete outcome is the functioning model (EICOMP). Further results have been obtained from the process of developing

Chapter 1

and constructing the model. A clearer picture of the overall problem is now available. The work on dynamic models reveals the excellent tools for conveying results within the modelling software itself. These tools are far superior to what may be provided as written texts like a dissertation.

From this overview we now turn to the energy structures and environmental concerns that form the basis from which change has to start.

Energy structures and environmental futures

“Without modern energy systems, society as we know it today would cease to exist. Providing adequate energy supplies at reasonable prices has been an integral part of modernisation and nation-building throughout the western world.”

T. Haugland, H. O. Bergesen and K. Roland, 1998

The purpose of this chapter is to present a broader picture of the existing energy sector. This is where change will have to start.

2.1 BACKGROUND

In this Chapter we follow Haugland et al's (1998) extensive review of the energy sector. These authors review the energy sector in Europe and in the former Eastern Europe from a political and economic perspective. This perspective complements the basis in technology otherwise in focus in this thesis. The authors present a broad social and political picture of the energy structure as it has evolved in Europe. They divide the picture since World War II into three main periods:

- 1945 - 1960s: *Reconstruction and industrialisation*
- 1960s - 1970s: *Security of supplies*
- 1970s - 1980s: *Environmental concerns*

Chapter 2

The discussion since then has increasingly come to focus on sustainability (*Our Common Future*, 1987 and *The Kyoto Protocol*, 1997).

Following WWII, some countries experienced large national monopolies through nationalisation of energy enterprises, with public ownership as the primary organisational pattern. In other European countries a decentralised structure was maintained, as it also was in the US. A main issue was simply to provide the required infrastructure to deliver an adequate energy supply. This led to capital intensive and large-scale vertical integration.

In the 60s and 70s, the geopolitical aspects of energy supply came more to the forefront. In particular, the supply of oil came into public focus. Steps were taken initially (Haugland et al, 1998:3 ff.) to balance the influence of the multinational oil companies. The oil crises in 1972 and in 1979 lead to a shift away from oil to generate electricity. The pressure on electricity generation was further intensified by the Three Mile Island incident in 1979 and the Chernobyl disaster in 1986. However, the basic supply structure, as well as the role of public authorities remained largely unchanged until the 1990s. From the 1970s onwards, the consciousness of environmental impact has been increasing. Prior to this, the shift from local combustion of coal to combustion in large plants with extensive cleanup-facilities, had lead to a general impression of environmental improvement. Then came the problems with acid rain, and with ground-level ozone, NO_x and photochemical smog. These problems could not be solved through improved end-of-pipe cleanup. Meadows et al (1992) make the following comment regarding the exceeding of sustainable limits:

“The human world can respond in three ways to signals that resource use and pollution emissions have grown beyond their sustainable limits. One way is to disguise, deny, or

confuse the signals: to build higher smokestacks, for instance, or to dump toxic chemicals secretly and illegally in someone else's territory; to overexploit fish or forest resources knowingly, claiming the need to save jobs or pay debts while in fact endangering the natural systems on which jobs and debt payments depend... ...A second way to respond is to alleviate the pressures from limits by technical or economic fixes without changing their underlying causes: to reduce the amount of pollution generated per mile of driving or per kilowatt of electricity generation... ...The third way to respond is to step back and acknowledge that the human socioeconomic system as currently structured is unmanageable... ...and therefore, to change the structure of the system."

D. H. Meadows et al, 1992:190-191

From the mid 80ies onwards the strategic challenges have been growing more complex, in part from the increasing local and regional environmental impact from fossil fuels, and in part because of increasing dependence on oil supplies from the Middle East. Also, in this period the problem of global warming has come to the forefront of political consciousness. Haugland et al (1998) identify five major challenges they believe will dominate energy development in Europe in the years to come:

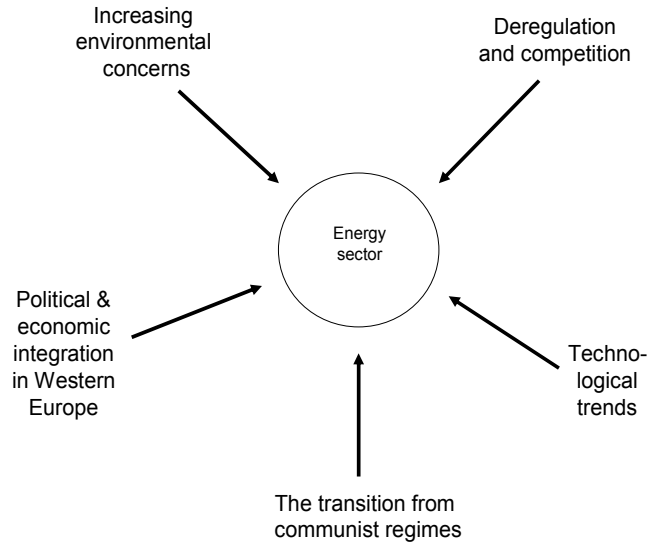


Figure 2.1.1 *Major challenges to the European energy future*

T. Haugland et al, 1998:6

The energy structure has three fundamental sides: a technological, an institutional and a socio-economic. The picture may be discussed in general terms only to a limited extent. For further understanding it is necessary to look at the major sources and carriers of energy. In the following sections we briefly discuss the sectors of oil, coal, natural gas and generation of electric power.

2.2 OIL

The supply of oil has clearly been important ever since the shift from coal to oil after World War I, and was a central strategic issue in World War II. Until the second part of the 1950s the supply of oil was in the hands of three European and five American giant, vertically organised companies (“the Majors”). Oil prices rose sharply in 1973-74, and from 1978 onwards, but have since 1986 stabilised somewhere below the prices following the first oil crisis

(Haugland et al, 1998:57). European oil consumption reached an all-time high of about 750 mtoe (8723 TWh, [heat]) in 1973, fell to about 600 mtoe in 1983-84, and has been growing steadily (exponentially) since then to 664 mtoe in 2000 and projected to 717 mtoe in 2010 (EU Energy and Transport – Trends to 2030:27). From a situation of almost exclusive imports, European production increased to about 40% in 1994 (Haugland et al, 1998:58). There has been a major transformation in the European oil market since the 1970s. From a situation where “the Majors” dominated, they have lost control of the vertical chains from production to distribution. Barriers to trade in oil and oil products have largely disappeared, with increasing international competition.

The early period was marked by a “social contract” between nations and the oil industry that protected the latter and made it possible for the oil companies to remain as reliable suppliers. There has been a shift towards markets and increased competition, where the markets appear able to maintain sufficient reliability of supply.

2.3 COAL

Whereas oil sees markets that will absorb all it can get, coal has met with a quite different situation. Since the early 1970s most German and French production has been uneconomic, and major parts of the production in Spain and in the UK would not have survived without protection. Coal was the most important energy source up until 1966, when oil took over (Haugland et al, 1998:73).

Since then, natural gas has increasingly become a competitor both for household heating and generation of electricity. The major use of coal since 1970 has been in the steel industry and in generation of electric power. West European consumption of coal provided 85 % of overall energy in 1954, peaked in 1956 at about 575 mtoe, and fell to 34 % of all energy in 1970. The coal share in total energy consumption in

Chapter 2

2002 was 14%, amounting to 217 mtoe (BP Statistical Review 2003:38).

Another side to this picture is seen from the workforce involved. In 1955 the German hard coal industry employed 600,000 people (Haugland et al, 1998:73-96). In 1994 this had sunk to 102,000. The industry has in later years remained at this level, delivering about 40 mtoe a year, mostly for electric power. Domestic hard coal is heavily subsidised: production costs are given as US\$ 160/tce (tonne coal equivalent) in 1994, as against import prices of US\$ 40/tce at the same time. On the other hand, Germany (in particular the former East Germany) has a considerable production of brown coal (lignite) that is not subsidised. All of this is used for production of electric power.

The British coal industry was nationalised after World War II. Since 1970 production has about halved, from 85 mill. tonnes to about 40 mill. tonnes. About half is used for generation of electricity, and the remainder for industrial purposes. In the mid-1990s some 70-80 % was produced in the UK, with the remainder imported. Following the national miner's strike and political fight with the Thatcher government, 120 out of 170 pits were closed between 1975 and 1992, and employment in the coal mining industry fell from 220,000 to 54,000. The British production of coal has continued to fall in recent years, as generation of electricity from natural gas in combined-cycle generation has taken over. This type of generation at the same time delivers low temperature heat for heating of buildings. The picture seen in Britain is representative of all of Europe, with security of supply remaining as the central argument for keeping up production of coal at all. It is more expensive, and more polluting, than all alternative sources of primary power. Haugland et al (1998) concludes that the security of supply is no longer a sufficiently weighty argument, and that the industry is subsidised annually to the equivalent of US\$ 60,000 per employee. Their comment on the time perspective is:

“In Central Europe, and particularly in Poland, a larger part of the coal production is economically viable. Still, there are strong economic and environmental arguments for reducing production drastically. In the short and medium term, however, both employment and security-of-supply considerations may work against any swift changes, in addition to the substantial retrenchment that has already taken place.”

T. Haugland et al, 1998:96

2.4 NATURAL GAS

Natural gas is the most environmentally friendly of the fossil fuels, with NO_x and carbon dioxide as the main effluents. It differs from oil and coal in that it is distributed mostly through pipeline grids and cannot be stored in large volumes. The use of gas started in Europe in the 1960s, grew rapidly in the 1970s, and became an international commodity with main pipelines from Algeria, the Netherlands, Norway and Russia from the 1980s onwards. In 2002 natural gas provided about 23% (351 mtoe) of all energy consumption in Europe (BP Statistical Review 2003:38) and is increasing, in particular for generation of electricity. Natural gas is an alternative to oil in many applications, in particular for stationary use, and eases the European dependence on Middle East oil.

Natural gas transmission and distribution have traditionally been state-regulated monopolies, although the organisational structure is changing. In Germany for instance, contracts for natural gas e.g. from Norway, as well as distribution to end-users, are handled by a set of private companies. The UK started out with a single state-owned company (British Gas) that has later been privatised and now competes with a set of importing firms, although it still remains by far the largest. Apart from the UK, the commercial structure of the

European natural gas industry has remained stable since the early 1970s. It remains dominated by a limited set of large transmission companies. Some of these are state-owned and some are private, or with a shared public and private ownership. All engage in social contracts with the respective national governments.

Haugland et al (1998) conclude with two main points: firstly, technological development makes natural gas increasingly applicable and economically attractive. Secondly, if consumption of natural gas develops along present trends, then Europe will sooner or later be facing major supply constraints. When this will happen, depends on the extent to which large new natural gas fields will be found, changes in patterns of consumption, and supply of oil to the transport sector.

2.5 ELECTRIC POWER

Electric power is not a source of primary energy, but is a way of mediating such primary energy to the host of applications it has in modern society. It is not easily stored, and the energy technology for electric power is deeply influenced by the need to adapt to changing consumption with time, and the difficulty and costs invoked in achieving this. Electric power and transmission have historically been highly regulated. However, since the early 1990s a pressure for change and deregulation has become clearer. The main policy objective has been to provide adequate supplies at lower costs. Electric power consumption increased fivefold in Western Europe from the 1950s to about 1975, with a rate of growth of 6.5 % annually. Prices fell throughout this period. The oil crises in 1973 and 1979 lead to a massive shift away from oil, to natural gas, and back to coal. At the same time heavy investments were made in nuclear power generation, particularly in France, but also in Germany and the UK, and to a lesser extent in Belgium, Spain and Sweden. However, the Chernobyl disaster in particular, and also the difficulties with

final storage of spent fuel, has more or less stopped the expansion of nuclear technology for now.

Norway is in a special position as almost all electric power is renewable hydro power, and also is easily adapted to varying demand. Furthermore, Norway has instituted a radical change in organisation of supply and distribution. Briefly, generation was separated from distribution, with common carriage for all generating facilities through the grid, and also for all end-users. An open market was instituted for suppliers and consumers, constituting respectively a spot market, a contract market, and a “regulation market”. This “electric power pool” is now extended to the Nordic countries (Nord Pool). Experience is at present being gained with the pros and cons of this regime.

2.6 FUTURE TRENDS

Haugland et al (1998) portray an overall framework for the energy sector. It is projected into two scenarios for the further development of the European Communities. The first portrays a “national rebound” with continuation of individual nations in dominant roles. The second scenario is a picture of liberalisation and trade in which political and economic integration proceeds rapidly in depth and throughout the community. In the first scenario trade in natural gas and electric power remains under national control, and research and development is directed so as to support the national energy industries. In the second scenario the top political priority is to stimulate energy markets.

In both scenarios it appears that technological and environmental forces will decide much of the final outcome; more so than the economic and political forces. Oil and natural gas found in new reservoirs will alleviate the extent to which Europe will become dependent on the Middle East. Likewise, the social consensus on the

environmental impact from the use of these fossil fuels will largely determine their pattern of use.

At present, global warming from CO₂ as a greenhouse gas, is an overriding concern for the acceptability to use carbon based fuels. However, there is still discussion of the issue of global warming.

For the European region the EU Directorate-General (European Energy and Transport, 2003:24) assumes a continued growth in GDP until 2030 at about 1.9% annually for the OECD region, and 2.9% for the world at large. For Europe the EU Green Paper (2000) argues that if no measures are taken in the next 20-30 years, 70% of the Union's energy requirements, as opposed to the current 50%, will be covered by imported products. Renewables currently provide 6% of the total supply and if nothing is done is expected to grow to 8%.

Energy information is often displayed as the relative importance of different primary sources, with 100% in each interval of time. This hides the increase in physical volume. Figure 2.6.1 illustrates this phenomenon:

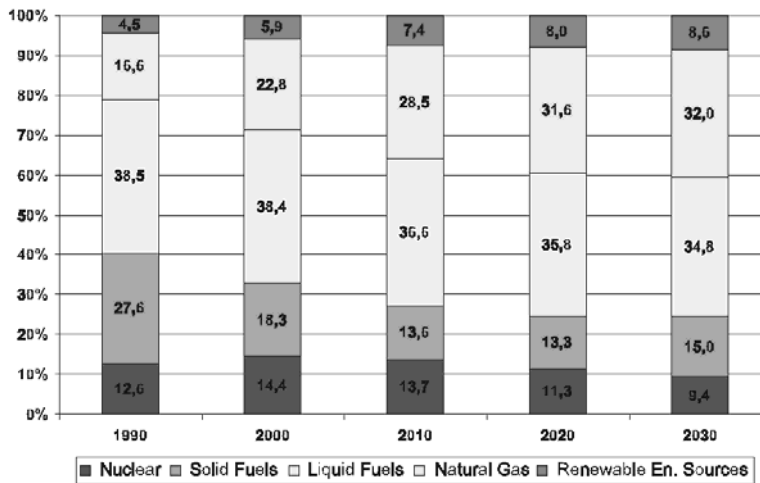


Figure 2.6.1 Structure of primary energy demand in EU, 1990 – 2030
European Energy and Transport, 2003:111

The column to the right (2030) in Figure 2.6.1 implies (with 1.9 % annual growth) twice the amount of energy compared to the leftmost column (1990). Much economic and political discussion seems not to realise the effect of this.

In a world perspective, as mentioned in the Introduction, there appears to be little danger of actually running out of fossil energy. However, historical energy consumption has increased four-fold from 1950 to 2000 (cf. Figure 1.1). Assuming constant energy intensity with annual growth in GDP of 2.9 %, this corresponds to another doubling of energy consumption within approximately 24 years. In the last few years a decrease in energy intensity is observed. The overall impression is however, that an all out effort towards sustainability is called for.

The situation appears even more serious when considering the current development in China and India. The BP Statistical Review 2003 reports growth in energy consumption in China in 2002 at close to 20 %, which as exponential growth amounts to a doubling in 3.5 years. The growth in coal consumption alone is reported at 28 %. This amounts to an increase in physical volume of 145 mtoe in one year. For comparison this corresponds to nearly half of Germanys total energy consumption in 2002! China also consumed twice as much oil as Germany. The competition over Middle East oil may well increase.

The energy structures outlined by Haugland et al (1998) appear commonsense and likely unless the effects of exponential growth are taken into account. The existing energy infrastructure further represents large investments in entrenched behaviour, in physical equipment and in knowledge. Not only technology, but also the behaviour of institutions leads to path dependence:

Chapter 2

“As modern technological systems are deeply embedded in institutional structures, these factors leading to institutional lock-in can interact with and reinforce the drivers of technological lock-in.”

T. J. Foxon, 2002:3

This creates networks that are difficult to break away from. It is believed that improved understanding of these network effects, how they interact, and how they influence economic and social structures of power - will be of outmost importance if we are to change a pattern that in the long run is not sustainable. A central part of this is for those involved with developing a hydrogen infrastructure to build coalitions with powerful established interests.

Hopefully it will be possible to transfer to a quite different energy regime. It will have to be much more parsimonious in its use of energy to achieve the products and services required by modern society, and based on principles that promise both sustainability and drastically reduced environmental impact. It is important that change occurs smoothly, rather than in abrupt shifts.

Many speak of hydrogen as the ultimate fuel in a sustainable future. They may well be right, but how do we get there? The changeover to sustainable energy practices is disappointingly slow. The new has to be phased in while at the same time the old must be phased out. This is J. A. Schumpeter's famous "creative destruction": stakeholders tend to defend their interests (cf. networks above). In the next chapter we seek to outline important elements of a “hydrogen future”.

Vision of a hydrogen future

*Problems worthy
of attack
prove their worth
by hitting back*

Piet Hein

Hydrogen, particularly in combination with fuel cells, is seen as a possible solution towards renewable energy and dramatically reduced environmental loads. It would make clean energy available from a wide set of primary resources and allow a smooth change-over from the present regime of carbon-based fossil fuels. There is excitement over a future 'Hydrogen Economy', but also doubt and considerable discussion. A glimpse of this is:

“Even as a few doubters question the economics and wisdom of this revolution, today’s stewards of conventional wisdom question not whether the hydrogen revolution will occur, but rather, the exact timing and sequence of events that will propel modern society to that shining hydrogenous city on the hill.”

D. G. Victor et al, 2003:1

This chapter has two main objectives: to identify the parameters that appear most important to future development and that are most amenable to change, and secondly to outline a picture of what a hydrogen economy could look like.

The present exposition follows the pattern from the Hypothesis II seminar in Grimstad, Norway in 1997 (Saetre, 1998), which presents a comprehensive overview of basic facts and processes. It is augmented by newer reviews and overviews, notably Carrette et al (2001), Argonne (2003) and Glöckner (2004). While there is continuing improvement in many fields, no ‘beyond doubt’ route to a hydrogen future stands out at present.

The idea of hydrogen as a modern-day fuel seems to have had its birth with the idea that engines running on hydrogen would emit only water vapour. There would be no more carbon monoxide, no polycyclic aromatic hydrocarbons (PAH), no volatile organic compounds (VOC) and no more soot and grime. There might still be some NO_x from internal combustion (cf. section 3.1), but overall, hydrogen would represent a drastic improvement compared to the pollution now experienced with carbon-based fuels and corresponding technology.

Hydrogen gained further momentum as it was realised that sustainable generation in large quantities could be possible directly from renewable energy sources. To add to this picture, dependable fuel cells were developed in the US ‘man-on-the-moon’ program, converting hydrogen directly to electric power and water vapour with higher efficiency than attainable by means of ordinary combustion engines (ICEs), working on the Carnot principle, within practical feasible temperatures (cf. section 3.5).

However, while this vista of environmentally clean, sustainable energy unfolded, it also became clear that there is a range of hard technological hurdles to be overcome. Provided these are surmounted, there is further the question of actually moving from one regime of technology to another without turning the economy on its head, and to achieve this while we still have fossil fuels to see us by.

Rosenberg (1996) points to the historical evidence for long times required for change:

“...It took some 40 years or so before electric power came to play a dominating role in manufacturing. History strongly suggests that technological revolutions are not completed overnight.”

N. Rosenberg, 1996:344, (in Landau et al, 1996)

Thus the benefits of hydrogen will require time before taking effect. In the following we briefly present and discuss the central issues involved.

3.1 HYDROGEN AS A FUEL

The first thing to appreciate is that hydrogen, like electricity, is merely an energy carrier. Therefore a kilowatt-hour provided by hydrogen will in theory require a kilowatt-hour from a primary energy source. In practice considerably more is required, due to inefficiencies in the processes that convert primary energy into hydrogen. As an energy carrier, hydrogen is only part in a sustainable energy regime to the extent that the primary energy resources are sustainable. However, as hydrogen may be generated efficiently from natural gas, this may create familiarity and thus ease the transition to renewable energy.

Stationary and mobile use of hydrogen differ in their requirements for end-use and storage, and it is useful to discern between these sectors. However, they share upstream facilities so that development and improvements in one sector will influence the other. Natural gas for stationary use is most often supplied through networks, with main and local pipelines consisting of compressors, lines and valves, ending in stationary outlets. These outlets are equipped with metering facilities to allow monitoring of consumption, and billing similar to

electric power. Much of this infrastructure is in place for stationary use of hydrogen as well. A substantial part of the technology and practical training is already in place both for suppliers and customers, although there is uncertainty as to the extent to which the existing ‘hardware’ may be used for hydrogen (cf. Iskov, 2004).

The existence of competence related to natural gas is an example of path dependence – it is easier to introduce hydrogen into stationary use since both the technological know-how and familiarity with combustible gas by users are in place.

The requirements of hydrogen as a fuel are the same as for carbon-based fuels: it must be available where needed, and it must make energy available in a suitable form. Hydrogen has the property of making energy available in two forms: directly as electric power by means of fuel cells, and as heat through reaction with oxygen. The latter is further described in section 3.5.

Hydrogen has the following main technical specifications:

<i>Lower heating value</i>	10,800	kJ/Nm ³
	120,000	kJ/kg
	3.00	kWh/Nm ³
Density (gas)	0.09	kg/Nm ³
Density (liquid)	70.9	kg/m ³

Table 3.1.1 *Practical properties of hydrogen*

C. J. Winter, J. Nitsch, 1988:XII

Lower heating value (LHV) is enthalpy minus heat of evaporation of water. One Nm³ of H₂ corresponds to 0.3 litre of diesel fuel, and to 0.68 litre of methanol. The energy stored in hydrogen is liberated as heat through the chemical reaction:



with a Gibb's energy of 237.2 kJ/mole (LHV=241.8 kJ/mole) corresponding to an electrolysis voltage of 1.23 V (e.g. Andreassen, 1998:95). The heat from this reaction may be utilised in an ordinary internal combustion engine (ICE) with small modifications to the engine itself (cf. below). The technical details of this have been worked out (cf. Skjølsvik, 2003). While bounded by the Carnot cycle limits to efficiency, ICEs provide a direct route to utilisation as hydrogen becomes available, based in familiar and affordable technology.

The most critical parameter for the use of hydrogen in ICEs, is in an environmental context the local release of NO_x. However, as hydrogen allows operation of engines under very lean conditions, this seems to be manageable: Hydrogen burns over a much larger air/fuel-ratio than petrol, allowing engines to run on correspondingly leaner mixtures. At air/fuel ratios above 2.0 there is practically no release of NO_x to the atmosphere. Figure 3.1.1 illustrates this clearly:

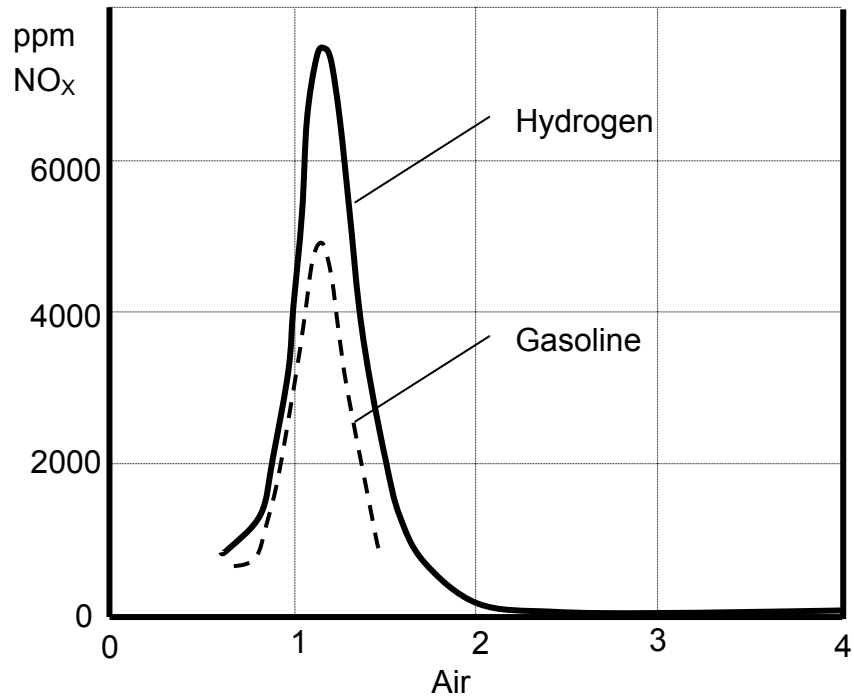


Figure 3.1.1 *NO_x v/s air-ratio for ICEs*

H. Buchner, 1995:178

Wong and Karim (1998:63) report on the concept of adding hydrogen to ICEs running on natural gas, in order to allow such engines to run on leaner mixtures and reduce NO_x-emissions drastically. Adding hydrogen to existing fuels again provides an opening for hydrogen to get started.

Cost of vehicles running on hydrogen relates in part to the cost of hydrogen itself, and in part to the cost of onboard storage. Figure 3.1.2 presents an overview of the relationship between various energy carriers from 1970 onwards, assembled by Victor et al (2003) based on a number of sources. The hydrogen prices are estimated from data on the value and volume of hydrogen shipments from the US Government Bureau of Census, United States Department of

Commerce (various years) and Chemical Economics Handbook Report by Suresh (2001). All other fuel data are from EIA. Further data on prices of hydrogen are given in ICCEPT (2002:92). In the early 1990s the cost of hydrogen was estimated at some € 0.30 per equivalent of 1 litre of gasoline, if produced from natural gas, and about € 1.50 if provided by a local electrolysis unit, using electric power based on coal (Buchner, 1995:185). The cost of onboard storage is discussed in more detail in section 3.4.

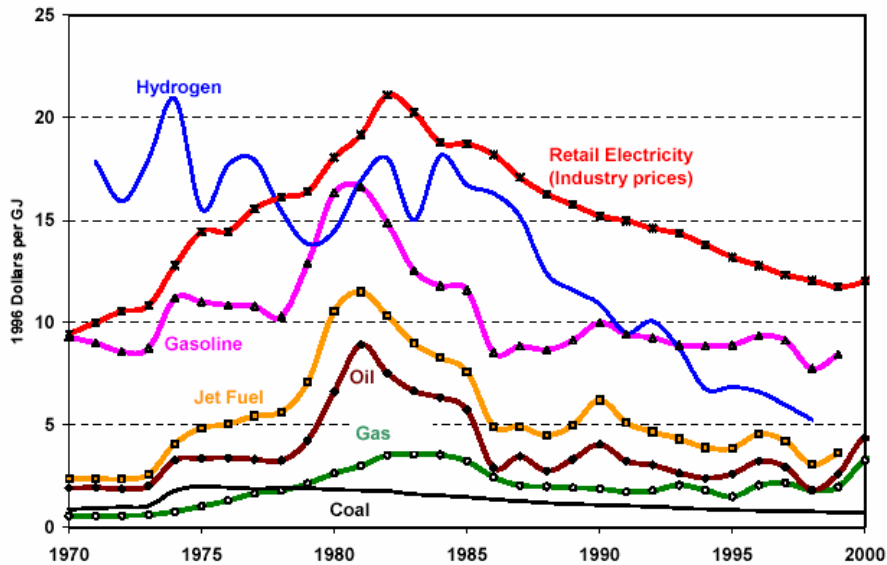


Figure 3.1.2 *Development in prices for Major Fuels and Energy Carriers, 1970 – 1999*

Figure 3.1.2 illustrates clearly that the price of hydrogen has shown a decreasing trend since 1985 and is getting more competitive. Buchner (1995:186) cites the main obstacles beyond costs, to be extensive regulations and standards dealing with H₂-based vehicles and filling facilities, and the insufficient supply infrastructure available until hydrogen-based vehicles gain a reasonable market volume. One needs, however, to keep the scales of usage in mind. Argonne (2003:11) gives the volume of hydrogen for the US road-based

transport sector, if hydrogen was to replace carbon-based fuels now, at 200 million tons a year. This is expected to grow to 265 million tons a year in 2020. If all natural gas produced each year in Norway was to be converted to hydrogen, it would cover about one third of this.

3.2 PRODUCTION OF HYDROGEN FROM FOSSIL FUELS

Our aim here is an overview in order to understand the most critical parameters involved. We follow in the main a review presented by Gaudernack (1998:75). While there is steady progress, no radical breakthrough has taken place in the six years since then. The area mostly pertains to well known large scale industrial processes. A particular issue is that efficient large scale production, based on natural gas in association with its recovery and clean-up, meets with the problem of transporting the hydrogen to where it will be used. If solutions are found, then large amounts of CO₂ from production of hydrogen might be reinjected into reservoirs for increased recovery of oil, providing a much better economy.

In 1997 some $5 \cdot 10^{11}$ Nm³ of hydrogen gas was produced world-wide, from several main sources:

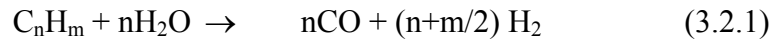
Origin	10 ⁹ Nm ³ /year	%
Natural gas	240	48
Oil	150	30
Coal	90	18
Electrolysis	20	4
Sum	500	100

Table 3.2.1 *World-wide production of hydrogen in 1997*

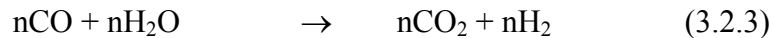
B. Gaudernack, 1998:75

As discussed in section 2.6 the trend is towards increased use of natural gas. This has a bearing on the choice of fossil fuels for making hydrogen as well. There are three principal processes involved in chemical generation of hydrogen:

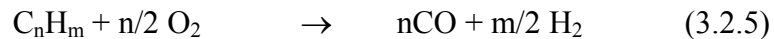
Steam reforming (SMR):



Shift reaction:



Partial oxidation (POX):



The sum reaction for (3.2.2) and (3.2.4) for methane is:



The SMR is strongly endothermic (222 kJ/mole CH₄ under standard conditions). Energy must be provided in order for the reaction to proceed from left to right, and the required reaction takes place under high temperature with a nickel catalyst. The transfer of heat to the reaction has about 50 % thermal efficiency. The shift reaction is exothermal and thus delivers energy in the form of heat. It operates at lower temperatures and pressures, and is not easily integrated with the steam reforming reaction. It furthermore fails to fully cover the energy needs of the steam reforming reaction. Partial oxidation (POX) in combination with the two first makes everything go together (Carrette et al, 2001:19-23). POX is commonly used alone in refineries to provide hydrogen for removal of sulphur and recovery of light fuels from heavy (high-carbon) oil residues.

As normally carried out, production of hydrogen from carbon fuels takes place in large plants producing large quantities of hydrogen. This requires particular competence in plant operation and would seem to require a separate network for distribution of hydrogen. Attempts have been made to develop smaller, integrated systems that require less space and are more amenable to distributed installations. Examples of this are the Topsøe heat exchange reformer, and the Katalco gas-heated reformer (Gaudernack, 1998:80-81). It appears possible that small plants developed along these lines could supply hydrogen on a distributed basis, from natural gas available through established networks.

A particular problem with the shift reaction (eq. 3.2.3 / 3.2.4) is that it leaves a residual of carbon monoxide (CO) in the hydrogen. This has to be removed as it poisons the catalysts of fuel cells working below 200 °C. Hydrogen produced through electrolysis does not have this problem (cf. section 3.5).

In summary, we see that large scale production of hydrogen from carbon-based fuels is well established technology, but in its present form requires a distribution network to make it available from large plants. Generation near the sites where natural gas is obtained would make feasible recovery of CO₂ for re-injection in wells, but the transport of hydrogen to where it is needed becomes even more difficult and expensive.

3.3 PRODUCTION OF HYDROGEN BY ELECTRO- AND PHOTOLYSIS

In contrast to the chemical route to hydrogen, electrolysis easily adapts to small-scale and local generation. Also, the hydrogen produced is very clean, and does not cause problems with poisoning of catalysts encountered in some forms of fuel cells. The headache is

the cost of electric power; without this, electrolysis would provide an almost perfect road to hydrogen. The principal operation of electrolysis is shown in Figure 3.3.1:

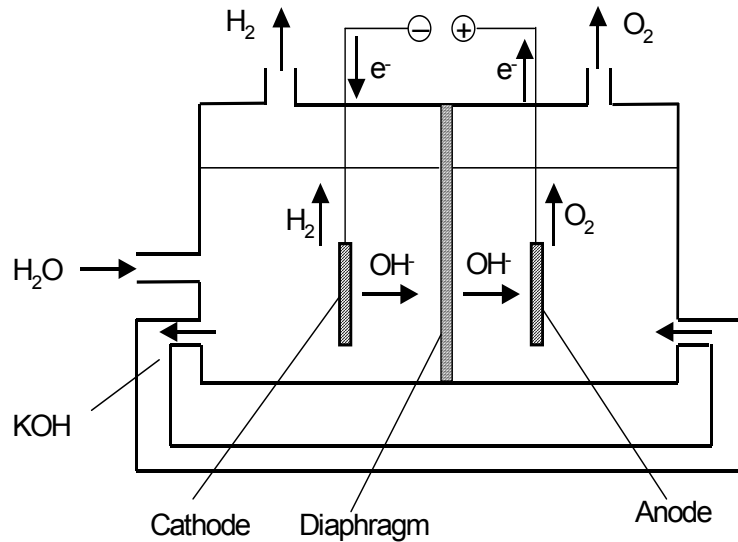
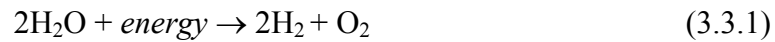


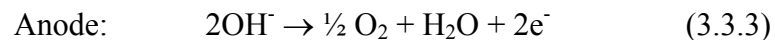
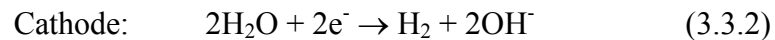
Figure 3.3.1 *Principal electrolysis cell*

K. Andreassen, 1998:93

Water is a poor conductor. In order to keep electric resistance low an electrolyte is used, which in larger units is commonly 20-30 % KOH, giving less problems with corrosion than acidic electrolytes. The overall reaction is:



Under alkaline conditions the electrode reactions are:



The diaphragm serves to keep hydrogen and oxygen apart. Many electrolysis cells are coupled together, both in series and in parallel. In series, each electrode serves as anode in one cell and as cathode in the next, with the DC current voltage applied to each end of the full electrolyser. Such units are called “filter press” electrolysers from their similarity in appearance to this type of equipment. The well-known “Hydro” electrolysers from Norsk Hydro ASA in Norway are built in this way.

A modern development is electrolysers with proton exchange membranes made from polymers. The polymers in question are commonly perfluoro sulphonic acid. The principle for these electrolysers is shown in Figure 3.3.2.

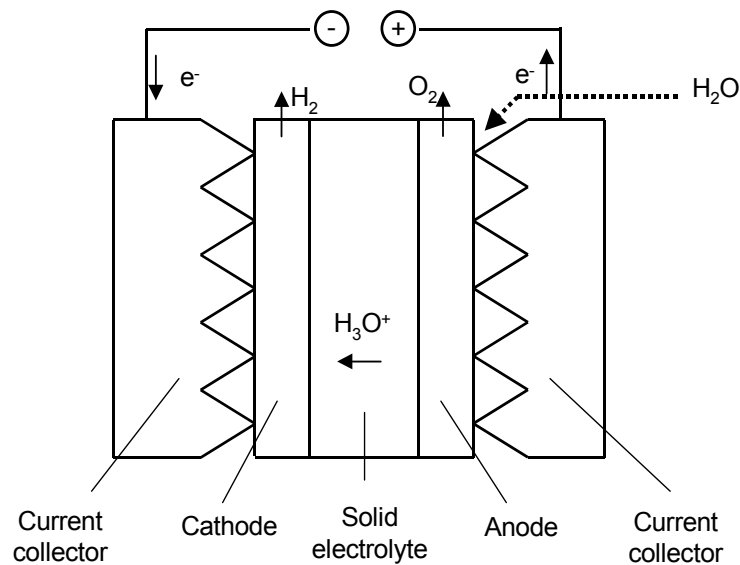
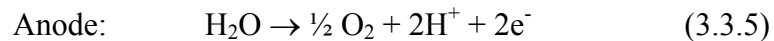
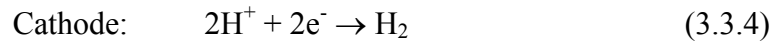


Figure 3.3.2 *Acid membrane electrolyte*

K. Andreassen, 1998:94

In this case the transfer of current through the cell is by means of protons, and the electrode reactions are:



It follows from these equations that water need only be supplied at the anode.

In these cells the solid electrolyte also serves to keep hydrogen and oxygen apart. The electrodes are placed in contact with the solid electrolyte, and consist of catalyst materials (cf. below). Ceramic materials may also serve to conduct protons, to allow operation at temperatures above 100 °C. Electrolysers with proton-conducting membranes give high cell efficiencies at high current densities, and thus allow compact construction. The main problem at present is high costs for the membranes and catalysts used in the electrodes.

The energy picture of electrolysis is determined through a set of voltages:

$$U = U_{\text{rev}} + \eta_{\text{C}} + \eta_{\text{A}} + j \cdot R^* \quad (3.3.6)$$

Where:

U_{rev} = the reversible voltage corresponding to the enthalpy of the reaction

η_{C} = the overvoltage at the cathode

η_{A} = the overvoltage at the anode

j = the current density

R^* = Σ electric resistance

A more detailed discussion is given by Andreassen (1998:95-97), and there is an extensive literature on practical construction and means for overcoming the over-voltages in particular. An example cited by Andreassen (ibid.:97) gives the following values:

Chapter 3

➤ Sum electric resistance:	$110.6 \cdot 10^{-6} \Omega$
➤ Reversible cell voltage (ΔG):	1.189 V
➤ Over-voltage ($j=0.2 \text{ A/cm}^2$)	0.325 V
➤ Ohmic resistance	0.221 V

This example gives an energy efficiency of $100 \cdot \frac{1.189}{1.189+0.325+0.221} = 68 \%$. Losses from over-voltage are 19 % and from electrical resistance 13 %. The example has an energy requirement of 4.23 kWh/Nm³ H₂ gas. Andreassen (ibid.:102) lists the following possibilities for increasing energy efficiency:

- *Amend the electrolysis conditions (lower current density, rise the pressure, rise the temperature, change the electrolyte).*
- *Reduce the resistance in the electrolyte layer between the electrodes. (diminish the gap, reduce the gas/electrolyte volume ratio, secure the electrolyte circulation).*
- *Reduce the resistance in the diaphragm (choose another diaphragm type).*
- *Reduce the resistance in the electrodes (the fore and back electrodes).*

In order to reduce over-voltages:

- *Amend the electrolysis conditions (lower the current density, rise the temperature).*
- *Improve/stabilise the activity of the electrode catalysts/active coatings).*

K. Andreassen, 1998:102

In general, the losses due to over-voltage and electric/ionic resistance appear as heat. This heat may be used to lower ΔG by operating at higher temperature. The limit in this regard is dictated by finding materials that will stand up to the temperature and operating

conditions. There is thus a broad agenda for research towards improved efficiency and practicality of large scale electrolysis.

In a long term perspective, the supply of hydrogen from sunlight may prove a route to a sustainable future. An extensive review is provided by Argonne (2003:14 ff.). The two central requirements are access to energy and keeping hydrogen and oxygen apart. Sunlight provides the required primary energy, and may be converted to electricity for electrolysis by means of photo-voltaic (PV) elements. Electrolysis will make hydrogen (and oxygen) available, well separated. The main problem at present is prohibitive cost, primarily for PV. Current cost for photovoltaic generation of electric power is cited by Argonne at about USD 3.50/peak watt, whereas break even with fossil fuels is at about USD 0.20/peak watt (ibid. 21).

Energy efficiency has long been at some 6-9 % for thin-film PV elements that are suitable for mass production, and some 20-23 % for elements cut from crystalline silicon (ibid.:15). A theoretical limit for Shockley-Queisser type devices (conventional silicon) is 32 %.

Recently efficiencies of more than 30 % have been announced in laboratory trials with new types of devices. There is, furthermore, hope that electrically conductive polymers may be engineered to function as PV-materials, to be produced efficiently in large quantities. The array of new approaches is extensive. While we may glimpse the outlines and the potential appears substantial, time will tell what the dominating technology will turn out to be in the future.

A different approach seeks to find a direct route from light to hydrogen. This is what plants do in photosynthesis, although the normal process accumulates the energy in the form of carbohydrates rather than as hydrogen. Cyano-bacteria that produce hydrogen are, however, known. Chemical approaches have tried to exploit the porphyrine structure utilised by plants (cf. e.g. Ion and Fara, 1998:

161), as well as the manganese-based catalyst for oxygenation of the water molecule. Living organisms also utilise these structures for photosynthesis and for binding of oxygen (haemoglobin and myoglobin). Artificial approaches must master the problem of keeping hydrogen and oxygen apart. A detailed treatment leads too far from the present work. A wider discussion and references are provided by Argonne (2003:24 ff.).

Although important, and perhaps essential in a longer perspective of time, the recovery of energy from the sun by means of hydrogen appears not to critically influence the development of a “hydrogen society” in the shorter term. This development is initially driven by the environmental improvement gained from hydrogen as a major energy carrier. As a carrier it may find new primary sources as needs dictate, and opportunities from ongoing research allow.

3.4 STORAGE AND DISTRIBUTION OF HYDROGEN

One of the main technological hurdles may be seen from the low volumetric energy density of hydrogen gas - it is only 3 kWh per Nm³ as discussed in section 3.1. While hydrogen, like natural gas, may be efficiently distributed in pipelines, this is scant help in the transport sector where the fuel must be carried aboard vehicles. Although the energy density is high on a weight basis, hydrogen at atmospheric pressure requires too much space to be conveniently carried about. As with generation of hydrogen, only a brief overview can be given here. We focus on the requirements for the transport sector as these are the most stringent, but there is considerable opportunity for synergism in technology between this sector and that of stationary usage of hydrogen.

A set of specifications for the “FreedomCAR” project in the US are as follows:

Targeted Factor	2005	2010	2015
Specific energy (MJ/kg)	5.4	7.2	10.8
Hydrogen (wt %)	4.5	6.0	9.0
Energy density (MJ/L)	4.3	5.4	9.72
System cost (\$/kg/system)	9	6	3
Operating temperature (°C)	-20/50	-20/50	-20/50
Cycle life-time (absorption/desorption cycles)	500	1000	1500
Flow rate (g/s)	3	4	5
Delivery pressure (bar)	2.5	2.5	2.5
Transient response (s)	0.5	0.5	0.5
Refuelling rate (kg/H ₂ /min)	0.5	1.5	2.0

Table 3.4.1 FreedomCAR Hydrogen System Targets

(Argonne, 2003:32)

Depending on type of car, a storage capacity of about 5-10 kg of H₂ is required. The problem of storage for transport may be approached in several ways:

- *Compression and storage in containers*
- *Liquefaction*
- *Absorption as metal hydrides*
- *Adsorption in high surface-area adsorbents*
- *“Hiding” in reversible chemical reactions*

Compression

This is the established practice. Storage at 300-500 bar is commonly specified for steel containers. Oil-free compressors are required for hydrogen to be used in low-temperature fuel cells. Storage of compressed hydrogen is feasible for heavy vehicles like buses, but is considered too heavy and cumbersome for private cars. Buchner (1995:180) gives as the equivalent to a 60 l gasoline tank, pressurised steel containers weighing 450 kg. Stronger and lighter containers are under development, but even these fall short of the operational requirements cited in Table 3.4.

The energy required for compression is given by the formula:

$$E \approx k \cdot R \cdot T \cdot \ln(P/P_0) \quad (3.4.1)$$

Where:

- k = a factor relating to compressor efficiency
- R = the universal gas constant
- T = Temperature
- P = final pressure
- P₀ = starting pressure

Compression from one to 300 bars requires about 4.3 kWh/kg H₂. When produced by electrolysis at high pressure (some 30 bar), the requirement for compression energy is reduced to a third, and the equipment is simplified as well (Glöckner, 2004).

Liquefaction

Buchner (1995) specifies a 100 kg insulated tank as equivalent to a 60 l gasoline tank. Liquefaction is well known for rocket fuel, but is difficult, expensive, and extremely demanding in terms of energy (rule of thumb: approximately 30% of the energy is lost). A useful role may exist in distribution similar to liquefied natural gas (LNG), and with larger vehicles, but liquid hydrogen hardly seems a technology amenable to widespread use in private vehicles. An interesting new development is the combination of liquefaction and pressurisation (Glöckner, 2004).

Metal hydride

Hydrogen interacts with metals in two different ways to form hydrides: interstitial compounds where protons slip into the existing atomic structure, and chemically bonded to form chemical compounds. Interstitial compounds have a storage capacity of 1.5 – 2.5 %wt, and hydrogen is absorbed and released by changing the

temperature in the range 10-80 °C, which is conveniently done by means of water as heat exchange medium. Hydrogen delivered via interstitial hydrides is very pure. The other side of the coin is expensive auxiliary systems and low thermal efficiency (30-35 %, Glöckner, 2004).

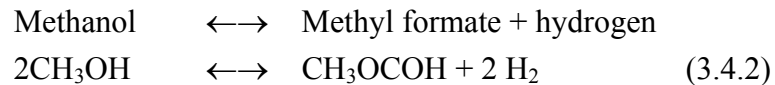
The other type of hydride is exemplified by compounds such as sodium- or lithiumborohydrides, and compounds with aluminium referred to as “alanates”. These have total hydrogen contents of 7-18.5 %wt (Argonne, 2003:40) and good recycling properties either directly or through a reaction with water (see e.g. Argonne, 2003:34 ff.). There is intense research activity in this area at present.

Nanostructured materials

This pertains to high-surface area substances such as activated carbon. Particular interest is at present focused on “nano-tubes”. A general introduction is given by Argonne (2003:42 ff.). Nanoscale hydrogen storage may be divided into two general classes. The first is “atomic” storage in which hydrogen gas molecules are split into individual atoms and react with the matrix. The second class is “non-dissociative” materials in which hydrogen molecules remain intact, but are held by van der Waals forces to lower their intrinsic vapour pressure. Adsorption in this way appears to be limited to about 2 %wt, but with good adsorption/desorption characteristics. The main function of absorbents is to lower the compression pressure for a given amount of gas. If successful, this would merge nicely with generation of hydrogen under pressure by electrolysis, as described in section 3.3. However, no distinct breakthroughs are reported thus far.

Hiding in chemical reactions

There is a wide variety in approaches to storage of hydrogen in chemical compounds. As one example, Chung et al (1998) presents the reversible chemical reaction:



If operated properly this reaction liberates one mole of hydrogen per mole of methanol, or $100 \cdot \frac{2}{32} = 6.5$ % by weight of hydrogen, which is much more than achieved by interstitial hydride systems. A system like this will depend critically on suitable catalysts and on good knowledge of thermodynamic parameters. This particular system has not become a preferred solution since being presented in 1997, but it (or similar ideas) may well succeed in the future. If successful they would provide a major agent for distribution and storage of hydrogen for transport purposes.

Argonne (2003:46 ff.) provides an overview of recommended areas for research, as well as extensive further references to the important and exiting field of hydrogen storage. An area of considerable interest is the use of hydrogen as a buffer between intermittent generation (e.g. by wind turbines) and varying consumption of electric power. However, depending on regional opportunities alternatives may compete with hydrogen, for example flow batteries (cf. Briskeby, 2002 and Korpaas, 2004).

3.5 FUEL CELLS

Fuels cells are essentially electrolysis run backwards. They take as input hydrogen and oxygen, and produce electric energy and water. Hoffmann (2001:145 ff.) traces the original invention to William Grove in England in 1839. He also describes work by Francis Bacon at Cambridge, England before WWII, and on propulsion of vehicles in the early 1960s in the US. Major work towards practical development took place in the various programs undertaken by NASA, and in particular as part of the Man-on-the-moon- and the Space Shuttle programs. However, in the early 1970s an impasse was

reached, due to a combination of technical difficulties and reduction in funding. Hoffmann (2001) gives the following resumé:

- *Hydrogen was the only really useful non-exotic fuel, but using it with relatively inexpensive nickel catalysts in an alkaline fuel cell required high temperatures and pressures, costly pressure vessels, and ancillary equipment.*
- *Alkaline fuel cells require very pure hydrogen. That was problematic when hydrogen was produced from common fuels such as natural gas or coal. Any residual CO₂ in the hydrogen reacts with the liquid alkaline electrolyte, gumming up the electrodes' microscopic pores and slowing the overall chemical reactions.*
- *The use of "dirty" commercial fuels plus CO₂ – containing air - as opposed to pure hydrogen and pure oxygen used on spacecraft – made the useful life of fuel cell systems (using construction materials commercially available at the time) too short for economical operation.*
- *In hindsight, it became clear that the close-knit community of fuel cell designers, engineers, and scientists, had "tended to oversell the merits of the fuel cell before really having come to terms with all the teething troubles of an immature technology... ...As a result of over-enthusiasm, deadlines were not met, and private funding for the fuel cell greatly declined. "*

P. Hoffmann, 2001:149

I believe this historical perspective is of interest in evaluating future prospects. Interest in fuel cells picked up following the 1973-1974 oil embargo, now with interest in both stationary and mobile applications. This interest almost exploded in the 1990s as the problems with global heating came into focus, associated with the use of carbon-based fuels.

There are five main types of fuel cells, with the following central characteristics:

Type	Operating Temperature	Electrolyte	Fuel	Application & notes	Efficiency
Alkaline fuel cell (AFC)	50-90 °C	Aqueous KOH	Pure hydrogen	Aerospace	50-60%
Polymer electrolyte membrane fuel cell (PEFC)	60-90 °C	Acidic polymer ion exchange membrane	Reformate hydrogen	Vehicles, aerospace, combined heat and power (CHP)	40-60%
Phosphoric acid fuel cell (PAFC)	200 °C	Phosphoric acid	Reformate hydrogen	Stationary power plant, CHP	35-40%
Molten carbonate fuel cell (MCFC)	650 °C	Alkali carbonates	Methane, reformate hydrogen	Stationary power plant CHP	45-60%
Solid oxide fuel cell (SOFC)	850-1000 °C	Yttria Stabilised zircona	Methane, coal gas reformate	Stationary power plant CHP	50-60%

Table 3.5.1 Major types of fuel cells

J. Garche, 1998:566

A fine overview of the current research front is given by Argonne (2003:54 ff.). Alkaline fuel cells were first developed by NASA. They have high power density, but are sensitive to CO₂ and to CO.

Polymer electrolyte fuel cells are mostly of the PEMFC type (Proton Exchange Membrane Fuel Cells) and are the ones most in focus for use in vehicles. They have high current density and allow compact construction, but rely on platinum as catalyst and are sensitive to CO and sulphur compounds. The platinum is a major cost factor.

Development of new polymers promises operation at temperatures approaching 200 °C, which both reduces the reliance on catalyst and removes the requirement for a separate system for handling of water.

PAFC tolerates CO₂, but require a platinum catalyst and are sensitive to CO. They have medium power density.

MCFC and SOFC are high temperature fuel cells that appear interesting for stationary applications. In particular SOFC is currently under intense study. SOFC will use both natural gas and CO as fuels, in addition to hydrogen.

Fuel cells offer two main promises: high efficiency and silent, non-polluting operation. The theoretical efficiency limit of fuel cells and ICEs is the same, but is obtained at different operating temperatures. The result is that the practical efficiency of fuel cells is higher (Halseid, 2004:15-42). The fuel cell limit efficiency is determined by:

$$\eta_{FC} = \Delta G / (\Delta G + T \bullet \Delta S) \quad (3.5.1)$$

which, as discussed in section 3.3, and as we see from Table 3.6, leads to practical efficiencies approaching 60 % within feasible operating temperatures. In contrast, alternative heat based processes are limited by the Carnot efficiency:

$$\eta_H = (T_1 - T_2) / T_1 \quad (3.5.2)$$

which in large scale conventional generation of electric power now has an average operational efficiency of about 38 % (IIASA CO2DB, 2003). Combined turbine/steam technology burning natural gas achieves about 55 %, as an example of modern, optimised heat technology. A modern trend is to combine power generation with low temperature heating of buildings. Clearly a combination of a SOFC

type fuel cell with Carnot cycle technology offers a very interesting prospect for efficiency.

At present two main areas stand out for growth of fuel cell technology and application. The first is small (1-20 kW) PEM cells used in transport and in combined heat/power (CHP) for homes. The second promising area is large-scale use of MCFCs and SOFCs for production of electric power from natural gas.

The main obstacles now are cost and operational lifetime. The following excerpt from Argonne (2003) illustrates the situation on costs:

“The use of fuel cells in the transportation sector could potentially provide substantial benefits and is perhaps the largest potential market for fuel cells. However, present technology cannot come close to the cost targets that will allow substitution of a fuel cell-based “engine” for an internal combustion engine. All current designs of fuel cells for this application are low-temperature (about 80 °C) technologies based on polymeric membranes as proton-conducting electrolytes. The need for low-temperature fuel cells in this sector is driven by system considerations – in particular, overall weight and volume, fuel conversion efficiency, fast start-up times, and long out-of-use periods. Current estimates of polymer proton exchange membrane (PEM) fuel cell costs, when extrapolated to mass production, exceed \$100/peak kW. To compete economically with the internal combustion engine, however, the cost must approach \$35/peak kW. Present-day costs for low-volume production are roughly \$3000/kW for hydrogen-based fuel cell systems. The need to reduce the costs by nearly two orders of magnitude underscores the long pathway to the goal of automotive application of fuel cells

(potential stationary applications can be viable at much higher costs in terms of dollars per kilowatt-hour).

The primary difficulties in achieving the automotive cost target lie in the materials used in the fuel cell stack. Polymer electrolyte membranes, precious metal catalysts (typically platinum or platinum alloys), gas diffusion layers, and bipolar plates account for 70% of the cost of the system. Any two of these are projected to cost more than the target for the full system unless researchers can achieve significant advances in power density and materials optimisation and durability.

Argonne, 2003:56

There is an interesting flaw in the above reasoning: if the ICE was just a cost item of \$35/peak kW then there is no reason to bother with fuel cells. However, the ICE has an extensive environmental impact and may imply insecure supplies of energy in the future – this is why we trouble with fuel cells in the first place. We should keep in mind that a successful solution to the cost problem of fuel cells carries a considerable amount of increased value to society.

We may start on the path to efficient and affordable fuel cells by seeking out areas in which they convey particular advantages. An example of a promising place to begin, might be the use of fuel cells for standby back-up power. This area, known as ‘primary power’ is characterised by few operational hours, a requirement for quick response, and high priced alternatives. Good points of entry for hydrogen related technology are further discussed in section 8.3.

3.6 THE PROCESS OF TRANSITION TO A HYDROGEN ECONOMY

The foregoing sections have briefly reviewed the main technological parameters of a possible hydrogen future. There is promise of drastically reducing environmental impact by increased use of

hydrogen. Even better, the hydrogen may be produced from renewables. There are, however, hurdles beyond technology that must be surmounted if a shift from fossil fuels to hydrogen is to take place. Victor et al (2003) compare the introduction of a hydrogen infrastructure, to former major changes in technology:

“The rate of change in canals was relatively rapid – a few decades – because they required little change in the vehicles. Floating barges were a simple derivative of floating ships already in use, and horses (also abundantly in use) were often deployed with ropes to tow the barge down the canal. Railroads took much longer as the infrastructure required not only the physical creation of the rails themselves but also complementary changes in steam locomotives and regulation, which followed massive speculation and a crash in railroad stocks reminiscent of our most recent crash in .com investments and in broadband...

...Telegraphs, by contrast, diffused rapidly within networks (the rights of way along railroads) that already existed...

...Would the shift to hydrogen infrastructures follow the pattern established by telegraphs – rapid diffusion within an existing rights-of-way infrastructure? Or will a hydrogen infrastructure follow the slower diffusion illustrated by most other transportation infrastructures, including railroads, automobile roads and gas pipelines? All told, the historical experience with infrastructures suggests that it is unlikely that hydrogen will diffuse as a dominant energy carrier more rapidly than about four decades – about the same time scale as railroads or the natural gas system.”

D.G. Victor, T. C. Heller and N. M. Victor, 2003:7-8

This agrees well with what is seen in Figure 1.1. Extensive time is required for change in basic energy structures. Under the heading “The Empire strikes back” the authors continue:

“The factors that determine the diffusion of a new technology lie not only in the intrinsic properties of the technology itself – such as the needs for complementary innovations in the infrastructure and the end uses, which would tend to slow down the rate of diffusion – but also at the incumbents...”

Ibid: 2003:8

This is a major reason that the study of innovation and transition becomes so multi- and interdisciplinary.

...Here we focus on two of hydrogen’s likely competitors: electricity (for stationary uses) and petroleum (for transportation). Measured by sheer volume of business, hydrogen is a mouse in this game...

...The incumbents can respond in at least two ways. Politically, they could respond by organizing themselves... even if society may benefit from a shift to hydrogen, the individual beneficiaries are unknown and few. In contrast, the possible losers are already known and already well organized through industrial organizations. Indeed, some of the “losers” may transform themselves into winners, and thus large energy companies are today also dabbling in hydrogen. But that dabbling makes sense not only as positioning for future commercial benefits but also to gain inside knowledge needed to mount a defence”

ibid. 2003:9

A somewhat different position is adopted by the European Union: in their Summary Report, June 2003, the EU High Level Group for Hydrogen and Fuel Cells have the following view regarding a European roadmap for hydrogen and fuel cells:

“Moving Europe away from its 20th century dependency on fossil fuels to an era powered by the complementary carriers, electricity and hydrogen, will require careful strategic planning. Hydrogen is not likely to be the only fuel for transport in the future. Moreover, maintaining economic prosperity during the transition period will involve maximising efficient use of various forms of fossil-based fuels such as natural gas, methanol, coal and synthetic liquid fuels derived from gas”.

EU High Level Group for Hydrogen and Fuel Cells, 2003

To understand change and dynamics towards a hydrogen infrastructure, it is simply not sufficient to study hydrogen in its own. The challenge becomes massive and complex, not by our choosing, but because reality is that complex. Attempts at simplifications run the risk of missing out relationships that may prove important in the future. Yet, the art of modelling is this process of simplification. The key is to aim for overall understanding.

In the next chapter we consider change. Apart from the technological hurdles, the main factors to enforce change are intensity of driving forces, political will, and ability to turn the need for change into practice.

Understanding change

"The eye sees only what the mind is prepared to comprehend."

Henri Bergson

“Change” is the fact that today differs from yesterday. Some change is to advantage and some is not. Some results from circumstances beyond our control, and some is of our own doing. Resistance to change is widespread, since almost all change is perceived by some to be to their disadvantage in one way or another. Many feel insecure in a changing and developing world. They cling to the perceived safety of an unchanging world even when change brings definite improvement.

The realisation of a hydrogen economy implies change. The purpose of this chapter is to better understand the concept of *change* in its own. A first focus is to understand change based in irreversible thermodynamics. We subsequently consider important factors governing change; in particular path dependence and feedback. Finally, we look at factors that form barriers to intended change, and factors that form driving forces for change. This is the fundamental idea in order to capture reality in the EICOMP model.

4.1 THE PHYSICAL NATURE OF CHANGE

A physical understanding of change seems the best way towards a more systematic treatment. In this way we identify deeper structures

that concern all forms of change. We need this deeper understanding to guide our choice of models as described in chapters 6 and 7. We discuss this in terms of linear- and nonlinear change, and then present a picture of change grounded in irreversible thermodynamics as first developed by Ilya Prigogine (1997).

The most readily appreciated situations are those with seeming stability. Such situations usually indicate stable equilibria where systems return to a starting point if disturbed. In physical systems this corresponds to states with minimum free energy (Prigogine, 1997:63 ff.). In human terms situations of this kind appear foreseeable and “safe”. They are, however, also without scope for growth and development.

A next step is situations where change is linear and incremental. A familiar case is the steady draining of the fuel tank as a car is running. But already this simple example reveals a characteristic of change: if the tank runs empty, the car stops and a whole new situation arise. We shall see that an underlying result is development of instability, and that this is a general feature whenever something moves systems away from their equilibrium (Prigogine 1997:66).

Another challenge is introduced with “exponential” change. Such change *doubles* a quantity in a fixed period of time, while people often tend to assume a quantity that increases by the same absolute amount per time interval. Exponential change is very common. It describes e.g. the growth of bank accounts/debts with time, or the growth of populations. The time for doubling is easily estimated: it is seventy divided by the percentage increase (e.g. an account with 7% interest doubles in $(70/7 = 10)$ years). The exponential relationship is discussed in more detail in section 6.2.1.

Like linear change, exponential change is smooth and incremental, and is perfectly foreseeable. However, over short periods of time it is

commonly mistaken for linear, or no significant change with time. The problem we encounter is that people are poor at estimating such relationships, and tend to grossly underestimate how fast output increases once it starts to become noticeable (Sterman, 2000:269). A well known example is the thickness of folded paper: an ordinary sheet of paper (0.1 mm thickness) is folded ten times. How thick is the final result? Common guesses are a few millimetres. In fact, the folding produces a wad that is $0.1 \cdot 2^{10}$ mm ($= 0.1 \cdot 1024$) ≈ 10 cm thick. Another example, given by Meadows et al (1972:29), is a lake in which water lilies grow and divide daily. One day the lake is fully covered with water lilies. The question is: how long was it since it was half full? The answer is *one* day (With a total of thirty days to cover 100%, 23 days are required to cover the first 1% of the lake).

We now turn to the understanding of change in irreversible, open thermodynamic systems, and follow Prigogine (1997). The fundamental laws of physics, i.e. Newton's laws of motion, relativity and quantum mechanics, all lack a direction of time. We may exchange "t" with "-t" with no other effect than reversing the direction of motion, which just relates to a frame of reference. Energy and all other fundamental properties remain unchanged. The only part of physics that gives a direction of time is the production of entropy in thermodynamic systems. Entropy is related to disorder. The second law of energy as formulated by Clausius (cf. Cropper, 2001:100 ff.) states that *entropy* increases with *irreversible* processes in closed systems. Irreversibility is a necessary requirement for all change, and also for all observation. In open systems there is a possibility of exporting entropy to the outside, resulting in decreased entropy and increased order (less randomness) where entropy is exported from. Such transport requires an input of work; i.e. the export must be coupled to a flow of energy. Physical situations (systems) in which entropy is exported, arise when thermodynamic systems are kept away from equilibrium by a critical distance, and a reinforcing condition exists:

“...with catalytic steps such as the production of an intermediate compound “Y” from a compound “X”, together with the production of X from Y.”

I. Prigogine, 1997:66 ff.

Prigogine refers to such situations (or systems) as *dissipative systems*. “Catalytic steps” is a chemical concept. However, it generalises to any kind of loop with positive feedback (cf. section 4.2). A dissipative system is in principle a positive feedback (reinforcing) loop that upholds a flow of energy to maintain the system far from equilibrium. Living organisms, firms and societies are all examples of dissipative systems.

Instability develops if either the supply of energy or the autocatalytic, reinforcing feedback mechanism is interfered with. The result of instability is *bifurcations*, as illustrated in Figure 4.1.1:

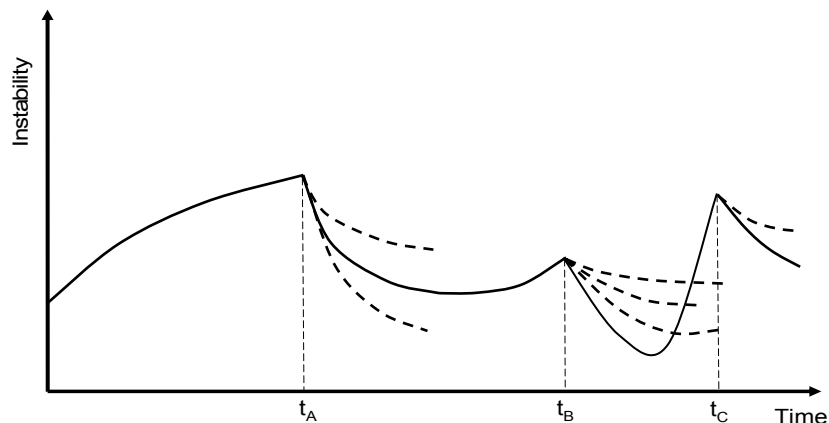


Figure 4.1.1 *Successive bifurcations (A,B,C) as instability develops in open thermodynamic systems with time*

The differences ($t_B - t_A$ etc.) indicate time required to develop instability. In bifurcations a system switches abruptly to some more stable configuration, whence new instability may start to develop. The outcome depends *both* on the system itself and on its surroundings. Equally important: the system may have multiple ways in which it may react, and it “chooses” one of these (cf. Arthur, 2000:3 ff.). Which one is chosen may be foretold only in terms of *probabilities*. With this type of change, nature is simply indeterminate to varying degree. Figure 4.1.1 illustrates a familiar behaviour, although we may not realise the underlying situation of developing instability. Change is no longer continuous and incremental. Our earlier example of the emptying fuel tank illustrates how instability develops. As long as there is fuel left in the tank, the reinforcing loop exists, but the result is sudden stop when the tank runs empty. The sudden stop represents a qualitative change in system properties, while the continuous emptying of the tank represents quantitative change. Such a bifurcation constitutes a qualitative change. Similar situations exist in economic systems, extensively discussed by Arthur (2000) and by Romer (1990). The EICOMP model is adapted to reflect reality in this sense.

A final point should be made: A pencil standing upright on its flat end represents an unstable equilibrium, and will topple over if only slightly pushed. How it will end up is not known exactly in advance, but is described again as an *intrinsic probability*. However, the probability increases in the direction in which the pencil was pushed, showing that we may influence the outcome of bifurcations.

To sum up: dissipative systems (e.g. energy systems) are subject to abrupt shifts in the form of bifurcations with developing instability. The instability results from interference with either the feedback mechanism or the flow of energy. The outcome of bifurcations depends both on a system itself, and on its surroundings (i.e. pushing

the pencil in a given direction). The outcome may be foretold in terms of (intrinsic) probabilities and not in exact terms.

4.2 FEEDBACK, PATH DEPENDENCE AND LOCK-IN

In the previous section we encountered the situation of *feedback*. This is simple as a concept: it denotes the basic idea that information about a process acts back to influence the same process. This creates a loop that is known as a *feedback loop*, and the information in question is referred to as “feedback”. It is a basic concept in control theory. A familiar example is a car engine whose output is controlled by the throttle. It is the throttle that decides the actual output, rather than the properties of the engine itself.

Feedback is an explicit element and tool used to control man-made processes, and forms central elements in dynamic models. Feedback situations are, however, intimate parts of reality, not least in social and biological systems. Therefore, as we incorporate feedback loops in system dynamic models (cf. section 6.7), we genuinely simulate important functional properties of real life: identification of feedback loops (both reinforcing and balancing) constitutes an important part of our understanding of the problem we seek to study.

We now turn to the notion of *path dependence*. Path dependence arises in systems dominated by positive feedback, or systems with irreversible processes. A typical result from path dependence is right-hand driving in some countries and left-hand driving in others. Both function just as well. An example of path dependence in markets is the well known case of VHS / Betamax as standard for videocassette recorders (Sterman, 2000:359). Even though Betamax was seen as a superior technology, VHS gained a slightly larger market share at an early stage and subsequently came to dominate the entire market. In energy systems the AC standard for electric power is another

example. Path dependence is ubiquitous and not limited to historic selections. It is continually created in real life.

Equally important to the adoption of some choice, is the fact that it *prevents* other alternatives from being chosen. This is clearly seen in chess, where a move will lead to a given continuation, and will prevent other courses from being taken. The move in question would often be one of several alternative moves, that might be just as preferable. The position on the board resulted from particular earlier moves and sequences of moves, made in the same situations of selection from several alternative possibilities.

The overall effect of path dependence is to introduce history as a parameter - reasonably enough - but it becomes much harder to foresee the outcome of change, particularly in longer perspectives of time. Even more crucial, path dependence determines remaining available options. Path dependence is described in a wider context of evolutionary models of economic development by Nelson (1995).

Arthur (1989), and Sterman (2000:349-406) provides an extensive introduction to the understanding of path dependence in the real world, and how it may be described and studied within system dynamic models. A characteristic feature of this type of models (cf. chapter 6.7) is the feasibility of representing reality as interacting sets of reinforcing and balancing feedback loops. Sterman (2000:365-380) lists a number of such reinforcing loops commonly seen in practice:

- *Product awareness increasing with sales, creating more sales.*
- *Spreading development costs over a larger volume reduces costs, to increase volume further.*
- *Economies of scale lowering production costs with increasing volume, to increase sales and further increase volume.*

Chapter 4

- *Network effects: the usefulness to customers increases with increasing volume, to increase sales and volume. Classic case: telephones.*
- *Complementary goods: increasing sales of a given computer system make it a better platform for software development, in turn increasing the competitiveness of the given computer system.*
- *Product differentiation: increased competitiveness of products allows higher prices while remaining competitive, making further funds available for differentiation.*
- *Market power: increasing clout towards supplier and channel firms to obtain selective advantages with respect to competitors, freeing resources for advertising etc..*
- *Mergers and acquisitions: consolidation of market dominance to increase dominance.*
- *Workforce quality and loyalty: the more superior the firm, the higher the wages and the better the career opportunities, and in turn the better people and workforce quality.*
- *Cost of capital: the more growth and earnings, the higher the market value of a firm and the easier to raise capital for further growth.*
- *Ambition and aspirations: achieving goals leads to increasing aspirations.*

Reinforcing loops have the general property of “locking in” to create lasting “entities”. Even when all paths are initially equally attractive, the symmetry is broken by microscopic noise and external perturbations. The positive feedbacks then amplify these small initial differences to macroscopic significance. Once a dominant design or standard has emerged, the costs of switching become prohibitive: the system has locked in.

With a good grasp of such models one understands their basis and thus what keeps these situations alive over extended periods of time.

This corresponds closely with the more general picture resulting from dissipative systems as described in section 4.1. It allows a better understanding of ways to effect change, as this depends on development of instability that opens for bifurcations. This is discussed in more detail in section 4.4.

There is a problem of information transfer to the reader with these highly complex issues. A better feeling for what is going on is obtained by actually attempting to construct and run models with positive and negative feedback. This is further discussed in section 6.9.

4.3 BARRIERS TO CHANGE

The concept of barriers is meaningful in relation to wanted change; i.e. to change that has some accepted reason for being set in motion, different from resisting change that is imposed for unaccepted reasons. In this section we look at barriers (or reasons for resistance) that fall in four categories:

- *Human costs*
- *Defence of existing investments*
- *Lack of capital and knowledge*
- *Lock-ins and catch-22 situations*

4.3.1 HUMAN COSTS

Human and social factors often lead to resistance against technological change. Technological change is linked with economic change, which in turn is closely associated with social change. The human reaction to change ranges from personal whims and attitudes to highly substantial reasons.

Such substantial reasons, which we refer to here as “human costs” range from insecurity and loss of influence in a job situation, to loss

of work, loss of income, and perhaps force families to move in order to find new work, and lead to social uprooting. These are powerful reasons for resisting change, and the threat of getting into such situations constitutes powerful barriers to change. We are talking about *threats*; these may be perceived as more or less real.

4.3.2 DEFENCE OF EXISTING INVESTMENTS

Existing investments relate to situations where investments have been made and cannot be unmade. However, investments are made in order to reap benefit. This may be undermined by changing conditions. A familiar situation is the investment in a product, where introduction of a new product by a competitor may make the first one obsolete and unable to compete. The capital invested in the first product is then lost. This tends to slow down and work against change. Sunk costs may be capital invested in physical equipment by firms, but may as well be in time and effort invested in learning new skills. If people, due to change, have to shift into new jobs, the value of original investments is reduced.

“...An existing technology often has significant ‘sunk costs’ from earlier investments, meaning that firms will be reluctant to invest in more sustainable alternatives”

T. J. Foxon, 2002:2

Stakeholders tend to defend their interests (i.e. existing investments). Introduction of something new becomes easier when what is to be replaced, departs in ways acceptable to both investors and operators/workers. The rate of change also matters. The more abrupt the changes in technology, the more reduction in lifetime of investments, the greater the lost capital.

4.3.3 LACK OF CAPITAL

The requirement for risk capital for change is well established, and is not discussed in further detail here. However, three points should be made:

1. *Projects compete for available capital*
2. *Projects are evaluated according to net present-value calculations*
3. *The cost of capital depends on associated risk and time*

The first point is important with respect to initiatives by society to effect desired change: if projects towards sustainable energy receive no particular attention, this means that such projects are treated like anything else in day-to-day markets. This is accentuated by the second point. Net present value is calculated in principle as the amount one would have to deposit in a bank account today, for the amount (with accumulated interest) to match a given value some specified time from now. A concise introduction is provided by Hanssen et al (1996:69 ff.). The higher the interest rate, the shorter time required to grow to the given value. Hence, the more short-term projects become preferred. Boyle (1996) gives a clear example of the effect of ordinary net present value calculations:

“The Puueo plant (in Hawaii), first operated in 1918, is still generating electricity. ...Long ago, the plant paid off it’s initial capital costs and completed its economic life. Therefore, according to a cost-benefit economic study, the electricity which the plant is producing has no value. But the fact is that the plant continues to produce ... the same power that it did 60 years ago and there should be some way to place a value on it... [Miyabara, 1981]

...There is much truth in the observation that all power producers wish they had invested in hydroelectricity 20 years

ago, but unfortunately can't afford to do so now – and that they held this same view 20 years ago!”

G. Boyle, 1996:223

Ordinary return-on-investment calculations do not distinguish between capital- and operating costs, and thus are blind to long term benefits. The interest rate may look quite different from a society- and a firm-based point of view.

As to the third point: the more transparent and “safe” a project is, the more easily it is financed (cf. section 5.5).

4.3.4 LACK OF KNOWLEDGE

The impact of knowledge on the ability to adapt to change, is a much discussed theme in the field of economic development. In the 1950ies Solow (1988) detected a significant discrepancy between productivity as predicted by work and capital in neoclassic economic theory, and productivity as actually observed in real life. He attributed this to increasing quality of the work-factor. Romer and others (cf. e.g. Mork, 1996) developed this into what is called the “endogenous” model/theory of economic growth. In this model, capital is seen as consisting of two parts: the traditional physical equipment, and *knowledge* as an entity that behaves differently (e.g. it does not wear with use, it increases with education and with learning curves, and it is only in part generated through own efforts. The EICOMP model reflects this distinction between knowledge and physical capital through attractiveness (cf. section 7.2). Knowledge creates a number of advantages: improved coordination, quicker ability for institutions to respond, improved quality etc. Modelling is thus an important input to competitiveness.

4.3.5 LOCK-IN AND CATCH 22 SITUATIONS

The concept of lock-in was described in section 4.2. Lock-ins create barriers to change somewhat like being at a local optimum, where

other and better optima exist, but where some “valley” must be crossed in order to reach the more advantageous maxima. As long as one is unable to cross these “valleys”, which represent the cost of switching, no change will take place.

Catch 22 depicts situations with blocking feedback. It is illustrated through the problem of developing a hydrogen infrastructure before there is a market with customers, and there are no customers because there is no infrastructure. In practice a project could pay its way once realised, but cannot be realised for lack of capital to pay for initial costs. This is a likely evaluation when interest rate is set too high in net present value calculations (cf. example by Boyle in section 4.3.3). It is essentially an initialisation problem that may be solved through coordination. One example of such coordination (by society) is to arrange for sufficiently low interest rates in net present value calculations for investments in sustainable energy. The catch 22 phenomenon is a powerful block to change.

4.4 HOW TO ACHIEVE CHANGE

The areas of endeavour most directly linked with achieving desired change are those of innovation and strategy. Both are large and complex fields that may only be briefly discussed here. Extensive introductions are available such as Grønhaug and Kaufmann (1988) on innovation, and Mintzberg et al (1998) on strategy. Innovation is generally understood as the adoption in volume of products and processes that are somehow *new* (cf. section 5.1.3). *Strategy* involves ways of identifying and reaching goals as optimally as possible. To achieve change means choosing between some set of alternative approaches. It is not sufficient to only identify technological potential. This guides the direction, but practical change must be carried out in the real world. Innovation may be used to achieve the improvements we seek (cf. further discussion in section 8.3).

4.4.1 DESCRIBING CHANGE BY MEANS OF ACTIVITIES

Change may be understood in terms of enhanced supply of goods and services, quantitatively and qualitatively. It may also be the upholding supply at current standards with less down-sides (e.g. less pollution), as well as changes in the underlying support structure, in order to uphold supply in a stable, sustainable manner.

Value sequences (cf. section 6.5) provide a systematic means of describing the creation of goods and services in terms of interacting activities. These activities are hierarchically subdivided to allow further description. All change may then be described as improvement of individual activities, or adoption of new activities to meet the same objective. However, any historical picture implies a snapshot of continuous development. This causes problems when we seek to describe change outside the individual firm: the “same” activity is carried out differently in individual firms and the structure of activities is perceived differently, even when the firms are producing essentially the same output.

4.4.2 MOTIVATION FOR CHANGE

Three themes seem to lie behind most motivation for change:

1. *The striving for fortunes*
2. *Removal of bottlenecks*
3. *Some change to avoid more dramatic change*

Striving for fortunes

Where there are prospects of lasting income, almost nothing appears to stop the capitalistic system from exploitation. This goes for farming of new land, for oil exploitation wherever oil may be found, for gold rushes and for international commerce. The general impression is that if conditions can be created to ensure lasting income from conversion to sustainable supplies of energy, then such conversion will come about almost by itself.

Removal of bottlenecks

If a particular resource has some kind of disadvantage, there will tend to be continuous effort to overcome this. This may in turn lead to proficiency that allows the utilisation of the resource in question to prosper. This overall phenomenon has come to be called the “bottleneck” mechanism for change (Sejersted, 1982:11 ff.). This is systematically applied in the EICOMP model.

Some change to avoid more dramatic change

This situation is simple in principle: as it becomes understood that “the world is heading for trouble” preventive action is taken. The main problem is that situations are judged differently by different people. Some ignore the warnings, there are different perceptions of how acute a problem is, and there is of course hesitancy as to what course of action will be most suitable. The latter is frequently connected to ownership. All these factors hold initiatives back, but as situations and problems become clearer, change takes place (cf. section 8.3). The overall situation of sustainability is clearly of this kind, with a clear warning bell sounded by Meadows et al (1972), more than thirty years ago.

The perception of change and its various aspects as presented in this chapter permeates this thesis. It is the fundamental understanding of change related to dissipative systems and bifurcations that is the basis for the development of the EICOMP model. Bifurcations imply behaviour that can only be described in terms of intrinsic probability. This in turn is behind uncertainty as described in the next chapter.

Facing uncertainty and risk

“I would like to begin with two generally accepted propositions about technological change: it is a major ingredient of long-term economic growth, and it is characterized by a high degree of uncertainty.”

N. Rosenberg, 1996

To channel change towards some desired outcome, as introduction of a hydrogen economy is an example, we need to grasp the influence of uncertainty and risk. The previous chapter discussed the fundamental aspects of change. We now look into the concepts of uncertainty and risk measured towards some comprehension of *value*, as value defines the direction in which we want change to take place, and allows us to rank alternatives. The notion of value is strongly dependent on *ownership*. This defines the frameworks of stakeholders. In the present work we distinguish in particular between the public sphere and private ownership.

Uncertainty may be seen in two different dimensions. The first is with respect to whether or not we will reach our intended goals. We call this *quantitative* uncertainty (cf. section 5.2.1). This form of uncertainty may be expressed in terms of outcomes (events), each with estimates of likelihood and consequences, to provide us with a quantitative understanding of uncertainty. The second dimension pertains to situations in which the idea of value changes qualitatively, for example leading to quite new products or services. We will refer

to this as *qualitative* uncertainty (cf. section 5.2.2). Section 5.3 briefly discusses the handling of risk versus probability and consequences. Section 5.4 presents the concept of scenarios, and discusses briefly how they are applied in modelling to help facing uncertainty. Section 5.5 looks at how the EICOMP model in principle may be used to reduce risk.

5.1 CONCEPTS AND CONSEQUENCES OF VALUE

For the transition towards sustainability, we will focus on three main meanings of *value* related to provision of products and services:

- *Value related to price*
- *Value related to competition*
- *Value related to technological effectiveness and efficiency*

5.1.1 VALUE RELATED TO PRICE

At first hand this is what you pay for goods in markets, which seems straightforward. On a deeper level it relates to the meaning of value on an individual, personal basis. It is a complex subject, reaching from personal whims and bounded rationality, to deep philosophical and moral issues. The only generalisation made here is that this deeper level often considers principal, long-term perspectives.

Heilbroner (1988) provides an excellent review and overview of ‘value’ as it has developed in an economic, philosophical and historical context. The central issue is the relationship between value and price. Heilbroner defines a central *problematic* of value as:

“The general problematic of value, as I see it, is the effort to tie surface phenomena of economic life to some inner structure or order.”

R. Heilbroner, 1988:105

Heilbroner discusses five different areas where a common understanding of value could be based:

1. *A normative basis for 'just' prices*
2. *Value based on the required amount of work to achieve some benefit*
3. *Value based on cost of production (the problem of rents and profit)*
4. *Value based on utility of outcome*
5. *Value of labour in exchange for benefit*

Heilbroner (1988) finds all of these lacking in various ways, with the greatest headache being a problem of commensurability: 'just' prices depend on the point of view of what is 'just'; 'work' may be pleasurable or drudgery; 'cost' results from a combination of work, rents and sharing of profits. Value based on utility is the dominating idea of value in present-day economic thinking, but in Heilbroner's (1988:127-129) view this fails in three major ways:

- *through the operational difficulty in reaching or agreeing on a price to reflect equilibrium between supply and demand*
- *through a conceptual difficulty of such an equilibrium*
- *through looking at just the market and demand, and forgetting the supply side*

Heilbroner summarises:

“From a view that sees utilities as the fundamental and irreducible building blocks of price, gross national product is a meaningless concept, the ‘summation’ of individual experiences of pleasure and pain. This has no more validity than the summation of the enjoyments of an audience at a concert.”

ibid. 129

Heilbroner believes the best bet for his problematic to be the fifth concept of economic value above, based on *abstract labour* of work *towards exchangeable commodities*. This can be traced to Marxian theory of value. The problem of commensurability is solved by this approach (ibid. 120-121). Overall, however, Heilbroner (1988) concludes that there exist no agreed-upon, basic meaning of ‘value’ that may be used for comparison on any general basis, even when limited to the economic sphere. As we still require a concept of value, we have to relate it to the practical situations in which it is needed.

5.1.2 VALUE RELATED TO COMPETITION

If the problematic of value as discussed by Heilbroner was solved, we could have used it everywhere we need to judge the value of something with respect to something else. Value in markets achieves commensurability as it relates directly to products and services directed at comparable needs of customers. The central function of a value here is to rank competing alternatives. A side to the understanding of value in this respect is *buyer value*, as used by Porter (1990:42-44) in his work on competition.

With a need to look beyond strictly economic issues, the present work uses a compound construct of *cost*, *quality* and *availability* leading to a concept of *attractiveness*. The combined construct is used with the introduction of Porter’s model of competition (cf. section 7.2). By introducing attractiveness it becomes possible to *internalise* issues that are otherwise external to economic modelling, like the issues of environmental quality and sustainability. In the EICOMP model this gives us a means for commensurable comparison that takes place in markets.

5.1.3 VALUE RELATED TO EFFECTIVENESS AND EFFICIENCY

Beyond the need for ranking in the context of competitiveness, ‘value’ serves to define *direction*. It is obtained in this sense through

seeing value as consisting in the *meeting of needs*. In the EICOMP model this is done through the concept of *value sequences* (cf. section 6.5) and goes beyond utility. Briefly, value sequences are grounded in the concepts of end-products and end-users, where the latter have needs that the end-products seek to meet. A sequence describes the need that its end-product aims at in terms of a hierarchical structure of activities, required to bring forth the end-product in question. These activities are *value-oriented*, i.e. they carry descriptions of their associated, “local” needs. The net result is that the value of any particular activity becomes related to the competitiveness of its final end-product, i.e. is described in terms of technology. The description is *qualitative* in the form of descriptions of products/services and effectiveness (how well a product meets its targeted need), and *quantitative* in terms of efficiency (output v/s input).

This description in terms of technology is important to the present work. It forms a basis for evaluation of subsystems in the model in terms of competitiveness, as elements of the value sequences they are part of.

5.2 UNCERTAINTY AND TECHNOLOGICAL CHANGE

Technological change is complex. Barriers to change are discussed in section 4.3. Since the work of Jorgenson, Abramovitz and Solow in the 1950ies technological change is understood to be of utmost importance for economic development (cf. Introduction and Landau et al, 1996:2-18). A modern picture of the overall process from ideas and knowledge to large-scale creation of value is illustrated in Figure 5.2.1:

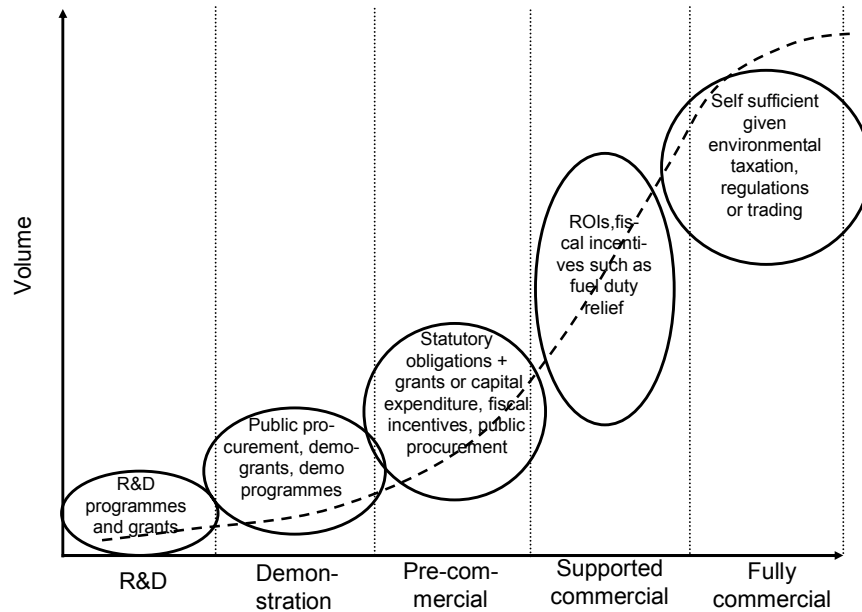


Figure 5.2.1 Modern understanding of significant innovation

Following ICCEPT/E4tech, 2003

Uncertainty in innovation relates to all sections outlined in Figure 5.2.1. Quantitative uncertainty relates to the success of an intended development, which Figure 5.2.1 implicitly implies. Qualitative uncertainty relates to the possibility that the innovation in question will give rise to development (and perhaps problems) in quite different directions (cf. section 5.2.3).

5.2.1 QUANTITATIVE UNCERTAINTY

In commercial life the quantitative uncertainty generally starts with an idea for a product (or service) that is perceived to have commercial value, i.e. to be sold in volume to form a basis for profit. The investment uncertainty in this situation is a combination of a number of sources:

1. *Whether the product achieves its technical specifications*
2. *Whether estimated costs prove realistic*
3. *Whether judgement of its market value holds*
4. *Whether the time to achieve estimated market penetration is reasonable*
5. *Whether the market volume will be large enough*
6. *Whether the product avoids substitutes and remains competitive for a reasonable period of time*
7. *Whether the product depends on simultaneous introduction of additional products/services by someone else (a 'chicken-and-egg' situation)*

Quantitative uncertainty is experienced by firms and is familiar in this setting. We will come back to point 7. Firms generally approach uncertainty by:

- *Not moving too far from existing experience, neither by way of technology nor markets*
- *Obtaining public support to reduce risk*
- *Sharing expenses (and profits) with others*

An important point should be raised: *It is normally not an option in the long run for any firm not to take risks.* Not to take any action generally increases the likelihood of being overrun by competitors.

Point 7 differs from the other points above by introducing dependence on positive interaction with others. This uncertainty raises a serious barrier to change (cf. section 4.3.4), and may in principle be reduced in three ways:

- *By means of vertical integration*
- *Through network cooperation and agreements by sets of firms*
- *Through assistance from public authorities*

The first two points appear to require sufficient prospects of economic gain and clear overview, to move from opportunity to reality. If the change is of public concern, such as the quest for sustainability, then it falls to the public authorities to create the required incentives and clarity. We now turn briefly to the public interest in reducing uncertainty.

5.2.2 QUANTITATIVE UNCERTAINTY IN THE PUBLIC SPHERE

The verdict from history is that most innovation attempts fail (cf. Rosenberg, 1996:334). When publicly motivated, change must take place either directly in the public sector, or through (new or existing) firms that begin to deliver new products and services. In principle it is irrelevant which firms succeed, unless national preferences matters. Public authorities may reduce uncertainty in several ways:

- *By creating dependable market conditions with economic incentives*
- *Through regulations that restrict existing practices*
- *Assist directly through engagements in the various phases of innovation (outlined in Figure 5.2.1)*
- *Through investments in general knowledge, including modelling to gain overview and direct assessment of uncertainty*
- *Through investment in public education*

5.2.3 QUALITATIVE UNCERTAINTY AND ROBUSTNESS

The qualitative uncertainty may lead to both negative and positive results. The first may be illustrated by the insecticide DDT, which is effective against mosquitoes and thus controls malaria and other disease vectored by mosquitoes. Following World War II, DDT saved millions of lives. However, as its potential for environmental damage became clear, this highly successful product had to be phased out. The commercial value of patent rights to DDT was high and

visible in the 1950s, while the concept of environmental incompatibility was literally unknown. This changed in 1972 as DDT was banned by the US Environmental Protection Agency.

The second type of qualitative uncertainty has to do with not understanding value, and thus missing opportunities. Rosenberg (1996) describes convincingly a wide range of such situations, from which we refer to the invention of the laser:

“But perhaps no single application of the laser has been more profound than its impact on telecommunications, where, together with fiber optics, it is revolutionizing transmission. The best trans-Atlantic telephone cable in 1966 could carry simultaneously only 138 conversations between Europe and North America. The first fiber optic cable, installed in 1988, could carry 40000. The fiber optic cables being installed in the early 1990s can carry nearly 1.5 million conversations [Wriston, 1992, pp. 43-44].

And yet it is reported that the patent lawyers at Bell Labs were initially unwilling even to apply for a patent on the laser, on the grounds that such an invention had no possible relevance to the telephone industry. In the words of Charles Townes [1968, p. 701], who subsequently won a Nobel Prize for his research on the laser, “Bell’s patent department at first refused to patent our amplifier or oscillator for optical frequencies because, it was explained, optical waves had never been of any importance to communications and hence the invention had little bearing on Bell System interests”.”

N. Rosenberg, 1996:336

[Townes, C. *Quantum Mechanics and Surprise in the Development of Technology*. Science, Feb. 16, 1968]

[Wriston, W. B. *The Twilight of Sovereignty*. Charles Scribner's Sons, New York]

The qualitative uncertainty may both create problems for intentional development, and reap advantages that were not planned. Provided that one allows for “impossible” breakthroughs, their consequences in the energy sector may in principle be modelled by EICOMP and similar approaches.

5.3 RISK VERSUS PROBABILITY AND CONSEQUENCES

Uncertainty leads to risk. Risk results from the combination of probability and consequences of events. A common approach to evaluate risk in engineering environments is to use an array as illustrated in Table 5.3.

	<i>Low consequence</i>	<i>Significant consequence</i>	<i>Serious consequence</i>	<i>Very serious consequence</i>
<i>Extremely seldom</i>				
<i>Seldom appearance</i>			X	
<i>Infrequent appearance</i>		X		
<i>Moderate appearance</i>	X			
<i>Frequent appearance</i>				
	<i>Acceptable</i>	<i>Alarm Area</i>	<i>Un-acceptable</i>	

Table 5.3 Risk assessment

Following M. Lehman, 2002

Lehman (ibid:7) makes the two main points with respect to risk management:

- *Quantify risks in terms of probability and consequence*
- *Compare results with risk acceptance criteria to determine whether or not risks are acceptable*

The overall outcome is a conscious handling of risk as induced by uncertainty.

An extensive literature pertains to the handling of uncertainty in relation to investments. The traditional method for evaluating investments has been to calculate net present value between the alternative of making an investment or not. If net present value is positive and greater than other alternatives, then go ahead. Dixit and Pindyck (1994:3) list three distinct characteristics of investments:

- *Investments are partially or completely irreversible*
- *There is uncertainty over future rewards from an investment*
- *Investments have leeway on timing that may be used to obtain further information*

On this basis Dixit and Pindyck (ibid.) argue that investments are to be understood as *options*, that actual investments carry a cost of losing the option, and that this opportunity-cost should be included in the financial evaluation. This idea of financial value of options puts emphasis on models to reduce uncertainty.

5.4 SCENARIOS AS TOOLS

This section introduces the concept of Scenarios as tools to develop people's mental models. Scenarios are interesting tools in the context of uncertainty because they help us to identify events and relationships. The approach goes well with modelling tools: scenarios may

serve as starting points for modelling, while the modelling helps to create structure and relate events to each other. Models assist in systematic and automatic handling of information which is often difficult in scenario development. For an introduction to scenarios see e.g. Schwartz (2001). Scenarios generally help towards “the art of taking the long view” as Schwartz (ibid.) puts it. The following provides a feeling for what is involved:

“How can you see, most clearly, the environment in which your actions will take place, and how those actions will fit with (or stand against) the prevailing forces, trends, attitudes and influences.

I wrote this book, initially, to show managers, business readers, and individuals how to begin using a method of investigating important decisions, a method that I and my close associates have found extremely useful for more than twenty years. And yet, this method is overlooked by most planning processes, usually because (although it includes analysis) it’s not “quantitative enough”. This method is the scenario...

...In a scenario process, managers invent and then consider, in depth, several varied stories of equally plausible futures. The stories are carefully researched, full of relevant detail, oriented toward real-life decisions, and designed (one hopes) to bring forward surprises and unexpected leaps of understanding. Together, the scenarios comprise a tool for ordering one’s perceptions. The point is not to “pick one preferred future”, and hope for it to come to pass (or, even, work to create it)... ...Nor is the point to find the most probable future and adapt to it... ...Rather, the point is to make strategic decisions that will be sound for all plausible futures.”

P. Schwartz, 2001:xiii-xiv

As an appendix Schwartz (ibid. 242 ff.) provides a useful recipe to develop scenarios. In the following we describe some main concepts, and the potential interaction between scenarios and modelling.

Scenario plots, building blocks and variables

The idea of *plots* comes from novels, plays and the movies. Stories have plots that tie the various scenes or situations together in sequences. In scenarios there are similar plots that tie together the parts of larger systems, and also largely define what those parts are. To get a good understanding of sequence is important. As in movies there is a limited set of recurring plots. A familiar example is the “Bright idea → Volume sales → Ensured income” that underlies the quantitative uncertainty discussed in section 5.2. The parts constitute the *building blocks* of a scenario, and they interact through a set of scenario *variables*. The latter express the central parameters that underlie the development with time in a given scenario, and thus help us to come to grips with the dynamic aspects. Here is an unclear area; scenarios are thought of in terms of both *situations*, which are static in nature, and in terms of *development* which is dynamic.

Internal consistency

One of the valuable sides to scenarios is the possibility of identifying internal consistency: parts interact through relationships that are smooth and likely, costs tend to be minimised, products and services to be improved are seen from the need they seek to meet. Companies (and bureaucracies) tend to grow and to increase their power and influence if at all possible. Development in which this does not happen, is inconsistent unless there are good reasons to hinder the normal outcome. Consequently, such reasons may be methodically sought for.

Another type of consistency relates to stability of the situation following a period of development. Stability in this sense means that one can stop to catch one’s breath – to evaluate a situation that is not

changing rapidly. A scenario is consistent in this sense if it proves possible to find stable situations in a changing world.

Scenarios guiding input to models

Dynamic modelling as in the present work needs to start from some defined situation as described above. Scenarios with stable situations are suitable in this respect: they portray understandable, carefully argued situations to start from. This means that one is not tied to any particular fixed “now”-situation, but may use alternative scenarios with specified outsets. Modelling (to the extent that the model reflects reality) will start from this outset, to identify likely development towards new situations characterised by stability.

Models as tools to create scenarios

We finally turn the argument around, to evaluate modelling as an aid to construction of scenarios. The model now assists in describing and understanding the underlying basis of a scenario, and in understanding the dynamic development towards new stable situations – new scenarios. The specification of model boundaries helps to delimit a scenario. In particular the dynamic approach in the present work helps to detect path dependencies that may steer development into unforeseen or intended channels.

In the present section we have looked, briefly, at the concept of scenarios and both how this may assist in modelling and how modelling assists in construction of scenarios. Scenarios are a powerful means of concretising vistas of the future. They also help to reveal the underlying assumptions that we, as human beings, have of the real world in which we live.

5.5 HOW THE EICOMP MODEL MAY REDUCE RISK

Dynamic models will, to the extent that they reflect reality, contribute to reduce uncertainty and thereby risk in two different ways: through

improved overall understanding of reality and through direct detection of areas of danger. Models that operate on a basis of probabilities will, in addition, help in gauging the probability for particular outcomes. The EICOMP model has this basis in probability from the background of dissipative systems and bifurcations (cf. section 4.2).

For the present discussion, it is useful to think in terms of a *response surface*, created by the model's output in response to changes in its input domain. Albeit possibly in multiple dimensions, one may think of a "surface" in the same sense as the terrain described by a geographic map, and with the value (cf. section 5.1) corresponding to the altitude in the map. Models designed for simulation figuratively display and characterise such response surfaces, to allow their description in terms of "peaks", "valleys", "plains", and "lakes". Optimising models will search for peaks, which correspond to local optima, and where the highest peak will be a global optimum. The optima in principle correspond to opportunities (H. Gether, 2000:127 and 130-131). Change (as discussed in Chapter 4) within this concept of response surfaces, is the movement on the response surface from one peak of opportunity to another. However, the valleys, plains and lakes are of equal interest. Lakes correspond to areas that are to be avoided (e.g. there are no degrees of bankruptcy); valleys correspond to situations that make moving from one peak to another more difficult, i.e. they prevent change. The plains represent areas corresponding to calm, uneventful existence.

The focus in the present section is how the EICOMP model may reduce risk. Under the requirement that EICOMP actually represents reality (cf. sections 8.1 and 8.2) and is used within its range of constraints, EICOMP helps to reduce risk in several ways:

- *By indicating peaks of opportunity*
- *By indicating routes leading from one peak to another*

Chapter 5

- *By identifying valleys that make moving away from a peak difficult (help to avoid “painting oneself into a corner”)*
- *By finding lakes and staying away from such areas*

The main theme in the present discussion is overview. This goes for assumptions as well, as they become revealed and consciously examined. Apart from these principal ideas towards the use of simulation models, an additional, very common approach is *sensitivity analysis* (cf. section 8.2).

The understanding developed in this chapter helps to handle value and uncertainty when modelling change and shifts in technology. Risk is reduced both through better qualitative understanding (i.e. what to include) and better judgement of probability of events. We now bring these concepts into the realm of modelling in the next chapter.

Models and modelling

“...The usefulness of models lies in the fact that they simplify reality, putting it into a form that we can comprehend.”

J. D. Sterman, 1991

The aim of this chapter is to provide an overview of basic sides to modelling, its limitations, and the central building blocks and concepts to help understand change and transition in energy systems. Models are helpful instruments in this process; therefore the birth of the EICOMP model. The use of models is pervasive in our thinking and our daily conduct. In some ways we experience models as simple and natural. Geographic maps are an example. In a number of other respects the concept of models is utterly complex and fraught with difficulty.

We use models in our communication with others. In this respect Meadows et al (1992) bring out an important point:

“... Every word in this book is a model. “Growth”, “population”, “forest”, “water” are just symbols, which stand for very complex realities. ...And of course our thoughts, and every person’s thoughts, are only models of the real world. Therefore we have a difficulty. We are about to talk about a formal model, a computer-based simulation of the world. For this model to be of any use we have to compare

it to the “real world”, but neither we nor our readers have an agreed-upon real world to compare it to. We only have the worlds of our mental models.”

D. H. Meadows et al, 1992:105

The dilemma that we interpret other models in terms of our personal mental models, leads to a requirement for careful description of what we do, the purpose of our efforts, and the results we obtain.

There are many different types of models and many different characteristics by which they may be distinguished and classified. Following an introduction to simpler models, we look first at the concept of system equations, and subsequently discuss in more detail the three kinds of models employed in the present thesis. These are respectively value sequences with end-products and hierarchical descriptions as introduced by H. Gether (2002), Porter’s (1980) model of competition, and system dynamics as introduced by Forrester in 1961 and significantly extended in recent years. The two last sections review the characteristics of major existing energy models, and discuss the importance and challenge in communicating dynamic behaviour.

6.1 BASIC CONCEPTS OF MODELLING

Geographic maps provide an example of a natural kind of model that is easy to understand, and at the same time lets us introduce a number of properties that are basic to all models. To begin with, a map has direct correspondence with the geographic terrain that it portrays. There are two *independent* input parameters in the form of the coordinates for North/South and East/West. By specifying these, information (for instance altitude) can be extracted about the specified position. The map may, however, return further and more complex information like the presence and position of cities, of distance between these, the existence of roads, and their quality. Such

scope and range of returned information is characteristic to models. Maps further reference sub-sets of all terrain, with boundaries corresponding e.g. to national borders or some other natural delimitation. Finally, a map returns information with a limited amount of detail. The boundaries, scope of returned information and the translation from input to output are all characteristic and important features of models in general. The impression of maps as natural and straightforward appears to result from several concomitant issues: input (i.e. coordinates) is straightforward, the output information is normally familiar, and there are no dynamic issues involved.

Through the example of maps we encounter some central characteristics of all models. We now consider such properties, starting with the concept of *system equations*. Such equations vary greatly. In geographic maps the system equations *translate* from the input (i.e. geographic coordinates) to the output (information about properties of terrain), and present detail or overview according to the scale of the map. All models generate output according to some input, as they *process* the input relative to the real world as described through the system equations. The scope of the input (the kind of input parameters and their allowed range) is called the *input domain* of the model. The range (and nature) of output is correspondingly referred to as the *output domain*.

6.2 APPROACHES TO MODELLING

As introduced above the system equations form the link between input and output domains. Section 6.2.1 discusses central principles of system equations and 6.2.2 discusses modelling approaches to complex systems.

6.2.1 BASICS OF SYSTEM EQUATIONS

We start with what is conceptually simple. The mathematical equations (functions) discussed in the following sections are well known and are not a subject by themselves. We use them to introduce aspects that are central to the application of all kinds of equations in modelling. A *system equation* is in principle anything that translates from an input into an output. It may consist of a set of simpler equations. In the same way that we used geographic maps as an example to introduce central concepts that apply to all models, we will use simple mathematical relationships to clarify general properties of system equations. Simplest of all is a description of something that does not vary, i.e. is constant:

$$y = c \quad (6.2.1.1)$$

We note that the nature of this value (i.e. the output domain) is determined by the dimension and unit of the constant "c". Some very important relationships belong to this category, like for instance the sum total of energy in a closed physical system, or the output from a value sequence that is limited by some bottleneck. In the latter case the outcome remains unchanged, determined by the capacity of the bottleneck.

Our next example is the linear relationship:

$$y = f(x) = k \cdot x + c \quad (6.2.1.2)$$

The main difference between equations 6.2.1.1 and 6.2.1.2 is that in the latter we have an input parameter "x" that is an "independent" variable. It may take on varying numeric values, and for each value the equation will return a "dependent" value for $f(x)$. In practice "x" has two direct constraints in the form of lower and upper bounds (x_1 , x_2), that determine the extension of the input domain. Similarly, there

are lower and upper constraints on the value that $f(x)$ may take, to correspondingly determine the extent of the output domain.

An important usage of the linear relationship is to express proportionality: something varies in direct proportion to something else. In the example of value sequences above, output varies proportionally to the capacity of the bottleneck, whereas it does not vary with respect to changes in other parts of the value sequence.

The linear relationship is the simplest model that allows us to introduce *time*, “ t ”, as an independent variable:

$$f(t) = k \cdot t + c \quad (6.2.1.3)$$

We use this simple relationship in mathematical terms to introduce the physical concept of change. Models that vary with time are generally referred to as *dynamic* models. However, although equation 6.2.1.3 does express how $f(t)$ varies with time, it does not describe the actual process of change: there is no “now” where things happen, and no “running” generation of history. Instead there is a static system as seen from the “outside”, where we observe a history and a future simultaneously. There is seemingly no difference between future and history (cf. also section 4.1).

In order to obtain an intrinsically dynamic model, we both need a “now” and a simultaneous recording of history. We need something that simulates “now”, in order to turn a *possible* future into a *recorded* history. When we record history we automatically introduce irreversibility (Prigogine, 1997:53). When the recorded history is used in the next time-step in a model, this model will reflect irreversibility. By differentiating equation 6.2.1.3 with respect to time we get a description of what happens at time *now*:

$$df(t)/dt = k \quad (6.2.1.4)$$

It is the interplay between an equation expressing variation with time, and its time derivative, that *together* form the fundamental tools to distinguish between now and history. In section 6.7.1 we show that system dynamics maintains both the derivative (“now”) and the history in form of integrals (“levels”), and they are kept carefully separated. See also section 4.1, which is directly concerned with a better understanding of the physical phenomena of change.

A final issue of system equations introduced here is *nonlinearity*. There are many versions of this. The mathematical expressions, $f(x) = x^2$ or $f(x,y) = x \cdot y$, express two common versions. A physical example of the first is the dependence of kinetic energy on velocity. The second expression describes the typical situation that outcome depends on more than one variable. A particularly important relation is the *exponential* function:

$$f(x) = c \cdot e^{kx} \quad (6.2.1.5)$$

This equation is nearly as familiar as the linear relationship given by equation 6.2.1.2. The first characteristic of the exponential function is that for each input value "x", it returns an output value $f(x)$ that is changed from some start value "c", and the kind of value returned is determined by the dimension (unit) of this initial value. "x" may typically be a measure of time to afford us a time dependent model:

$$f(t) = c \cdot e^{kt} \quad (6.2.1.6)$$

Note, however, that like the linear model (eq. 6.2.1.3), it describes the static picture from a starting time to an end-time, rather than turning a possible future into a recorded history.

A system with feedback exemplifies exponential nonlinearity. The following coupled equations represent a single feedback loop:

$$y_t = x_t \cdot \Delta t \quad (6.2.1.7)$$

$$x_t = k \cdot y_{t-1} \quad (6.2.1.8)$$

Here, the variable x_t depends on the previous function value y_{t-1} and the function value y_t depends on the variable x_t . If the constant k is positive, this is a positive feedback loop. If k is negative, it is a negative feedback loop.

The exponential relationship occurs in practice whenever output is some fixed percentage of a starting amount (e.g. the interest on a loan). Apart from being nonlinear, there are two particular reasons for looking closer at this relationship. The first is the property that the initial amount doubles (or halves if $k < 1.0$) in *fixed* intervals of time. The doubling time is:

$$t_d = \ln(2) \cdot 100 / x\% \quad (= 0.6931 \cdot 100 / x\% = 69.31 / x\%)$$

$$t_d \approx 70 / x\% \quad (6.2.1.9)$$

where $\ln(2)$ is the natural logarithm of 2, and "x%" is the percentage increase. Thus a loan with 10 % interest doubles in $70/10 = 7$ years, 7% interest doubles in 10 years, and 2.3% increase per year doubles a population within 30 years (i.e. within one human generation). The second reason why the exponential relationship is important is that it is notoriously badly estimated by people in everyday life (cf. e.g. Sterman, 2000:269).

The presence of nonlinearity has a number of consequences:

- *There may be (multiple) optima within the range of the input/output domains*
- *Integration of nonlinear functions generally lead to eigenvalues or separate, particular behaviour*

- *Functions are not necessarily reversible ("you can't unscramble an egg")*
- *Nonlinearity leads to dynamic complexity (cf. section 6.2.2)*
- *Mathematical treatment of sets of nonlinear equations soon becomes intractable*

This complexity has led to the application of linear relationships even where nonlinear descriptions would be appropriate. Thus Sterman (2000) cites the following observation:

“Up until now most economists have concerned themselves with linear systems, not because of any belief that the facts were so simple, but rather because of the mathematical difficulties involved in nonlinear systems...[Linear systems are] mathematically simple, and exact solutions are known. But a high price is paid for this simplicity in terms of special assumptions which must be made.”

Paul A. Samuelson, 1947:288

Cited from J. D. Sterman, 2000:551

6.2.2 COMPLEX SYSTEMS

Complex systems are systems with many degrees of freedom. The basics discussed above, like the constant, linear and exponential relationships allow us to understand and evaluate a host of situations and relationships in daily life. Even so, we encounter situations too complex to be handled in this way:

“We frequently talk about side effects as if they were a feature of reality. Not so. In reality, there are no side effects, there are just effects... ...Unanticipated side effects arise because we too often act as if cause and effect were always closely linked in time and space. But in complex systems such as an urban center or a hamster (or a business, society, or

ecosystem) cause and effect are often distant in time and space.”

J. D. Sterman, 2000:11

The complexity that we have to deal with is of two kinds. The first kind has to do with tangled relationships that may yet be handled by means of some functional description with one (or some few) independent variables. The second kind arises because a set of factors act together to create the final outcome, with indistinct relationships between these factors besides their contribution to the particular situation. Our field of inquiry concerns this latter kind of complexity:

“Most people think of complexity in terms of the number of components in a system or the number of combinations one must consider in making a decision. The problem of optimally scheduling an airline’s flights and crews is highly complex, but the complexity lies in finding the best solution out of an astronomical number of possibilities. Such needle in a haystack problems have high levels of combinatorial complexity (also known as detail complexity). Dynamic complexity, in contrast, can arise even in simple systems with low combinatorial complexity... ..Dynamic complexity arises from the interactions among the agents over time”

J. D. Sterman, 2000:21

There are two main approaches to modelling systems with multiple interacting factors. The first is to increase the number of independent variables. However, with more variables (i.e. “dimensions”) it becomes increasingly difficult to ensure their independence. The second principal approach to complex modelling is to describe systems as sets of subsystems that interact with each other. This is the approach chosen in the present work.

After identifying some general properties which pertain to many real world situations, we now discuss modelling approaches as tools to handle complex systems in focus in the present work. The overall modelling approach should be based on two main criteria. First we need to select approaches which give acceptable representation of reality. Secondly we must choose approaches which return the kind of information and knowledge we are seeking. There are several principal types of approaches, of which we will comment briefly on the following:

- *Linear programming*
- *Decision analysis*
- *Porter's model of competition*
- *Hierarchical structures of activities in value sequences*
- *Simulation*
- *System dynamics*

Linear programming

In this approach systems are described by means of sets of linear relationships. The main advantage of this type of representation is that it is well suited for optimisation, typically in situations with scarce resources that are brought together to obtain some wanted result. A linear programming model consists of an objective function, and a set of constraints. The objective function is either minimised or maximised. The constraints describe the boundaries of a solution space. The constraints may be equations or inequalities. The most efficient algorithm for finding the optimal solution to most linear programming problems is the Simplex Algorithm (e.g. Ravindran et al, 1987:35ff.). The objective function has the form:

$$Z = c_1x_1+c_2x_2+\dots+ c_nx_n \quad (6.2.2.1)$$

where a set of constraints on the x'es apply:

“In recent years, statisticians, engineers, economists, and students of management have placed increasing emphasis on decision making under conditions of uncertainty. This area of study has been called statistical decision theory, Bayesian decision theory, decision theory, decision analysis, and many other things. We shall henceforth use the term decision analysis as a matter of expediency and as an anchor point in discussing this topic.”

A. Ravindran et al, 1987:221

Decision analysis is a modelling technique that is also widely used in operations research. It introduces a systematic and clearly defined set of terms (concepts/tools) for approaching complex situations, and the idea of multiple, local decisions as the “independent variables”. Central concepts are:

- *The decision maker* (the agent that actually makes decisions)
- *Alternative courses of action* (to be explicitly specified)
- *Events* (situations *outside* the control of the decision maker)
- *Probability* (of an event taking place)
- *Consequences* (benefit or payback from an alternative)
- *Expected value of perfect information (EVPI)* (to judge whether further information should be sought)
- *Maximising expected value (MEV)* (criterion for choice of alternative)
- *Minimising Expected Opportunity Loss (EOL)* (criterion for choice of alternative)

Armed with these tools, situations in the real world may be modelled as payoff- and regret matrices, or more pictorially as *decision trees*:

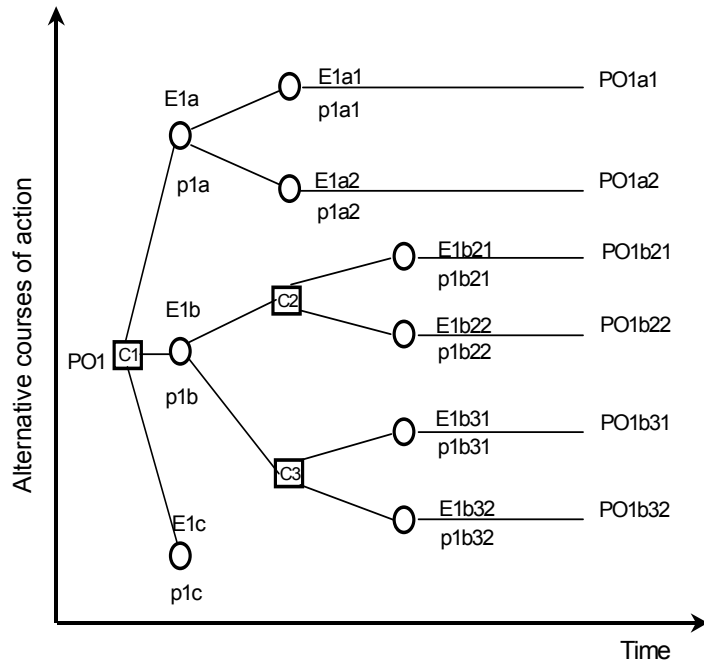


Figure 6.2.2.1 *Illustration of decision tree with two instances of decisions*

(□ = decision, ○ = event, p = probability, PO = payoff)

H. Gether, 2002:157

Decision analysis requires that a set of alternative courses is worked out and that each consists of a set of events. Each event has an estimate for the probability of its occurrence, and an estimate of payoff. This latter estimate is *from the point of view of the interested party* (i.e. the decision maker). The tree is evaluated from the right, and decisions are made so as to obtain the best combination of payoff and of probability for obtaining this payoff. This method, or way of thinking, describes the real world as a set of alternative courses. It deviates from strict deterministic behaviour by allowing the *probability* that something will actually occur. This is a very important step. Decision analysis also introduces an element of

human ingenuity both in discovering and describing alternative courses of action, and in providing estimates of probability and payoff. Like linear programming the element of optimisation is present, but it becomes “decentralised” in the sense that it applies to individual decisions that are made towards an overall optimum.

The main limitation of decision analysis appears to be what is also its strength, namely that individual elements (in the form of alternatives) are evaluated in isolation (before being summed up). As each alternative course requires careful setting out, and probabilities and payoffs for all decisions and events must be estimated, the process soon becomes work-intensive. Decision analysis therefore appears most amenable where there are few alternative courses of action, and with important decisions that deserve the proper attention.

It is also possible to use optimisation algorithms to analyse situations where the future is uncertain. This is commonly referred to as stochastic programming. Through stochastic programming it is possible to automatically optimise decisions in many decisions analysis problems. Still, the future scenarios must be estimated from statistical data, or derived from the decision maker’s knowledge.

Porter’s model of industry competition

The main function of this model is to give a better understanding of how competition in markets functions with a multiplicity of decision makers (i.e. buyers), and how individual firms may approach competition in a context of strategy.

Competition leads to implicit optimisation through minimisation of value sequence costs and in the EICOMP model through maximisation of attractiveness. It is described in more detail in section 6.6.

Hierarchical structures of activities in value sequences

Hierarchies are a valuable way of structuring information. In the context of end-products, hierarchical structures of activities called value sequences are a means of optimising routes to end-products. This is described in more detail in section 6.5. The value of hierarchical structures lies in their ability to structure information in a rational way in order to obtain clear descriptions of perceived reality. A very clear example is given by Dawkins (1986):

“If I ask an engineer how a steam engine works, I have a pretty fair idea of the general kind of answer that would satisfy me. Like Julian Huxley I should definitely not be very impressed if the engineer said it was propelled by ‘force locomotif’. And if he started boring on about the whole being greater than the sum of its parts, I would interrupt him: ‘Never mind about that, tell me how it works’. What I would want to hear is something about how the parts of an engine interacts with each other to produce the behaviour of the whole engine. I would initially be prepared to accept an explanation in terms of quite large subcomponents, whose own internal structure and behaviour might be quite complicated and, as yet, unexplained. The units of an initially satisfying explanation could have names like fire-box, boiler, cylinder, piston, steam governor. ...Given that the units each do their particular thing, I can then understand how they make the whole engine move. Of course I am then at liberty to ask how each part works. Having previously accepted the fact that the steam governor regulates the flow of steam, and having used this fact in my understanding of the whole engine, I now turn my curiosity on the steam governor itself.”

R. Dawkins, 1986:15

Dawkins refers to this type of description as “hierarchical reductionism”. A second important property of hierarchies is that

they allow an understanding of how the root of the hierarchy may be optimised through adopting and adapting of individual activities in the hierarchical structure, provided the activities are associated with some measure of value that relate to the activity's parent activity. As shown by H. Gether (2002), this is closely associated with how we understand biological evolution (Goldberg, 1989).

Simulation

Simulation is central to the present work. Reasons for adopting simulation (Ravindran et al, 1987) are:

1. *Simulation makes it possible to study and experiment with the complex internal interactions of a given system whether it be a firm, an industry, an economy, or some subsystem of one of them.*
2. *Through simulation, one can study the effects of certain information, organisational, and environmental changes of the operation of a system by making alterations in the model of the system and by observing the effects of these alterations on the system's behaviour.*
3. *A thorough observation of the system being simulated may lead to a better understanding of the system and to suggestions for improving it, which would otherwise go undetected.*
4. *Simulation can be used as a pedagogical device for teaching both students and practitioners basic skills in theoretical analysis, statistical analysis, and decision making.*
5. *The experience of designing a computer simulation model may be more valuable than the actual simulation itself. The knowledge obtained in designing a simulation study frequently suggests changes in the system being simulated. The effect of these changes can then be tested via simulation before implementing them on the actual system.*

6. *Simulation of complex systems can yield valuable insight into which variables are more important than the others in the system and how these variables interact.*
7. *Simulation can be used to experiment with new situations about which we have little or no information, so as to prepare for what may happen.*
8. *Simulation can serve as a 'pre-service test' to try out new policies and decision rules for operating a system, before running the risk of experimenting on the real system.*
9. *For certain types of stochastic problems the sequence of events may be of particular importance. Information about expected values and moments may not be sufficient to describe the process. In these cases, simulation methods may be the only satisfactory way of providing the required information.*
10. *Simulation analysis can be performed to verify analytical solutions.*
11. *Simulation enables one to study dynamic systems in either real time, compressed time or expanded time.*
12. *When new elements are introduced into a system, simulation can be used to anticipate bottlenecks and other problems that may arise in the behaviour of a system.*

A. Ravindran et al 1987:376

Cited from Naylor, T. H. (1971)

There is a gliding transition from a position of *creating* a mechanism for imitating the real world and operate this to see what *will happen*, to a position in which a model or mechanism is *proposed* and to see if what happens in the real world corresponds to the proposed mechanism. In the first case we seek to use what we already know, while in the second we seek an understanding of nature. (See further description in section 6.7).

System dynamics

System dynamics thinking introduces *causality*, *feedback* and *delay* (between levels and rates, cf. section 6.7). It further provides a mechanism for creating an artificial “*now*” with recording of history. In this way an intrinsically dynamic structure is obtained. Thus system dynamics is dynamic par excellence. It is fundamentally different from the dynamic concept described by equations 6.2.1.3 and 6.2.1.6 (i.e. the linear and exponential variation of something with time). System dynamics include the mechanisms that actually perform change, whereas equations like the exponential development with time (eq. 6.2.1.6) only report the overall outcome. System dynamics is the main methodological basis of the present work, combined with Porter’s model of competition, and hierarchical value sequences. The principles of system dynamics are treated in detail in section 6.7.

As we move from basic systems to complex ones, we encounter a range of system equations and methods geared towards different needs. The methods commonly used in operations research, like linear programming and decision analysis, have the advantage of producing optimal solutions to many problems. However, they require careful consideration of their areas of validity. In complex, nonlinear and dynamic situations other tools are needed.

6.3 THE RELATIONSHIP BETWEEN MODELS AND REALITY

The concept of modelling relies on a correspondence between a model and a situation in the real world that this model seeks to imitate. Our focus in this section is a closer examination of this correspondence. It is an area that philosophically has caused great difficulty. What is “reality” in its own right, as opposed to our perceptions through the signals that we obtain from it? This is the essential issue in Immanuel Kant’s concept of “The thing in it’s own”. A clear exposition of these problems is given by Russel

(1912:9-30). As we saw in the introduction to this Chapter, Meadows et al (1992) takes the position that we interpret the real world through mental models in our heads. Eigen and Winkler (1981) see a requirement for more than one person to experience the same, as a necessary position:

“...Interpretation is a process that takes place in our brains, and as a rule it is a co-operative effort. “Science” interprets. That means that a number of minds agree that in a given phenomenon there is something that occurs with regularity, can be reproduced, and can be traced back to recognizable causes, something, indeed, that can be interpreted.”

M. Eigen and R. Winkler, 1981:21

A related problem has to do with *determinism*. That something is deterministic means that it follows unconditionally from something else. Direct relationships such as equations 6.2.1.2/6.2.1.3 and 6.2.1.5/6.2.1.6 have this property, whereas decision analysis treats this issue in terms of probabilities of events. Determinism meets with deep problems in several ways. One is the concept of “free will”:

“Common sense inclines, on the one hand, to assert that every event is caused by some preceding events, so that every event can be explained or predicted. ... On the other hand,... common sense attributes to mature and sane human persons ... the ability to choose freely between alternative possibilities of acting”

K. Popper, cited from Prigogine, 1997:1

This “Dilemma of determinism” arises as any execution of free will appears to others as randomness. The only possible action is to reason about a person’s propensity to choose in some particular way, and to express this in terms of a probability for a given decision. To the

extent that the person matches this probability, his/her free will is restricted.

The basic laws of science, such as Newton's laws, have this property of being deterministic. In "The End of Certainty" (1997), Ilya Prigogine presents ways of coping with this, and relates under what conditions determinism breaks down (chapters 3, 4, 5 and 6). In general, the outcome is that any result from a model (and from understanding of how reality "works") comes to us in terms of probabilities.

In practical life it is more concrete. Models are made for a purpose and may be interpreted and understood in terms of this purpose. Yet, as expressed by J. D. Sterman (2000):

"A main purpose of modeling is to design and test policies for improvement. To do so, the client must have confidence that the model will respond to policies the same way the real system would. Fitting the logistic curve (or any model) to a data set does not identify the specific feedback processes responsible for the dynamics. The ability of a model to fit the historical data by itself provides no information at all about whether its response to policies will be correct."

J. D. Sterman, 2000:330

All results from models require care in evaluation with respect to reality.

Input to models takes the form of specifying "something" within the model's system equations and within its input- and output domains. This specification is in the field of data and information. In the next section we study this in more detail.

6.4 SPECIFICATION OF INFORMATION TO MODELS

Models vary widely in their requirements to information input, the format of this input, and how extensively and flexibly they allow description of reality. Simple system equations like the linear or exponential relationships have a numeric description of the input domain (in the form of upper and lower limits on the independent variable), a textual description of the output domain (i.e. what the output is and how it is to be used), and a textual and numeric specification of the parameters of the system equation itself.

Complex systems require specification of information to be tailored to the particular model in question. This implies a translation into some particular format.

In linear programming the description of reality is formulated through linear constraints with variables, constants and coefficients. In decision analysis reality is formulated as sets of events and alternative courses of action with probabilities and payoffs. In Porter's model of industry competition it is formulated as products and market segments, and in value sequences as descriptions of activities and sub activities. In system dynamics (cf. section 6.7.1) reality is described as rates and levels of flows with couplings between the flows. These flows and couplings represent the dynamic behaviour of a set of activities. In the EICOMP model individual flows reflect activities in hierarchies. Thus description relates to both activities and to rates and levels. In addition there is information relating to markets and competition.

6.5 VALUE SEQUENCES, END-PRODUCTS AND ACTIVITIES

Value sequences are hierarchical structures that open for studying and evaluating important activities together, and at the same time seeing each individual activity in context (Gether, 2002:205-207). We

describe this in considerable detail as it is a fundamental concept both to the approach in this thesis and in the EICOMP model itself.

Value sequences rest on the idea of *end-products*. These are products and services that are either consumed by *end-users* or are tools to create further end-products. End-products pay for all activities leading to them, and therefore decide the success or demise of firms involved in the supply sequences of the end-products. The firms, therefore, have a common interest in the competitiveness of their common end-product. For example, buildings are end-products, while bricks are not. Bricks are instead materials that are consumed in the processes of bringing forth end-products. Value sequences are independent from individual firms, and are systematically described as substructures of *value-oriented activities*. A final activity delivers the end-product to the customer. This activity is built up from a limited set of sub-activities (e.g. design, production, marketing, etc.). These activities in turn have their own sub-activities, to a level where organisation and technological function is fully described. The full concept of a value-oriented activity is illustrated in Figure 6.5.1:

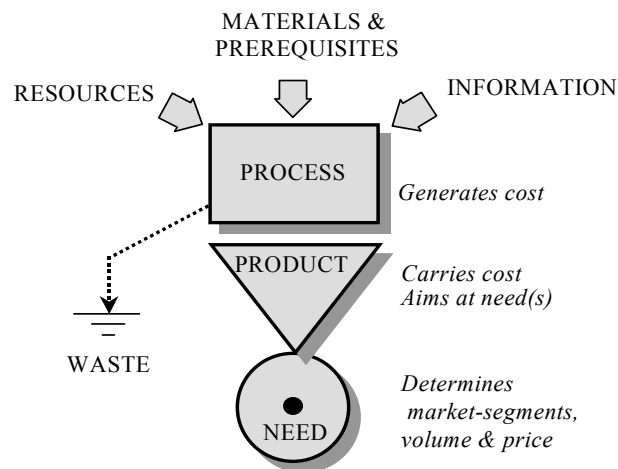


Figure 6.5.1 *The concept of a value-oriented activity*

H. Gether, 2002:184

Like activities in project- and production planning, value-oriented activities carry information about the product/service that they deliver. They also carry information about the process that each activity carries out. Costs are incurred through the deployment of resources and eventually through losses and waste. These costs are carried further downstream via the various activities in the value sequence. Consumption of materials in an activity does not incur costs in this activity, these costs originate instead in the activity that provides the material. However, the costs of materials are carried on via the required activities, eventually to the end-product of the value sequence.

The property that sets value-oriented activities apart from traditional activities in for example production- or project planning, is that all activities carry descriptions of *value* (for further discussion of value see section 5.1). Value is described through the *need* that the activity seeks to fulfil, which in the hierarchical structure translates to the contribution each activity makes to its parent activity. In this way a connected sequence of descriptions is created from detailed operations, to the need that the end-product itself seeks to meet. The final activity of the value sequence translates the meeting of needs into a measure of money in *markets*.

End-products constitute the direct supply of “consumables” to society, and are thereby directly connected to the standard of living. A direct link is therefore created between standard of living and technology. This technology becomes described through the activities in the value sequences.

Value sequences vary with time as individual activities may be improved or substituted with new activities that meet the same need(s) in better ways or with less costs. Such substitution of activities constitutes a process of overall improvement in bringing

forth the end-products, and perforce describes improvement towards a higher standard of living. In the EICOMP model multiple value sequences and their corresponding end-products compete towards end-user needs. The model thereby dynamically brings in improved activities. As competition becomes central, we look further into essential features of industry competition in the next section.

6.6 PORTER'S MODEL OF INDUSTRY COMPETITION

Competition is the second main concept and building block applied in the approach of this thesis, and is an integral part of the EICOMP model. The competitive framework opens for commensurable comparison of industries even though these provide different kinds of energy services.

Competition is a central concept in the field of business strategy. It is concerned with situations with multiple decision makers in markets. A vast literature on strategy (e.g. Mintzberg et al, 1998) is concerned with the problem of what determines the success of some firms, and the demise of others. In the preface to his book “The Competitive Advantage of Nations” M. E. Porter (1990) describes this central issue:

“Why do some social groups, economic institutions, and nations advance and prosper? This subject has fascinated and consumed the attention of writers, companies and governments for as long as there have been social, economic and political units. In fields as diverse as anthropology, history, economics and social science, there have been persistent efforts to understand the forces that explain the questions presented by the progress of some entities and the decline of others.”

M. E. Porter, 1990: xi

The essence in this is that those who prosper are those who win out in industry competition. In 1980 Porter introduced a model for understanding competition that outside the “plain quarrel” between directly competing contestants, introduced the ancillary role of suppliers and buyers, threats of substitutes, and of new entrants:

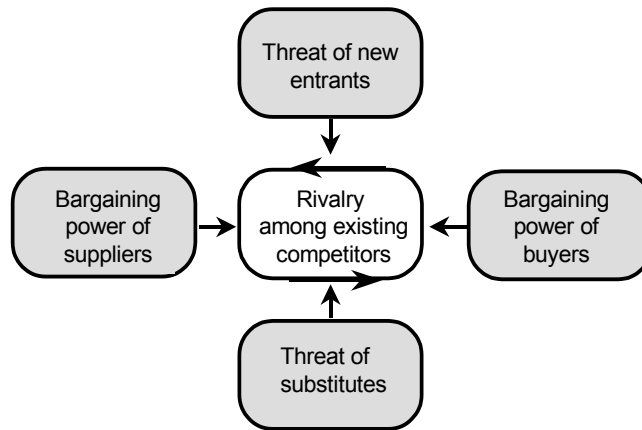


Figure 6.6.1 Porter’s model of industry competition

Following M. E. Porter, 1990:35

Threat of new entrants and threat of substitutes set upper limits to prices that may be charged. Bargaining power of suppliers and buyers take away some of the margins between prices and costs that would otherwise accrue to the existing suppliers. An extensive literature applies to the understanding of the rivalry between existing competitors, not least from Porter, who introduced the concepts of “cost leadership”, “product differentiation” and “market focus” (cf. e.g. Porter, 1990:39 ff.). An important issue is the concept of *buyer value* (cf. section 5.1.2). This pertains to the image the customer has of what she buys, and is often considered in terms of cost, quality and availability. How this is incorporated in the EICOMP model is described in section 7.2.

Value sequences and competition help us to describe structures and mechanisms of the real world. In the following section we turn to its dynamic behaviour, where system dynamics is used as methodology.

6.7 SYSTEM DYNAMICS WITH CAUSALITY, FEEDBACK AND DELAY

System dynamics is the third main concept and building block in the approach of this thesis. It provides us with the means to handle causal relationships, feedback and delay. From this approach a basis is developed for the integral dynamic structure in the EICOMP model, which handles competition among value sequences. As system dynamics is so central to the present work, we present it in a detailed manner.

System dynamics was invented by J. W. Forrester in the late 1950ies, based on control theory. His book “Industrial Dynamics”, from 1961, is an excellent introduction to its central concepts and background. Over the years it has grown into a distinct discipline, and although widely known, has remained in some ways curiously isolated, for example with respect to such fields of endeavour as operations research. In recent years two separate factors appear to contribute to its much wider dissemination and use. The first is the availability of new sophisticated software tools like “Vensim” from Ventana Systems, Boston, U.S., and “Powersim” from Powersim Software, Bergen, Norway. The latter has come to be the main platform of the present work. The second factor is the work “*Business Dynamics – Systems Thinking and Modeling for a Complex World*” by J. D. Sterman (2000), presently of MIT. This provides an extensive introduction to modelling in general, and to stumbling blocks and merits, in a wide area of applications. It has been fundamental also to the present work. The present author came to know about system dynamics through reading the works “*The Limits to Growth*” and “*Beyond the Limits; Global Collapse or a Sustainable Future*” by

Meadows et al (1972, 1992). These books discuss global sustainability based on system dynamics thinking and analysis.

6.7.1 BASIC MECHANISM

In this section we concentrate on the actual working of system dynamics. The basic concepts for describing reality are a system of *levels* (reservoirs) and *rates* (of flows) going into and out of these reservoirs. In addition there are *auxiliary variables* to make intermediate information available. Information from the levels, rates and auxiliary variables are connected to channel information back into rates. This is a very helpful metaphor for what is basically the Euler method of numerical integration of systems of nonlinear differential equations (cf. e.g. Sterman, 2000:903-911). In some environments the expressions “stocks” and “flows” are used instead of levels and rates, respectively. Difference in inflow or outflow from a level (reservoir), or change in either of them, respectively, will cause a change in the level with temporal development. This introduces *delay*. Each rate is described by equations that control the flows into, or out from a reservoir (this may be understood as control of valves). The equations may be of any mathematical form, linear or nonlinear, and may have other levels and rates as input variables. An important point is that all levels and constants in equations must be initiated to some value at the start of a simulation. This leads to the requirement for determination of *boundary conditions*, which is the way of specifying the input domain as discussed in section 6.2. Boundary conditions for the EICOMP model are described in more detail in section 7.5.

A system dynamics model consists of an overall system equation. The system equation consists of simpler equations, and defines the model structure. The equations in the present work represent activities in a set of value sequences. The basic building blocks in system dynamics are presented in Figure 6.7.1.1:

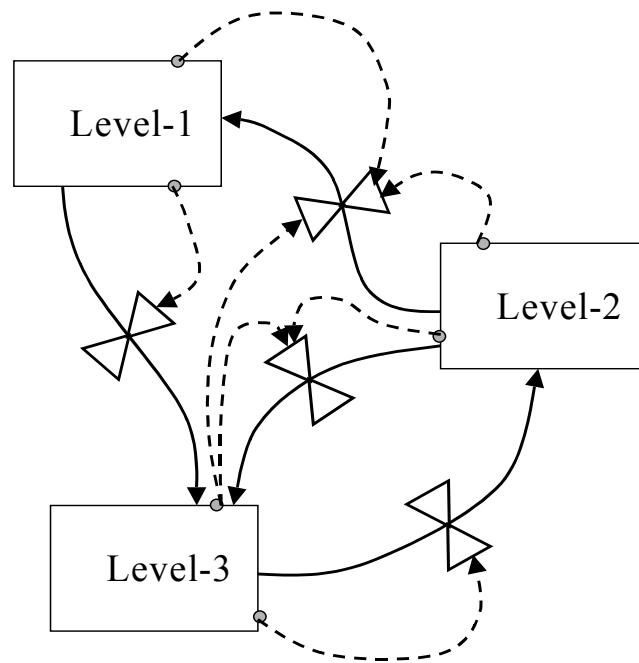


Figure 6.7.1.1 Basic system dynamics model structure

Following J. W. Forrester, 1961:67

Explanation of graphic symbols:

- Multiple levels (rectangles)
- Rates (continuous lines) that transport the contents between these levels
- Rate controls / decision rules (shown as valves)
- Information channels (stippled) that send information about levels to rate controls

Forrester (1961:69) refers to the rate controls as “decisions”, although they may be continuous functions (i.e. there is no limit to some set of alternative courses, as we saw with decision analysis).

Levels

Levels represent accumulations within a given system, like inventories, or H₂, oil or natural gas in a part of a value sequence. Levels are the quantities that result from the difference between inflows and outflows over time. They represent the *memory* of the system.

While a simulation is running, *all* change in levels comes from rates. This means that *all memory* and continuity from the past to the future *exist* in data about levels.

Rates

Rates represent the actual, instant situation of what is going on, and provide the running description of *now*. The rates are determined by the current levels, auxiliary variables and constants. They are analogous to the settings of the “valves”.

Computation

We now turn to the actual mechanism system dynamics applies to calculate, and to create an artificial “now” (cf. background in section 6.2.1). This is illustrated in Figure 6.7.1.2:

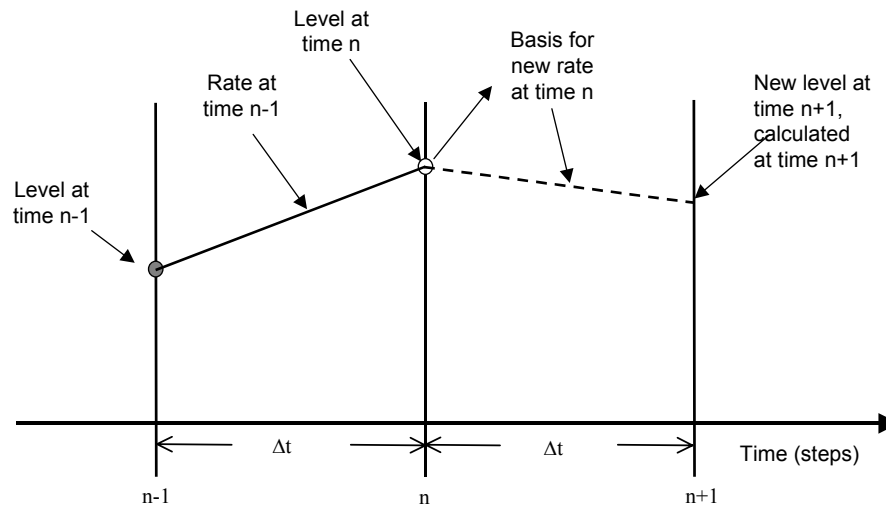


Figure 6.7.1.2 Basic computation in system dynamics

The procedure described in Figure 6.7.1.2 corresponds directly to the Euler method of numeric integration. Our concern here is primarily to show how this procedure lets us introduce an “artificial now” and a direction of time that is only realised when a historic reference is available.

The computation operates with time-slices Δt , small enough for change within a time-slice to be considered linear (for a discussion of size of Δt see discussion by Sterman, 2000:907 ff.). At any one stage, the three times $n-1$, n and $n+1$ are considered, with “ n ” representing “now”. All computation is with respect to time n . At n , levels, rates and auxiliary variables at $n-1$ have already been calculated and are known. The next operation is to calculate *levels* at time n , based on levels and rates calculated at $n-1$. With levels known at time n , we have a basis for calculation of rates and auxiliary variables for the next time-slice ($n \rightarrow n+1$). We then shift the indexes so that n becomes $n-1$ and $n+1$ becomes n , and repeat the operation.

This careful consideration of when information is available with respect to *n* (i.e. “now”) has lead Forrester (1961) to an important general conclusion about knowledge of the future: *we simply don't have such knowledge.*

“There are no exceptions to this unavailability of future information. Forecasts are not future information; they are present concepts about the future, based on information available in the present and the past.”

J. W. Forrester, 1961:74

In sum, we see as the basis for system dynamics, the elegant shifting between levels and rates. It gets around the traditional problem of simultaneous equations and obtains sorely needed simplification of the requirements of mathematics to describe the real world. As a result, it becomes possible to target, and to cope with, systems with a complexity that otherwise are out of reach. In the next section we look at how features of system dynamics may be instituted to take advantage of this mastering of complexity.

6.7.2 NETWORKS, CAUSAL LOOPS AND HIERARCHIES

Systems of equations soon become difficult to deal with. Adding dynamic behaviour with delay and feedback increases this difficulty substantially. Sterman (2000) describes this:

“The robustness of the misperceptions of feedback and the poor performance they cause are due to two basic and related deficiencies in our mental model. First, our cognitive maps of the causal structure of systems are vastly simplified compared to the complexity of the systems themselves. Second, we are unable to infer correctly the dynamics of all but the simplest causal maps.The heuristics we use to judge causal relations lead systematically to cognitive maps that ignore

Chapter 6

feedbacks, multiple interconnections, nonlinearities, time delays, and the other elements of dynamic complexity.”

J. D. Sterman 2000: 27-28

In this section we look at a set of ideas that help to gain overview and manage.

Interconnected networks

The first tool is called “interconnected networks” and was also introduced by Forrester (1961:70). The basic idea is to keep flows apart and congruent: there is one flow of hydrogen, one of oil, one of money, one of orders etc. This assures qualitative distinction between flows, and consistency in units. For modelling of industrial activity Forrester introduced six networks:

- *Materials network*
- *Orders network*
- *Money network*
- *Personnel network*
- *Capital equipment network*
- *Interconnecting information network*

While the “materials network” connects levels and rates of materials, “orders” do the same for orders, etc., the information network has the special property that it may hook levels on to rates for all of the other networks, as well as have feedback loops within the information network itself. The information network is in a position superior to the other networks and integrates everything into a whole.

Causal loop diagrams

Rather than focusing on the activities of a system directly, one may regard feedback loops as tools in themselves. In the real world they are ubiquitous and important to behaviour in all systems. Investigating such loops is essential to system dynamics analysis and

is a very powerful tool to understand dynamic behaviour. Feedback loops may be studied in *causal loop diagrams*. Sterman (2000:137 ff.) gives a broad introduction to this concept. Figure 6.7.2.1 illustrates the approach.

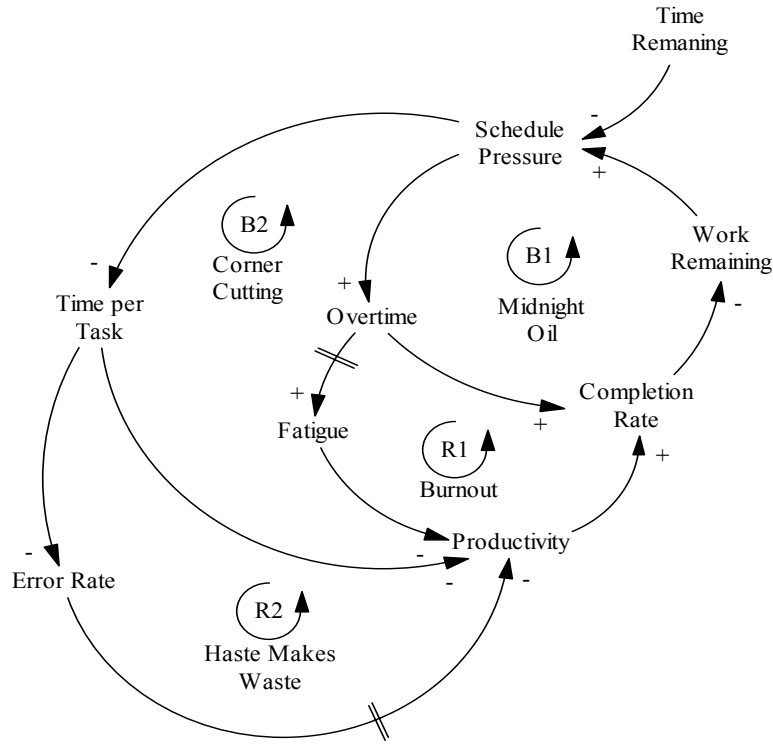


Figure 6.7.2.1 *Causal loop diagram and notation* (double crossing lines denote delay)

J. D. Sterman, 2000:149

The figure shows a causal diagram developed by engineers and managers in a work shop designed to explore the causes of late delivery for their organisation’s design work. The causal diagram represents the behaviour of the engineers trying to complete a project against a deadline. Loops return either a signal that opposes current change, or a signal that amplifies it. The first is known as a *balancing loop* (negative feedback), identified by a minus sign or a “B”. This kind of loop seeks to maintain a system in balance (i.e. as it is).

The other kind of loop is *reinforcing* (positive feedback, i.e. “feed forward”, “autocatalytic”) and is denoted by an “R” or a plus sign. It’s behaviour is quite different from the balancing loop: it will grow exponentially (cf. section 6.2.1) until meeting some constraint. Reinforcing loops can give rise to situations of *lock-in* (cf. section 4.2). A central characteristic of complex models is their tendency to display feedback loops. The identification of such loops through construction and actual running of models provides important outcomes of modelling. System dynamics is an excellent tool for such analysis: even the simple system displayed in Figure 6.7.2.1 includes a number of loops. Such is also the case in the current thesis.

Sterman (2000) stresses the point that loops are to be causal. This requirement goes back to the understanding of the real world that one seeks to model; it is not a property of the model itself:

“A system dynamics model must mimic the structure of the real system well enough that the model behaves in the same way the real system would. Behaviour includes not only replicating historical experience, but also responding to circumstances and policies that are entirely novel. Correlations among variables reflect past behaviour of a system. Correlations do not represent the structure of a system. If circumstances change, if previously dormant feedback loops become dominant, if new policies are tried, previously reliable correlations may break down. Your models and causal diagrams must include only those relationships you believe capture the underlying structure of the system. Correlations among variables will emerge from the behaviour of the model when you simulate it.”

J. D. Sterman, 2000:141

Hierarchies in system dynamics models

Hierarchies were described in section 6.5, and by good fortune the software supplier Powersim Software AS released a system dynamics software with entirely new capabilities for hierarchical structure:

“With the introduction of model hierarchy, Powersim Studio 2003 allows you to divide your simulation model into smaller submodels that hide away unnecessary implementation details. As each submodel can contain its own diagram book, modeling these submodels is just as easy as creating any other model. In addition, submodels can be created from components in any simulation project. This makes the duplication of existing model structures extremely easy! Utilizing model hierarchy allows you to make more abstract models, divide your model into subsections that are easier to maintain and model, and start reusing model structures from project to project”.

Powersim, Bergen, Norway, May 2003

The new software allows much better overview. This is vital to simplifying communication of comprehensive modelling tasks (cf. section 6.9). The present work would hardly have been feasible without this new development in software. The option of hierarchical operation of system dynamics models means that value sequences may now be integrated with system dynamics. To our knowledge, the present thesis is the first to attempt this.

6.7.3 ENDOGENOUS AND EXOGENOUS VARIABLES

In complex systems particular care is required with respect to what is inside and what is outside a model. Therefore, the important concepts of endogenous and exogenous variables are introduced. *Endogenous* variables generate the dynamics of a system through the interaction of variables and agents inside the model, and are in turn affected by this dynamic behaviour. *Exogenous* variables are variables that

influence the dynamics of the system, but they remain unaffected by the modelling effort. There is no feedback from the model to any exogenous variable. At each time-step each endogenous variable assumes a certain value according to the influence from the system. The values assumed by a variable for each time-step constitute a *time-path* which describes the variables' development with time.

To incorporate endogenous relationships means building required feedback structures. This rapidly leads to increased complexity and effort. In order to keep models as small and simple as possible, input would preferably be described exogenously. However, this means that any feedback from the area of concern in the model to the world outside is assumed to be zero. Sterman (2000:95-96) emphasises the importance of careful consideration of such assumptions. If there exist any significant feedbacks in reality, then the boundary of the model must be expanded and the variables in question must be modelled endogenously. A more detailed discussion of these questions related to the EICOMP model is given in section 7.5.

Thus far in Chapter 6 we have looked at important concepts and tools for modelling, to help us understand and study change in energy infrastructures. We have obtained a background to allow us to comment on features of existing energy related models. As a general observation the majority of such models are oriented towards modelling the operation of existing energy systems or extending these (e.g. whether to invest in a new power plant or transmission line), rather than modelling dynamic change from one kind of system to another. Furthermore, the majority of existing models are not concerned with communicating dynamic behaviour. These aspects are in focus in the last two sections of Chapter 6.

6.8 EXAMPLES OF EXISTING ENERGY MODELS

The motivation of this section is to give examples of common principles and approaches in existing energy models. These models are often referred to as “analytical tools”, “climate-” or “energy-economy” models. The literature of existing energy models is simply overwhelming (Wainwright and Mulligan, 2004; Huntington and Weyant, 2002; Bunn and Larsen, 1997; Fiddaman, 1997). When studying different models, the model structure is often only partially described. This is also true for underlying simplifications and assumptions. The result is to varying degree “black-boxes”, and uncertainty about how well the models correspond to reality.

A study of related work and models has been carried out to see what is available and to learn from earlier experience. No enquiry has been found that pertains directly to hydrogen technology and infrastructure and a focus on dynamic change in energy infrastructure.

Huntington and Weyant (2002) survey 16 economic models of energy supply and demand that have been used for global climate change analysis. While diverse, these models have a common concept of market-clearing equilibrium prices to balance production- and consumption for different types of fuels. Also, in order to make availability of energy endogenous, all models are global. Economic models generally compute the price of carbon that would be required to keep emission levels at some predetermined level. Most models solve a set of equations to find prices corresponding to equilibrium in supply and demand.

Huntington and Weyant (2002) find that the 16 models fall in three main categories according to representation of energy consumption. They term these “Fuel Supplies & Demand by Sector”, “Energy Technology Detail” and “Carbon Coefficients” respectively. The models may further be grouped according to the principal economy models they are based on, namely “Aggregate Production/Cost

Function”, “Multi-sector General Equilibrium”, and “Multi-sector Macro-economic”.

Models based on carbon-coefficients look at the aggregated effect of carbon on the economy, and are similar to models of aggregate energy use. In addition they omit inter-industry relations. Another category of models is centred on the energy sector in the overall economy, with supply and consumption of different kinds of energy as major parameters. These models (of which the “MARKAL” is an example, cf. below) explicitly represent capital stock turnover and introduction of new technology in the energy sector, but aggregate the rest of the economy. GDP is included in both types of models. Some of these models include the concept of *economic regions*. A third class of models combine multiple economic regions within a general equilibrium framework, and focus on interactions of firms and consumers in multiple sectors (e.g. in varying industries). This class of models mostly ignore economic parameters like unemployment and financial market responses. A fourth class of models combines elements of the first two classes, to present multi-region, multi-sector economic models with an explicit energy sector. They provide detail on capital stock turnover, energy efficiency and fuel switching.

Huntington and Weyant (2002) note that baseline conditions (technological, economic and political), the opportunities to replace fossil fuels, the costs of the turnover, and the dynamics of technological change, are central to determining a model’s response to climate change.

We now look briefly at three particular models to bring out other approaches and typical concepts and features.

The MARKAL model

MARKAL (MARKet ALlocation) is an energy system model based on linear programming originating from the 1970s. MARKAL was developed in a cooperative multinational project over a period of more than three decades by the Energy Technology Systems Analysis Programme (ETSAP) of the International Energy Agency (IEA). It holds a strong position in power generation and grid/network analysis. In this area there are many linear relationships, and change has proved to be small and incremental near some equilibrium (cf. ETSAP, 2004).

MARKAL has proven to be a powerful tool to find optima in known production/operation regimes. It is however, criticised for being used to foretell what's best to do in the future, where this approach is not robust, and for being applied outside its boundaries.

As one might expect, the MARKAL community has sought to respond to some of the critique. There are now several offshoots available and a variety of accessory programmes. Some have nonlinear features, but the fundamental structure and thinking appear to remain LP- and equilibrium based (Bakken, 2002:35-41). Size and complexity has grown considerably and a high price is paid because it has become a challenge to communicate the rationale of the model. A main problem seems to be that it is difficult to see what is caused by endogenous model structure and what is caused by boundaries and assumptions.

The TIMER model

The TARGETS-IMAGE Energy Regional model (TIMER) is a global energy simulation model developed at the “RIVM” (Rijksinstituut voor volksgezondheid en milieu) in the Netherlands. The main objective of TIMER is “to analyse the long term dynamics of energy conservation and the transition to non-fossil fuels within an integrated modelling framework”. It is used to generate scenarios for

the Intergovernmental Panel on Climate Change (IPCC). The TIMER model represents an extensive effort over many years (de Vries et al, 2001). At present the model is implemented for 17 *world regions*. TIMER further divides the economy in each region into five *sectors*: industrial, transport, residential, services and “other”, and has eight different *energy carriers*. The model is calibrated to reproduce the major world energy trends over the years 1971-1995, and is used to construct scenarios (cf. section 5.4) over the forthcoming century. TIMER is reported to calculate development year by year.

The model is built around the demand for *energy services* (useful energy) in the various regions. The demand varies with changes in population and economic activity under a special distribution of *energy intensity* within individual sectors. The actual shape of this distribution function is a major determinant of the demand for energy, and constitutes an important scenario parameter. The distribution function has information on *income elasticity* for energy (i.e. change in energy services per unit of change in activity).

The demand for energy services is adjusted by an “Autonomous Energy Efficiency Increase” (AEEI) to account for observed historical decrease in energy intensity. It is further corrected according to a “Price-Induced Energy Efficiency Improvement” (PIEEI) to include the effect of rising energy costs to consumers, and to reflect learning effects. The overall demand is subsequently distributed among energy carriers according to end-use energy costs, where such costs for individual carriers result from primary fuel costs, taxes and conversion costs and efficiencies. The relative difference between energy carriers is an important model parameter. From this outset of total energy consumption and distribution between carriers, environmental impact is calculated by means of a particular sub-module.

The TIMER model is built from five sub models:

- *The Energy Demand model*
- *The Electric Power Generation model*
- *The Solid Fuel model*
- *The Liquid Fuel model*
- *The Gaseous Fuel model*

The Energy Demand model was briefly described above. The Electric Power Generation model divides total demand for electric power into the four sub-carriers: hydropower, thermal, nuclear and renewable. Hydropower is exogenous. The remaining difference is met by thermal power and renewable generation through an allocation process described by a multinomial logit formulation (i.e. S-shaped curve) based on costs, and including a factor of preferences. The latter allows environmental concerns to be taken into account. The three fuel models describe supply and inter-conversion relationships. TIMER has the following inputs:

- *Regional population*
- *Regional macro-economic activity levels*
- *Submodel assumptions*
- *Energy intensity development*
- *Technology development*
- *Resource availability and fuel preferences*

and main outputs:

- *Use of primary and secondary energy carriers and feedstocks*
- *Production of energy carriers*
- *Energy-related and industrial emission of greenhouse gases and atmospheric pollutants*
- *Demand for modern and traditional bio-fuels*

As currently described TIMER has no structure to reflect hydrogen as an energy carrier and hydrogen related end-user technologies. It is recently reported that hydrogen is being incorporated as an energy carrier.

The ECS / IIASA models

Another significant effort is carried out under the Environmentally Compatible Energy Strategies Project (ECS) at the International Institute for Applied Systems Analysis (IIASA, 2003), Austria. Output from this project is also scenario projections for the World Energy Council and for the IPCC. Their effort is centred around a suite of modelling tools, of which the topmost is a scenario generator (SG). This relies on a model for energy supply strategy alternatives and their eventual environmental impact, called "MESSAGE". In turn this is integrated with a top-down macro economic model (MACRO). A particular module called "MAGICC" translates scenarios into climate impact. The last one, "ERIS" is a module to examine the effects of different representations of technological improvement. It provides a simplified multiregional representation of the global electricity market, supplied by a number of electricity generation technologies. It includes exogenous cost trends, and endogenous experience curves. A particularly valuable side to the efforts at IIASA is an extensive database (CO2DB) that is made freely available to third parties, and that is part of the data input in the present work as well.

The scope of the approaches above is similar. A principal difference between the ECS project models and the TIMER model is that the former is an optimisation approach, while TIMER (as part of the IMAGE scenarios) is a simulation model.

Perspective on energy/climate modelling

A challenge arising from the modelling examples is summed up by Baldwin (2002) as a representative of decision makers dealing with energy- and environment concerns:

- **Emphasis on building Tools, not writing static reports**
- **Build Institutional Memory:** (on-line) databases, case studies, analytical tools, etc
- **Web-based where possible and appropriate:** enable widespread, open access to tools; interactive in real time; promote collaborative development and systematic learning. Serve as a catalyst and forum for energy technology and policy analysis.
- **Interactive and Dynamic:** adjustable, not passive and static; data entry forms to develop estimates for key parameters, databases of case studies, other info resources
- **Adjustable:** allow input assumptions and even some algorithms to be adjusted by the user in order to test parameter sensitivities, and to develop alternative scenarios.
- **Collaborative; build user communities.** Design open architecture to allow contributors to develop tools for system, “delphi” estimates of parameters, other collaboration; gather/synthesize data at low cost.
- **Modular system(s):** provide early usability, increased flexibility, and ongoing adaptability and extensibility to new user needs and to build new capabilities.
- **Systematic and rigorous.** Coupled with its transparency and baselines, provides the means for comparing alternative scenarios and analytical approaches.
- **Keep it Simple; emphasize transparency over (false) precision**

Maintain skepticism and critically review models.

S. Baldwin, 2002

(Sam Baldwin, Chief Technology Officer and Member, Board of Directors, U.S. Department of Energy)

The models described above illustrate the complexity involved in energy-environment related issues. They also illustrate that models soon become large and unwieldy, making it difficult to reach and include decision makers. In particular, the aspects of dynamic change are hard to convey.

6.9 COMMUNICATING DYNAMIC BEHAVIOUR

In this section we focus on the demanding and critical issue of transferring information from modelling to people, and in particular to decision makers. Why develop new models, and do they get any better? Understanding of nonlinear dynamics, path dependence, irreversibility, disequilibrium-, economic- and innovation theory is improving, and to a larger extent can be combined and studied together as computers and software become more available and powerful. This creates a benefit from developing new models. In particular, interactive and dynamic models are useful for promoting improved mental models and learning together with stakeholders. However, as models become bigger and more complex, the main objective may be compromised: to be able to communicate and share new knowledge. The need for clearness and comprehensibility is essential to convey understanding to others. This is further described by Sterman (2000) who focuses on *transparency* and states the following:

“As computers become ever faster and more capable, simulations will not merely run faster, but the nature of the models, simulation software, and ways of interacting with models will be transformed. Among the tools future simulation software will include are:

- *Automated mapping of parameter space*
- *Automated sensitivity analysis*
- *Automated extreme condition testing*
- *Automatic, interactive parameter estimation, calibration, and policy optimization*
- *Automated identification of dominant loops and feedback structure*
- *Visualisation of model behaviour”*

J. D. Sterman, 2000:895-899

The present work has sought to attain some of this through the capabilities of Powersim Studio 2003 software (cf. last of section 6.7.2)

To achieve the features above, requirements will also have to be made with respect to the models themselves. They will have to be manageable by decision makers and executives, or the understanding and confidence in the results will disappear. This view has been the guiding star in the development of the EICOMP model as presented in Chapter 7.

Chapter 6

The Energy Infrastructure Competition Model

Guiding star:

To represent reality well

The purpose of this chapter is to describe the construction and operation of The Energy Infrastructure Competition Model (EICOMP). Sterman (2000:89) stresses the importance of *purpose* of any modelling effort. The purpose of the EICOMP model is to better understand the role of *change* in an energy system and how to actually achieve it. Special attention is directed towards hydrogen related technology and hydrogen as an energy carrier in energy infrastructures.

As described by Prigogine (1997:5) a fundamental starting point is that we live in a world of *becoming*: the road is created as we proceed, and gets shaped along the way. Where we end up depends on the paths we have already chosen and on path dependencies continuously broken and created. A central focus of the modelling effort is how transition plays out due to a system's structure, causality and feedback. The goal is to be able to identify and understand central mechanisms as a basis for policy making, not in an uphill struggle (where there are no driving forces, no positive feedbacks, or insurmountable balancing feedbacks), but in a regime of fair wind. This will always be an ongoing process. The ambition is to achieve practical change.

The current model also aims at promoting a common ground for discourse. This implies particular emphasis on transparency, i.e. that the model can be accessible beyond the range of specialist modellers.

What situation would we like to find ourselves in in the decades to come? Traditional projections and extrapolations have short term perspectives and change takes time (cf. Figure 1.2, and sections 3.6 and 4.3). Do we have a goal of a sustainable world and, if so, what does this goal entail? To add to the complexity, a goal of sustainability involves many different stakeholders: individuals, authorities, private firms, research communities, non-governmental organisations (NGO) etc. What means do we have, and what policy- and decision making needs to be done to get to where we want?

Modelling is a principal way of obtaining necessary knowledge. The current effort is inspired by Meadows et al (1992):

“All models, including the ones in our heads, are a little right, much too simple, and mostly wrong. How do we proceed in such a way as to test our models and learn where they are right and wrong? How do we speak to each other as fellow modelers, with the appropriate mixture of scepticism and respect? How do we stop playing right/wrong games with each other, and start designing right/wrong tests for our models against the real world?”

D. H. Meadows, D. L. Meadows and J. Randers, 1992:231

In this light the following sections describe the process of selecting the ‘right tools for the job’, and how these tools are applied. This includes overall rationale, structure, boundaries, variables, and data sources in the construction of the EICOMP model. The documentation is intended to make it available and reproducible, and to build confidence in the model and its results.

The challenge at hand is multifaceted and complex. The corresponding theory belongs in several scientific fields that do not fit neatly into one standard discipline. This is an important part of the problem, and also makes communication difficult. As outlined in the preceding chapters, a number of academic disciplines are applied together to promote an overall understanding of the underlying structures.

7.1 MODELLING TOOLS

The purpose of this section is to describe the thinking behind the tools selected to build the EICOMP model. We discuss this in the light of four themes:

- *The importance of purpose and regional focus*
- *The requirement for correspondence with reality*
- *Transparency of the overall model*
- *The embedding of models in their surroundings*

Importance of purpose and regional focus

Purpose is a main tool for deciding what has to be included in the model and what may be left out. For simplicity and transparency the least possible should be included. At the same time Forrester (1992) observes that:

“A model for simulating dynamic system behaviour requires formal policy descriptions to specify how individual decisions are to be made. Flows of information are continuously converted into decisions and actions. No plea about the inadequacy of our understanding can excuse us from estimating decision-making criteria. To omit a decision point is to deny its presence – a mistake of far greater magnitude than any errors in our best estimate of the process”

In order to focus the present work we concentrate on central energy-related decision points and their interactions. A next step is to delimit the model to a geographic region. A main argument for a regional model is that change has to start somewhere; it cannot start everywhere at once. However, once started it subsequently diffuses. Local conditions (anywhere, at any time) influence whether or not a process of change will get off the ground. A region may be quite small, or a larger economic region – the point is to focus on the problem of change, rather than modelling an entire system.

Correspondence with reality

This is key to any modelling effort. It requires that we specify what we believe are the essential sides to reality, and subsequently find ways to implement structure and relationships to obtain the corresponding behaviour in our model. With respect to the background of the present work the important aspects of reality were argued in Chapter 1 and specified in Table 1.1. To recap, central factors involved with change are seen to be:

- *Competition*
- *Market preference and choice beyond costs*
- *Bounded rationality*
- *Uncertainty and risk*
- *Irreversibility*
- *Increasing returns*
- *Path dependence*
- *Feedback*
- *Defence of existing investments*
- *Delay*
- *Nonlinear behaviour*

These phenomena/features are captured in the EICOMP model in the following way: the requirement that the model simulates competition is the central argument for introducing Porter's model of industry competition (cf. section 6.6). Market preference and choice beyond costs, along with bounded rationality, is included through a concept of attractiveness that allows description of multifaceted decision patterns. How this works is described in section 7.2. Uncertainty (cf. section 5.2) is inherently captured through the output from dynamic path dependencies, while risk may also be expressed as part of the attractiveness criteria.

The next point on the list is irreversibility. It is captured by combining "now" and "history" (cf. section 6.2.1): The mechanisms in system dynamics (cf. section 6.7.1) maintain the derivative ("now") and the history separate from each other. History is recorded in the form of integrals (levels). As history is recorded for each time-step (the passage of time), a *possible* future is turned into *existing* history. The recorded history, in turn, influences the next set of decisions in the following time-step. This is an irreversible process, creating a successive set of events. Irreversibility opens for incorporation of dissipative systems and bifurcations as described in section 4.1.

The remaining points of our requirement list for matching reality are all provided through the inherent properties of system dynamics (cf. section 6.7) underlying the EICOMP model.

Transparency of the overall model

A fundamental requirement is that the overall system equation of EICOMP is divided into a set of subunits, each with distinct correspondence to reality and with clear interconnections. It is achieved through value sequences leading to end-products (cf. section 6.5), and is described in increasing detail in sections 7.2, 7.3 and 7.4.

The embedding of models in their surroundings

How models exist in general, and interact with a surrounding world is illustrated in Figure 7.1.1:

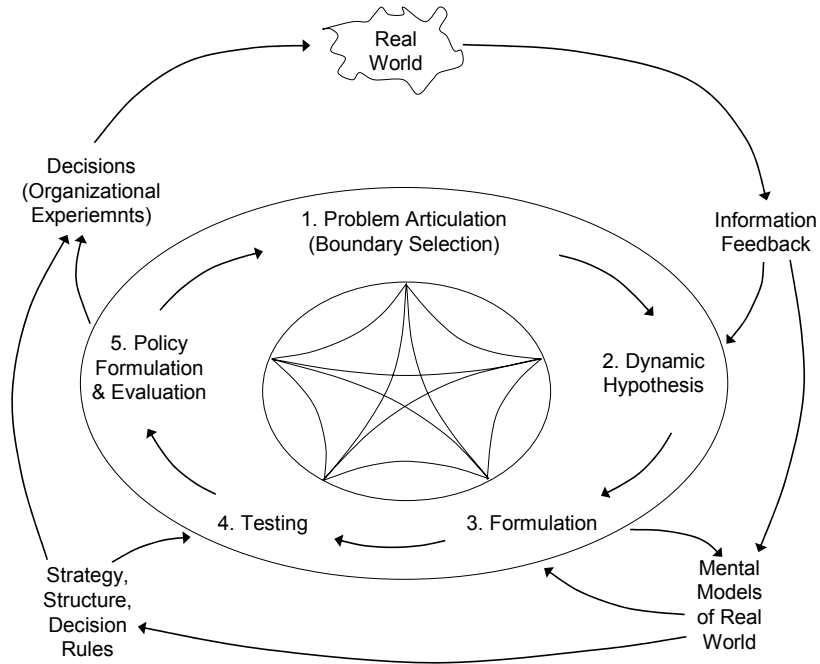


Figure 7.1.1 *The modelling process is embedded in the dynamics of the real world system*

Following J. D. Sterman, 2000:88

Figure 7.1.1 illustrates three important sides to modelling: the iterative process of building models, the boundary between a model and the real world (as seen from the model), and the overall, interactive process of using the model to better understand the real world. The boundaries of a model are important, as choice of boundaries strongly influences the design of the model and its use and behaviour. The boundaries further determine the possibility of

simplification, and they are important to achieve transparency. We discuss central sides to these issues in section 7.5.

The third side to modelling is the interactive use of models. This is an inherent ingredient in the overall effort. As conceived in EICOMP there is an active interplay in real time between the model, used as a ‘flight-simulator’, and the persons involved. Intrinsic dynamic behaviour is mentally difficult. The result is that most of our understanding is based on static perceptions (Croson and Donohue, 2002:74-82). Through dynamic interaction with a tool like EICOMP, a better understanding is obtained of why things turn out the way they do - often counterintuitive to what the static perception was suggesting (Sterman 2000:11, section 6.2.2). By understanding how the dynamics ‘works’ through studying the relative strength of feedback loops as the model is running, the participants (that should include decision makers and policy makers) improve their mental models. The use of EICOMP in this way opens for new ideas, to include knowledge and thinking (‘human ingenuity’) beyond what is present at any time in the model itself. In this way qualitatively nonlinear elements, like new ideas, play a role as the model and modelling is embedded in overall analysis and insight.

In the next sections we turn to the rationale of the EICOMP model and how it actually works.

7.2 MODEL RATIONALE

(For an overview map while reading this section please see the flip-out chart at the back (Appendix 2), together with a listing of module abbreviations (Appendix 1). Embedded in the chart is flow of information, from the markets through the infrastructure to the primary energy sources, and back).

In this section we describe the principles of operation of EICOMP. Both transport and stationary energy services are included. Stationary end-use is differentiated into appliances (high quality energy) and heating (low quality energy). A central feature in the model is to be able to study the interplay between the transport- and stationary sectors and how the main energy carriers (liquid fuel, natural gas, hydrogen and electricity) with corresponding end-user technologies operate together in an energy system for a region. The model is carefully based on the physical principles of energy- and mass conservation. This includes efficiencies, corresponding losses and carbon content per energy unit for all energy carriers throughout all value sequences.

A given modelling effort applies to a region. This region is described through a certain area, population, economic strength, reference need for transport and stationary energy services, and an energy infrastructure with initial capacity to meet demand in the year 2000. The region acts in accordance with international energy markets and can procure/obtain energy from the outside according to available infrastructure, or from local renewable energy.

In the EICOMP model there are two kinds of value sequence (see section 6.5 for description of value sequences). A primary type are value sequences that deal with the flow of energy (including losses) through *capacities* from primary resources to end-user technologies that meet end-user needs. We refer to these as *primary value sequences* (PVS). They consist of strings of *modules* (cf. Figure 7.2.1 for example refining/liquid transport/liquid storage/liquid ICE), and shift dynamically depending on the competitiveness of individual modules. The modules are characterised by what they do (e.g. liquid transport), and each module represents a technology with an associated capacity and efficiency. The second kind of value sequence deals with the supply of capacity for each module, and inherently handles cost of capacity, and delay in changing capacity,

in each module. We refer to these as *module value sequences* (MVS). The modules are generic and consist of *substructures* that are also generic (described in more detail in section 7.4).

The principal structure of primary value sequences (PVS) and module value sequences (MVS) are illustrated in Figure 7.2.1:

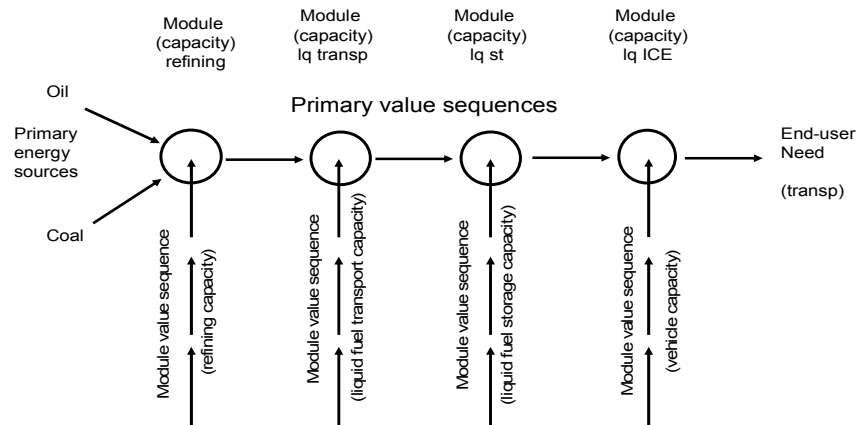


Figure 7.2.1 *Principal structure of the two types of value sequences in EICOMP*

The EICOMP model has seven central (out-of-region) primary energy resources and three local ones (cf. flip-out chart). Energy from these resources meets a variety of needs in the end-user markets, both stationary and related to transport. Each allowed ‘route’, from primary energy resources to end-user consumption, constitutes a PVS. In the example in Figure 7.2.1 there are two PVSeS, one connected with oil and the other with coal, and each delivering liquid fuels to the end-user markets. The flip-out chart shows the layout and connections considered in the full EICOMP model. There are a large number of potential PVSeS. EICOMP elucidates the flow in and among the various PVSeS as they change with time. The varying

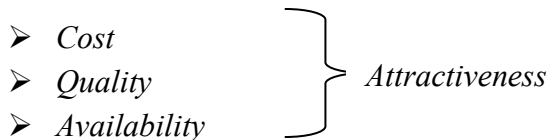
capacity of each module and the volume of energy flowing through the module (reported as “energy asked for” and “energy supplied”) are central scenario variables (see section 5.4). The volume in “Iq ICE” (cf. top right of flip-out chart) will for example report the development with time of volume of vehicles with internal combustion engines, under varying conditions of overall energy supply and technological development. Likewise, the “FC” module will give the development of fuel cell powered transport.

We now describe at a topmost level how the EICOMP model operates. Calculations are carried out for any given “now”. The model ‘begins’ in the end-user markets to the right in the flip-out chart. A population in a region can choose from a number of end-user technologies (options) to fulfil their needs (e.g. transport needs). This generates requests for corresponding energy. The choice entails both a capacity (such as a car) and flow of energy through this capacity. As one goes upstream (towards the left) further options are available. The modules here sum requests from downstream (initially from other market requests), and pass the requirement to the next upstream module(s). The requests eventually reach the primary energy resources (see more detail in section 7.3). From the leftmost (most upstream) modules the calculations return back downstream, now telling each downstream module (still at the same “now”) how much it will receive relative to what it ordered. How the choices are made is discussed below.

A module may receive all its “ordered” from its upstream module(s), or perhaps less from one, due to upstream constraints for this value sequence. It may then obtain more from other available upstream suppliers. The module supplies the next downstream module(s) with what it eventually obtained. If the downstream module(s) received less than requested, the inadequate availability is recorded. The upstream module also records that its capacity is too low to meet requests. The latter information becomes input to investments in new

capacity. Additionally, if the attractiveness of an alternative value sequence is more competitive, the end-user markets will favour this to increase infrastructure capacity. These calculations are carried out for all modules as input for the next time step. The calculations for individual modules are described in detail in section 7.4.

We now turn to how choices between options are made in individual modules. Such options exist when more than one alternative for upstream supply is available (e.g. “h transp” can obtain hydrogen from both the reforming- and electrolysis modules). Choices involve the concept of competition (cf. section 6.6). Like in the real world technologies (i.e. the modules) compete against each other. The competition is based on a set of criteria:



These competition criteria together form an overall *attractiveness*. We look at the parameters of attractiveness in more detail:

Cost

Cost is a compound of how much it costs to buy and install the required technology (investment cost), how much it costs to own this capacity (fixed cost), and how much it costs to operate this capacity (operating cost) including the energy cost for the required fuel.

Quality

Quality is based on fulfilment of expectations (For example when you buy a new car you expect it to be better than the old one, cf. discussion of value, section 5.1). Quality also includes local environmental impact and global environmental impact. Environmental concerns can directly affect the attractiveness of end-user technologies (i.e. relative comparison with the other

alternatives). This means the unit of attractiveness can be odd, but it represents people's opinions and is always commensurable between alternatives.

In upstream modules the quality parameter is converted to influence cost. For example the global warming concern may lead to a CO₂ tax. This in turn will alter the cost structure according to carbon content and efficiency, resulting in less competitiveness for carbon-based energy carriers. In the case of hydrogen, demand can either be met by reforming or by electrolysis (central or local, see flip-out chart). Hydrogen delivered from the reforming module will become more expensive due to the CO₂ tax, while hydrogen from electrolysis based on renewable electric power, will become more competitive. Reforming is again subdivided into the primary energy sources natural gas and coal, with different carbon content. The same CO₂ tax will shift reforming towards natural gas with its lower carbon content.

Availability

Availability is based on capacity, utilisation of this capacity and time delay to change the capacity. It is an important part of forming the overall attractiveness. As illustrated above, the module recording inadequate availability from upstream modules will get a reduced availability 'score' and thus attractiveness in the next time step. Availability also reflects defence of existing investments through better attractiveness for available capacities (cf. section 4.3.2).

The end-user technologies represent the 'face' of several PVSes supporting what the end-user technologies require in order to provide services to the end-users. This means that the end-user technologies embody the costs, quality and availability of the overall PVSes. What the end-users see is the performance of the total PVS, reflecting the performance of the combined technologies of the involved modules.

The overall end-user demand for energy services is based on the following: the population has an expected total attractiveness reference for their needs. If the actual system attractiveness from all value sequences comes out differently from expected, demand will reflect this (e.g. if something is cheaper than expected then people tend to buy or use more, and if it feels costly people moderate themselves).

We now consider the incorporation of bounded rationality, market preferences and choice beyond costs. The most attractive technology (module) will be chosen the most. However, people have different preferences, there are limitations in availability, and information is limited and imperfect. The overall result is that not only the most attractive is chosen. A distribution of choice could for example be: #1: 94%, #2: 5% and #3: 1%. Several technologies will be chosen and more than one module may thus increase its volume. This is important as new technologies have to start somewhere, and *learning-by-doing* gets started even though the volume initially is small. The criteria for learning are applied throughout the EICOMP model in a generic manner (for detailed description of the learning mechanism, see Sterman, 2000:338).

As time develops in the model, the selection criteria are endogenously applied on the structure and its connections, and a certain scenario is played out based on the concrete assumptions of the current run. Different ‘paths’ (PVSes) are created for each energy carrier. This opens for studying all parts of an energy system’s performance with time while these parts interact dynamically. As described in section 7.8, graphs (of volume and capacity for individual modules and virtually anything else) are readily available to display the performance of the total system, an industrial sector, or particular modules. This access to information and results in a flexible and structured manner is of great help in communicating ‘what the model does’ to an audience while running the model in real

time. The main strategy of identifying reinforcing- and balancing loops and their strength over time, is also easier to communicate in real time dynamic graphics than what may be conveyed in ‘static’ reports.

In the following section we present a more detailed description of the main structure and parts of the model.

7.3 MAIN STRUCTURE AND PARTS

The present section describes the overall structure of the EICOMP model.

7.3.1 OVERVIEW OF STRUCTURE

The EICOMP model is presented in a top down manner. The overall system is visualised in Figure 7.3.1.1:

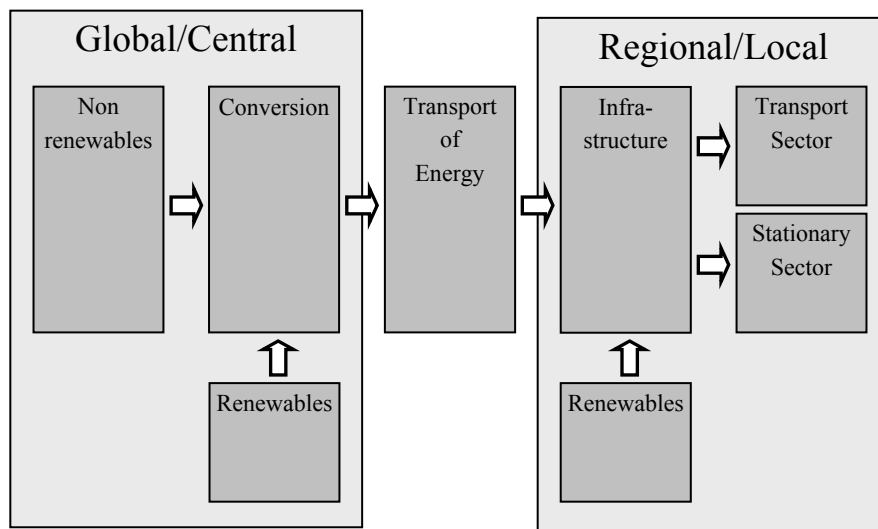


Figure 7.3.1.1 Model overview (for detailed model map with modules see flip-out chart at the back)

Information about end-user demand goes from the right, upstream to the energy sources at the left, and the information about the corresponding supply of energy to be consumed returns to the market sectors.

The total system consists of 5 main sectors:

- *Consumption sector (transport and stationary)*
- *Regional infrastructure*
- *Energy transport*
- *Conversion of resources*
- *Energy resources*

The current model has 33 modules (see overview in Appendixes 1 and 2). As described in section 7.2 each module represents a particular technology related to an energy carrier. The mode of operation for all modules is generic, with the possibility of further information detail as required. Each module is built from a set of substructures that are also generic, to describe energy flow and losses, capacity, costs and interfaces with upstream and downstream connections. All this is connected by means of positive- and negative feedback loops as described in section 6.7.

Consumption sector (transport and stationary markets)

A region is represented with an end-user need for energy services. This is divided into a transportation sector and a stationary sector (see flip-out chart). The stationary sector distinguishes between high- and low quality energy services.

Regional infrastructure

The population in a region can chose among 9 end-user technologies to meet their need for transport, and 5 end-user technologies to meet the stationary demand. Of the latter, 3 cover both appliances and

Chapter 7

heating with one type of technology, while the last two consist of a combination of electricity for appliances and liquid fuel- or natural gas for heating. The 14 end-user technologies are differentiated by the type of energy that is converted and the technology and processes applied to convert energy into useful services. The region also has a corresponding infrastructure to support these end-user technologies, described in 5 modules. Of these, 4 are distribution stations according to the type of energy carrier and required conversion/storage, and the last is a module for local electrolysis.

Energy transport

The interface between the region and the outside has 5 modules for transport of energy, distinguished by type of energy carrier (liquid fuel, natural gas, hydrogen, electricity and hot water).

Conversion

The primary energy is refined (made transferable) in 4 conversion modules: refining, reforming, central electrolysis and generation of electricity. The latter are divided into the three categories: peak, off-peak and base load, according to responsiveness. Hydro-power has the highest responsiveness (and thus attractiveness) while nuclear power is considered to be useful only for base load.

Energy sources

The region's access to energy is differentiated between *central* (out-of-region) non-renewables and renewables, and *local* renewables. The first two require a large-scale transport system (grid, pipelines or bulk transport) for transfer from where it is generated to where it is consumed. The reason for the differentiation between local- and central renewables is to capture local variation and advantages. At the same time competition from a central/global market is reflected. For example, a coal power plant will sell electric power to whoever is paying the most, as long as there is grid capacity to transfer and distribute the energy. Similarly, an oil refinery buys its crude oil in a

world market, and if demand and prices are higher outside the region, this will influence the prices in the region.

7.3.2 SPECIFICATION OF REGION AND DATA INPUT

A data-set for Germany has been used as a first region to analyse. Germany was chosen because the infrastructure is representative for many regions, and there are few anomalies (e.g. Norway with its dominating hydro power and oil industries is not representative for comparison with other regions). To choose Germany was also attractive because Germany already has a substantial initiative to change their energy system, and has started the transition towards large-scale use of renewable energy. The Germans also seem more environmentally conscious and access to data is good.

To build a data-set for another region is not a major operation (expected to require in the order of a few days work). Most of this work relates to adjusting initial capacities, opportunities/limitations for renewable energy and evaluation of any special circumstances. In the next level of hierarchy we describe how modules consist of substructures and how these interact.

7.4 SUBSTRUCTURES WITH MODE OF OPERATION

We now turn to the “interior” of the modules. As stated above there are 33 different modules in the current version of EICOMP. All modules have the same basis of substructures with identical principles of operation. We therefore limit our description to explanatory examples. All modules capture the following:

- *Energy flow with corresponding losses through a capacity*
- *Change in the capacity*
- *The costs associated with the capacity*

Where the modules have multiple upstream options for where to get the energy (e.g. to get hydrogen from reforming or electrolysis, or electricity from the central grid or from local renewable generation), or where there are several downstream ‘customers’ (e. g. natural gas is consumed both in the transport- and stationary sector), the modules have two additional substructures that function as interfaces between modules with multiple options:

- *Procurement*
- *Sales*

Through understanding how one module operates with the decision principles described in section 7.2, it is possible to understand the full model in a rational way (cf. requirement for interactive use and transparency in section 7.1). There are no principal differences in how the modules operate, but some differences exist between the five *sectors* as the level of information detail increases towards the end-user technologies and markets.

A principal description of a module’s substructures is given below. Symbols are explained in Table 7.4. Standardisation of colours and symbols are used to help readability (Myrtveit, 2003):




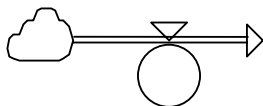



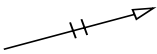
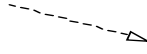
	<i>Constants</i>
	<i>Variables</i>
	<i>Levels</i>
	<i>Flow rates (can only be connected to levels). "Cloud" signifies unlimited source or sink</i>
	<i>Information flows</i>
	<i>Global parameter used several places (cross)</i>
	<i>Matrices containing several parameters (double edge)</i>
	<i>Information delay</i>
	<i>Initialisation value</i>
<i>Pink</i>	<i>Indicates flow of energy</i>
<i>Blue</i>	<i>Costs</i>
<i>Green</i>	<i>Investments</i>
<i>Orange</i>	<i>Time constants</i>
<i>Gray</i>	<i>Ratio/percentage</i>

Table 7.4 Explanation of symbols in the substructures

As an example, module # 4 (lq ICE hyb - vehicles based on liquid fuel internal combustion engines and electrical drive trains) is used. Figure 7.4.1 illustrates the substructure for energy flow and corresponding losses through a capacity:

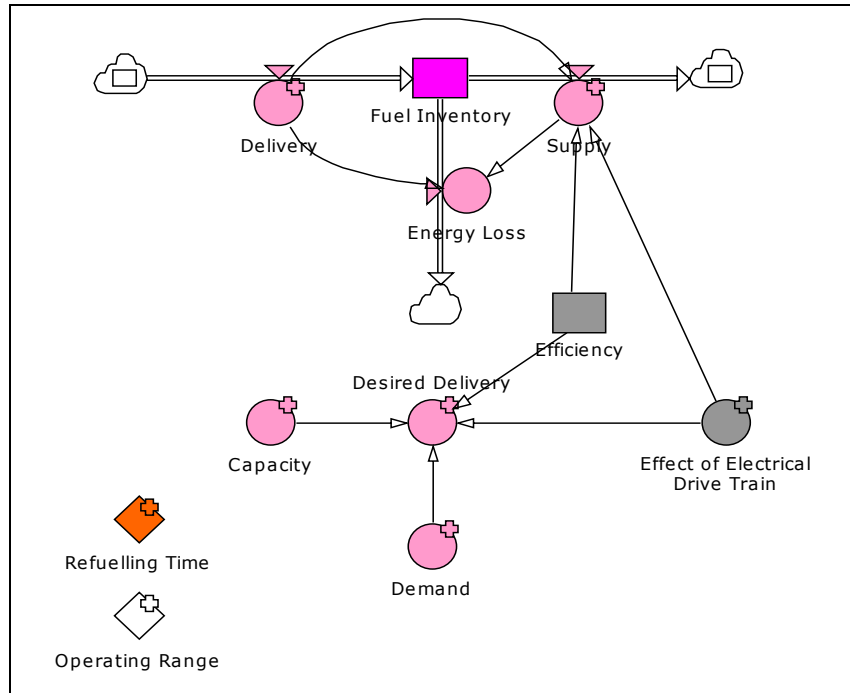


Figure 7.4.1 Substructure for energy flow

Information from the transport market (Demand) is input to calculate desired delivery from the next upstream module (in this case module #17: lq st). Desired delivery is calculated based on this signal, the module's capacity and the capacity's efficiency. Desired delivery performs similar calculations all the way up through each PVS to the final sector, energy resources. Concurrently this is calculated for all other energy carriers. Corresponding to the various capacities upstream, energy is then returned to the module (# 4 lq ICE hyb), and with increasing costs as one goes back downstream towards the end-users. The delivered energy is consumed in the process of providing the demanded transport services, and a loss is calculated based on the

module's efficiency. Since there are several modules demanding energy from the same energy carrier, there is no guarantee that the requested amount of energy will actually be delivered (the more change there is in consumption the more this is amplified).

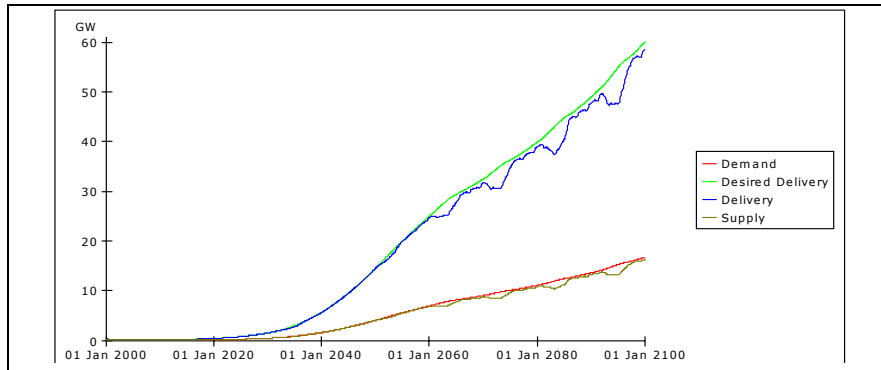


Figure 7.4.2 *Generic example of demand and actual delivery of energy*

Figure 7.4.2 illustrates the demand and delivery for the module, and the information transmitted through upstream modules to get desired delivery. The gap between the two sets of curves illustrates the loss due to low efficiency in module # 4 (80% loss in hybrid ICE end-user technology).

The demand is endogenously generated based on the module's attractiveness. The 'Effect of Electrical Drive Train' (cf. Figure 7.4.1) is also an endogenous variable, functioning as a learning curve (i.e. improves as cumulative installed capacity increases from any of the technologies with electrical drive trains (modules #: 4, 5, 7, 8, 10 and 11)).

The constants 'Refuelling Time' and 'Operating Range' are examples of features describing the end-user technology. These features feed information into the attractiveness of the value sequence (and technology it is part of). This attractiveness then feeds information

back into desired capacity. “Any” feature one may think of that will affect the value sequences’ attractiveness, could be added. However, many features have similar effects, and care should be exercised not to compromise transparency of the model.

Change in a module’s capacity is captured in the substructure in Figure 7.4.3:

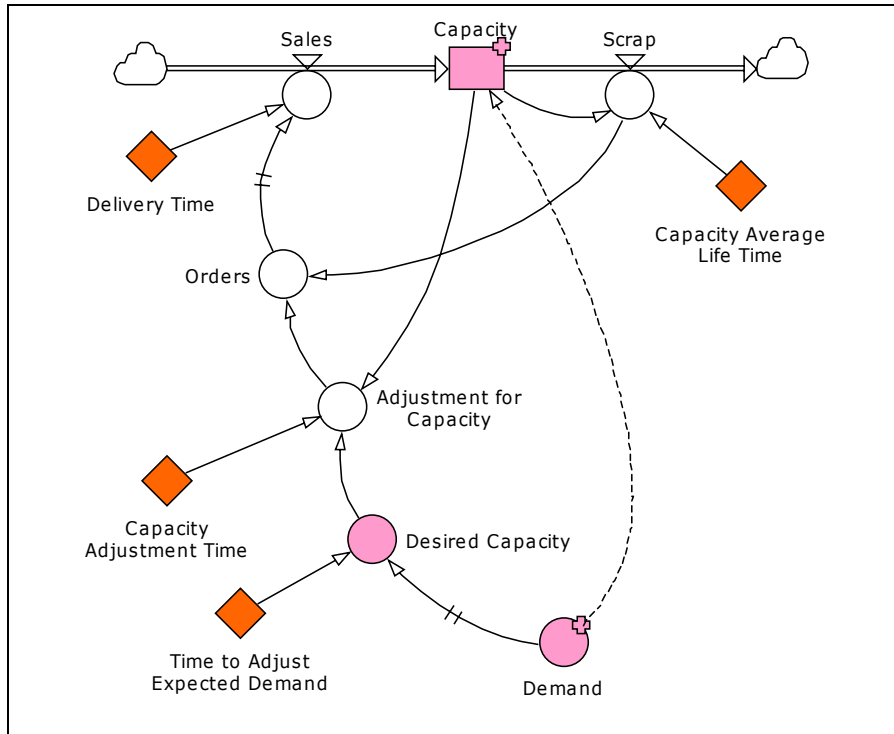


Figure 7.4.3 *Substructure for capacity*

In the module, desired capacity is calculated based on downstream demand and relative attractiveness. Based on the actual capacity and desired capacity, an adjustment for capacity is calculated. Together with information on scrapping of capacity (capital wear) orders for new capacity are calculated. Due to the time it takes to adjust capacity, delays are introduced in the system. Delays occur due to the time it takes to realise that the current capacity is insufficient, and to

the time it takes from an order is placed until capacity is installed (Sales). Further upstream in the value sequences this can be several years e.g. building a pipeline, refinery or power plant.

If demand for energy exceeds the existing capacity, then this capacity becomes a bottleneck. It will then be increased in the following time steps. At the same time, the end-users ‘remember’ the inadequate availability, and the attractiveness of the value sequence becomes reduced.

The costs associated with using or changing the capacity are described in Figure 7.4.4:

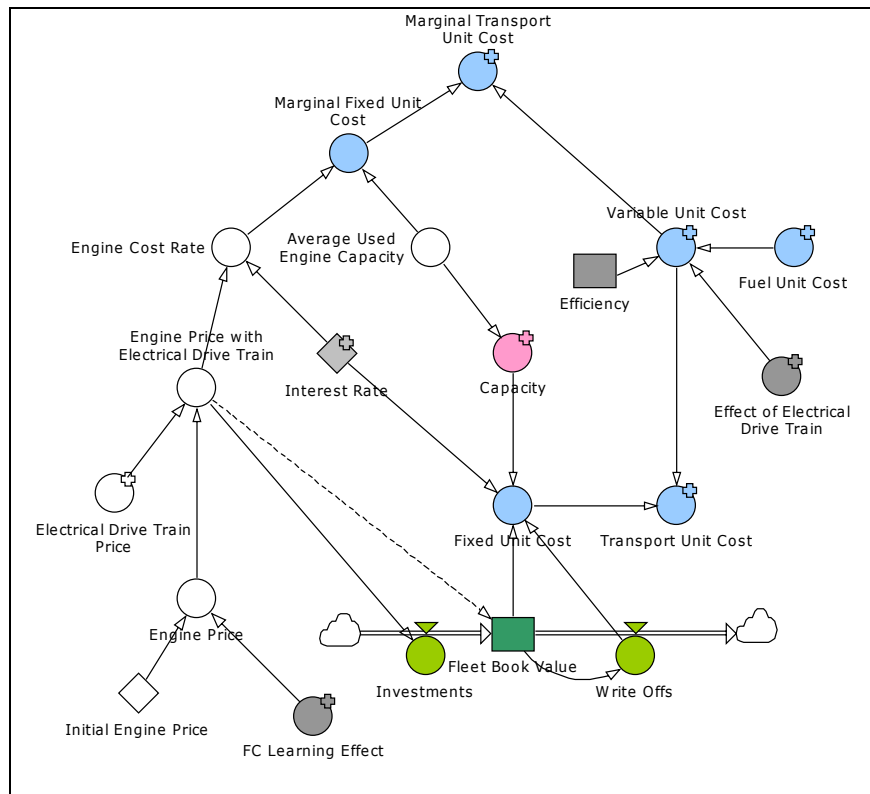


Figure 7.4.4 Substructure for costs

As an example, module # 8 (ng RFC - vehicles based on natural gas reforming fuel cell) is used. Three kinds of costs are calculated in the module, as input to the module's attractiveness. Information on fuel unit cost for this primary value sequence is received from the upstream module (in this case #18 ng st). This information is based on the total costs from upstream modules involved in this value sequence. The 'Fuel Unit Cost' together with 'Efficiency' and 'Effect of Electrical Drive Train' is used to calculate 'Variable Unit Cost'. This is the cost related to the amount of energy needed for using the capacity. The 'Variable Unit Cost' is further used with 'Fixed Unit Cost' to calculate 'Transport Unit Cost', which expresses the cost of owning and utilising the existing capacity. A final cost in the module is the 'Marginal Transport Unit Cost'. This expresses the marginal cost of increasing this module's capacity. It is different from 'Transport Unit Cost' (which is an average unit cost of the existing capacity), because technological improvement results in reduced cost for the next installed capacity compared to the average cost of the existing capacity. Investment in the capacity is calculated from the 'Engine Price' and 'Electrical Drive Train Price' and is summed up in the 'Fleet Book Value' level. 'Engine Price' starts with an 'Initial Engine Price' of the fuel cell, but is reduced as the 'FC Learning Effect' comes into play.

As already stated, modules with the option to obtain energy from several upstream modules, or to distribute to several downstream modules, require two additional substructures (Procurement and Sales). These are described in Figures 7.4.5 and 7.4.6:

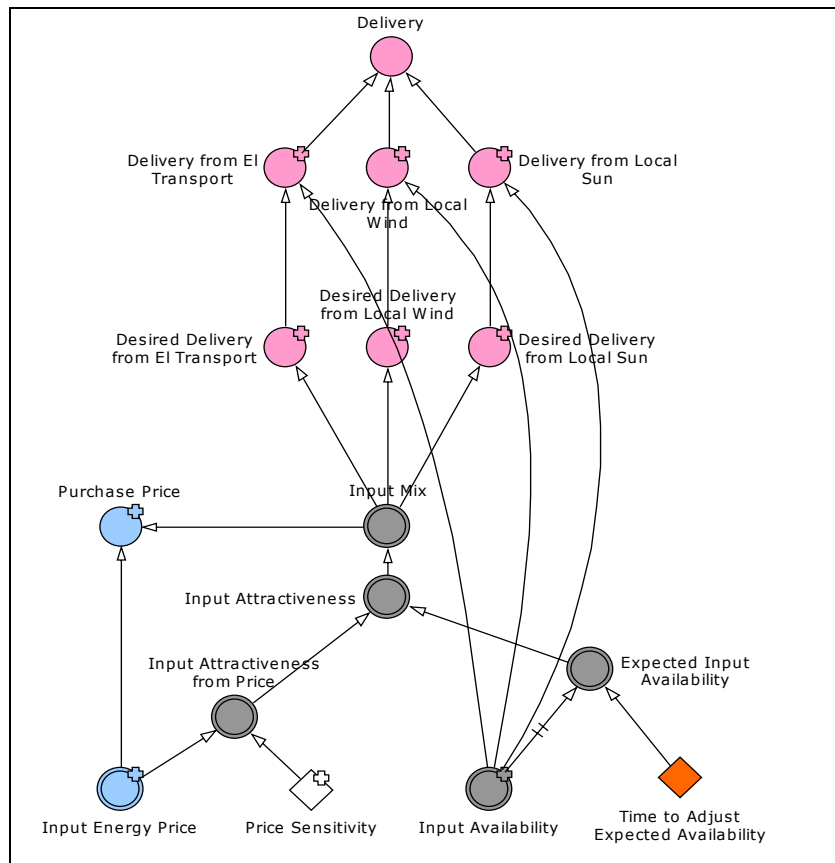


Figure 7.4.5 Substructure for *upstream procurement*

This substructure (module #21 electrolysis (local) is used as an example) combines expected input price and expected availability from the upstream modules, to calculate desired delivery of energy. The actual delivery is then calculated from the actual availability.

Similarly, to distribute to downstream modules, a generic ‘sales’ module is illustrated in Figure 7.4.6:

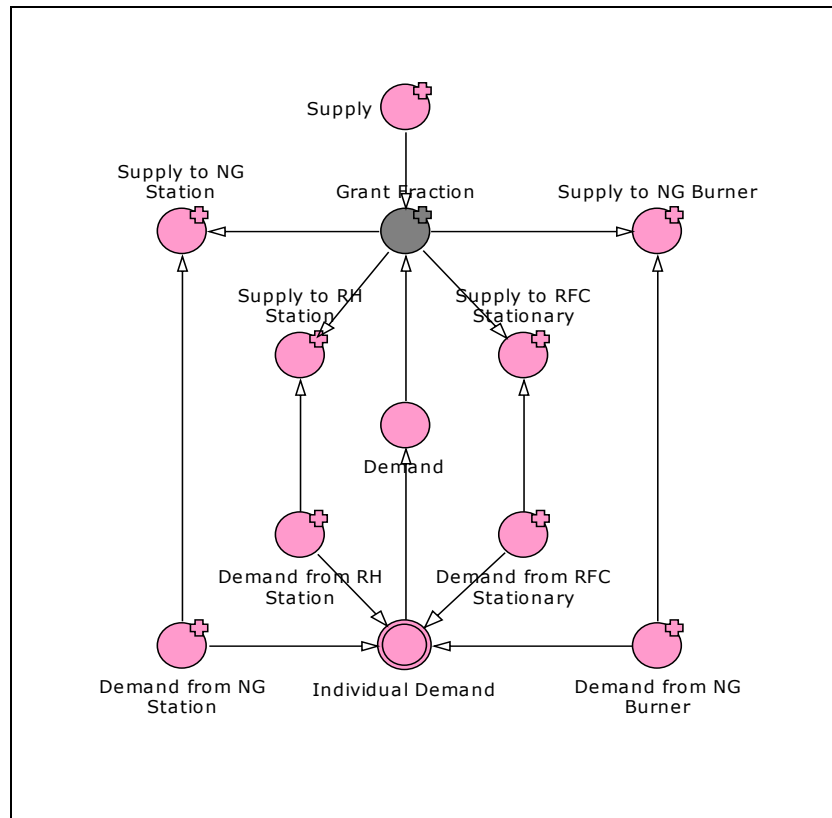


Figure 7.4.6 Substructure for *downstream sales*

This substructure (module #23 (ng transp) is used as an example) illustrates the interface between the module providing natural gas to the region and downstream ‘customers’. The individual demands are added to give a total demand that is used to send a signal further upstream to obtain desired total delivery. The actual delivery is granted proportionally according to the individual downstream desired demands.

Details on the remaining modules and substructures are best inspected interactively on the computer. This allows easy excursions to neighbouring structures and modules, which is practical as a reference to ‘where you are’ when using (and examining) this multidimensional model. It further facilitates display of constants and

equations (with units and connections) and corresponding references and explanations, where this is deemed necessary. The overall transparency and readability of the model is greatly assisted in this way (cf. section 7.8.1).

We now turn to the description of the model boundaries.

7.5 MODEL BOUNDARIES AND ASSUMPTIONS

In this section we describe the boundaries and underlying assumptions in the EICOMP model. The boundary is the interface between the model proper and its surroundings, as illustrated in Figure 7.1.1. The boundaries are implemented in the model's overall system equation and its input and output domains. As described in section 6.1 the system equation reflect the inner workings of the model. The input domain is the description of the outside environment by way of exogenous variables, while the output domain is given by the opportunities for inspection of endogenous variables. In the EICOMP model these opportunities are pervasive.

Sterman (2000:86) lists the following issues related to boundary specification:

1. **Theme selection:** *What is the problem? Why is it a problem?*
2. **Key variables:** *What are the key variables and concepts we must consider?*
3. **Time horizon:** *How far in the future should we consider? How far back in the past lie the roots of the problem?*
4. **Dynamic problem definition (reference modes):** *What is the historical behaviour of key concepts and variables?*

The *theme selection* is the situation of ever increasing use of non-renewable fossil fuels coupled with problems of supply and environmental impact. The broad background to this is described in

Chapter 1. The concrete issue behind the EICOMP is a better understanding of change towards sustainable practices.

The *key variables* related to the overall problem are considered to be the various technologies (capacities) with corresponding networks (value sequences), and the flow of energy (and related losses) through existing and prospective technologies, in order to ensure adequate energy services.

The *time horizon* is seen to be the next few decades. In the development of EICOMP the timeframe 2000 – 2100 has generally been used, although the key focus with respect to change is seen to be the next 10 – 20 years. A new era of decision making has to begin now in order to make a better tomorrow. The situation in year 2000 is taken as a general starting point with reported input data for the modelling effort. History is generally specified through the reference needs and end-user preferences (cf. Table 7.5.1).

The *dynamic problem* is well illustrated in Figure 1.1 – the overall consumption of energy has had exponential growth for the last two centuries. Figure 1.1 relates to the US. As described in section 2.6 there is now rapid increase in consumption of fossil fuels, particularly in South-East Asia.

Sterman (2000:97-98) emphasises the usefulness of model boundary charts to communicate the boundaries of a model and underlying assumptions. Table 7.5.1 lists the main variables used in the current version of EICOMP (cf. section 6.7.3 for discussion on endogenous and exogenous variables):

Endogenous	Exogenous	Excluded
Consumption of energy	<i>All constants in the model</i>	Inventory of energy
Investments	Initial end-user capacities	Global warming
Costs	Initial end-user costs	
Energy imports	Initial experience	
Energy production	Population	
Energy demand	Primary energy costs	
Energy supply	Reference transport need	
Attractiveness	Reference energy need for appliances	
Availability	Reference energy need for heat	
Technological change	End-user preferences	
Capacities	Interest rates	
Capacity utilisation	Energy policies	
Energy loss in primary value sequences	Reference capacity efficiencies	
Global CO ₂ emissions	Carbon content per energy unit	
Regional CO ₂ emissions	Nuclear phase out start time	
	Nuclear phase out duration	

Table 7.5.1 *Model boundary chart*

The numeric values on exogenous variables will vary with the chosen region (cf. section 7.1). Exogenous variables may be specified in the model in three different ways:

Chapter 7

- *As constants*
- *As system equations (heuristic- or mathematical functions)*
- *As time-series constructed by interactive graphs*

Constants

Specification of constants is done interactively: in practice by “opening” (clicking on) the pertinent symbol in the substructure map, and entering a unit and a numerical value. The interactive feature is important: rather than handling constants in separate tables, the constants are integrated with the graphic structure for the corresponding technology. In this way values are connected to the actual technology, giving the structure a realistic and transparent dimension.

System equations

Instead of a specifying a constant, a function $f(t)$ may be set up. This may be a system equation or a mathematical function based on heuristics. A value is subsequently calculated for each time-step, and used as input to the endogenous modelling. Common functions are the linear and exponential variation with time, described in section 6.2.1 (eq. 6.2.1.3 and 6.2.1.6). The function may be changed at any time between simulation runs. The only requirement besides being a sensible relationship, is that it is mathematically defined through the time interval.

Time-series constructed by interactive graphs

This is a particular feature with the “Powersim” software used to build the EICOMP model and user interface. In a control panel, time-series are specified by literally drawing it on the computer screen. For each time-step in a simulation run, the corresponding value is read from the curve. Again the operation is interactive: a curve may be drawn and redrawn at any time.

In sum, the specification of exogenous variables is interactive and highly flexible. This is helpful when studying the effects these assumptions have on the model behaviour.

The underlying assumptions in the EICOMP model have their basis in the thinking and concepts as described in the previous sections. Correspondence with reality (cf. section 7.1) and selection according to attractiveness (cf. section 7.2) are central tools. Some more concrete assumptions and constraints require further discussion:

- *In Figure 7.4.6 the mechanism of distributing energy to downstream modules was described. The actual delivery of energy is granted proportionally according to the individual downstream desired demands. There is no differentiation with respect to buyer power among different downstream modules.*
- *Costs of non-renewable primary energy are treated as exogenous variables. They may, however, be specified interactively through the flexible time-series. The rationale behind this simplification is based on the following: the purpose of the present work is not to model/predict future energy prices. It is rather to study the effect of such changes. A model of future energy prices would soon become a complicated modelling effort in its own, and the focus would become the economical issues rather than technological change (cf. section 6.8). Costs of renewable energy are calculated endogenously. The chosen solution allows the investigation of different price-developments in the future.*
- *For the special case of nuclear power, with its strong political debate, an exogenous mechanism forcing a politically decided phase out is available in the model.*

- *For renewable energy (both central and local) there are constraints to how much capacity that can be installed according the region being modelled. This covers available hydropower potential, area for wind turbines and solar based technologies.*

Documenting a model is hard work, but also rewarding as new insight is developed as each ‘stone is turned’. It is important in several aspects, and poses a particular challenge in documenting the dynamics of the current model. This is discussed further in the next section.

7.6 THE CHALLENGE OF DOCUMENTING THE EICOMP MODEL

Our objective in this section is to discuss the documentation of the EICOMP model. Such documentation raises two particular issues. The first is that time-reflecting models necessarily refer to, or reflect real time. Any “starting-line” is therefore arbitrary. A background referring to the millennium will change as the model is updated in the following years. The other issue of documentation is that modern computer technology allows practical handling of extensive systems that are best documented interactively within the systems themselves. The main requirement then, is that those who need the documentation must have practical access to the system. Again, at least for models like the EICOMP, this is readily available with modern technology. It is a practical necessity to gain an overview of the dynamic issues and the hierarchical structures.

Model documentation could serve several different purposes:

- *Documentation to make the model available to those who would gain from using it*
- *Documentation to help new instances use the model*

➤ *Documentation as a thesis dissertation*

The first kind needs to focus on what the model achieves, with a general introduction to operation. Presentations and descriptive papers are typical documentation of this kind. The second kind is in the line of user manuals and interactive help-functions built into the model. The focus of the third kind of documentation is to give an adequate description of the background and purpose of the model, and how it operates, without drowning in detail. The present thesis attempts to document primarily background and purpose, the principles of operation, and the relevance between the purpose and the output from the model. The details of operation, as well as the data that describe the starting line for modelling, is made available as interactive documentation within the model itself, in the hope that most of the detail will be easier to get at in this way. Detailed information communicates easier/better when the reference structure is available.

In the current version of the EICOMP model the amount of detail may be summed as follows:

- *5 sectors*
- *33 modules*
- *3-5 substructures in each module*
- *1467 variables defined in the substructures*
- *3435 values involved with these variables in relation to where they are based and their connections*

While extensions to the model in the form of new modules and structure require active programming and specialist knowledge, the accessibility to the modules (with all substructures and variables) is open and straightforward.

7.7 THE CHALLENGE OF DATA QUALITY

This section describes the basis of data that make up the “starting-line” for the current modelling effort. It refers to the region being modelled, Germany, as introduced in section 7.3.

Input data are vital to any modelling effort. Both in terms of quality and the amount of data required. Data quality is important, but difficult to assure for a number of reasons as discussed further below. Data quantity however, is not difficult. In fact, most modellers agree, often in hindsight after a project, that too much time and resources were spent on finding data rather than discussing carefully what data they would really need to carry out the work. The usual story is that 90% (or at least a large part) of the project’s budget and time is spent on collection of data, with the result that the actual modelling work and analysis do not receive enough attention. Hence, the project’s overall quality may become substantially reduced, and delays and budget overrun usually comes with it (Sterman, 2000).

The development of EICOMP has been an ongoing process for more than two years. Time has mostly been spent on developing operating principles and model structure. By refining the model structure several times and putting off the data gathering process to the final part of the project, the required amount of input data has been substantially reduced.

Finding good quality data, however, has proved to be a demanding task. A number of reasons are listed below, but the challenge remains. The lack of good quality data makes it difficult to communicate and compare modelling results, and is an important part of the difficulty when discussing and comparing different hydrogen related technologies and strategies.

There are a number of challenges involved in finding good quality data:

1. *A massive amount of data is published and circulated. Much is of poor quality (several factors are listed below) or without sufficient references*
2. *A majority of the reported data is circulated by and among academics. However, few academic institutions actually operate power plants or large scale industrial equipment. What are the original source and circumstances?*
3. *While private firms operate power plants and industrial equipment, operating data are often considered 'strategic secrets'. There is often little enthusiasm for sharing this knowledge. Especially, data on costs tend to suffer under this. (To be fair it is not always easy to measure/calculate relevant data)*
4. *Assumptions behind measurements and estimates are often not reported, while such assumptions may heavily influence the outcome*
5. *What's included in the reported data is often not specified. A typical example relates to efficiency, where auxiliary equipment is left out or not specified*
6. *Economic data are seldom adjusted for inflation, or are listed without reference to what year the measurement was made (cf. Consumer Price Index, 2003)*
7. *Aggregated, average or marginal data is often not differentiated or specified*
8. *What's included in future estimates is often not described. What are the assumptions behind estimates of future prices or effects of learning?*
9. *Data may sometimes be published with a hidden agenda to promote a particular view*

The present study has data input from:

- The International Institute for Applied Systems Analysis (IIASA), CO2DB database (CO2DB, 2003)
http://www.iiasa.ac.at/Research/ECS/docs/data_index.html
- National Renewable Energy Laboratory (NREL, 2003):
http://www.nrel.gov/analysis/power_databook/
- B. Sørensen (2001), Roskilde University, report (in Danish: Hydrogen as an energy carrier – scenarios for future use of hydrogen in the Danish energy system)
<http://mmf.ruc.dk/energy/>
- EU Energy and Transport in Figures, Statistical pocketbook 2002:
http://europa.eu.int/comm/energy_transport/etif/index.html

Sensitivity analysis is helpful when using data and evaluating their quality and importance. However, it soon becomes labour-intensive and makes communication of modelling demanding. Through discussions and interactive workshops with stakeholders and relevant people, input data may be improved. In this process, the value lies in the learning and improved mental models, and not in producing ‘specific results’ to argue a particular view.

7.8 MODEL OUTPUT

This section aims to describe the output from the model. It is closely related to the interface between the model and its surroundings as illustrated in Figure 7.1.1, and is essentially a process of communication. As in any communication it takes a sender and a receiver. There must be someone to listen, and it takes a prepared mind for the message to come across. In the following we describe

first the process of sending, i.e. the procedure of output, and subsequently turn to the prepared mind and the ability to listen.

7.8.1 PROCEDURE OF OUTPUT

Output requires specification of what the modeller is interested in. This is done in a hierarchical manner through several levels. The first level consists of a screen display corresponding to the flip-out chart, showing the various modules and the flow of information between them. Clicking on a module opens it, to provide access to the substructures belonging to this module. Selecting a substructure in turn opens the individual variables embedded in an information flow structure (Key variables are described in section 7.5).

Output from variables is obtained by clicking on their symbol (Their value and unit is also automatically displayed on the screen). Endogenous variables (and the exogenous variables defined as time-series) vary through a modelling “run”, and take the form of *time paths*. These variables are inspected by means of a highly flexible tool for creating *graphs*. A graph may either show the time path of one or more variables through a single run, or display the time paths of a particular variable through a series of runs as exogenous variables or assumptions are varied. An example of a graph is shown in Figure 7.4.2. Graphs may be set to display continuously as the model is running (which is educational and entertaining) or it may be set to display as the run finishes, to increase speed of computation.

A modelling run is started simply by clicking on the “run” button. Typical time of a run in the first mode is about 10 sec. on a standard computer.

7.8.2 THE PREPARED MIND

When communicating there are (at least) two central rules to describe the success of the outcome: that the true/actual message is conveyed and that there are no misunderstandings (e.g. causality versus

correlation). Encouragement, motivation and enthusiasm may be of crucial importance to make this come true. The “prepared mind” is the mind carefully framing questions, on the basis of a clearly stated purpose. The purpose of the modelling work in the present thesis is described above. This purpose is not to simply model a system, nor to develop a balance sheet of parameters, but the system is a learning tool to better understand the feedback structures of our societies’ energy services. Using the model in this way is an ongoing and continuous learning process. It is certainly not a “spectator sport”. The learning takes place through interactive participation and use of the model, and through careful observation of the dynamics as the game/model plays.

This thesis focuses in particular on the modules involved with hydrogen. The EICOMP model, however, models the overall flow in all major energy carriers and development in all involved capacities (and related technologies). It may be used to study, say, the development of liquid fuel ICE (module lq ICE at the top right of the flip-out chart) under various assumptions of oil price or concerns with greenhouse gas development. Similarly, the EICOMP model may be used to study the future development of natural gas transport (ng transp).

The EICOMP model is developed with Powersim Studio 2003, academic version. All documentation (including the model with all parameters and initial conditions as it currently exists) is available. Please send an e-mail to: kaare.gether@ntnu.no

The Powersim software is available at: <http://www.powersim.com/>

We now turn to the final chapter where central aspects to the present work and the EICOMP model are discussed.

Discussion and conclusions

“William James used to preach the “will to believe”. For my part, I should wish to preach the “will to doubt”. ... What is wanted is not the will to believe, but the wish to find out, which is the exact opposite.”

Bertrand Russell

The background of the thesis is essentially the two principal questions:

- *What is the likely influence on the climate, environment and health of people over the next hundred years or so if current practice continues?*
- *How do we institute change towards sustainability?*

The first question is the field of resource- and climate modelling. The present work originates in, and deals with, the last question.

Looking ahead is difficult. As related in Chapter 5, there is both quantitative and qualitative uncertainty involved. The latter, inherently nonlinear, is the most difficult to deal with. This is further complicated by path dependence and delay. Path dependence is pervasive in society and a main reason for inertia and barriers to learning and change. History shows surprising examples of how long it takes before even highly motivated change takes place. Thus, Sterman (2000:19-20) relates how it took two and a half centuries to get rid of the scourge of scurvy in the British Navy and merchant

fleets - and all that was required was a supply of lemons. Learning, in this case, was almost unbelievably slow despite the life and death importance of a solution, and despite clear evidence. We might think of this as the quaint old days, yet:

“The delays in learning for many pressing problems remain woefully long. In most settings we lack the ability to run experiments, and the delays between interventions and outcomes are much longer. As the rate of change accelerates throughout society, learning remains slow, uneven and inadequate.”

J. D. Sterman, 2000:20

Ours is a world more tangled and multifaceted than the problem of supplying seafarers with an adequate diet. There is a complex picture where several factors play a role, and multiple restraining- and driving forces are active at the same time and interact dynamically. The inherent nonlinearity aggravates the complexity. Modelling is required to navigate in this maze.

Such modelling should reflect reality (cf. Table 1.1) and dynamic behaviour, and interact with the environment where decisions are made. Decision making based on dynamic modelling improves the results, and the interactive approach provides local information into the modelling effort. Furthermore, this interactive approach must be transparent enough to allow insight into the overall dynamic picture of ongoing change, and provide information on central mechanisms. Through the approach of identifying the strength of positive- and negative feedback loops in a dynamic energy infrastructure, the EICOMP model achieves this, as described in Chapter 7.

The current chapter is concerned with the following themes:

- *Is the overall approach suitable to its purpose?*

- *Critical aspects to the modelling approach*
- *Correspondence between the EICOMP model and the real world*
- *Strategy and how to achieve desired change*

8.1 APPROACH AND RELEVANCE TO PURPOSE

The focus in this section is twofold: The suitability of the overall approach and whether or not the thesis, including the EICOMP model (assuming it works as intended), actually provides useful information towards achieving change. The aim of the modelling effort is to seek out what will be important, and to rule out what is not.

A main departure from the energy models reviewed in section 6.8 is the focus on change. As described in Chapter 4 change is often met by resistance. However, it also opens new avenues to creation of value, and new possibilities and opportunities arise for firms to take advantage of this.

Change inherently involves a dynamic aspect and a need to describe elements interacting in complex structures. This has come to influence the present work and its results. At the start of the work, this focus on change was by no means clear; it has emerged as the work proceeded and has led to the development described in the previous chapters. As illustrated in Figure 7.1.1, the work is iterative and embedded in the dynamic behaviour of the real world.

8.1.1 CRITIQUE OF APPROACH

In this section we discuss critical aspects to the current approach. Does the PhD work provide the information actually required to induce and manage practical change? Response, thus far, is encouraging and a number of aspects are further discussed below. The ultimate goal would be to tell where money is to be made (and not made) with respect to a hydrogen-based sustainable society. One

of the results from the present work, the EICOMP model, may in this regard be seen as a toolbox to provide information.

A number of approaches are possible to most modelling efforts. Furthermore, results may vary widely depending on how a particular model is constructed, which underlying assumptions are made and how these are applied. To illustrate, Sterman (2000) provides an informative example relating approach and purpose:

“During the congressional debate over the North American Free Trade Agreement (NAFTA) in the early 1990s, proponents of NAFTA argued that free trade would boost the incomes and standard of living of all trading partners. The traditional economic theory of comparative advantage suggests trade benefits both parties because each can produce more of what it is best at and trade for the rest, instead of producing all the goods and services it consumes with lower efficiency. NAFTA opponents, however, argued that companies would divert capital investment from the US, with its high wages and comparatively strict environmental regulations, to Mexico, destroying US jobs and harming the environment. They argued that capital mobility would lead to a ‘race to the bottom’ that would drag down wages, safety standards, and environmental quality as different countries competed for factories and jobs.

Dozens of econometric models were used to predict the effects of NAFTA on the health of the US, Canadian, and Mexican economies and their results were used to buttress the arguments in the debate. The vast majority of these models suggested NAFTA would be a boon to all three economies, with little or no short-run costs. The models were used to argue that concerns over capital flight from the US were misplaced.

The Congressional Budget Office (1993) examined 19 models used to make forecasts of the impact of NAFTA. Of the 19 models, 14 did not consider investment flows at all. Of the five that did, four assumed there would be no impact on US investment. Implicitly, they assumed all NAFTA-induced investment in Mexico would come from nations other than the US. The one study that treated investment endogenously concluded that NAFTA would transfer \$2.5 billion per year in investment from US to Mexico, resulting in a loss of about 375,000 US jobs over 5 years.

Following standard practice since the days of Smith and Ricardo, most NAFTA modelers assumed that capital and labour were fixed and immobile. Goods flowed between nations in trade but capital and labour could not. The theory of comparative advantage works under these conditions. But when capital is mobile, comparative advantage operates because businesses will locate in the region with the greatest absolute advantage. Assuming immobile capital eliminated important feedbacks from the model boundary, feedback that changed the outcome of the policy analysis.”

J. D. Sterman, 2000:862

The example illuminates two central points: the current thesis does not follow a “standard practice”, but is based on the concepts outlined in Table 1.1 (e.g. see multiple equilibria described in section 8.3). Secondly, a main reason to encourage the interactive approach is to make assumptions available for review and prevent models being used to argue particular positions.

One of the most critical issues is to what extent models reflect reality. There are two principal sides to this, one related to our perception of the real world, and the other to the behaviour of the model itself. Models *have to* reflect reality. If they do not, then whatever conclusions are drawn from the modelling have no foundation. An

example of this kind of situation is the criticism of the classic/neo-classic economic model (cf. section 8.1.3). The major criteria for our perception of reality were set out in Chapter 1, and how the EICOMP model seeks to achieve this was explained in Chapter 7. Chapters 1-5 seek to give a broad background of our understanding of reality, and the interface between modelling and reality is treated in Chapter 6.

The thesis introduces a new approach. There are several reasons why this new approach has been developed:

- *Unravelling of problems through modelling*
- *Different focus and purpose*
- *Transparency, interactivity and flexibility*
- *Reflecting new technology and opportunities*
- *The value in multiple approaches*
- *The difficulty of documenting models*
- *Improved opportunities in modelling*

Unravelling of problems through modelling

When working with complex systems, one of the main requirements is some means for systematising information. Knowledge needs to be placed into categories or classes, and these in turn must be placed in relation to each other. This is precisely what *construction* of models achieves. Construction of models (mentally or in some formal manner) is a natural way of systematic intellectual work. Hence, in order to understand something, modelling is essential. It would appear that this is a main reason for the overwhelming number of models already around.

Different focus and purpose

Rather than studying the effects from energy use on economies, or environmental/climate consequences, the current work focuses on change and dynamic interaction between end-user technologies and required energy infrastructures.

Transparency, interactivity and flexibility

It is important to get a clear understanding of a model's assumptions and their influence on the results and to avoid 'black-box' modelling. This is approached through the structured and open source code and graphics (cf. section 7.4). Also, transparency, interactivity and flexibility are vital to make the modelling locally available for decision-makers. The aim of transparency is further discussed in section 8.1.2.

Reflecting new technology

As new technology is developed and new opportunities arise, it should be reflected in the modelling. In EICOMP it has been a particular focus to include hydrogen related technology and infrastructure, to study the interaction/competition between natural gas, hydrogen and electricity.

The value in multiple approaches

Huntington and Weyant (2002) underline that the 16 models involved in their review (cf. section 6.8) supported each other in illuminating the overall problem of climate change, or the relationship between energy and national economies. Of the various models examined in the context of the present work, the one closest to the EICOMP model is the TIMER model developed in the Netherlands (cf. section 6.8). The TIMER operates as part of a larger system (the IMAGE system), and seeks to model environmental consequences from use of energy in a world perspective. In contrast the EICOMP model focuses on understanding factors underlying change and how to achieve this. It seems clear that the TIMER and the EICOMP models will assist each other in helping to better understand the overall challenge.

The difficulty of documenting models

In addition to documenting a model to build confidence, as discussed in section 7.6, documentation needs to make the modelling effort available to others. This requires that those who might be interested get to know about it. Furthermore, the purpose and nature of the model must be sufficiently described. These issues would be eased by means of systematic schemes for classification of models. However, as such classification is not in common use, the effort of learning about other work may be comparable to, or even exceed the effort of developing a new model. Not least is the problem of finding out what existing models take into account, and what are left out.

Improved opportunities in modelling

Theory in related fields such as nonlinear dynamics, path dependence, irreversibility, disequilibrium-, economic- and innovation theory is advancing. As computers and software become more available and powerful the new/improved knowledge can to a larger extent be combined and studied together, encouraging new development (cf. section 6.9).

These reasons together form a strong motive for developing new models rather than extending existing ones. A last issue is that existing models are usually owned by someone, and access to a model may not be free.

8.1.2 INFORMATION RELATED TO CHANGE

Fundamental to the present work is that significant change requires *firms* to deliver new and different products and services to achieve sustainability. These firms have to earn money to remain in business. In addition, required organisation, technology and standards must be in place. Given a sound economic basis, then knowledge, established technology and organisation are mostly a question of time and supply of capital. However, new solutions raise new requirements and questions. To illustrate (cf. Figure 1.2 and section 4.3.3): based on present estimates in Norway, a natural gas power plant might take 5

years of planning and construction, and will have a service life of some 30-40 years before corrosion- and safety issues require decommissioning. A wind farm providing the same energy supply might require about the same time for construction, and would cost up to three times as much with present technology (Tande, 1999). However, it would have *no costs* for primary energy and require substantially less replacement of the original investment to significantly extend lifetime. Once discounted, sustainable energy from wind (or hydropower) only have operational and maintenance costs. As these tend to be reasonable, such sustainable energy becomes an almost free public good in the long run. What is the proper economic goal: medium-term income for a firm while natural gas is reasonably cheap, or long-term, low-cost sustainable energy, but at a higher entrance fee?

Information provided by EICOMP

The EICOMP model essentially provides information on a set of modules that each represents distinct technologies in the existing energy infrastructure network and end-user markets. The primary information from the model is how attractiveness, capacity, and requested and delivered volume of energy vary with time for each module. The information about each individual technology depends on the interplay between all the technologies in the overall infrastructure network, under the influence of the exogenous specifications. Associated with this primary information is information on existing- and required investment, operating costs, conversion efficiency and time required for change in capacity in any module.

As described in sections 7.1 and 7.2 the interaction between modules/ technologies is described by means of competition, where this competition is based on a concept of attractiveness rather than on traditional costs. This attractiveness is a compound of costs, quality and availability. The “unit of attractiveness” may vary, but is

everywhere congruent within competition for the same service. Comparison in this way is commensurable throughout the EICOMP model. This has been a challenge particularly for the market modules for stationary energy services, since appliances and heating may be provided by different mixes of technologies. The model reflects high- and low quality energy services and generation of electric power for peak-, off-peak and base loads. It has proved feasible in the EICOMP model to organise information (“hide away detail”) in structured hierarchies and to establish commensurable offerings in markets.

Learning is reflected where meaningful (with improvement according to increased cumulative capacity volume in modules) in the form of reduction in both marginal- and unit costs. In turn this leads to increased attractiveness.

The user of EICOMP may study any technology in a detailed manner, as the dynamics of the model plays out. In addition, the effect of changes in exogenous variables, like prices of primary energy, interest rates, life times of equipment or conversion efficiencies etc. may be studied.

Exogenous variables like costs of primary energy may be obtained from other models to reflect particular scenarios of future development. This is incorporated in the form of time series as input to the EICOMP model. Our overall conclusion at this point is that the information provided by the EICOMP model on timing, utilisation of capacity, competitiveness and demand for capital is relevant to fundamental aspects of change.

The goal of transparency

Transparency has two objectives: to bring modelling close to decision makers (without requirements for special knowledge), and to create confidence in the results. A number of factors contribute:

- *Mental correspondence between model and reality*
- *Generic principle of operation to ease learning*
- *Generic implementation of structure to ease learning*
- *Structured overview of information*
- *Recognising technology in the model as in the real world*
- *Subdivision of the model into parts with clear purposes (cf. citation from Dawkins, section 6.2.2)*
- *Easy access to substructures for inspection and testing of sensitivity to change*

Considerable efforts have gone into achieving transparency. The most demanding issues are the three first points above. Mental correspondence with reality and generic construction is fundamental to the multi- and interdisciplinary character of the current study.

A common ground for discourse

Similarly to transparency, a need for a common ground for discourse stems from the multi- and interdisciplinary nature of sustainability. The requirement for good communication may be seen through the underlying structure of value sequences. A sustainable economy will consist of a set of value sequences that provide end-users with what they need, with the underlying activities carried out in a sustainable fashion. If a single activity fails there will be no end-product and no satisfaction of end-user needs. Nor will there be any upstream flow of money to ensure a sound economy for participating firms. This means that all the individual activities must do their job, all the time. In ongoing change it means that a large number of activities are simultaneously shifting in their characteristics, yet are required to uphold their function. In this complex picture we clearly need a common ground to ensure good communication between relevant stakeholders. The EICOMP model assists this communication by giving a systematic overview, structuring concepts and in a clear way organising the interplay between activities of the complex network. If

change is to become significant in the overall picture then functional bottlenecks must be identified and dealt with.

Optimisation through markets

The EICOMP model is built and used as a simulation model. Yet the simulation of competition in markets leads the model towards “most attractive” pathways according to the running attractiveness mechanism. This is optimisation in the same sense as Adam Smith’s “invisible hand”. In view of situations with increasing returns (cf. section 8.1.3) this optimisation is “local” (in optimisation parlance), and will lock in to the “nearest” optimum. The main point to be made here is that the principle of operation in EICOMP leads to an inherent search for attractiveness; it is not blind simulation without direction. This influences the character of the information obtained from the modelling, in particular with respect to path dependence that impede change away from locked-in optima (e.g. the regime of ICE-technology).

8.1.3 POSITIVE FEEDBACK AND INCREASING RETURNS

In this section we discuss in more detail some of the consequences of increasing returns (i.e. positive feedback, cf. section 4.2) in the EICOMP model (cf. sections 7.2, 7.4 and 7.5). To recap, these situations of increasing returns result from network effects, learning, increased efficiency, and lowered costs with increasing volumes of the various modules. The modern understanding of the phenomenon of increasing returns is discussed e.g. by Krugman (1979) and by Moxnes (1992), and described by Arthur (2000) as follows:

“I would update Marshall’s insight by observing that the parts of the economy that are resource-based (agriculture, bulk-goods production, mining) are still for the most part subject to diminishing returns... ...The parts of the economy that are knowledge-based, on the other hand, are largely subject to increasing returns. Products such as computers,

pharmaceuticals, missiles, aircraft, automobiles, software, telecommunications equipment, or fiber optics are complicated to design and manufacture. They require large initial investments in research, development, and tooling, but once sales begin, incremental production is relatively cheap... ..Increased production brings additional benefits: producing more units means gaining more experience in the manufacturing process and achieving greater understanding of how to produce additional products even more cheaply. Moreover, experience gained with one product or technology can make it easier to produce new products incorporating similar or related technologies.”

W. B. Arthur, 2000:3

Conventional economic theory (e.g. the classic/neo-classic model and the endogenous model) has as one of its basic outsets an equilibrium between supply and demand in markets. This equilibrium results from the law of diminishing returns (Arthur, 2000:6-7) and is the basis for formation of price according to the principles of marginality. It is also behind the argument that free markets will find the “best possible” (i.e. global optimum), and is a central argument behind ‘laissez-faire’ economic policies. Arthur (ibid.) argues the opposite:

“Positive-feedback economics, on the other hand, finds its parallels in modern physics. Ferromagnetic materials, spin glasses, solid-state lasers and other physical systems that consist of mutually reinforcing elements show the same properties as the economic examples I have given. They “phase lock” into one of many possible configurations; small perturbations at critical times influence which outcome is selected, and the chosen outcome may have higher energy (that is, be less favorable) than other possible end states....

...In this new view, initially identical economies with significant increasing returns sectors do not necessarily select the same paths. Instead they eventually diverge. To the extent that small events determining the overall path always remain beneath the resolution of the economist's lens, accurate forecasting of an economy's future may be theoretically, not just practically, impossible. Steering an economy with positive feedbacks into the best of its many possible equilibrium states requires good fortune and good timing – a feel for the moments when beneficial change from one pattern to another is most possible. Theory can help identify these states and times, and it can guide policymakers in applying the right amount of effort (not too little but not too much) to dislodge locked-in structures”

W. B. Arthur, 2000:11-12

The problem that the economy could lock into lasting situations with high unemployment is a key issue raised by Keynes (1936). Such lock-ins correspond to local optima. There is now ample evidence that markets often (perhaps most of the time) do not attain optimal allocation of resources.

In the context of economic development the equilibrium-based classic model has been severely criticised particularly on two accounts: that equilibrium provides no driving force to explain actually observed change, and that in reality there is lack of information, time and ability to maximise properly (“bounded rationality”, Simon, 1996:87-88). However, without equilibrium between supply and demand, much of the deep theoretical insight thought to be gained through the classic and neo-classic economic models loses its footing. Still, in much conventional, day-to-day economic thinking these models linger on.

As described by Arthur above, high-technology products/services tend to show increasing returns. Arthur (2000) and colleagues have proved that systems with increasing returns generally will have multiple equilibria (for a concrete example see *ibid.* 2000:7). One of these tentative equilibria (optima) will come to be chosen in real life. There is no guarantee that the actually chosen equilibrium will correspond to a global optimum. The choice between equilibria takes place as dynamics play out, and follows a course that depend both on the system's own properties, and on circumstances that develop along the way (cf. Figure 4.1.1). This is the reason why the properties of irreversibility, nonlinearity and feedback are central to the present work.

8.1.4 MODELLING OF IRREVERSIBLE BEHAVIOUR

Reality is irreversible. Correspondence with reality therefore need to model irreversibility to simulate dissipative systems and bifurcations as discussed in Chapter 4. The arguments for capturing irreversibility in the EICOMP model are treated in sections 6.2.1, 6.7.1 and 7.1. In section 6.2.1 we point out how traditional graphs and corresponding functions involving time, simultaneously depict a past and a future, with some arbitrary "now". Such models simply indicate "from-the-outside" how things vary with time. In section 6.7.1 we discuss how system dynamics with its internal mechanism has a "now" converting possible futures into experienced history. This is modelling of dynamic behaviour much closer to what is observed in reality. The overall perspective is discussed in section 7.1. The physical background and understanding allow us the perspective of irreversibility. Irreversibility as described here is not confined to the EICOMP model, but is inherently present in all system dynamic models. It is easy to see in nonlinear behaviour. The importance of irreversibility is that it relates to congruence between model and reality. Systems with feedback and delay are inherently nonlinear. Sterman (2000:129-133) gives a brief introduction to how such

nonlinear dynamics relates to behaviour characterised by chaos. Chaos is closely connected with dissipative systems and bifurcations.

8.1.5 CATEGORIES OF DYNAMICS

Another phenomenon related to irreversibility is different kinds of dynamic behaviour. We have experienced difficulties in communication and discussions relating to dynamic behaviour while the present work has been in progress. There appears to be agreement that all models that involve development with time, may be called “dynamic”. Yet there are differences between systems like the exponential relationship with time (equation 6.2.1.6) and the actual experience of irreversible systems with direction of time as discussed in section 4.1. To illustrate: the exponential relationship is independent of path dependence – we get a certain function value at any given “t”, no matter what the previous input/output was. In contrast, with an irreversible mechanism, as in the EICOMP model, path dependence is created because the function value is influenced by the situation in previous time-steps. We distinguish between two distinct categories of dynamics:

1. *Dynamics observed from the outside*
2. *Dynamics actually experienced*

In the first category there is no direction of time and thereby no difference between “history” and “future”. The second category is quite different: there is a history with corresponding consequences, a “now” where decisions are made, and a multitude of unrealised (potential) futures. Keeping these categories apart is important when attempting to capture real dynamic behaviour, such as path dependence.

8.1.6 A BASIS FOR DETAILED STUDIES

From the outset of seeing hydrogen as part of an existing and developing overall energy infrastructure, various parts of this

infrastructure will be subject to further studies. One result from the approach in the current work is the generic and flexible framework to represent technologies in an infrastructure network. This framework provides a robust foundation for extensions. For example, increased specification of particular technologies may be added, with the integrated infrastructure remaining intact. As a concrete case related to the EICOMP model, “refining” (module # 27, Appendix 1) might be subdivided to distinguish between production of liquid fuels from crude oil or from coal, without any change to the overall supply and flows of energy.

8.2 CRITIQUE OF THE EICOMP MODEL

What makes a good model? Our concern in this section is to build confidence that the EICOMP model is sound. It is generally a challenge to test dynamic models. We follow Sterman’s attitude and advice:

“Validation and verification are impossible. Many modelers speak of model ‘validation’ or claim to have ‘verified’ a model. In fact, validation and verification of models is impossible. The word ‘verify’ derives from the Latin verus – truth; Webster’s defines ‘verify’ as ‘to establish truth, accuracy, or reality of’. ‘Valid’ is defined as ‘having a conclusion derived correctly from premises... Valid implies being supported by objective truth.’

By these definitions, no model can ever be verified or validated. Why?... ...all models, mental or formal, are limited, simplified representations of the real world....

...If validation is impossible and all models are wrong, why then do we bother to build them? As a leader you must recognize that you will be using a model – mental or formal – to make important decisions. Your choice is never whether to use a model but only which model to use. Your

responsibility is to use the best model available for the purpose at hand despite its inevitable limitations. The decision to delay action in the vain quest for a perfect model is itself a decision, with its own set of consequences. Experienced modelers likewise recognize that the goal is to help their clients make better decisions, decisions informed by the best available model. Instead of seeking a single test of validity models either pass or fail, good modelers seek multiple points of contact between the model and reality [cf. Table 1.1 and section 7.1] by drawing on many sources of data and a wide range of tests. In stead of viewing validation as a testing step after a model is completed, they recognize that theory building and theory testing are intimately intertwined in an iterative loop [cf. Figure 7.1.1]. Instead of presenting evidence that the model is valid, good modelers focus the client on the limitations of the model so it can be improved and so clients will not misuse it.”

J. D. Sterman, 2000:846,850

We now turn to describe tests performed to the EICOMP model. Since any real world hydrogen infrastructure development is open, the EICOMP model has to be wide-ranging. Furthermore, as discussed in section 8.1.3 – 8.1.5, due to the nature of positive feedback, irreversibility and Category 2 dynamics, there is no proof of correctness in replicating history (cf. discussion in section 6.3). Yet there are definite notions of whether something is right or wrong. Rather than asking for verification and validation we will seek to ensure that the EICOMP model is *reasonable*. We have sought to test the EICOMP model in this sense in a number of ways:

- *Convergence*
- *Obeying mass- and energy balances*
- *Dimensional consistency*
- *Testing extreme conditions*

- *Control of shifts in expected direction*
- *Unexpected patterns of behaviour*
- *Study behaviour of competition*
- *Shift between electrolysis and reforming for supply of hydrogen free from carbon monoxide for PEM fuel cells (cf. section 3.5)*
- *Sensitivity analysis*

Convergence

The model runs without any error from uninitiated variables, division by zero or similar, and present consistent and reasonable behaviour. With current specification of delays the model converges.

Obeying mass- and energy balances

The model conforms to these conservation laws.

Dimensional consistency

The model requires specification of dimensions for all variables. The software assists in dimensional analysis.

Testing extreme conditions

The model behaves reasonably when subjected to extreme shocks and parameters such as zero demand, no available energy, varying efficiencies etc. To illustrate, if all carbon-based energy is heavily taxed or is made unavailable, the model shifts towards obtaining all the energy it can get from renewable sources. Testing these extreme conditions consistently produce reasonable behaviour.

Control of shifts in expected direction

All capacities, delivered volumes and learning effects consistently follow increasing/decreasing relative attractiveness.

Unexpected patterns of behaviour

Do variables one would expect to change, remain constant with significant shifts for example in oil price? Or the other way round: is there unexpected variation in variables such as capacity when population and energy prices are held constant? Testing thus far has revealed some errors that have been corrected. This is a work intensive process due to the flexibility of the model. As more experience is gained this testing becomes increasingly efficient.

Study behaviour of competition

All competition situations based on attractiveness behave reasonably.

Shift between electrolysis and reforming for supply of hydrogen free from carbon monoxide for fuel cells

This is a test of a fairly complex calculation of attractiveness. The results are reasonable.

The tests discussed above have helped to substantially reduce the probability of drastic or systematic error in the EICOMP model. On the other hand there is no guarantee that all constants expressing assumptions in the model have the most realistic values. As all exogenous input may be changed and experimented with through sensitivity analysis, it is possible to study their effect:

Sensitivity analysis

Sterman (2000) refer to three types of sensitivity related to assumptions: *numerical*, *behaviour mode* and *policy* sensitivity:

- ***Numerical sensitivity*** exists when a change in assumptions changes the numerical values of the results. For example, changing the strength of the word of mouth feedback in an innovation diffusion model will change the growth rate for the new product. All models exhibit numerical sensitivity.

- *Behaviour mode sensitivity* exists when a change in assumptions changes the patterns of behaviour generated by the model. For example, if plausible alternative assumptions changed the behaviour of a model from smooth adjustment to oscillation or from s-shaped growth to overshoot and collapse, the model would exhibit behaviour mode sensitivity.
- *Policy sensitivity* exists when a change in assumptions reverses the impacts or desirability of a proposed policy. If cutting prices boosted market share and profitability under one set of assumptions but led to ruinous price wars and bankruptcy under another, the model would exhibit policy sensitivity.

J. D. Sterman 2000:883

Sensitivity tests as proposed by Sterman have been performed with respect to variables considered important (e.g. for varying primary energy prices and CO₂ taxation) for numerical-, behaviour mode- and policy sensitivity. All results observed thus far appear reasonable. However, for behaviour mode- and policy sensitivity this effort soon changes from model development towards practical application and understanding.

8.2.1 CHOICE OF STRUCTURE

The concepts and related structure of the EICOMP model is extensively described in sections 7.1 - 7.4. This structure reflects the correspondence between concepts in reality and in the model. The main rationale is that there exists a set of interacting technologies in the real world that together meet end-user needs. The model (in its current version) simulates this in the form of 33 distinct modules that represent these technologies. The modules/ technologies interact through competition guided by attractiveness and further reflect bounded rationality. The model is confined to a region, based on the main argument that change has to begin somewhere and initially is

influenced by local conditions. This is principally important in the presence of positive feedback (cf. section 8.1.3), and allows for study of regional opportunities for creation of value.

This is different from the usual topics encountered in discussions of a future hydrogen technology and economy, which concentrates on subjects like production, transport and storage of hydrogen, and technologies like reforming, electrolysis and fuel cells alone. The reason for this difference in approach is the aim of understanding change, and thereby the necessity of both considering the overall existing energy infrastructure and alternative technologies.

An important aspect to the overall model structure is the specification of model boundaries. Sterman (2000:95-96) stresses that variables that must be expected to reflect conditions that are modelled, must be made endogenous. This is presumably the reason why many efforts to model change in climate, are global (e.g. the MARKAL and TIMER models, cf. section 6.8). This way of thinking would lead to the primary energy costs being modelled endogenously. However, the main reasoning is that the purpose in the present work is not to model/predict future energy prices (cf. section 7.5). Leaving these parameters exogenous makes it possible to model development under varying cost regimes.

8.2.2 DEGREE OF AGGREGATION

Aggregation is central to any modelling effort and determines the balance between overview and detail, both key to transparency. As introduced in Chapter 7, Figure 7.3.1.1 illustrates how the EICOMP model is divided into a central (global) and a regional (local) part, and further into five main sectors spanning from end-user markets to primary energy resources. Each sector is divided into a set of modules that represent the ability to relate further detail. Sterman (2000) has the following view on the degree of detail:

“As a rule of thumb, clients generally want to see more detail in a model than the modeller thinks is needed, and modelers, in turn, generally overestimate the detail necessary to capture the dynamics of interest. Of course, the amount of detail needed to capture the dynamics relevant to the client’s purpose and the amount of detail needed to give the client confidence in the results are two different things... [citing Roberts, 1977/1978]: ... ‘You must provide the level of detail that causes [the client] to be persuaded that you have properly taken into account his issues, his questions, his level of concerns. Otherwise he will not believe the model you have built, he will not accept it, and he will not use it’.”

J. D. Sterman, 2000:217

[Roberts, E. (1977/1978): *Strategies for effective implementation of complex corporate models*, Interfaces 7(5)]

From the underlying structure of value sequences and the generic structure, it is methodologically feasible to extend the EICOMP model for particular studies, and in this way obtain more detailed information (cf. section 8.1.6). Also, the current approach is sufficiently computationally effective to allow the modelling to be close to executive levels. It has been the view in the present work that challenges requiring particular details are best tailored in close cooperation with clients.

8.2.3 FURTHER DEVELOPMENT TO THE EICOMP MODEL

In this section we inspect some opportunities for improvement to the EICOMP model. The present version is primarily based in knowledge from constructing the model, whereas future development to the EICOMP model is expected to come from practical simulation experience, going deeper into particular situations and issues. Some concrete suggestions for improvement may be made:

- *Improvement of EICOMP's control panel for user-friendly operation (for example incorporate flexible time-series for interest rates).*
- *Continue sensitivity analysis to improve correspondence between exogenous model parameters and reality (and also to study the influence of parameter accuracy).*
- *Extend the model to explicitly display regional creation of value.*
- *Extend "Hub" solutions to improve transparency and flexibility.*
- *Consider price effects due to bottlenecks (in order to identify opportunities for profitability).*

Improvements are also expected in the software currently in use that will help to improve the EICOMP model and its use (cf. section 6.9). As an example, the introduction of hierarchal structure was a major improvement (cf. section 6.7.2). Some further improvements could be:

- *Automatic stepping through variables*
- *Algorithm for visualising 3-dimensional response surfaces*
- *Access to optimisation (e.g. by means of genetic algorithms)*

The genetic algorithm should be capable of handling optimisation in the complex response surfaces encountered in dynamic, nonlinear simulation (Goldberg, 1989:2-7).

8.3 STRATEGY AND HOW TO ACHIEVE DESIRED CHANGE

Our focus in this section is how to institute change towards sustainability. Section 4.4.2 describes various forms of motivation for change. However, to attain sustainability the introduction of measures in time and required order appears to be severely underestimated. Furthermore, change takes time. We have to start now in order to

have solutions ready for when we need them. We base most of our discussion on considerations from the basis developed in chapters 1 - 6. Our main starting points are:

- *To achieve sustainability requires change away from our present energy infrastructure.*
- *Such change implies new opportunities for creation of value.*
- *To actually achieve change we simultaneously need to consider how to develop a new energy infrastructure, and how to leave the existing one with as low human- and financial costs as possible. Separate policies are needed for investments already made, and for guiding new investments into sustainable channels desired by society.*
- *The EICOMP model will, by simulating technologies and infrastructure as it changes, act as a searchlight to illuminate technical, economic, environmental and societal aspects along the way.*
- *As change will normally start in a local context, this searchlight must be available to support evaluation of technological and economic issues for creation of value in a local manner.*
- *Decisions are correspondingly local in nature, with many actors and stakeholders, and with local access to information and local limitations in this access. In this setting we need local modelling tools.*

These aspects are further discussed in the following sections.

8.3.1 DECISIONS AND POLICY MAKING

Change is more than meets the eye: what do we base our decisions on? Least amount of hassle and friction and the easiest path to reach some sort of consensus, or towards best possible results based on careful analysis and thinking? The latter is vital to avoid unintended consequences elsewhere in the system:

“You cannot meddle with one part of a complex system from the outside without the almost certain risk of setting off disastrous events that you hadn’t counted on in other, remote parts. If you want to fix something you are first obliged to understand...the whole system... Intervening is a way of causing trouble.”

Lewis Thomas, 1974:90

Cited from J. D. Sterman, 2000:8

Furthermore, policy making should be directed at causes rather than symptoms:

“Policies directed at alleviating the symptoms of a problem usually fail because they trigger compensating feedbacks, feedbacks that undercut the intended effects of the policy. The compensating loops arise because other actors, with their own goals, respond to changes in the state of the system in such a way as to offset the intended effects of the policy. While each individual loop may be weak, the combined effect can often compensate completely for any policy directed at the symptom of a problem. Directing policies at the symptoms of a problem is like trying to squeeze a balloon to make it smaller. Whenever you squeeze, the air pressure increases, expanding some other part of the balloon so its volume remains about the same.

Why then do so many policies focus on alleviating the symptoms of difficulty? We focus on symptoms because so much of our experience is with simple systems in which cause and effect are closely related in time and space, in which symptom and cause are obvious.”

J. D. Sterman, 2000:189

Modelling is our best way to understand complex systems well enough to properly capture the relationships between cause and effect. Moreover, it is important to capture features beyond cost to design efficient policies.

Proposal

Identify areas with low risk and low expense as policy instruments. An example of this has been to open bus lanes (designated carriageways for public transport) for electrical vehicles in Oslo. This has proven highly successful at virtually no cost.

8.3.2 VALUE TODAY – VALUE TOMORROW

Motivation for change relates fundamentally to our appreciation of value. This applies in particular to value now compared to value at some time in the future (Portney and Weyant, 1999). Our conventional way of dealing with such questions is through calculation of net present value, which compares a good and its cost/benefit now and in the future, in terms of money (cf. section 4.3.3). This is used to compare alternative investments. The results from net present value calculations depend strongly on the applied rate of interest. This rate of interest is speculative and uncertain in the time perspectives under consideration for energy technologies. As a result net present value calculations in this context are more or less reduced to a number exercise. Yet, this is a real impediment to change. Lasting value is not credited in net present value calculations, meaning that the value of sustainability is not accounted for. The overall result is a future with less economic growth than if long-term investment had been made. New, sustainable technology is forced to compete with existing cost-optimised technology. Firms cannot achieve change on this basis.

Proposal

A natural proposal is that society stipulates the use of low interest rates in net present value calculations for projects related to desired courses towards sustainability.

8.3.3 POSITIVE FEEDBACK AND STRATEGIC OPPORTUNITIES

There is an intriguing relationship between increasing returns caused by positive feedback and strategic opportunities. Business strategy may be understood as optimisation (H. Gether, 2002). Strategy analysis is the process of locating optima and of finding the best way to reach them. In this picture optima exist in the real world, separate and independent from individual firms, but representing strategic opportunities. The optima relate to end-products, and are in the realm of technology and organisation (ibid. 2002:138). As there are multiple optima (cf. section 8.1.3) they are local in nature. They are separated from each other by “valleys” so that transfer from one optimum to another implies a transfer cost. One optimum will be better than the rest and thereby be a global optimum. However, optima are dynamic and may wax and wane with time (e.g. due to development in technology). Hence a global optimum at one point in time may not necessarily remain global at a later point. If a firm remains close to an optimum that increases it is more likely to prosper. Similarly, if close to a shrinking optimum the firm is more likely to succumb, unless it is able to shift to a better one.

The point to be made is the following: *there is a striking analogy between optima in this strategy sense, and the multiple equilibria as described by Arthur.*

Because the EICOMP model embodies increasing return situations, it will in principle (Arthur, 2000:6) imply multiple equilibria. If these equilibria correspond to strategic optima, it follows that the EICOMP model will help to identify such optima. As we vary assumptions (i.e. exogenous input) we may see whether or not the model locks in to

some fixed setting, which would correspond to an equilibrium and hence to a local strategic optimum. We may in principle find multiple optima by modelling. The practical experience with the EICOMP model in this direction is as yet only in its beginning.

At present the correspondence between multiple equilibria resulting from positive feedback, and strategic optima remains a hypothesis. We cannot as yet prove this correspondence - only point to the striking analogy and to the exciting vistas it opens. If correct, it means that we may use EICOMP and similar models to search for optima under dynamic change, and also to point out future targets. As firms get close to optima they increasingly attain competitiveness.

Proposal

Put the EICOMP- and similar models to use in order to detect and characterise positive feedback situations, and to study how nonlinearities change the dominant feedback structure.

8.3.4 INTERMITTENT SUPPLY OF RENEWABLE ENERGY

Another issue related to sustainability results from the intermittent, stochastic nature of renewable energy sources such as solar- and wind power. Present technology cannot store energy in sufficient degree to guarantee continuous supply, and stand-by facilities are therefore required. These are in most areas fossil-fuelled power plants. The required stand-by facilities will, however, due to the renewable power supply, have fewer operation hours available to carry fixed costs. The electric power generated in this way therefore becomes more expensive, even when primary energy costs remain at the present levels. This effect is clearly demonstrated by the EICOMP model. This is yet another financial barrier to change, and one that seems to be underestimated by people wanting to move towards a sustainable society as soon as possible.

Proposal

To reduce the uncertainty from intermittent renewable energy generation, policy making should be directed towards coordination of efforts and ownership. This entails that companies based on traditional fossil fuel technology must incorporate renewable energy production according to their capacities, rather than competing directly with the sustainable initiatives.

8.3.5 INVESTMENTS ALREADY MADE

Even more critical, and hard to bring into the discussion on how to develop a sustainable society, is the following issue: the activities that most people agree are polluting and degrade our environment, are provided by companies that rely for their income and existence on exactly these activities. The protection of these investments, and how they absorb and re-allocate capital back into the same technology, appears as a serious stumbling block to change (cf. section 3.6). How may this be approached?

For guiding new investments towards sustainable technologies the perhaps most important step appears to be avoiding challenging existing investments. Owners of existing investments will oppose policy making that reduce the value of those investments. To guide new investments in the right direction, incentives have to be clear to investors. There should also be a clear communication that new, non-sustainable investments will be punished through taxation or in other ways, to give a predictable framework in which firms operate.

Proposal

Introduce incentives to decommission old polluting equipment, to be replaced with sustainable alternatives. It is important to reduce pollution in particular where people live. This may be encouraged by means of more public transport, renewed in the following way:

diesel → natural gas → syngas → hydrogen.

(syngas = natural gas + hydrogen)

AND where the incremental improvements are credited all the way.

8.3.6 NEW OPPORTUNITIES FOR CREATION OF VALUE

The overall outcome of creation of value is the safe and efficient supply of end-products to a population. On the other hand, in a situation of competition the central objective for firms is to attain competitiveness and survive in the market place. A firm will make money while engaged in value sequences with competitive end-products. Additional opportunities may arise from situations with insufficient supply. Modelling may allow these to be foreseen and taken advantage of. An example of this type of situation is seen in Figure 7.4.2.

A common view of change as extra and unwelcome cost and effort makes change an uphill struggle almost no matter how necessary it might appear. However, different kinds of change also embody new opportunities for creation of value. Hence the present work reflects:

- *Innovation (technological and institutional)*
- *New possibilities and business development*
- *Increased end-user demand (particularly in transport)*
- *Concern over resource depletion and insecurity of energy supply with increasing prices of existing energy services*
- *Willingness to pay for environmental improvement*
- *Change in external conditions (e.g. population- and economic growth in Asia, and competition from other regions)*

Furthermore, there is growing understanding of innovation and requirements for new technology to succeed, as illustrated in Figure 5.2.1. New technology must provide comparable qualities at a lower cost, or superior qualities at an acceptable cost in order to succeed.

Chapter 8

To achieve large volume innovation, attitude and communication are crucial. Important positive feedbacks in this regard are:

- *Increased word-of-mouth to increase knowledge about new products and services*
- *Increased familiarity and proficiency in use that lead to increased end-user value and acceptance.*

As an example of successful local creation of value related to sustainability, the development of the wind turbine industry in Denmark and its positive spillover effects is often cited.

Proposal

See introduction of sustainable energy practices in the same light of creation of value as other efforts towards innovation and economic growth. Society should particularly support this area for the combination of creating jobs and bringing forth sustainable development at the same time.

EPILOGUE

The problems of unrestrained growth have been recognised since the time of Thomas Malthus. Another eye-opener came with “Limits to Growth” by Meadows et al in 1972. This work essentially demonstrated that continuing growth combined with end-of-pipe approaches to curb pollution, will lead to overshoot and collapse of modern society. The oil crisis in 1973 underlined the message by showing how dependent the western world had become on imported crude oil, particularly from the politically unstable Middle East.

In the light of exponential growth (and people’s often poor understanding of the resulting consequences and the time it takes) we may be running short of time to get alternatives ready. Impacts on the global economy may be tolerated, but consequences may prove far

more serious. Are we willing to run this risk? This is not an attempt to prediction of doom, but as stated by Meadows et al (1992):

“...Here are some common biases and simplifications, verbal traps, untruths that we have run into frequently in discussing limits to growth. We think they need to be pointed out and avoided, if there is ever to be clear thinking about the human economy and its relationship to the earth.

- *Not: A warning about the future is a prediction of doom.*
- *But: A warning about the future is a recommendation to follow a different path.”*

D. H. Meadows et al, 1992:229

Growth is above all related to economic growth and increase in populations, with access to fresh water and energy as main constraints. While fresh water is still described as renewable, fossil energy is not. The overall message is that we have to change our current practice of energy consumption.

Moreover, there is a special reason why this is of great importance now: if the developed world adopts sustainable technology and makes it competitive, before the developing countries have invested too much in conventional technology and infrastructure, this creates a *window of opportunity* for the developing world to have an eased transition to sustainable practices. This window of opportunity is really worth striving to take advantage of.

Achieving sustainability is a challenge. However, it is a dilemma that the seriousness of the situation at present tends to go unheeded. A different attitude is called for:

“I find it far more irresponsible - and at the same time paralysing - to uphold an optimism that allows one to let

Chapter 8

things proceed calmly as usual, in the conviction that problems will eventually be overcome through scientific research and new technology, and through the balance of supply and demand in free markets”

Georg Henrik von Wright, philosopher, University of Cambridge

References

Andreassen, K. (1998):

Hydrogen Production by electrolysis

In Saetre, T. O. (ed), 1998, p 91-102 (cf. below)

Argonne (2003):

Basic Research Needs for the Hydrogen Economy

*Report of the Basic Energy Sciences Workshop on Hydrogen
Production, Storage and Use.*

US Department of Energy,

Argonne National Laboratory, May 2003

Arthur, W. B. (1989):

*Competing Technologies, Increasing Returns, and Lock-in by
Historical Events*

Economic Journal, 99: 116-131

Arthur, W. B. (2000):

Increasing Returns and Path Dependence in the Economy

The University of Michigan Press, Ann Arbor.

Baldwin, S. (2002):

Renewable Energy Modeling

Presentation at the Renewable Energy Modeling Series

Meeting 1, October 2, 2002

Policymakers' and Users' Perspectives on Renewable Energy
Modeling

http://www.epa.gov/cleanenergy/renew_series.htm

Bakken, B. H., M. M. Belsnes and P. I. Hansen (2002):

References

Spesifikasjon av analyseverktøy for transportsystemer med multiple energibærere

(In Norwegian: *Specification of analytical tools for transport systems with multiple energy carriers. Technical report*)

SINTEF, Trondheim, Norway

Boyle, G (1996):

Renewable Energy; Power for a Sustainable Future

Oxford University Press in association with the Open University

BP Statistical Review of World Energy (2003):

A consistent and objective series of historical energy market data

<http://www.bp.com/>

Briskeby, S.T. (2002):

Energilagring i Red/Oks-system

(In Norwegian: Storage of energy in red/ox systems)

Institutt for materialteknologi, NTNU, Trondheim, Norway

Buchner, H. (1995):

Hydrogen and Other Alternative Fuels for Air and Ground Transportation

In Pohl, H.W. (ed) 1995, (cf. below)

John Wiley & Sons

Bunn, D. W. and E. R. Larsen (ed.) (1997):

Systems Modelling for Energy Policy

John Wiley & Sons

Carrette, L., K.A. Friedrich and U. Stimming (2001):

Fuel Cells - Fundamentals and Applications

Fuel Cells, 1: 5-39

Wiley InterScience

Consumer Price Index* (2003):

U.S. Department of Labor, Bureau of Labor Statistics, Consumer Price Indexes, <http://www.bls.gov/cpi/>

Inflation Calculator: <http://data.bls.gov/cgi-bin/cpicalc.pl>

* Data source

Cropper, W. H. (2001):

Great Physicists: The Life and Times of Leading Physicists from Galileo to Hawking

Oxford University Press, New York

Croson, R. and K. Donohue (2002):

Experimental Economics and Supply-Chain Management Interfaces, 32: 74-82

Dawkins, R (1986):

The Blind Watchmaker

Penguin Books, London

Dixit, A. K. and R. S. Pindyck (1994):

Investment under Uncertainty

Princeton University Press, New Jersey

Eigen, M. and Winkler, R. (1981):

Laws of the Game. How principles of Nature Govern Chance

Alfred A. Knopf Inc., New York

ETSAP (2004):

Energy Technology Systems Analysis Programme

Annex VIII: Exploring Energy Technology Perspectives

<http://www.etsap.org/main.html>

EU Green Paper (2000):

References

Towards a European strategy for the security of energy supply
Commission of the European Communities, Brussels
http://europa.eu.int/comm/off/green/index_en.htm

EU Energy and Transport in Figures (2002):

Statistical pocketbook

ISBN 92-894-4165-8

http://europa.eu.int/comm/energy_transport/etif/index.html

European Energy and Transport (2003):

Trends to 2030

European Commission, Directorate-General for Energy and Transport, ISBN 92-894-4444-4

http://europa.eu.int/comm/energy/index_en.html

EU High Level Group for Hydrogen and Fuel Cells (2003):

Hydrogen Energy and Fuel Cells

Summary report, European Commission, Community Research.

(Direct reference: D. Hart, Imperial College, London.)

Fiddaman, T. (1997):

Feedback Complexity in Integrated Climate-Economy Models.

PhD Thesis, MIT Sloan School of Management, Cambridge, MA

Forrester, J. W. (1961):

Industrial Dynamics

The M.I.T. Press, Cambridge, Mass., USA.

Forrester, J. W. (1992):

Policies, decisions and information sources for modeling

European Journal of Operational Research, 59: 42-63

Foxon, T. J. (2002):

Technological and institutional 'lock-in' as a barrier to sustainable innovation

ICCEPT Working Paper

<http://www.iccept.ic.ac.uk/public.html>

Garche, J. (1998):

Application of small scale fuel cells in combined heat/power cogeneration

In Saetre, T. O. (ed), 1998: 563-574 (cf. below)

Gaudernack, B. (1998):

Hydrogen production from fossil fuels

In Saetre, T. O. (ed), 1998: 75 (cf. below)

Gether, H. (2002):

Strategy as a Contribution to Prosperity and Sustainability

Overall Optimisation of Technological Value Sequences

Doctor Technicae Thesis, Norwegian University of Science and Technology, Trondheim, Norway

Glöckner, R. (2004):

Lagring av Hydrogen

(In Norwegian: *Storage of Hydrogen*)

Personal communication

IFE, Norway.

Goldberg, D. E. (1989):

Genetic Algorithms in Search, Optimisation, and Machine Learning

(16th reprint), Addison-Wesley, Reading, MA, USA.

Grübler, A., N. Nakicénović and D. G. Victor (1999):

References

Dynamics of energy technologies and global change
Energy Policy 27: 247-280

Grønhaug, K. and G. Kaufmann (ed.) (1988):
Innovation: A cross-disciplinary Perspective
Norwegian University Press, London, UK.

Halseid, R. (2004):
*Ammonia as Hydrogen Carrier. Effects of Ammonia on Polymer
Electrolyte Membrane Fuel Cells*
PhD Thesis, Norwegian University of Science and Technology,
Trondheim, Norway

Hanssen, S. O., J. V. Thue, Ø. Skarstein, F. O. Gjerstad and V.
Novacovic (1996):
ENØK i Bygninger. Effektiv Energibruk
(in Norwegian: Energy Conservation in Buildings. Efficient use of
Energy)
SINTEF/Universitetsforlaget, Oslo, Norway (in Norwegian).

Haugland, T., H. O. Helgesen and K. Roland (1998):
Energy Structures and Environmental Futures in Europe
Oxford University Press, New York

Hoffmann, P. (2001):
*Tomorrow's Energy. Hydrogen, Fuel Cells and the Prospects for a
Cleaner Planet*
The MIT Press, London

Huntington, H. G. and J.P. Weyant (2002):

Modeling Energy Markets and Climate Change Policy

EMF OP52

Energy Modeling Forum

452 Terman Center, Stanford University, Stanford, CA

<http://www.stanford.edu/group/EMF/publications/order.htm>

IIASA CO2DB (2003):

CO2DB database

The International Institute for Applied Systems Analysis (IIASA)

http://www.iiasa.ac.at/Research/ECS/docs/data_index.html

IIASA (2003):

Environmentally Compatible Energy Strategies (ECS) Project

The International Institute for Applied Systems Analysis (IIASA)

<http://www.iiasa.ac.at/Research/ECS/docs/index.html>

ICCEPT (2002):

Assessment of Technological Options to Address Climate Change

A Report for the Prime Minister's Strategy Unit

<http://www.iccept.ic.ac.uk/a5-1.html#work>

ICCEPT/E4tech (2003):

The UK Innovation Systems for New and Renewable Energy

Technologies

Final Report

http://www.dti.gov.uk/energy/renewables/policy/renewables_innovation_review.shtml

Ion, R. M. and L. Fara (1998):

Photophysical and photochemical properties of some porphyrins used in the photosensitized reduction of water

In Saetre, T. O. (ed), 1998: 161-167 (cf. below)

Iskov, H. (2004):

References

The use of the natural gas distribution system for pure hydrogen
Dansk Gasteknisk Center a/s, (Personal communication)

Keynes, J. M. (1936):

The General Theory of Employment, Interest and Money
Macmillan, Cambridge, UK

Korpaas, M. (2004):

Distributed Energy Systems with Wind Power and Energy Storage
PhD Thesis, Norwegian University of Science and Technology,
Trondheim, Norway

Krugman, P. (1979):

Increasing returns, monopolistic competition, and international trade
Journal of International Economics, 9(4):469-479

Landau, R., T. Taylor and G. Wright (1996):

The Mosaic of Economic Growth
Stanford University Press, Stanford, Cal., US

Lehman, M. (2002):

Risk Management for CO₂ Capture, Transport and Storage
Det Norske Veritas, DNV Research, Oslo, Norway. Presentation,
IBC's Carbon Sequestration Conference, London, May 15., 2002

Meadows, D. H., D. L. Meadows and J. Randers (1992):

Beyond the Limits; Global Collapse or a Sustainable Future.
Earthscan Publications, London

Mintzberg, H., B. Ahlstrand and J. Lampel (1998):

Strategy Safari, A Guided Tour through the Wilds of Strategic Management.

The Free Press, New York

Mork, K. A. (1996):

Videregående Makroøkonomi

(in Norwegian: Advanced Macro-economics), 2nd ed.

Cappelen Akademisk Forlag, Oslo, Norway

Mosekilde, E. (1996):

Topics in nonlinear dynamics

Application to Physics, Biology and Economic Systems

World Scientific, London

Moxnes, E. (1992):

Positive feedback economies and the competition between “hard” and “soft” energy supplies

Journal of Scientific and Industrial Research 51: 257-265.

Myrtveit, M (2003):

Cognitive perception

Personal communication

Dynaplan

National Renewable Energy Laboratory (NREL) (2003):

Power Databook

http://www.nrel.gov/analysis/power_databook/

Naylor, T. H. (1971):

Computer simulation experiments with Models of Economic Systems

Wiley, New York

Pohl, H.W. (ed) (1995):

References

Hydrogen and Other Alternative Fuels for Air and Ground Transportation
John Wiley & Sons, New York

Porter, M. E. (1980):

Competitive Strategy, Techniques for Analyzing Industries and Competitors
(Reprint 1998)
The Free Press, London

Porter, M. E. (1990):

The Competitive Advantage of Nations
The MacMillan Press Ltd, London

Portney, P. R. and J. P. Weyant (1999):

Discounting and Intergenerational Equity
Resources for the Future, Washington DC

Prigogine, I. (1997):

The End of Certainty. Time, Chaos, and the New Laws of Nature
The Free Press, London

Ravindran, A., D. T. Phillips and J. J. Solberg (1987):

Operations Research, Principles and Practice
John Wiley & Sons, New York.

Romer, P. M. (1990):

Endogenous technological change
Journal of Political Economy 98(5, Part 2): 71-102

Rosenberg, N. (1996):

Uncertainty and Technological Change
In Landau et al (ed), 1996: 344 (cf. own reference).

Russel, B. (1912):

The Problems of Philosophy

Oxford University Press, UK.

Norwegian Ed.: Gyldendal "Fakkell", (Reprint 1964), Oslo
Norway.

Saetre, T. O. (ed.) (1998):

Hydrogen Power: Theoretical and Engineering Solutions

Kluwer-Academic Publishers, London

Schwartz, P. (2001):

*The Art of the Long View. Planning for the Future in an Uncertain
World.*

John Wiley & Sons, Chichester, UK

Sejersted, F. (ed.). (1982):

Vekst Gjennom Krise; Studier i Norsk Teknologihistorie

(In Norwegian: Growth through Crisis; Studies in Norwegian
History of Technology).

Universitetsforlaget, Oslo, Norway

Simon, H.A., (1978):

Nobel Banquet Speech

<http://www.nobel.se/economics/laureates/1978/simon-speech.html>

Simon, H.A., (1996):

Models of my life

The MIT Press, Cambridge, MA.

Skjølsvik, K.O., (2003):

References

Status, muligheter og utfordringer forbundet med motorteknologi basert på hydrogen

(In Norwegian: *Status, opportunities and challenges in engine technology for hydrogen*)

HYNOR Conference, Stavanger,

<http://www.hynor.no/nyheter/20031201.php>

+ Personal communication

Solow, R. M., (1988):

Growth Theory: An exposition.

Nobel Prize Lecture., 1987,

Oxford University Press, New York

Sterman, J. D. (1991):

A Sceptic's Guide to Computer Models

In Barney, G. O. et al (ed.), *Managing a Nation: The Microcomputer Software Catalog.*

Westview Press, 209-229. Boulder, CO:

Sterman, J. D. (2000):

Business Dynamics: Systems Thinking and Modeling for a Complex World

Irwin/McGraw-Hill, New York

Stiglitz, J. (2002):

Globalization and its discontents

Penguin Books, UK

ISBN 0-713-99664-1

Sørensen, B. (2001):

Scenarier for samlet udnyttelse af brint som energibærer i Danmarks fremtidige energisystem
(In Danish: *Hydrogen as an energy carrier – scenarios for future use of hydrogen in the Danish energy system*)
Roskilde University
<http://mmf.ruc.dk/energy/>

Tande, J. O. G (1999):
Likeverdig sammenligning av energiteknologier
(In Norwegian: *Commensurable comparison of energy technologies. Technical Report*
SINTEF, Trondheim, Norway

Victor, D. G., T. C. Heller and M. Victor (2003):
Political Economy and the Hydrogen Revolution
Working Paper #17,
Program on Energy and Sustainable Development,
Stanford University, CA.
<http://pesd.stanford.edu>

de Vries, B. J. M., D. P. van Vuuren, M. G. J. den Elzen and M. A. Janssen (2001):
The Targets IMage Energy Regional (TIMER) Model
Technical Documentation
RIVM report 461502024 2001

Wainwright, J. and M. Mulligan (ed.) (2004):
Environmental Modelling: Finding Simplicity in Complexity
John Wiley & Sons

Winter, C. J. and J. Nitsch (1988):
Hydrogen as an Energy Carrier. Technologies, Systems, Economy.
Springer-Verlag, Berlin

Wong, Y. K. and G. A. Karim (1998):

References

A kinetic examination of the effects of the presence of some gaseous fuels and preignition reaction products with hydrogen on its autoignition characteristics in engines

In Saetre, T. O. (ed) 1998: 63 (cf. above).

Appendix 1

<i>Consumption sector</i>			
	transp	transport sector	1
	stationary appliances heating	stationary sector	2
<i>Regional infrastructure</i>			
	lq ICE	liquid fuel internal combustion engine	3
	lq ICE hyb	liquid fuel internal combustion engine w/ electrical drive train	4
	lq RFC	liquid fuel reforming fuel cell	5
	ng ICE	natural gas internal combustion engine	6
	ng ICE hyb	natural gas internal combustion engine w/ electrical drive train	7
	ng RFC	natural gas reforming fuel cell	8
	h ICE	hydrogen internal combustion engine	9
	FC	hydrogen fuel cell	10
	EI	electrical vehicle	11
	RFC	stationary natural gas reforming fuel cell	12
	FC	stationary hydrogen fuel cell	13
	el.transf.	electrical transforming	14
	el.transf.	electrical transforming for appliances	15
	lq burner	electrical transforming for heating	16
	el.transf.	electrical transforming for appliances	16
	ng burner	electrical transforming for heating	17
	lq st	liquid fuel station	17
	ng st	natural gas fuel station	18
	ref/h st	natural gas reformed onsite to hydrogen fuel station	19
	h st	hydrogen fuel station	20
	electrolysis (local)	local conversion of electricity to hydrogen	21
<i>Energy transport</i>			
	lq transp	liquid fuel transport	22
	ng transp	natural gas transport	23
	lgt	* link gas transport (only to simplify readability)	
	h transp	hydrogen transport	24
	lht	* link hydrogen transport (only to simplify readability)	
	el transp	electrical transmission	25
	let	* link electrical transmission (only to simplify readability)	
	hot water transp	hot water transportation (from generating of electricity)	26
<i>Conversion</i>			
	refining	conversion of oil and coal to liquid fuel	27
	reforming	conversion of natural gas and coal to hydrogen	28
	electrolysis (central)	central conversion of electricity to hydrogen	29
	generating	conversion of natural gas, coal, nuclear, hydro, wind and sun to electricity	30
<i>Energy sources</i>			
	Non renewables	energy resources (oil, natural gas, coal and nuclear)	31
	Renewables (central)	central energy sources (hydro, wind and sun)	32
	Renewables (local)	local energy sources (bio, wind and sun)	33

Appendix 2

