THE TREATMENT OF UNCERTAINTY IN AIRPORT STRATEGIC PLANNING

J.H. Kwakkel

THE TREATMENT OF UNCERTAINTY IN AIRPORT STRATEGIC PLANNING

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,

in het openbaar te verdedigen
op maandag 6 december 2010 om 15.00 uur

door
Jan Hendrik KWAKKEL

Ingenieur Technische Bestuurskunde
geboren te Dronten

Dit proefschrift is goed gekeurd door de promotoren:
Prof.dr. W.E. Walker
Prof.dr. G.P. van Wee

Copromotor:
Dr.ir. V.A.W.J. Marchau

Samestelling promotiecommissie:

Rector Magnificus	voorzitter
Prof.dr. W.E. Walker	Technische Universiteit Delft, promotor
Prof.dr. G.P. van Wee	Technische Universiteit Delft, promotor
Dr.ir. V.A.W.J. Marchau	Technische Universiteit Delft, copromotor
Prof.dr. A.C. Petersen	London School of Economics, United Kingdom
Prof.dr. J.G. de Wit	Universiteit van Amsterdam
Prof.dr. R. Curran	Technische Universiteit Delft
Prof.dr.ir.W.A.H. Thissen	Technische Universiteit Delft

TRAIL Thesis Series nr. T2010/13, the Netherlands TRAIL Research School

TRAIL Research School
PO Box 5017
2600 GA Delft
The Netherlands
T +31 15 2786046
E info@rstrail.nl
I www.rstrail.nl

ISBN: 978-90-5584-138-7

Copyright:© 2010 by J.H. Kwakkel
All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission form the author.

Printed in the Netherlands

Preface

This dissertation is the result of a research project that was started four years ago, although one can trace it back even further. Six years ago, I started working as a research assistant for prof. dr. Warren Walker on a Fifth Framework project for the European Commission. This project aimed at developing a Decision Support System (DSS) for supporting Airport Strategic Planning. While working on this project, I became ever more interested in model-based decision support. When Warren presented me with the opportunity to continue my research on Airport Strategic Planning, I gladly seized the opportunity. This dissertation is the final result of my continued research on Airport Strategic Planning.

The dissertation is focused on the treatment of uncertainty in the long-term planning of airports. The central contribution of this thesis is that it outlines an approach for handling uncertainty in Airport Strategic Planning better than the current dominant approach. It addresses what is meant by uncertainty, and provides a generic typology of uncertainty in model-based decision support. It presents a generic approach for the treatment of the different types of uncertainty and illustrates this approach by applying it to the long-term development of an airport. Finally, it provides computational evidence that a plan based on the outlined approach can more successfully steer the development of an airport despite the presence of various uncertainties than the current dominant approach. The research reported on has been made possible in part by funding from the Transport Research Center Delft and a Sixth Framework project for the European Commission.

The main body of this thesis consists of a set of journal papers. These papers have been co-authored with various people. In addition to Warren as promoter and Vincent Marchau as co-promoter, I have collaborated with Scott Cunningham and Jan-Willem van der Pas on the papers found in this dissertation. The joint research with Scott has taken on a life of its own. We have attended multiple conferences together, frequently discuss both our individual as well as our joint research interests over coffee, and still have various research ideas we want to jointly work on in the future. Although not a co-author of any of the papers found in this dissertation, my joint work with Roland Wijnen, first at the faculty of Aerospace Engineering and later at the faculty of Technology, Policy and Management, has been highly influential on the final results. Discussions with Roland while traveling to Aachen and back in the context of the aforementioned project for the European Commission contributed to two of the five papers. Moreover, his willingness to share the code of his HARMOS DSS allowed me perform the computational experiments reported on in Chapter 6. As my second promoter, Bert van Wee was involved primarily at the early stages, during my formulation of the research plan, and near the end, during the completion of the dissertation. I have appreciated his practical attitude and constructive criticism.

The research was carried out in the Policy Analysis section of the faculty of Technology, Policy and Management. This group brings together methodologically interested researchers that work on a wide spectrum of policy problems. Although the focus of my research was on long-term airport planning, my interest was not confined to this single application domain. As such, the Policy Analysis section provided an environment in which it was possible to get alternative perspectives on the treatment of uncertainty from different application domains.

J.H. Kwakkel
Delft, October 2010

Content

List of Figures

List of Tables

1 Introduction

1.1 Background

The aviation industry operates in a fast changing environment. At the end of the 1970's, the air transport industry was liberalized and privatized in the U.S.A. Europe followed in the 1990's. As a result of this privatization and liberalization, the air transport industry has undergone unprecedented changes, exemplified by the rise of airline alliances and low cost carriers (Burghouwt, 2007; de Neufville & Odoni, 2003; Forsyth, 1998). Parallel to this, the aviation industry has witnessed increasing environmental awareness, which has resulted in more attention being paid to the negative external effects of aviation, such as noise and emissions, and, since 9/11, safety and security are also of more concern. It is likely that the aviation industry will become even more dynamic in the coming years, for example because of the recently signed U.S.A.-EU Open Skies treaty. In addition to these changes in the air transport industry itself, outside influences such as the oil price, flu epidemics, and financial and economic woes further add to the volatility of aviation demand development. All these changes together pose a major challenge for airports. They have to make investment decisions that will shape the future of the airport for many years to come, taking into consideration the many uncertainties that are present.

1.1.1 Positioning the research

Airport strategic planning (ASP) is an activity carried out by an airport that produces plans for the medium- to long-term development of the airport. In general, strategic planning is defined as 'the managerial activities that produce fundamental decisions and actions that shape and guide what the organization is, what it does, and why it does it' (Bryson, 1995). In order to support strategic planning in general, various different activities can be employed, ranging from detailed quantitative studies to meetings with stakeholder (Mayer, van Daalen, & Bots, 2004). In the context of ASP, the dominant approach is Airport Master Planning (AMP) (Burghouwt, 2007; de Neufville & Odoni, 2003).

As will be outlined below in more detail, AMP relies heavily on models for estimating future demand levels and calculating the expected performance of the airport given these future demand levels (de Neufville & Odoni, 2003). That is, these various models serve an advisory purpose, which suggests that AMP is dominated by 'design and recommend' activities (Mayer, et al., 2004; Yücel & Van Daalen, 2009). This domination of model-based design and recommend activities does not imply that other types of activities aimed at engaging stakeholders, clarifying their values, and mediating conflicts are not employed in the context of the current practice of ASP. The FAA, the American Air Transport Regulator, has specific guidelines for stakeholder engagement (FAA, 2005) and for example van Eeten (1998) reports a case study of the various frames involved in the long-term decisionmaking process about the future of Amsterdam Airport Schiphol. However, AMP as the current approach for ASP, is still predominantly characterized by object-oriented activities (Mayer, et al., 2004) focusing on the airport system and its measurable inputs (e.g. demand) and outputs (e.g. economic performance, noise, emissions).

This thesis will stay within the confines of the object-oriented activities that are currently dominant in ASP. Therefore, no attention will be given to subject-oriented activities that focus on the interaction between the various actors. Instead, this research focuses on exploring the limits uncertainty imposes on the object-oriented activities that currently dominate in ASP, what the negative consequences are of the current approach, and how these negative consequences can be overcome by employing these object-oriented activities in a different way. That is, this research is committed to outlining a critical rational style decision support for ASP within the limits imposed by uncertainty.

1.1.2 The current approach to airport strategic planning: airport master planning

The current dominant approach for the long-term development of an airport is AMP. AMP is a formalized, structured planning process that results in a Master Plan that 'presents the planner's conception of the ultimate development of a specific airport' (ICAO, 1987). As such, the focus in AMP is on the development of plans and not on the decisionmaking process about the plans. In the United States, the FAA has set up strict guidelines for an AMP study (FAA, 2005). Internationally, reference manuals of IATA and books about airport planning by leading scholars heavily influence AMP practices (e.g. de Neufville & Odoni, 2003; IATA, 2004; ICAO, 1987).

The goal of a Master Plan is to provide a blueprint that will determine the future development of the airport (Burghouwt & Huys, 2003; Dempsey, Goetz, & Szyliowicz, 1997). As such it describes the strategy of an airport operator for the coming years, without specifying operational concepts or management issues. A Master Plan covers both the aeronautical developments (i.e. runways, terminals) and non-aeronautical developments (e.g. real estate, commercial activities, and retail developments) of the airport. The time horizon covered in a Master Plan can vary, depending on the situation of the airport for which the Master Plan is being developed, but in general a time horizon of 20 to 30 years is used (FAA, 2005).

AMP follows a strict linear process (de Neufville & Odoni, 2003):

- Analyze existing conditions
- Make an aviation demand forecast
- Determine facility requirements needed to accommodate this forecasted demand
- Develop and evaluate several alternatives to meet these facility requirements
- Develop the preferred alternative into a detailed Master Plan

AMP, as the main way to shape and determine the long-term development of an airport, has proven to be ineffective, as can be seen for example in planning failures at Amsterdam Airport Schiphol, Denver International Airport, Boston Logan Airport, and Montréal Mirabel Airport. In 1995, a plan for the long-term development of Amsterdam Airport Schiphol was accepted. This plan had a time horizon of 20 years, but was obsolete in 1999, due to the unanticipated rapid growth of aviation demand (Kwakkel, Walker, & Marchau, 2007). The new Denver Airport was developed because of anticipated growth, which did not materialize. The new airport ended up with fewer air transport movements than took place at the old airport (Dempsey, et al., 1997; Szyliowics & Goetz, 1995). Boston Logan planned and started the construction of a new runway in the early 1970's, but, due to unanticipated changes in regulations and strong stakeholder opposition, they were unable to open this runway until 2006 (Cidell, 2004; Kwakkel, et al., 2007; Nelkin, 1974, 1975). Montréal Mirabel Airport was constructed in 1975 and was forecast to handle 40 million passengers by 2025. However, the airport failed to attract significant travel and was closed for passenger traffic in 2004 (Canadian Press, 2006). Given the ongoing transition of the aviation industry from a state-owned and state-run enterprise to a market situation, with its associated changes in how the public and the government view the aviation industry, the number and severity of the uncertainties is only expected to increase. In light of this, Master Planning becomes even less appropriate for long-term airport planning.

1.1.3 The challenge of uncertainty for airport master planning

AMP has been unsuccessful in planning the future development of airports. As the examples of Amsterdam Airport Schiphol, Denver International Airport, Boston Logan Airport, and Montréal Mirabel Airport illustrate, plans become quickly obsolete and are not robust with regard to the future. In other words, uncertainty (e.g. aviation demand, regulatory context, technological breakthroughs) is a key source of problems in ASP. In this section, we explore in more depth how uncertainty is currently treated in AMP, why this treatment is inadequate, and what this implies for alternative treatments.

Aviation demand forecasting forms the basis for a Master Plan. An aviation demand forecast can be a forecast for the number of passengers, the tons of goods, or the number of air transport movements, although the forecast usually contains information concerning all three. For example, the forecasts used for the plan for the long-term development of Schiphol in 1995 were forecasts of aviation demand for 2015 in terms of passengers. Given assumptions about the composition of the fleet and the average number of passengers on the various types of airplanes, this was translated into forecasts for air transport movements. By comparing a forecast with the existing conditions at an airport, an assessment can be made whether there is a need for new or expanded facilities. As such, aviation demand forecasting is the main way in which uncertainties about the future context in which an airport operates are handled. The basic concept of developing an aviation demand forecast is simple: past trends, based on time series data and/or theories about underlying mechanisms, are identified and extrapolated forward. In mathematical terms, a relationship between independent variables (X_1, X_2, …, X_n) and the dependent variable (Y) is developed that matches aviation demand observed in the past. In its simplest form, only a single equation is used to model the relationship between the independent and dependent variables. However more sophisticated mathematical models and simulation models are also used. The resulting model is used for extrapolation in order to obtain a forecast for the year of interest (FAA, 2001).

Forecasting in general has come under increasing criticism. The criticisms can be split into two categories: forecasting failure due to bias and forecasting failure due to uncertainty.

Forecaster bias contributes to forecast failure in several ways. Forecasters often integrate political wishes into their forecasts (Flyvbjerg, Bruzelius, & Rothengatter, 2003). Forecasts by project promoters may be even more biased, since the promoter has an interest in presenting the project in as favorable a light as possible (Flyvbjerg, et al., 2003).

Forecasting failure due to uncertainty manifests itself in several ways. As pointed out by Flyvbjerg et al. (2003), discontinuous behavior of the phenomena we try to forecast, unexpected changes in exogenous factors, unexpected political activities, and missing realization of complementary policies are important reasons for forecasting failure. Ascher (1978) sees faulty core assumptions as a prime reason for forecasting failure. Faulty core assumptions refers to the fact that, since the phenomenon we are trying to forecast is not completely understood, forecasters have to make assumptions about the data they need, the formulas to be used, etc. (Porter, Roper, Mason, Rossini, & Banks, 1991). This problem of faulty core assumptions often translates into criticism of the models that are used for forecasting. With respect to data, there are also several uncertainties. Forecasters often have a poor database that has internal biases caused by the data collection system, and they use data from their home countries (instead of the local areas) for calibrating their models (Flyvbjerg, et al., 2003). In addition, forecasters have a tendency to misjudge the relevance of (recent) data (Porter, et al., 1991). Despite these problems with data, forecasters still rely heavily on historic data for testing the adequacy of a forecast. However, there are in principle an infinite number of formulas possible that can match the given historical data. Moreover, in order to forecast a dependent variable Y based on a formula $Y = f(X_1, X_2, \ldots, X_n)$, forecasts are needed for the future values of the *n* independent variables. Instead of forecasting a single variable, one ends up forecasting *n* variables. Even if the problems associated with forecaster bias are addressed, forecasting failure due to uncertainty means that forecasting can always go wrong. By looking at the past and assuming that past behavior will continue into the future, uncertainties leading to trend breaks are overlooked, which, often, are the uncertainties with the largest impacts on the system.

In the case of aviation demand forecasting, forecasting failure due to uncertainty is of specific importance. Over the last twenty years, the aviation industry worldwide has undergone exceptional changes. It has moved from a heavily regulated, state-owned, state-operated industry, towards a fully privatized industry. Currently, aviation transport in the US and Europe is largely privatized, while other regions in the world are moving in this direction as well. The net result of this privatization is that there have been unprecedented changes in the air transport sector, exemplified by the KLM-Air France merger, the rise of airline alliances, the US-EU Open Skies treaty, the rise of low cost carriers, and fierce competition between airports in order to attract carriers. Burghouwt (2007) has studied how airline networks evolved in Europe over time during these changes and concludes that air traffic demand is becoming more volatile and more uncertain, implying that forecasting air traffic demand for specific airports is becoming ever more problematic.

In the process of forecasting, scenarios can be quite useful. An important reason for using scenarios is the aforementioned problem that information on the independent variables is needed. These frequently are based on long-term scenarios developed for that specific purpose by national agencies. However, there are several problems with the way in which this is generally done. First, scenarios from national agencies are general purpose scenarios that have to be adapted for use to the specific case. For instance, in the Netherlands, the economic scenarios developed by the Netherlands Bureau for Economic Policy Analysis (Centraal Planbureau, or CPB) are used as one of the inputs for air transport demand forecasting for

Amsterdam Airport Schiphol. This creates varies problems related to interpretation (Ortúzar & Willumsen, 2002). Second, forecasting suffers from garbage in garbage out: if the scenarios that are used for some of the inputs suffer from defects, so will the forecasts derived from them (Ortúzar & Willumsen, 2002). Third, scenarios are often trend based rather than trend break (Annema & de Jong, 2009). That is, in the scenarios it is assumed that the future will be more or less a continuation of the past. The difference between the scenarios is only how strong the continuation will be. However, trend based scenarios can grossly underestimate the multiplicity of plausible futures that are consistent with the available knowledge and information (Lempert, Popper, & Bankes, 2003). Fourth, scenarios can be used in various ways. Originally, scenarios were meant for 'what if' analysis to support the development of static robust policies (i.e., they would do fairly well under all of the scenarios). However, frequently when using scenarios in forecasting, one of the scenarios is chosen as the most likely future and the others are neglected (de Neufville & Odoni, 2003). Thus, the resulting policy would do well if that one scenario happened, but not necessarily if one of the others did.

Apart from the fact that aviation demand forecasting is highly problematic in light of the many uncertainties that are present, there are several additional reasons that make aviation demand forecasting as the main way to treat uncertainty in AMP inadequate. First, very often, although multiple aviation demand forecasts are generated, only one is used as the foundation for the Master Plan (de Neufville & Odoni, 2003). The airport Master Plan is then designed based on this specific forecast. By making only a single forecast, however, one runs the risk of severely underestimating the range within which future aviation demand might develop. Second, there are many uncertainties present when developing plans for the long-term development of an airport. Aviation demand is only one such uncertainty. Other uncertainties include, among others, regulatory developments, technological developments, and demographic developments. Airport Master Planning does not generally consider these other uncertainties (de Neufville & Odoni, 2003). Third, even if these uncertainties were considered, surprises or “black swans” can still happen. Fourth, the Master Plan that results from the AMP process has a blueprint character (Burghouwt & Huys, 2003). It presents a construction plan that envisions the maximum development of the airport and guides the capital investments in its facilities (Kazda & Caves, 2000). As such, it generally does not consider changing conditions or conditions significantly different from those presented by the forecast. Consequently, the Master Plan leaves little room for adapting to changing conditions during the implementation phase. As a result, the Master Plan is static in nature and leaves little room for adapting to changing conditions.

In practice, airport planners are now increasingly trying to deal with the inefficacy of AMP in various ways. There are (formal) procedures in place for updating the Master Plan; since the Master Plan often consists of a list of semi-independent capital investment projects, there is some room to postpone or speed up projects. Sometimes, it is also possible to solve problems that emerge when the real world deviates significantly from the anticipated world underlying the Master Plan by making operational adjustments. For example, the Australian Airport Act specifies that the Master Plan needs to be formally updated every five years. In this way, stepwise adaptation to changing conditions is realized. Despite the well-known problems surrounding the development of the Denver Airport, Dempsey et al. (1997) praise the technical design of the airport for its modular character, which allows for the gradual development of the airport facilities depending on how the actual conditions evolve. Operational procedures at Schiphol were implemented that allow for the use of one of the finger piers for both Schengen and non-Schengen passengers. By changing the setting of certain doors, passengers are guided to immigration or not. This allows the airport to adapt to changing patterns of demand. Still, the ways in which airport planners cope with uncertainty and find creative workarounds mainly operate within the confines of a static Master Planning approach based on a limited appreciation for the multiplicity of futures.

1.2 Problem statement, research goals, and research questions

We summarize the preceding discussion: a key problem in Airport Strategic Planning is the inadequacy of airport Master Planning for the long-term development of airports. For a while, this has been recognized in the literature, and an alternative to AMP is called for (Burghouwt, 2007; de Neufville, 2000). AMP is inadequate for the long-term development of an airport because the resulting plan is not robust with respect to future developments. This lack of robustness is the result of the fact that (a) very few uncertainties are addressed – usually only aviation demand uncertainties; (b) only a single trend based extrapolation for future demand is considered, instead of a range of plausible demands including demands based on trend breaks; and (c) a Master Plan is static, in that it does not have any mechanisms in place for reacting to changing conditions.

An alternative approach for ASP that is discussed in the literature is based on adaptability and flexibility. Instead of trying to predict future demand, which is known to be very volatile, it is recommended that plans should be able to cope with a range of demand levels. To realize this, a variety of techniques and approaches, such as real options, experimentation, flexible strategic planning, scenarios, and adaptive policymaking, have been put forward in the airport planning literature (Burghouwt, 2007; de Neufville, 2000; Walker, 2000; Walker, Rahman, & Cave, 2001). In light of this, our research goals are to design an alternative approach for ASP that overcomes the weaknesses of AMP, test the performance of this new approach, and compare its performance to AMP. The central question this thesis aims to answer is:

How can uncertainties in ASP be treated in a better way than is currently done in AMP?

In order to answer this central question, several sub questions need to be answered, namely

1. *What is meant by uncertainty?* This question aims at making transparent the various meanings associated with the term uncertainty, revealing terms that are used in a similar way, and identifying the various contexts or discourses that are internally coherent in their use of language when discussing uncertainty. This question is motivated by the fact that uncertainty is an ambiguous concept that has various shades of meaning and is used in different ways across the sciences. In order to develop an approach for the treatment of uncertainty in ASP, we first need to clarify what is meant by uncertainty.
2. *How can uncertainties be classified and analyzed for model-based decision support?* This question aims to synthesize the wide array of existing conceptual frameworks, taxonomies etc. that are available in the model-based policy analysis literature, while taking into account the various ways in which uncertainties are discussed as identified in response to Question 1. The focus on model-based policy analysis is motivated by the observation that the current ASP practice −AMP− is heavily reliant on model-based decision support.
3. *How can the uncertainties currently prevalent in ASP be treated better?* The aim of this question is to design an approach for ASP that is tailored to addressing the uncertainties prevalent in ASP. Such a design will be based on an analysis of the uncertainties prevalent in ASP using the framework resulting from Question 2 and an analysis of the available approaches for the treatment of uncertainty.

4. *How can the efficacy of the new approach for ASP be assessed?* To our knowledge, there is no established methodology for assessing the efficacy of new infrastructure planning approaches in general. This question aims at outlining such a general methodology and then tailoring it to the specific context of ASP.
5. *How does the efficacy of the proposed approach for the treatment of uncertainty in ASP compare to the efficacy of AMP?* The design for a new approach for ASP that results from answering Question 3 needs to be tested using the methodology developed as an answer to Question 4. This question aims at assessing the efficacy of the new planning approach in comparison to AMP.

1.3 Research methodology

Throughout this research, a variety of methods have been employed. Below, we provide the background for each research question and the methods that have been used for answering each individual question.

1.3.1 What is meant by uncertainty?

Uncertainty is prevalent in practically all modern day decisionmaking. The financial crisis that emerged in 2008 and the ensuing recession remind us of the prevalence of uncertainty in decisionmaking about the future. But economic or financial uncertainty is just one source of uncertainty. Uncertainties surrounding climate change, technological developments, the price of oil, etc., all have profound impacts on the expected consequences of the different options that are considered in any decisionmaking problem. When the expected consequences of alternative courses of action are uncertain, identifying good decisions is hard. Although uncertainty might appear to be a relatively modern phenomenon, uncertainty has always been a problem, specifically when the future is involved. From early days when men started to plant crops, for example, uncertainty about the weather had an impact on when to start planting and when to harvest. Given this long history of decisionmaking under uncertainty, it is not surprising that a wide variety of analytical tools and techniques have emerged that offer decision support.

A variety of conceptual schemes, definitions, and typologies of uncertainty have been put forward in different scientific fields (Funtowicz & Ravetz, 1990; Knight, 1921; Morgan & Henrion, 1990; e.g. Refsgaard, van der Sluijs, Brown, & van der Keur, 2006; van Asselt, 2000; van der Sluis, 1997; Walker, et al., 2003). For example, in risk analysis the distinction between aleatory and epistemic uncertainty emerged (Helton, 1994; Hoffman & Hammonds, 1994). Epistemic uncertainty denotes the lack of knowledge or information in any phase or activity of the modeling process. Aleatory uncertainty denotes the inherent variation associated with the physical system or the environment under consideration. Others have tried to clarify where uncertainty manifests itself in the form of a source or location of uncertainty (Morgan & Henrion, 1990; Walker, et al., 2003), and still others have tried to classify the severity of the uncertainty in the form of a level of uncertainty (Courtney, 2001; van Asselt, 2000; Walker, et al., 2003). That is, where does the uncertainty manifest itself along the continuum ranging from deterministic knowledge to total ignorance?

Walker et al. (2003) have attempted to integrate these different frameworks, typologies, and taxonomies in an overarching framework that could provide a starting point for the systematic treatment of uncertainty in model-based decision support. They define uncertainty as ‘any departure from the unachievable ideal of complete determinism’ and identify three different dimensions of uncertainty that together characterize uncertainty. Following earlier literature, we will refer to this framework as the W&H framework after the first two authors (Walker and Harremoës) (e.g Aslaksen & Myhr, 2007; Gillund, Kjølber, Krayer von Krauss, & Myhr, 2008; Krayer von Krauss, 2005).

The term 'uncertainty' has a plethora of meanings and connotations. The W&H framework has been criticized on the ground that it fails to do justice to this (Norton, Brown, & Mysiak, 2006). When one uses the term, the context shapes which specific meaning and connotation is used. Furthermore, words such as ignorance, doubt, unsureness, risk, ambiguity, imprecision, and randomness all capture some element of what is meant by uncertainty. If one aims at providing an integrative framework for classifying uncertainty, an analysis of the diversity of meanings and concepts associated with uncertainty, and the contexts in which these meanings and concepts are used, is called for (Norton, et al., 2006). The importance of performing such an analysis is underscored by Gillund et al. (2008), who observed that the W&H framework is often difficult to understand for researchers with a different background from that of the authors. Related, Skeels et al. (2008) point out that an empirical evaluation of the success of the W&H framework in harmonizing the language with respect to uncertainty in model-based decision support is missing.

In order to assess how and in what directions the W&H typology needs to be improved in order to capture the aspects of uncertainty relevant to strategic planning and decisionmaking, it is necessary to first survey the diversity of terms and meaning associated with uncertainty and the situations in which these usages arise. We used science mapping for this. Science maps "attempt to find representations of the intellectual connections within the dynamically changing system of scientific knowledge (Small, 1997)". Such maps are made for a variety of purposes both theoretical and applied. Science maps have been used for knowledge discovery (Swanson, 1987), research evaluation (Healey, Rothman, & Hoch, 1985; Noyons, Moed, & Raan, 1999), as a technique for managing the information explosion (Borner, Chen, & Boyack, 2005), and as part of the "science of science" (Small & Garfield, 1985; Small & Griffiths, 1974). In addition, science maps may be used to enhance the dialog between various scientific disciplines, and to explore the interfaces between related fields of knowledge (Small, 1997).

A variety of techniques has been used for science mapping. The oldest techniques are based on co-citation analysis (Bauin, 1986; Marshakova, 1973; Small, 1973). Co-citation analysis involves measuring the relatedness of documents based on shared citation patterns. A different category of techniques is based on co-word analysis. Co-word analysis is based on judging the semantic similarity of research fields. It has its origins in Bauin (1986). Factor analytic techniques such as latent semantic indexing (LSI) are dominant for co-word analysis (Borner, Chen, & Boyack, 2003; Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990). In order to identify where and how uncertainty is discussed, co-word analysis is appropriate, for it reveals the latent semantics.

1.3.2 How can uncertainties be classified and analyzed for model-based decision support?

Uncertainty is increasingly important and controversial in a variety of domains, such as environmental science (Van der Sluijs, 2007), water management (Pahl-Wostl, et al., 2007), and transport planning (Marchau, Walker, & van Duin, 2009). The W&H framework was put forward in order to offer a common basis for uncertainty in model-based policy analysis across these domains and has been applied in these various domains. However, these applications have resulted in numerous modifications and changes, which were motivated on the one hand by application domain related issues, and on the other hand by perceived weaknesses and problems with the original framework. Now, a wide variety of frameworks, all derived from the original framework, are available. This is inconsistent with the purpose and intention of the original framework.

Instead of integration and harmonization, we observe a new proliferation of various typologies, taxonomies, and frameworks for analyzing uncertainty. For many reasons, it would be desirable to have a generally agreed upon framework for analyzing and communicating uncertainties (Walker, et al., 2003). This question aims at assessing the variety of frameworks that has emerged and synthesizing them into a new and improved common framework for uncertainty analysis in model-based policy analysis. In order to achieve this aim, we performed an extensive literature review, based on citation analysis (De Sola Price, 1963, 1965; Small, 1973, 1995, 2003). From this literature review, we extracted the key problems and criticisms. In light of this, we modified the framework.

1.3.3 How can the uncertainties currently prevalent in airport strategic planning be treated better?

Uncertainties pose a significant challenge to infrastructure planning in general and ASP in specific. The dominant approach in infrastructure planning was to ignore the uncertainties or to try and reduce them (Dempsey, et al., 1997; Marchau, et al., 2009; McDaniel & Driebe, 2005; Quade, 1982; Van Geenhuizen, Reggiani, & Rietveld, 2007; van Geenhuizen & Thissen, 2007). However, such approaches suffer from the problem that they focus on those uncertainties that are "among the least of our worries; their effects are swamped by uncertainties about the state of the world and human factors for which we know absolutely nothing about probability distributions and little more about the possible outcomes" (Quade, 1982). Similarly, Goodwin and Wright (2010) (p. 355) demonstrate that "all the extant forecasting methods – including the use of expert judgment, statistical forecasting, Delphi and prediction markets – contain fundamental weaknesses." And Popper, et al. (2009) state that the traditional methods "all founder on the same shoals: an inability to grapple with the long-term's multiplicity of plausible futures." In response to this, various new planning approaches have been put forward (e.g. de Neufville, 2000, 2003; Dewar, 2002; Dewar, Builder, Hix, & Levin, 1993; Holling, 1978; Lempert, 2002; Lempert, et al., 2003; Walker, et al., 2001). These approaches emphasize the need for a more thorough analysis of the uncertainties, and suggest that flexibility and adaptability are of central importance in coping with the uncertainties.

The starting point for this research was that AMP as the current approach to ASP is inadequate in light of the many diverse and severe uncertainties airports are facing. More specifically, AMP is inadequate for ASP because the resulting plan is not robust with respect to future developments. This lack of robustness is the result of the fact that (a) very few uncertainties are addressed – usually only aviation demand uncertainties(b) only a single trend based extrapolation future for future demand is considered, instead of a range of plausible demands including demands based on trend breaks; and (c) a Master Plan is static in that it does not have any mechanisms in place for reacting to changing conditions. In order to develop an alternative approach, the first step is the identification of alternative treatments for uncertainty in ASP in the airport planning literature through literature review. Next, these alternatives are to be compared qualitatively on several criteria that follow from the problem statement, such as

- the planning approach should consider many different types of uncertainties, in addition to demand uncertainties;

- the planning approach should consider many different plausible futures; including these with trend breaks
- the resulting plan should be robust across the different futures;
- the resulting plan should be flexible

The results of this comparison were analyzed in order to assess whether it might be possible to design an improved approach for ASP by combining ideas from these alternatives. This comparison showed that a synthesis of the approaches was possible and this was pursued.

1.3.4 How can the efficacy of the new approach for airport strategic planning be assessed?

Public decisions on infrastructure have a profound impact on society. They are expensive, shape traveling patterns, influence economic development etc.. The careful planning of these infrastructures is therefore of great importance. However, the new planning approaches that have been put forward for handling these uncertainties, have seen limited application (Hansman, Magee, De Neufville, Robins, & Roos, 2006). One reason for this is that the validity of these new planning approaches has not been explored in depth (Hansman, et al., 2006). Given the profound importance of infrastructure to society, the careful assessment of the efficacy of new planning approaches is needed.

There is a gap in the literature on how to evaluate planning approaches. Hansman et al. (2006) highlight the importance of testing new infrastructure planning approaches, but immediately suggest controlled real world application. In light of the importance of infrastructure to society and the risks of failure, the soundness of this suggestion is put into question. Furthermore, as Dewar et al. (1993) highlight, there is a methodological problem with controlled real world applications: in general, it is not possible to implement both a traditional plan and a plan based on a new planning approach in order to compare the performance of both plans. In light of this, we conclude that there is currently no established approach for evaluating the efficacy of new infrastructure planning approaches.

In order to develop an approach for evaluating the efficacy of new infrastructure planning approaches, a literature review on evaluating techniques in various disciplines was carried out. Of particular interest for this literature review were disciplines related to infrastructure planning, such as design disciplines. There exists a literature on validating design methods. This literature focuses on developing formal approaches for assessing new design approaches. Another body of literature that was useful was the literature on assessing the efficacy of new medical treatments. Here too, new treatments need to be rigorously assessed before applying them in practice. We argue that these bodies of literature can meaningfully inform the thinking about how to assess the efficacy of new infrastructure planning approaches. So, in order to develop a methodology for evaluating planning approaches, a literature review was carried out, focused on bodies of literature that are facing similar questions about establishing the efficacy of approaches or treatments. The reviewed literature was analyzed in order to construct a methodology.

1.3.5 Is the proposed approach for the treatment of uncertainty in airport strategic planning indeed preferable to airport master planning?

As argued in answering research question 4, there is a range of techniques and methods available for gathering evidence about the efficacy of new infrastructure planning approaches in general. One of these techniques relies on the use of models as a substitute for the real world combined with Exploratory Modeling and Analysis (Agusdinata, 2008; Bankes, 1993).

EMA is a research methodology to analyze complex and uncertain systems (Agusdinata, 2008; Bankes, 1993). It can be contrasted with a *consolidative modeling* approach, in which all the existing knowledge about a system is consolidated in a model that is subsequently used as a surrogate for the real world system (Hodges, 1991; Hodges & Dewar, 1992). The consolidative approach is valid only when there is sufficient knowledge at the appropriate level and of adequate quality available. When dealing with long-term infrastructure planning, these conditions are not met, so using such a consolidative approach might produce erroneous results (Dewar & Wachs, 2006; Marchau, et al., 2009; Van Geenhuizen, et al., 2007). However, in such situations there still is a wealth of knowledge and information available that supports a set of structurally different models across a range of parameter values. EMA aims at offering support for exploring this set of models across the range of plausible parameter values and drawing valid inferences from this exploration (Agusdinata, 2008; Bankes, 1993). In the context of EMA, a computer model can be used as a platform for computational experiments, as lab equipment that maps specific inputs into output about system behavior (Bankes, 2009; Bankes, Lempert, & Popper, 2002). Using models as lab equipment has implications for model design: models need to be modular so that a variety of hypotheses about system structure can be implemented, tested, and compared (Bankes, 2009).

The basic approach for testing a new approach to ASP, using EMA was: (i) develop a fast and simple model of an airport; (ii) generate an ensemble of future worlds; (iii) specify a traditional Master Plan and an Adaptive Plan based on the new approach; and (iv) calculate and compare the efficacy of both plans across an ensemble of future worlds using the fast and simple model. By comparing the performance of a traditional Master Plan and the Adaptive Plan as calculated by the fast and simple model, one can reason about how these plans would behave in the real world. However, given that ASP is decisionmaking about the future, there are significant uncertainties present. There is simply not enough knowledge to accurately forecast the future, the models that are used in ASP are often contested, and there are a variety of value-systems involved from the different stakeholders that are also bound to change over time. The performance of the two plans derived needs to be assessed across these uncertainties. Therefore, when using simulation models for assessing the efficacy of the new approach to ASP, EMA is an appropriate method for EMA can be used exactly for that purpose.

1.4 Outline of the thesis

The body of this thesis consists of five papers that have been submitted, accepted, are forthcoming, or have already appeared in a scientific peer-reviewed journal. The thesis closes with a conclusions and reflections chapter. Each of the papers addresses one of the research questions. As a result, there is some overlap in content between the various papers. Generally, the content of an earlier paper reappears in summarized form in the later papers. Below, we provide a brief summary of each paper.

- KWAKKEL, J.H. & CUNNINGHAM, S.W. (2009) Managing Polysemy and Synonymy in Science Mapping using the Mixtures of Factor Analyzers Model. *Journal of the American Society for Information Science and Technology*, 60(10), 2064-2078.

One way of elucidating what is meant with uncertainty is to reveal the latent semantic structure of the scientific literature on uncertainty. This paper presents a new method for mapping the latent semantic structure of science and applies this method to the scientific literature of 2006 on uncertainty. This dataset is chosen because it was the most recent complete year in ISI and contains over 12000 abstract. The dataset was expected to be heterogeneous with respect to word use. Therefore, a new method was needed, since the existing techniques for mapping the latent semantics struggle with datasets that are heterogeneous in their word use (Borner, et al., 2003; Deerwester, et al., 1990). It is assumed that the different researchers, working on the same set of research problems, will use the same words for concepts central to their research problems. Therefore, different research fields and disciplines should be identifiable by different words, and the pattern of co-occurring words. However, in natural language, there is quite some diversity as a result of fact that many words have multiple meaning; in addition, the same meaning can be expressed by using different words. Therefore, the traditional factor analytic and cluster analytic techniques that are used for mapping the latent semantics of science are inadequate if such polysemous and synonymous words are present. Instead an alternative model, the mixtures of factor analyzers (MFA) model, is utilized. This model extends the traditional factor analytic model by allowing multiple centroids of the dataset. We argue that this model is structurally better suited to map the semantic structure of science. The model is illustrated by a case study of the uncertainty literature sampled from the ISI Web of Science. The MFA model is applied with the goal of discovering multiple, potentially incommensurate, conceptualizations of uncertainty in the literature. In this way, the MFA model helped in creating understanding of the use of language in science, which can benefit multidisciplinary research, interdisciplinary understanding, and assist in the development of multidisciplinary taxonomies of science. The emphasis in this paper is on the method, which is motivated by the journal where the paper appeared and the fact that methodological innovations were necessary in order to reveal the various contexts where uncertainty is discussed and what words are used in these various contexts for discussing uncertainty.

- KWAKKEL, J.H., WALKER, W.E. & MARCHAU, V.A.W.J. (2010) Classifying and Communicating Uncertainties in Model-Based Policy Analysis. *International Journal of Technology, Policy and Management*, 10(4), 299-315.

Uncertainty is of paramount importance in modern day decisionmaking. In response to this, it has been suggested that policy analysts have to be more careful in communicating the uncertainties that are inherent in policy advice. In order to support policy analysts in identifying uncertainties and communicating these uncertainties to decisionmakers, an uncertainty matrix was proposed by Walker, et al. (2003), which synthesized various taxonomies, frameworks, and typologies of uncertainties from different decision support fields. Since its publication, this framework has been applied to different domains. As a result, the framework has undergone changes resulting in a proliferation of uncertainty frameworks. This proliferation runs counter to the purpose of the original framework. This paper presents an extensive review of the literature that builds on Walker, et al. (2003). In light of this, a synthesis is presented, which can be used to assess and communicate uncertainties in model-based policy analysis studies.

- KWAKKEL, J.H., WALKER, W.E. & MARCHAU, V.A.W.J. (2010) Adaptive Airport Strategic Planning. *European Journal of Transportation and Infrastructure Research*, 10(3), 227-250.

This paper presents an alternative approach to AMP. It presents a literature review of three alternative approaches that have been put forward in the ASP literature. These three approaches share the basic idea that in order to handle uncertainty better in ASP, an adaptive approach that is flexible and over time can adapt to the changing conditions under which an airport most operate is called for. Based on this, it concludes that these approaches are complementary and that it might be worthwhile to combine the three into a new, adaptive approach to ASP. A design that integrates the key ideas from the three alternative approaches is presented and illustrated with a case based on Amsterdam Airport Schiphol.

- KWAKKEL, J.H., PAS, J.W.G.M.V.D. (under review) Evaluation of Infrastructure Planning Approaches: An Analogy with Medicine

This paper discusses the evaluation of new infrastructure planning approaches such as Adaptive Airport Strategic Planning and adaptive policymaking. These new planning approaches have been put forward in response to the challenges of deep uncertainty about the future. However, these approaches up till now have seen little real world applications. One important reason for this lack of application is that the efficacy of these approaches has not been established yet. In turn, this is due to the problem that there is no agreed upon method for proving the efficacy of a new planning approach. In this paper, we will draw an analogy to medical research and development in order to outline a methodology for establishing the efficacy of new infrastructure planning approaches. We discuss how the well-established methodology for evaluating new medical treatments can be adapted to evaluating new infrastructure planning approaches. We illustrate the resulting evaluation methodology by outlining an evaluation strategy for adaptive policymaking. It is concluded that the well-established methodology from medicine can successfully be used to inform the evaluation of infrastructure planning approaches.

- KWAKKEL, J.H., WALKER, W.E. & MARCHAU, V.A.W.J. (under review) Assessing the Efficacy of Adaptive Airport Strategic Planning: Results from Computational Experiments.

In this paper, we apply EMA to assess the efficacy of AASP compared to AMP across a large range of possible futures, for the case of Amsterdam Airport Schiphol. The results show that, given the same uncertainties, the range of outcomes from the Adaptive Plan is smaller than that of the Master Plan. So, the Adaptive Plan will expose an airport less to negative outcomes. Furthermore, in those cases in which the Master Plan produces preferable outcomes, the difference in performance compared to the Adaptive Plan is rather small. Moreover, AMP is better than AASP for only a small range of future conditions. These three findings together suggest that AASP minimizes the downside risk without significantly reducing the upside potential. As such, AASP should be preferred to AMP for ASP.

References

Agusdinata, D. B. (2008). *Exploratory Modeling and Analysis: A promising method to deal with deep uncertainty.* Ph.D. thesis, Delft University of Technology, Delft.

Annema, J. A., & de Jong, M. (2009). Leren van de Geschiedenis van de Toekomst [Learning from the History of the Future]. *Vervoerwetenschap,* June.

Ascher, W. (1978). *Forecasting: An Appraisal for Policy Makers and Planners*. Baltimore: Johns Hopkins University Press.

Aslaksen, I., & Myhr, A. I. (2007). "The worth of a wildflower" Precautionary perspectives on the environmental risk of GMOs. *Ecological Economics,* 60(3), 489-497.

Bankes, S. (1993). Exploratory Modeling for Policy Analysis. *Operations Research,* 4(3), 435-449.

Bankes, S. (2009). Models as Lab Equipment: Science from Computational Experiments. *Computational & Mathematical Organization Theory,* 15(1), 8-10.

Bankes, S., Lempert, R., & Popper, S. (2002). Making Computational Social Science Effective. *Social Science Computer Review,* 20(4), 377-388.

Bauin, S. (1986). Aquaculture: A Field by Bureaucratic Fiat. In M. Callon, J. Law & A. Rip (Eds.), *Mapping the Dynamics of Science and Technology*. London: MacMillan Press Ltd.

Borner, K., Chen, C., & Boyack, K. W. (2003). Visualizing Knowledge Domains. *Annual Review of Information Science and Technology,* 37(1), 179-255.

Borner, K., Chen, C., & Boyack, K. W. (2005). Visualizing Knowledge Domains. *Annual Review of Information Science and Technology,* 37(1), 179-255.

Bryson, J. M. (1995). *Strategic Planning for Public and Nonprofit Organizations*. San Francisco, California: Jossey-Bass.

Burghouwt, G. (2007). *Airline Network Development in Europe and its Implications for Airport Planning*. Burlington, Vermont: Ashgate Publishing Company.

Burghouwt, G., & Huys, M. (2003). Deregulation and the Consequences for Airport Planning in Europe. *DISP,* 154(3), 37-45.

Canadian Press. (2006). Mirabel Airport to be Turned into Amusement Park. Retrieved 24-07-2008, from http://www.ctv.ca/servlet/ArticleNews/story/CTVNews/20060221/mirabel_park_060221?s_name=&no_ads=

Courtney, H. (2001). *20|20 Foresight: Crafting Strategy in an Uncertain World*. Boston: Harvard Business School Press.

de Neufville, R. (2000). Dynamic Strategic Planning for Technology Policy. *International Journal of Technology Management,* 19(3-5), 225-245.

de Neufville, R. (2003). Real Options: Dealing With Uncertainty in Systems Planning and Design. *Integrated Assessment,* 4(1), 26-34.

de Neufville, R., & Odoni, A. (2003). *Airport Systems: Planning, Design, and Management*. New York: McGraw-Hill.

De Sola Price, D. J. (1963). *Little Science, Big Science*. New York: Columbia University Press.

De Sola Price, D. J. (1965). Networks of Scientific Papers. *Science,* 149(3683), 510-515.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, D. (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science,* 41(6), 391-407.

Dempsey, P. S., Goetz, A. R., & Szyliowicz, J. S. (1997). *Denver International Airport: Lessons Learned*. New York: McGraw-Hill.

Dewar, J. A. (2002). *Assumption-Based Planning: A Tool for Reducing Avoidable Surprises*. Cambridge: Cambridge University Press.

Dewar, J. A., Builder, C. H., Hix, W. M., & Levin, M. H. (1993). *Assumption-Based Planning: A Planning Tool for Very Uncertain Times* (Report MR-114-A). Santa Monica: RAND.

Dewar, J. A., & Wachs, M. (2006). Transportation Planning, Climate Change, and Decisionmaking Under Uncertainty *Proceedings of the Workshop Conference on Climate Change and the US Transportation*. Washington, DC: National Academies of Sciences and Engineering.

FAA. (2001). *Forecasting Aviation Activity by Airport*. Washington DC: US Department of Transportation.

FAA. (2005). *Advisory Circular 150/5070-6B: Airport Master Plans*. Washington DC: US Department of Transportation.

Flyvbjerg, B., Bruzelius, N., & Rothengatter, W. (2003). *Megaprojects and Risk: An Anatomy of Ambition*. Cambridge: Cambridge University Press.

Funtowicz, S. O., & Ravetz, J. R. (1990). *Uncertainty and Quality in Science for Policy*. Dordrecht: Kluwer Academic Publishers.

Gillund, F., Kjølber, K., Krayer von Krauss, M. P., & Myhr, A. I. (2008). Do Uncertainty Analyses Reveal Uncertainties? Using the Introduction of DNA Vaccines to Aquaculture as a Case. *Science of the Total Environment,* 407, 185-196.

Goodwin, P., & Wright, G. (2010). The limits of forecasting methods in anticipating rare events. *Technological Forecasting and Social Change,* 77, 355-368.

Hansman, R. J., Magee, C., De Neufville, R., Robins, R., & Roos, D. (2006). Research agenda for an integrated approach to infrastructure planning, design, and management. *International Journal of Critical Infrastructures,* 2(2-3), 146-159.

Healey, P., Rothman, H., & Hoch, P. K. (1985). An Experiment in Science Mapping for Research Planning. *Research Policy,* 15, 233-251.

Helton, J. C. (1994). Treatment of Uncertainty in Performance Assessment for Complex Systems. *Risk Analysis,* 14(4), 483-511.

Hodges, J. S. (1991). Six (or so) Things You Can Do With a Bad Model. *Operations Research,* 39(3), 355-365.

Hodges, J. S., & Dewar, J. A. (1992). *Is it You or Your Model Talking? A Framework for Model Validation* (R-4114-AF/A/OSD). Santa Monica: RAND.

Hoffman, F. O., & Hammonds, J. S. (1994). Propagation of Uncertainty in Risk Assessment: the Need to Distinguish between Uncertainty Due to Lack of Knowledge and Uncertainty Due to Variability. *Risk Analysis,* 14(4), 707-712.

Holling, C. S. (1978). *Adaptive Environmental Assessment and Management*. New York: John Wiley & Sons.

IATA. (2004). *Airport Development Reference Manual*. Montreal: IATA.

ICAO. (1987). *Airport Planning Manual, Part 1, Master Planning*. Montreal, Canada: ICAO.

Kazda, A., & Caves, R. E. (2000). *Airport Design and Operation*: Elsevier Science.

Knight, F. H. (1921). *Risk, Uncertainty and Profit* (1957 ed.). Mineola, New York: Dover Publication.

Krayer von Krauss, M. P. (2005). *Uncertainty in Policy Relevant Sciences.* Ph.D. thesis, Technical University of Denmark, Lyngby

Kwakkel, J. H., Walker, W. E., & Marchau, V. A. W. J. (2007, June 21-23). *Dealing with Uncertainty in Airport Strategic Planning.* Paper presented at the 11th Air Transport Research Society 2007 World Conference, Berkeley.

Lempert, R. J. (2002). A New Decision Sciences for Complex Systems. *Proceedings of the National Academy of Sciences of the United States of America,* 99(3), 7309-7313.

Lempert, R. J., Popper, S., & Bankes, S. (2003). *Shaping the Next One Hundred Years: New Methods for Quantitative, Long Term Policy Analysis* (Report MR-1626-RPC). Santa Monica: RAND.

Marchau, V. A. W. J., Walker, W. E., & van Duin, R. (2009). An adaptive approach to implementing innovative urban transport solutions. *Transport Policy,* 15(6), 405-412.

Marshakova, I. V. (1973). A System of Document Connections Based on References. *Scientific and Technical Information Serial of VINITI,* 6, 3-8.

Mayer, I. S., van Daalen, C., & Bots, P. W. G. (2004). Perspectives on policy analyses: a framework for understanding and design. *International Journal of Technology, Policy and Management,* 4(2), 169-191.

McDaniel, R. R., & Driebe, D. J. (Eds.). (2005). *Uncertainty and Suprise in Complex Systems: Questions on Working the Unexpected*: Springer.

Morgan, M. G., & Henrion, M. (1990). *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge: Cambridge University Press.

Norton, J. P., Brown, J. D., & Mysiak, J. (2006). To What Extent, and How, Might Uncertainty be Defined? Comments Engendered by "Defining Uncertainty: a Conceptual Basis for Uncertainty Management in Model-Based Decision Support" Walker *et al*. Integrated Assessment 4: 1, 2003. *Integrated Assessment,* 6(1), 82-88.

Noyons, E. C. M., Moed, H. F., & Raan, A. F. J. v. (1999). Integrating Research Performance Analysis and Science Mapping. *Scientometrics,* 46(3), 591-604.

Ortúzar, J. d. D., & Willumsen, L. G. (2002). *Modeling Transport* (3nd ed.). Chichester, England: Wiley.

Pahl-Wostl, C., Sendzimir, J., Jeffrey, P., Aerts, J., Berkamp, G., & Cross, K. (2007). Managing change towards adaptive water management through social learning. *Ecology and Society,* 12(2), 30 [online].

Popper, S., Griffin, J., Berrebi, C., Light, T., & Min, E. Y. (2009). *Natural Gas and Israel's Energy Future: A Strategic Analysis Under Conditions of Deep Uncertainty* (TR-747-YSNFF). Santa Monica, California: RAND.

Porter, A. L., Roper, A. T., Mason, T. W., Rossini, F. A., & Banks, J. (1991). *Forecasting and Management of Technology*. New York: John Wiley and Sons.

Quade, E. S. (1982). *Analysis for Public Decisions* (2nd ed.). New York: Elsevier Science Publishing Co.. Inc.

Refsgaard, J. C., van der Sluijs, J. P., Brown, J., & van der Keur, P. (2006). A Framework for Dealing with Uncertainty due to Model Structure Error. *Advances in Water Resources,* 26(11), 1586-1597.

Skeels, M., Bongshin, L., Smith, G., & Robertson, G. (2008). *Revealing Uncertainty for Information Visualization.* Paper presented at the Working Conference on Advanced Visual Interfaces, Napoli, Italy.

Small, H. (1973). Co-Citation in the Scientific Literature: A New Measure of the Relationship Between Two Documents. *Journal of the American Society for Information Science,* 24, 265-269.

Small, H. (1995). *Navigating the Citation Network*. Paper presented at the 58th Annual Meeting of the American Society for Information Science: Forging new Partnerships in Information.

Small, H. (1997). Update on Science Mapping: Creating Large Document Spaces. *Scientometrics,* 38(2), 275-293.

Small, H. (2003). Paradigms, Citations, and Maps of Scence: A Personal History. *Journal of the American Society for Information Science and Technology,* 54(5), 394-399.

Small, H., & Garfield, E. (1985). The Geography of Science: Disciplinary and National Mappings. *Journal of Information Science,* 11, 147-159.

Small, H., & Griffiths, B. C. (1974). The Structure of Scientific Literature 1: Identifying and Graphing Specialties. *Science Studies,* 4, 17-40.

Swanson, D. R. (1987). Two Medical Literatures that are Logically but not Bibliographically Connected. *Journal of the American Society for Information Science,* 38(4), 228-233.

Szyliowics, J. S., & Goetz, A. R. (1995). Getting Realistic About Megaproject Planning: The Case of the New Denver International Airport. *Policy Sciences,* 28(4).

van Asselt, M. B. A. (2000). *Perspectives on Uncertainty and Risk*. Dordrecht: Kluwer Academic Publishers.

Van der Sluijs, J. P. (2007). Uncertainty and Precaution in Environmental Management: Insights from the UPEM conference. *Environmental Modelling & Software,* 22(5), 590-598.

van der Sluis, J. P. (1997). *Anchoring Amid Uncertainty: On the Management Of Uncertainties in Risk Assessment of Anthropogenic Climate Change.* University of Utrecht, Utrecht.

van Eeten, M. J. G. (1998). *Dialogues of the Deaf: Defining New Agendas for Environmental Deadlocks*. Dekft: Eburon.

Van Geenhuizen, M., Reggiani, A., & Rietveld, P. (2007). New Trends in Policymaking for Transport and Regional Network Integration. In M. Van Geenhuizen, A. Reggiani & P. Rietvelt (Eds.), *Policy Analysis of Transport Networks* (pp. 1-16). Alsdershot: Ashgate.

van Geenhuizen, M., & Thissen, W. A. H. (2007). A Framework for Identifying and Qualifying Uncertainty in Policy Making: The Case Of Intelligent Transport Systems. In M. van Geenhuizen, A. Reggiani & P. Rietvelt (Eds.), *Policy Analysis of Transport Networks* (pp. 101-118). Aldershot: Ashgate.

Walker, W. E. (2000). *Uncertainty: The Challenge for Policy Analysis in the 21st Century* (P-8051). Santa Monica: RAND

Walker, W. E., Harremoës, J., Rotmans, J. P., Van der Sluijs, J. P., Van Asselt, M. B. A., Janssen, P. H. M., et al. (2003). Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment,* 4(1), 5-17.

Walker, W. E., Rahman, S. A., & Cave, J. (2001). Adaptive Policies, Policy Analysis, and Policymaking. *European Journal of Operational Research,* 128(2), 282-289.

Yücel, G., & Van Daalen, C. (2009). An Objective-Based Perspective on Assessment of Model-Supported Policy Processes. *Journal of Artificial Societies and Social Simulation,* 12(4), [online].

2 Managing polysemy and synonymy in document collections using the mixtures of factor analyzers model

This chapter has originally been published as KWAKKEL, J.H. & CUNNINGHAM, S.W. (2009) Managing Polysemy and Synonymy in Science Mapping using the Mixtures of Factor Analyzers Model. *Journal of the American Society for Information Science and Technology*, 60(10), 2064-2078.

Abstract

A new method for mapping the semantic structure of science is described. We assume that different researchers, working on the same set of research problems, will use the same words for concepts central to their research problems. Therefore, different research fields and disciplines should be identifiable by different words, and the pattern of co-occurring words. However, in natural language, there is quite some diversity as a result of fact that many words have multiple meaning; in addition, the same meaning can be expressed by using different words. We argue that traditional factor analytic and cluster analytic techniques are inadequate for mapping the semantic structure if such polysemous and synonymous words are present. Instead an alternative model, the mixtures of factor analyzers (MFA) model, is utilized. This model extends the traditional factor analytic model by allowing multiple centroids of the dataset. We argue that this model is structurally better suited to map the semantic structure of science. The model is illustrated by a case study of the uncertainty literature sampled from data from the ISI Web of Science. The MFA model is applied with the goal of discovering multiple, potentially incommensurate, conceptualizations of uncertainty in the literature. In this way, the MFA model can help in creating understanding of the use of language in science, which can benefit multidisciplinary research, interdisciplinary understanding, and assist in the development of multidisciplinary taxonomies of science.

2.1 Introduction

We describe a new approach to map science. We are interested in the way in which different researchers, working in different disciplines and fields use language. Understanding the use of language in science can benefit multidisciplinary research, interdisciplinary understanding, and assist in the development of multidisciplinary taxonomies of science. Synonymy and polysemy present a challenge to effective science mapping. Synonymy occurs when different words mean the same thing; polysemy occurs when the same word means different things. Clearly, both situations occur frequently within scientific language.

In order to overcome the problems posed by polysemy and synonymy for mapping the semantics of science, the mapping should be based on latent semantics (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990). Following Deerwester et al. (1990), we assume that there is a latent semantic structure underlying the data about word usage. We will argue that this hidden semantic structure can be revealed using the mixtures of factor analyzers (MFA) model (Ghahramani & Hinton, 1997). This model simultaneously clusters and factors the data, thus combining the relative merits of these techniques. As a result, different clusters of similar word usage, and the underlying semantics for these clusters are simultaneously derived. As will be discussed below, this procedure can circumvent the problems of polysemy and synonymy. In addition, the MFA model is attractive for mapping the semantics of science, since it has a proven record in reducing high-dimensioned data (McLachlan, Bean, & Peel, 2002; McLachlan & Peel, 2000; McLachlan, Peel, & Bean, 2003).

In this paper we first discuss previous research in the field of science mapping. Next, we discuss in more detail the MFA model. Third, we will compare the MFA model with clustering and factoring, and illustrate the MFA model using data sampled from the ISI web of science pertaining to the uncertainty literature. We close with several directions for further research, for both the application of science mapping to the uncertainty literature, and the development of methods and techniques for science mapping.

2.2 Previous research

Science maps "attempt to find representations of the intellectual connections within the dynamically changing system of scientific knowledge (Small, 1997)." Such maps are made for a variety of purposes both theoretical and applied. Science maps have been used for knowledge discovery (Swanson, 1987), research evaluation (Healey, Rothman, & Hoch, 1985; Noyons, Moed, & Raan, 1999), as a technique for managing the information explosion (Borner, Chen, & Boyack, 2005), and as part of the "science of science" (Small & Garfield, 1985; Small & Griffiths, 1974). In addition, science maps may be used to enhance the dialog between various scientific disciplines, and to explore the interfaces between related fields of knowledge (Small, 1997, p. 276). This review takes a closer look at previous research in science mapping, the most frequently used representational forms, challenges of validity, and the mapping of semantics. The review suggests that science mapping is an appropriate technique for creating multidisciplinary taxonomies of science; the field is in need of a new class of models which are approachable by non-experts while also offering a high fidelity representation of scientific activity; and the problems of synonymy and polysemy might be best overcome by combining factor analysis and cluster analysis within a unified approach.

2.2.1 Science maps and interdisciplinary taxonomies of science

A variety of techniques have been used for science mapping. The oldest techniques are based on co-citation analysis (Bauin, 1986; Marshakova, 1973; Small, 1973). Co-citation analysis involves measuring the relatedness of documents based on shared citation patterns. A newer technique which is conceptually related to co-citation analysis is the use of hyperlinks for science mapping.

A different category of techniques is based on co-word analysis. Co-word analysis is based on judging the semantic similarity of research fields. The use of co-word analysis for science mapping has its origins in a 1986 book on science studies (Bauin, 1986). Co-word analysis has been extensively applied; the work by Peters and van Raan is exemplary (Peters & van Raan, 1993a, 1993b). Co-word and co-citation analysis are not mutually exclusive. For example, Braam et al. (1991a, 1991b) combine co-word and co-citation techniques to develop comprehensive maps of science.

McCain (McCain, 1989) provides an early and authoritative comparison and contrast between the co-citation and co-word techniques. The author carefully explicates the underlying assumptions of each technique, and further suggests that the two techniques are actually measuring qualitatively different kinds of similarities between documents. McCain (McCain, 1989) therefore concludes that co-word and co-citation techniques are best used for different, although complementary purposes.

2.2.2 Representational forms

Regardless of whether co-citation or co-word analysis is used, a wide range of representational forms are available for analysts in creating maps of science. The review to follow focuses most carefully on factoring and clustering techniques in bibliometrics and in information science. The words "factoring" and "clustering" are used to informally designate a wide variety of related models. Other reviews have described factoring models as "positional" and clustering models as "relational." This has the merit of not confusing technical words with informal designations of model families. Nonetheless in this paper we adopt the factoring and clustering designation, as these two approaches are unified in the model to follow. Several good surveys of mapping techniques for the field of information retrieval and data visualization are available in the literature (Borner, et al., 2005; Lin, 1997). These surveys include factoring and clustering, but also consider a wider variety of models.

Factoring techniques have been most extensively used in co-word analysis. Greenacre (1984) suggests this practice stemmed from structural linguistics, and its use of correspondence analysis. French semioticians informed both structural linguistics, and the field of actor network analysis (Callon, Law, & Rip, 1986). Two techniques within the broad category of factor models are multidimensional scaling and matrix decomposition. These techniques all emphasize the role of data reduction in the representation of co-word data. Multidimensional scaling is now the technique most commonly used within co-word analysis (Bauin, 1986; Noyons, et al., 1999). Techniques of matrix decomposition, such as singular value decomposition, are also seeing significant usage (Borner, et al., 2005). One significant subclass of these models are known as latent semantic indexing or latent semantic analysis (Borner, et al., 2005; Deerwester, et al., 1990).

Clustering has been used most extensively in co-citation analysis (Small & Griffiths, 1974). Analysis techniques within the broad category of clustering include techniques such as single link clustering and k-means analysis. The main purpose of clustering techniques is the classification of objects into meaningful sets by forming groups of highly similar entities. One weakness of cluster analytic techniques (with the exception of mixture models) is that they do not allow for cross classification. Entities are allocated to only one specific cluster. For example, if one has a dataset with two distinct groups of citations, and a couple of articles that share a number of citations with both groups, how are these articles to be clustered?

Increasingly, both clustering and factoring models are being used in tandem to create more comprehensive and comprehensible models of science (Borner, et al., 2005; Small, 1997, 1998, 1999). One example is the generative topographic map (Bishop, Svensen, & Wiliams, 1998). The GTM is a constrained clustering model. The full data set is presumed to be well-structured such that key features of the data can be captured using a small set of latent features. These latent features may vary gradually and continuously throughout the full data set, creating a non-linear manifold through the data. Clusters are constrained to fall in a grid-like pattern along this manifold. Clusters are needed for the inevitable situation of a real point of data falling a significant distance away from the latent manifold. Each cluster is associated with a common noise model, which serves as a distance metric by which the noisy data can be mapped to a unique location on the latent manifold.

GTM has been used for data visualization purposes. The non-linear character of GTM provides a technique for exposing otherwise hard-to-find features in the data for closer inspection. Likewise, the low-dimensional character of GTM makes it highly suitable for visualization purposes. GTM first emerged in the bibliometrics community in the form of self-organizing maps. Self-organizing maps are a heuristic model of neural computing which predates GTM (Kohonen, 1995). Because self-organizing maps lack an explicit model of noise in the data they provide no metric for comparison within or across model formulations (Bishop, et al., 1998). Nonetheless, self-organizing maps have been previously applied in bibliometrics to co-linking analysis of hyperlink patterns on a collection of Web pages (Faba-Perez, Guerrero-Bote, & Den Moya-Anegon, 2003).

2.2.3 Validity

If science maps are to be useful in creating a shared vision of scientific activity between disparate research fields, special care will be needed in the validation of the maps. Both external and internal measures of validity have been examined in the past (McCain, 1989; Noyons, et al., 1999). Healey et al. (1985) demonstrate the apparent paradox faced by science mappers: if the map adequately represents the field of science as it is known to experts, then it is uninteresting; if the map departs from the perspective of the experts, then it is questionable. Hicks (1987) is particularly critical of this interface between the internal representation of the data, and its external interpretation and presentation to decision-makers and domain experts. As a result of this challenge to science map interpretation, historically science maps have faced significant challenges when facing external validation.

The best way to meet this external validation challenge may be a heightened effort for internal validity in the maps. Science maps must consistently and uniformly map represented components back to their source data, or will continue to face challenges from experts. Models which provide such maps are also known as generative models. A generative model is a representation, subject to uncertainty about the specific content of the data, of how the data is structured and generated. A generative model enables one to test specific hypotheses about the quality of the data under the model. It also represents a specific and reproducible “map” between the data and its representation in model form.

The GTM is an example of such a generative model. The MFA model is another. These models offer explicit hypotheses about the semantics of scientific output. GTM and MFA both make explicit assumptions about the structure of science, while incorporating probabilistic uncertainty given incomplete and possibly noisy input data.

2.2.4 Mapping the latent semantics of science

Words are a partial indicator of scientific content. Semantics (the patterns of words) are an even richer indicator of content. Problems remain with the use of modeling semantics. One such problem involves modeling hierarchical structure. Another problem, and the chief problem addressed in this work, is polysemy. Polysemy is an obvious problem for information scientists because they are dealing with diverse document collections, written by different authors for different purposes and different intended audiences. The assumption of a shared semantic structure across all texts may not be valid.

Co-citation analyses have been used to produce homogeneous collections of data. This relieves some of the problem of polysemy. Yet, interesting questions remain about the shared use of language. In addition, such a perspective prohibits examining interdisciplinary use of language, and the larger scale structure of science. Scientific content interacts in complex ways with disciplinary structures, professional networks, and institutional systems. Our vision of science mapping is the production of uniform measures of knowledge so that these extended questions can be pursued from a firm footing of content.

Our purpose in this paper is to examine scientific language as a measure of scientific content. The current problem with this approach, as we see it, is polysemy. Thus our paper strives to incorporate heretofore unexplored methods for dealing with polysemy. Equivalent analyses of science maps, starting from a basis of citation analyses, can and should be produced. Previous attempts to use citation data to address polysemy have been, in our opinion, overly limited in the generality of the results.

Traditional co-word analysis relies predominantly on factor analytic techniques. Underlying this is the assumption that there is some latent, hidden, structure to the data. Through factor analytic techniques this latent structure can be revealed. A well-known example of this line of thinking, from the field of document retrieval, is latent semantic indexing (Deerwester, et al., 1990). Deerwester et al. (1990) postulate that there is a latent semantic structure to data about word usage that can be revealed through singular value decomposition. This latent semantic structure can be used for information retrieval purposes in order to improve the effectiveness of a search. Instead of searching directly against the documents, the search uses the latent semantics and returns those articles that score higher on the latent semantics associated with the specific search terms.

Latent semantic indexing is in spirit very close to the research reported on in this paper. Deerwester et al. (1990) are primarily occupied with the problems of synonymy and polysemy for document retrieval, whereas we are interested in these problems for science mapping. Synonymy is loosely defined as the fact that there are many ways to refer to the same object. Polysemy is understood as the fact that people use the same term but it takes on a different meaning in different contexts. Latent semantic indexing can address the problem of synonymy nicely, but “it offers only a partial solution to the polysemy problem” (Deerwester, et al., 1990, p. 405), the reason being that differences in meaning are averaged out by singular value decomposition. From this, we conclude that factor analytic techniques are suited to identify the latent semantic structure of science and can address the problem of synonymy; however they fail to address the problem of polysemy.

The use of cluster analytic techniques for the identification of latent semantics is more problematic, since clustering approaches are considered to be inadequate to capture the rich semantics of most datasets (Deerwester, et al., 1990). However, cluster analytic techniques do offer a solution to the problem of polysemy. Polysemous words can be identified by their co-occurrence with two different sets of words. If one wants to cluster articles that contain polysemous words, what is of relevance is the set of other words the article also uses. Articles that use words from one specific set and the polysemous word will thus be clustered.

An alternative to mapping the latent semantics of science directly is the work by van Braam et al. (1991a, 1991b) and van den Besselaar and Heimeriks (2006). Braam et al. (1991a, 1991b) use co-citation analysis for the identification of the structure of science and complement this with co-word analysis in order to identify the concepts that are used. Van den Besselaar and Heimeriks (2006), in contrast, look at word-citation co-occurrence. In this way, they mitigate the problems posed by polysemy, synonymy, and noise in word usage by individual scientists. This paper expands on the work of these authors by seeking a purely structural model for managing document heterogeneity.

2.2.5 Conclusion

The literature review suggests that polysemy and synonymy can be handled by combining clustering and factoring, drawing upon their relative strengths. The work of Deerwester et al. (1990) suggests that a factor analytic technique offers a solution to the problem of synonymy and can reveal the latent semantic structure of science. In contrast, cluster analytic techniques are able to sort out polysemous words, but are too simplistic to capture the rich semantics of scientific language. Interestingly, it is increasingly commonplace to combine both clustering and factoring models to create high fidelity models of science. In this paper, we will explore a specific instance of such a combined model that might offer a potential solution to the problems of synonym and polysemy. The chosen model is also a generative model, thus it also addresses the problem of validity.

We draw from the literature review a number of points for further consideration. Our desire to map the semantics of science is entirely consistent with previous research into science maps as a tool for creating consistent metrics of scientific content. Co-word analysis is a technique in science mapping with a strong conceptual foundation in the comparative analysis of semantics and document similarity. The field is challenged by both the internal and external validity of the science mapping approach. Internally consistent models with rigorous assumptions for the comparison of documents are required. In addition, existing co-word analysis techniques (i.e. factoring and clustering) offer only partial solutions to the problems of polysemy and synonymy.

2.3 Method

The literature review presented in the previous sections suggests that science mapping is in need of a new class of models which are approachable by non-experts while still offering a high fidelity representation of scientific activity. Further the review suggests that the problems of synonymy and polysemy might be overcome by a combined method of factor analysis and cluster analysis. In this section, we present a generative model that combines clustering and factoring. We will argue that when this model is applied to co-word data it can overcome the problems of synonymy and polysemy.

The section begins with a qualitative illustration demonstrating the modeling challenges of polysemy and synonymy. We then describe the origins of the MFA model, and justify its use in this context. The formal description of the model begins with the notation and terminology used in the article. The inputs, outputs and equations of the formal model are described. The section concludes with a discussion of the model estimation and identification.

2.3.1 A qualitative illustration of MFA

Before we provide the technical details of the MFA model, we will illustrate how the MFA model works with an example. We begin with a logical demonstration of a situation where clustering or factoring alone would be inadequate. As will be discussed, both clustering and factoring methods have the expressive power to represent the data, but not necessarily in the most parsimonious fashion. An argument is presented regarding the consequences of excessive parameterization of the data. Finally, we show how the MFA model would deal with the same situation that would cause clustering and factoring to fail.

Consider the diagram presented in figure 2-1, below. The figure shows a three-dimensional vector space of documents where each document is assigned a location in space based upon the frequency of occurrence of three words. In this example there are words (represented by w1, w2, w3) and factors (f1, f2) and clusters (c1 and c2). Most documents lie close to the plane represented by combinations of w1 and w2. This plane is shaded in the diagram. As can be seen the document set contains two distinct collections each with its own characteristic term usage. One cluster centroid is represented by the point c1, and the other cluster centroid is represented by c2. Since there are more documents in c1, the shared center c3 is closer to c1 than c2.

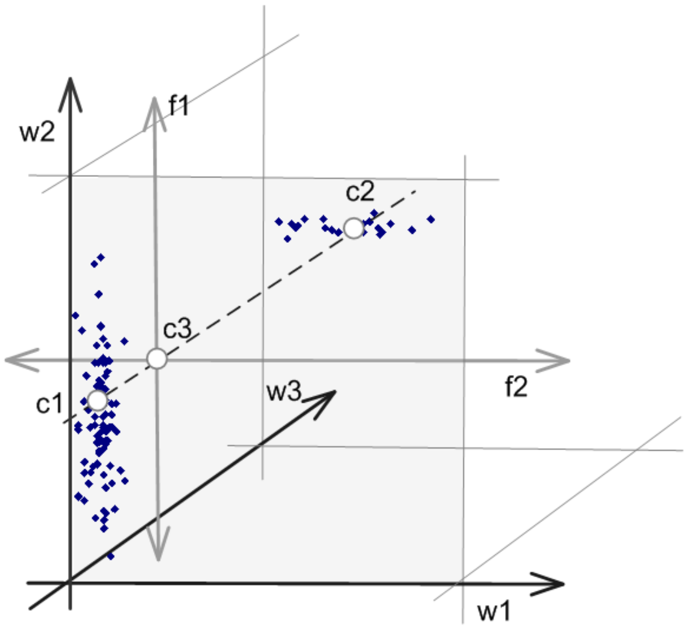


Figure 2-1: Structural challenges to the clustering and factoring model

A factor analytic approach will begin at the center of the data (c3), seeking vectors which efficiently span the document space. A reasonable span would be the factors represented by f1 and f2. The documents in cluster 1 will score highly on factor 1, while the documents in cluster 2 will load highly on factor 2. However, there are several problems with such results. First, although the documents in cluster 1 lie along the f1 dimension and the documents in cluster 2 lie along the f2 dimension, both have a displacement. For cluster 1 there is a small displacement along f2. For cluster 2 there is a high, displacement along the f1 dimension. This displacement must be included in the factor scores of every document in both c1 and c2. It would be far more parsimonious to simply tag the entirety of a cluster with a shared displacement. The resultant inefficiency of a factor solution increases with the number of documents in the collection. Second, the semantics of cluster 2 incorporate entirely new semantics (f2) not common to the other cluster.

A clustering solution has obvious merit in this circumstance. Documents may be tagged with membership in the c1 or c2 clusters respectively. A representation of the data then requires only a vector for the centroid of each of the clusters. Nonetheless in this example neither cluster is positioned far along the w3 space. Word 3 is largely irrelevant to the document semantics; thus there are wasted parameters here. The matter is minor when there are two clusters and only one extraneous dimension. The resultant inefficiency of a clustering solution increases with the number of words in the collection. To put this point differently, the process of clustering can be made more efficient if the dimensions that are not relevant to clustering are eliminated. The problem is how to identify these superfluous dimensions.

Figure 2 below, shows how the MFA model would deal with this specific case. Two clusters consisting of their own distinct semantic dimension efficiently models this space of documents. Cluster 1 focuses on those document rich in semantics f1. Cluster 2 focuses on those documents rich in semantics f2. The characteristic rates of term usage are efficiently captured by the cluster centroids c1 and c2. Overall, the MFA solution (in this manufactured example) is characterized by a great deal of parsimony in the use of parameters.

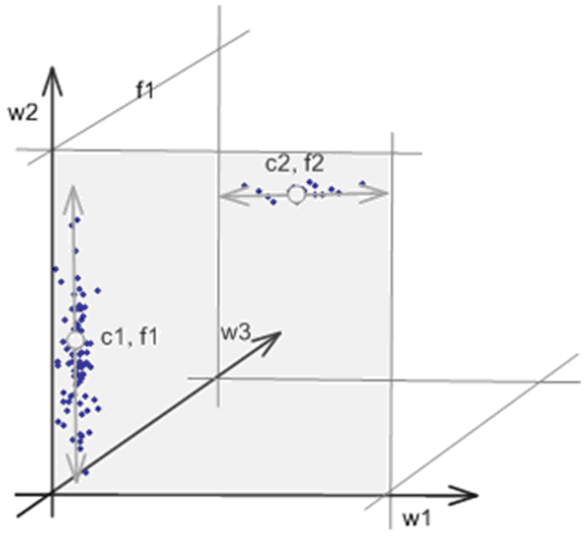


Figure 2-2: Example MFA solution

2.3.2 An explicit model of synonymy and polysemy

At this point, in the context of the MFA model, we may venture an explicit model of synonymy and polysemy. Synonymy occurs whenever two words within a collection are used in an interchangeable manner. Deerwester et al. (1990) describe this as the "latent semantic structure" of the document colleagues. This is a sufficient model of synonymy within a homogeneous collection of documents.

Polysemy occurs whenever there are underlying differences in rates of word usage which cannot be explained in light of the semantics. Neither clustering nor factoring alone will be sufficient since, while clustering can explain underlying differences in word usages, it cannot explain semantics. Likewise factoring can explain semantics, but all differences in word usage are viewed as the sole result of latent semantics. Thus, unlike in the clustering model, excess usages of any word must be accompanied by a similar excess of related words. We hypothesize that this model of polysemy best characterizes a real-world dataset involving uncertainty, which is discussed in more detail further below.

2.3.3 Origins and justification of the method

The mixtures of factor analyzers (MFA) model (Ghahramani & Hinton, 1997) generalizes the factor analytic model to allow multiple centroids of the data set. In contrast, the standard factor analysis procedure starts with the mean-centering of the data. As will be argued MFA is a viable model of the data when there are polysemous and synonymous key words and phrases in the dataset. The MFA model, to our knowledge, has not been previously applied to science mapping. A number of features of the model make it attractive for such an application. The model has a proven record in reducing high-dimensioned data. The joint simultaneous estimation of factors and clusters extends and expands prior work in the field, the form of the model is unusually effective in integrating diverse representations of the data, and it is a generative model that consistently and uniformly maps source data to its components

The MFA model is used in general for exploring high dimensional data sets (McLachlan & Peel, 2000; McLachlan, et al., 2003). Particular successes have been achieved in informatics (Internet data) and bioinformatics (genetic and neurological data). The model has also been extensively used in image recognition tasks including face recognition, image and sound segmentation, and handwriting recognition (Kumagi & Utsugi, 2004; Yamamoto, Nankaku, Miyajima, Tokuda, & Kitamura, 2005; Zhoe & Mobasher, 2006). The model is valued for its capabilities for integrating diverse representations of data into a “single, coherent global coordinate system” (Teh & Roweis, 2002). The combination of factors and clusters allow a more heterogeneous representation of the data than either model used alone.

The MFA model combines factoring and clustering just like the GTM, and therefore the two models share some similarities. In particular, MFA models with a high cluster dimensionality and a low factor dimensionality will approach a GTM model. GTM and MFA are only two among many generative models which integrate features of clustering and factoring. The most appropriate model for science mapping purposes will only become apparent as this family of models are rigorously explored and tested.

2.3.4 Terminology and notation

The following section formally describes the generative model of the data. We first specify our hypothesis about how the data is generated by use of a probability model. Then, a procedure for parameterizing the model given the data is derived from the probability model. Thus, the following discussion involving a generative model is a hypothesis about the data. Confirming or denying the hypothesis requires evaluating the quality of the model against the data. The discussion begins with the hypothesis, describes a strategy for confirming or denying the generative hypothesis, then explores the evidence for the hypothesis using a case.

The formal discussion is adapted from Gharamani and Hinton (1997). We provide additional details of the generative model of the MFA model, which is implied but not explicitly given by this source. Some of the technical language of factor analysis, in naming some of the associated matrices, is borrowed from Harman (1976). First, a short description of the formal model is provided. Then, a more detailed justification of the model is provided following the elements of the generative model as presented in equation 1.

As an interim step, the major quantities of the generative model are named and dimensioned (table 1). Parameters of the mixtures of factor analyzers model include the number of factors (f) and the number of clusters (c) chosen. Characteristics of the data include the number of words used in the indexing (w); later in the discussion, the number of documents (d) will also be referenced. The table provides the symbol, and the informal name of the quantity. The table also includes the nature of the variable (deterministic or stochastic), and the dimensions of the variable. The argument proceeds with the notation, a short summary of the model, and only then is a detailed justification and dimensioning of the variables is provided.

Table 2-1: Components of the mixtures of factor analyzers model

Symbol	Name	Nature	Dimension
X	Document	Stochastic	[1 · w]
U	Uniqueness	Stochastic	[1 · w]
ω	Responsibilities	Stochastic	[1 · c]
z-tilde	Augmented Factor Scores	Deterministic	[(f+1) · c]
Λ-tilde	Augmented Factor Loadings	Deterministic	[(f+1) · w · c]

First, a short summary of the formal model is provided. Consider a vector of words sampled from a single document of a collection. Let this vector be represented by the row vector x, dimensioned 1 x words. The generative model may be repeated to generate content for the entirety of the collection; we need describe only the generation of a single document.

There is a component of word usage which is explainable, and a component which is intrinsically variable in nature. The intrinsic variability of word usage is represented by the variable u, the *uniqueness*. The predictable character of word usage is based upon clusters and factors. Let there be c clusters, indexed by the variable j. Word usage is explained by the sum of a number of clusters. Variable ω, which assigns the document to one or more clusters, are known as the responsibilities. Each cluster is associated with one or more *factor loadings* (Λ). An individual document expresses the latent semantics according to its *factor score* (z). The model, formally described below, combines both factoring and clustering. In special circumstances, the MFA is reducible to either.

$$x = u + \sum_{j=1}^{c} \omega_j \tilde{z}_j \tilde{\Lambda}_j \tag{1}$$

2.3.5 Formal description of the model

The uniquenesses u is a row vector dimensioned 1 x w. The uniquenesses are distributed according to the multivariate normal distribution, with a mean of zero, and a covariance matrix Ψ (equation 2). This model assumes no covariance, and a fixed variance per word. Thus, the matrix Ψ is diagonal.

$$u \sim \mathcal{N}(0, \Psi) \tag{2}$$

The discussion now turns to the structural component of the model. A categorical random variable (ω) is used to select one (and only one) of the clusters to generate a row of the data. The variable ω describes the full allocation of the document to one of the clusters of the model. This variable, called the responsibility, constitutes an index variable, where one element of the row is one, and the remainder of the elements is zero. The values of the variable are indexed as ω_j for each j from 1 to c. The categorical variable is parameterized by the prior probabilities (π) which describe the fixed proportions in which the model draws from each of the clusters (equation 3).

$$\omega \sim \mathscr{C}(\pi) \tag{3}$$

This paper uses a categorical variable to model these responsibilities because the variable is necessary to fully specify a recipe for generating the data. This usage departs from Ghahramani and Hinton (1997) and McLaughlin and Peel (2000) who treat the mixture proportions as fixed quantities. The resultant generative model of these authors describes the "expected average" of the entire set, rather than the model given here which randomly generates an expected document from associated clusters. While conceptually the two models are distinct, estimation of the two models is identical.

It is important to make a distinction between the data generating process, and the model estimation process, when considering the document assignments and responsibilities. While we may say that a document was generated from one and only one cluster, in practice it is often uncertain as to which of several plausible clusters might have generated the document. Thus, a soft assignment is used in estimation. A given document is therefore assigned in mixed proportion to its likelihood of cluster membership.
Each cluster is associated with f factors. Factors consist of a factor score (z) which represents the unique expression of that document along the given factor loadings, and a factor loading (Λ) which is common across all documents in the cluster. Factor scores and factor loadings are matrices (further dimensioned below).

Factor scores are interpreted by Deerwester et al. (1990) as document latent semantics; the authors of this paper also adopt and endorse this concept. The factor scores are distributed according to the multivariate normal, with a mean of zero, and the identity matrix as the covariance. Thus factor scores are orthonormal (Equation 4).

$$z \sim \mathscr{N}(0, I) \tag{4}$$

Each factor is associated with a cluster in the model. Clusters and factors introduce a second and third dimension to the model, in addition to words. The necessary multiplications are fully determined by the summations and indexes in the model. This paper follows one convention for dimensioning document and cluster metadata; others are possible.

The factor is centered about the centroid of that cluster. For the purposes of notation, it is helpful to augment the factor loading and factor score matrices. This useful artifice permits the matrices to contain both the factor loadings and factor origins within a single matrix (Equation 5). The additional element in both matrices is the factor origin. The factor origin (μ) is a vector with dimensions 1 x w for each cluster j.

$$\tilde{z}_j = \begin{bmatrix} z_j & 1 \end{bmatrix}$$

$$\tilde{\Lambda}_j = \begin{bmatrix} \Lambda_j \\ \mu_j \end{bmatrix} \tag{5}$$

The factor scores (Λ_j) are matrices dimensioned (f+1) x w for each cluster j and the factor loadings (z_j) are matrices dimensioned 1 x (f+1) for each cluster j. In equation 1 matrix multiplication of factor scores and factor loadings occurs. The resultant product is a row vector , dimensions 1 x w for each cluster j. This is further summed by clusters (from j = 1 to c) to produce the required index of word usage by document, dimensioned (1 x w). The responsibility ω_j is a scalar, valued either zero or 1. Therefore one and only one of the cluster explanations contributes to the final row vector.

The model given in Equation 1 reduces to cluster analysis when the factor loadings are all zero. The model reduces to factor analysis when there is only a single cluster in the data. The maximum likelihood solution of the model requires that the position of this cluster be at the mean of the data. Variations on MFA (and cluster analysis, and factor analysis) are possible depending on the modeling of the noise and covariance (Ψ). A model involving fewer parameters involves making the noise isotropic (uniform variance across words). Other models allowing a richer covariance structure are also possible.

2.3.6 Model outputs

Concluding this discussion of the model, the outputs of the MFA model are further examined. Parameters of the mixtures of factor analyzers model include the number of factors (f) and the number of clusters (c) chosen. As previously noted, characteristics of the data include the number of documents (d) and the number of words used in the indexing (w). Specific outputs for documents, collections, and words are created by the model. The document classification includes an assignment of the documents to one or more or the clusters. The word classification identifies the latent semantics associated with each word, and also the unique variance associated with each word. Words are polysemous, and may therefore rate multiple latent semantic dimensions. The unique variance shows the residual error associated with modeling each word. There are two components to the collection classification. Every collection is a cluster; every cluster has a centroid representing the mean rate of term usage. Furthermore, every collection has f factors describing the semantics of that cluster. Each document in the collection is scored along the associated semantic space.

The four outputs of the model are summarized in the table below. In Table 2, the analysis is replicated across documents. Two of the quantities vary by document (and therefore bear document in their name); the other two are independent of documents. The document assignment and the document augmented factor scores extend the parameters of the generative model with additional document entries.

Table 2-2: Outputs of the mixtures of factor analyzers model (in this table d are the documents, c are the clusters, f are the factors, and w are the words)

Output name	Output dimensions
Document Responsibilities (ω)	[d · c]
Document Augmented Factor Scores (z)	[d · (f+1) · c]
Augmented Factor Loadings (Λ)	[(f+1) · w · c]
Uniquenesses (u)	[1 · w]

A new, further elaborated semantic model of documents naturally leads to a richer space of outputs for knowledge management. The document responsibilities are the resolution of d draws from the categorical variable (ω). Subject to uncertainty these draws may involve a soft assignment of a given document to multiple clusters. The document responsibilities are useful for dividing the document into separate collections. Uniquenesses may be used for model diagnosis. The document augmented factor scores provide the necessary information to use

MFA for information retrieval. A novel output from the MFA model is the augmented factor loadings. This output places the mean rate of word usage alongside the latent semantics by clusters. Word usage and semantics are conceptually distinct; both are of interest for science mapping. This space is similar to previous “science maps” but is augmented by the recognition that the data occurs in distinct clusters or collections. An example of this output is discussed in greater detail in the case that follows.

2.3.7 Model estimation

Equations 1, 2, and 5 imply a probabilistic hypothesis about the data. The equations project an f-dimensional normal distribution into a higher w-dimensional space through multiplication. Then the equations add a second multivariate distribution, dimensioned (1 by w). The resultant product is distributed according to a multivariate normal. This hypothesis is based upon the propagation equations of normal distributions, and the stated assumptions of the model.

$$X^{\omega} \sim \mathcal{N}(\Lambda z, \Sigma)^{\omega} \quad (6)$$

$$\Sigma = \Lambda^T \Lambda + \Psi$$

We introduce a superscript notation to segment the data by clusters. Recall that the responsibilities (ω) are index variables which assign each document to the appropriate cluster. For a given cluster, the mean of the data is determined by the cluster-appropriate factor scores and loadings (Λz). For a given cluster, the covariance of the distribution is related to the factor loadings. The variance of the distribution is a function of both the factor loadings and the uniqueness. Thus, the resultant density function of the data is described in equation 6.

Model fitting entails finding the maximum likelihood estimates of the parameters (u, π, z, Λ) given the data (X). Evaluating likelihood of the model given the data involves first computing the random values (given the assumed model parameters) which are needed to reproduce the data. The likelihood of these values is then computed under the multivariate normal density function. The full likelihood results from performing this calculation for each cluster, and then summing the total probability of the model across clusters using the appropriate responsibilities (ω) and mixtures (π). There are various model estimation techniques which direct the search for suitable model parameters. The goal of these estimation techniques is to search for parameters in a manner which will improve − or at least not worsen − the model likelihood at each step of the search. We use Ghaharamani and Hinton’s (1997) implementation of the expectation maximization algorithm (Dempster, Laird, & Rubin, 1977) to estimate the MFA model.

Effectively, these likelihood equations entail computing the distance of any given document from a centroid. The distance metric is defined by the factors, and the available centroids are defined by the clusters. Since there is potentially a different distance metric for each cluster, the problem does not readily decompose into a separate factor analysis problem and a separate cluster analysis problem. Degenerate cases do exist where identification of factors and clusters are decomposable, although this is not generally the case.

2.3.8 Model identification and validation

The mixtures of factor analyzers model is a generative model. A generative model contains a deterministic component, which describes structure, and a stochastic component, which acknowledges structural and parametric uncertainty. The deterministic component consists of a set of assumptions which must be confronted given the evidence of real data. The stochastic component consists of a set of probability distributions which are used to *parameterize* the uncertainty inherent in the data.

There are multiple sources of inherent uncertainty in the data. The nature of this uncertainty may be intrinsic variability – there are multiple means available to the author for expressing the same content. Given another chance to write the author may choose to express themselves in several related ways. The nature of the uncertainty might be prejudicial – the content in a database is only a sample of the full knowledge domain available to the researchers. Thus, there can be errors of both precision and recall in the sampling of this literature. The nature of this uncertainty might be epistemic – the actual, complete knowledge of the authors is unavailable to us as we attempt to reconstruct maps of science. Once we have acknowledged that the data is uncertain, it becomes clear that exact reproduction of the data is neither feasible nor desirable.

The actual specification of the data is subject to these ever-present sources of uncertainty. The real data can be viewed as one possible realization out of the many which are possible. Recall that the generative model contains a stochastic component. By using this stochastic component of the model, and by accepting as a given the model assumptions, the analyst can assign a level of confidence that any given instance of the data matches our understanding of the generative process. The confidence that a given result could have been generated by a probability model is measured by likelihood.

The goal of model estimation is to select the most likely parameterization of the model subject to these sources of uncertainty. There are multiple techniques available for model estimation. All these techniques involve maximizing likelihood subject to the structure of the model, and the available parameterization of uncertainty. The algorithm used for estimating the model presented in this paper uses an expectation maximization procedure, which was originally employed as a technique for treating incomplete data. Expectation maximization procedures are hill-climbing algorithms: they are guaranteed over the course of successive steps to increase the likelihood of the model under the available data.

A more extensive model specification does not necessarily lead to a better model. A more elaborate model has more parameters which can be adjusted to fit the data. Extensively specified models are subject to over-fitting. Therefore a higher likelihood does not necessarily indicate a more robust model: the model may become invalidated when more data becomes available. We therefore place a premium on succinct descriptions of the data, which are less likely to be subject to over-fitting. A robust model is a succinct, yet likely description of the data.

There are three perspectives on seeking a robust model. Some researchers use split sample training: they reserve part of the data set for fitting the model, and then evaluate the quality of the model using the remaining sample. This procedure requires a lot of data, or alternatively, relatively few models for testing. Other researchers create a "beauty contest" where alternative models are evaluated using user requirements or expert judgment. This procedure potentially introduces a degree of subjectivity into the definition of robust or effective models, and requires a substantial input of human judgment.

Another perspective on model robustness is to seek metrics of model quality. One such metric of quality is Akaike's Information Criteria (AIC) (Akaike, 1974). This metric rewards high likelihood models, while penalizing models according to the number of free parameters assumed by the model structure. In case of the MFA model, the free parameters are the outputs presented in Table 2. A drawback of AIC is its lack of a theoretical orientation. Nonetheless the metric has been repeatedly proven to assist researchers in selecting robust models subject to uncertainty. The AIC bears more than a passing resemblance to Occam's razor – a heuristic which has served science well for many centuries. Assumptions made about the data should be questioned. As stated earlier, our confidence in model results are conditioned on the correctness of the structure. The structure of the data should be questioned given concerns about over-fitting the data, and a general interest in seeking the best explanation of the data. The stochastic character of the data should also be questioned since we are seeking a likely explanation of the data. This entails minimizing, as much as possible, the stochastic component of the model by increasing ever more effective explanations of the data. The AIC supports the evaluation of competing explanations of the data, allowing us to simultaneously evaluate both the structural and uncertainty issues with the model.

2.4 Case

In order to test the performance of the MFA model, a heterogeneous dataset is needed. That is, the data should contain polysemous and synonymous words sampled from different scientific disciplines. For reasons of other research interests, we have chosen to use a dataset containing articles that discuss uncertainty. This dataset does meet the criterion of being highly heterogeneous for the literature on uncertainty is very diverse. Further, it is interesting to do some exploratory analysis on this data to determine if the collections, as identified by the MFA model, correspond to known scientific disciplines.

One traditional definition of uncertainty is by Knight (1921), who made a distinction between risk and uncertainty. In a situation of risk the probability that a specific event will occur is known. In a situation of uncertainty in contrast the probabilities of a specific event are unknown. Many other definitions, classifications, frameworks, and taxonomies have been proposed for harmonizing (parts of) this literature (e.g. Funtowics & Ravetz, 1990; Morgan & Henrion, 1990; van Asselt, 2000; van der Sluis, 1997). To further complicate the uncertainty literature, different words are used across different scientific disciplines to denote uncertainty, such as doubt, unsureness, ambiguity, imprecision, ignorance, and risk. As a result the uncertainty literature is highly heterogeneous and many relevant words and concepts are polysemous and used differently across the sciences. As such, this dataset present a good test case for the MFA model presented in this paper.

Recently, an attempt has been made to integrate these different frameworks, typologies, and taxonomies in an overarching framework that could provide a starting point for the systematic treatment of uncertainty in decision support (Walker, et al., 2003). This framework of Walker et al.(2003) has been criticized, because it overlooks the diversity of terms and meaning associated with uncertainty and the situation in which these usages arise (Norton, Brown, & Mysiak, 2006). Furthermore the framework is limited because it only explores the modeler's view of uncertainty, overlooking the decisionmakers view of uncertainty (Norton, et al., 2006; van Asselt, 2000). A novel notion that emerged in the uncertainty literature, namely deep uncertainty, can serve as an example of this point. Deep uncertainty is defined as a condition in which the decisionmaker does not know, or multiple decisionmakers cannot agree on the system model, the prior probabilities for the uncertain parameters of the system model and/or the value function (Lempert, Popper, & Bankes, 2002). Central to this definition is the decision makers' perspective on uncertainty, and as such this notion cannot be captured in the framework of Walker et al. (2003). In light of these considerations, it is concluded that the framework does not capture all aspects of uncertainty that are relevant in strategic planning and policy making. By mapping the usage of the stem uncertain across the sciences using the MFA model, an overview of the different distinct semantics within the uncertainty literature can be generated. This overview can help in improving existing taxonomies and typologies, position taxonomies and typologies better, and perhaps support the development of new taxonomies and typologies.

With the case study of the uncertainty literature, we aim to answer the following questions:

- Which model structure − clustering, factoring, or MFA − is preferable for mapping the heterogeneous dataset?
- Given a preferred model structure, what is the preferable parameterization of the model?
- What kind of results does the MFA model generate when applied for science mapping purposes?

2.4.1 Data

The data used for this case comes from the ISI Web of Science database. A search on "uncertain*" for the publication year 2006 was carried out and all records, 12,889 in total, where downloaded as flat text files. These text files where subsequently processed using a Perl script. This script was designed such that it was able to analyze any element of an ISI record; including keywords, titles, and ISI subject categories. All records containing abstracts, 12,603 records in total, were selected, and the single word occurrence per abstract was counted. Stop words where removed based on a stop word list containing 328 stop words. The final step was the printing of the top 100 most frequently occurring non stop words to a text file as input for further analysis. This file contained the words, the sum of occurrences, and a row for each article, identified by the ISI unique article identifier (i.e. the UT tag) with the frequency of the different words per article.

We deliberately choose not to use any stemming. First, we hypothesize that the choice between words such as uncertain, uncertainty, and uncertainties, will reflect a real difference in the conceptualization of uncertainty. In case of stemming, this distinction is lost and we will not be able to test this hypothesis. Second, the MFA model is presented as a model that is equipped to deal with polysemous and synonymous words. If stemming is used, the noise in the data is reduced, and the MFA model will not be tested as ruinously as possible.

2.4.2 Exploratory analysis

Table 3, below, shows the frequency of the 15 most frequently occurring words. Unsurprisingly, the most frequently occurring term is uncertainty.

Table 2-3: Top 15 most frequently occurring words

Word	Frequency
uncertainty	9504
model	8109
data	7214
results	6643
using	5655
used	5032
method	4149
uncertainties	4075
analysis	4051
system	3861
based	3736
patients	3652

Figure 3 shows a histogram of the sum of words per article. This is a partial indicator of the adequacy of the choice of words used in indexing. This histogram shows that the majority of the abstracts, 9486 to be precise, are between 10 and 39. In other words, the top 100 most frequently occurring words in the dataset does result in a reasonable indicator of what words are used by each article.

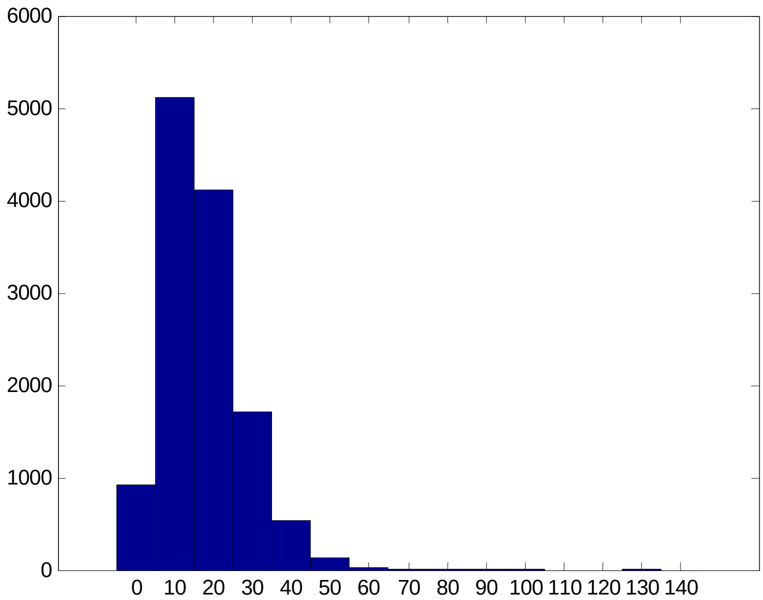


Figure 2-3: Histogram of sum of words indexed per article

2.4.3 Model selection and estimation

In this paper we use a metric approach for model evaluation in the form of Aikake's Information Criterion. First, we compare between different structural models, by contrasting the mixture of factor analysis models with its two simpler, component models: factor analysis, and Gaussian mixture models. Second, we use AIC to determine the best parameterization of the model. The use of AIC in both tasks is motivated by the fact that the diffuse content of knowledge inherent in science mapping makes user evaluation impractical. We aim at comparing three different structural models and for each model multiple different parameterizations. The large set of alternative models which must be formulated given our current understanding of the structure of science limits the viability of split sample testing. The AIC provides confidence that the model that will be chosen, given this data, is robust.

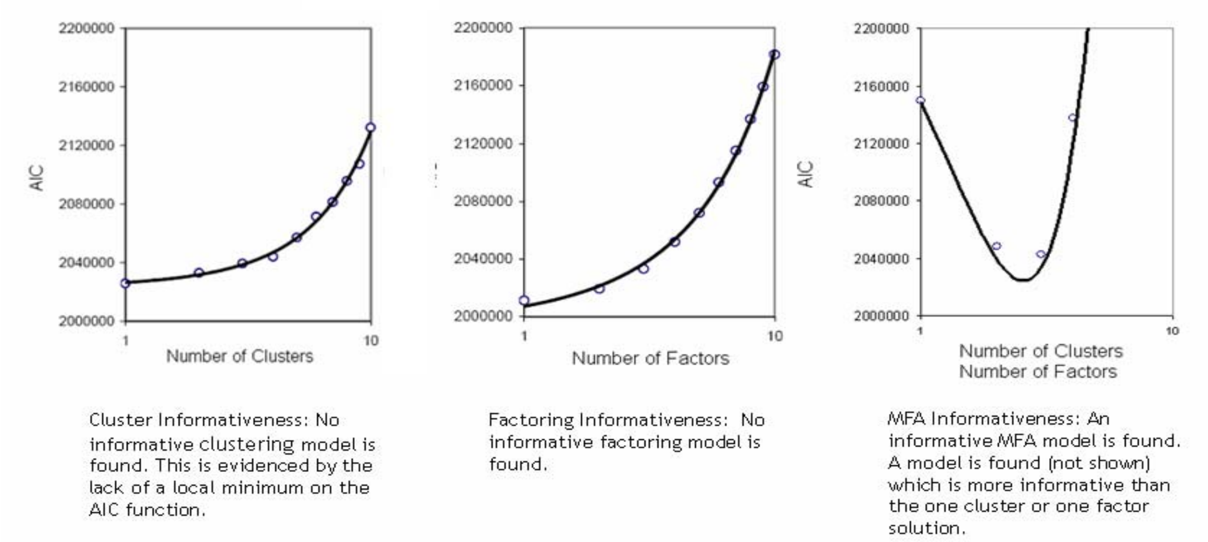


Figure 2-4: AIC comparison plots

Figure 2-4 shows three AIC plots: one for a range of Gaussian mixture models, one for a range of factor analysis models of the data, and one for the MFA model. A logarithmic scaling on the x-axis is used to better display the range in model informativeness across the prospective models. For ease of presentation this last graph examines the subset of MFA models where the same number of clusters and factors are chosen. The cluster AIC plot shows the most informative available model is the lowest dimensioned model. The factor plot also shows that the most informative available model is the lowest dimensioned model. In other words, neither clustering nor factoring can find an informative model. In contrast, the MFA model shows that the maximally informative model is in the space of two or three clusters and factors. A possible explanation for these results is that the data is too noisy or complexly structured for a Gaussian mixture model or factor analysis model taken separately.

A more elaborate exploration of the parameters of the mixtures of factor analyzers is offered. We systematically explored the space between one and ten factors and one and ten clusters. For each combination of clusters and factors, the model was run ten times, each time with a different initialization. An exploration of the solutions revealed that the single best solution in the exploration was a 3 clusters, 2 factors solution with an AIC of 1.839E+06. Since the number of clusters and factors are not identical, this solution is not shown in the graph in figure 4. Next, the average result of each run was used to generate a contour plot as shown below (Figure 2-5). We used the average, instead of the best run, because this results in a more robust solution that can be replicated more easily. The center of gravity or lowest point, is somewhere near the 3 cluster 3 factor solution, hence this combination is chosen for further analysis. From Figure 2-4 and Figure 2-5, we conclude that the combined factor and cluster representation is more informative than either the Gaussian mixture model or the factor analysis model.

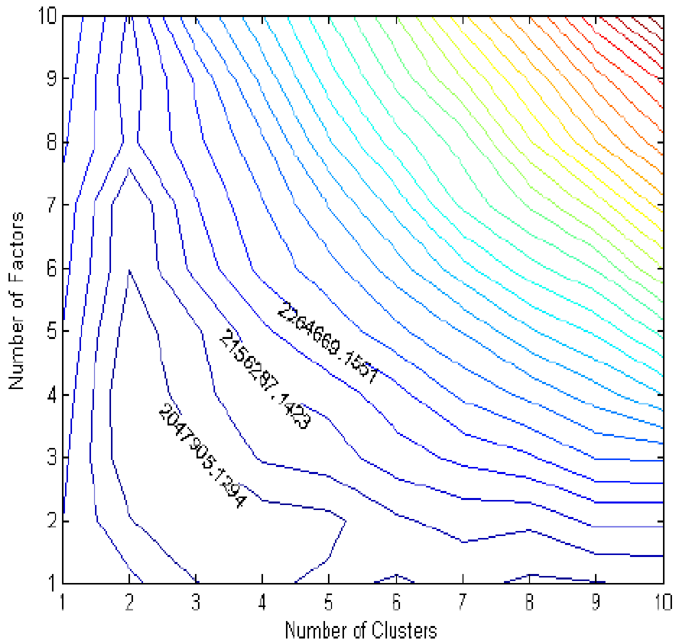


Figure 2-5: AIC contour plot

Use of the AIC metric provides a number of evaluative insights into the model. A comparison of alternative parameterizations of the model demonstrates that a robust, low-dimensional representation of this data is possible using a MFA model. The null hypothesis is that this data is randomly structured. The AIC metric shows that there is clearly a structure to the data. The lower AIC of the MFA compared to a Gaussian mixture model and a factor analysis model show that the stochastic component of the data is smaller in case of MFA model. This provides evidence for the claim that the added model complexity of the MFA model is warranted. Our assumptions about the heterogeneous character of the data are not falsified.

2.4.4 Results

A three cluster-three factor solution was extracted. The clustering produces a “soft” assignment of documents to clusters. Most documents are assigned exclusively to one cluster and not another, although about 20% of the documents do fall between multiple clusters. Hereafter each of the three major document sets are labeled as cluster I, cluster II and cluster III.

The ISI categorization system can be used for descriptive purposes to categorize the contents of each of the clusters. Table 4 shows the leading ISI categories for the documents in the dataset. Each of the ISI categories is assigned to one or more clusters. Some ISI categories mix between one or more of the clusters. The table also displays the intersection between the clusters, showing the ISI categories with a mixed assignment. Almost 80% of the sample fit cleanly into one of the three main clusters while 20% has a mixed assignment.

Table 2-4: ISI categorization of clusters

	Cluster I	Cluster II	Cluster III
Cluster I	55.6% Natural Sciences Environmental Sciences Astronomy & Astrophysics Meteorology & Atmospheric Sciences		
Cluster II	5.8% Environmental Sciences Engineering, Electrical & Electronic Engineering, Civil	14.5% Applied Sciences Engineering, Electrical & Electronic Automation & Control Systems Environmental Sciences	
Cluster III	13.7% Environmental Sciences Public, Environmental & Occupational Health Medicine, General & Internal	1.1% Environmental Sciences Public, Environmental & Occupational Health Radiology, Nuclear Medicine & Medical Imaging	9.7% Medicine Oncology Medicine, General & Internal Clinical Neurology

The three factors associated with each cluster are examined in greater detail in the table below (Table 5). The table shows the three clusters (i.e. natural sciences, applied sciences, and medicine). Each cluster has three factors extracted by the MFA models. These factors, the latent semantics, are described as "fields" in reference to an oft-used unit of scientific organization and indicate different meanings of the stem uncertain as present within each cluster. Further, the alignment of ISI categories with MFA clusters suggests that there may be disciplinary causes for the different vocabularies evidenced in the data. We briefly return to this point in the conclusions.

Representative words for each of the fields were selected using the following procedure. For each word, we identified the factor on which it has the highest factor loading, and the word was assigned to that factor. This procedure resulted in some factors with few words. The overall factor loadings for these factors were reviewed and additional words were added (these words are shown in italics in the table). The table below shows the heavily loading terms for each factor, the additional words for some factors are shown in italic. In light of the heavily loading terms, we assigned the names to the factors. A more detailed exploration of this table can be found in Kwakkel and Cunningham (2008).

Table 2-5: Clusters and factors

Cluster	Factors	Heavily loading terms
Natural Sciences	Engineered Systems	function, method, system, control, proposed, systems problem, given, fuzzy, state, linear, measurement, standard
	Natural Systems	species, uncertain control, management, decision
	Models	model, data, results, using, used, uncertainties, based models, methods, parameters, large, conditions, developed, distribution, estimates, error, surface, parameter, estimate, study, effects, important, significant, changes, response, change, time, uncertainty, analysis, different, use, water, number, process, simulation, set, approach, performance
Applied Sciences	Risk Assessment	management, robust, design assessment, risk
	Statistical Mechanics	rate, energy, theories set, systems, linear
	Policy Evaluation	values, high, temperature, low, studies, treatment, effect, research, evidence, quality, levels, risk, potential, field, assessment, information, decision, measurements, measured
Medicine	Diagnosis	similar, obtained, range, observed, total, cases, patients, compared, p, group
	Longitudinal research	value, mean, years, factors, clinical, estimated, disease, associated
	(quasi) Natural experimentation	case, order, experimental, mass

We deliberately choose not to use any stemming and hypothesized that the choice between words such as uncertain, uncertainty, and uncertainties, will reflect a real difference in the conceptualization of uncertainty. Now that the MFA model is specified we can return to this hypothesis. Table 6 shows the correlations between the factor loadings of the words “uncertain,” “uncertainty,” and “uncertain” within each cluster. From this table, we conclude that there are clear semantic differences between these three words. In cluster one, “uncertainty” and “uncertainties” are highly correlated while they have a strong negative correlation with “uncertain.” In cluster two, “uncertainty” in contrast is negatively correlated with “uncertainties” and “uncertain.” Cluster three is similar to cluster one. Our choice not to apply stemming appears to be justified in light of this. In addition, not only does this show that there are semantic differences between these three words, the fact that the semantic difference differs from one cluster to another reaffirms that there are clear semantic differences between the three clusters.

Table 2-6: Correlations between factor loadings within clusters

		Uncertainty	Uncertainties	Uncertain
Cluster 1	**Uncertainty**	1.000	0.998	-1.000
	Uncertainties	0.998	1.000	-1.000
	Uncertain	-1.000	-1.000	1.000
Cluster 2	**Uncertainty**	1.000	-0.192	-0.494
	Uncertainties	-0.192	1.000	0.948
	Uncertain	-0.494	0.948	1.000
Cluster 3	**Uncertainty**	1.000	0.882	-0.994
	Uncertainties	0.882	1.000	-0.929
	Uncertain	-0.994	-0.929	1.000

2.4.5 Discussion

The sample is dominated by the natural sciences literature. However, the applied science cluster and medical cluster also represent substantial, and semantically distinct, views on uncertainty. The natural science cluster contains three distinct fields of uncertainty research. The field of engineered systems is primarily occupied with the systems control. In contrast, researchers working in the field of natural systems are occupied with complexity and dynamics and how this affects the potential for control and steering. The third field works on the development of models and is occupied with model based uncertainty.

The applied sciences cluster is further specified by the fields of risk assessment, statistical mechanics, and policy evaluation. The field of risk assessment involves the applied discipline of estimating the risk of accident or failure in a technical environment. The field of statistical mechanics is occupied with the application of statistical and nonlinear mathematics. The field of policy evaluation focuses on the assessment of human interventions. Uncertainty is used here mainly to denote uncertainty about the relationships between the intervention and the effects and uncertainty about what to measure. The uncertainty fields within the applied sciences are quite distinct, and show little overlap with the fields of uncertainty seen in the medical and natural sciences.

The medicine cluster shows a third distinct semantics. It consists of the field diagnosis, longitudinal research, and (quasi) natural experiments. The diagnosis field works on understanding the pathology of diseases, how diseases can be identified, and how it should be treated. The longitudinal research field focuses on longitudinal case studies with many cases. In contrast, the third field focuses on few and more exploratory case studies.

In this case, we explored only the usage of terms consisting of one word. For a more in depth insight into the uncertainty literature, the analysis of terms consisting of more than one word would be of great relevance. For example, we hypothesize that the names of tools and techniques for the treatment of uncertainty consist of more than one word. An exploration of these terms, using the result presented here (i.e. cluster allocation of articles and the factor loadings) would allow for the identification of methods and techniques used for the treatment of the different notions of uncertainty. In this way the leading methods for the analysis and treatment of uncertainty for each semantically distinct view of uncertainty could be identified.

The data reveal substantial differences in the meaning of key terms including risk and uncertainty. We argue that this reveals real differences in the way the scientific discourse differs across fields of science. Furthermore, the comparison between the MFA model, the factor analysis model, and the Gaussian mixture model revealed that both the factor analysis model and the Gaussian mixture model are unsuitable for representing this dataset. In particular, a mixture of different representations is demonstrated to be a more informative model for analysis than either factoring or clustering used in isolation.

2.5 Conclusions and further research

This work utilized a model which jointly estimated a model of clustering and factoring. However both the clustering and the factoring model used in the MFA model assume normally distributed data. Bensman (2005) highlights the growing understanding in information science that other, non-normal distributions such as the Poisson, gamma, and negative binomial distribution are often better representations of underlying processes. This paper uses the normal distribution as an approximation to the underlying probability processes. The model should be revised to incorporate Poisson or a negative binomial model, which as Coleman (1964) argues are sound representations of contagion and information spread in society.

Additional evaluation of the mixtures of factor analyzers model is needed. In this paper, we presented the MFA model for dealing with heterogeneous document collections containing synonymous and polysemous words. We used the MFA for mapping the semantics of science and illustrated this with semantics of the uncertainty literature. The current model should be confronted with alternative sets of data. In addition a comparative analysis of the model with other competing explanations of the data is needed. Given these evaluative results, competing explanations of the data should be at least as structurally rich as the mixtures of factor analysis model. Also, for comparative purposes, new research should adopt a framework utilizing likelihood and the AIC. Two challenges are present: there are multiple possible ways of combining clustering and factoring models. Only a few of these combinations have been systematically documented; fewer still have been applied to science mapping. Richer structures, regardless of whether they are cluster or factor-based, should be explored. A second challenge lies in the fact that there are multiple metrics for model comparison: this limits the capacity to compare results between studies.

The case of the uncertainty literature used in this paper demonstrated a real-world, highly heterogeneous collection of documents. The MFA assists in disaggregating the literature into a set of semantically homogeneous collections. The results are suggestive: these collections may correspond to the respective disciplinary orientation of scientists. Thus, this work is part of an extended effort within science mapping to provide reliable, reproducible measures of scientific content. Such measures inform both science policy and the sociology of science. More work is needed to probe the relationship between the ascription of scientific discipline, and the use of selected scientific language. Cunningham (2000) evaluates this relationship through interviews with scientists. This work finds that scientific language is as least as telling as journal classification systems in appraising the disciplinary orientation of scientists.

In this paper we have described a new approach to map the semantics of science. This new approach is motivated by interest in how different researchers, working in different disciplines and field use language. Understanding how scientists use language can benefit multidisciplinary research, interdisciplinary understanding, and aid the development of taxonomies of science. Data about language usage, however, is heterogeneous in character due to polysemy, synonymy, and natural variance in the selection of words by individual scientists.

Approaches for handling synonymy and polysemy have been proposed in the literature. Most prominent, latent semantic indexing, a method intended for document retrieval, aimed at addressing this joint problem. In this paper, we argued that either clustering or factoring, if used independently, is inadequate in overcoming the problems of polysemy and synonymy. Factoring can be used for synonymy, but fails when facing polysemy. Clustering approaches, in contrast, can offer support for identifying polysemous words and still allocate the documents that make use of these words to the cluster with which they have the most in common as reflected by the other words that are also used. However, clustering struggles to capture the rich semantics of language. By combining these two approaches in some way, an integral solution to the problems might be found.

We extensively discussed and illustrated one possible model that combines clustering and factoring, namely the mixtures of factor analyzers. This model was first proposed in the literature by Ghahramani & Hinton (1997), but has not yet been used for science mapping purposes. This model expands the traditional factor analysis model by allowing multiple centers in the dataset. We have discussed the mathematical formulation underlying this model and what this implies for science mapping. We illustrated conceptually how the model can handle both synonymy and polysemy. Finally, we demonstrated the model using a case study.

In the case study, we first compared the MFA model with its constituent models, namely a cluster analytic technique called the mixture of Gaussians and traditional factor analysis. Using Aikake's information criterion, we selected both the best model structure and the best parameterization of that structure. Neither the mixture of Gaussians nor traditional factor analysis could generate a meaningful model. In contrast, the MFA model could both identify a meaningful model and the AIC of this model was lower, indicating a higher in formativeness, then the best model of the mixture of Gaussians or traditional Factor Analysis. The case thus showed that the MFA model can handle heterogeneous data and translate this data into a semantic map of science. Additional evaluation of the MFA model is needed. In this paper, we presented the MFA model and illustrated it using a single case, where we compared it to a mixture of Gaussians and traditional factor analysis. The MFA model should both be confronted with alternative sets of data, and with alternative model structures that combine factoring and clustering or other richer structures which are not factor or cluster based. For comparative purposes, new research should adopt a framework for comparison that utilizes likelihood and the AIC.

The authors appreciate the comments of two anonymous reviewers, as well as the comments of their colleague Jill Slinger.

References

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transaction on Automatic Control,* 19(6), 716-723.

Bauin, S. (1986). Aquaculture: A Field by Bureaucratic Fiat. In M. Callon, J. Law & A. Rip (Eds.), *Mapping the Dynamics of Science and Technology*. London: MacMillan Press Ltd.

Bensman, S. J. (2005). Urquhart and Probability: The Transition From Librarianship to Library and Information Science. *Journal of the American Society for Information Science,* 56(2), 189-214.

Bishop, C. M., Svensen, M., & Wiliams, C. K. I. (1998). The Generative Topographic Mapping. *Neural Computation,* 10(1), 215-234.

Borner, K., Chen, C., & Boyack, K. W. (2005). Visualizing Knowledge Domains. *Annual Review of Information Science and Technology,* 37(1), 179-255.

Braam, R. R., Moed, H. F., & van Raan, A. F. J. (1991a). Mapping of Science by Combined Co-Citation and Word Analysis. (I) Structural Aspects. *Journal of the American Society for Information Science,* 42(4), 233-251.

Braam, R. R., Moed, H. F., & van Raan, A. F. J. (1991b). Mapping of Science by Combined Co-Citation and Word Analysis. II: Dynamical Aspects. *Journal of the American Society for Information Science,* 42(4), 252-266.

Callon, M., Law, J., & Rip, A. (Eds.). (1986). *Mapping the Dynamics of Science and Technology*. London: MacMillan Press Ltd.

Coleman, J. S. (1964). *Introduction to Mathematical Sociology*. New York: The Free Press.

Cunningham, S. W. (2000). *The Content Evaluation of British Scientific Research.* Ph.D. thesis, University of Susex.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, D. (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science,* 41(6), 391-407.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum Likelihood From Incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B (Methodological),* 39(1), 1-38.

Faba-Perez, C., Guerrero-Bote, V. P., & Den Moya-Anegon, F. (2003). Data mining in a closed Web environment. *Scientometrics,* 58(3), 623-640.

Funtowics, S. O., & Ravetz, J. R. (1990). *Uncertainty and Quality in Science for Policy*. Dordrecht: Kluwer Academic Publishers.

Ghahramani, Z., & Hinton, G. E. (1997). *The EM algorithm for mixtures of factor analyzers* (Technical Report CRG-TR-96-1): Department of Computer Science, University of Toronto, .

Greenacre, M. J. (1984). *Theory and Application of Correspondence Analysis*. London: Academic Press.

Healey, P., Rothman, H., & Hoch, P. K. (1985). An Experiment in Science Mapping for Research Planning. *Research Policy,* 15, 233-251.

Hicks, D. (1987). Limitations of Co-Citation Analysis as a Tool for Science Policy. *Social Studies of Science,* 17(2), 295-316.

Knight, F. H. (1921). *Risk, Uncertainty and Profit* (1957 ed.). Mineola, New York: Dover Publication.

Kohonen, T. (1995). *Self-Organizing Maps*. Berlin: Springer-Verlag.

Kumagi, T., & Utsugi, A. (2004). Removal of Artifacts and Fluctuations from MEG Data by Clustering Methods. *Neurocomputing,* 62, 153-160.

Kwakkel, J. H., & Cunningham, S. W. (2008). *The Semantics of the Uncertainty Literature.* Paper presented at the Portland International Conference on Management of Engineering and Technology, Capetown, South Africa.

Lempert, R., Popper, S., & Bankes, S. (2002). Confronting Surprise. *Social Science Computer Review,* 20, 420-439.

Lin, X. (1997). Map displays for information retrieval. *Journal of the American Society for Information Science,* 48, 40-54.

McCain, K. W. (1989). Descriptor and Citation Retrieval in the Medical Sciences Literature: Retrieval Overlaps and Novelty Distribution. *Journal of the American Society for Information Science,* 40(2), 110-114.

McLachlan, G. J., Bean, R. W., & Peel, D. (2002). A Mixture Model-Based Approach to the Clustering of Microarray Expression Data. *Bioinformatics,* 18(3), 413-422.

McLachlan, G. J., & Peel, D. (2000). *Finite Mixture Models*: John Wiley & Sons.

McLachlan, G. J., Peel, D., & Bean, R. W. (2003). Modelling High-Dimensional Data by Mixtures of Factor Analyzers. *Computational Statistics & Data Analysis,* 41(3-4), 379-388.

Morgan, M. G., & Henrion, M. (1990). *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge: Cambridge University Press.

Norton, J. P., Brown, J. D., & Mysiak, J. (2006). To What Extent, and How, Might Uncertainty be Defined? Comments Engendered by "Defining Uncertainty: a Conceptual Basis for Uncertainty Management in Model-Based Decision Support" Walker *et al*. Integrated Assessment 4: 1, 2003. *Integrated Assessment,* 6(1), 82-88.

Noyons, E. C. M., Moed, H. F., & Raan, A. F. J. v. (1999). Integrating Research Performance Analysis and Science Mapping. *Scientometrics,* 46(3), 591-604.

Peters, H. P. F., & van Raan, A. F. J. (1993a). Co-word-based Science Maps of Chemical Engineering. Part I: Representations by Direct Multidimensional Scaling. *Research Policy,* 22(1), 23-45.

Peters, H. P. F., & van Raan, A. F. J. (1993b). Co-word-based Science Maps of Chemical Engineering. Part II: Representations by Combined Clustering and Multidimensional Scaling. *Research Policy,* 22(1), 47-71.

Small, H. (1997). Update on Science Mapping: Creating Large Document Spaces. *Scientometrics,* 38(2), 275-293.

Small, H. (1998). A General Framework for Creating Large-Scale Maps of Science in Two or Three Dimensions: The SciViz System. *Scientometrics,* 41(1-2), 125-133.

Small, H. (1999). Visualizing Science by Citation Mapping. *Journal of the American Society for Information Science,* 50(9), 799-813.

Small, H., & Garfield, E. (1985). The Geography of Science: Disciplinary and National Mappings. *Journal of Information Science,* 11, 147-159.

Small, H., & Griffiths, B. C. (1974). The Structure of Scientific Literature 1: Identifying and Graphing Specialties. *Science Studies,* 4, 17-40.

Swanson, D. R. (1987). Two Medical Literatures that are Logically but not Bibliographically Connected. *Journal of the American Society for Information Science,* 38(4), 228-233.

Teh, Y. W., & Roweis, S. (2002). Automatic Alignment of Local Representations. *Neural Information Processing Systems,* 15.

van Asselt, M. B. A. (2000). *Perspectives on Uncertainty and Risk*. Dordrecht: Kluwer Academic Publishers.

van den Besselaar, P., & Heimeriks, G. (2006). Mapping Research Topics Using Word-Reference Co-Occurances: A Method and an Exploratory Case Study. *Scientometrics,* 68(3), 377-393.

van der Sluis, J. P. (1997). *Anchoring Amid Uncertainty: On the Management Of Uncertainties in Risk Assessment of Anthropogenic Climate Change.* University of Utrecht, Utrecht.

Walker, W. E., Harremoës, J., Rotmans, J. P., Van der Sluijs, J. P., Van Asselt, M. B. A., Janssen, P. H. M., et al. (2003). Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment,* 4(1), 5-17.

Yamamoto, H., Nankaku, Y., Miyajima, C., Tokuda, K., & Kitamura, T. (2005). Parameter Sharing in Mixture of Factor Analyzers for Speaker Identification. *IEICE Transactions on Information and Systems,* E88-D(3), 418-424.

Zhoe, Y. K., & Mobasher, B. (2006). Web User Segmentation Based on a Mixture of Factor Analyzers. *Lecture Notes in Computer Science,* 4082, 11-20.

3 Classifying and communicating uncertainties in model-based policy analysis

This chapter has originally been published as KWAKKEL, J.H., WALKER, W.E. & MARCHAU, V.A.W.J. (2010) Classifying and Communicating Uncertainties in Model-Based Policy Analysis. *International Journal of Technology, Policy and Management*, 10(4), 299-315.

Abstract

Uncertainty is of paramount importance in modern day decisionmaking. In response to this, it has been suggested that policy analysts have to be more careful in communicating the uncertainties that are inherent in policy advice. In order to support policy analysts in identifying uncertainties and communicating these uncertainties to decisionmakers, an uncertainty matrix was proposed by Walker, et al.(2003), which synthesized various taxonomies, frameworks, and typologies of uncertainties from different decision support fields. Since its publication, this framework has been applied to different domains. As a result, the framework has undergone changes resulting in a proliferation of uncertainty frameworks. This proliferation runs counter to the purpose of the original framework. This paper presents an extensive review of the literature that builds on Walker, et al. (2003). In light of this, a synthesis is presented, which can be used to assess and communicate uncertainties in model-based policy analysis studies.

3.1 Introduction

Uncertainty is increasingly important and controversial in a variety of domains, such as environmental science (Van der Sluijs, 2007), water management (Pahl-Wostl, et al., 2007), and transport planning (Marchau, Walker, & van Duin, 2009). A framework that aims at offering a common basis for uncertainty in model-based policy analysis across these domains was put forward by Walker, Harremoës, et al. (2003). We will refer to this as the W&H

framework, which integrated previous research on uncertainty (e.g. Funtowicz & Ravetz, 1990; Morgan & Henrion, 1990; van Asselt, 2000; van der Sluis, 1997) into a single coherent taxonomy for uncertainty classification and analysis in model-based decision support. The emphasis of the framework was on providing a common vocabulary for classifying uncertainties in a model. The wider context of the study, and the decisionmaking problem for which the study was conducted, were not explicitly considered.

The W&H framework has been applied in a variety of domains. However, these applications have resulted in numerous modifications and changes, which were motivated on the one hand by application domain related issues, and on the other hand by perceived weaknesses and problems with the original framework. Now, a wide variety of frameworks, all derived from the original framework, are available. This is inconsistent with the purpose and intention of the original framework.

Instead of integration and harmonization, we observe a new proliferation of various typologies, taxonomies, and frameworks for analyzing uncertainty. For many reasons, it would be desirable to have a generally agreed upon framework for analyzing and communicating uncertainties (Walker, et al., 2003). First, consistency in terminology will facilitate communication among policy analysts. Second, it will provide for better communication between policy analysts and decisionmakers. Third, a generally agreed upon and accepted framework for analyzing and classifying uncertainty that adequately captures the multi-dimensional character of uncertainty will support policy analysts in assessing the impact of the various uncertainties on the decisionmaking problem at hand, offering a starting point for identifying and prioritizing effective ways of handling the uncertainties. The observed proliferation of frameworks runs counter to the achievement of these benefits. Therefore, there is a need to assess the variety of frameworks that has emerged and, where possible, synthesize these into a new and improved common framework for uncertainty analysis in model-based policy analysis.

This paper aims at assessing the variety of frameworks that has emerged and synthesizing them into a new and improved common framework for uncertainty analysis in model-based policy analysis. In order to achieve this, we will answer the following questions: How has the original W&H framework evolved? What problems were encountered and how have they been addressed? What modifications have been made in the subsequent literature and what motivated these changes? In order to answer these questions, we performed an extensive literature review, based on citation analysis. We also report on this literature review. In conclusion, we present a framework that synthesizes the existing literature. We observe that the W&H framework has been expanded to classify uncertainties in the wider process. However, we believe that process categories are often context dependent; generic categorization might, therefore, not fit any specific case. That is, if researchers use a different process for the development and use of models for model-based decision support, some of the categories that apply to one process may not apply to another.

The structure of this paper is as follows. Section 2 presents details on the original framework. Section 3 presents the results of the literature review. Section 4 presents the new and improved framework. Section 5 contains our closing remarks.

3.2 The W&H framework

In this section, we introduce the original W&H framework, which aimed at providing a conceptual basis for the systematic treatment of uncertainty in model-based decision support activities by highlighting the overlaps among a large number of typologies (e.g. Funtowicz & Ravetz, 1990; Hodges, 1987; Morgan & Henrion, 1990; van Asselt, 2000; van der Sluis, 1997). Uncertainty is here defined as any departure from the unachievable ideal of complete determinism (Walker, et al., 2003). Uncertainty is not simply a lack of knowledge, since an increase in knowledge might lead to an increase of knowledge about things we don't know and thus increase uncertainty (van Asselt, 2000).

The problem the W&H framework addresses is that despite the general acknowledgment of the presence of uncertainty by both policymakers and policy analysts, there is limited appreciation for the multidimensionality of uncertainty, the relative magnitudes, and approaches for dealing with uncertainty. The framework is intended first and foremost to facilitate communication among policy analysts. Second, the framework should enhance communication about uncertainty among policy analysts, policymakers, and stakeholders by providing the foundations for a common language. Third, a better understanding and appreciation of the different dimensions of uncertainty can support the identification and the prioritization of uncertainties and help in the selection of the appropriate treatment of a given uncertainty. Walker et al. (2003) hope that uncertainty analyses will be conducted along the lines of the uncertainty framework they propose, and that these analyses will be applied on a routine basis when communicating about the results of a model-based decision support study, which should help improve communication among the actors involved in the study and using the results of the study, and should ultimately improve the quality of model-based decision support.

Walker et al. (2003) identify three different dimensions of uncertainty. The first dimension is the location of uncertainty – where is the uncertainty located in the systems analysis framework? The second dimension is the level of uncertainty. This is an expression of the degree or severity of the uncertainty, which can range from deterministic knowledge to total ignorance. The third dimension is the nature of the uncertainty. The phenomena about which we are uncertain can either be due to our lack of knowledge about the phenomena (i.e. epistemic uncertainty) or due to the inherent variability in the phenomena (i.e., uncertainty inherent in their nature). Table 3-1 shows the resulting framework.

Table 3-1: The original W&H framework

Location		Level				Nature	
		Statistical Uncertainty	Scenario Uncertainty	Recognized Ignorance	Total Ignorance	Epistemic Uncertainty	Stochastic Uncertainty
Context	Natural, technological, economic, social, and political setting						
Inputs	System data						
	Driving forces						
Model	Model structure						
	Computer implementation						
Parameters							
Model Outputs							

Source: Walker, et al. (2003)

3.3 Evolution of the W&H framework

In order to generate an overview of how the W&H framework has evolved, we performed a citation analysis of the articles that cite Walker et al. (2003), using Google Scholar, Scopus, and ISI Web of Science. In total, through January 2010, we identified 125 unique references to Walker et al. (2003). In some cases we identified both a journal and conference version of the same paper. In these cases, we included only the journal version of the paper in the analysis.

The resulting database of articles was analyzed using citation analysis and social network analysis techniques. Specifically, we identified citation patterns and deduced the different bodies of literature that built on the W&H framework, using scripts developed in the programming language Python and visualized using the social network analysis tool Pajek. Below we first discuss applications of the unmodified W&H framework. Next, we discuss various modifications. We close this section with a summary of problems identified in the literature.

3.3.1 Applications of the Original W&H framework

First, several applications of the original framework have been reported in the literature. One body of literature comes from the Ph.D. work of Krayer von Krauss (2005). He operationalized the framework and used it in case studies focused on transgene silencing and herbicide-tolerant oilseed crops (Krayer von Krauss, Casman, & Small, 2004; Krayer von Krauss, Kaiser, Almaas, van der Sluijs, & Kloprogge, 2008). More recently, this operationalization of Krayer von Krauss has been used for studying uncertainty about the effects of the introduction of DNA vaccines to aquaculture (Gillund, Kjølber, Krayer von Krauss, & Myhr, 2008). A second area of application of the W&H framework is in assessing uncertainty in models used for water level predictions (Warmink, Booij, Van der Klis, & Hulscher, 2007; Xu, Holzhauer, Booij, & Sun, 2008; Xu & Tung, 2008). A third area of application of the original framework can be found in transportation research, where it has been used to study uncertainty in traffic forecasts and long-term transportation planning (de Jong, et al., 2005; J.H. Kwakkel, Walker, & Wijnen, 2008; Salling & Leleur, 2006, 2007).

3.3.2 Modifications and extensions of the W&H framework

The W&H framework has been used as a starting point for their own uncertainty analysis frameworks by various researchers. Table 3-2 summarizes these different modifications. The "Changes" column summarizes the main changes from the original framework. The "Motivation" column specifies the motivation for the changes. The "Sources" column specifies the literature where the changes can be found. Below, we discuss Table 3-2 in more detail.

Table 3-2: Overview of modifications and extensions of the W&H framework

Changes	Motivation	Sources
Focus on perceived uncertainties, reworking of the location dimension into a source dimension	Object of study was the role of perceived uncertainties on entrepreneurial decisionmaking	(Meijer, 2008; Meijer, Hekkert, Faber, & Smits, 2006; Meijer, Hekkert, & Koppenjan, 2007)
Starts from the framework of (Meijer, 2008) and changes the level dimension using (Courtney, 2001)	Level dimension was perceived to be unclear	(Fijnvandraat, 2008)
Adds two dimensions (qualification of knowledge base and value-ladenness of choice), and expands location dimension with additional locations.	Dimensions were implicit in (Walker, et al., 2003); now they are made explicit. Furthermore, the object of study is the entire model-based decision support process, making these dimensions more relevant.	(Janssen, Petersen, van der Sluijs, Risbey, & Ravetz, 2005; Van der Sluijs, et al., 2003)
A reworking of the entire framework, The location dimension is modified, making sharper distinctions between the conceptual model, the mathematical model, and its implementation. Similarly, the level dimension is reworked, and two dimensions related to pedigree (Funtowicz & Ravetz, 1990) are added.	The framework was intended as a philosophical grounding for the RIVM guidance system. It is a synthesis of (Funtowicz & Ravetz, 1990) and the W&H framework, based on philosophy of science considerations pertaining to the role of simulation models in the scientist's toolbox.	(Petersen, 2006)
Replaces level dimension with a taxonomy of uncertainties based on (Brown, 2004) and renames location to source. Furthermore, uncertainty is redefined as perceived uncertainties	The framework is aimed at supporting people involved in Integrated Water Resource Management with performing uncertainty analyses, in order to clarify to the decisionmakers how reliable the outcomes of the study are.	(Refsgaard, 2007; Refsgaard, Henriksen, Harrar, Scholten, & Kassahun, 2005; Refsgaard, van der Sluijs, Brown, & van der Keur, 2006; Refsgaard, van der Sluijs, Hojberg, & Vanrolleghem, 2007)
Case based operationalization of the "Refsgaard" framework, by combining it with (Van Asselt & Rotmans, 2002) and (Dewulf, Craps, Bouwen, Taillieu, & Pahl-Wostl, 2005). The location dimension is reworked in order to achieve this. The most important change here is the explicit consideration of the various conceptual frames.	Focus is on the entire modeling process, not merely on the uncertainties in the model. Quite some emphasis is placed on ambiguity arising from various different ways of framing.	(van der Keur, et al., 2008)

Meijer et al. adapted the original W&H framework in order to classify perceived uncertainties in socio-technological transformations (Meijer, 2008; Meijer, et al., 2006; Meijer, et al., 2007). As a result of this shift in the object of study, Meijer et al. (2006) made modifications to the location dimension and redefined the framework to apply it to study perceived uncertainties. Fijnvandraat (2008) adapted this typology of perceived uncertainties in order to understand the role of uncertainties and risk in infrastructural investment, with specific attention being given to broadband roll-out. She changed the scale used for describing the level of uncertainty to one introduced by Courtney (2001).

At the Dutch National Institute for Public Health and the Environment (RIVM), a guidance system for uncertainty analysis was developed based on the W&H framework (Janssen, et al., 2005; Van der Sluijs, et al., 2003). The scope of this guidance system extends beyond uncertainty assessment in model results. It covers the entire process of environmental assessment, including issues such as problem framing, stakeholder participation, and indicator selection (Janssen, et al., 2005; Van der Sluijs, Petersen, Janssen, Risbey, & Ravetz, 2008). Because of this increase in scope, the W&H framework was expanded by adding two dimensions – qualification of the knowledge base and value-ladenness of the choices – and categories were added to the location dimension that are related to the wider process of environmental assessment. In parallel to the development of the guidance system for uncertainty assessment and communication in environmental assessment studies at RIVM (Janssen, et al., 2005), a philosophical reflection on uncertainty was conducted by Petersen (Petersen, 2006). The goal of this reflection was to shed light on the role of simulations in science, and provide a philosophical foundation for the aforementioned guidance system. The resulting matrix can be understood as combining the distinction of Funtowicz and Ravetz (1990) between sorts of uncertainty and sources of uncertainty with the uncertainty matrix of Janssen et al. (2005).

Refsgaard et al. adapted the W&H framework for use in integrated water resource management (IWRM) (Refsgaard, 2007; Refsgaard, et al., 2005; Refsgaard, et al., 2006; Refsgaard, et al., 2007). Uncertainties range from ambiguity in defining problems and goals to uncertainty in data and models. All these uncertainties need to be considered during the management process of a modeling study To assist people involved in modeling studies for IWRM with performing such an uncertainty analysis, Refsgaard et al. (2007) propose a terminology and framework for uncertainty analysis. The terminology and framework they propose combines the W&H framework with the taxonomy of imperfect knowledge of Brown (2004). An application of the approach suggested by Refsgaard et al. (2007) can be found in van der Keur et al. (2008). Van der Keur et al. (2008) combine this approach with work by van Asselt & Rotmans (2002) in the field of uncertainty in integrated assessment and Dewulf, et al. (2005) on the importance of multiple frames or ambiguity of different actors in the environmental management process. By combining these three different elements, van der Keur et al. (2008) are able to identify uncertainty in all the stages of the IWRM process, which in turn can contribute to the development of adaptive management (Pahl-Wostl, Moltgen, Sendzimir, & Kabat, 2005; van der Keur, et al., 2008).

3.3.3 Problems and criticism of the W&H framework

Over the course of the evolution of the W&H framework, problems, research questions, and criticism have emerged. Two issues stand out: (1) a shift to perceived uncertainties and the role of frames, and (2) various modifications of the level dimension. The W&H framework focuses explicitly on the modeler's perspective on uncertainty, purposely avoiding the perspective of decisionmakers. Walker et al. (2003) focused on offering a step towards addressing the problem uncertainty poses in model-based decision support by providing a framework that can serve as a common language for scientists, decisionmakers, and stakeholders. The emphasis of this matrix is on the modelers. However, as is also acknowledged by Walker et al. (2003), there is a decisionmaker's perspective on uncertainty that also deserves attention. The attention given in the literature to perceived uncertainties, ambiguity, and the role of framing can all be understood as attempts to address this issue.

A second problem is that the typology overlooks the diversity of terms and meanings associated with uncertainty and the situations in which these usages arise. The term uncertainty has a plethora of meanings and connotations. When one uses the term, the context shapes which specific meaning and connotation is used. Related to this, Norton et al. (2006) claim that many of these different meanings can be expressed by using terms that are in some way synonymous with the term uncertainty. Words such as ignorance, doubt, unsureness, risk, ambiguity, imprecision, and randomness all capture some element of what is meant by uncertainty. If one aims at providing an integrative framework for classifying uncertainty, an analysis of the diversity of meanings and concepts associated with uncertainty, and the contexts in which these meanings and concepts are used, is called for (Norton, et al., 2006). The importance of performing such an analysis is underscored by Gillund et al (2008), who observed that the W&H framework is often difficult to understand for researchers with a different background from that of the authors. Related, Skeels et al. (2008) point out that an empirical evaluation of the success of the W&H framework in harmonizing the language with respect to uncertainty in model-based decision support is missing. An analysis of the diversity of terms and discourses in which uncertainty is discussed across the sciences has recently appeared, but its implications for uncertainty typologies in general and the W&H typology in particular are not explicitly addressed (J. H. Kwakkel & Cunningham, 2009). We speculate that the fact that so much of the subsequent literature has made modifications to the level dimension, which in principal need not be application domain specific, can be explained by the fact that there is such a large variety of words, terms, and concepts that can be used to discuss uncertainty.

3.4 Towards a synthesis

Across the different directions of development of the W&H framework, dimensions were added, existing ones re-conceptualized, and a shift to include perceived uncertainties took place. Looking across these different developments, we observed two main issues: the importance of framing and problems associated with the level dimension. We now reflect on these two issues in more detail, with the goal of presenting a new framework that addresses these two issues, while retaining consistency with earlier work.

3.4.1 Rethinking the level dimension

The level dimension focuses on the degree of uncertainty. That is, where does the uncertainty manifest itself along the continuum ranging from deterministic knowledge to total ignorance? This aspect of uncertainty might very well have the longest history, dating back to philosophical questions debated among the ancient Greeks about the certainty of knowledge and perhaps even further. Its modern history begins around 1921 when Knight made a distinction between risk and uncertainty (Knight, 1921; Rechard, 1999; Samson, Reneke, & Wiecek, 2009). According to Knight (1921), risk denotes the calculable and thus controllable part of all that is unknowable. The remainder is the uncertain – incalculable and uncontrollable. Luce and Raiffa (1957) adopted these labels to distinguish between decisionmaking under risk and decisionmaking under uncertainty. Similarly , Quade (1982) makes a distinction between “stochastic” uncertainty and “real” uncertainty. According to Quade (1982), stochastic uncertainty includes frequency-based probabilities and subjective (Bayesian) probabilities. Real uncertainty covers the future state of the world and the uncertainty resulting from the strategic behavior of other actors. Often, attempts to express the degree of certainty and uncertainty have been linked to whether or not to use probabilities, as exemplified by Morgan and Henrion (1990), who make a distinction between uncertainties that can be treated through probabilities and uncertainties that cannot. Uncertainties that cannot be handled probabilistically include model structure uncertainty and situations in which experts cannot agree upon the probabilities. These are the more important and hardest to handle types of uncertainties (Morgan, 2003). These kinds of uncertainties are now sometimes referred to as deep uncertainty (R. J. Lempert, Popper, & Bankes, 2002), or severe uncertainty (Ben-Haim, 2006).

The literature review has shown that various alternative conceptualizations of the level dimension have been put forward. However, most of these suffer from a problem that originates in the W&H framework. The labeling of the level dimension in Walker et al. (2003) is open to multiple interpretations, for it includes the names of families of techniques for treating uncertainty as a measure for expressing our degree of uncertainty. Reconceptualizations of the level dimension that do not address this basic problem suffer from this same defect (e.g. Refsgaard, 2007; Van der Sluijs, et al., 2003). The other alternative, pursued by (Fijnvandraat, 2008), to use numerical labels, following (Courtney, 2001), is also not free of problems. Using numbers for expressing the degree or level of uncertainty can be viable – at least, numbers are free of methodological or disciplinary connotations, in contrast to terms like probability, scenario, or risk. However, the specifications given by Courtney (2001) of the four levels of uncertainty cannot be directly applied to all the locations identified in the W&H framework and subsequent literature.

A careful reconceptualization of the level dimension is, therefore, in order. If this reconceptualization is to be useful, it needs to be based on unambiguous terms that are known to scientists working in various model-based decision support fields. Only in this way can the observation of Gillund (2008) about the difficulty in explaining the framework to others be overcome. To some, the idea of different levels of uncertainty might appear strange. However, as is clear from the foregoing literature, numerous researchers do in fact acknowledge that there are different levels of uncertainty (e.g. Knight, 1921; Luce & Raiffa, 1957; Morgan & Henrion, 1990; Quade, 1982). Broadly speaking, the level of uncertainty is the assignment of likelihood to things or events. In some cases the likelihood or plausibility of a thing or event can be expressed using numbers, but in other cases more imprecise labels are used, such as more likely, less likely, or equally likely. Instead of arguing for or against this practice, we take this practice as a given.

We define four levels of uncertainty. Table 3-3 shows the resulting levels of uncertainty, their description, and some examples. The least uncertain is Level 1 uncertainty, or shallow uncertainty. In case of Level 1 uncertainty, probabilities can be used to specify the likelihood or plausibility of the uncertain alternatives. In case of Level 2 uncertainty, or medium uncertainty, alternatives can be enumerated and rank ordered in terms of their likelihood, but how much more likely or less likely cannot be specified. This level of uncertainty is encountered when one is able to enumerate alternatives and is able to say whether they are more likely, equally likely, or less likely, without being able or willing to quantify this further. In case of Level 3 uncertainty, or deep uncertainty, alternatives can be enumerated, but for various reasons, such as that decisionmakers or experts cannot agree or don't know, even a rank ordering is ruled out. The strongest form of uncertainty is Level 4 uncertainty, or recognized ignorance. Here, alternatives cannot even be enumerated. However, even when alternatives cannot be enumerated, merely keeping open the possibility of being wrong or of being surprised is still possible.

Table 3-3: The four levels of uncertainty

Level of Uncertainty	Description	Examples
Level 1 (shallow uncertainty)	Being able to enumerate multiple alternatives and being able to provide probabilities (subjective or objective)	Being able to enumerate multiple possible futures or alternative model structures, and specify their probability of occurring
Level 2 (medium uncertainty)	Being able to enumerate multiple alternatives and being able to rank order the alternatives in terms of perceived likelihood. However, how much more likely or unlikely one alternative is compared to another cannot be specified	Being able to enumerate multiple possible futures or alternative model structures, and being able to judge them in terms of perceived likelihood
Level 3 (deep uncertainty)	Being able to enumerate multiple alternatives without being able to rank order the alternatives in terms of how likely or plausible they are judged to be	Being able to enumerate multiple possible futures or specify multiple alternative model structures, without being able to specify their likelihood
Level 4 (recognized ignorance)	Being unable to enumerate multiple alternatives, while admitting the possibility of being surprised	Keeping open the possibility of being wrong or being surprised

Thus, the level dimension tries to capture differences in the types of scales that are used in practice when assigning likelihood to things or events. The purpose of the theory of scales of measurement was to recognize the various forms of measurement that exist, specify the formal mathematical properties of these different forms, make the rules for the assignment of numerals for each form explicit, and identify the statistical operations that were permissible for each (Stevens, 1946). We argue that the scales of measurement also apply to the different levels of uncertainty. Relating scales of measurement to the levels of uncertainty has several advantages over the alternatives identified in the existing literature. The scales are reasonably well known across different scientific disciplines. They are well and clearly defined. The labels used to denote the different scales of measurement are free of methodological connotations. Together these advantages should overcome the identified difficulties with existing forms of labeling of the different levels of uncertainty.

Table 3-4 shows the different types of measurement scales. The simplest type is the nominal or categorical scale. Here, if numbers are used, they are just names; letters or other labels could be used as well. The only operation is to assess whether things belong to the same class. This scale of measurement corresponds to Level 3 uncertainty. The second scale of measurement is the ordinal scale. On top of determination of equality, one can also rank order the objects that are measured using an ordinal scale. This scale of measurement corresponds to Level 2 uncertainty. The third scale of measurement – the interval scale – allows one to not only express greater or less, but also how much greater or less something is. To our knowledge, no interval scale has been used in practice to express the degree of uncertainty. If numbers are used and intervals can be specified, probabilities (a ratio scale) are the measurement of choice. The ratio scale also has an absolute zero point. This scale corresponds to Level 1 uncertainty, which is generally measured in the form of probabilities. Only Level 4 uncertainty has no equivalent scale of measurement, which is explained by the fact that in case of Level 4 there are no objects that can be measured.

Table 3-4: Levels of measurement

Level of Measurement	Mathematical Structure	Basic Empirical Operation	Characteristics
Nominal / categorical	Unordered set	Determination of equality	Classification into categories, no relations between categories
Ordinal	Totally ordered set	Determination of greater or less	In addition to the above, ordering of categories based on degree to which they possess some characteristic. No information on magnitude of difference
Interval	Affine line	Determination of equality of intervals or differences	In addition to the above, ordering of categories, but with information on magnitude of difference, but no absolute zero
Ratio	Field	Determination of equality of ratios	In addition to the above, absolute zero

Which scale of measurement is used depends on the empirical character of the thing to be measured, on the analysis that one wants to perform on the measurements, and on the procedural choices made during the measurement (Stevens, 1946). So, too, with the level of uncertainty: the scale to be used depends on the uncertainty and on procedural and methodological choices. For example, the state of the world thirty years from now is highly uncertain and limits the potential for using probabilities, specifically if the decision problem at hand is contested. That is, decisionmakers and/or experts cannot agree (R.J. Lempert, 2002; O'Hagan & Oakley, 2004). Some analysts might then choose to treat it as a Level 3 uncertainty, with equally likely scenarios; others might prefer to treat is as a Level 2 uncertainty, with rank ordered scenarios. Which is correct, and even whether there is a correct approach, is beyond the purpose of this paper. That is, the uncertainty matrix aims at offering analysts a tool for communicating about uncertainties. It is not intended to be normative in the sense of specifying the correct approach for handling the identified uncertainties. To quote the author that proposed the scales of measurement “this proposed solution to the semantic problem is not meant to imply that all scales belonging to the same mathematical group are equally precise or accurate or useful or "fundamental." Measurement is never better than the empirical operations by which it is carried out, and operations range from bad to good. Any particular scale . . . may be objected to on the grounds of bias, low precision, restricted generality, and other factors, but the object should remember that these are relative and practical matters and that no scale used by mortals is perfectly free of their taint" (Stevens, 1946, p. 680).

3.4.2 Perceived uncertainties and the importance of framing

There are multiple stakeholders involved in almost all model-based decision support activities. These different stakeholders have a heterogeneous background. This implies that different opinions, experiences, expectations, values, and forms of knowledge are present. In such situations, there is no obvious correct way of framing the decision problem at hand (Dewulf, et al., 2005; van Eeten, 1998). This plurality can give rise to unclarity, misunderstandings, and value conflicts. Therefore, it is necessary to account for these differences in frames when attempting to address uncertainty in model-based decision support activities in general (Brugnach, Dewulf, Pahl-Wostl, & Taillieu, 2008; Funtowicz & Ravetz, 1990; van Eeten, 1998). This issue in relation to the W&H typology has been explored in depth by Brugnach et al. (2008) with respect to natural resource management. However, as for example the work of van Eeten (1998) shows, these same framing related uncertainties are also present in other types of policy analysis activities, such as airport development. Therefore, we see the need to generalize the work of Brugnach et al. (2008) and explicitly include ambiguity as an additional category of the nature dimension.

Ambiguity is added to nature as a separate category to highlight the importance of how the same data can be interpreted differently by different actors depending on differences in frames and values. Ambiguity is defined as uncertainty arising from the simultaneous presence of multiple frames of reference about a certain phenomenon (Brugnach, et al., 2008; Dewulf, et al., 2005). Ambiguity is treated as a third kind of nature of uncertainty, along with ontic and epistemic uncertainty. Walker et al. (2003) argue that the nature of the uncertainty matters in choosing a strategy for handling uncertainty. For example, if the nature of the uncertainty is ontic, more research will not help. Similarly, if the nature is ambiguity, methods that aim at integrating frames and support joint sense making are appropriate, while scientific research based on accepting a single frame are not (Brugnach, et al., 2008; van Eeten, 1998).

3.4.3 The synthesized framework

We started this paper with the observation that a framework intended to harmonize language across different fields of model-based decision support had resulted in a new variety of frameworks with differences in terminology. This is a threat to the expected benefits of such a harmonized and generally accepted framework, such as improving the adequacy of communication among policy analysts, between policy analysts and decisionmakers, and the identification and prioritization of uncertainties and their treatment. We set out to review the literature in order to present a harmonized framework that reunites the different strands of research. The framework we presented is limited to classifying uncertainty about models and their inputs. We observe that the W&H framework has been expanded to also classify uncertainties in the wider policy analysis process. However, we believe that categories related to this wider policy analysis process are often context dependent, so a generic categorization might not fit the specific case. That is, if researchers use a different process for the development and use of models for model-based decision support, some of the categories that apply to one process may not apply to another. As such, the scope of the synthesized framework is the same as the scope of the W&H framework.

Given the emphasis on model-based decision support, the locations can be specified in quite some detail, as outlined in Walker et al. (2003). We combined the specification of location as given in Walker et al. (2003) with the one of Petersen (2006). Petersen (2006) focused on simulation models in general, without the specific model-based decision support focus. We slightly modified his specification of the location dimension in order to bring back this decision-oriented focus. So, we consider the following locations:

- **System boundary** is the demarcation of aspects of the real world that are included in the model from those that are not included. The system boundary is often determined by the problem and the chosen framing (Walker, et al., 2003). This location is called ‘context’ by Walker et al. (2003), while Petersen (2006) includes it under ‘conceptual model’.
- The **conceptual model** specifies the variables and relationships inside the boundaries. The conceptual model gives the interpretation of the computer model. Petersen (2006) has this as a separate category, while Walker et al. (2003) include it under model uncertainty.
- The **computer model** is the implementation of the conceptual model in computer code. With respect to this model, a distinction is often made between the **model structure** and the model parameters. The latter can be further subdivided into **parameters inside the model** that are fixed and the **input parameters to the model** that can be changed to reflect different external developments and/or different policies. Compared to Petersen (2006), we refine his generic parameter category by subdividing it along lines outlined by Walker et al (2003).
- **Input data** uncertainty deals with the uncertainties associated with determining the values for the different parameters both inside the model and as inputs to the model. These are often estimated based on empirical data, or derived from other models. Petersen (2006) calls this ‘model inputs; and Walker et al. (2003) call this ‘inputs’. In order to differentiate between model parameters and the data used in choosing their values, we call this ‘input data’.
- **Model implementation** uncertainty deals with the uncertainties arising from the implementation of the model in computer code. It addresses uncertainties related to bugs and errors in the code, or hardware errors such as the well-known numerical fault in early Pentium chips. Petersen (2006) calls this ‘technical model implementation’ and Walker et al. (2003) call this ‘computer implementation’.
- **Processed output data** is where the different uncertainties accumulate after propagating through the model complex. To highlight the fact that the output is often post-processed before being shown to decisionmakers, we follow Petersen (2006)in calling this ‘processed output data’.

By combining the location dimension, with the reconceptualization of the level dimension and the expanded nature dimension, a new synthesized uncertainty matrix can be made. This matrix is shown in Table 3-5.

Table 3-5: Synthesized uncertainty matrix

Location		Level				Nature		
		Level 1: Shallow uncertainty	Level 2: Medium uncertainty	Level 3: Deep uncertainty	Level 4: Recognized Ignorance	Ambiguity	Epistemology	Ontology
System boundary								
Conceptual Model								
Computer Model	Model structure							
	Parameters inside the model							
	Input parameters to the model							
Input Data								
Model Implementation								
Processed Output Data								

3.5 Conclusions

We started from the observation that uncertainty is increasingly important and controversial in a variety of scientific fields. To support scientists and decisionmakers in communicating about uncertainty, it has been suggested by different researchers that the policy advice should be accompanied by an uncertainty analysis that will clarify the quality and certainty of the conclusions based on the knowledge base underlying the policy advice. The W&H framework was proposed to provide a conceptual basis for the systematic treatment of uncertainty and improve the management of uncertainty in model-based decision support.

Over the course of this paper, we traced how the W&H framework has evolved and presented a synthesis of this evolution in the form of a new uncertainty matrix. We observed that the framework has evolved significantly since its inception along a variety of lines. Notable expansions include the extensions by van der Sluijs et al. (2003) and Refsgaard et al. (2007), who both tried to expand the framework in order to make it applicable to the analysis of uncertainties in the overall model-based decision support process. In order to make it applicable, a shift towards the perception of uncertainty took place. Independently, a very similar shift can be observed in the work of Meijer (2008), who not only uses the notion of perceived uncertainties – just like Refsgaard et al. (2007) – but also replaces the location dimension with a source dimension – just like Refsgaard et al. (2007). In line with this shift to perceived uncertainties, the role of framing in relation to uncertainty has seen increasing attention, resulting in adding an additional type to the nature dimension in the form of ambiguity. Furthermore, we observe that a variety of researchers have struggled in communicating about uncertainty through the framework. In fact Gillund et al (2008), observed that it is ironic that a framework intended to facilitate communication requires quite some explanation and elaboration. The labels in the original typology for the different aspects are not intuitively clear.

These observations have been taken into account in the new framework. This synthesis explicitly addresses the importance of multiple frames by including ambiguity as a new type of nature of uncertainty. Furthermore, the level dimension has been reconceptualized, drawing on the well-established theory of measurement scales. No longer are names of methods or techniques used as labels on this dimension.

We hope that the presented synthesis, along with continuing work in the various fields where the W&H framework has been used, will help to promote systematic reflection on uncertainty and facilitate discussions among policy analysts, decisionmakers, and other stakeholders by offering a common vocabulary.

References

Ben-Haim, Y. (2006). *Information-Gap Decision Theory: Decisions under Severe Uncertainty* (2 ed.). New York: Wiley.

Brown, J. D. (2004). Knowledge, Uncertainty and Physical Geography: Towards the development of methodologies for questioning belief. *Transactions of the Institute of British Geographers,* 29(3), 367-381.

Brugnach, M., Dewulf, A., Pahl-Wostl, C., & Taillieu, T. (2008). Towards a Relational Concept of Uncertainty: about Knowing Too Little, Knowing Too Differently, and Accepting Not to Know. *Ecology and Society,* 13(2), 30. [online].

Courtney, H. (2001). *20|20 Foresight: Crafting Strategy in an Uncertain World*. Boston: Harvard Business School Press.

de Jong, G., Pieters, M., Miller, S., Daly, A., Plasmeijer, R., Graafland, I., et al. (2005). *Uncertainty in traffic forecasts: literature review and new results for the Netherlands* (WR-268-AVV). Santa Monica: Rand.

Dewulf, A., Craps, M., Bouwen, R., Taillieu, T., & Pahl-Wostl, C. (2005). Integrated Management of Natural Resources: Dealing with Ambiguous Issues, Multiple Actors and Diverging Frames. *Water Science & Technology,* 52(6), 115-124.

Fijnvandraat, M. L. (2008). *Shedding Light on the Black Hole the roll-out of broadband access networks by private operators.* Ph.D. thesis, Delft University of Technology, Delft, the Netherlands.

Funtowicz, S. O., & Ravetz, J. R. (1990). *Uncertainty and Quality in Science for Policy*. Dordrecht: Kluwer Academic Publishers.

Gillund, F., Kjølber, K., Krayer von Krauss, M. P., & Myhr, A. I. (2008). Do Uncertainty Analyses Reveal Uncertainties? Using the Introduction of DNA Vaccines to Aquaculture as a Case. *Science of the Total Environment,* 407, 185-196.

Hodges, J. S. (1987). Uncertainty, Policy Analysis, and Statistics. *Statistical Science,* 2(3), 259-257.

Janssen, P. H. M., Petersen, A. C., van der Sluijs, J. P., Risbey, J. S., & Ravetz, J. R. (2005). A Guidance for Assessing and Communicating Uncertainty. *Water Science & Technology,* 52(6), 125-131.

Knight, F. H. (1921). *Risk, Uncertainty and Profit* (1957 ed.). Mineola, New York: Dover Publication.

Krayer von Krauss, M. P. (2005). *Uncertainty in Policy Relevant Sciences.* Ph.D. thesis, Technical University of Denmark, Lyngby

Krayer von Krauss, M. P., Casman, E. A., & Small, M. J. (2004). Elicitation of Expert Judgments of Uncertainty in the Risk Assessment of Herbicide-Tolerant Oilseed Crops. *Risk Analysis,* 24(6), 1515-1527.

Krayer von Krauss, M. P., Kaiser, M., Almaas, V., van der Sluijs, J. P., & Kloprogge, P. (2008). Diagnosing and Prioritizing Uncertainties According to their Relevance for Policy: The case of transgene silencing. *Science of the Total Environment,* 390(1), 23-34.

Kwakkel, J. H., & Cunningham, S. W. (2009). Managing polysemy and synonymy in science mapping using the mixtures of factor analyzers model. *Journal of the American Society for Information Science and Technology,* 60(10), 2064-2078.

Kwakkel, J. H., Walker, W. E., & Wijnen, R. A. A. (2008, 6-10 July). *The Treatment of Uncertainty in Airport Strategic Planning: The Case of Schiphol Airport's Long-Term Vision.* Paper presented at the 12th Air Transport Research Society World Conference, Athens, Greece.

Lempert, R. J. (2002). A New Decision Sciences for Complex Systems. *Proceedings of the National Academy of Sciences of the United States of America,* 99(3), 7309-7313.

Lempert, R. J., Popper, S., & Bankes, S. (2002). Confronting Surprise. *Social Science Computer Review,* 20, 420-439.

Luce, R. D., & Raiffa, H. (1957). *Games and Decisions*. New York: Wiley.

Marchau, V. A. W. J., Walker, W. E., & van Duin, R. (2009). An adaptive approach to implementing innovative urban transport solutions. *Transport Policy,* 15(6), 405-412.

Meijer, I. S. M. (2008). *Uncertainty and Entrepreneurial Action: the role of uncertainty in the development of emerging energy technologies.* Ph.D. thesis, University of Utrecht, Utrecht.

Meijer, I. S. M., Hekkert, M. P., Faber, J., & Smits, R. E. H. M. (2006). Perceived Uncertainties Regarding Socio-Technological Transformations: towards a Typology. *International Journal of Foresight and Innovation Policy,* 2(2), 214-240.

Meijer, I. S. M., Hekkert, M. P., & Koppenjan, J. F. M. (2007). How Perceived Uncertainties Influence Transitions; the Case of Micro-CHP in the Netherlands. *Technological Forecasting and Social Change,* 74(4), 519-537.

Morgan, M. G. (2003). Characterizing and Dealing With Uncertainty: Insights from the Integrated Assessment of Climate Change. *The Integrated Assessment Journal,* 4(1), 46-55.

Morgan, M. G., & Henrion, M. (1990). *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge: Cambridge University Press.

Norton, J. P., Brown, J. D., & Mysiak, J. (2006). To What Extent, and How, Might Uncertainty be Defined? Comments Engendered by "Defining Uncertainty: a Conceptual Basis for Uncertainty Management in Model-Based Decision Support" Walker *et al*. Integrated Assessment 4: 1, 2003. *The Integrated Assessment Journal,* 6(1), 82-88.

O'Hagan, A., & Oakley, J. E. (2004). Probability is perfect, but we can't elicit it perfectly. *Reliability Engineering and System Safety,* 85(1-3), 239-248.

Pahl-Wostl, C., Moltgen, J., Sendzimir, J., & Kabat, P. (2005). *New Methods for Adaptive Water Management under Uncertainty - the NeWater Project*. Paper presented at the EWRA conference.

Pahl-Wostl, C., Sendzimir, J., Jeffrey, P., Aerts, J., Berkamp, G., & Cross, K. (2007). Managing change towards adaptive water management through social learning. *Ecology and Society,* 12(2), 30 [online].

Petersen, A. C. (2006). *Simulating nature: A philosophical study of computer-simulation uncertainties and their role in climate science and policy advice.* Ph.D. thesis, Universiteit Amsterdam.

Quade, E. S. (1982). *Analysis for Public Decisions* (2nd ed.). New York: Elsevier Science Publishing Co.. Inc.

Rechard, R. P. (1999). Historical Relationship between Performance Assessment for Radioactive waste Disposal and Other Types of Risk Assessment. *Risk Analysis,* 19(5), 763-807.

Refsgaard, J. C. (2007). *Hydrological Modelling and River Basin Management.* Ph.D. thesis, University of Copenhagen, Copenhagen.

Refsgaard, J. C., Henriksen, H. J., Harrar, W. G., Scholten, H., & Kassahun, A. (2005). Quality assurance in model based water management - review of existing practice and outline of new approaches. *Environmental Modelling & Software,* 20(10), 1201-1215.

Refsgaard, J. C., van der Sluijs, J. P., Brown, J., & van der Keur, P. (2006). A Framework for Dealing with Uncertainty due to Model Structure Error. *Advances in Water Resources,* 26(11), 1586-1597.

Refsgaard, J. C., van der Sluijs, J. P., Hojberg, A., & Vanrolleghem, P. A. (2007). Uncertainty in the environmental modelling process - A framework and guidance. *Environmental Modelling & Software,* 22(22), 1543-1556.

Salling, K. B., & Leleur, S. (2006). *Assessment of Transport Infrastructure Projects by the Use of Monte Carlo Simulation: the CBA-DK model.* Paper presented at the Winter Simulation Conference, Monterey, California.

Salling, K. B., & Leleur, S. (2007). *Appraisal of Airport Alternatives in Greenland by the use of Risk Analysis and Monte Carlo Simulation.* Paper presented at the Winter Simulation Conference, Washington, DC.

Samson, S., Reneke, J. A., & Wiecek, M. M. (2009). A Review of Different Perspectives on Uncertainty and Risk and an Alternative Modeling Paradigm. *Reliability Engineering and System Safety,* 94(2), 558-567.

Skeels, M., Bongshin, L., Smith, G., & Robertson, G. (2008). *Revealing Uncertainty for Information Visualization.* Paper presented at the Working Conference on Advanced Visual Interfaces, Napoli, Italy.

Stevens, S. S. (1946). On the Theory of Scales of Measurement. *Science,* 103(2684), 667-680.

van Asselt, M. B. A. (2000). *Perspectives on Uncertainty and Risk*. Dordrecht: Kluwer Academic Publishers.

Van Asselt, M. B. A., & Rotmans, J. (2002). Uncertainty in Integrated Assessment Modelling. From Positivism to Pluralism. *Climate Change,* 54(1-2), 75-105.

van der Keur, P., Henriksen, H. J., Refsgaard, J. C., Brugnach, M., Pahl-Wostl, C., Dewulf, A., et al. (2008). Identification of Major Sources of Uncertainty in Current IWRM Practice, Illustrated for the Rhine Basin. *Water Resource Management,* 22(11), 1677-1708.

Van der Sluijs, J. P. (2007). Uncertainty and Precaution in Environmental Management: Insights from the UPEM conference. *Environmental Modelling & Software,* 22(5), 590-598.

Van der Sluijs, J. P., Petersen, A. C., Janssen, P. H. M., Risbey, J. S., & Ravetz, J. R. (2008). Exploring the quality of Evidence for Complex and Contested Policy Decisions. *Environmental Research Letters,* 3(024008), 1-9.

Van der Sluijs, J. P., Risbey, J. S., Kloprogge, P., Ravetz, J. R., Funtowicz, S. O., Corral Quintana, S., et al. (2003). *RIVM/MNP Guidance for uncertainty assessment and communication: detailed guidance*. Utrecht: Copernicus Institute for Sustainable Development and Innovation.

van der Sluis, J. P. (1997). *Anchoring Amid Uncertainty: On the Management Of Uncertainties in Risk Assessment of Anthropogenic Climate Change.* University of Utrecht, Utrecht.

van Eeten, M. J. G. (1998). *Dialogues of the Deaf: Defining New Agendas for Environmental Deadlocks*. Dekft: Eburon.

Walker, W. E., Harremoës, J., Rotmans, J. P., Van der Sluijs, J. P., Van Asselt, M. B. A., Janssen, P. H. M., et al. (2003). Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment,* 4(1), 5-17.

Warmink, J. J., Booij, M. J., Van der Klis, H., & Hulscher, S. J. M. H. (2007, 12-15 November). *Uncertainty in Water Level Predictions due to Various Calibrations.* Paper presented at the International Conference on Adaptive and Integrative Water Management, Basel, Switzerland.

Xu, Y. P., Holzhauer, Booij, M. J., & Sun, H. Y. (2008). A two-step approach to investigate the effect of rating curve uncertainty in the Elbe decision support system. *Journal of Zhejiang University - Science a,* 9(9), 1229-1238.

Xu, Y. P., & Tung, Y., K. (2008). Decision-making in Water Management under Uncertainty. *Water Resource Management,* 22(5), 535-550.

4 Adaptive airport strategic planning

This chapter has originally been published as KWAKKEL, J.H., WALKER, W.E. & MARCHAU, V.A.W.J. (2010) Adaptive Airport Strategic Planning. *European Journal of Transportation and Infrastructure Research*, 10(3), 227-250.

Abstract

Airport Strategic Planning (ASP) focuses on the development of plans for the long-term development of an airport. The dominant approach for ASP is Airport Master Planning (AMP). The goal of AMP is to provide a detailed blueprint for how the airport should look in the future, and how it can get there. Since a Master Plan is a static detailed blueprint based on specific assumptions about the future, the plan performs poorly if the real future turns out to be different from the one assumed. With the recent dramatic changes occurring in the context in which an airport operates (e.g., low cost carriers, new types of aircraft, the liberalization and privatization of airlines and airports, fuel price developments, the European Emission Trading Scheme), the uncertainties airports face are bound to increase. Hence, there is a great need for finding new ways to deal with uncertainty in ASP. An alternative direction is to develop an adaptive approach that is flexible and over time can adapt to the changing conditions under which an airport most operate. Three adaptive alternatives to AMP have been discussed in the literature. This paper explores these three alternative approaches. Based on this, it concludes that these approaches are complementary and that it might be worthwhile to combine the three into a new, adaptive approach to ASP. A design that integrates the key ideas from the three alternative approaches is presented and illustrated with a case based on Amsterdam Airport Schiphol.

4.1 Introduction

The aviation industry operates in a fast changing environment. At the end of the 1970's, the air transport industry was liberalized and privatized in the U.S.A. Europe followed in the

1990's. As a result of this privatization and liberalization, the air transport industry has undergone unprecedented changes, exemplified by the rise of airline alliances and low cost carriers (de Neufville & Odoni, 2003; Forsyth, 1998). Parallel to this, the aviation industry has witnessed increasing environmental awareness, which has resulted in more attention being paid to the negative external effects of aviation, such as noise and emissions, and, since 9/11, safety and security are also of more concern. It is likely that the aviation industry will become even more dynamic in the coming years, for example because of the recently signed U.S.A.-Europe Open Skies treaty. All these changes together pose a major challenge for airports. They have to make investment decisions that will shape the future of the airport for many years to come, taking into consideration the many uncertainties that are present.

The current dominant approach for the long-term development of an airport is Airport Master Planning (AMP) (de Neufville & Odoni, 2003). AMP is a formalized, structured planning process that results in a Master Plan that 'presents the planner's conception of the ultimate development of a specific airport'(ICAO, 1987). As such, the focus in AMP is on the development of plans and not on the decisionmaking process about the plans. Admittedly, the decisionmaking process is interwoven with the AMP process, but for analysis purposes we focus here on the AMP process. In the United States, the FAA has set up strict guidelines for an AMP study (FAA, 2005). Internationally, reference manuals of IATA and books about airport planning by leading scholars heavily influence AMP practices (de Neufville & Odoni, 2003; IATA, 2004; ICAO, 1987).

The goal of a Master Plan is to provide a blueprint that will determine the future development of the airport (Burghouwt & Huys, 2003; Dempsey, Goetz, & Szyliowicz, 1997). As such it describes the strategy of an airport operator for the coming years, without specifying operational concepts or management issues. A Master Plan covers both the aeronautical developments (i.e. runways, terminals) and non-aeronautical developments (e.g. real estate, commercial activities, and retail developments) of the airport. The time horizon covered in a Master Plan can vary, depending on the situation of the airport for which the Master Plan is being developed, but in general a time horizon of 20 years is used (FAA, 2005). AMP follows a strict linear process (de Neufville & Odoni, 2003):

- Analyze existing conditions
- Make an aviation demand forecast
- Determine facility requirements needed to accommodate this forecasted demand
- Develop and evaluate several alternatives to meet these facility requirements
- Develop the preferred alternative into a detailed Master Plan

AMP, as the main way to shape and determine the long-term development of an airport, has proven to be ineffective, as can be seen for example in planning failures at Amsterdam Airport Schiphol, Denver International Airport, Boston Logan Airport, and Montréal Mirabel Airport. In 1995, a plan for the long-term development of Amsterdam Airport Schiphol was accepted. This plan had a time horizon of 20 years, but was obsolete in 1999, due to the unanticipated rapid growth of aviation demand (Kwakkel, Walker, & Marchau, 2007). The new Denver Airport was developed because of anticipated growth, which did not materialize. The new airport ended up with fewer air transport movements than took place at the old airport (Dempsey, et al., 1997; Szyliowics & Goetz, 1995). Boston Logan planned and started the construction of a new runway in the early 1970's, but, due to unanticipated changes in regulations and strong stakeholder opposition, they were unable to open this runway until 2006 (Cidell, 2004; Kwakkel, et al., 2007; Nelkin, 1974, 1975). Montréal Mirabel Airport was constructed in 1975 and was forecast to handle 40 million passengers by 2025. However, the airport failed to attract significant travel and was closed for passenger traffic in 2004 (Canadian Press, 2006). Given the ongoing transition of the aviation industry from a state-owned and state-run enterprise to a market situation, with its associated changes in how the public and the government view the aviation industry, the number and severity of the uncertainties is only expected to increase. In light of this, Master Planning becomes even less appropriate for long-term airport planning. In response to this problem of the inadequacy of AMP for ASP, the identification and analysis of alternative approaches has become more urgent. Several alternatives to AMP have been proposed in the literature (Burghouwt, 2007; de Neufville, 2000; Kwakkel, et al., 2007). However, these new approaches are not fully developed, have not been applied in practice, have not yet been compared with each other, nor has their performance compared to AMP been assessed (Burghouwt, 2007). The aim of this paper is to describe and compare the available alternatives to AMP discussed in the airport planning literature, and to synthesize these alternatives into a single approach to ASP that is better equipped to deal with the many and diverse uncertainties airport planners face in the long-term development of an airport.

Following the earlier literature on long-term airport planning, we approach airport planning from a 'research and analyze' perspective (Mayer, van Daalen, & Bots, 2004). Therefore, this paper does not consider the stakeholder and actor related problems associated with airport planning. In order to achieve the aim of comparing and synthesizing the alternative suggested approaches into a single approach to ASP that overcomes the problems associated with APM, we performed a literature review of the problems associated with APM. From these problems we derive criteria that alternative approaches have to meet to be a suitable alternative to APM. This literature review is presented in Section 2. Next, we performed a literature review of the alternative approaches to APM that have been put forward in the literature. We compare these approaches based on criteria derived at the end of Section 2. This review and comparison of alternatives to APM is presented in Section 3. From this comparison it is concluded that all alternatives have different strengths and weaknesses. A better approach can be created by synthesizing the alternatives into a single approach. This synthesized approach is presented in Section 4. Section 5 illustrates this new approach via a case focused on Amsterdam Airport Schiphol. Section 6 contains a discussion. Section 7 presents the conclusions.

4.2 Uncertainty in airport strategic planning

AMP has been unsuccessful in planning the future development of airports. As the examples of Amsterdam Airport Schiphol, Denver International Airport, Boston Logan Airport, and Montréal Mirabel Airport illustrate, plans become quickly obsolete and are not robust with regard to the future. In other words, uncertainty (e.g. aviation demand, regulatory context, technological breakthroughs, and stakeholder behavior) is a key source of problems in ASP. In this section, we explore in more depth how uncertainty is currently treated in AMP, why this treatment is inadequate, and what this implies for alternative treatments.

The main way in which uncertainty is handled in AMP is through aviation demand forecasting. Forecasting has a long history in transport planning in general. The different techniques that are used in forecasting have been debated for a long time. A full review of this literature is beyond the scope of this paper; instead, we focus on the aviation planning literature.

4.2.1 The challenge of uncertainty for airport master planning

The aviation demand forecast is the basis for a new Master Plan. An aviation demand forecast can be a forecast for the number of passengers, the tons of goods, or the number of air transport movements, although the forecast usually contains information concerning all three. For example, the forecast used for the plan for the long-term development of Schiphol in 1995 was a forecast of aviation demand for 2015 in terms of passengers. Given the average number of passengers on an airplane, this was translated into a forecast for air transport movements. By comparing the forecast with the existing conditions at an airport, an assessment can be made whether there is a need for new or expanded facilities. As such, aviation demand forecasting is the main way in which uncertainties about the future context in which an airport operates are handled. The basic concept of developing an aviation demand forecast is simple: past trends, based on time series or theories about underlying mechanisms, are identified and extrapolated forward. In mathematical terms, a relationship between independent variables (X_1, X_2, …, X_n) and the dependent variable (Y) is developed that matches aviation demand observed in the past. The resulting model is then used for extrapolation in order to obtain a forecast for the year of interest (FAA, 2001).

Forecasting in general has come under increasing criticism. The criticisms can be split into two categories: forecasting failure due to bias and forecasting failure due to uncertainty. Forecaster bias contributes to forecast failure in several ways. Forecasters often integrate political wishes into their forecasts (Flyvbjerg et al., 2003). Forecasts by project promoters may be even more biased, since the promoter has an interest in presenting the project in as favorable a light as possible (Flyvbjerg, Bruzelius, & Rothengatter, 2003).

Forecasting failure due to uncertainty manifests itself in several ways. As pointed out by Flyvbjerg *et al.* (2003), discontinuous behavior of the phenomena we try to forecast, unexpected changes in exogenous factors, unexpected political activities, and missing realization of complementary policies are important reasons for forecasting failure. Ascher (1978) sees faulty core assumptions as a prime reason for forecasting failure. Faulty core assumptions refers to the fact that, since the phenomenon we are trying to forecast is not completely understood, forecasters have to make assumptions about the data they need, the formulas to be used, etc. (Porter, Roper, Mason, Rossini, & Banks, 1991). With respect to data, there are also several uncertainties. Forecasters often have a poor database that has internal biases caused by the data collection system, and they use data from their home countries (instead of the local areas) for calibrating their models (Flyvbjerg, et al., 2003). In addition, forecasters have a tendency to misjudge the relevance of (recent) data (Porter, et al., 1991). Despite these problems with data, forecasters still rely heavily on historic data for testing the adequacy of a forecast. However, there are an infinite number of formulas possible that can match the given historical data. Related to this is the fact that, in order to forecast a dependent variable Y based on a formula $Y = f(X_1, X_2, \ldots, X_n)$, forecasts are needed for the future values of the n independent variables. Instead of forecasting a single variable, one ends up forecasting n variables. Even if the problems associated with forecaster bias are addressed, forecasting failure due to uncertainty means that forecasting can always go wrong. By looking at the past and assuming that past behavior will continue into the future, uncertainties leading to trend breaks are overlooked, which, in most cases, are the uncertainties with the largest impacts on the system.

In the case of aviation demand forecasting, forecasting failure due to uncertainty is of specific importance. Over the last twenty years, the aviation industry worldwide has undergone exceptional changes. It has moved from a heavily regulated, state-owned, state-operated industry, towards a fully privatized industry. Currently, aviation transport in the US and Europe is largely privatized, while other regions in the world are moving in this direction as well. The net result of this privatization is that there have been unprecedented changes in the air transport sector, exemplified by the KLM-Air France merger, the rise of airline alliances, the US-EU Open Skies treaty, the rise of low cost carriers (LCC), and fierce competition between airports in order to attract carriers. Burghouwt (2007) has studied how airline networks evolved in Europe over time during these changes and concludes that air traffic demand is becoming more volatile and more uncertain, implying that forecasting air traffic demand for specific airports is becoming ever more problematic.

Apart from the fact that aviation demand forecasting is highly problematic in light of the many uncertainties that are present, there are several additional reasons that make aviation demand forecasting as the main way to treat uncertainty in AMP inadequate. First, usually only a single aviation demand forecast is generated. The airport Master Plan is designed based on this specific forecast. By making only a single forecast, however, one runs the risk of severely underestimating the range within which future aviation demand might develop. Second, there are many uncertainties present when developing plans for the long-term development of an airport. Aviation demand is only one such uncertainty. Other uncertainties include, among others, regulatory developments, technological developments, and demographic developments. Aviation demand forecasting does not consider these, and, as a result, these other uncertainties are often ignored in the AMP process. Third, the Master Plan that results from the AMP process has a blueprint character (Burghouwt & Huys, 2003). The plan is drafted during the planning phase and is then handed over for implementation. During the implementation phase, the plan is implemented without much consideration for changing conditions. As a result, the Master Plan is static in nature and leaves little room for adapting to changing conditions. An analysis of the current long-term planning process of Amsterdam Airport Schiphol revealed that many uncertainties in addition to demand are not explicitly treated, and of those uncertainties that are addressed, most are addressed by making specific assumptions that are just estimates, rather than ranges of values (Kwakkel, Walker, & Wijnen, 2008).

4.2.2 Criteria for a new planning approach

We can summarize the preceding discussion by saying that AMP is inadequate for the long-term development of airports, because the resulting plan is not robust with respect to future developments. This lack of robustness is the result of the fact that (a) very few uncertainties are addressed – usually only aviation demand uncertainties; (b) only a single future demand is considered, instead of a range of plausible demands; and (c) a Master Plan is static. In light of this, an alternative planning approach for long-term development that would deal better with the many uncertainties airport planners face should be designed to meet several criteria:

- the planning approach should consider many different types of uncertainties, in addition to demand uncertainties;
- the planning approach should consider many different plausible futures;
- the resulting plan should be robust across the different futures;
- the resulting plan should be flexible.

Admittedly, there are many other criteria that a planning approach for the long-term development of an airport should meet, which relate more to implementation of the new planning approach. For example, the approach should be easy to execute, not require too many resources, not be too time consuming, consider the different stakeholders that are affected, contain arrangements for stakeholder involvement, and so on. However, in this paper we are concerned primarily with the problems uncertainty causes for the long-term planning of airport development. Hence, we are interested in finding a new approach that can address uncertainty better than the current Master Planning approach. For the purposes of this paper, therefore, we do not consider these additional criteria (but, we do keep them in mind).

4.3 Adaptive approaches for airport strategic planning

Initial ideas on adaptive policies are found early in the 1900s. Dewey (1927) put forth an argument proposing that policies be treated as experiments, with the aim of promoting continual learning and adaptation in response to experience over time (Busenberg, 2001). Early applications of adaptive policies, called Adaptive Management, can be found in the field of environmental management (Holling, 1978). Motivated by the complexity of the environmental system, managers resort to controlled experiments aimed at increasing their understanding of the system (McLain & Lee, 1996). Or, as Lee (1993) puts it, adaptive policies are ‘designed from the outset to test clearly formulated hypotheses about the behavior of an ecosystem being changed by human use’.

A recent development that is related to adaptive policies is the discussion about ‘deep uncertainty’ and its implications for the development of robust long-term policies. Decisionmaking under deep uncertainty is understood as a situation in which decisionmakers do not know or cannot agree on a system model, the prior probabilities for the uncertain parameters of the system model, and/or how to value the outcomes (R.J. Lempert, Popper, & Bankes, 2002). Lempert (2002) presents exploratory modeling as a method for the systematic analysis of large ensembles of potential futures. This method can be used to identify key uncertainties that influence policy performance, thus enabling the policymakers to improve the robustness of their policies. The main area of application of exploratory modeling has been in the field of climate change (R. J. Lempert, Popper, & Bankes, 2003).

Recently, Walker et al. (2001) developed a structured, stepwise approach for adaptive policymaking. This approach differs from adaptive approaches in the field of environmental management in that the key sources of uncertainty are external forces outside the control of the policymakers, instead of arising out of the complexity of the system the policymakers are trying to manage. Since the sources of uncertainty are different, the approach also differs in several important respects from Adaptive Management. Most importantly, the approach advocates not only the development of a monitoring system but also the pre-specification of responses when specific trigger values are reached. This is now called “planned adaptation”.

Scientific work in the field of adaptive policies starts from the explicit recognition of the many and severe uncertainties decisionmakers face. Instead of predicting what will happen, which is impossible in light of these uncertainties, these researchers try to develop policymaking approaches that allow implementation to begin prior to the resolution of all major uncertainties, with the policy being adapted over time based on new knowledge. Adaptation is an innovative way to proceed with the implementation of long-term (transport) polices despite the uncertainties. These policymaking approaches make adaptation explicit at the outset of policy formulation. Thus, the inevitable policy changes become part of a larger, recognized process and are not forced to be made repeatedly on an *ad-hoc* basis. Adaptive policies combine actions that are time urgent with those that make important commitments to shape the future, preserve needed flexibility for the future, and protect the policy from failure. In case of ASP, there are three alternatives to AMP discussed in the airport planning literature, all based on concepts of flexibility and adaptability. These three are de Neufville’s Dynamic Strategic Planning (de Neufville, 2000, 2003; de Neufville & Odoni, 2003), Adaptive Policymaking (Kwakkel, et al., 2007) , and Flexible Strategic Planning (Burghouwt, 2007). Below we discuss these three approaches in more detail.

4.3.1 Dynamic Strategic Planning

Dynamic Strategic Planning (DSP) offers a new approach to AMP, although it is still based on traditional systems analysis, which is also at the heart of AMP (de Neufville & Odoni, 2003). DSP is an approach for making plans, particularly for infrastructure, that can be easily adjusted over time to the actual situation and conditions. In this way, bad situations can be avoided and opportunities can be seized. The resulting dynamic strategic plan defines a flexible development over several stages; it commits only to a first stage, and then proposes different developments in the second and subsequent stages. DSP recognizes that the future cannot be anticipated accurately, and hence that all forecasts will be wrong. Therefore, a plan should build in flexibility to deal effectively with a range of futures. In DSP, this flexibility is created through real options (de Neufville, 2000). An option is a right, but not an obligation, to take an action for a certain cost at some time in the future, usually for a predetermined price and a given period (de Neufville, 2003). A well-known example of a real option is to make a land use reservation (this is also known as 'land banking'). Such a reservation offers planners the option to expand infrastructure in the future if this turns out to be needed.

There is no clear prescribed process for performing DSP, although there are seven distinct categories of methods and activities that together will result in a dynamic strategic plan. These are (de Neufville, 2000):

1. Modeling: this activity should result in one or more models of the technical system and its performance.
2. Optimization: this activity should result in an overview of different cost-effective means for achieving specified levels of results.
3. Estimation of probabilities: since the performance of a system in the future cannot be forecast, it is necessary to estimate the range of values for key system parameters and the likely probability distributions for these parameters.
4. Decision Analysis: by combining the results from the previous three activities, a Decision Analysis for the set of choices can be carried out.
5. Sensitivity Analysis: this activity should make sure that the outcome of the Decision Analysis is robust with regard to changes in parameter values.
6. Evaluation of Real Options: this activity should focus on identifying cost-effective real options that increase the flexibility of the plan. These can then be inserted into the Decision Analysis.
7. Analysis of implicit negotiation: the implementation of a plan is to a large extent dependent on the support of relevant stakeholders. This activity aims at analyzing the stakeholders and their possible behavior. The results are to be taken into account when thinking about the implementation of the plan that is developed through activities 1-6.

4.3.2 Adaptive Policymaking

Adaptive Policymaking (APM) is proposed as a generic approach for the treatment of uncertainty. It recognizes that, in a rapidly changing world, fixed static policies are likely to fail. Over time, however, we learn, reducing the uncertainty. To plan effectively in such a changing world, therefore, we should plan adaptively and allow for this learning (Walker, 2000; Walker, et al., 2001). The APM process is split into two phases: a thinking phase, during which the adaptive policy is developed, and an implementation phase, during which the policy is implemented, its performance monitored, and the policy adapted if necessary. During the thinking phase, a basic policy is designed and subsequently analyzed for vulnerabilities (i.e. plausible events or developments that would hamper the performance of the plan). The identified vulnerabilities are screened on the level of uncertainty. The relatively certain vulnerabilities are taken into account in the basic policy by including mitigating actions that should be taken when starting the implementation of the basic policy. For some of the uncertain vulnerabilities, hedging actions are implemented to make the basic policy more robust. In addition, a monitoring system is created for uncertain vulnerabilities, and actions are prepared to be taken when the monitoring reveals that specific vulnerabilities have manifested themselves. During the implementation phase, events unfold, the signposts are monitored, and defensive or corrective actions are taken if necessary. The implemented policy remains active as long as the signposts signify that the policy is on course to achieve its intended outcomes. Otherwise, a reassessment of the policy is necessary.

4.3.3 Flexible Strategic Planning

Flexible Strategic Planning (FSP) has been suggested as an alternative to traditional AMP by Burghouwt (2007). He suggests that, in light of the inability to forecast future traffic accurately as a result of the increasing volatility of aviation demand and airline network development, a more flexible and pro-active planning style is necessary. FSP draws heavily on DSP, but adds to this the notion of pro-active planning. An airport should try and shape the future through its own actions. In order to realize a flexible strategic plan for an airport, FSP relies on real options, scenario style robustness, back casting, contingency planning, monitoring, experimentation, and diversification. The discussion in Burghouwt (2007), however, is very brief. Exactly how FSP should work and how it could be applied in practice remain open issues. Burghouwt (2007) explicitly acknowledges this and adds that there is little empirical evidence to support a flexible adaptive approach; the creation of flexibility and adaptability is often difficult in light of the stakeholders affected by the airport, and more sophisticated tools are needed to support airport planners using the flexible approach than are needed for traditional AMP.

4.3.4 A comparison of the three approaches

Table 4-1 below gives an overview of some key characteristics of the three Adaptive Planning approaches. As a first element for comparison, we consider the focus of each approach. In light of the criteria for a new planning approach specified in Section 2.2, it is also relevant to identify the types of uncertainty that are considered in the three alternative planning approaches (Criterion 1), whether multiple futures are considered (Criterion 2), whether the resulting plan is robust (Criterion 3), and how flexibility is guaranteed in the plan (Criterion 4). Given our stated goal of proposing a new Airport Strategic Planning approach, it is also relevant to take into consideration to what extent these the three approaches provide a clear planning process. Table 4-1 can be used as a starting point for analyzing, comparing, and identifying a promising alternative planning approach to AMP.

Table 4-1: Comparison of three approaches for Adaptive Planning

Aspect	Dynamic Strategic Planning	Adaptive Policy Making	Flexible Strategic Planning
Focus	Flexibility in a plan created through real options	Starts from a vision of the decisionmaker and creates a plan for realizing this vision and protecting it from failure	Extends the focus of DSP by adding pro-active planning and contingency planning
Types of uncertainties considered	Emphasis on demand uncertainty, but other types of uncertainties could be considered as well via real options	Any uncertainty can be considered	Emphasis on demand uncertainties as driven by airline network developments, but in principle open to all types of uncertainties
Consideration of different futures	Via a staged development	Via hedging and mitigating actions	Via scenario robustness
Robustness of the resulting plan	No direct consideration of robustness, but a range of futures can be handled via real options	Explicit consideration of increasing robustness of plan via hedging and mitigating actions	Explicit consideration via use of scenarios
Flexibility of resulting plan	Flexibility of plan is guaranteed via real options	Flexibility of the plan is addressed via the establishment of a monitoring system and pre-specification of responses	Flexibility of the plan is guaranteed via real options and contingency planning
Planning process	Seven categories of activities specified, but their relationships to each other and how they constitute a planning process remain unclear	Has a clear planning process, with a distinction between a thinking phase and an implementation phase	No clear process is specified

As can be seen in Table 4-1, all three approaches meet the criteria specified in Section 2.2. The treatment of uncertainty in all three approaches moves beyond demand forecasting and makes use of additional techniques, such as real options, hedging, scenarios, and pro-activeness. These techniques all aim at making the plan more robust with respect to uncertainty about the future. The three approaches can consider all types of uncertainty, although DSP and FSP focus mainly on demand uncertainties. The three approaches can also be used to consider multiple different futures, although the way in which this is done differs. FSP explicitly includes the idea of multiple futures, since it intends to make use of scenarios. DSP considers multiple futures by only committing to a first stage of development while preparing different actions for future stages of development. In addition, with its insistence on forecasting failure, DSP also emphasizes the need for multiple forecasts based on different assumptions about future external developments. The idea of committing to a first set of actions while preparing others in advance can also be found in APM and is its main way of dealing with multiple futures. The three approaches differ with respect to the presence or absence of a clear planning process. FSP does not provide any description of a process. DSP provides only several categories of activities. Only APM has a clear process and framework that, if followed carefully, will result in a complete Adaptive Plan that can be implemented by policymakers.

The idea of committing to a first set of actions while preparing others in advance is of specific importance in infrastructure planning and development, because of the time it takes to build new infrastructure. For example, in the case of airport development, implementing a new runway can take ten years or more. Over this period however, significant changes can occur that would render the investment superfluous. Adaptive approaches suggest that, where possible, the investments should be phased. To continue the example of runway expansion, in the first phase, the land could be acquired. Next, if the runway still appears to be necessary, the groundwork could be carried out. If, after this, the runway is still required, the next phase of construction could start. By phasing the infrastructure investment, the risks of superfluous investments can at least be partly mitigated. By phasing the development of new infrastructure, the flexibility to respond to changing conditions is retained, reducing the risk of unnecessary investments.

Based on Table 4-1, we conclude that the three approaches are all capable of dealing with the many and diverse uncertainties airport planners face, although this capability is realized in different ways. DSP uses real options as the main mechanism to create a flexible plan. APM forces planners to consider many and diverse uncertainties and to prepare for these in advance through hedging and mitigating actions. The successful execution of the plan is also taken into account via the pre-specified monitoring system. FSP is perhaps the broadest in terms of the available ideas and notions for the treatment of uncertainty, with its discussion of robustness, hedging, diversification of revenues, and its insistence on pro-activeness. FSP is, however, also the least developed planning approach in terms of its operationalization. In light of the different angles by which the approaches address uncertainty, it is important to note that these angles are not contradictory. Real options, for example can be used as a means for creating a mitigating or hedging action in the context of APM. In light of all the above, it appears that it might be possible to design an improved approach for ASP by combining ideas from these three approaches. This synthesis can draw on the relative strengths of the different approaches, such that the resulting synthesis is better equipped to overcome the weaknesses of AMP than any of the three approaches individually. Another benefit of designing a single approach is that researchers can concentrate their efforts on further developing and testing this single approach instead of spreading their efforts over all of them or focusing on one of them. This is especially relevant in light of the fact that all three approaches are still in their conceptual stages and significant work is required before any of the approaches can be used in practice, as noted by Burghouwt (2007) and (Hansman, Magee, De Neufville, Robins, & Roos, 2006).

4.4 A synthesized approach to adaptive airport strategic planning

In light of the fact that only APM has a well-developed planning process, we use it as the starting point for developing an integrated adaptive approach for ASP, which we call Adaptive Airport Strategic Planning (AASP). The main idea from DSP is the real options concept. Real options in the context of APM can be used as a means to create adaptive actions (e.g. hedging actions). A key idea of FSP that is not explicitly part of APM is the notion of proactive planning. APM can be expanded to cover this in a straightforward manner by recognizing that uncertain future developments can be two sided. Some external changes can cause a policy to fail, while other changes can make a policy more successful. So, the future presents a strategic planner both with vulnerabilities that can cause a policy to fail and with opportunities that can improve the policy's success. Pro-activeness can then be integrated into APM by including actions that try to shape the nature of a vulnerability, and by including actions that aim at taking advantage of opportunities when they present themselves. Other relevant ideas from FSP are robustness and contingency planning, although both are already incorporated in APM. Robustness is covered in the form of hedging and mitigating actions. Contingency planning is present in the form of the monitoring system and its associated actions that are triggered if threshold values are reached on the signposts. Figure 4-1 presents the expanded APM framework. Note that in this paper the term 'policy' does not refer to government policies, such as regulations, but refers to airport plans or strategies, such as adding a runway or building a new terminal. In the context of adaptive policymaking, such plans can also contain actions that prepare the airport for the future without directly changing the system. This is in contrast to traditional policy analysis, in which policies are the set of forces within the control of the actors in the policy domain that affect the structure and performance of the system (Walker, 2000).

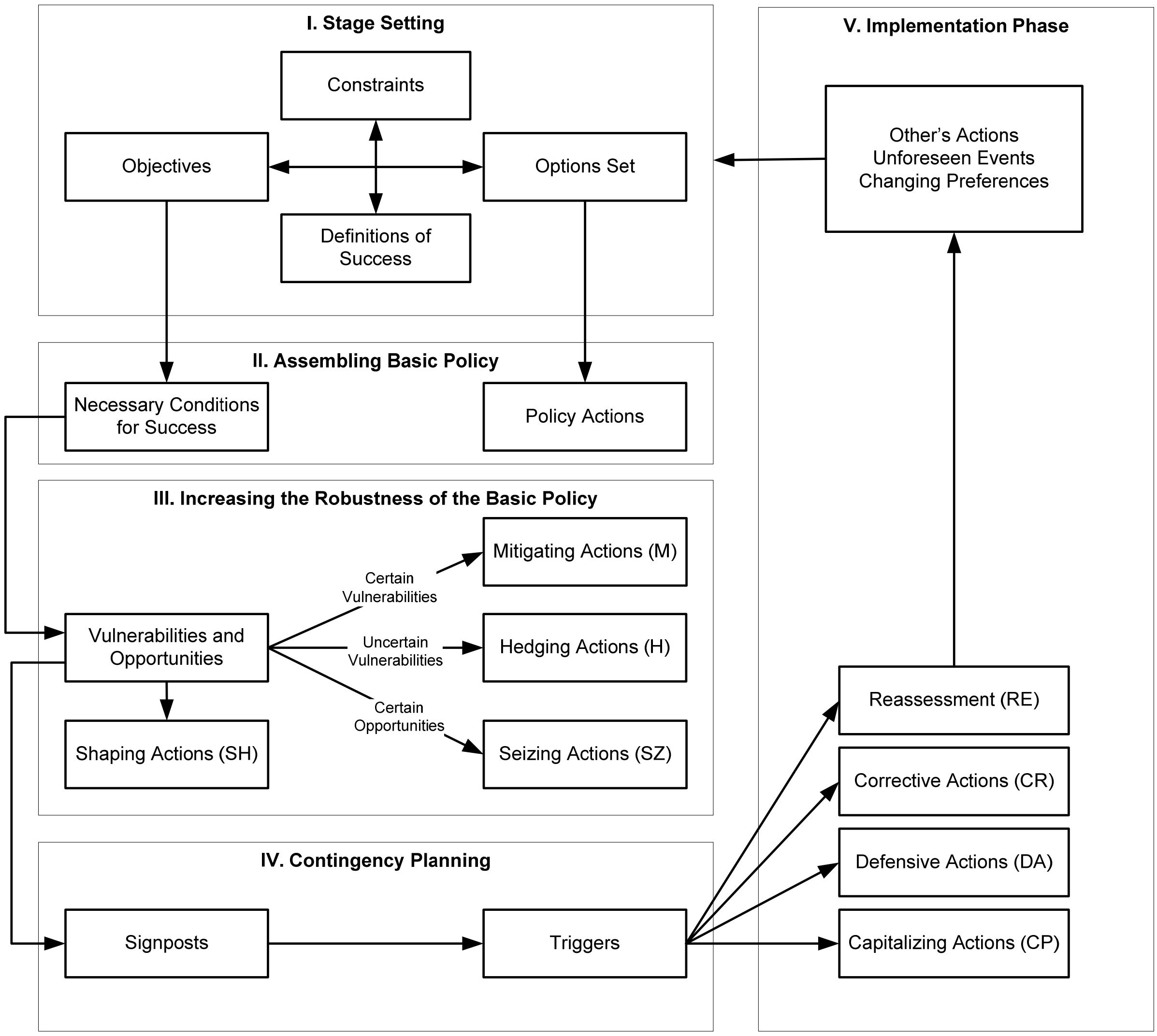


Figure 4-1: The steps of Adaptive Airport Strategic Planning (AASP)

4.4.1 Step I (stage setting) and step II (assembling the basic policy)

Both the first and second steps are similar to the current steps in AMP. The first step constitutes the stage-setting step (e.g. analyzing the existing conditions of an airport). This step involves the specification of objectives, constraints, and available policy options (e.g. expand the terminal, build a new terminal, add a new runway, or extend an existing runway). This specification should lead to a definition of success, in terms of the specification of desirable outcomes (e.g. desired noise levels, number of houses in noise contours, number of air transport movements served at the airport, minimum average delay of aircrafts). In the next step, a basic policy is assembled. It involves (a) the specification of a promising policy and (b) the identification of the conditions needed for the basic policy to succeed.

4.4.2 Step III (robustness)

In the third step of the adaptive policymaking process, the robustness of the basic policy is increased. This step is based on identifying in advance the vulnerabilities and opportunities associated with the basic policy, and specifying actions to be taken in anticipation or in response to them. The key element of this step is the identification of vulnerabilities and opportunities. Vulnerabilities are possible developments that can degrade the performance of a policy so that it is no longer successful. Opportunities are developments that can increase the success of the policy. For example, an important vulnerability of most airport Master Plans is that demand turns out to be lower than anticipated, rendering investment in capacity expansion superfluous. But, demand might also develop rapidly, allowing the airport to expand faster then anticipated. In this case, the same uncertain external development can be both a vulnerability and an opportunity. There are two basic ways of preparing a policy for vulnerabilities and opportunities, either by taking actions now, or by preparing actions in advance that can be taken in the future if necessary (the latter is considered in Step IV). There are four different types of actions that can be taken in advance in anticipation of specific contingencies or expected effects of the basic policy:

- *mitigating actions (M)* – actions to reduce the *certain* adverse effects of a policy;
- *hedging actions* (H) – actions to spread or reduce the risk of *uncertain* adverse effects of a policy;
- *seizing actions (SZ)* – actions taken to seize certain available opportunities;
- *shaping actions (SH)* – actions taken to reduce the chance that an external condition or event that could make the policy fail will occur, or to increase the chance that an external condition or event that could make the policy succeed will occur.

Mitigating actions and hedging actions prepare the basic policy for potential adverse effects and in this way try to make this policy more robust. Seizing actions are actions taken now to change the policy in order to seize available opportunities. In contrast, shaping actions are pro-active and aim at affecting external forces in order to reduce the chances of negative outcomes or to increase the chances of positive outcomes. As such, shaping actions aim not so much at making the plan more robust, but at changing the external situation in order to change the nature of the vulnerability or opportunity. For example, marketing is an attempt to increase the demand for a given product. In this way, one tries to prevent insufficient demand for the product (Dewar, 2002). Real options can be used as a technique for any of these four types of actions. For example, if an airport plans to expand its terminal capacity, it faces the vulnerability of insufficient demand. A real option's design of the terminal (e.g. in a modular way) is then a hedging action against insufficient demand.

4.4.3 Step IV (contingency planning)

Even with the actions taken in advance, there is still the need to monitor the performance of the policy and take action if necessary. In the fourth step, the policy is further expanded via contingency planning, in which the robust basic policy is further enhanced by including adaptive elements. The first element of the contingency plan is the identification of signposts. Signposts specify information that should be tracked in order to determine whether the policy is achieving its conditions for success. The starting point for the identification of signposts is the set of vulnerabilities and opportunities specified in Step III. Critical values of signpost variables (triggers) are specified, beyond which actions should be implemented to ensure that a policy keeps moving the system in the right direction and at a proper speed. Some of these actions might be prepared in advance and might require a change to the basic policy. To continue our example of the terminal from Step III, the development of demand is something the airport should monitor closely. In light of how the demand develops, new modules can be added to the modular terminal. If demand grows rapidly, the terminal can easily be expanded, while if demand does not grow as fast as anticipated, further extensions can be delayed. As is shown by this example, real options can also form part of the contingency planning part of the plan. Again, the opportunity side of the vulnerabilities should also be considered in this step.

There are four different types of actions that can be triggered by a signpost:

- *defensive actions (DA)* – actions taken *after the fact* to clarify the policy, preserve its benefits, or meet outside challenges in response to specific triggers that leave the basic policy remains unchanged;
- *corrective actions (CR)* – adjustments to the basic policy in response to specific triggers;
- *capitalizing actions (CP)* – actions taken *after the fact* to take advantage of opportunities that further improve the performance of the basic policy;
- *reassessment (RE)* – a process to be initiated or restarted when the analysis and assumptions critical to the policy's success have clearly lost validity.

4.4.4 Step V (implementation)

Once the basic policy and additional actions are agreed upon, the final step involves implementing this entire plan. In this step, the actions to be taken immediately (from Step II and Step III) are implemented and a monitoring system (from Step IV) is established. Then time starts running, signpost information related to the triggers is collected, and policy actions are started, altered, stopped, or expanded. After implementation of the initial mitigating, hedging, seizing, and shaping actions, the adaptive policymaking process is suspended until a trigger event occurs. As long as the original policy objectives and constraints remain in place, the responses to a trigger event have a defensive or corrective character – that is, they are adjustments to the basic policy that preserve its benefits or meet outside challenges. Sometimes, opportunities are identified by the monitoring system, triggering the implementation of capitalizing actions. Under some circumstances, neither defensive nor corrective actions might be sufficient to save the policy. In that case, the entire policy might have to be reassessed and substantially changed or even abandoned. If so, however, the next policy deliberations would benefit from the previous experiences. The knowledge gathered in the initial adaptive policymaking process on outcomes, objectives, measures, preferences of stakeholders, etc., would be available and would accelerate the new policymaking process.

4.5 Application of Adaptive Airport Strategic Planning to the case of Schiphol airport

In this section, we illustrate the approach outlined in the previous section through a case. For an effective illustration, a single in-depth case is preferred over several small cases. Given that AASP is intended to improve upon AMP under conditions of uncertainty, the case needs to have a multitude of different uncertainties. These uncertainties should cover the full range of uncertainties to which airports around the world are exposed. We choose to use the current challenges Schiphol faces in its long-term development as our case. As outlined below, Schiphol faces a range of uncertainties that could affect the airport in different ways. In addition, we are familiar with the current situation of Schiphol: the uncertainties the airport currently faces have been studied recently (Kwakkel, et al., 2008), and a multitude of policy documents from multiple stakeholders is available (e.g. CPB, KiM, NMP, & RP, 2007; Provincie Noord-Holland, 2007; Rijksoverheid, 2009; Schiphol Group, 2007; Schiphol Group & LVNL, 2007; V&W, 2007; V&W & VROM, 2007).

Aviation demand has experienced unprecedented growth since the early 1990's, fuelled by privatization and liberalization of the aviation industry. Amsterdam Airport Schiphol has benefited from this growth and has evolved into one of the European Union's major hubs. Since 1990, Schiphol has expanded its runway system and its terminal. Parallel to the increasing number of passengers and flights handled at Schiphol, negative external effects have also increased, resulting in regulations concerning noise, emissions, and third-party risk.

Currently, Schiphol's position as a hub within Europe is under pressure. In 2006, Schiphol was surpassed by Madrid's Barajas Airport and now ranks as Europe's fifth airport in terms of air transport movements. The merger of Air France and KLM has resulted in the threat that KLM, Schiphol's hub carrier, which is responsible for 52% of the scheduled aircraft movements at the airport, might move a significant portion of its operations to Charles de Gaulle Airport. The other major airports in Europe are planning on expanding their capacity or are developing dual airport systems, while Schiphol's capacity is under threat of being reduced do to climate change induced deterioration of wind conditions. Together, this makes the long-term planning for Schiphol both urgent and problematic.

In the remainder of this section, we illustrate how each of the steps of the new adaptive policymaking approach might be applied to the case of the long-term development of Schiphol. The purpose of this extensive case is two-fold. First, it is intended to illustrate the adaptive approach described in Section 4 and clarify how the concepts could be applied in practice. Second, it serves as a first face-validation of the outlined approach. To give the reader a sense of how this approach could work in the real world, we use the example of a real airport (Schiphol) instead of a made-up airport. However, to make the approach clear and understandable, the example simplifies some of the key challenges Schiphol faces. Therefore, this case should not be understood as presenting a realistic plan for the long-term development of Schiphol. It is merely an example loosely based on real policy issues and policy debates that policymakers are currently facing with respect to the long-term development of an airport.

4.5.1 Step I: specification of objectives, constraints, and available policy options

The Schiphol Group is primarily interested in medium- to long-term developments until 2020. As outlined in its current long-term vision (Schiphol Group & LVNL, 2007), the main goals of the Schiphol Group are: (1) to create room for the further development of the network of KLM and its Skyteam partners, and (2) to minimize (and, where possible, reduce) the negative effects of aviation in the region. Underlying the first goal is the implicit assumption that aviation will continue to grow. However, in light of recent developments such as peak oil and the financial crisis, this assumption is questionable. It might be better to rephrase this first goal more neutrally as 'retain market share'. If aviation in Europe grows, Schiphol will have to accommodate more demand in order to retain its market share, while if aviation declines, Schiphol could still reach its goal of retaining market share.

There are several types of changes that can be made at Schiphol in order to achieve its goals of retaining market share and minimizing the negative effects of aviation. Schiphol can expand its capacity by using its existing capacity more efficiently and/or building new capacity. It can also expand its capacity or use its existing capacity in a way that mitigates the negative effects of aviation. More explicitly, among the policy options that Schiphol might consider are:

1. Add a new runway
2. Add a new terminal
3. Use the existing runway system in a more efficient way, in order to improve capacity
4. Use the existing runway system in a way that minimizes noise impacts
5. Move charter operations out of Schiphol (e.g., to Lelystad)
6. Move Schiphol operations to a new airport (e.g., in the North Sea)
7. Invest in noise insulation

Some of these policies can be implemented immediately (e.g., using the existing runway system in a more efficient way). For others, an adaptive approach would be to begin to prepare plans and designs (e.g., for a new runway), but to begin actual building only when conditions show it to be necessary (i.e., when it is triggered). The various options can, of course, be combined. The changes that can be made are constrained by costs, spatial restrictions, public acceptance, and the landside accessibility of Schiphol. The definition of success includes that Schiphol maintains its market share and that living conditions improve compared to some reference situation (e.g. number of people affected by noise within a specified area).

4.5.2 Step II: basic policy and its conditions for success

A basic policy might be to immediately implement existing plans for using the runways more efficiently (option 3) and in a way that reduces noise impacts (option 4). It might also include all policy options that focus on planning capacity expansions, without beginning to build any of them (i.e. options 1, 2, and 5). A final element of the basic policy would be option 7: invest in noise insulation. The choice for only planning capacity expansions but not yet building them is motivated by the fact that Schiphol is currently constrained by the environmental rules and regulations, not by its physical capacity. This also motivates the choice for options 3 and 4, which together can reduce the negative externalities of aviation.

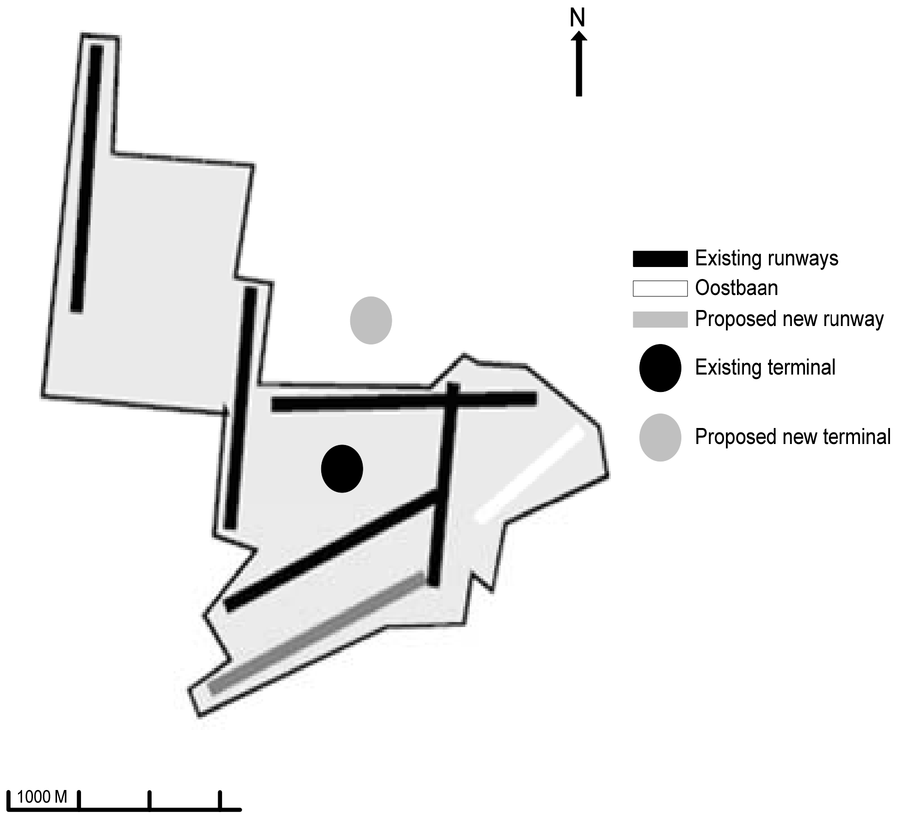


Figure 4-2: The planned extensions to Amsterdam Airport Schiphol

The most discussed option for the new runway is to place it parallel to one of the existing runways – the Kaagbaan. There are several arguments for this choice. First, it would improve the capacity of Schiphol under crosswind conditions in case of a southwesterly storm, making the peak-hour capacity throughout the year more sustainable. This is of particular importance given climate change induced changes in wind regime. Currently, the Oostbaan is used for landings under these crosswind conditions. However, incoming aircraft that use the Oostbaan have to come in over the center of Amsterdam, which produces significant noise nuisance. Furthermore, the Oostbaan cannot handle the larger passenger aircraft. Schiphol would prefer to replace the Oostbaan with the new parallel Kaagbaan. This runway would create less noise nuisance, can handle larger aircraft, can be operated independently from the current Kaagbaan, and can, if necessary, also be used for take-off operations.

The current terminal design of Schiphol is based on a single terminal concept. The available space for further expanding the existing terminal is, however, limited. Therefore, if Schiphol wants to expand its terminal capacity, a new terminal on a different location has to be developed. The basic policy could include a plan for a separate dedicated terminal for LCC's located away from the existing terminal (black circle in Figure 4-2). This is an option that Schiphol is currently considering.

In addition to these capacity expansions, Schiphol might begin to develop plans to move charter operations to Lelystad airport, which would reduce noise around Schiphol and increase Schiphol's capacity for regular flight operations. In order to realize such a move, Lelystad Airport would need to be expanded considerably, so planning should be started right away. Charter operations should then be moved there as soon as possible. In the short run this would create additional capacity and reduce noise at the edges of the night, which is favorable for Schiphol, because the current noise regulation system heavily penalizes flights in the evening (19.00-23.00) and during the night (23.00-07.00).

In light of Schiphol's goals (retaining market share and minimizing the negative effects of aviation (Schiphol Group & LVNL, 2007)), several necessary conditions for the success of the basic policy can be specified:

- Schiphol should retain its current market share
- The population affected by noise and the number of noise complaints should not increase
- Schiphol's competitive position in terms of available capacity in Europe should not decrease
- Schiphol's landside accessibility should not deteriorate

4.5.3 Step III: vulnerabilities and opportunities of the basic policy, and anticipatory actions

The long-term development of Schiphol is complicated by the many and diverse trends and developments that can affect Schiphol. These developments and trends present both opportunities and vulnerabilities. Some of these vulnerabilities are relatively certain. These are given in Table 2. Two certain vulnerabilities are resistance from stakeholders and a reduction of the landside accessibility. The mitigating actions for addressing these vulnerabilities are very similar to actions currently being discussed by the Government (V&W & VROM, 2007). A shaping action for the vulnerability of landside accessibility is investment in research. In addition to vulnerabilities, there are currently also some opportunities available to Schiphol. First, recent work shows the potential for 'self-hubbing'. Self-hubbing means that passengers arrange their own flights and routes, using low cost carriers or a variety of alliances, in order to minimize costs and/or travel time. Schiphol has a great potential for attracting such self-hubbing passengers because it connects 411 European cities. Schiphol can seize this opportunity by developing and implementing services tailored to self-hubbing passengers, such as services for baggage transfer and help with acquiring boarding passes. Furthermore, Schiphol could take into account walking distances between connecting European flights when allocating aircraft to gates. A second opportunity is presented by the fact that airports in general, and Schiphol in particular, are evolving into 'airport cities'. Given the good transport connections available, an airport is a prime location for office buildings. Schiphol can seize this opportunity by investing in non-aeronautical landside real estate development.

Table 4-2: Certain vulnerabilities, and responses to them

Vulnerabilities and Opportunities	Mitigating (M), Shaping (SH) and Seizing (SZ) Actions
Reduction of the landside accessibility of the airport.	**M:** develop a system for early check-in and handling of baggage at rail-stations **SH:** invest in R&D into the landside accessibility of the Randstad area.
Resistance from Schiphol stakeholders (e.g. environmental groups, people living around Schiphol)	**M:** develop plans for green areas to compensate for environmental losses. **M:** offer financial compensation to residents in the high noise zone
Rise of self-hubbing	**SZ:** design and implement a plan for supporting self-hubbing passengers with finding connection flights, transferring baggage, and acquiring boarding passes
Rise of the airport city	**SZ:** Diversify revenues by developing non-aeronautical landside real estate

Not all vulnerabilities and opportunities are certain. The real challenge for the long-term development of Schiphol is presented by the uncertain vulnerabilities and opportunities. Table 3 presents some of the uncertain vulnerabilities together with possible hedging (H) and shaping actions (SH) to take right away to handle them. The vulnerabilities and opportunities can be directly related and categorized according to the success conditions specified in the previous step. With respect to the success condition of growing demand, air transport demand might develop significantly different from what is hoped and anticipated. Schiphol can respond to this development by making Lelystad airport suitable for handling non−hub−essential flights. Another vulnerability is that KLM might decide to move a significant part of its operations to Charles de Gaulle. This will leave Schiphol without its hub carrier, significantly reducing demand, and changing the demand to origin-destination demand. Schiphol could prepare for this vulnerability by making plans for adapting the terminal to the requirements of an O/D airport and by diversifying the carriers that serve Schiphol. Schiphol can also try to directly affect KLM by investing in a good working relationship, reducing the chance that KLM will leave. Currently, there is an ongoing debate about the future of the hub-and-spoke network structure. Due to the Open Sky agreements and the development of the Boeing 787, long-haul low-cost, hub bypassing, and self-hubbing become plausible, resulting in the emergence of long-haul low-cost carriers and increasing transfer between short−haul low−cost, and long−haul carriers (both LCC and legacy carriers). Schiphol can prepare for this by developing a plan to change its current terminal to serve a different type of demand and by taking these plausible developments into consideration when designing the new LCC terminal and its connection with the existing terminal. If a transformation to international origin-destination traffic and/or a no-frills airport is needed, this plan can be implemented, making sure that the transformation can be achieved quickly

The second success condition is that the population affected by noise and the number of noise complaints should not increase. Vulnerabilities and opportunities associated with this condition are that the current trend of decrease of environmental impact of aircraft changes, the population density in the area affected by noise increases, and the valuation of externalities (predominantly noise) by the large public changes. If the current trend of decreasing environmental impact slows down, the area affected by noise will not continue to shrink if demand stays the same. If demand increases, it is possible that the area affected by noise will also increase. On the other hand, the trend could also accelerate, giving Schiphol the opportunity to expand the number of flights that is handled. Given the potential impact of this trend, Schiphol should try and shape its development by investing in R&D and negotiate with Air Traffic Control about testing noise abatement procedures, such as continuous descent approaches. If the population density changes, the situation is similar. If it increases, the number of people affected by noise will increase, while if it decreases, the number of people affected by noise will decrease. Schiphol can try and shape this development by negotiating with surrounding communities about their land use planning and invest in research that can make the area affected by noise smaller. It can also hedge against a growing population density by starting to test noise abatement procedures outside peak hours. This will make the area affected by noise smaller. Thus even if the population density increases, the total number of people affected will not increase. A third uncertainty is how the valuation of noise will change in the future. If noise will be considered more of a nuisance, complaints are likely to go up, and vice versa. Schiphol could try to affect this valuation by branding the airport as environmentally friendly and support the development of an emission trading scheme that also includes aviation.

The third success condition is that Schiphol's competitive position in terms of available capacity in Europe does not decrease. Schiphol is vulnerable to the capacity developments at other airports in Europe. The major hubs in Europe are all working on expanding their capacities, either by adding runways and expanding terminals, or by moving non-hub-essential flights to alternative airports in the region. Schiphol should monitor these developments closely and, if necessary, speed up its capacity investments. A second vulnerability is the robustness of Schiphol's peak-hour capacity across weather conditions. Under southwesterly wind conditions, Schiphol's hourly capacity is almost halved, resulting in delays and cancellations. If (e.g., due to climate change) these wind conditions were to become more frequent, Schiphol would no longer be able to guarantee its capacity. Schiphol should hedge against this by having plans ready for building the sixth runway.

Table 4-3: Uncertain vulnerabilities and opportunities, and responses to them

Vulnerabilities and Opportunities	Hedging(H) and Shaping(SH) Actions
Retain market share	
Demand for air traffic grows faster than forecast.	**H**: Prepare Lelystad airport to receive charter flights
Demand for air traffic grows slower than forecast.	**SH:** Advertise for flying from Schiphol
Collapse or departure of the hub carrier (KLM) from Schiphol.	**H:** Prepare to adapt Schiphol to be an O/D airport. **H:** Diversify the carriers serving Schiphol **SH:** Develop a close working relation with KLM
Rise of long-haul low-cost carriers	**H:** Design existing and new LCC terminal to allow for rapid customization to airline wishes
Rise of self-hubbing, resulting in increasing transfers among LCC operations	**H:** Design a good connection between the existing terminal and the new LCC terminal, first with buses, but leave room for replacing it with a people mover
Population affected by noise and the number of noise complaints should not increase	
Maintain current trend of decrease of environmental impact of aircraft	**SH:** *N*egotiate with air traffic control on investments in new air traffic control equipment that can enable noise abatement procedures, such as the continuous descent approach **SH:** Invest in R&D, such as noise abatement procedures
Increase in the population density in area affected by noise	**H:** Test existing noise abatement procedures, such as the continuous descent approach, outside the peak periods (e.g. at the edges of the night) **SH:** Negotiate with surrounding communities to change their land use planning **SH:** Invest in R&D, such as noise abatement procedures
Change in the valuation of externalities by the public	**SH:** Invest in marketing of the airport to brand it as an environmentally friendly organization **SH:** Join efforts to establish an emission trading scheme
Schiphol's competitive position in terms of available capacity in Europe does not decrease	
Other major airports in Europe increase capacity	No immediate action required
Development of wind conditions due to climate change	**H:** Have plans ready to quickly build the sixth runway, but do not build it yet. If wind conditions deteriorate even further, start construction

4.5.4 Step IV: contingency planning

Step IV sets up the monitoring system and identifies the actions to be taken when trigger levels of the signposts are reached. The vulnerabilities and opportunities are those presented in Table 3. Table 4 shows the signpost to be set up for each vulnerability and each opportunity, and the possible responsive actions in case of a trigger event. The numbers used as triggers are for illustrative purposes only. For example, if demand increases twice as fast as expected, this presents an opportunity and triggers capitalizing actions. If demand grows 25% slower then anticipated, this presents a threat to the policy. In reaction, investments in capacity are delayed or even cancelled. If demand fully breaks down or explodes, the policy should be reassessed.

Table 4-4: Contingency planning

Vulnerabilities and Opportunities	Monitoring and Trigger System	Actions (Reassessment (RE), Corrective (CR), Defensive (DA), Capitalizing (CP))
Retain market share		
Demand for air traffic grows faster than forecast.	Monitor the growth of Schiphol in terms of passenger movements, aircraft movements (and related noise and emissions), if double demand (trigger) take CP-action. If demand explodes, take RE-action.	**CP:** Begin to implement the plan for the new terminal and the new runway **RE**: Reassess entire policy
Demand for air traffic grows slower than forecast.	Monitor types of demand. If overall demand is decreasing by half of forecast, take D-actions. If demand fully breaks down, take RE-action. If transfer rate decreases below 30% take CR-action.	**DA:** Delay investments, and reduce landing fees **RE:** Reassess entire policy **CR:** Cancel terminal capacity expansions
Collapse or departure of the hub carrier (KLM) from Schiphol.	Monitor the network of KLM-Air France, if 25% of flights are moved take DA-action, if 50% take CR-action, if 80% or more take R-action.	**DA:** Diversify the carriers that fly from Schiphol **CR:** Switch airport to an O/D airport by changing terminal **RE:** Reassess entire policy
Rise of long haul low cost carriers	Monitor development of the business model of low cost carriers. If long-haul LCC carriers make profit for 2 years take CP-action.	**CP:** Attract long haul LCC by offering good transfer between LCC terminal and existing terminal and/or by offering wide body aircraft stands at the LCC terminal
Rise of self-hubbing, resulting in increasing transfers between LCC operations	Monitor transfer rate among LCC flights and between LCC and legacy carriers. If transfer rate becomes more then 20%, take CP-action.	**CP:** Expand transfer capabilities between the new LCC terminal and the existing terminal
Population affected by noise and the number of noise complaints should not increase		
Maintain current trend of decrease of environmental impact of aircraft	Monitor noise footprint and emissions of the fleet mix serving Schiphol and of the new aircraft entering service. If there is an increase of noise or emissions of 10%, take CR-action.	**CR**: Change landing fees for environmentally unfriendly planes
Increase in the population density in area affected by noise	Monitor population affected by noise. If population affected by noise increases by 2%, take DA-action; by 5%, take CR-action; by 7.5%, take R-action. If population density decreases by 2%, take CP-action.	**DA:** Expand insulation program and explain basic policy again **CR:** Slow down of growth by limiting available slots **RE:** Reassess entire policy **CP:** If the population density decreases, make new slots available.
Change in the valuation of externalities by the large public	Monitor the complaints about Schiphol. If complaints increase by an average of 5% over 2 years, take DA-action. If complaints increase by an average of 10% or more over 2 years, take CR-action.	**DA:** Increase investments in marketing and branding **CR:** Slow down the growth of Schiphol by limiting the available slots
Schiphol's competitive position in terms of available capacity in Europe does not decrease		
Other major airports in Europe increase capacity	Monitor declared capacity for the major airports in Europe. If declared capacity is up by 25%, take D-action.	**DA:** Speed up expansions
Development of wind conditions due to climate change	Monitor the prevailing wind conditions throughout the year. If for 2 years in a row the number of days with cross-wind conditions exceeds 50, take D-action.	**DA:** Begin to implement the plan for the new runway

4.5.5 Step V: implementation

In the implementation phase, the plan is implemented. This plan consists of the basic policy specified in Step II, the actions specified in Table 2 and Table 3, and the monitoring system specified in Table 4. Note that the new runway being planned in the basic policy is not built yet, but can be built when necessary in light of demand increases or capacity increases at other major European airports. As such, it is a real option. The same is true of the new terminal. All the preparatory work should be started, including the clearing of the land, relocation of the current facilities on the location to other places, and putting in place the required utilities (e.g. electricity, sewers, water, space for a connection to the existing terminal, connections to the highway system and the rail system). Construction should begin if triggered by demand developments or capacity developments at other airports.

During the implementation phase, Schiphol monitors the development. Schiphol might experience faster growth than anticipated in the plan. The signposts might indicate that Schiphol is maintaining its position as a major airport for the Skyteam alliance and its partners; however, the boundaries set for safety, the environment, and quality of life, and spatial integration with its surroundings might be violated. Construction of the new terminal can start. In addition, actions need to be taken to defend the policy with respect to the negative external effects. The noise insulation program can be expanded and more investment can be made in branding and marketing that aim at explaining the policy. If these actions prove to be insufficient, the noise insulation program can be expanded, Schiphol should start to buy out residents that are heavily affected by noise, and increase landing fees for environmentally unfriendly planes. If this still is insufficient, Schiphol should consider limiting the number of available slots, especially during the night and edges of the night. If these actions are still insufficient, either because demand grows very fast or because the environmental impact grows too fast, the policy should be reassessed. If this option is chosen, the decisionmakers would reiterate through the adaptive policymaking steps in order to develop a new (adaptive) policy.

4.6 Discussion

The design of AASP as outlined in this paper was motivated by the observation that it was possible to design an improved approach for ASP by combining ideas from APM, DSP, and FSP. The combined design draws on the relative strengths of the different approaches. Compared to APM, the main improvement is that AASP explicitly considers opportunities and also includes pro-active actions (i.e. shaping actions). Compared to FSP, AASP provides a systemic framework in which the many ideas (e.g. pro-activeness, opportunities, robustness, contingency planning) are integrated in a coherent stepwise approach. Compared to DSP, AASP contains many more ways to handle uncertainty in addition to real options. We, therefore, conclude that AASP is indeed an improvement over the three separate approaches. Subsequent research should therefore try to further develop and improve upon AASP. Below we outline some questions and issues that we think are of key relevance.

In Section 4.2.2, we introduced criteria that alternatives to AMP should at least meet in order to be considered a possible viable alternative. How does the approach outlined in this paper hold up to these criteria? First, the planning approach should consider many different types of uncertainties, in addition to demand uncertainties. As is shown by the case application, other uncertainties, such as uncertainty about future climate change and its impact on wind regimes, can be considered. There is no reason why the approach should be restricted to demand uncertainties only. Therefore, we conclude that AASP meets this first criterion. The second criterion is that the planning approach should consider many different plausible futures. As can be seen in the case, the approach allows for the consideration of different futures through the identification of a wide range of vulnerabilities. The third criterion is that the resulting plan should be robust across the different futures. At this time, we have insufficient tools to formally assess the robustness of the plan outlined above. However, given that the plan contains hedging and mitigating actions, it is plausible to assume that the plan is reasonably robust. The final criterion is that the resulting plan should be flexible. Flexibility of the plan is guaranteed via the monitoring system and its associated actions. For example, the plan allows for using Lelystad airport in the future, but it does not determine it. There is, thus, flexibility. In light of this, we conclude that the outlined approach does meet the four criteria outlined in Section 4.2.2. However, a more thorough assessment of the efficacy of the approach is needed.

New infrastructure planning approaches for handling the full range of uncertainties have seen limited application (Hansman, et al., 2006). One reason for this is that the validity and efficacy of these new planning approaches has not been explored in depth (Hansman, et al., 2006). There is currently no best practice for evaluating the efficacy of new planning approaches (Dewar, Builder, Hix, & Levin, 1993; Hansman, et al., 2006). In establishing the efficacy of new infrastructure planning approaches one faces a methodological problem for ”nothing done in the short term can ‘prove’ the efficacy of a planning methodology; nor can the monitoring, over time, of a single instance of a plan generated by that methodology, unless there is a competing parallel plan.” (Dewar, et al., 1993). However, Frey and Dym (2006), suggest that by drawing an analogy with the evaluation and testing of new medicine, a methodology can be developed. For testing new infrastructure planning approaches, this analogy implies that evidence can be gathered through a variety of methods, including simulation gaming, computational experiments using exploratory modeling and analysis (Bankes, 1993), and face validation with experts (Kwakkel, Cunningham, & Pas, 2009). Currently, we are working along these lines to assess the efficacy of the approach outlined in this paper in more detail.

A possible objection to Adaptive Planning in general is that it is too costly. For example, an airport cannot afford to buy up pieces of land without knowing whether they will use the land. However, this objection overlooks an important issue. Namely, the fact that the costs associated with the possibility to adapt is a form of insurance. That is, one pays a price in order to prevent larger costs in the future. So, in the case of the airport, the costs associated with not being able to build an additional runway will be extremely high if the airport runs into capacity limitations. The price that one is willing to pay for such insurance will have to be determined on a case by case basis. However, in most cases the cost can be rather low. As for example, instead of buying the land, a spatial land use reservation can be sufficient in case of the airport. Such a land use reservation will prevent the ground from being developed in a way that will prevent future use for a runway, without the airport currently having to buy the land.

There are two other open issues with AASP. First, long-term airport development is embedded in a lengthy policymaking process with many actors. Reaching decisions in such a network of actors can be very difficult. There are, therefore, strong incentives for creating a package deal that addresses many diverse issues related to the long-term development of an airport. The actors might very well prefer to reach a decision with apparently clear consequences instead of an adaptive decision, the effect of which will depend on how the external world develops. For example, airport planners might prefer to come to an agreement to start building a new runway now, rather than agreeing that a new runway can be built five years from now if certain levels of demand are reached. Second, since there are many actors involved, even if one succeeds in developing and agreeing on an adaptive policy, there is no guarantee that in the future the actors will live up to the current agreement. For example, future changes in the government can render current decisions superfluous. This is exemplified by the decision of the newly-elected government of the Netherlands to cancel the privatization of Schiphol in 2007. It might be possible to further develop the adaptive approach to incorporate these two open issues into the planning process.

4.7 Conclusions

The current dominant approach to the long-term development of an airport is AMP. In AMP, only demand uncertainties are taken into account. Demand uncertainties are treated via aviation demand forecasting. The final product of AMP, the Master Plan, is a static blueprint that will determine the future development of the airport. In general, this approach has not been very successful, mainly due to the many uncertainties airports face. Nowadays, airports operate in an increasingly uncertain context, rendering AMP even less appropriate.

Current research into strategic planning suggests that, in light of many uncertainties, planners should strive for flexibility or adaptability. Instead of trying to predict what will happen, which is impossible in light of these uncertainties, this research has focused on planning approaches that allow implementation to begin prior to the resolution of all the major uncertainties. Over time, the policy can be adapted as new information becomes available. Adaptation is made an explicit element of the policy development. With respect to the long-term development of airports, three different adaptive approaches have been proposed in the literature. Each of these three approaches emphasizes a different aspect of Adaptive Planning. DSP emphasizes real options, APM provides a detailed framework for the development of Adaptive Plans, and FSP covers a broad spectrum of planning concepts that together result in a thorough treatment of the many and diverse uncertainties airports face. The fact that each of these approaches emphasizes a different aspect of Adaptive Planning also suggests that they can be integrated into a single adaptive airport strategic planning approach.

We chose to use APM as our starting point and extended it to incorporate pro-active actions that aim at seizing opportunities and attempting to shape the external forces. The resulting approach is labeled Adaptive Airport Strategic Planning. It is a stepwise approach. First, similar to AMP, objectives, constraints, and available options are identified. Next, a basic plan is drafted and the conditions for its success are enumerated. Third, the basic plan is examined to reveal its vulnerabilities and opportunities. Where possible, actions that can be taken now to protect the plan against vulnerabilities are added to the plan. Similarly, actions aimed at seizing the available opportunities, thus enhancing the performance of the basic plan, are also added to the plan. For the remaining vulnerabilities, a monitoring system, triggers, and responsive actions are designed. The resulting plan consists of a set of actions that will be taken directly, and an adaptive part that consists of the planned adaptations and a monitoring system that will trigger the planned adaptations. The resulting approach was illustrated using the current debate about the long-term development of Schiphol as a starting point.

References

Ascher, W. (1978). *Forecasting: An Appraisal for Policy Makers and Planners*. Baltimore: Johns Hopkins University Press.

Bankes, S. (1993). Exploratory Modeling for Policy Analysis. *Operations Research,* 4(3), 435-449.

Burghouwt, G. (2007). *Airline Network Development in Europe and its Implications for Airport Planning*. Burlington, Vermont: Ashgate Publishing Company.

Burghouwt, G., & Huys, M. (2003). Deregulation and the Consequences for Airport Planning in Europe. *DISP,* 154(3), 37-45.

Busenberg, G. J. (2001). Learning in Organizations and Public Policy. *Journal of Public Policy,* 21(2), 173-189.

Canadian Press. (2006). Mirabel Airport to be Turned into Amusement Park. Retrieved 24-07-2008, from http://www.ctv.ca/servlet/ArticleNews/story/CTVNews/20060221/mirabel_park_060221?s_name=&no_ads=

Cidell, J. L. (2004). *Scales of Airport Expansion: Globalization, Regionalization and Local Land Use*: Center for Transportation Studies, University of Minnesota.

CPB, KiM, NMP, & RP. (2007). *Advies over de lange termijnverkenning Schiphol [Advice concerning the long term exploration Schiphol]*. Retrieved from http://www.cpb.nl/nl/pub/cpbreeksen/notitie/28sep2007_2/notitie.pdf.

de Neufville, R. (2000). Dynamic Strategic Planning for Technology Policy. *International Journal of Technology Management,* 19(3-5), 225-245.

de Neufville, R. (2003). Real Options: Dealing With Uncertainty in Systems Planning and Design. *Integrated Assessment,* 4(1), 26-34.

de Neufville, R., & Odoni, A. (2003). *Airport Systems: Planning, Design, and Management*. New York: McGraw-Hill.

Dempsey, P. S., Goetz, A. R., & Szyliowicz, J. S. (1997). *Denver International Airport: Lessons Learned*. New York: McGraw-Hill.

Dewar, J. A. (2002). *Assumption-Based Planning: A Tool for Reducing Avoidable Surprises*. Cambridge: Cambridge University Press.

Dewar, J. A., Builder, C. H., Hix, W. M., & Levin, M. H. (1993). *Assumption-Based Planning: A Planning Tool for Very Uncertain Times* (Report MR-114-A). Santa Monica: RAND.

Dewey, J. (1927). *The Public and its Problems*. New York: Holt and Company.

FAA. (2001). *Forecasting Aviation Activity by Airport*. Washington DC: US Department of Transportation.

FAA. (2005). *Advisory Circular 150/5070-6B: Airport Master Plans*. Washington DC: US Department of Transportation.

Flyvbjerg, B., Bruzelius, N., & Rothengatter, W. (2003). *Megaprojects and Risk: An Anatomy of Ambition*. Cambridge: Cambridge University Press.

Forsyth, P. (1998). The Gains from the Liberalisation of Air Transport. *Journal of Transport Economics and Policy,* 32(1), 73-92.

Frey, D. D., & Dym, C. L. (2006). Validation of design methods: lessons from medicine. *Research in Engineering Design,* 17(1), 45-57.

Hansman, R. J., Magee, C., De Neufville, R., Robins, R., & Roos, D. (2006). Research agenda for an integrated approach to infrastructure planning, design, and management. *International Journal of Critical Infrastructures,* 2(2-3), 146-159.

Holling, C. S. (1978). *Adaptive Environmental Assessment and Management*. New York: John Wiley & Sons.

IATA. (2004). *Airport Development Reference Manual*. Montreal: IATA.

ICAO. (1987). *Airport Planning Manual, Part 1, Master Planning*. Montreal, Canada: ICAO.

Kwakkel, J. H., Cunningham, S. W., & Pas, J. W. G. M. v. d. (2009). *Evaluation of Infrastructure Planning Approaches: An Analogy with Medicine*. Paper presented at the 2nd Next Generation Infrastructure Conference: 'Developing 21st Century Infrastructure Networks'.

Kwakkel, J. H., Walker, W. E., & Marchau, V. A. W. J. (2007, June 21-23). *Dealing with Uncertainty in Airport Strategic Planning.* Paper presented at the 11th Air Transport Research Society 2007 World Conference, Berkeley.

Kwakkel, J. H., Walker, W. E., & Wijnen, R. A. A. (2008, 6-10 July). *The Treatment of Uncertainty in Airport Strategic Planning: The Case of Schiphol Airport's Long-Term Vision.* Paper presented at the 12th Air Transport Research Society World Conference, Athens, Greece.

Lee, K. (1993). *Compass and Gyroscope: Integrating Science and Politics for the Environment*. Washington: Island Press.

Lempert, R. J. (2002). A New Decision Sciences for Complex Systems. *Proceedings of the National Academy of Sciences of the United States of America,* 99(3), 7309-7313.

Lempert, R. J., Popper, S., & Bankes, S. (2002). Confronting Surprise. *Social Science Computer Review,* 20(4), 420-439.

Lempert, R. J., Popper, S., & Bankes, S. (2003). *Shaping the Next One Hundred Years: New Methods for Quantitative, Long Term Policy Analysis*. Santa Monica: Rand.

Mayer, I. S., van Daalen, C., & Bots, P. W. G. (2004). Perspectives on policy analyses: a framework for understanding and design. *International Journal of Technology, Policy and Management,* 4(2), 169-191.

McLain, R. J., & Lee, R. G. (1996). Adaptive Management: Promises and Pitfalls. *Environmental Management,* 20(4), 437-448.

Nelkin, D. (1974). *Jetport: the Boston Airport Controversy*. New Jersey: Transaction Books.

Nelkin, D. (1975). The Political Impact of Technical Expertise. *Social Studies of Science,* 5(1), 35-54.

Porter, A. L., Roper, A. T., Mason, T. W., Rossini, F. A., & Banks, J. (1991). *Forecasting and Management of Technology*. New York: John Wiley and Sons.

Provincie Noord-Holland. (2007). *Vestigingslocaties Schiphol, een globale verkenning voor de lange termijn*. Retrieved from www.noord-holland.nl/Images/65_105492.pdf

Rijksoverheid. (2009). *Luchtvaartnota, Concurrerende en duurzame luchtvaart voor een sterke economie [Air Transport White Paper Competitive and Sustainable Air Transport for a Strong Economy]*.

Schiphol Group. (2007). *Lange Termijn Visie op de Ontwikkeling van de Mainport Schiphol: Een Wereldwijd Netwerk Voor een Concurrerende Randstad [Long-Term Vision on the Development of the Mainport Schiphol]*.

Schiphol Group, & LVNL. (2007). *Verder Werken aan de toekomst van Schiphol en de regio [continu working on the future of Schiphol and the region]*.

Szyliowics, J. S., & Goetz, A. R. (1995). Getting Realistic About Megaproject Planning: The Case of the New Denver International Airport. *Policy Sciences,* 28(4).

V&W. (2007). *Nieuwsbrief Schipholbeleid nr 12 [newsletter Schiphol policy, nr 12]*: Ministry of Transport.

V&W, & VROM. (2007). *Lange Termijn Verkenning Schiphol, probleemanalyse [Long term exploration Schiphol: problem exploration]*: Ministry of Transport, Ministry of Housing, Spatial Planning and the Environment.

Walker, W. E. (2000). Policy Analysis: A Systematic Approach to Supporting Policymaking in the Public Sector. *Journal of Multicriteria Decision Analysis,* 9(1-3), 11-27.

Walker, W. E., Rahman, S. A., & Cave, J. (2001). Adaptive Policies, Policy Analysis, and Policymaking. *European Journal of Operational Research,* 128(2), 282-289.

5 Evaluation of infrastructure planning approaches: an analogy with medicine

KWAKKEL, J.H., PAS, J.W.G.M.V.D. (under review) Evaluation of Infrastructure Planning Approaches: An Analogy with Medicine.

Abstract

This paper discusses the evaluation of new infrastructure planning approaches. These new planning approaches have been put forward in response to the challenges of deep uncertainty about the future. These approaches emphasize the need for flexibility of the system in order to enable the plan to adapt to changing conditions However, these adaptive approaches up till now have seen little real word applications. One important reason for this lack of application is that the efficacy of these approaches has not been established yet. In turn, this is largely due to the problem that there is no agreed upon method for proving the efficacy of a new planning approach. In this paper, we will draw an analogy to medical research and development in order to outline a methodology for establishing the efficacy of new planning approaches. We discuss how the well-established methodology for evaluation new medical treatments can be adapted to evaluating new planning approaches. We illustrate the resulting evaluation methodology by outlining an evaluation strategy for a specific new planning approach. It is concluded that the well-established methodology from medicine can successfully be used to inform the evaluation of infrastructure planning approaches

5.1 Introduction

Policymakers, business leaders, and governmental planners in general have to face a wide variety of uncertainties that will affect to outcomes of their plans or policies. Moreover, many of these uncertainties are beyond their direct control. Such strategic decisionmaking problems are now frequently labeled as decisionmaking under deep uncertainty (Lempert, Popper, & Bankes, 2003; Volkery & Ribeiro, 2009; Walker, Marchau, & Swanson, 2010).

Decisionmaking under deep uncertainty can be defined as the situation in which the decisionmaker does not know, or multiple decisionmakers cannot agree on, the system model, the prior probabilities for the uncertain parameters of the system model and/or the value function(Lempert, et al., 2003). Over the last couple of years, a growing body of literature has emerged that focuses on deep uncertainty, the problems it creates for decisionmaking, and how to handle or treat it (e.g. Ben-Haim, 2006; de Neufville, 2003; Dewar, 2002; Lempert, et al., 2003; Walker, et al., 2010).

One particular field were deep uncertainty is of prime importance is infrastructure planning. Infrastructures have a long life time, alter the world around them in unpredictable ways, and bring together a variety of stakeholders and decisionmakers across geographical scales and levels of the governmental hierarchy. In the past, the dominant approach in infrastructure planning was to largely ignore the uncertainties (Dempsey, Goetz, & Szyliowicz, 1997; Marchau, Walker, & van Duin, 2009; Quade, 1982; Van Geenhuizen, Reggiani, & Rietveld, 2007; van Geenhuizen & Thissen, 2007). In the face of deep uncertainties about the future, the resulting infrastructure plans, based on a limited exploration of the uncertainties are likely to fail. In response to this, recently various new planning approaches have been put forward. These approaches emphasize on the one hand the need for a more thorough integrated forward-looking analysis (Swanson, et al., 2010) of the uncertainties through techniques such as exploratory modeling and analysis (Agusdinata, 2008; Lempert, et al., 2003), bounce casting (Kahan, et al., 2004), scenarios in various forms (Bradfield, Wright, Burt, Cairns, & van der Heijden, 2005; Varum & Melo, 2010b). On the other hand, because of the limited potential of these techniques in anticipating rare events (Goodwin & Wright, 2010), there is a growing interest in flexibility and adaptiveness in plans where a strategic vision of the future is combined with short-term actions and a framework that can guide future actions (Albrechts, 2004; Walker, et al., 2010; Walker, Rahman, & Cave, 2001). With respect to infrastructures, the need for flexibility or adaptiveness is often translated into real options (de Neufville, 2003).

Although the toolbox of the infrastructure planner has gradually expanded, a lot of progress needs to be made in incorporating methods such as scenarios, exploratory modeling, Delphi studies, and techniques aimed at adaptivity and flexibility such as real options analysis into processes of policy design, choice and implementation (Hansman, Magee, De Neufville, Robins, & Roos, 2006; Varum & Melo, 2010b; Volkery & Ribeiro, 2009; Walker, et al., 2010). It has been suggested that this lack of application is at least partly due to the limited assessment of the validity and applicability of these new planning approaches (Hansman, et al., 2006; Walker, et al., 2010). Given the profound importance of infrastructure to society, the careful assessment of the efficacy of new planning approaches is needed.

There is a gap in the literature on how to evaluate planning approaches. Hansman et al. (2006) highlight the importance of testing new planning approaches, but immediately suggest controlled real world application. In light of the importance of infrastructure to society and the risks of failure, the soundness of this suggestion is put into question. Furthermore, as Dewar et al. (1993) highlight, there is a methodological problem with controlled real world applications: in general, it is not possible to implement both a traditional plan and a plan based on a new planning approach in order to compare the performance of both plans. In light of this, we conclude that there is currently no established approach for evaluating new infrastructure planning approaches. This paper aims at sketching such an approach by drawing an analogy with the evaluation of new medical treatments. It is argued that before performing real world evaluations of new planning approaches, a wide array of other techniques can be employed.

The goal of this paper is to sketch an approach for testing the efficacy of new infrastructure planning approaches. In Section 2, we elaborate on the treatment of uncertainty in infrastructure planning. In Section 3, we discuss the medical analogy. Section 4 presents the implications of this analogy for the evaluation of a specific new infrastructure planning approach. Section 5 contains our concluding remarks.

5.2 The treatment of uncertainty in infrastructure planning

Deep uncertainties about the future pose a significant challenge to infrastructure planning. The traditional approach in infrastructure planning has been to ignore the uncertainties or to try and reduce them (Dempsey, et al., 1997; Marchau, et al., 2009; McDaniel & Driebe, 2005; Quade, 1982; Van Geenhuizen, et al., 2007; van Geenhuizen & Thissen, 2007). Planners extrapolated past trends forward and developed static blueprint plans for achieving the desired goals. However, for a multitude of reasons, such forecasts are practically always wrong, since the future that materializes differs significantly from the forecasted future (Ascher, 1978; Batty & Torrens, 2005; Flyvbjerg, Bruzelius, & Rothengatter, 2003; Kristóf, 2006; Porter, Roper, Mason, Rossini, & Banks, 1991). Furthermore, such approaches suffer from the problem that they focus on those uncertainties that are “among the least of our worries; their effects are swamped by uncertainties about the state of the world and human factors for which we know absolutely nothing about probability distributions and little more about the possible outcomes” (Quade, 1982). Similarly, Goodwin and Wright (2010 p. 355) demonstrate that “all the extant forecasting methods – including the use of expert judgment, statistical forecasting, Delphi and prediction markets – contain fundamental weaknesses.” And Popper, et al. (2009) state that the traditional methods “all founder on the same shoals: an inability to grapple with the long-term’s multiplicity of plausible futures.” Any infrastructure plan designed on the basis of one or a few forecasts or a small set of assumptions about the future performs poorly as a result, and unplanned ad-hoc adaptations are needed to rectify this.

Table 5-1: Categories of policy approaches under uncertainty (adapted from Agusdinata, 2008)

	Static Rigid Policy	Static robust policy	Dynamic Flexible Policy
Policy approaches	Predict the future and implement optimal policy for that future	Identify plausible futures and find policies that produce acceptable results across these futures. Hedge against contingencies	Adapt policy over time as conditions change and learning takes place

Table 5-1: gives an overview of different types of approaches to policymaking under uncertainty. Recently, an alternative paradigm to the static policy approaches has emerged. This paradigm holds that in light of the deep uncertainties, one needs to plan dynamically and build in flexibility (de Neufville, 2000; IISD, 2006; Lempert, 2002; Lempert & Groves, 2010; Swanson, et al., 2010; Walker, et al., 2001). The initial ideas for this paradigm are found almost a century ago. Dewey (1927) put forth an argument proposing that policies be treated as experiments, with the aim of promoting continual learning and adaptation in response to experience over time (Busenberg, 2001). Early applications of adaptive policies, can be found in the field of environmental management (Holling, 1978). Motivated by the complexity of the environmental system, environmental managers resort to controlled experiments aimed at increasing their understanding of the system (McLain & Lee, 1996). Policies are designed

from the outset to test clearly formulated hypotheses about the behavior of an ecosystem being changed by human use (Lee, 1993). A similar attitude is also advocated by Collingridge (1980) with respect to the development of new technologies. Given ignorance about the possible side effects of technologies under development, he argues that one should strive for correctability of decisions, extensive monitoring of effects, and flexibility. More recently, Walker et al. (2001) developed a structured, stepwise approach for dynamic adaptation. This approach advocates that plans should be adaptive: one should take only those actions that are non-regret and time-urgent and postpone other actions to a later stage. In order to realize this, it is suggested that a monitoring system and a pre-specification of responses when specific trigger values are reached should complement the basic plan. The resulting plan is thus flexible and can be adapted to how the future unfolds.

Dynamic flexible policymaking in one form or another has received attention in various disciplines. In infrastructure planning, the need for adaptiveness and flexibility is increasingly recognized. For example, in air transport, the developments of the last decade including various terrorist attacks, SARS, Mexican flu, and the second Gulf war, have highlighted this need. Combine this with the impacts of privatization and liberalization, the rise of airline alliances, merges, and take-offers, and it is obvious that it is next to impossible to plan for the long-term development of an airport based on a prediction of the size and composition of future demand. In response to these uncertainties, the need for Dynamic Adaptive Planning is forcefully argued (Burghouwt, 2007; de Neufville, 2000; de Neufville & Odoni, 2003; Kwakkel, Walker, & Marchau, 2010; Walker, et al., 2001). Another argument for dynamic adaptation in the transport domain comes from the side of research on transport innovations. The implementation of innovations, such as advanced driver assistance systems and innovative approaches for intra-city logistics, is hampered by various uncertainties, both about the technology to be implemented and uncertainties surrounding the system. Dynamic flexible implementation plans have been put forward as a way to overcome these problems (Marchau & Walker, 2003; Marchau, et al., 2009). In other domains, the need for adaptivity and flexibility is argued on very similar grounds. For example, in integrated river basin management, the omnipresence of uncertainties in both the environmental system and the societal system is used as an argument for adaptivity and flexibility (Pahl-Wostl, Moltgen, Sendzimir, & Kabat, 2005; Pahl-Wostl, et al., 2007). Policies related to climate change are yet another area where dynamic adaptation and flexibility are put forward as the approach that should be used for policy design (Dessai, Hulme, & Lempert, 2009; Wardekker, de Jong, Knoop, & Van der Sluijs, 2010).

Dynamic adaptive policymaking approaches impose various requirements on the system on which they act and on the design of the policy or plan. In order to be able to adapt a policy to changing circumstances, the policy should be flexible. In case of infrastructure planning, this implies that adaptive policies require flexible infrastructures. To illustrate this with a simple example: in order for an airport to be able to respond to changes in demand, the investment in new facilities should be postponed as long as possible. Moreover, the time between investment and the availability of the new facilities should be minimized. For terminal design, one way of achieving this is by having a modular terminal that can easily and quickly be expanded. Land banking or land reservations and zoning are techniques that can be employed to retain the possibility to add new runways. The increasing research interest in Real Options, as such forms of flexibility are sometimes named, reveals the interest in the engineering world for approaches for handling uncertainty in long-term engineering projects.

The kind of flexibility that is needed for a given system will depend critically on the uncertainties that decisionmakers are facing. So, in case of airport planning, if the uncertainty is about the size of future demand, the required flexibility will be different from a situation in which the uncertainty is about the composition of demand. Uncertainty in the size of demand requires the ability to easily and quickly expand or contract when necessary. The composition of demand requires the flexible use of facilities. If the majority of travel is international, a different terminal layout is needed compared to the situation in which the majority of travel is national (or Schengen, in the case of large parts of Europe). Another relevant dimension of flexibility relates to time. The flexibility required to cope with short-term shocks is quite different from the flexibility to adapt to changing conditions over a longer lifetime. The flexibility of airports to accommodate passengers for a long time inside the terminal because of a volcanic eruption is quite different from the flexibility that is required to cope with the aging of people that are flying. Regardless of the various kinds of flexibility or dimensions of flexibility, there is an intimate relation between the technical design of an infrastructure and the extent to which a policy that guides the development of the infrastructure can dynamically adapt to changing circumstances.

From the foregoing, we conclude that uncertainty is a central problem in long-term infrastructure planning. A large body of literature exists that argues that in order to handle these uncertainties, infrastructure planning needs to shift from the static rigid policymaking paradigm to the dynamic flexible policymaking paradigm. Despite the discussion in the literature, these dynamic flexible approaches for policymaking have seen limited real world applications (Hansman, et al., 2006; IISD, 2006; Lempert & Groves, 2010; McCray, Oye, & Petersen, 2010). It has been suggested that one of reasons for this limited application is the fact that evidence for the efficacy of these approaches is lacking (Dewar, et al., 1993; Hansman, et al., 2006). Their efficacy is currently mainly argued on theoretical grounds. Other sources of evidence, particularly empirical evidence, are needed. This demand for evidence for the efficacy of new planning approaches is in turn motivated by the importance of infrastructure to society, using new unproven planning approaches poses a significant risk. In establishing the efficacy of new infrastructure planning approaches, however, one faces a methodological problem for ”nothing done in the short term can ‘prove’ the efficacy of a planning methodology; nor can the monitoring, over time, of a single instance of a plan generated by that methodology, unless there is a competing parallel plan” (Dewar, et al., 1993). In light of the above, we argue that a methodology for testing the efficacy of new infrastructure approaches is needed.

5.3 Medical treatment– uncertainty treatment in planning analogy

In the foregoing section, we concluded that there are various approaches available for addressing the problems of deep uncertainty in infrastructure planning, but that evidence for their efficacy is lacking. In this section, we address this problem. We turn to the experience and methodology that has emerged in medical research and development, for it has addressed the problem of assessing the efficacy of new treatments with a considerable degree of success. Other fields that could also be worthwhile to explore for providing insights include engineering design evaluation and design evaluation. Here also, researchers have grappled with finding methods for evaluating the efficacy of the design process and the resulting designs. They looked, among others, to medicine and the validation of decision support systems for insight (Frey & Dym, 2006).

Before turning to the analogy, we first demarcate the object of evaluation more sharply. In the evaluation literature, one important distinction is that between product and process (Walls, Widmeyer, & El Sawy, 1992, 2004) or between plan, process, and product (Verschuren & Hartog, 2005). If the object of evaluation is the plan, the evaluation focuses on the assessment of the quality of the design on paper (Verschuren & Hartog, 2005). Process evaluation focuses on the procedures for the construction of the design (Verschuren & Hartog, 2005; Walls, et al., 2004). Product evaluation involves the assessment of the value of the created artifact and its short- and long-term impacts after its creation (Verschuren & Hartog, 2005). With respect to approaches for policymaking, the product is the final policy that is implemented. An extensive literature exists on policy evaluation, and this will therefore not concern us here. The plan is the drafted policy, while the process is the process of drafting the policy. In the context of this paper, we are interested in the evaluation of both the plan and the process.

A second important distinction is that between kernel theories and design theories (Gregor & Jones, 2007; Kuechler & Vaishnavi, 2008; Walls, et al., 1992, 2004). Design theory focuses on how to do something and so shows how theories from a specific science can be put to practical use. These latter types of theories are called kernel theories (Gregor & Jones, 2007; Walls, et al., 1992). The concept of design theory goes back to the work of Simon (Simon, 1996). Adopting the distinction between design theory and kernel theory, the various approaches that have been put forward to address deep uncertainty in infrastructure planning can be considered to be design theories, for they specify how to achieve 'well' performing plans under conditions of deep uncertainty. As a design theory, a given approach for handling deep uncertainty in general makes statements both on how to structure the *process* of designing the plan, as well as statements about the structural composition of the design (i.e. the *plan*).

5.3.1 Introduction of the analogy

Analogies express recognition of similarities between otherwise dissimilar things. Reasoning by analogy allows for adopting an alternative perspective that can produce novel insights (Porter, et al., 1991). In the case of medical treatment versus uncertainty treatment, we use the analogy to sketch an approach for testing new infrastructure planning approaches. The key similarity is that in both cases there is a new treatment of which the efficacy needs to be established prior to using the treatment in daily practice. Not carefully assessing the diverse effects of the new treatment poses a risk either to humans (in the case of medical treatments), or to society (in the case of infrastructure planning).

Researchers in medicine have extensive experience and a well-established methodology for validating new medical treatments. Although the process of medicine evaluation is subject to continuous improvement and refinement, it shares attributes with the evaluation of new infrastructure planning approaches. The emphasis in medical research and development is on the development of new treatments. These new treatments should improve the health of people. Before they can be applied in practice, their efficacy needs to be exhaustively tested. Specific attention is also given to the potentially harmful unintended effects of new treatments. The goal of evaluating new medical treatments is to assess whether the new treatment is beneficial, under what conditions it should be applied, and when it should not be applied.

The main goal of research into new infrastructure planning approaches is the development of approaches for successfully handling deep uncertainty. These new approaches should result in infrastructure plans that produce adequate results regardless of how the future turns out. Given the importance of infrastructure planning to society, new planning approaches should be rigorously tested. Attention should be given to whether the new approaches are preferable to existing approaches, under what conditions these new approaches are applicable, and when these new approaches should not be used.

Given the similarity in questions that need to be addressed in evaluating medical treatments and evaluating new infrastructure planning approaches, we explore the analogy in more depth in the next section. Of course, there are profound differences. For example, medical treatments in general require little interpretation on part of the medical doctor. When the treatment is useful, in what dose a medicine can be prescribed, or how to perform a specific surgery is fully specified. In contrast, the new infrastructure planning approaches leave significant room for case-based operationalization and interpretation. Despite these obvious differences, the similarity in questions that need to be addressed in both evaluating new medical treatments and in evaluating new infrastructure planning approaches leads us to belief that pursuing the analogy will prove useful

5.3.2 The analogy in detail

In this section we explore the analogy in detail. Table 5-2 summarizes the analogy by providing an overview of the types of evidence that are used in medical research and development and how these same types of evidence can be applied to the evaluation of infrastructure planning approaches. The evidence is hierarchy ordered, from the most rudimentary evidence to the most detailed form of evidence. This table is adapted from (Frey & Dym, 2006), who apply the medical analogy to design method evaluation. In the subsections below, each level of evidence is discussed in more detail.

Table 5-2: Types of evidence used to develop and validate medical treatments and infrastructure planning approaches (adapted from Frey & Dym, 2006)

Evidence used to develop/validate medical treatment	Evidence used to develop/validate infrastructure planning approaches
Theory	Theories (e.g. decision science, cognitive science, political science, organizational behavior, policy analysis)
Animal Models	Computational experiments of plans across an ensemble of futures Simulation gaming with students
In Vitro Experiments	Simulation gaming with actual decision makers
Natural Experiments	Case studies of successful long-term infrastructure plans
Clinical Trials	Controlled field application of planning approach

Theory

The development of new medical treatments occurs in close interaction with theoretical developments in biology, biochemistry, and the life sciences. New theoretical insights spur innovations in treatments. Vice versa, successful but incompletely explained treatments further theoretic insight. This is intimately related to distinction between kernel theories and design theories introduced earlier. The theoretical developments in biology, biochemistry, and the life sciences are the kernel theories, while the treatments developed in medicine embody design theories. Similarly, new approaches for infrastructure planning are developed based on theoretic insights from a diverse number of fields, including decision science, cognitive science, political science, organizational behavior, and policy analysis. Thus, theories from these fields are the kernel theories for the new infrastructure approaches.

Animal models

Animal models (often mice) are developed in medical research and development to display specific physiological characteristics that are of interest to the research. New treatments are subsequently first tested on these animal models. Animal models in medical research and development bear only a partial resemblance to the real, human, subject. Moreover, animal models are frequently engineered to exhibit desired characteristics. For example, mice are treated to grow cancers or genetically altered to exhibit behavior comparable to ADHD. Animal models are comparable to simulation models frequently employed in engineering. Like animal models, simulation models bear a partial resemblance to the actual world and, just like animal models, engineering models can be tailored to resemble a desired aspect of the world in the desired way.

In the context of assessing the efficacy of new infrastructure planning approaches, computer models can be used as surrogates for the real world system. With respect to the evaluation of plans, for example, the performance of traditional and new style plans as calculated by the computer models can be compared and one can reason about how these plans would behave in the real world. However, given that infrastructure planning is decisionmaking about the future, there are significant uncertainties present. There is simply not enough knowledge to accurately forecast the future, the models of large infrastructure systems are often contested, and there are a variety of value-systems from the different stakeholders that are also bound to change over time. The performance of the plans derived from the different planning approaches needs to be assessed across these uncertainties. Therefore, when using simulation models as animal models for assessing the efficacy of infrastructure planning approaches, Exploratory Modeling and Analysis (EMA) is an appropriate method (Agusdinata, 2008; Bankes, 1993).

EMA is a research methodology to analyze complex and uncertain systems (Agusdinata, 2008; Bankes, 1993). It can be contrasted with a *consolidative modeling* approach, in which all the existing knowledge about a system is consolidated in a model that is subsequently used as a surrogate for the real world system (Hodges, 1991; Hodges & Dewar, 1992). The consolidative approach is valid only when there is sufficient knowledge at the appropriate level and of adequate quality available. When dealing with long-term infrastructure planning, these conditions are not met, so using such a consolidative approach might produce erroneous results (Batty & Torrens, 2005; Dewar & Wachs, 2006; Marchau, et al., 2009; Van Geenhuizen, et al., 2007). However, in such situations there still is a wealth of knowledge and information available that supports a set of structurally different models across a range of parameter values. EMA aims at offering support for exploring this set of models across the range of plausible parameter values and drawing valid inferences from this exploration (Agusdinata, 2008; Bankes, 1993). In the context of EMA, a computer model can be used as a platform for computational experiments, as lab equipment that maps specific inputs into output about system behavior (Bankes, 2009; Bankes, Lempert, & Popper, 2002). Using models as lab equipment has implications for model design: models need to be modular so that a variety of hypotheses about system structure can be implemented, tested, and compared (Bankes, 2009).

If the object of evaluation is the process, animal models can be employed in infrastructure planning and research through simulation gaming with students in infrastructure planning (Mayer & Veeneman, 2002). This is in line with what Schön (1983) called a practicum. Students are not the actual decisionmakers but reflect them to a certain extent, and the laboratory setting bears a partial and incomplete resemblance to the real world. The use of exploratory modeling and simulation gaming with students can provide preliminary evidence for the efficacy of new infrastructure planning approaches.

In vitro experiments

In vitro experiments focus on applying treatments to pieces of human tissue outside the human body. In this way, there is more potential to control the various exogenous variables, thus allowing the research to focus on the identification of specific biochemical mechanisms. Moreover, in vitro experimentation avoids the risk of harming human subjects. However, the gain in controlling the exogenous conditions does limit the extent to which findings can be generalized. In vitro experiments in infrastructure planning can take the form of simulation gaming with actual decisionmakers. Simulation gaming presents a safe laboratory environment. The actual decisionmakers are removed from their normal context and placed in a highly controllable environment. Within this safe environment, new infrastructure planning processes can be tested on actual decisionmakers. In vitro experimentation allows for learning both on part of the researchers developing the new infrastructure planning approaches, and also on part of the potential future users of these approaches. In vitro experiments in infrastructure planning would thus expand on simulation gaming with students and moves the experiments ever closer to real world application.

Natural experiments

Natural experiments are experiments without controlled administration of treatment. In medicine, these are often used to study the effects of long-term exposure to specific conditions. For example, to establish the relationship between asbestos and cancer, smoking and cancer, or the long-term health effects of medication for chronic illnesses, natural experiments are ideally suited. For such cases, it is not possible or desirable to perform clinical experiments. So, instead, the consequences of uncontrolled treatment or exposure are studied in a systematic manner. In order to move from mere association to causality, Hill (1966) specifies several aspects of the association that are of great relevance: strength of the association, consistency of the association across different studies using different methods and different populations, the temporal relationship, and the coherence with other available knowledge and evidence. To support the medical profession in this, meta-analysis has become increasingly important. Meta-analysis focuses on the statistical analysis of large collections of analysis results from individual studies for the purpose of integrating the findings (DerSimonian & Laird, 1986; Glass, 1976).

In infrastructure planning, natural experimentation could be achieved through utilizing "found" experiments, where, by history or chance, experimental groups have been exposed to different conditions. The causal outcomes of these conditions can then be explicated by observation and direct or indirect measurement. Endogenous growth theory, which postulates that economic growth is dependent upon technical progress – and that the prior conditions for technical progress are deserving of explanation – has been an avid consumer of natural experimentation techniques (Poot, 2000). Infrastructure delivery has been a part, although not the exclusive part, of these studies. The development of a set of natural experiments based on the systematic study of multiple cases of long-term infrastructure planning projects is urgently needed, for it can serve as a basis for providing evidence for the efficacy of new planning approaches, the revision of theory, and the fine tuning of existing methods. Performing a series of natural experiments or performing meta-analyses of the available case studies will greatly enhance the empirical base of theory and will guide theory improvement. This is analogous to medicine where past cases form an indispensable empirical base for theory and theory improvement (Hill, 1966).

Clinical trails

Clinical trials are the final test of the effects of new medical treatments. By and large, they are regulated by national and international agencies that regulate which treatments and medication can be used (e.g European Medicines Agency, 1998; U.S. Department of Healt and Human Services - Food and Drug Administration, 1998). In clinical trials, human subjects are administered the treatment or a placebo under controlled conditions, the effects are monitored, and the results are recorded. Usually, the trial is double blind, meaning that neither the patient nor the medical doctor knows whether the treatment is real or a placebo. Moreover, the results of a clinical trial need to be independently substantiated. It is debatable whether a full clinical trial of new infrastructure planning approaches is possible. Most importantly, the effect of new planning approaches is critically dependent on external forces (e.g. state of the economy, technological developments) that cannot be controlled. One could argue that even if the protocols for new planning approaches are rigorously applied (and a control group can be created that is denied these protocols) the context of infrastructure development is too localized and too socialized (and therefore inherently variable across cases) to believe we can ever succeed in performing a controlled experiment.

However, one of the main drivers in medical research for double blind trials is the fact that the placebo effect poses a significant threat to validity. Arguably, a sort of placebo effect can also affect the perception of stakeholders and decisionmakers when using a new planning approach. Its mere newness might alter the perception of the planning approach. But it is questionable to assume that the placebo effect is as much a problem for testing new planning approaches as it is for testing new medical treatments. Therefore, the need for double blind trials in testing new infrastructure planning approaches is diminished. A form of experimentation that does not rely on double blind trails is therefore a defensible technique for assessing the efficacy of new infrastructure planning approaches.

One type of experimentation that could be of great benefit in assessing the efficacy of new infrastructure planning approaches is the use of pilot projects. Broadly speaking, pilot projects are experimental projects undertaken in a field setting (Lee, 1999; Weiss, 1975). More precisely, pilot projects are the means of applying new approaches (the innovation) in a confined field setting to learn about the innovation-context interaction and to use these lessons for improving the innovation or adjusting the context (e.g. management practices) (Vreugenhil, Slinger, Thissen, & Ker Rault, in press). The use of pilot projects in assessing the efficacy of innovative infrastructure planning approaches is attractive, for it provides a space for innovation, communication, evaluation, creativity, and learning, while resources can often relatively easy be streamlined (Hoogma, Kemp, Schot, & Truffer, 2002; Vreugenhil, et al., in press). Pilot projects often play an important role in preparing for the wide scale introduction of an innovation, for the innovation can mature, the involved actors can learn about the innovation, and this knowledge can diffuse (Raven, 2007; Rogers, 2003). Three uses of pilot projects can be distinguished: research pilots focusing on exploring and evaluating the innovation; management pilots focusing on communication, problem mitigation, policy implementation, and insurance; and, political-entrepreneurial pilots focusing on providing incentives for adopting the innovation, on advocating the innovation, or for political games (Vreugenhil, et al., in press). Each of these three uses of pilot projects can be used for assessing the efficacy of new infrastructure planning approaches, the research usage being the most important from the perspective of the evaluation of new approaches for handling deep uncertainty in infrastructure planning.

5.4 Explication of the analogy for Adaptive Policymaking

In this section, we explore the implications of the analogy for Adaptive Policymaking (APM). APM is an example of an infrastructure planning approach belonging to the dynamic flexible paradigm. The key idea of APM is to only take those actions that are time urgent, and combine these with those that make important commitments to shape the future, preserve needed flexibility for the future, and protect the policy from failure. Figure 5-1 shows the APM framework. In short, in Step I, the existing conditions of an infrastructure system are analyzed and the goals for future development are specified. In Step II, the way in which this is to be achieved is specified. This basic plan is made more robust through four types of actions specified in step III: mitigating actions are actions to reduce the *certain* adverse effects of a plan; hedging actions are actions to spread or reduce the risk of *uncertain* adverse effects of a plan; seizing actions are actions taken to seize certain available opportunities; and shaping actions are actions taken to reduce the chance that an external condition or event that could make the plan fail will occur, or to increase the chance that an external condition or event that could make the plan succeed will occur. Even with the actions taken in step III, there is still the need to monitor the performance of the plan and take action if necessary. This is called contingency planning and is specified in step IV. Signposts specify information that should be tracked in order to determine whether the plan is achieving its conditions for success. Critical values of signpost variables (triggers) are specified, beyond which actions should be implemented to ensure that the plan keeps moving the system in the right direction and at a proper speed. There are four different types of actions that can be triggered by a signpost: defensive actions are taken to clarify the basic plan, preserve its benefits, or meet outside challenges in response to specific triggers that leave the basic plan remains unchanged; corrective actions are adjustments to the basic plan; capitalizing actions are actions trigger to take advantage of opportunities that improve the performance of the basic plan; and a reassessment of the plan is initiated when the analysis and assumptions critical to the plan's success have clearly lost validity. Step V is the actual implementation. In this step, the actions to be taken immediately (from Step II and Step III) are implemented and a monitoring system (from Step IV) is established. Then time starts running, signpost information related to the triggers is collected, and actions are started, altered, stopped, or expanded in response to this. After implementation of the initial actions, the implementation of other actions is suspended until a trigger event occurs. The central claim of APM is that a more successful coherent development of infrastructure can be achieved by planning for adaptation instead of having to make ad-hoc adaptations and modifications to existing plans. For a more detailed explanation of this framework, see Kwakkel et al (2010), Marchau et al. (2009), and Walker et al. (2001).

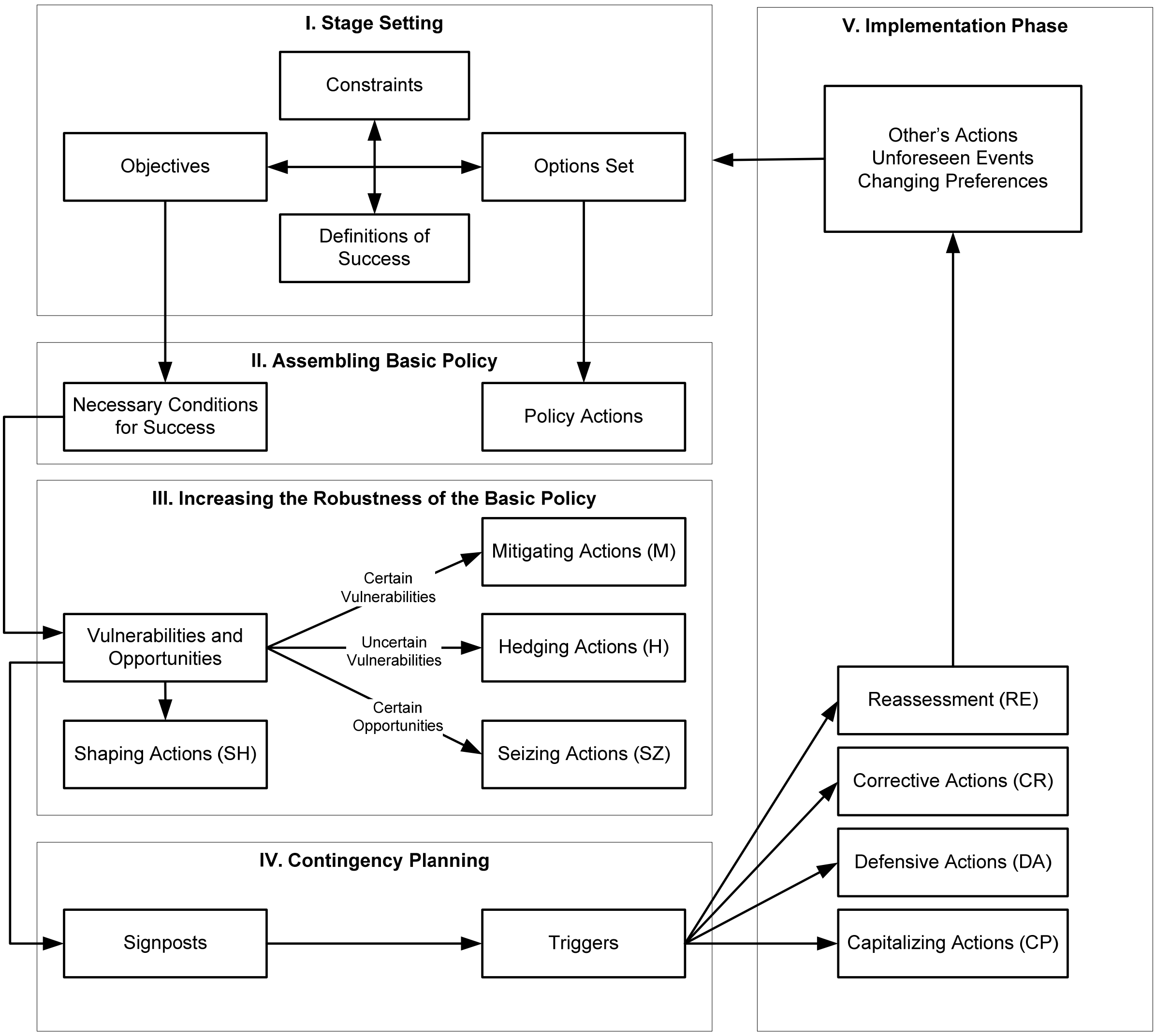


Figure 5-1: Steps in adaptive policy making (adapted from Kwakkel, et al., 2010)

In line with the distinction between process and plan introduced in the preceding section, we draw a distinction between the process of developing an Adaptive Plan – the Adaptive Policymaking Process (APP), and the performance of the Adaptive Policy (AP). Table 5-3 present an application of the medical analogy to both APP and AP. Below, we discuss each row of this table in more detail and take stock of the evidence that has been gathered on the efficacy of APM for each of the sources of evidence.

Before turning to the detailed discussion of the analogy with respect to APM, we first specify the main criteria on which APM has to be evaluated. Important categories of evaluation criteria include the activity of developing the Adaptive Plan, the resulting plan, the use of the plan, and the effects of the plan (Thissen & Twaalfhoven, 2001). APP coincides with the activity of developing plans. Here, relevant criteria include the appropriateness of the methods and techniques that are employed, the validity of the results from these techniques, the transparency of the decisionmaking process and the cooperation and involvement of stakeholders. AP coincides with the effects of the plan. Effect criteria focus on evaluating the effects of the plan, such as what is the range of expected outcomes, under what conditions would a traditional plan outperform an Adaptive Plan and vice versa.

Table 5-3: The medical analogy applied to APM

Medicine	Adaptive Policymaking Process	Performance of the Adaptive Policy
Theory	Significant theoretical work is still needed on the process of developing adaptive policies. Questions include the institutionalization of APM, the interaction of actors in an APM process, and the employment of existing analytic techniques and methods in developing adaptive polices.	There appears to be a consensus in the literature that flexible adaptive policies are better then static policies. Opinions differ on the type of flexibility or adaptability. One question is how different adaptive policies turn out, compared to static policies where ad hoc modifications occur in response to real world events.
Animal Models	Simulation gaming with students. The focus of the game would be on the development of an adaptive policy. Initial results of a workshop with both students (3 groups) and colleagues (1 group) appear promising.	Simulations of an adaptive policy compared to traditional policy across a large ensemble of plausible future worlds. Initial results appear promising. Face-validation of adaptive policies by experts. This is planned to be executed in the near future Simulation gaming with students. Let a group of players play out a static policy against a scenario; let another group play out an adaptive policy against a scenario.
In Vitro Experiments	Simulation gaming with actors involved in infrastructure planning to develop to adaptive policy. The workshop that has been executed with students and colleagues is planned to take place with real word actors.	Simulation gaming as under animal models, but with actors that are actually involved in infrastructure planning.
Natural Experiments	Case studies of successful long-term decisionmaking processes, interviews with people involved in long-term decisionmaking. How do they do it, how do they deal with uncertainty.	Case studies of successful long-term decisions, interviews with people involved in long-term decisionmaking. Why was the policy successful, how can decisionmakers improve the success of a policy?
Clinical Trials	Controlled application of process to real world infrastructure planning problems in the form of pilot projects is needed.	Controlled implementation of adaptive policies for carefully selected infrastructure planning problems (i.e. pilot projects) is needed.

Theory

The kernel theories in which APM is rooted can be found in the policy analysis literature. They are derived from Assumption Based Planning (Dewar, et al., 1993) and systems analysis (Quade, 1982). In addition, the need for flexibility and adaptability in response to uncertainty is now well established, although opinions differ on the nature of the flexibility (Walker et al., 2001, Holling, 1978, de Neufville, 2000, Collingridge, 1980). As highlighted in the literature review, there is an extensive body of literature that presents theoretical arguments for the efficacy of dynamic flexible policies, such as AP. However, with respect to APP, there is still significant work needed, since there is no clear process for designing adaptive policies. Important questions that need to be addressed here include

- How do existing and new techniques and methods, such as forecasting and scenario analysis, fit into APM? In order to answer this question, one can draw on the generic policy analysis literature as well as the various bodies of literature dedicated to specific methods and techniques. For example, there is literature available on using scenarios as an early warning system (e.g. Botterhuis, van der Duin, de Ruijter, & van Wijck, 2010), which could meaningfully inform the development and design of a monitoring system. For a more thorough overview of the various purposes to which scenario planning can be adapted see (Varum & Melo, 2010a).
- How should the monitoring system with its associated triggers be institutionalized? Here one can draw on theories from political science and organizational theory, perhaps also neo institutional economics with respect to the design of incentives.
- How can adaptive policies be developed for multi actor problems? Here organization theory, decision science, and political science are the sciences from which kernel theories can be derived.

Animal models

Some work on validating adaptive policies through animal models has taken place. A first evaluation of APM along these lines can be found in (Kwakkel, Walker, & Marchau, 2007). A comparison is presented between a real world static policy and a fictitious adaptive version of the same policy. This comparison suggests that the adaptive version of the plan could have produced preferable results. A more rigorous analysis along the same lines is to quantitatively explore the performance of a static rigid policy versus a dynamic flexible policy across a large ensemble of futures through EMA. This has been completed recently for a case related to airport strategic planning (Kwakkel, et al., 2010).

In this recent case, an ensemble of simulation models was used in a series of computational experiments. By comparing the performance of an Adaptive Plan and a Traditional Plan as calculated by the computer models, one can reason about how these plans would behave in the real world. The basic approach for testing APM was: (i) develop a fast and simple model of the system of interest; (ii) generate an ensemble of future worlds; (iii) specify a traditional Plan and an Adaptive Plan; and (iv) calculate and compare the performance of both plans across an ensemble of future worlds using the fast and simple model. For the fast and simple model, several tools from the Federal Aviation Administration (the US regulator for air transport) for calculating airport performance were used to calculate airport performance in terms of noise, emissions, external safety, and capacity, The fast and simple model was complemented with a set of components that represent different exogenous developments, such as demand, technological development, and demographics. This set of components specifies the ensemble of future worlds. The two plans that were used, were based on the available policy documents discussing the long-term development of Amsterdam Airport Schiphol For the traditional Plan, a static rigid version of the current proposal was used. For the Adaptive Plan, various ideas put forward in the policy documents were utilized using the APM framework resulting in a flexible dynamic plan.

The EMA study revealed that traditional planning would perform better than APM only if the future demand were to fall into a narrow bandwidth. Conversely, if there is even small uncertainty about future demand, the adaptive plan is to be preferred. This is explained by the fact that the traditional plan is optimized for a specific future world. If the future world is very close or identical to this world, the traditional plan outperforms the adaptive plan. If we not only consider demand uncertainty, but also other uncertainties such as demographics and technological developments, the advantages of the adaptive plan over the traditional plan increase even more. This follows from the observation that APM has a smaller range of outcomes for the various outcomes of interest compared to traditional planning. That is, APM reduces the range of the expected outcomes given a range of uncertainties as compared to traditional planning.

If the focus is on evaluating the plan, the simulation game would focus on the performance of a static plan versus the performance of an adaptive plan given a scenario. Simulation gaming could also be employed for evaluating the process. If the emphasis is on evaluating the process, the focus of the game would be on APP and how this process is experienced by the participants in the game.

In addition to assessing the efficacy of AP using an ensemble of simulation models, we recently started testing APP using students (3 groups) and colleagues (1 group). Until now, in APM research, adaptive policies have been developed in an ad-hoc manner and represent the expert opinion of the authors of the papers. As a first step to overcome this problem, we conducted an APM workshop. The workshop was designed as a structured workshop, which consisted of 3 steps:

1. Pre-workshop phase: all the participants received decision support information in the form of a cost-benefit analysis in advance of the workshop, and were asked to read this before the workshop.
2. Selecting a promising basic policy: after an introduction on APM and the policy problem (how to implement innovative traffic safety measures), the participants were asked to select a basic policy as a group, drawing on the decision support information.
3. Enhancing the robustness of the basic policy: drawing upon well-established techniques such as SWOT (Strengths, Weaknesses, Opportunities and Treats) analysis (Ansoff, 1965) and TOWS (Treats, Opportunities, Weaknesses and Strengths) analysis (Weihrich, 1982), vulnerabilities and opportunities were identified. In order to support the discussion about how to handle these vulnerabilities and opportunities, a flowchart was used. Figure 5-2 shows an example.

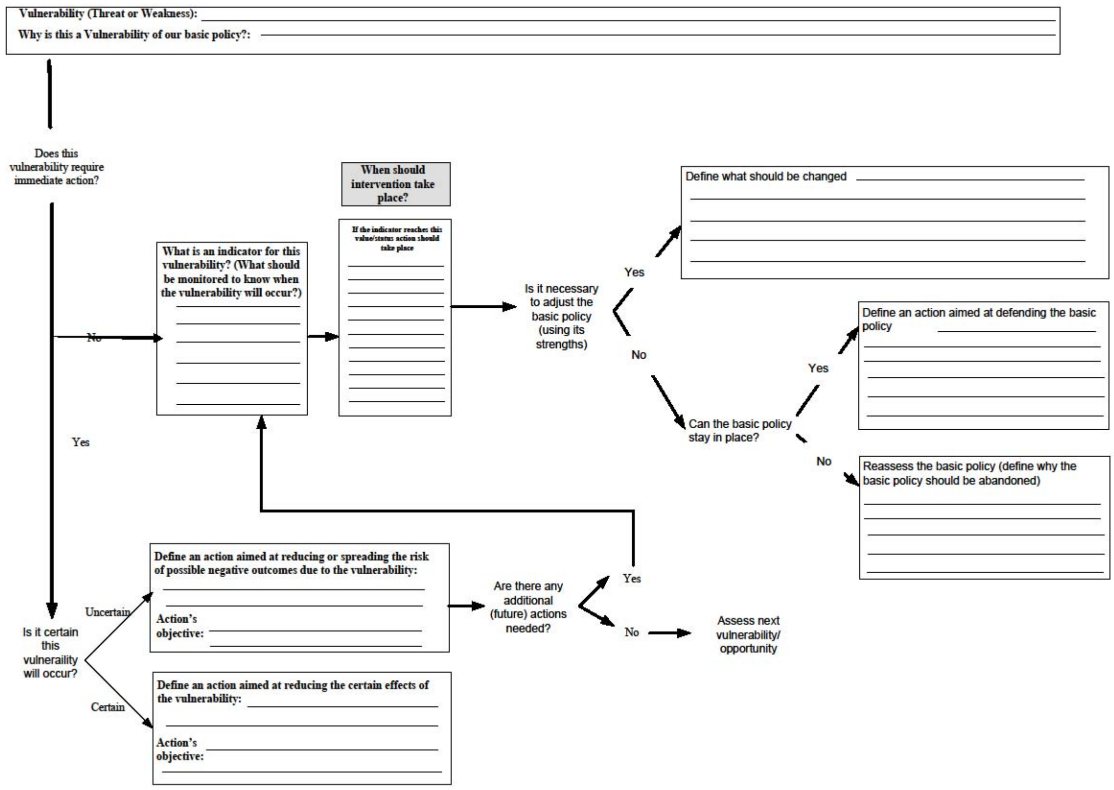


Figure 5-2: The flowchart used for specifying how to handle a specific vulnerability

Initial results suggest that both students and colleagues, neither of whom where intimately familiar with APM, were able to quite quickly agree on a comprehensive adaptive policy. More workshops, with other groups including real word decisionmakers, are needed to provide firmer evidence on the efficacy of developing adaptive policies through such workshops. The resulting plans of these workshops should then be analyzed to assess the extent to which these workshops result in policies that can cope with a wide range of uncertainties. Currently, it is planned to use expert opinion for this.

In vitro experiments

Characteristic of in vitro experiments is that human tissue is used outside the human body. The analogous case in testing APM would be to have the actual decisionmakers brought into a safe laboratory environment. Simulation gaming with these decisionmakers, along the same lines as discussed under animal models, could then be carried out. This would provide further evidence about the efficacy of both APP and AP. With respect to APP, the aforementioned workshop is planned to be executed with real-world actors. The results of this workshop are expected to give rich insight into the efficacy of developing adaptive policies using the APM framework.

Natural experiments

Over the course of history, examples can be found of successful long-term infrastructure planning. These examples could be studied in depth, through archival analysis, interviews with actors that were involved, and literature reviews. With respect to evaluating AP, such cases should be studied with respect to the effects of the policy. What can we learn from the effects of these policies with respect to adaptivity? From the perspective of evaluating APP, the emphasis should be on clarifying the process that led to the chosen policy. Understanding this process could offer supportive evidence for the efficacy of the adaptive policymaking process. Currently, the RAND Pardee Center for Longer Term Policy Analysis is working along these lines by identifying promising cases.

Clinical trials

The extent to which the analogue to clinical trials in evaluating infrastructure planning approaches can be achieved is debatable. With respect to AP, clinical trials are problematic, for the performance of the plan is to a large extent conditional on uncontrollable external developments. Still, it might be possible to have controlled implementation of both adaptive and traditional policies for carefully selected typical infrastructure planning problems, allowing for a comparison. For APP, the situation is less problematic. Here, pilot projects can be of great use. They would allow researchers to evaluate the new infrastructure planning process in confined but real-world conditions. This can generate evidence for the efficacy of the approach or lead the way for improving the process. However, as it stands, APM research is still quite removed from real-world implementation. There is preliminary evidence from animal models that plans using the outlined principles are preferable over static rigid plans and that developing plans in the outlined workshop format is effective. However, in vitro evidence is still lacking, and theoretical issues with respect to APP abound. Before seriously considering executing pilot projects, these other issues need to be addressed first.

5.5 Closing remarks

New approaches for handling deep uncertainty in infrastructure planning have been proposed. However, they have seen limited real world application, in part due to limited testing. This makes testing new approaches a key area of research. Testing the efficacy of these new planning approaches is, however, problematic for there is a methodological problem: it is almost impossible to compare a traditional approach with an innovative approach for a nearly identical infrastructure planning problem.

We presented a medical analogy and argued that in developing a methodology for testing new infrastructure planning approaches we can learn from the well-established practice of evaluating new medical treatments. The efficacy of medical treatments is evaluated by assessing different categories of evidence. Evidence comes from theory, animal models, in vitro experiments, natural experiments, and controlled clinical trials. Similarly, in evaluating new infrastructure planning approaches, evidence can be derived directly from theories. Evidence form animal models can be generated through Exploratory Modeling and Analysis and simulation gaming with students. In vitro experiments are possible through simulation gaming with actual decisionmakers. Natural experiments can take the shape of quasi-natural case studies of successful long-term infrastructure plans. Clinical trials are the most problematic category, but carefully crafted research oriented pilot projects might be the closest analogue available to researchers working on innovative dynamic adaptive infrastructure planning approaches.

We have further elaborated the medical analogy by specifying how each category of evidence can be used for validating APM. This showed how the medical analogy results in a checklist that can be used to structure and assess the available evidence for a specific innovative new infrastructure planning approach. With respect to APM, it revealed that the main evidence for the efficacy of APM is theoretical. However, there is emerging computational evidence and evidence from workshops with students that corroborates the theoretical evidence. Moreover, the analogy also revealed that the emphasis of APM research has been on the Adaptive Plan and not on the process of developing such plans. The theoretical evidence and computational experiments focus on the former, while only the workshop addresses the latter. Thus, the medical analogy allows one to spot gaps in the available evidence and thus it can also be used to guide research efforts. For APM, research is clearly needed with respect to the process of developing adaptive policies.

This medical analogy should be further elaborated and explored, paying close attention to the threats to validity. Medicine has a well-established approach to the evaluation of new medical treatments. This approach is, in our opinion, also applicable to the evaluation of new infrastructure planning approaches. Many of the issues and problems that one faces when evaluating new medical treatments also pose a problem to the evaluation of new infrastructure planning approaches. As such, the development of a methodology for evaluating new infrastructure planning approaches can benefit greatly from the accumulated knowledge and experience from medicine. In this paper, we introduced the analogy and explored some of its implications. We highlighted how the different categories of evidence could apply to infrastructure planning. In doing so, we only briefly touched upon the extensive literature on how one moves from evidence to valid conclusions and what threats to validity play a role in this. Exploring the threats to validity of the different categories of evidence is a major direction for further elaboration of the outlined evaluation methodology.

References

Agusdinata, D. B. (2008). *Exploratory Modeling and Analysis: A promising method to deal with deep uncertainty*. Ph.D. thesis, Delft University of Technology, Delft.

Albrechts, L. (2004). Strategic (spatial) planning reexamined. *Environment and Planning B: Planning and Design,* 31(5), 743-758.

Ansoff, H. I. (1965). *Corporate Strategy: An Analytic Approach to Business Policy for Growth and Expansion*. New York: McGraw-Hill.

Ascher, W. (1978). *Forecasting: An Appraisal for Policy Makers and Planners*. Baltimore: Johns Hopkins University Press.

Bankes, S. (1993). Exploratory Modeling for Policy Analysis. *Operations Research,* 4(3), 435-449.

Bankes, S. (2009). Models as Lab Equipment: Science from Computational Experiments. *Computational & Mathematical Organization Theory,* 15(1), 8-10.

Bankes, S., Lempert, R., & Popper, S. (2002). Making Computational Social Science Effective. *Social Science Computer Review,* 20(4), 377-388.

Batty, M., & Torrens, P. M. (2005). Modelling and prediction in acomplex world. *Futures,* 37(7), 745-766.

Ben-Haim, Y. (2006). *Information-Gap Decision Theory: Decisions under Severe Uncertainty* (2 ed.). New York: Wiley.

Botterhuis, L., van der Duin, P., de Ruijter, P., & van Wijck, P. (2010). Monitoring the future. Building an early warning system for the Dutch Ministry of Justice. *Futures,* 42(5), 454-465.

Bradfield, R., Wright, G., Burt, G., Cairns, G., & van der Heijden, K. (2005). The origins and evolution of scenario techniques in long range business planning. *Futures,* 37(8), 795-812.

Burghouwt, G. (2007). *Airline Network Development in Europe and its Implications for Airport Planning*. Burlington, Vermont: Ashgate Publishing Company.

Busenberg, G. J. (2001). Learning in Organizations and Public Policy. *Journal of Public Policy,* 21(2), 173-189.

Collingridge, D. (1980). *The Social Control of Technology*. London, UK: Frances Pinter Publisher.

de Neufville, R. (2000). Dynamic Strategic Planning for Technology Policy. *International Journal of Technology Management,* 19(3-5), 225-245.

de Neufville, R. (2003). Real Options: Dealing With Uncertainty in Systems Planning and Design. *Integrated Assessment,* 4(1), 26-34.

de Neufville, R., & Odoni, A. (2003). *Airport Systems: Planning, Design, and Management*. New York: McGraw-Hill.

Dempsey, P. S., Goetz, A. R., & Szyliowicz, J. S. (1997). *Denver International Airport: Lessons Learned*. New York: McGraw-Hill.

DerSimonian, R., & Laird, N. M. (1986). Meta-Analysis in Clinical Trials. *Controlled Clinical Trials,* 7, 177-188.

Dessai, S., Hulme, M., & Lempert, R. (2009). Do We Need Better Predictions to Adapt to a Changing Climate? *EOS,* 90(13), 111-112.

Dewar, J. A. (2002). *Assumption-Based Planning: A Tool for Reducing Avoidable Surprises*. Cambridge: Cambridge University Press.

Dewar, J. A., Builder, C. H., Hix, W. M., & Levin, M. H. (1993). *Assumption-Based Planning: A Planning Tool for Very Uncertain Times* (Report MR-114-A). Santa Monica: RAND.

Dewar, J. A., & Wachs, M. (2006). Transportation Planning, Climate Change, and Decisionmaking Under Uncertainty *Proceedings of the Workshop Conference on Climate Change and the US Transportation*. Washington, DC: National Academies of Sciences and Engineering.

Dewey, J. (1927). *The Public and its Problems*. New York: Holt and Company.

European Medicines Agency. (1998). *Note for Guidance on Statistical Prinicples for Clinical Trials*: European Medicines Agency.

Flyvbjerg, B., Bruzelius, N., & Rothengatter, W. (2003). *Megaprojects and Risk: An Anatomy of Ambition*. Cambridge: Cambridge University Press.

Frey, D. D., & Dym, C. L. (2006). Validation of design methods: lessons from medicine. *Research in Engineering Design,* 17(1), 45-57.

Glass, G. V. (1976). Primary, Secundary and meta-analysis of research. *Educational Researcher,* 5(10), 3-8.

Goodwin, P., & Wright, G. (2010). The limits of forecasting methods in anticipating rare events. *Technological Forecasting and Social Change,* 77, 355-368.

Gregor, S., & Jones, D. (2007). The anatomy of a design theory. *Journal of the Association for Information Systems,* 8(5), 312-335.

Hansman, R. J., Magee, C., De Neufville, R., Robins, R., & Roos, D. (2006). Research agenda for an integrated approach to infrastructure planning, design, and management. *International Journal of Critical Infrastructures,* 2(2-3), 146-159.

Hill, A. B. (1966). The Environment and Disease: Association or Causation. *Proceedings of the Royal Society of Medicine,* 58, 295-300.

Hodges, J. S. (1991). Six (or so) Things You Can Do With a Bad Model. *Operations Research,* 39(3), 355-365.

Hodges, J. S., & Dewar, J. A. (1992). *Is it You or Your Model Talking? A Framework for Model Validation* (R-4114-AF/A/OSD). Santa Monica: RAND.

Holling, C. S. (1978). *Adaptive Environmental Assessment and Management*. New York: John Wiley & Sons.

Hoogma, R., Kemp, R., Schot, J., & Truffer, B. (2002). *Experimenting for Sustainable Transport. The approach of Strategic Niche Management*. London: Spon Press.

IISD. (2006). *Designing Policies in a World of Uncertainty, Change and Surprise - Adaptive Policy-Making for Agriculture and Water Resources in the Face of Climate Change – Phase I Research Report*. Winnipeg: International Institute for Sustainable Development.

Kahan, J. P., Botterman, M., Cave, J., Robinson, N., Shoob, R., Thomson, R., et al. (2004). *Cyber Trust and Crime Prevention: Gaining Insight from Three Different Futures*: Prepared for Foresight Directorate, Office of Science and Technology, UK.

Kristóf, T. (2006). Is it Possible to make Scientific Forecasts in Social Sciences. *Futures,* 28(5), 561-574.

Kuechler, B., & Vaishnavi, V. (2008). On theory development in design science research: anatomy of a research project. *European journal of Information Systems,* 17, 489-504.

Kwakkel, J. H., Walker, W. E., & Marchau, V. A. W. J. (2007). *Dealing with Uncertainty in Airport Strategic Planning*. Paper presented at the 11th Air Transport Research Society World Conference.

Kwakkel, J. H., Walker, W. E., & Marchau, V. A. W. J. (2010). Adaptive Airport Strategic Planning. *European Journal of Transportation and Infrastructure Research,* 10(3), 227-250.

Lee, K. (1993). *Compass and Gyroscope: Integrating Science and Politics for the Environment*. Washington: Island Press.

Lee, K. (1999). Appraissing Adaptive Management. *Conservation Ecology,* 3(2), 3.

Lempert, R. J. (2002). A New Decision Sciences for Complex Systems. *Proceedings of the National Academy of Sciences of the United States of America,* 99(3), 7309-7313.

Lempert, R. J., & Groves, D. G. (2010). Identifying and Evaluating Robust Adaptive Policy Responses to Climate Change for Water Management Agencies in the American West. *Technological Forecasting and Social Change,* 77(6), 960-974.

Lempert, R. J., Popper, S., & Bankes, S. (2003). *Shaping the Next One Hundred Years: New Methods for Quantitative, Long Term Policy Analysis* (Report MR-1626-RPC). Santa Monica: RAND.

Marchau, V. A. W. J., & Walker, W. E. (2003). Dealing with Uncertainty in Implementing Advanced Driver Assistance Systems: An Adaptive Approach. *Integrated Assessment,* 4(1), 35-45.

Marchau, V. A. W. J., Walker, W. E., & van Duin, R. (2009). An adaptive approach to implementing innovative urban transport solutions. *Transport Policy,* 15(6), 405-412.

Mayer, I. S., & Veeneman, W. (2002). *Games in a world of Infrastructures*. Delft, the Netherlands: Eburon Publishers.

McCray, L. E., Oye, K. A., & Petersen, A. C. (2010). Planned adaptation in risk regulation: an initial survey of US environmental, health, and safety regulation. *Technological Forecasting and Social Change,* 77(6), 951-959.

McDaniel, R. R., & Driebe, D. J. (Eds.). (2005). *Uncertainty and Suprise in Complex Systems: Questions on Working the Unexpected*: Springer.

McLain, R. J., & Lee, R. G. (1996). Adaptive Management: Promises and Pitfalls. *Environmental Management,* 20(4), 437-448.

Pahl-Wostl, C., Moltgen, J., Sendzimir, J., & Kabat, P. (2005). *New Methods for Adaptive Water Management under Uncertainty - the NeWater Project*. Paper presented at the EWRA conference.

Pahl-Wostl, C., Sendzimir, J., Jeffrey, P., Aerts, J., Berkamp, G., & Cross, K. (2007). Managing change towards adaptive water management through social learning. *Ecology and Society,* 12(2), 30 [online].

Poot, J. (2000). A Synthesis of Empirical Research on the Impact of Government on Long-Run Growth. *Growth and Change,* 31, 516-546.

Popper, S., Griffin, J., Berrebi, C., Light, T., & Min, E. Y. (2009). *Natural Gas and Israel's Energy Future: A Strategic Analysis Under Conditions of Deep Uncertainty* (TR-747-YSNFF). Santa Monica, California: RAND.

Porter, A. L., Roper, A. T., Mason, T. W., Rossini, F. A., & Banks, J. (1991). *Forecasting and Management of Technology*. New York: John Wiley and Sons.

Quade, E. S. (1982). *Analysis for Public Decisions* (2nd ed.). New York: Elsevier Science Publishing Co.. Inc.

Raven, R. P. J. M. (2007). Niche Accumulation and Hybridisation Strategies in Transition Processes towards a Sustainable Energy System: An Assessment of Differences and Pitfalls. *Energy Policy,* 35(4), 2390-2400.

Rogers, E. (2003). *Diffusion of Innovation* (5 ed.). New York: The Free Press.

Schön, D. A. (1983). *The Reflective Practioner*. New York: Basic Books.

Simon, H. A. (1996). *The Sciences of the Artificial* (3rd ed.). Cambridge Massachusetts: The MIT Press.

Swanson, D., Barg, S., Tyler, S., Venema, H., Tomar, S., Bhadwal, S., et al. (2010). Seven tools for creating adaptive policies. *Technological Forecasting and Social Change,* 77(6), 924-939.

Thissen, W. A. H., & Twaalfhoven, P. G. J. (2001). Towards a Conceptual Structure for Evaluating Policy Analytic Activities. *European Journal of Operational Research,* 129(3), 627-679.

U.S. Department of Healt and Human Services - Food and Drug Administration. (1998). *Guidance for Industry - Providing Clinical Evidence of Effectiveness for Human Drugs and Biological Products*: U.S. Department of Healt and Human Services - Food and Drug Administration.

Van Geenhuizen, M., Reggiani, A., & Rietveld, P. (2007). New Trends in Policymaking for Transport and Regional Network Integration. In M. Van Geenhuizen, A. Reggiani & P. Rietvelt (Eds.), *Policy Analysis of Transport Networks* (pp. 1-16). Alsdershot: Ashgate.

van Geenhuizen, M., & Thissen, W. A. H. (2007). A Framework for Identifying and Qualifying Uncertainty in Policy Making: The Case Of Intelligent Transport Systems. In M. van Geenhuizen, A. Reggiani & P. Rietvelt (Eds.), *Policy Analysis of Transport Networks* (pp. 101-118). Aldershot: Ashgate.

Varum, C. A., & Melo, C. (2010a). Direction in Scenario Planning Literature - A review of the past decades. *Futures,* 42(3), 355-369.

Varum, C. A., & Melo, C. (2010b). Directions in scenario planning literature - A review of the past decades. *Futures,* 42(4), 355-369.

Verschuren, P., & Hartog, R. (2005). Evaluation in Design-Oriented Research. *Quality & Quantity,* 39, 733-762.

Volkery, A., & Ribeiro, T. (2009). Scenario planning in public policy: understanding use, impacts and the role of institutional context factors. *Technological Forecasting and Social Change,* 76(9), 1198-1207.

Vreugenhil, H. S. I., Slinger, J. S., Thissen, W., & Ker Rault, P. A. (in press). Pilot Projects in Water Management. *Ecology and Society*.

Walker, W. E., Marchau, V. A. W. J., & Swanson, D. (2010). Addressing deep uncertainty using adaptive policies: Introduction to section 2. *Technological Forecasting and Social Change,* 77(6), 917-923.

Walker, W. E., Rahman, S. A., & Cave, J. (2001). Adaptive Policies, Policy Analysis, and Policymaking. *European Journal of Operational Research,* 128(2), 282-289.

Walls, J. G., Widmeyer, G. R., & El Sawy, O. A. (1992). Building an Information Systems Design Theory for Vigilant EIS. *Information Systems Research,* 3(1), 35-59.

Walls, J. G., Widmeyer, G. R., & El Sawy, O. A. (2004). Assessing Information System Design Theory in Perspective: How Useful was our 1992 Initial Rendition? *Journal of Information Technology Theory and Application,* 6(2), 43-58.

Wardekker, J. A., de Jong, A., Knoop, J. M., & Van der Sluijs, J. P. (2010). Operationalising a Resilience Approach to Adapting an Urban Delta to Uncertain Climate Changes. *Technological Forecasting and Social Change,* 77(6), 987-998.

Weihrich, H. (1982). The TOWS Matrix - a Tool for Situational Analysis. *Long Range Planning,* 60(April).

Weiss, C. (1975). Evaluation Research in the Political Context. In E. L. Struening & Guttentag (Eds.), *Handbook of Evaluation Research* (pp. 13-26): Sage Publication Inc.

6 Assessing the efficacy of Adaptive Airport Strategic Planning: results from computational experiments

KWAKKEL, J.H., WALKER, W.E. & MARCHAU, V.A.W.J. (under review) Assessing the Efficacy of Adaptive Airport Strategic Planning: Results from Computational Experiments.

Abstract

This paper assesses the efficacy of a dynamic adaptive planning approach for guiding the long-term development of infrastructure. The efficacy of the general approach is tested on the specific case of airport strategic planning. Utilizing a fast and simple model of an airport, and a composition of small models that can generate a wide spectrum of alternative futures, the performance of a dynamic adaptive plan is compared to the performance of a static, rigid implementation plan across a wide spectrum of conceivable futures. These computational experiments reveal that the static rigid plan outperforms the dynamic adaptive plan in only a small part of the spectrum. Moreover, given the wide array of possible futures, the dynamic adaptive plan has a narrower spread of outcomes then the static rigid plan, implying that the dynamic adaptive plan exposes planners to less uncertainty about its future performance despite the wide variety of uncertainties that are present. These computational results confirm theoretical hypotheses in the literature that dynamic adaptive planning approaches are more efficacious for planning under uncertainty.

6.1 Introduction

Policymakers, business leaders, and governmental planners face a wide variety of uncertainties that can substantially affect the outcomes of their plans or policies. Moreover, many of these uncertainties are beyond their direct control. Such strategic decisionmaking problems are now frequently labeled as decisionmaking under deep uncertainty (Lempert, Popper, & Bankes, 2003; Volkery & Ribeiro, 2009; Walker, Marchau, & Swanson, 2010). Decisionmaking under deep uncertainty can be defined as the situation in which the decisionmaker does not know, or multiple decisionmakers cannot agree on, the system model, the prior probabilities for the uncertain parameters of the system model, and/or the value function (Lempert, et al., 2003). Over the last few years, a growing body of literature has emerged that focuses on deep uncertainty, the problems it creates for decisionmaking, and how to handle it (e.g. Ben-Haim, 2006; de Neufville, 2003; Dewar, 2002; Lempert, et al., 2003; Walker, et al., 2010).

One particular field in which deep uncertainty is of prime importance is infrastructure planning. The dominant approach in infrastructure planning has been to largely ignore the uncertainties (Dempsey, Goetz, & Szyliowicz, 1997; Marchau, Walker, & van Duin, 2009; Quade, 1982; Van Geenhuizen, Reggiani, & Rietveld, 2007; van Geenhuizen & Thissen, 2007). In the face of deep uncertainties about the future, the resulting infrastructure plans, based on a limited exploration of the uncertainties, have a great chance of failing. In response to this problem, various new planning approaches have been put forward. On the one hand (Swanson, et al., 2010), these approaches emphasize the need for a more thorough integrated forward-looking analysis of the uncertainties through techniques such as exploratory modeling and analysis (Agusdinata, 2008; Lempert, et al., 2003), bounce casting (Kahan, et al., 2004), and scenarios in various forms (Bradfield, Wright, Burt, Cairns, & van der Heijden, 2005; Varum & Melo, 2010). On the other hand, because of the limited capability of these techniques for anticipating rare events (Goodwin & Wright, 2010), there is a growing interest in flexibility and adaptability in plans in which a strategic vision of the future is combined with short-term actions and a framework that can guide future actions (Albrechts, 2004; Walker, et al., 2010; Walker, Rahman, & Cave, 2001).

Although the toolbox of the infrastructure planner has gradually expanded, more progress needs to be made in incorporating methods such as scenarios, exploratory modeling, Delphi studies, and techniques aimed at adaptivity and flexibility, such as real options analysis, into processes of policy design, choice, and implementation (Hansman, Magee, De Neufville, Robins, & Roos, 2006; Varum & Melo, 2010; Volkery & Ribeiro, 2009; Walker, et al., 2010). It has been suggested that the limited use of these methods and techniques is at least partly due to the lack of evidence with respect to the validity and applicability of these new planning approaches (Hansman, et al., 2006; Walker, et al., 2010), or, even stronger, that planning research focuses on the generation of new abstract concepts and then stops (Straatemeier, Bertolini, & Brömmelstroet, 2010). Given the profound importance of infrastructure to society, a careful assessment of the efficacy of new planning approaches is needed. Moreover, testing new planning concepts allows for the continuous improvement of these new concepts, thus closing the experiential learning loop (Straatemeier, et al., 2010).

When discussing the evaluation of new planning approaches, many researchers envision some form of controlled real world application (Hansman, et al., 2006). Similar to Straatemeier et al.(2010), we conceive of planning research as a design science. However, in distinction to them, we argue that there are other approaches besides the case study approach they advocate that can usefully be employed to assess the efficacy of new planning approaches. Often, the claim that each infrastructure problem is different serves to limit the potential of properly comparing different planning approaches on a like to like basis. Implicit in this claim is the assumption that something akin to randomized double blind controlled trails, well known from medicine, are the hallmark for assessing the efficacy of new treatments and approaches. However, a more thorough analysis of the practice of medicine would reveal that clinical trials are only the final stage in a much longer process of establishing the efficacy of new medical treatments. Prior to clinical trials, a sequence of theories, animal models, and in-vitro experiments have already been applied. We argue that these same forms of establishing the efficacy of new medical treatments can be adapted to the issue of assessing the efficacy of new planning approaches. In particular, we see the use of animal models in clinical research as analogous to the use of computer models. Applying this analogy, we use a computer model of an infrastructure system in a series of computational experiments to assess the efficacy of a new planning approach compared to a more traditional approach.

The structure of this paper is as follows. In Section 2, we elaborate on our methodology. In Section 3, we describe various infrastructure planning approaches, paying specific attention to those that emphasize the need for flexibility and adaptivity in coping with deep uncertainty. Section 4 describes the computational experiments and their results. Section 5 discusses these results. Section 6 presents our conclusions.

6.2 Methodology

In establishing the efficacy of new infrastructure planning approaches, one faces a methodological problem, for ”nothing done in the short term can ‘prove’ the efficacy of a planning methodology; nor can the monitoring, over time, of a single instance of a plan generated by that methodology, unless there is a competing parallel plan” (Dewar, Builder, Hix, & Levin, 1993). In this section, we address this problem. We turn to the experience and methodology that has emerged in medical research and development, for it has addressed the problem of assessing the efficacy of new treatments with a considerable degree of success.

Before turning to the analogy, we first demarcate the object of evaluation more sharply. In the evaluation literature, one important distinction is that between product and process (Walls, Widmeyer, & El Sawy, 1992, 2004) or between plan, process, and product (Verschuren & Hartog, 2005). If the object of evaluation is the plan, the evaluation focuses on the assessment of the quality of the design on paper (Verschuren & Hartog, 2005). Process evaluation focuses on the procedures for the construction of the design (Verschuren & Hartog, 2005; Walls, et al., 2004). Product evaluation involves the assessment of the value of the created artifact and its short- and long-term impacts after its creation (Verschuren & Hartog, 2005). The distinction between plan, process, and product corresponds to the distinction between the plan as written (plan), the process of drafting the plan (process), and the effects of the implemented plan (product). In this paper we focus on evaluating a final plan that utilizes innovative abstract concepts related to flexibility and adaptiveness.

Analogies express recognition of similarities between otherwise dissimilar things. Reasoning by analogy allows for adopting an alternative perspective that can produce novel insights (Porter, Roper, Mason, Rossini, & Banks, 1991). In the case of medical treatment versus uncertainty treatment, we use the analogy to sketch an approach for testing new infrastructure planning approaches. The key similarity is that in both cases there is a new treatment whose efficacy needs to be established prior to using the treatment in daily practice. Not carefully assessing the diverse effects of the new treatment poses a risk either to humans (in the case of medical treatments) or to society (in the case of infrastructure planning).

Researchers in medicine have extensive experience and a well-established methodology for validating new medical treatments. Although the process of evaluating medical treatments is subject to continuous improvement and refinement, it shares attributes with the evaluation of new infrastructure planning approaches. The emphasis in medical research and development is on the development of new treatments. These new treatments should improve the health of people. Before they can be applied in practice, their efficacy needs to be exhaustively tested. Specific attention is also given to the potentially harmful unintended effects of new treatments. The goal of evaluating new medical treatments is to assess whether the new treatment is beneficial, under what conditions it should be applied, and when it should not be applied. The main goal of research into new infrastructure planning approaches is the development of approaches for successfully handling deep uncertainty. These new approaches should result in infrastructure plans that produce adequate results regardless of how the future turns out. Given the importance of infrastructure planning to society, new planning approaches should be rigorously tested. Attention should be given to whether the new approaches are preferable to existing approaches, under what conditions these new approaches are applicable, and when these new approaches should not be used.

Table 6-1 summarizes the analogy by providing an overview of the types of evidence that are used in medical research and development and how analogous types of evidence can be applied to the evaluation of infrastructure planning approaches. The evidence is hierarchically ordered, from the most rudimentary evidence to the most detailed form of evidence.

Table 6-1: Types of evidence used to develop and validate medical treatments and infrastructure planning approaches (adapted from Frey & Dym, 2006)

Evidence used to develop/validate medical treatment	Evidence used to develop/validate infrastructure planning approaches
Theory	Theories (e.g. decision science, cognitive science, political science, organizational behavior, policy analysis)
Animal Models	Computational experiments of plans across an ensemble of futures
In Vitro Experiments	Simulation gaming with actual decisionmakers
Natural Experiments	Case studies of successful long-term infrastructure plans
Clinical Trials	Controlled field application of planning approach in the form of pilot projects

In this paper, we utilize computational experiments to provide evidence for the efficacy of adaptive infrastructure planning approaches. We, therefore, elaborate on this aspect of the medical analogy in some more detail. Animal models (often mice) are developed in medical research and development to display specific physiological characteristics that are of interest to the research. New treatments are subsequently first tested on these animal models. Animal models in medical research and development bear only a partial resemblance to the real human subject. Moreover, animal models are frequently engineered to exhibit desired characteristics. For example, mice are treated to grow cancers or genetically altered to exhibit behavior comparable to Attention-Deficit/Hyperactivity Disorder (ADHD). Animal models are comparable to simulation models frequently employed in engineering. Like animal models, simulation models bear a partial resemblance to the actual world, and, just like animal models, engineering models can be tailored to resemble a desired aspect of the world in the desired way.

In the context of assessing the efficacy of new infrastructure planning approaches, computer models can be used as surrogates for the real world system. With respect to the evaluation of plans, for example, the performance of a static rigid plan as calculated by a computer model can be compared with the performance of a dynamic adaptive plan, and one can then reason about how these plans would behave in the real world. However, given that infrastructure planning is decisionmaking about the future, there are significant uncertainties present. There is simply not enough knowledge to accurately forecast the future, the models of large infrastructure systems are often contested, and there are a variety of value-systems from the different stakeholders that are also bound to change over time. The performance of the plans derived from the different planning approaches needs to be assessed across these uncertainties. We have found that, when using simulation models as animal models for assessing the efficacy of infrastructure planning approaches, Exploratory Modeling and Analysis (EMA) is an appropriate method (Agusdinata, 2008; Bankes, 1993) for carrying out this assessment.

EMA is a research methodology to analyze complex and uncertain systems (Agusdinata, 2008; Bankes, 1993). It starts from the explicit recognition that, when researching such systems, there are profound and diverse uncertainties present that will significantly affect the outcomes of a model. Consolidative modeling approaches are thus ruled out on principle for studying such systems (Hodges, 1991; Hodges & Dewar, 1992). Since long-term infrastructure planning deals with complex systems, and these systems operate in a highly uncertain environment, the use of a consolidative approach is ruled out (Batty & Torrens, 2005; Dewar & Wachs, 2006; Marchau, et al., 2009; Van Geenhuizen, et al., 2007). However, in such situations there still is a wealth of knowledge and information available that supports a set of structurally different models across a range of parameter values. EMA aims at offering support for exploring this set of models across the range of plausible parameter values and drawing valid inferences from this exploration (Agusdinata, 2008; Bankes, 1993).

6.3 The treatment of uncertainty in infrastructure planning

Deep uncertainties about the future pose a significant challenge to infrastructure planning. One dominant approach in infrastructure planning has been to largely ignore the uncertainties or to try and reduce them (Dempsey, et al., 1997; Marchau, et al., 2009; McDaniel & Driebe, 2005; Quade, 1982; Van Geenhuizen, et al., 2007; van Geenhuizen & Thissen, 2007). Planners forecast the future situation by extrapolating past trends forward and develop static blueprint plans for achieving their desired goals in this situation. However, for a multitude of reasons, such forecasts are practically always wrong, since the future that materializes differs significantly from the forecasted future (Ascher, 1978; Batty & Torrens, 2005; Flyvbjerg, Bruzelius, & Rothengatter, 2003; Kristóf, 2006; Porter, et al., 1991). More enlightened approaches advocate robustness. That is, the plan should perform well in a few foreseeable alternative futures (called 'scenarios'). However, both of these approaches suffer from the problem that they focus on those uncertainties that are "among the least of our worries; their effects are swamped by uncertainties about the state of the world and human factors for which we know absolutely nothing about probability distributions and little more about the possible outcomes" (Quade, 1982). Similarly, Goodwin and Wright (2010) (p. 355) demonstrate that "all the extant forecasting methods – including the use of expert judgment, statistical forecasting, Delphi and prediction markets – contain fundamental weaknesses." And Popper, et al. (2009) state that the traditional methods "all founder on the same shoals: an inability to grapple with the long-term's multiplicity of plausible futures." Any infrastructure plan designed on the basis of a few forecasts or a small set of assumptions about the future is likely to perform poorly, and unplanned ad-hoc adaptations are needed to improve its performance.

In response to the deficiencies of traditional planning, an alternative planning paradigm has emerged. This paradigm holds that, in light of the deep uncertainties, one needs to plan dynamically and build in flexibility (Albrechts, 2004; de Neufville, 2000; IISD, 2006; Lempert, 2002; Lempert & Groves, 2010; Swanson, et al., 2010; Walker, et al., 2001). According to this paradigm, the solution to planning under uncertainty is to create a shared strategic vision of the future, commit to short-term actions, and establish a framework to guide future actions. (Albrechts, 2004; Walker, et al., 2001). A plan that embodies these ideas allows for the dynamic adaptation of the plan over time to meet the changing circumstances. This planning paradigm, in one form or another, has been receiving increasing attention in various disciplines. In infrastructure planning, the need for adaptivity and flexibility is increasingly recognized. For example, in air transport, the developments of the last decade, including various terrorist attacks, SARS, Mexican flu, and the second Gulf war, have highlighted this need. Combine this with the impacts of privatization and liberalization, the rise of airline alliances, mergers, and take-overs, and the emergence of new players in the industry, such as low cost carriers, and it is obvious that it is next to impossible to plan for the long-term development of an airport based on a prediction of the size and composition of future demand. In response to these uncertainties, the need for dynamic adaptive planning has been forcefully argued (Burghouwt, 2007; de Neufville, 2000; de Neufville & Odoni, 2003; Kwakkel, Walker, & Marchau, 2010; Walker, et al., 2001). A similar line of reasoning can also be found with respect to port development (Taneja, Walker, Ligteringen, van Schuylenburg, & van der Plas, 2010). Another argument for dynamic adaptation in the transport domain comes from research on transport innovations. The implementation of innovations, such as advanced driver assistance systems and innovative approaches for intra-city logistics, is hampered by a variety of uncertainties, including uncertainties about the technology to be implemented and about the future structure of the transport system itself. Dynamic flexible implementation plans have been put forward as a way to overcome these problems (Erikson & Weber, 2008; Marchau & Walker, 2003; Marchau, et al., 2009; van Zuylen & Weber, 2002). In other domains, the need for adaptivity and flexibility is argued on very similar grounds. For example, in integrated river basin management, the omnipresence of uncertainties in both the environmental system and the societal system is used as an argument for adaptivity and flexibility (Pahl-Wostl, Moltgen, Sendzimir, & Kabat, 2005; Pahl-Wostl, et al., 2007). Policymaking with respect to climate change is yet another area in which dynamic adaptation and flexibility are suggested as the appropriate approach for policy design (Dessai, Hulme, & Lempert, 2009; Smith, 1997; Wardekker, de Jong, Knoop, & Van der Sluijs, 2010).

Figure 6-1 shows a framework that operationalizes the high level outline of the new planning paradigm, which we call dynamic adaptive planning (DAP). In short, in Step I, the existing conditions of an infrastructure system are analyzed and the goals for future development are specified. In Step II, the way in which this is to be achieved is specified. This basic plan is made more robust through four types of actions, which are specified in Step III: *mitigating actions* are actions to reduce the *certain* adverse effects of a plan; *hedging actions* are actions to spread or reduce the risk of *uncertain* adverse effects of a plan; *seizing actions* are actions taken to seize certain available opportunities; and *shaping actions* are actions taken to reduce the chance that an external condition or event that could make the plan fail will occur, or to increase the chance that an external condition or event that could make the plan succeed will occur. Even with the actions taken in Step III, there is still the need to monitor the performance of the plan and take action if necessary. This is called contingency planning, and is specified in Step IV. *Signposts* specify information that should be tracked in order to determine whether the plan is achieving its conditions for success. Critical values of signpost variables (*triggers*) are specified, beyond which actions should be implemented to ensure that the plan keeps moving the system in the right direction and at a proper speed. There are four different types of actions that can be triggered by a signpost: *defensive actions* are taken to clarify the basic plan, preserve its benefits, or meet outside challenges in response to specific triggers that leave the basic plan remains unchanged; *corrective actions* are adjustments to the basic plan; *capitalizing actions* are actions trigger to take advantage of opportunities that improve the performance of the basic plan; and a *reassessment* of the plan is initiated when the analysis and assumptions critical to the plan's success have clearly lost validity.

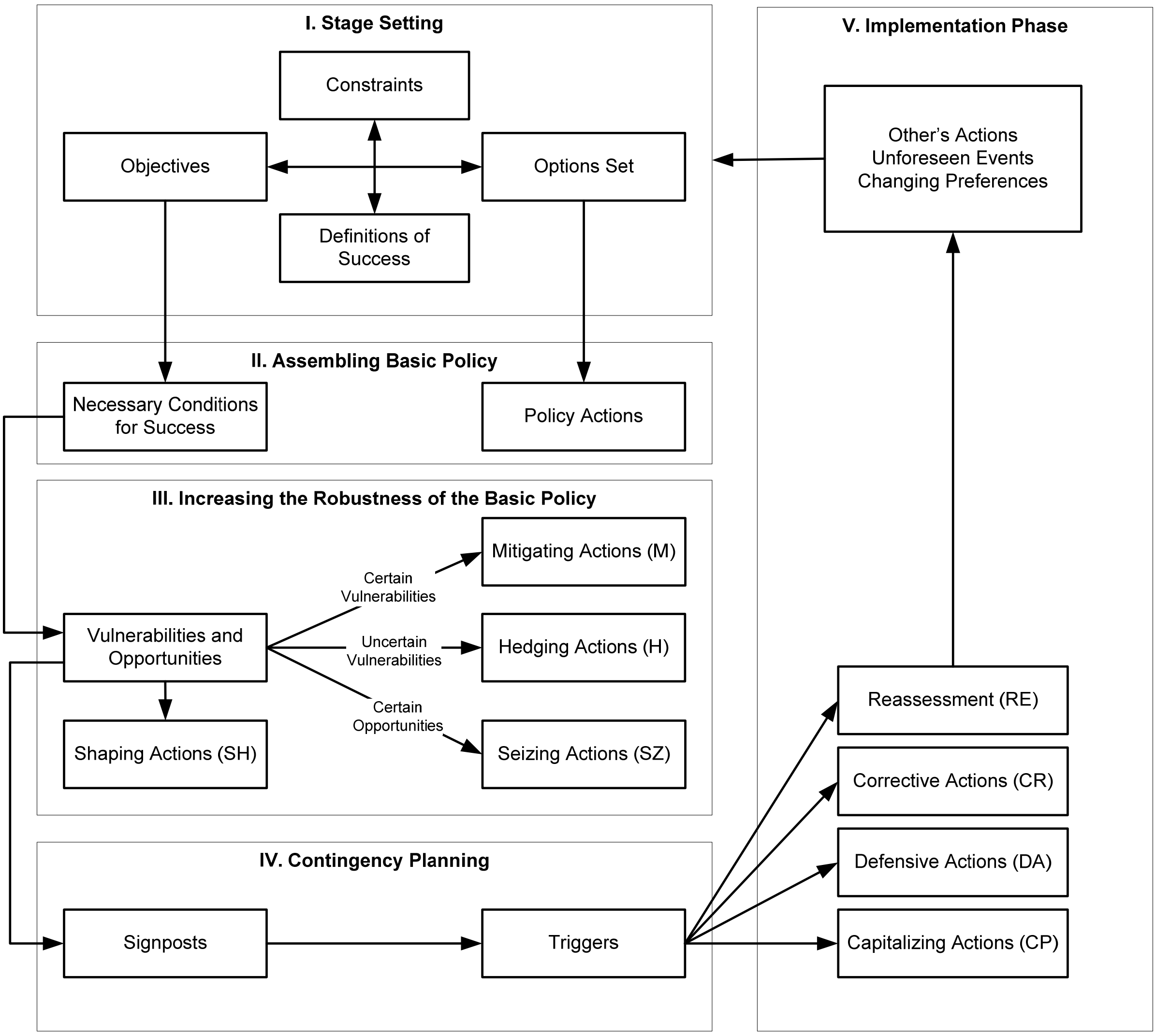


Figure 6-1: The steps of dynamic adaptive planning (Kwakkel, et al., 2010)

Step V is the actual implementation. In this step, the actions to be taken immediately (from Step II and Step III) are implemented, and a monitoring system (from Step IV) is established. Then time starts running, signpost information related to the triggers is collected, and actions are started, altered, stopped, or expanded in response to this information. After implementation of the initial actions, the implementation of other actions is suspended until a trigger event occurs. For a more detailed explanation of this framework, see Kwakkel et al. (2010), Marchau et al. (2009), and Walker et al. (2001).

From the foregoing, we conclude that uncertainty is a central problem in long-term infrastructure planning. A large body of literature exists that argues that in order to handle these uncertainties, infrastructure planning needs to shift from the static rigid policymaking paradigm to the dynamic adaptive policymaking paradigm. Despite the discussion in the literature, these dynamic flexible approaches for policymaking have seen limited real world applications (Hansman, et al., 2006; IISD, 2006; Lempert & Groves, 2010; McCray, Oye, & Petersen, 2010; Walker, et al., 2010). It has been suggested that one of reasons for this limited application is the fact that evidence for the efficacy of these approaches is lacking (Dewar, et al., 1993; Hansman, et al., 2006; Walker, et al., 2010). In terms of Table 6-1, the efficacy of new planning approaches is mainly argued on theoretical grounds. However, other sources of evidence are needed. This demand for evidence for the efficacy of new planning approaches is in turn motivated by the importance of infrastructure to society. Using new unproven planning approaches poses a risk, and evidence of the efficacy of new infrastructure planning approaches can be useful in convincing decisionmakers to use them. Moreover, the increasing evidential base for the efficacy of new planning approaches can help to convince planners to start adopting these new approaches.

6.4 Computational Experiments

6.4.1 Background

In this paper, we apply EMA to a specific case in order to compare the efficacy of a static rigid plan and a dynamic adaptive plan. The two plans are tested across a wide range of uncertainties, because the handling of uncertainties by a given planning approach determines the efficacy to a large extent. The results are not meant to be definitive, but are suggestive of the benefits that might accrue in practice from using dynamic adaptive planning. We choose to use the current challenges Amsterdam Airport Schiphol is facing in its long-term development as our case. As outlined below, Schiphol faces a range of uncertainties that could affect the airport in different ways. In addition, we are familiar with the current situation of Schiphol. The uncertainties the airport currently faces have been studied recently (Kwakkel, Walker, & Wijnen, 2008), a multitude of policy documents from multiple stakeholders is readily available, the data necessary to quantify a model for calculating airport performance metrics is also available, and an outline for a dynamic adaptive plan for the long-term development of Schiphol has recently been presented (Kwakkel, et al., 2010).

Aviation demand has experienced unprecedented growth since the early 1990’s, fuelled by privatization and liberalization of the aviation industry. Amsterdam Airport Schiphol has benefited from this growth and has evolved into one of the European Union’s major hubs. Since 1990, Schiphol has expanded its runway system and its terminal. Parallel to the increasing number of passengers and flights handled at Schiphol, negative external effects have also increased, resulting in regulations concerning noise, emissions, and third-party risk. This situation causes increasing tension between capacity, environment, and safety at and around the airport. Currently, Schiphol’s position as a hub within Europe is under pressure. In 2006, Schiphol was surpassed by Madrid’s Barajas Airport and now ranks as Europe’s fifth airport in terms of both air transport movements and passenger movements (Schiphol Group, 2010). The merger of Air France and KLM has resulted in the threat that KLM, Schiphol’s hub carrier, which is responsible for the majority of the scheduled aircraft movements at the airport, might move a significant portion of its operations to Charles de Gaulle Airport. The other major airports in Europe are planning on expanding their capacity or are developing dual airport systems. Together, this makes the long-term planning for Schiphol both urgent and problematic.

6.4.2 The Two Plans

Currently, a variety of stakeholders, such as the Ministry of Transport, the Schiphol Group, municipalities around Schiphol, KLM, and Netherlands Air Traffic Control, are in the process of drafting a plan for the long-term development of Schiphol. We chose to use a planning horizon of thirty years (CPB, KiM, NMP, & RP, 2007). In this case study, we approach the problem from the perspective of this network of actors that is responsible for the governance of Amsterdam Airport Schiphol. The main goals of this governance network are: (1) to create room for the further development of the network of KLM and its Skyteam partners, and (2) to minimize (and, where possible, reduce) the negative effects of aviation in the region (Schiphol Group & LVNL, 2007). There are several types of changes that are currently being considered by the governance network in order to achieve these goals. The physical capacity can be expanded by using its existing runways and terminals more efficiently and/or building new capacity. More explicitly, among the options that are considered:

1. Add a new runway parallel to one of the existing runways
2. Move charter operations out of Schiphol (i.e., to Lelystad and Eindhoven, which have a planned capacity of roughly 70,000 operations per year (Rijksoverheid, 2009; Schiphol Group, 2007))
3. Limit available slots

For the static rigid plan (which we call the Master Plan), we assume that Schiphol will add the new runway and that it will become operational in 2020. Furthermore, up to 70,000 operations will be moved away from Schiphol. We assume that this will be done over the course of five years, from 2015 to 2020. No slot limitation will be implemented, because it is assumed that there is enough environmental capacity available to accommodate the expected demand.

The dynamic adaptive plan (which we call the DAP) is adapted from (Kwakkel, et al., 2010). For the DAP, the basic plan includes planning for all the infrastructure options without beginning to build any of them. The basic plan is made more robust through the actions outlined in Table 6-2. The contingency plan is outlined in Table 6-3. With respect to the reassessment actions, if these are triggered they will be recorded, but the model run will still be completed and the outcomes for that run will be recorded as normal. With respect to the actions that are taken to influence technological development, we assume they are taken; the uncertainty about their effectiveness is included in the model through the ability to generate a range of possible technological improvements via the scenario generator.

Table 6-2: Increasing the robustness of the basic plan

Vulnerabilities and Opportunities	Hedging (H) and Shaping (SH) Actions
Demand for air traffic grows faster than forecast.	**H**: Prepare Lelystad and Eindhoven airport to receive charter flights
Increase in the population density in area affected by noise	**H:** Test existing noise abatement procedures such as the continuous descent approach, outside the peak periods (e.g. at the edges of the night) **SH:** Maintain land use reservation that allows for building the new runway
Maintain current trend of decrease of environmental impact of aircraft	**SH:** Negotiate with air traffic control on investments in new air traffic control equipment that can enable noise abatement procedures such as the continuous descent approach **SH:** Invest in R&D, such as noise abatement procedures
Development of wind conditions due to climate change	**H:** Have plans ready to quickly build the sixth runway, but do not build it yet. If wind conditions deteriorate even further, start construction

Table 6-3: The contingency plan

Vulnerabilities and Opportunities	Monitoring and Trigger System	Actions (Reassessment (RE), Corrective (CR), Defensive (DA), Capitalizing (CP))
Demand for air traffic grows faster than forecast.	Monitor the growth of Schiphol in terms of aircraft movements. If this exceeds 450.000 operations, start building the new runway. The new runway becomes available five years after this trigger is reached. If demand approaches 510.000 aircraft movements, activate CR action. If it exceeds 510.000, trigger RE.	**CP:** Begin to implement the plan for the new runway **CR**: Move a portion of the operations to Lelystad and Eindhoven. **RE**: Reassess entire plan
Increase area affected by noise	Monitor area affected by noise. If area affected by noise increases by 20% compared to start year, take DA-action; by 50%, take CR-action; by 75%, take RE-action. If area decreases by 20%, take CP-action.	**DA:** Slowdown of growth by limiting available slots **CR:** Slowdown of growth by limiting available slots even more **RE:** Reassess entire plan **CP:** Make new slots available.
Development of wind conditions due to climate change	Monitor the usage percentage of the cross-wind runway, If this increases by more than 10 percent compared to the start year, take DA action.	**DA:** Begin to implement the plan for the new runway. If this action is taken, the new runway becomes available five years later.

In this comparison, we focus on comparing the performance of the plans. We assume that the plans are executed as written. Uncertainties regarding the implementation are not taken into account. In practice, airport planners under a Master Planning regime do sometimes defer or speed up aspects of the static rigid plan, implement (formal) updates to the plan, and try and use operational changes to accommodate the divergence between forecasted and real world circumstances. This behavior is not included in the comparison. These types of adaptations to the plan that can occur in practice will reduce the difference in performance between the Master Plan and the DAP. So our comparison of the two provides an indication of the results for the best case for the DAP. In short, the comparison focuses on the theoretical differences between a static rigid plan and a dynamic adaptive plan, not their real world implementations. We return to this point in Section 6.5.

6.4.3 The ensemble of models

In order to quantitatively explore and compare the performance of the two plans, one or more models are needed. In this specific case, we chose to have a variety of *generators*, while we use a single model for calculating airport performance. This approach is motivated by the fact that there is relatively minor uncertainty about the internal functioning of an airport, while there is significant uncertainty about future developments. Figure 6-2 shows the basic structure of the model that is used. A given model structure consists of several generator components and the airport performance analysis component. For each of the generators, several structurally different versions are available.

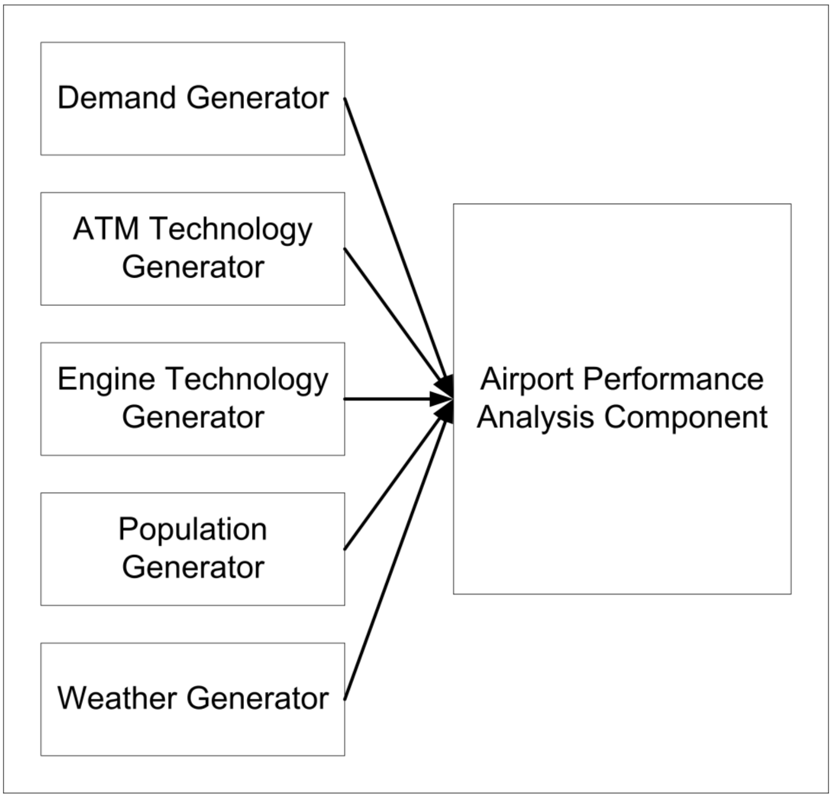


Figure 6-2: Conceptual design of the model

Airport performance analysis component

The airport performance analysis component of the model is based on the computational core of the HARMOS decision support system for airport strategic planning (Heblij & Wijnen, 2008; Wijnen, Visser, & Walker, 2009; Wijnen, Walker, & Kwakkel, 2008), with the exception that a different model is used for calculating noise. The selection of tools is motivated by the purpose of FASMAPA: to allow for a strategic quick scan of the performance of alternative plans. Macroscopic tools have therefore been used (de Neufville & Odoni, 2003; Stamatopoulos, Zografos, & Odoni, 2004). Table 6-4 specifies the tools that are used. FASMAPA has been implemented using the object oriented programming language Python. For the details on the validation see the references in Table 6-4. For more details on the integration of these tools and its validation, see the aforementioned HARMOS literature and Kwakkel et al. (2009).

Table 6-4: Tools integrated in FASMAPA

Airport Performance Aspect	Tool
Capacity	FAA Airfield Capacity Model (FCM) − an extension of the classic Blumstein model (Agusdinata, 2008; de Neufville & Odoni, 2003; FAA, 1981). It is a macroscopic tool for estimating the maximum hourly throughput capacity of the runway system and final approach. It does not consider aprons, taxiways, and the terminal area airspace.
Noise	Area Equivalent Method (AEM)– a model that approximates Integrated Noise Model (INM) results (FAA, 2008). Compared to INM, AEM does not consider flight paths. Other than this, the way of calculating the size of noise contours is the same.
Emissions	Emission Dispersion Modeling System (EDMS) − the FAA required tool for emission analysis (FAA, 2009), EDMS consists of an emissions inventory system that calculates the sum of emissions from various sources including aircrafts, Auxiliary Power Units, roadways etc. It also contains a component for dispersion analysis, taking weather etc. into account. FASMAPA only covers the emission inventory part.
Third Party Risk	Methodology developed by the National Air Traffic Services (NATS) for third-party risk (Cowell, et al., 2000; Cowell, Gerrard, & Paterson, 1997) − the NATS methodology has been extended to apply to multiple runways (Heblij, 2004, Heblij and Wijnen, 2008). The NATS methodology calculates the probability and effect of a crash for a given location relative to a runway in light of historical data about the crash frequency for different aircraft categories and the different parts of the landing take-off cycle.

The generators

FASMAPA focuses on calculating airport performance for a given set of inputs on a yearly basis. By providing input for several years, airport performance over time can be calculated. In order to generate these input parameters, FASMAPA has been complemented with a set of generators that generate its inputs. A specific combination of generator components could be called a ‘scenario generator’ (Lempert, et al., 2003). This scenario generator allows for generating demand volumes, wind conditions, technological developments, and changes in demographic patterns around the airport. With respect to the different generators that make up the scenario generator, each of them uses different structural assumptions. Table 6-5 summarizes the uncertainties that can be explored with FASMAPA and the scenario generator. Where information on the ranges was available, the source is given. The remainder reflect numbers encountered in various policy documents and expert opinions, which are deemed plausible. With respect to the various numbers about the future size of demand, these are inspired by the travel time budget analyses of (Schafer, Heywood, Jacoby, & Waitz, 2009; Schafer & Victor, 2000). Analogously to scenario planning, what is important in EMA is not to provide detailed predictions of what will happen, but to use ranges that are plausible.

Table 6-5: Overview of the uncertainties

Types of Uncertainties	Variations Examined
Model structure uncertainties	
Demand	– Exponential growth – Logistic growth – Logistic growth followed by logistic decline
Wide body vs. narrow body aircraft mix	– Linear change in ratio of wide body – Logistic change in ratio of wide body
Population	– Logistic growth – Logistic growth to a maximum followed by logistic decline
ATM technology	– Exponential performance increase – logistic performance increase
Engine technology (noise/emissions)	– Exponential performance increase – logistic performance increase
Parameterization of structural uncertainties	
Exponential demand growth	Ratio over 30 years, range:1-1.06: this corresponds to 0% growth in air transport movements per year and 6% growth in air transport movements per year
Logistic demand growth	Demand over 30 years, range: 390000-800000 air transport movements
Logistic demand growth, followed by logistic decline	– Demand over 30 years, range: 250000-375000 air transport movements – Year in which maximum will be reached, range: 2012-2020 – Demand in year of maximum, range: 400000-435000 air transport movements
Linear change in wide body ratio (De Haan, 2007)	Ratio of wide body to narrow body aircraft, range: 0.05-0.5; this corresponds to 5% and 50% wide body air transport movements respectively
Logistic change in wide body ratio (De Haan, 2007)	Ratio of wide body to narrow body aircraft, range: 0.05-0.5; this corresponds to 5% and 50% wide body air transport movements respectively
Logistic population growth (de Jong & Hilderdink, 2004)	Population size in 30 year, range: 17-20 million people
Logistic population growth to max followed by decline (de Jong & Hilderdink, 2004)	– Population size in 30 year, range:12-16 million people – Population size at maximum, range: 16-17 million people – Year maximum will be reached, range: 2010-2020
Exponential ATM technology development	ATM technology improvement in 30 years, range: 0.85-1, this implies a reduction in separation standards of 15% and 0% respectively
Logistic ATM technology development	ATM technology improvement in 30 years, range: 0.85-1, this implies a reduction in separation standards of 15% and 0% respectively
Exponential engine technology development (De Haan, 2007)	Engine technology improvement technology improvement over 30 years, range: 0.65-1; this implies a reduction in emissions of between 45% and 0% respectively
Logistic engine technology development (De Haan, 2007)	Engine technology improvement technology improvement over 30 years, range: 0.65-1; this implies a reduction in emissions of between 45% and 0% respectively
Parametric uncertainties	
Weather (KNMI, 2006)	Percentage of change in days with severe wind conditions per year, range: -1% - +4%; this implies between 1% increase and 4% decrease of availability of affected runway configurations.

6.4.4 Results

There are two specific questions we answer in this section:

– Question 1: given the ranges of uncertainties and the different models, what is the range of outcomes that both the Master Plan and the DAP can generate? If the DAP is any good, its range of outcomes should be smaller; or, put differently, the risks associated with the DAP should be smaller.

- Question 2: given the most favorable conditions for a Master Plan, how well does the DAP do in comparison? This question is motivated by the idea that, if an adaptive plan is to be attractive to decisionmakers, it should perform about equal to the static rigid plan or better, even under those conditions that most favor the rigid plan.

As outcome indicators, we used LDN (a metric for averaging day and night flights by penalizing night flights) for noise, Average Casualty Expectance (ACE) for third-party risk, Practical Annual Capacity (PANCAP) utilization for the airport's physical capacity, latent demand for unaccommodated demand, and cumulative CO emissions as a proxy for environmental impacts. LDN and ACE are well established metrics. Similarly, PANCAP is a high level metric for a quick assessment of capacity on a yearly basis (de Neufville & Odoni, 2003). Ideally, insight into delays should also be given, however the FCM tool does not generate this. To create some high level insight into the ratio of capacity and demand, we calculate the PANCAP – accommodated demand ratio. This is defined as the minimum of the demand and the maximum annual throughput capacity given opening hours of the airport, divided by the PANCAP. If demand is higher than the maximum annual throughput capacity, this portion of the demand will not be accommodated and becomes part of the latent demand. Admittedly, this ratio is a rough proxy for delays, but given that the presented comparison is at a strategic level, overly detailed models are inappropriate (Stamatopoulos, et al., 2004). Such detailed calculations only create a false sense of precision, since their true values are overwhelmed by the uncertainties. Moreover, if a more detailed delay analysis were included, additional assumptions with respect to among others the distribution of planned flights over the day, the deviation between actual and scheduled arrivals and departures, and the distribution of flights over the year would be needed. The choice for CO as a proxy for overall environmental impacts is motivated by the available data (the ICAO engine emissions data bank) pertaining to the emissions of the various engines during the respective parts of the Landing-Take Off (LTO) cycle. This databank has data only for NO_x, CO, SO^2, and VOC.

Question 1 above reduces to solving a boundary problem. For each outcome of interest, the upper and lower bound across both the model structures and parametric uncertainties needs to be determined. Technically, this requires the use of a non-linear optimization algorithm. For each outcome of interest, across each model structure and its parameter ranges, and across both plans, the minimum and maximum should be determined. To fully explore this would require 1344 separate optimizations (7 outcomes of interest * 48 scenario generators * 2 plans * upper bound + lower bound), regardless of the problem of local versus global optima. However, these need not all be explored. For example, the lower bound on the cumulative CO-emissions will be found in that model structure that has the lowest number of flights. Similar arguments can be presented for the other outcomes, reducing the number of optimization problems significantly. The results of the overall analysis are shown in Table 6-6. From this table, it can be concluded that the DAP has a smaller bound on all outcome indicators except for latent demand. The results for latent demand are explained by the fact that the DAP has triggers in place to limit the size of the noise contour. If these are triggered, less demand is accommodated, thereby increasing latent demand. The high value of the PANCAP – accommodated demand ratio for the upper bound of the Master Plan is explained by the large discrepancy between maximum annual throughput capacity, which determines the maximum accommodated demand, and PANCAP. Such a high figure implies that the airport is heavily capacity constrained and is experiencing severe delays.

Table 6-6: Performance bounds of the Master Plan and the DAP

Outcome indicators	Lower Bound Master Plan	Lower Bound DAP	Upper Bound Master Plan	Upper Bound DAP
Size of 65 LDN contour after thirty years (km^2)	13.2	10.2	63.8	47.4
Max. size of 65 LDN contour (km^2)	17.9	17.9	63.8	47.7
Cumulative ACE (ACE)	0.9	1.1	2.7	2.3
PANCAP - accommodated demand ratio after thirty years	0.25	0.89	2.48	1.1
Max PANCAP - accommodated demand ratio	0.9	0.52	2.48	1.1
Accumulated latent demand (flights)	0	0	5,058,504	8,290,622
Cumulative CO emission (kg)	21,520.9	19,773.9	195,729.1	103,899.5

In order to identify the difference in performance between the DAP and the Master Plan, we follow an approach similar to Agusdinata (2008) and Lempert et al. (2003). First, we identify the combination of uncertain parameters under which the Master Plan performs the best compared with the DAP. So, we try to find the best case for the Master Plan compared with the DAP. Once this point is identified, all uncertain parameters apart from demand growth and the wide body ratio are fixed to their values at this point. The choice for demand growth and wide body ratio is motivated by the observation that the main uncertainties in airport strategic planning are about the size and composition of future demand (Burghouwt, 2007; de Neufville & Odoni, 2003). A full factorial design is generated for the wide body ratio and demand growth per year, with 21 samples for each, resulting in 441 cases. For each case, the performance difference is calculated.

In order to determine the performance difference, we use ACE, PANCAP - accommodated demand ratio, latent demand, size of the 65 LDN contour, and cumulative CO emissions as outcome indicators. Next, these indicators are normalized, using the maxima and minima from Table 6-6, so that they scale between 0 (bad) and 1 (good). That is, the actual outcomes are mapped to a unit interval in order to make them comparable. So, for example the maximum size of the noise contour (63.8 km^2) is mapped to 0, and the minimum size of the noise contour (10.2 km^2) is mapped to 1. The five normalized outcome indicators together are a performance vector that describes the performance of a plan. We then define the performance of a plan as the length of the performance vector, using the Euclidian norm. The performance difference between the two plans then becomes the difference in length between the performance vector of the Master Plan and the performance vector of the DAP.

The resulting performance differences are shown in Figure 6-3. Grayscale is used to indicate the value of the performance difference. If this value is below 0, the rigid plan is 'better' than the adaptive plan. From this figure, we conclude that, even under the conditions that *most* favor the rigid plan, the Master Plan is only slightly better than the DAP. Furthermore, the Master Plan is better only in a relatively small area. So, if the wide body ratio and/or the demand growth deviate slightly from those that are the best for the Master Plan, the Master Plan will perform worse than the DAP.

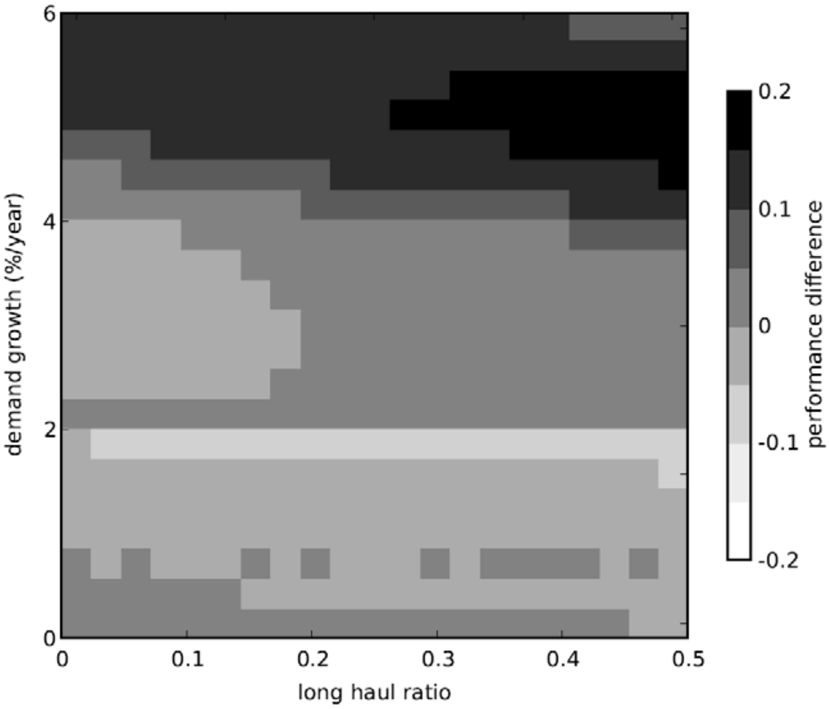


Figure 6-3: Performance difference of the Dynamic Adaptive Plan compared to the Master Plan for that combination of uncertain parameters that most favor the Master Plan

Figure 6-4 shows the same type of analysis as Figure 6-3 with a single important difference. Here, the conditions that *most* favor the DAP have first been identified. This figure has a structure very similar to Figure 6-3, suggesting that the other uncertainties have indeed only a minor influence on the performance of the plan. Furthermore, it shows again that the Master Plan performs slightly better than the DAP only in a narrow bandwidth of demand growth.

Figure 6-3 and Figure 6-4 cover only growing demand. However, these figures do suggest that if demand were to decline, the DAP would also be preferable. This makes sense since, in those cases, the new runway would not be built, resulting in double the utilization compared to the Master Plan (see also Table 6-6). To test this hypothesis, we looked at the model structure that covers declining demand and explored it using the same approach as used to generate Figure 6-3 and Figure 6-4. We did not find a single case in which the Master Plan would perform better than the DAP.

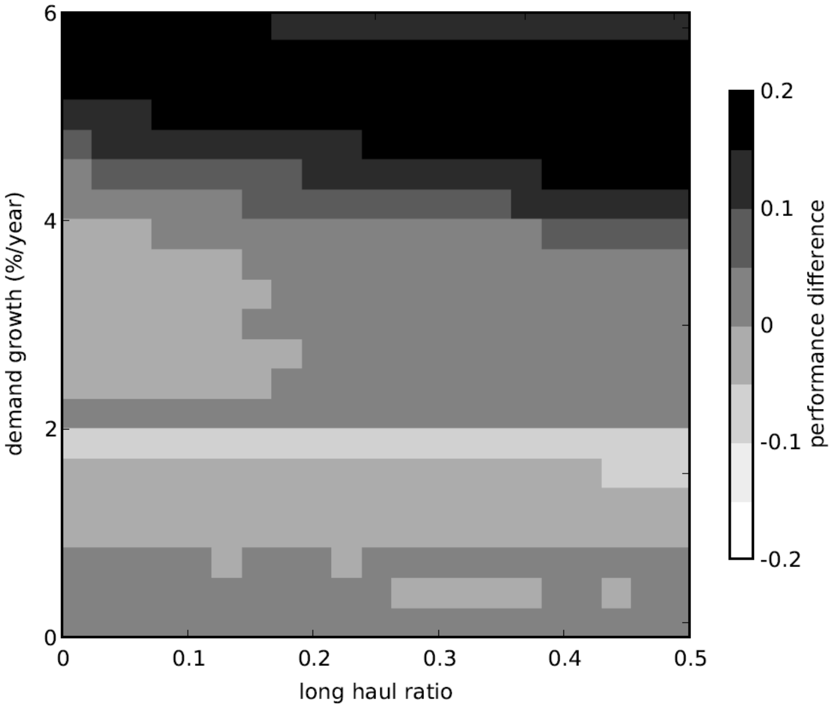


Figure 6-4: Performance difference of the DAP compared to the Master Plan for that combination of uncertain parameters that most favor the DAP

As a final step, summarizing the foregoing analyses, we generated Figure 6-5 and Figure 6-6. For these figures, the values of all the uncertain parameters were fixed to their values at the point that is the most positive for the Master Plan. Next, for Figure 6-5, which covers growing demand, we varied the demand growth parameter between 1 and 1.06 with a step size of 0.001. The two extreme values of the range correspond to a zero percent increase and a six percent increase in aviation demand per year, respectively. For Figure 6-6., which covers the case of declining demand, we varied the demand between 250,000 and 375,000 air transport movements in thirty years. The y-axis shows the performance difference, defined in the same way as in Figure 6-3 and Figure 6-4. Above 0 on the y-axis, the DAP performs better. Below 0 on the y-axis, the Master Plan performs better. As is clear from these figures, the Master Plan performs better in only a very narrow range of demand growth. Only if the growth in demand is between roughly 0% and 2% per year will the Master Plan perform better. Based on Figures 6.4-6.6 and Table 6-6), we conclude that given even medium levels of uncertainty, the DAP will perform better than the Master Plan. The results of these analyses thus corroborate the hypothesis that DAP is to be preferred over Master Planning.

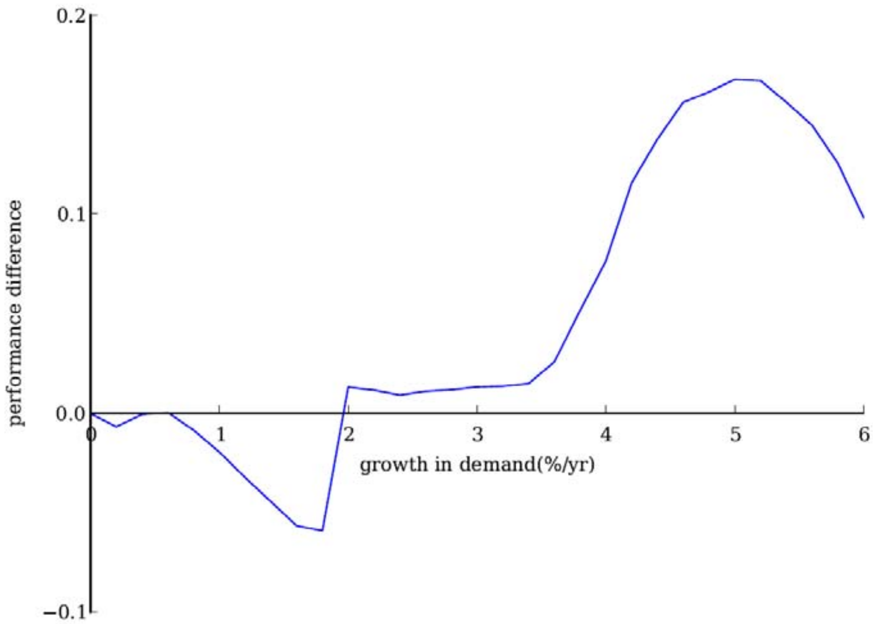


Figure 6-5: Performance difference of the DAP compared to the Master Plan for the situation most favorable to the Master Plan, with demand varied between 0% growth per year and 6% growth per year

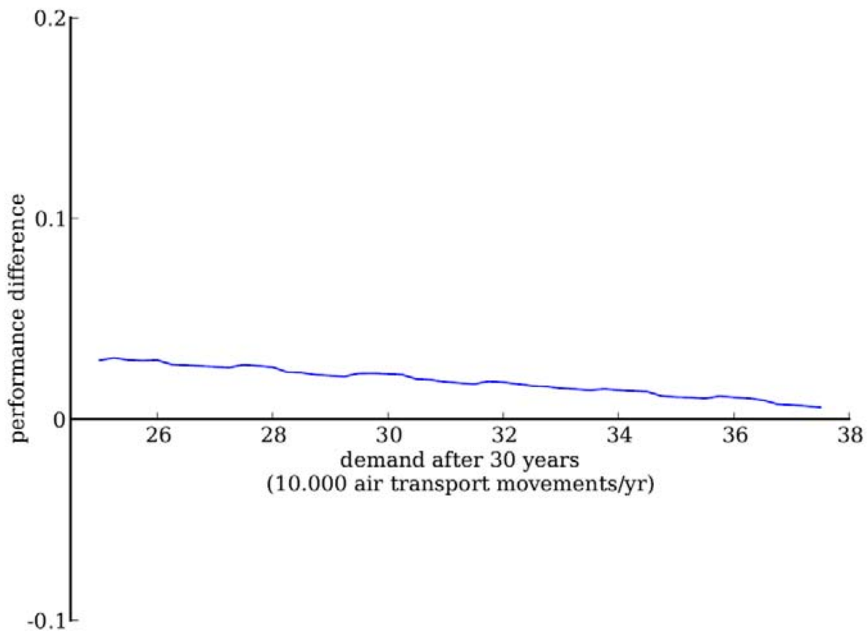


Figure 6-6: Performance difference of the DAP compared to the Master Plan for the situation most favorable to the Master Plan, with demand varied between 250,000 and 375,000 air transport movements in thirty years

6.5 Discussion of Results

From the foregoing, we conclude that, for this case, rigid static planning is preferable to dynamic adaptive planning only if the future demand were to fall into a narrow bandwidth. Conversely, if there is even small uncertainty about future demand, adaptive planning is to be preferred. Burghouwt (2007) has shown that aviation demand in Europe has become much more volatile since the introduction of privatization and liberalization. A similar argument can also be made for the United States. This implies that there is significant uncertainty about the size and composition of future demand. Hence, based on the computational experiments reported above, we conclude that, in privatized and liberalized aviation markets, adaptive planning is preferable to static planning.

This does not mean that airport planners should immediately switch to dynamic adaptive planning. First, our experiments represent an idealized case. We compare a static rigid plan with an adaptive plan, without any attention to subsequent actions to modify the static plan. In practice, frequently parts of a static plan are deferred or sped up, operational changes are made, and, if necessary, formal updates to the plan are submitted. Our experiments have not taken such *ad hoc* deviations from and (formal) updates to the rigid plan into account. It is expected that if such *ad hoc* modifications to the static plan are accounted for in the experiments, the difference in performance between the two plans would be smaller.

However, as argued, there is currently limited (computational) evidence for the efficacy of innovative infrastructure planning approaches. This paper provides such evidence for a theoretical case of a static plan and an adaptive plan. These results suggest that adaptive planning based on a clear vision for the future, a set of actions to be taken immediately, and a framework for guiding future actions, has significant merit when planning under uncertainty. However, further research is needed to bring the case closer to real world practice. One way of achieving this is by including the implementation process of the static plan in the model. One option would be to use some form of an agent based model to represent the various actors, rules, and institutions. Alternatively, or in addition, simulation gaming could be combined with computational experiments to reveal how the implementation process might evolve. These extensions are, in fact, implied by the medical analogy introduced in Section 6.2. As argued there, evidence for the efficacy of new infrastructure planning approaches comes from multiple sources, only one of which is computational experimentation.

A second reason for not immediately switching to adaptive airport planning is that we looked only at the case of Amsterdam Airport Schiphol. How generalizable are our findings? Schiphol is a relatively large airport in a privatized and liberalized market. Some of the uncertainties that the decisionmakers are facing are particular to that airport (e.g. wind conditions), but most others are faced by most large airports in the world (e.g. demand volatility, technological developments, and demographics). Given that the difference in airport performance is most sensitive to demand variations, we believe that the findings of this single case can be generalized to those airports that also operate in a privatized and liberalized environment.

Third, the analysis did not cover the costs of either version of the plan. For a more comprehensive comparison, these costs should be taken into account. Since the presented analysis was from the perspective of the different stakeholders involved in the governance of Schiphol's long-term development, however, costs are less of an issue than if the analysis were carried out from a business perspective. Costs in the presented case would be only one among the six performance indicators. Furthermore, in those situations in which the new runway was part of the Master Plan but was not necessary, the DAP would likely have saved expenses, while in those cases in which a trigger for building the runway would have been reached, the cost difference between the DAP and the Master Plan would have been small. These two considerations together imply that including costs in the analysis would have made the DAP look better in those situations in which the new runway would not be necessary, while in those situations in which it would be necessary, it would fail to differentiate between the plans. Adding cost to the analysis, therefore, would not alter the conclusions reached about the efficacy of AASP.

Finally, we hypothesize that our findings for the airport case apply as well to other cases of transport infrastructure planning:

- Transport infrastructure planning generally faces similar (deep) uncertainties about transport demand, technological developments, impacts on outcomes of interest, acceptance among stakeholders, etc.
- Forecast-based planning and/or static robust planning are currently the standard approaches for planning new transport infrastructures.
- As such, similar computational experiments are expected to show that a static rigid plan (e.g., for a new road, railway, or port) will outperform a dynamic adaptive plan in only a small part of the wide spectrum of possible futures.
- Hence, a dynamic adaptive infrastructure plan should expose planners to less uncertainty about its future performance than a static rigid plan, despite the wide variety of uncertainties that are present.

References

Agusdinata, D. B. (2008). *Exploratory Modeling and Analysis: A promising method to deal with deep uncertainty*. Ph.D. thesis, Delft University of Technology, Delft.

Albrechts, L. (2004). Strategic (spatial) planning reexamined. *Environment and Planning B: Planning and Design*, 31(5), 743-758.

Ascher, W. (1978). *Forecasting: An Appraisal for Policy Makers and Planners*. Baltimore: Johns Hopkins University Press.

Bankes, S. (1993). Exploratory Modeling for Policy Analysis. *Operations Research*, 4(3), 435-449.

Batty, M., & Torrens, P. M. (2005). Modelling and prediction in acomplex world. *Futures*, 37(7), 745-766.

Ben-Haim, Y. (2006). *Information-Gap Decision Theory: Decisions under Severe Uncertainty* (2 ed.). New York: Wiley.

Bradfield, R., Wright, G., Burt, G., Cairns, G., & van der Heijden, K. (2005). The origins and evolution of scenario techniques in long range business planning. *Futures*, 37(8), 795-812.

Burghouwt, G. (2007). *Airline Network Development in Europe and its Implications for Airport Planning*. Burlington, Vermont: Ashgate Publishing Company.

Cowell, P. G., Foot, P. B., Gerrard, R. J., Kent, D., Manson, S. M., & Rivoire, A. (2000). *A methodology for calculation individual risk due to aircraft accidents near airports*. London, Uk.: National Air Traffic Services Ltd.

Cowell, P. G., Gerrard, R., & Paterson, D. S. (1997). *A Crash Location Model for Use in the Vicinity of Airports*. London, UK: National Air Traffic Services Ltd.

CPB, KiM, NMP, & RP. (2007). *Advies over de lange termijnverkenning Schiphol [Advice concerning the long term exploration Schiphol]*. Retrieved from http://www.cpb.nl/nl/pub/cpbreeksen/notitie/28sep2007_2/notitie.pdf.

De Haan, A. R. C. (2007). *Aircraft Technology's Contribution to Sustainable Development*. Ph.D. thesis, Delft University of Technology, Delft.

de Jong, A. H., & Hilderdink, H. B. M. (2004). *Lange-Termijn Bevolkingsscenario's voor Nederland [Long-Term Population Scenarios for the Netherlands]*. Den Haag: CBS, RIVM, CPB, RPB, SCP, NiDi.

de Neufville, R. (2000). Dynamic Strategic Planning for Technology Policy. *International Journal of Technology Management,* 19(3-5), 225-245.

de Neufville, R. (2003). Real Options: Dealing With Uncertainty in Systems Planning and Design. *Integrated Assessment,* 4(1), 26-34.

de Neufville, R., & Odoni, A. (2003). *Airport Systems: Planning, Design, and Management*. New York: McGraw-Hill.

Dempsey, P. S., Goetz, A. R., & Szyliowicz, J. S. (1997). *Denver International Airport: Lessons Learned*. New York: McGraw-Hill.

Dessai, S., Hulme, M., & Lempert, R. (2009). Do We Need Better Predictions to Adapt to a Changing Climate? *EOS,* 90(13), 111-112.

Dewar, J. A. (2002). *Assumption-Based Planning: A Tool for Reducing Avoidable Surprises*. Cambridge: Cambridge University Press.

Dewar, J. A., Builder, C. H., Hix, W. M., & Levin, M. H. (1993). *Assumption-Based Planning: A Planning Tool for Very Uncertain Times* (Report MR-114-A). Santa Monica: RAND.

Dewar, J. A., & Wachs, M. (2006). Transportation Planning, Climate Change, and Decisionmaking Under Uncertainty *Proceedings of the Workshop Conference on Climate Change and the US Transportation*. Washington, DC: National Academies of Sciences and Engineering.

Erikson, E. A., & Weber, K. M. (2008). Adaptive Foresight: navigating the complex landscape of policy strategies. *Technological Forecasting and Social Change,* 75(4), 462-482.

FAA. (1981). *Upgraded FAA Airfield Capacity Model*. Washington D.C.: Federal Aviation Administration.

FAA. (2008). *FAA Aerospace Forecast Fiscal Years 2008-2025*: U.S. Department of Transportation, Federal Aviation Administration.

FAA. (2009). *Emissions and Dispersion Modeling System User's Manual*. Washington, DC: Federal Aviation Administration, Office of Envronment and Energy.

Flyvbjerg, B., Bruzelius, N., & Rothengatter, W. (2003). *Megaprojects and Risk: An Anatomy of Ambition*. Cambridge: Cambridge University Press.

Frey, D. D., & Dym, C. L. (2006). Validation of design methods: lessons from medicine. *Research in Engineering Design,* 17(1), 45-57.

Goodwin, P., & Wright, G. (2010). The limits of forecasting methods in anticipating rare events. *Technological Forecasting and Social Change,* 77, 355-368.

Hansman, R. J., Magee, C., De Neufville, R., Robins, R., & Roos, D. (2006). Research agenda for an integrated approach to infrastructure planning, design, and management. *International Journal of Critical Infrastructures,* 2(2-3), 146-159.

Heblij, S. J., & Wijnen, R. A. A. (2008). Development of a Runway Allocation Optimisation Model for Airport Strategic Planning. *Transportation Planning and Technology,* 31(2), 201--214.

Hodges, J. S. (1991). Six (or so) Things You Can Do With a Bad Model. *Operations Research,* 39(3), 355-365.

Hodges, J. S., & Dewar, J. A. (1992). *Is it You or Your Model Talking? A Framework for Model Validation* (R-4114-AF/A/OSD). Santa Monica: RAND.

IISD. (2006). *Designing Policies in a World of Uncertainty, Change and Surprise - Adaptive Policy-Making for Agriculture and Water Resources in the Face of Climate Change –*

Phase I Research Report. Winnipeg: International Institute for Sustainable Development.

Kahan, J. P., Botterman, M., Cave, J., Robinson, N., Shoob, R., Thomson, R., et al. (2004). *Cyber Trust and Crime Prevention: Gaining Insight from Three Different Futures*: Prepared for Foresight Directorate, Office of Science and Technology, UK.

KNMI. (2006). *Climate Change Scenarios 2006 for the Netherlands*. de Bild, the Netherland: KNMI.

Kristóf, T. (2006). Is it Possible to make Scientific Forecasts in Social Sciences. *Futures,* 28(5), 561-574.

Kwakkel, J. H., Walker, W. E., & Marchau, V. A. W. J. (2010). Adaptive Airport Strategic Planning. *European Journal of Transportation and Infrastructure Research,* 10(3), 227-250.

Kwakkel, J. H., Walker, W. E., & Wijnen, R. A. A. (2008, 6-10 July). *The Treatment of Uncertainty in Airport Strategic Planning: The Case of Schiphol Airport's Long-Term Vision.* Paper presented at the 12th Air Transport Research Society World Conference, Athens, Greece.

Kwakkel, J. H., Wijnen, R. A. A., Walker, W. E., & Marchau, V. A. W. J. (2009, 6-10 July). *A Fast and Simple Model for a Quick Scan of Airport Performance.* Paper presented at the 13th Air Transport Research Society World Conference, Abu Dhabi, United Arab Emirates.

Lempert, R. J. (2002). A New Decision Sciences for Complex Systems. *Proceedings of the National Academy of Sciences of the United States of America,* 99(3), 7309-7313.

Lempert, R. J., & Groves, D. G. (2010). Identifying and Evaluating Robust Adaptive Policy Responses to Climate Change for Water Management Agencies in the American West. *Technological Forecasting and Social Change,* 77(6), 960-974.

Lempert, R. J., Popper, S., & Bankes, S. (2003). *Shaping the Next One Hundred Years: New Methods for Quantitative, Long Term Policy Analysis* (Report MR-1626-RPC). Santa Monica: RAND.

Marchau, V. A. W. J., & Walker, W. E. (2003). Dealing with Uncertainty in Implementing Advanced Driver Assistance Systems: An Adaptive Approach. *Integrated Assessment,* 4(1), 35-45.

Marchau, V. A. W. J., Walker, W. E., & van Duin, R. (2009). An adaptive approach to implementing innovative urban transport solutions. *Transport Policy,* 15(6), 405-412.

McCray, L. E., Oye, K. A., & Petersen, A. C. (2010). Planned adaptation in risk regulation: an initial survey of US environmental, health, and safety regulation. *Technological Forecasting and Social Change,* 77(6), 951-959.

McDaniel, R. R., & Driebe, D. J. (Eds.). (2005). *Uncertainty and Suprise in Complex Systems: Questions on Working the Unexpected*: Springer.

Pahl-Wostl, C., Moltgen, J., Sendzimir, J., & Kabat, P. (2005). *New Methods for Adaptive Water Management under Uncertainty - the NeWater Project*. Paper presented at the EWRA conference.

Pahl-Wostl, C., Sendzimir, J., Jeffrey, P., Aerts, J., Berkamp, G., & Cross, K. (2007). Managing change towards adaptive water management through social learning. *Ecology and Society,* 12(2), 30 [online].

Popper, S., Griffin, J., Berrebi, C., Light, T., & Min, E. Y. (2009). *Natural Gas and Israel's Energy Future: A Strategic Analysis Under Conditions of Deep Uncertainty* (TR-747-YSNFF). Santa Monica, California: RAND.

Porter, A. L., Roper, A. T., Mason, T. W., Rossini, F. A., & Banks, J. (1991). *Forecasting and Management of Technology*. New York: John Wiley and Sons.

Quade, E. S. (1982). *Analysis for Public Decisions* (2nd ed.). New York: Elsevier Science Publishing Co.. Inc.

Rijksoverheid. (2009). *Luchtvaartnota, Concurrerende en duurzame luchtvaart voor een sterke economie [Air Transport White Paper Competitive and Sustainable Air Transport for a Strong Economy]*.

Schafer, A., Heywood, J. B., Jacoby, H. D., & Waitz, I. A. (2009). *Transportation in a Climate-Constrained World*. Cambridge, Massachusetts: MIT Press.

Schafer, A., & Victor, D. G. (2000). The future mobility of the world population. *Transportation Research Part A: Policy and Practice*, 34(3), 171-205.

Schiphol Group. (2007). *Lange Termijn Visie op de Ontwikkeling van de Mainport Schiphol: Een Wereldwijd Netwerk Voor een Concurrerende Randstad [Long-Term Vision on the Development of the Mainport Schiphol]*.

Schiphol Group. (2010). *Feiten & Cijfers 2009 (Facts & Figures 2009)*. Schiphol, Nederland.

Schiphol Group, & LVNL. (2007). *Verder Werken aan de toekomst van Schiphol en de regio [continu working on the future of Schiphol and the region]*.

Smith, J. B. (1997). Setting priorities for adapting to climate change. *Global Environmental Change,* 7(3), 261-266.

Stamatopoulos, M. A., Zografos, K. G., & Odoni, A. (2004). A Decision Support System for Airport Strategic Planning. *Transportation Research Part C: Emerging Technologies,* 12(2), 91-117.

Straatemeier, T., Bertolini, L., & Brömmelstroet, M. (2010). An experiential approach to research in planning. *Environment and Planning B: Planning and Design,* 37, 578-591.

Swanson, D., Barg, S., Tyler, S., Venema, H., Tomar, S., Bhadwal, S., et al. (2010). Seven tools for creating adaptive policies. *Technological Forecasting and Social Change,* 77(6), 924-939.

Taneja, P., Walker, W. E., Ligteringen, H., van Schuylenburg, M., & van der Plas, R. (2010). Implications of an Uncertain Future for Port Planning. *Maritime Policy & Management,* 37(3), 221-245.

Van Geenhuizen, M., Reggiani, A., & Rietveld, P. (2007). New Trends in Policymaking for Transport and Regional Network Integration. In M. Van Geenhuizen, A. Reggiani & P. Rietvelt (Eds.), *Policy Analysis of Transport Networks* (pp. 1-16). Alsdershot: Ashgate.

van Geenhuizen, M., & Thissen, W. A. H. (2007). A Framework for Identifying and Qualifying Uncertainty in Policy Making: The Case Of Intelligent Transport Systems. In M. van Geenhuizen, A. Reggiani & P. Rietvelt (Eds.), *Policy Analysis of Transport Networks* (pp. 101-118). Aldershot: Ashgate.

van Zuylen, H. J., & Weber, K. M. (2002). Strategies for European innovation policy in the transport field. *Technological Forecasting and Social Change*, 69(9), 929-951.

Varum, C. A., & Melo, C. (2010). Directions in scenario planning literature - A review of the past decades. *Futures,* 42(4), 355-369.

Verschuren, P., & Hartog, R. (2005). Evaluation in Design-Oriented Research. *Quality & Quantity,* 39, 733-762.

Volkery, A., & Ribeiro, T. (2009). Scenario planning in public policy: understanding use, impacts and the role of institutional context factors. *Technological Forecasting and Social Change,* 76(9), 1198-1207.

Walker, W. E., Marchau, V. A. W. J., & Swanson, D. (2010). Addressing deep uncertainty using adaptive policies: Introduction to section 2. *Technological Forecasting and Social Change,* 77(6), 917-923.

Walker, W. E., Rahman, S. A., & Cave, J. (2001). Adaptive Policies, Policy Analysis, and Policymaking. *European Journal of Operational Research,* 128(2), 282-289.

Walls, J. G., Widmeyer, G. R., & El Sawy, O. A. (1992). Building an Information Systems Design Theory for Vigilant EIS. *Information Systems Research,* 3(1), 35-59.

Walls, J. G., Widmeyer, G. R., & El Sawy, O. A. (2004). Assessing Information System Design Theory in Perspective: How Useful was our 1992 Initial Rendition? *Journal of Information Technology Theory and Application,* 6(2), 43-58.

Wardekker, J. A., de Jong, A., Knoop, J. M., & Van der Sluijs, J. P. (2010). Operationalising a Resilience Approach to Adapting an Urban Delta to Uncertain Climate Changes. *Technological Forecasting and Social Change,* 77(6), 987-998.

Wijnen, R. A. A., Visser, H. G., & Walker, W. E. (2009). Support for Airport Strategic Planning: Balancing an airport's economic and social performance. In M. E. Castille & P. B. Laurage (Eds.), *Airports: performance, risks and problems*. New York: Nova Science Publishers.

Wijnen, R. A. A., Walker, W. E., & Kwakkel, J. H. (2008). Decision Support for Airport Strategic Planning. *Transportation Planning and Technology,* 31(1), 11-34.

7 Conclusions and reflections

7.1 Overview of presented research

This thesis presented research on how the handling of uncertainties in Airport Strategic Planning (ASP) could be improved. First, a detailed analysis of the various meanings of uncertainty across the sciences was carried out. This revealed that there is no well-established shared meaning even in the various scientific disciplines that are involved in informing decisionmakers. Therefore, there was a need to specify more precisely what was meant with uncertainty in this thesis. In order to do so, we reviewed the literature that builds on the W&H framework (Walker, et al., 2003) − a framework that aimed at providing a common language for uncertainty classification and analysis in model-based decision support. This review revealed various problems with the W&H framework and potential solutions to these problems. We integrated these into a single improved framework for uncertainty analysis and classification. In light of this theoretical understanding of what was meant with uncertainty, we turned to studying the role of uncertainties in ASP, how they are currently handled, what the weaknesses and problems were with this approach, and what kinds of ideas were available on how to overcome this. This resulted in an approach called Adaptive Airport Strategic Planning (AASP). This approach is a synthesis of the various solutions that were put forward in the (airport) planning literature to improve the treatment of uncertainty in ASP. As a final step, we wanted to provide evidence for the efficacy of AASP as compared to traditional AMP. However, there was no generally agreed upon methodology readily available for doing this. Therefore, we first developed such a methodology, based on insights from medicine and design methods evaluation, where research copes with the same problem of providing evidence for the efficacy of a new treatment − either a new treatment for an illness or sickness, or a new treatment for handling design problems. In light of this methodology, we provided computational evidence for the efficacy of AASP in comparison to AMP.

The aim of this chapter is to present the key findings of this thesis, to reflect on the research, and to outline a research agenda for adaptive policymaking. First, the key findings are presented as answers to the research questions. The reflection is split into a reflection on the individual papers and an overall reflection. The chapter ends with an outline of a research agenda.

7.2 Answering the research questions

7.2.1 What is meant by uncertainty?

Uncertainty is an ambiguous term. It has different meanings depending on the context in which it is used. In order to address this problem, we analyzed the language of scientists as found in the abstracts of papers when they use words such as uncertain, uncertainty, and uncertainties. The aim of this analysis was to reveal the various meanings that uncertainty has in the sciences and where these meanings are used. Before performing an in depth analysis, we first explored the data. One of the most striking features of the data is that words that are claimed to be synonymous with uncertainty, such as ambiguity, ignorance, doubt, and indeterminacy, are rarely used by scientists when they speak about uncertainty. Only the word risk is used with quite some frequency. This strongly suggests that these words, apart from the word risk, are in fact not synonymous in the language of scientists. The claim that e.g. ambiguity and uncertainty have roughly the same meaning cannot be confirmed as far as the language of the scientist goes.

The next step was to analyze the data using a statistical model called the mixtures of factor analyzers (MFA). This model expands traditional factor analysis by allowing multiple centroids. When the MFA model is applied to data about word usage, it identifies clusters that are internally coherent in their word use, and for each cluster it identifies the latent semantic factors − or meanings. In order to select the appropriate parameterization of the MFA model, we used a metric approach for model evaluation, in the form of Aikake's Information Criterion (AIC) (Akaike, 1974). The AIC provides confidence that the model that will be chosen, given these data, is robust. Based on the AIC, a three cluster−three factor solution was extracted.

A detailed analysis of the ISIS categories, an analysis of the leading documents as expressed by their factor scores and the identification of the heavily loading terms for each cluster and factor, resulted in Table 7-1. This table implies that there are three dominant scientific discourses in which uncertainty is used. Each discourse is characterized by three latent semantic dimensions. These latent dimensions specify the various meanings of uncertainty. So, for example, the model-based latent dimension has one particular way of using the term uncertainty; this usage is clearly distinct from the usage described by the risk assessment latent assessment.

The results of this analysis confirm that there is not a single definition of the term uncertainty across the sciences. That is, there are nine (3 clusters times 3 factors) answers to the question 'what is meant by uncertainty?' Moreover, and problematic for this research, the various disciplines that provide decision support use uncertainty in different ways. For example, the model-based meaning of uncertainty is quite different from the meaning in risk assessment or policy evaluation. For this research, this implied that there was a clear need to specify in more detail what was meant with uncertainty. This is pursed in answering the next research question.

Table 7-1: Contexts, latent semantics, and indicative words

Corpora	Latent Semantic Dimensions	Heavily loading terms
Natural Sciences	Engineered Systems	Function, method, system, control, proposed, systems problem, given, fuzzy, state, linear, measurement, standard
	Natural Systems	Species, uncertain, control, management, decision
	Models	Model, data, results, using, used, uncertainties, based models, methods, parameters, large, conditions, developed, distribution, estimates, error, surface, parameter, estimate, study, effects, important, significant, changes, response, change, time, uncertainty, analysis, different, use, water, number, process, simulation, set, approach, performance
Applied Sciences	Risk Assessment	Management, robust, design, assessment, risk
	Statistical Mechanics	Rate, energy, theory, set, systems, linear
	Policy Evaluation	Values, high, temperature, low, studies, treatment, effect, research, evidence, quality, levels, risk, potential, field, assessment, information, decision, measurements, measured
Medicine	Diagnosis	Similar, obtained, range, observed, total, cases, patients, compared, p, group
	Longitudinal research	Value, mean, years, factors, clinical, estimated, disease, associated
	(quasi) Natural experimentation	Case, order, experimental, mass

7.2.2 How can uncertainties be classified and analyzed for model-based decision support?

ASP relies heavily on model-based decision support that is offered by various disciplines such as economics, aerospace engineering, and even psychology. Since the answer to the first research question revealed that there is no shared meaning of uncertainty even in the various scientific fields that inform decisionmaking, it was necessary to specify more clearly what is meant with uncertainty in this research. This problem had already been acknowledged in the literature to a certain extent. It has been suggested by different researchers that policy advice should be accompanied by an uncertainty analysis that will clarify the quality and certainty of the conclusions based on the knowledge base underlying the policy advice. Among others, the W&H framework (Walker, et al., 2003) was proposed to provide a common conceptual basis for this.

We chose to use the W&H framework as a starting point for specifying how uncertainty can be classified and analyzed for model-based decision support. The W&H framework was deemed a good starting point, for it focused on model-based decision support. Moreover, the W&H framework brought together experts from various different scientific fields. As such, it was expected that the framework already tried to harmonize some of the disparities in use and meaning of the term uncertainty that were identified in answer to research question 1.

In order to assess whether the W&H framework itself could be used directly, we analyzed its evolution using citation analysis. From this, we learned that a shift towards the perception of uncertainty took place. In line with this shift to perceived uncertainties, the role of framing in relation to uncertainty has seen increasing attention, resulting in adding an additional type to the nature dimension in the form of ambiguity. Furthermore, we observed that a variety of researchers have struggled in communicating about uncertainty through the framework. In fact, Gillund et al. (2008) observed that it is ironic that a framework intended to facilitate communication requires quite some explanation and elaboration. Particularly problematic for various researchers was the way the W&H framework had conceptualized the level of uncertainty − it uses ambiguous methodological names.

In light of the results from the citation analysis, it was decided that the W&H framework in its original form could not be used directly. However, the analysis also resulted in the identification of various suggestions for improving the framework. We harmonized these suggestions in the form of a new framework that can be understood as being the next incarnation of the W&H framework. The resulting framework presents the conceptualization of uncertainty that has guided the remainder of the research.

According to resulting framework, uncertainty is a three dimensional concept:

- Location: where is the uncertainty located? Uncertainty can be located in any of the various aspects of a policy problem: (a) the future world, (b) the model of the relevant system for that future world, (c) the outcomes from the system, and (d) the weights that the various stakeholders place on the various outcomes of the system
- Nature: is the uncertainty due to a lack of knowledge, due to inherent variability in the real world system, or due to different ways of framing?
- Level: how severe is or what is the degree of the uncertainty? Broadly speaking, the level of uncertainty is the assignment of likelihood to things or events. So, uncertainty for model-based policy analysis means that there are multiple plausible alternative future worlds, system models, outcomes, and/or weights. In some cases the likelihood or plausibility of these alternatives can be expressed using numbers, but in other cases more imprecise labels are used, such as more likely, less likely, or equally likely. We define four levels of uncertainty: We speak of Level 1 uncertainty, or shallow uncertainty, when one is able to enumerate multiple alternatives and is able to provide probabilities (subjective or objective). We speak of Level 2 uncertainty, or medium uncertainty, when one is able to enumerate multiple alternatives and able to rank order the alternatives in terms of perceived likelihood. However, how much more likely or unlikely one alternative is compared to another cannot be specified. We speak of Level 3 uncertainty, or deep uncertainty, when one is able to enumerate multiple alternatives without being able or willing to rank order the alternatives in terms of how likely or plausible they are judged to be. Finally, we speak of Level 4 uncertainty, or recognized ignorance, when one is unable to enumerate multiple alternatives, while admitting the possibility of being surprised.

7.2.3 How can the uncertainties currently prevalent in airport strategic planning be treated better?

The central problem this research addresses is that AMP is inadequate for the long-term development of airports. The inadequacy of AMP is the results of the fact that AMP addresses only a subset of the uncertainties that are relevant. More specifically, AMP focuses mainly on demand uncertainties and treats these as Level 1 or Level 2 uncertainties. That is, probabilities and/or a few trend-based most likely scenarios are used to capture the uncertainties surrounding demand. In the airport planning literature three main alternatives have been put forward: Dynamic Strategic Planning, Adaptive Policymaking, and Flexible Strategic Planning. Of these three, the first two approaches have been suggested as generic approaches for handling uncertainties in long-term decisionmaking. From a comparison of the three approaches, it became clear that these approaches are to a large extent complementary. Therefore, it was decided that a synthesis of the three approaches that draws on their relative strengths should be pursued.

Figure 7-1 presents a synthesis of the three approaches, which we called Adaptive Airport Strategic Planning (AASP). AASP is the proposed approach for improving the treatment of uncertainties in ASP. In short, in Step I, the existing conditions of an airport are analyzed and the goals for future development are specified. In Step II, the way in which this is to be achieved is specified. This basic plan is made more robust through four types of actions specified in Step III: mitigating actions are actions to reduce the *certain* adverse effects of a plan; hedging actions are actions to spread or reduce the risk of *uncertain* adverse effects of a plan; seizing actions are actions taken to seize certain available opportunities; and shaping actions are actions taken to reduce the chance that an external condition or event that could make the plan fail will occur, or to increase the chance that an external condition or event that could make the plan succeed will occur. Even with the actions taken in Step III, there is still the need to monitor the performance of the plan and take action if necessary. We call this contingency planning following e.g. Burghouwt (2007). This is specified in Step IV. Signposts specify information that should be tracked in order to determine whether the plan is achieving its conditions for success. Critical values of signpost variables (triggers) are specified, beyond which actions should be implemented to ensure that the plan keeps moving the system in the right direction and at a proper speed. There are four different types of actions that can be triggered by a signpost: defensive actions are taken to clarify the basic plan, preserve its benefits, or meet outside challenges in response to specific triggers that leave the basic plan remains unchanged; corrective actions are adjustments to the basic plan; capitalizing actions are actions triggered to take advantage of opportunities that improve the performance of the basic plan; and a reassessment of the plan is initiated when the analysis and assumptions critical to the plan's success have clearly lost validity. Step V is the actual implementation. In this step, the actions to be taken immediately (from Step II and Step III) are implemented and a monitoring system (from Step IV) is established. Then time starts running, signpost information related to the triggers is collected, and actions are started, altered, stopped, or expanded in response to this. After implementation of the initial actions, the implementation of other actions is suspended until a trigger event occurs. The central claim of AASP is that a more successful coherent development of infrastructure can be achieved by planning for adaptation instead of having to make ad-hoc adaptations and modifications to existing plans For a more detailed explanation of this framework, see Kwakkel et al (2010), Marchau et al. (2009), and Walker et al. (2001).

In contrast to AMP, AASP covers the four levels of uncertainty as specified in response to the previous research question. Level 1 and 2 uncertainty are considered jointly in Steps I and II. These two steps are comparable with AMP. Step III focuses on increasing the robustness of the plan through static actions taken at the start of the implementation. This corresponds to the use of for example scenario planning, which can be used to handle Level III uncertainty about the future world. The fourth step in AASP focuses on contingency planning. It prepares the basic plan for surprises and builds in mechanisms to handle these surprises. This fits with flexible and adaptive approaches that have been advocated for handling Level 4 uncertainty. AASP thus is capable of handling all the levels of uncertainty.

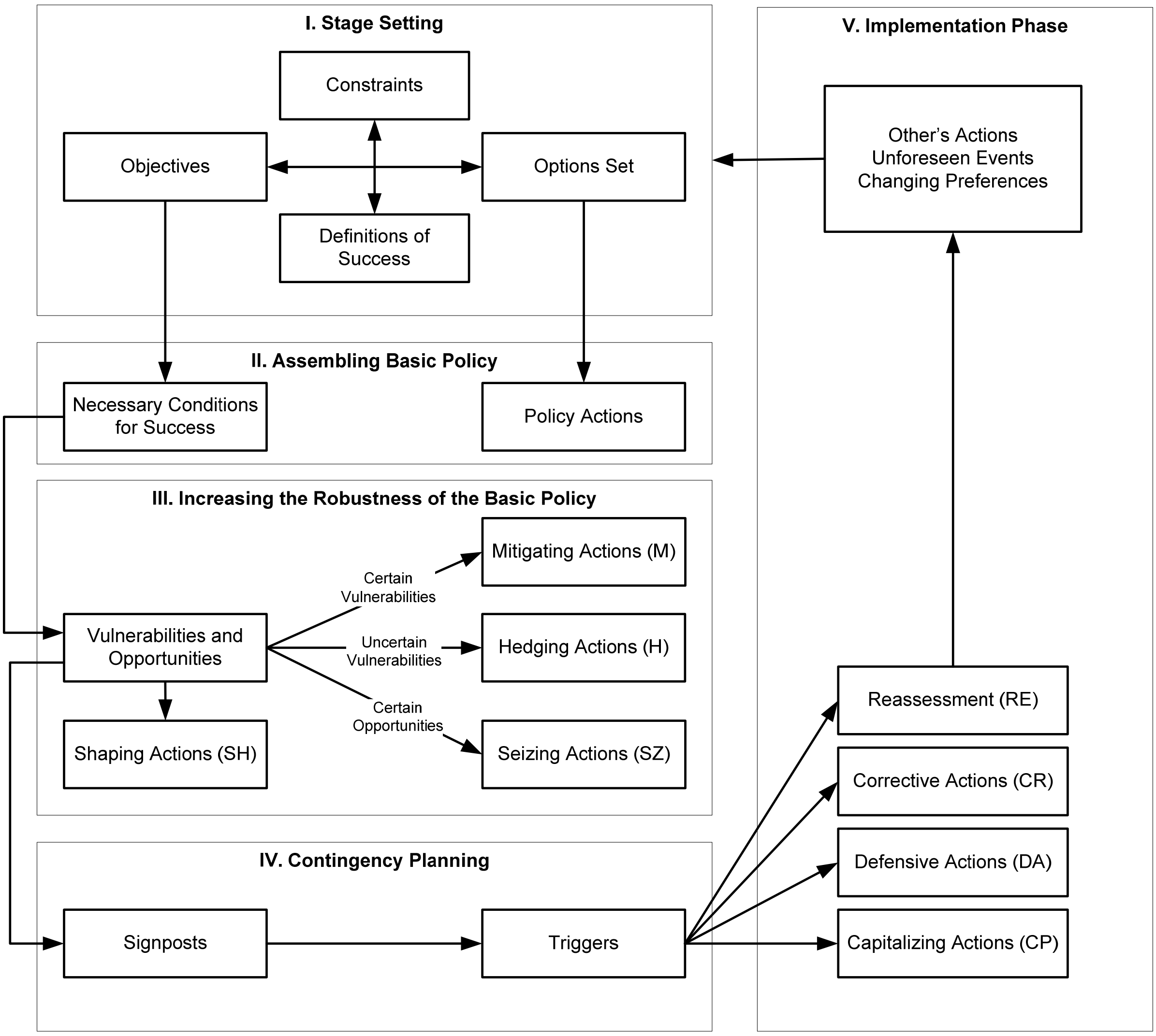


Figure 7-1: The steps of AASP

7.2.4 How can the efficacy of adaptive airport strategic planning be assessed?

Now that we have specified an alternative way of handling uncertainty in ASP, the next question is to test this alternative approach and compare its efficacy with AMP. However, there is currently no best practice for evaluating the efficacy of new planning approaches (Dewar, Builder, Hix, & Levin, 1993; Hansman, Magee, De Neufville, Robins, & Roos, 2006). Moreover, in establishing the efficacy of new infrastructure planning approaches one faces a methodological problem for "nothing done in the short term can 'prove' the efficacy of a planning methodology; nor can the monitoring, over time, of a single instance of a plan generated by that methodology, unless there is a competing parallel plan" (Dewar, et al., 1993).

In order to assess the efficacy of AASP, a methodology is needed. In order to arrive at a methodology we drew an analogy with medicine. Medicine has a well-established approach for gathering evidence about the efficacy of new treatments. Evidence for the efficacy of medical treatments is derived from five sources. *Theoretical developments* in disciplines such as biology, biochemistry, and the life sciences suggest new treatments. *Animal models* are used as a surrogate for the human body in early tests of new treatments. Human tissue is used *in vitro* to assess the efficacy of the treatment, without risking harm to human subjects.

Natural experiments are utilized when controlled experimentation is not possible (e.g. to study to long-term health effects of a treatment) (Glass, 1976; Hill, 1966). *Clinical trials* are the final test for new treatments and are subject to strict rules to minimize various threats to validity (European Medicines Agency, 1998; U.S. Department of Healt and Human Services - Food and Drug Administration, 1998). Frey and Dym (2006) argue that this approach can be used to inform testing design approaches. More specifically, they suggest that these same five sources of evidence can also be used to provide evidence for the efficacy of design approaches. We argue that this same analogy can also be adopted for assessing the efficacy of AASP.

7.2.5 Is the proposed approach for the treatment of uncertainty in airport strategic planning indeed preferable to Airport Master Planning?

In order to assess the efficacy of AASP, an ensemble of simulation models was used in series of computational experiments. This use of simulation models as a stand in for the real system corresponds to the use of animal models in medicine. Animal models in medical research and development only bear a partial resemblance to the real, human, subject. As such they are comparable to simulation models frequently employed in engineering. By comparing the performance of an Adaptive Plan and a Master Plan as calculated by the computer models, one can reason about how these plans would behave in the real world. After this, the next steps for assessing the efficacy of AASP would be to perform in vitro experiments, identify and analyze natural experiments, and set up pilot projects.

The basic approach for testing AASP was: (i) develop a fast and simple model of the system of interest; (ii) generate an ensemble of future worlds; (iii) specify a Master Plan and an Adaptive Plan; and (iv) calculate and compare the performance of both plans across an ensemble of future worlds using the fast and simple model. For the fast and simple model, several tools from the Federal Aviation Administration (the US regulator for air transport) for calculating airport performance were used to calculate airport performance in terms of noise, emissions, external safety, and capacity The fast and simple model for calculating airport performance was complemented with a set of components that represent different exogenous developments, such as demand, technological development, and demographics. This set of components specifies the ensemble of future worlds. The two plans that were used were based on the available policy documents discussing the long-term development of Amsterdam Airport Schiphol (e.g. CPB, KiM, NMP, & RP, 2007; Provincie Noord-Holland, 2007; Schiphol Group, 2007; Schiphol Group & LVNL, 2007; To70, 2007; V&W, 2007; V&W & VROM, 2007). For the Master Plan, a static version of the current proposal was used. For the Adaptive Plan, various ideas put forward in these policy documents were utilized using the AASP framework. In practice, airport planners try and cope with uncertainty through various *ad hoc* modifications to the Master Plan, these types of adaptations of the Master Plan were not taken into account. So, our experiments presented the best case for the alternative approaches to AMP.

There are two specific questions we tried to answer using EMA. First, given the ensemble of future worlds, what is the range of outcomes that both the Master Plan and the Adaptive Plan can generate? Second, given the most favorable conditions for a Master Plan, how well does the Adaptive Plan do in comparison? The first question reduces to solving a boundary problem. For each outcome of interest, the upper and lower bound across the ensemble of future worlds was determined, using a constrained non-linear optimization algorithm. The results of the overall analysis showed that an Adaptive Plan has a smaller bound on all but one outcomes of interest. The one outcome of interest that has a larger range is latent demand, which is explained by the fact that under extreme growth of demand, the Adaptive Plan is more aggressive in restricting the demand that can be accommodated by the airport.

In order to answer the second question, we need to identify the difference in performance between an Adaptive Plan and a Master Plan. We identified the future world under which the Master Plan performs the best compared with the Adaptive Plan and fixed all the parameters describing this future world apart from demand growth to their values in this world. The choice for demand growth was determined by the observation that the main uncertainties in airport long-term planning are about the size and composition of future demand (Burghouwt, 2007; de Neufville & Odoni, 2003). Next, we varied the demand growth between zero percent per year and six percent per year. From this analysis, it is concluded that the Master Plan performs better than the Adaptive Plan in only a very narrow range of demand growth. Only if the growth in demand is between roughly 0% and 2% per year will the Master Plan perform better. This is explained by the fact that the Master Plan is optimized for this demand development.

From the foregoing, we concluded that, for this theoretical case, Master Planning is preferable to Adaptive Planning only if the future demand were to fall into a narrow bandwidth. Conversely, if there is even small uncertainty about future demand, the Adaptive Plan is to be preferred. Burghouwt (2007) has shown that aviation demand in Europe has become much more volatile since the introduction of privatization and liberalization. A similar argument can also be made for the United States. This implies that there is significant uncertainty about the size and composition of future demand. Hence, based on the computational experiments reported above, we conclude that, in privatized and liberalized aviation markets, AASP would be the preferred planning methodology. If we not only consider demand uncertainty, but also other uncertainties such as demographics and technological developments, the advantages of AASP over AMP increase even more, as can be deduced from the fact that AASP by and large has a smaller bandwidth on the various outcomes of interest.

7.3 Reflection

This section presents a reflection on the research. First, a detailed reflection on the separate papers is presented. Next, a reflection on the overall research and how the various papers add up to a single story is presented. This section closes with a research agenda for adaptive policymaking.

7.3.1 Reflections on the individual papers

− KWAKKEL, J.H. & CUNNINGHAM, S.W. (2009) Managing Polysemy and Synonymy in Science Mapping using the Mixtures of Factor Analyzers Model. *Journal of the American Society for Information Science and Technology*, 60(10), 2064-2078.

The aim of the research underlying this paper was to reveal the various meanings of uncertainty across the sciences and identify in which disciplines these meanings are used. However, the available techniques for mapping the latent semantics were inadequate for achieving this goal. Hence, the emphasis of the research shifted to first developing a method that could do so. The resulting paper reflects this methodological shift. The identification of the various meanings of uncertainty is in the paper delegated to the case study that is used as a proof of concept of the proposed method. Still, this case study does provide great insight, albeit at a high level, into the nine dominant ways in which uncertainty is discussed in the sciences. Retrospectively, it is not surprising that the mapping of the literature is rather generic. The data were sampled from across all the sciences indexed in ISI. A more focused sampling of literature could have provided a more focused picture. However, how to design a query that gives this focus? There is an inferential circle from prior ideas about the data, to the analysis of the data, in light of which the prior ideas about the data have to be updated. A more focused query limits the potential for this inferential circle. From the perspective of the overall research, however, a more thorough exploration of the results is desirable. That is, for each of the nine ways, explore and describe how uncertainty is discussed, identify exemplars, and elaborate on how uncertainty is conceptualized. Such an exploration can enhance the high-level picture that the current mapping paints. It can also enhance the validity of the result. Some initial steps in this direction have already been taken (Kwakkel & Cunningham, 2008a, 2008b). In addition to exploring the case results in more depth, expanding the data to cover multiple years and to cover data from other sources is important. Adding more years would allow the study of the dynamics over time of the various meanings of uncertainty. Adding alternative data sources would enhance the richness of the mapping. Specifically, if non-scientific data sources were added it would be possible to compare the scientific usages of the term uncertainty with the non-scientific usages, revealing potential incommensurabilities and possible sources for misunderstanding.

− KWAKKEL, J.H., WALKER, W.E. & MARCHAU, V.A.W.J. (2010) Classifying and Communicating Uncertainties in Model-Based Policy Analysis. *International Journal of Technology, Policy and Management*, 10(4), 299-315.

This paper aims to synthesize the wide array of existing conceptual frameworks, taxonomies, etc. that are available in the model-based policy analysis literature on classifying and communicating uncertainties. As a starting point, it used the W&H framework and the literature that cites this paper. It has been claimed that such a framework that integrates and tries to harmonize alternative conceptualizations results in a loss of insight (Meijer, 2007). In opposition, we contend that attempts to integrate various conceptualizations into a single coherent overarching framework in fact bring out the richness. Attempts to harmonize frameworks reveal both how the frameworks differ and how they correspond, thus creating new insights. Moreover, conceptual frameworks are not static. They are put forward, utilized, modified, and expanded throughout time. A prime example of how attempts to harmonize and integrate have increased richness is presented by physics. Before the advent of the modern time with Galileo, Kepler, and Newton, it was believed that the rules that govern the stars and the rules that govern what goes on underneath the moon have nothing in common. The successful attempts by Newton, Kepler, and Galileo to bring ever more phenomena into a common explanatory framework have propelled physics into ever more esoteric, ever richer explorations into the inner workings of the universe. It can only be hoped that the framework outlined in this paper will be used, expanded, modified, and eventually discarded in favor of a more comprehensive framework that provides even richer insights into the various aspects of uncertainty. That such a more comprehensive framework is desirable is evidenced by the fact that the current framework only captures uncertainty in model-based policy analysis. However, there are also other styles of policy analysis in which uncertainty plays its role. Furthermore, the current framework addresses only the policy analyst's perspectives on uncertainty; a more comprehensive framework should also address the decisionmaker's perspective.

− KWAKKEL, J.H., WALKER, W.E. & MARCHAU, V.A.W.J. (2010) Adaptive Airport Strategic Planning. *European Journal of Transportation and Infrastructure Research*, 10(3), 227-250.

The aim of this paper was to outline how uncertainty could be treated better in ASP. The resulting framework expands the existing adaptive policymaking framework in several ways. Most importantly, it allows for the consideration of the available opportunities explicitly.

Before this paper, the relevance of opportunities was acknowledged in the adaptive policymaking literature. However, it was never made explicit how opportunities could be fit into the existing framework. This has now been rectified. In addition, the resulting framework integrates similar efforts found in the (airport) planning literature and thus harmonizes the literature on handling uncertainty in ASP. This paper limited itself to the literature on adaptiveness and flexibility that can be found in the airport planning literature. Two of the three approaches on which it builds are, however, more generic approaches. Retrospectively, the resulting framework might have benefited from expanding the scope of literature that was taken into account. Related, the paper presents the framework as being focused solely on ASP. An alternative presentation would have been to present the framework as a generic approach for policymaking under uncertainty, and applying it to a specific domain, namely ASP. Furthermore, in light of more recent insights, one thing that would have been good to stress more in the paper is that adaptation according to AASP is a dynamic process. In current debates about climate change, the notion of adaptation is often used in a more restrictive static manner. That is, how should we reorganize (aspects) of a system once so that it can cope with future climate change and its consequences. This can be called static adaptation. In contrast, AASP is a form of dynamic adaptation. Adding the word 'dynamic' here highlights the fact that the adaptation is over time, in response to new information gathered through the monitoring system. Therefore, we now prefer to speak of dynamic adaptation, dynamic adaptive policymaking, and dynamic adaptive airport strategic planning.

- KWAKKEL, J.H., PAS, J.W.G.M.V.D. (under review) Evaluation of Infrastructure Planning Approaches: An Analogy with Medicine.

When the research resulting in this thesis started, one of the first questions that was put forward was how one could know whether one way of handling uncertainty was indeed preferable over another way of handling uncertainty. Not having a direct answer, and literature and discussions with experts conforming this to be a troublesome question, it was turned into a research question. The answer to the question was found when someone suggested to take a look at Frey and Dym (2006). Their motivation for using a medical analogy for informing the literature on assessing the efficacy of new design approaches was very comparable to the problem encountered in assessing the efficacy of new approaches for handling uncertainty. By coupling this analogy with the available literature on EMA, a solution for this research could be crafted. Retrospectively, the answer to the questions might appear trivial. The problem being we focused too much on the point raised by Dewar et al. (1993) that "nothing done in the short term can 'prove' the efficacy of a planning methodology; nor can the monitoring, over time, of a single instance of a plan generated by that methodology, unless there is a competing parallel plan". This point however focuses solely on real world application. The realization that there are various sources of evidence available at different levels of generality resolved the problems. Moreover, this insight paved the way for the fifth paper, where we used a series of computational experiments as a source of evidence for the efficacy of AASP in comparison to AMP.

- KWAKKEL, J.H., WALKER, W.E. & MARCHAU, V.A.W.J. (under review) Assessing the Efficacy of Adaptive Airport Strategic Planning: Results from Computational Experiments.

This paper provides evidence from series of computational experiments for the efficacy of AASP. This evidence clearly shows that that AMP is preferable to AASP only for a tiny range of the uncertainties. Outside this tiny range, AASP is more preferable. This evidence is important and can be useful in convincing real world decisionmakers to seriously consider using AASP. However, it is only one source of evidence. The evidence focuses only on comparing the plans as written, and uncertainties related to the decisionmaking process, while the real world messiness of the implementation phase are completely ignored. That is, in the real world, if the conditions change, deviations from the Master Plan can and do occur; resulting in frequently large discrepancies between the vision of the airport as outlined by the Master Plan and that which is actually implemented. These ad-hoc unplanned deviations from the Master Plan are not considered by the experiments, biasing the results against Master Planning. However, this merely implies that more research on this is needed. Moreover as specified by the medical analogy, other sources of evidence are available in addition to computational experiments. These need to be pursued. Furthermore, research into how to develop adaptive policies is needed. Currently, the adaptive policies that are used as illustrations in most of the literature are developed in a non-systematic manner from the perspective of the authors. Research into structural ways of developing adaptive policies and what techniques and methods can be employed is needed, for this can turn the current outline of principles into a clear structured process. Finally, research into how the dynamics in the decisionmaking process can change as result of implementing dynamic adaptive policies is needed. The success or failure of dynamic adaptive policies depends critically on the monitoring system and the pre-specified triggers and actions. How to institutionalize this monitoring system is a key problem. In Section 7.4 we discuss the main directions for further research into dynamic adaptive policies in more detail.

7.3.2 Reflections on the overall thesis

The starting point of this thesis was the observation that Airport Master Planning is inadequate for the long-term development of airports, primarily because of how it handles uncertainty. For a while, this has been recognized in the literature, and an alternative to AMP is called for (Burghouwt, 2007; de Neufville, 2000). In light of this, our research goals were to design an alternative approach for ASP that overcomes the weaknesses of AMP, test the performance of this new approach, and compare its performance to AMP. The research presented in this thesis has done just that: a synthesis of the ideas in the airport planning literature on how uncertainty could be handled better in ASP has been put forward and we have provided computational evidence for its efficacy. This evidence strongly suggests that AASP is preferable over AMP even under conditions of medium uncertainty.

What are the implications of this research for infrastructure planning in general? Although this thesis focused on airports, uncertainty is a problem also for long-term planning on other infrastructures. Adaptive policymaking has been suggested as a solution for handling uncertainty in long-term infrastructure planning in general. AASP is an improvement on adaptive policymaking. Therefore, we are convinced that AASP can also be applied to long-term planning of other types of infrastructures. As a matter of fact, it has already informed research into handling uncertainty in the long-term development of seaports (Taneja, Walker, Ligteringen, van Schuylenburg, & van der Plas, 2010). Since seaports and airports are nodes in an infrastructure network, this suggests that AASP is applicable to nodes in infrastructure systems. Adaptive policymaking in general has been applied to other infrastructure planning cases, such as the implementation of road pricing (Marchau, Walker, & van Wee, 2010) and intelligent speed adaptation (Agusdinata, Marchau, & Walker, 2007; Marchau, et al., 2009), the development of a maglev based rail transport system(Marchau, et al., 2010), and in relation to innovative urban infrastructure systems (Marchau, et al., 2009). Admittedly, how the underlying system needs to be engineered for it to support AASP will be different from one transport mode to another. In case of an airport, land use reservations and modular terminals are relatively straightforward ways of engineering the flexibility required for AASP. In case of e.g. a seaport, the technical ways are quite likely to be different.

What are the implications of this research for policy analysis? Here the answer is similar to the previous. AASP, although presented in this research as a specific approach for handling uncertainties in ASP, can be utilized also in other non-infrastructure related domains. AASP in essence is a next version of the generic adaptive policymaking framework (Walker, et al., 2001). As such, the scope for application is not different from the scope of adaptive policymaking. The AASP framework, like APM, can be used for any long-term policy making problem. This also implies that the accumulated evidence for the efficacy of AASP can be understood as being case specific evidence for the efficacy of adaptive policymaking in general. Of course the uncertainties will differ from case to case and the technical way in which adaptivity is created will differ from domain to domain, and the efficacy of an Adaptive Plan needs to be assessed in light of this, but still this one case does provide computational evidence for the efficacy of adaptive policymaking.

What are the implications of this research for policymaking and governance under conditions of uncertainty? By and large, if decisionmakers face deep uncertainties, they should strongly consider adopting dynamic adaptation as a key design principle for new policy. The current approaches that are used for policymaking under these conditions have revealed an inability to deal with the multiplicity of plausible futures. Furthermore, the increasing connectedness and interactions among different societal subsystems at ever larger distances enhances the problem of surprises, or 'black swans'. That is, the decisionmakers have to recognize that such surprises are liable to happen. Dynamic adaptive policymaking is a most promising recourse for making decisions, developing policies, and societal governance in response to these deep uncertainties and potential surprises.

One aspect of the reported research that can be considered to be problematic is the apparent disjoint between papers 1 and 2, and papers 3-5. The first two papers focus on uncertainty and together paint a rich picture of the meaning of uncertainty across the sciences in general and for model-based decision support specifically. The last three papers focus on airport planning, how uncertainty should be handled there, and whether this new way of handling it actually works. This begs the question what the missing link is. In the original proposal, the idea was to develop a taxonomy for uncertainty classification and communication, apply this to airport planning to identify the key uncertainties, and then propose a way of handling these better. A similar line of reasoning, utilizing a modified version of the W&H framework in order to argue for adaptiveness, can also be found in integrated river basin management (Pahl-Wostl, et al., 2007). This middle part of this line of reasoning was carried out in a conference paper, but was deemed to be methodologically not adequate enough to be submitted to a journal (Kwakkel, Walker, & Wijnen, 2008). Although this paper was not submitted to a journal, it did inform papers 3-5, for it set up the Schiphol case study. The identified uncertainties were the starting point for the adaptive policy presented in paper 3. This adaptive policy is also utilized in paper 5. Furthermore, over the course of the research, it became clear that adaptive approaches for handling uncertainty in policymaking make broad generic claims. They are not, and need not be tailored to specific uncertainties. Therefore, a detailed analysis of the key uncertainties in ASP became less urgent.

7.4 A research agenda for Dynamic Adaptive Policymaking

This thesis has been about adaptive policymaking with a specific focus on how it can be used for ASP to overcome the problems encountered there due to uncertainty. Over the course of this thesis, the existing adaptive policymaking framework has been expanded to consider opportunities in addition to vulnerabilities, and the link between the various levels of uncertainty and how these levels are handled by the various steps of adaptive policymaking has been brought out more clearly. Furthermore, this thesis provides computational evidence for the efficacy of adaptive policymaking in the context of ASP. There are several key research challenges for adaptive policymaking that need to be addressed next.

In this thesis, and in the available adaptive policymaking literature, adaptive policies are presented primarily for clarification. These adaptive policies are developed in an ad hoc manner and represent the expert opinion of the authors of the papers. For adaptive policymaking to become a useful planning approach, the first research challenge is to specify in more depth how the various steps can be carried out and which methods and techniques can be employed in each of the steps. That is, adaptive policymaking needs to move from being a high level concept captured in a flowchart, to being a detailed planning approach. It is believed that many of the available traditional tools, such as forecasting, scenarios, and exploratory modeling, can be of great use in the various steps of adaptive policymaking. However, exactly how these tools can be employed for the purpose of developing an adaptive policy needs to be researched.

The second research challenge focuses on the institutional and decisionmaking implications of adaptive policymaking. The latter focuses on the problem of how to agree on an adaptive policy. On the one hand, adaptive policymaking may facilitate agreement on a policy since it reduces uncertainties about future outcomes. On the other hand, at the time of policy formulation, it is uncertain which parts of the adaptive policy will be implemented. So, adaptive policies reduce outcome uncertainties, but create new uncertainties about what will be implemented. Moreover, given that the policy is not implemented at the start in its entirety, there is the possibility of renegotiation in the future. To overcome this uncertainty, the institutional arrangements and safeguards with respect to the monitoring system and the triggering of future actions is a key research challenge. Only if there are institutional arrangements in place that safeguard the execution of triggers, can this problem be overcome. A related institutional problem is that it has been argued that most institutions, particularly governmental ones, have a limited ability to cope with deep uncertainty. A simple example of this is the often heard complaint that governments do not look beyond the next election. Yet another institutional problem is that the implementation of adaptation can be hindered by several institutional and social complexities, such as too many policy domains, too many administrative levels, too fragmented and rigid regulation and budgets, too detailed planning and budget allocation procedures, lack of awareness, insufficient learning capacity of key players, etc. According to Lempert & Light "Today's policymakers generally lack the tools and the institutions that can … identify priority long-term decisions". For the successful application of adaptive policymaking, the institutional arrangements that are necessary for Adaptive Plans to be agreed upon and implemented are a key research topic.

A third research challenge is assessing the efficacy of adaptive policymaking in comparison to traditional policymaking. This thesis has provided computational evidence from a single case for its efficacy. More research on this is needed. The presented medical analogy can guide this effort. Most importantly, next to computational evidence, additional sources of evidence, such as in vitro experimentation using simulation gaming, need to be exploited. Most importantly, addressing this research challenge will help in improving the validity of adaptive policymaking. But, as a beneficial spin-off, the accumulated evidence can also be used to convince decisionmakers and stakeholders that they should adopt adaptive policymaking as their main way of doing long-term planning.

A final research challenge is the assessment of the costs and benefits of adaptive policies in comparison to traditional policies. Cost-benefit analysis is used for many decisionmaking problems to support decisionmaking. Adaptive policies can be more costly than traditional policies, since the monitoring system and the required flexibility inside the system on which the policy act create extra costs. In the context of ASP, the ability to build or not build a runway, or the ability to quickly expand a terminal, require changes to the airport. For the runway, land-use reservations are needed. For the terminal, some form of flexible design using real options is required. The additional apparent costs of adaptive policies need to be offset against less uncertainty about the benefits. An in-depth analysis of the costs and benefits of adaptive policies as compared to traditional static policies is required.

How can these four research challenges be addressed? With respect to the first challenge, the first step is to make an inventory of methods and techniques that are used in current planning practice. Textbooks on policy analysis, systems analysis, (technology) forecasting, and (transport) planning are an ideal starting point for such an inventory. Next, the characteristics of the various tools and methods in the inventory need to be specified. The relevant characteristics can be deduced from the adaptive policymaking framework: each step in this framework requires certain kinds of knowledge, and each tool provides certain kinds of knowledge. The resulting mapping can reveal which tool can employed where in the adaptive policymaking process. The second research challenge focuses on the institutional implications of adaptive policymaking. Addressing this research challenge could start from the literature on institutional design (e.g. Koppenjan & Groenewegen, 2005; Ostrom, 1990; Williamson, 1997). This information might be combined with cooperative game theory and potentially some form of agent based modelling. With respect to the third challenge (the efficacy of adaptive policymaking), the answer is provided in Chapter 5. evidence can come from various sources, and research on adaptive policymaking should utilize these various sources. To increase the empirical embedding of adaptive policymaking, a focus on quasi-natural case studies is advised. The fourth research challenge focuses on assessing the costs and benefits of adaptive policies compared to traditional policies. A combination of Exploratory Modeling and Analysis and techniques currently employed to value Real Options executing in a way similar to the case in Chapter 7 is my suggested approach for addressing this challenge. These four research challenges are not of equal importance. Research challenges 1 appear to me as the most important. Significant questions still exist with respect to the institutional implications of adaptive policymaking and the available literature has largely ignored this issue. Also, the research on adaptive policymaking needs to move beyond ad-hoc cases for illustrating adaptive policies and should start to address the questions related to how to actually develop adaptive policies.

Apart from these various research challenges, another avenue of work for researchers working on dynamic adaptation is the communication of their research findings to a wider, societal audience. Recent policy discussions, such as the health care debate in the United States, how to address the financial crisis, and the ongoing debate about how to address climate change, highlight the problems society at large faces due to uncertainty, while also revealing the need for innovative ways of handling these. In the health care debate in the United States, the idea of a trigger was put forward. The financial crisis brought the problem of uncertainty, and the vulnerability of society due to this, clearly to the fore, while the ongoing struggle to address climate change shows how uncertainty can create deadlocks in decisionmaking. Dynamic adaptation points a way forward that does not rely on reducing uncertainties such as those surrounding climate change. It can potentially help to prevent another major financial crisis by explicitly considering ways in which a policy can fail. It could have helped in crafting a more solid version of the health care trigger idea that might have made it into law. Regardless, these three examples reveal the potential benefits to society from dynamic adaptation research. Communicating to society on such issues and suggesting how dynamic adaptation can be of benefit is an important area of work for researchers working on dynamic adaptation.

References

Agusdinata, D. B., Marchau, V. A. W. J., & Walker, W. E. (2007). Adaptive Policy Approach to Implementing Intelligent Speed Adaptation. *IET Intelligent Transport Systems,* 1(3), 186-198.

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transaction on Automatic Control,* 19(6), 716-723.

Burghouwt, G. (2007). *Airline Network Development in Europe and its Implications for Airport Planning*. Burlington, Vermont: Ashgate Publishing Company.

CPB, KiM, NMP, & RP. (2007). *Advies over de lange termijnverkenning Schiphol [Advice concerning the long term exploration Schiphol]*. Retrieved from http://www.cpb.nl/nl/pub/cpbreeksen/notitie/28sep2007_2/notitie.pdf.

de Neufville, R. (2000). Dynamic Strategic Planning for Technology Policy. *International Journal of Technology Management,* 19(3-5), 225-245.

de Neufville, R., & Odoni, A. (2003). *Airport Systems: Planning, Design, and Management*. New York: McGraw-Hill.

Dewar, J. A., Builder, C. H., Hix, W. M., & Levin, M. H. (1993). *Assumption-Based Planning: A Planning Tool for Very Uncertain Times* (Report MR-114-A). Santa Monica: RAND.

European Medicines Agency. (1998). *Note for Guidance on Statistical Prinicples for Clinical Trials*: European Medicines Agency.

Frey, D. D., & Dym, C. L. (2006). Validation of design methods: lessons from medicine. *Research in Engineering Design,* 17(1), 45-57.

Gillund, F., Kjølber, K., Krayer von Krauss, M. P., & Myhr, A. I. (2008). Do Uncertainty Analyses Reveal Uncertainties? Using the Introduction of DNA Vaccines to Aquaculture as a Case. *Science of the Total Environment,* 407, 185-196.

Glass, G. V. (1976). Primary, Secundary and meta-analysis of research. *Educational Researcher,* 5(10), 3-8.

Hansman, R. J., Magee, C., De Neufville, R., Robins, R., & Roos, D. (2006). Research agenda for an integrated approach to infrastructure planning, design, and management. *International Journal of Critical Infrastructures,* 2(2-3), 146-159.

Hill, A. B. (1966). The Environment and Disease: Association or Causation. *Proceedings of the Royal Society of Medicine,* 58, 295-300.

Koppenjan, J. F. M., & Groenewegen, J. (2005). Institutional design for complex technological systems. *International Journal of Technology, Policy and Management,* 5(3), 240-257.

Kwakkel, J. H., & Cunningham, S. W. (2008a). The Semantics of the Uncertainty Literature. *2008 Portland International Conference on Management of Engineering & Technology, Vols 1-5*, 2165-2175.

Kwakkel, J. H., & Cunningham, S. W. (2008b). Techniques and Methods for Uncertainty Management. *2008 Portland International Conference on Management of Engineering & Technology, Vols 1-5*, 1085-1097.

Kwakkel, J. H., Walker, W. E., & Marchau, V. A. W. J. (2010). Adaptive Airport Strategic Planning. *European Journal of Transportation and Infrastructure Research,* 10(3), 227-250.

Kwakkel, J. H., Walker, W. E., & Wijnen, R. A. A. (2008, 6-10 July). *The Treatment of Uncertainty in Airport Strategic Planning: The Case of Schiphol Airport's Long-Term Vision*. Paper presented at the 12th Air Transport Research Society World Conference, Athens, Greece.

Marchau, V. A. W. J., Walker, W. E., & van Duin, R. (2009). An adaptive approach to implementing innovative urban transport solutions. *Transport Policy,* 15(6), 405-412.

Marchau, V. A. W. J., Walker, W. E., & van Wee, G. P. (2010). Dynamic Adaptive Transport Policies for Handling Deep Uncertainty. *Technological Forecasting and Social Change,* in press.

Meijer, M. (2007). *Intertwining Uncertainty Analysis and Decision-Making about Drinking Water Infrastructure*. Ph.D. thesis, Delft University of Technology, Delft.

Ostrom, E. (1990). *Governing the Commons. The Evolution of Institions for Collective Action*. Cambridge: Cambridge University Press.

Pahl-Wostl, C., Sendzimir, J., Jeffrey, P., Aerts, J., Berkamp, G., & Cross, K. (2007). Managing change towards adaptive water management through social learning. *Ecology and Society,* 12(2), 30 [online].

Provincie Noord-Holland. (2007). *Vestigingslocaties Schiphol, een globale verkenning voor de lange termijn*. Retrieved from www.noord-holland.nl/Images/65_105492.pdf

Schiphol Group. (2007). *Lange Termijn Visie op de Ontwikkeling van de Mainport Schiphol: Een Wereldwijd Netwerk Voor een Concurrerende Randstad [Long-Term Vision on the Development of the Mainport Schiphol]*.

Schiphol Group, & LVNL. (2007). *Verder Werken aan de toekomst van Schiphol en de regio [continu working on the future of Schiphol and the region]*.

Taneja, P., Walker, W. E., Ligteringen, H., van Schuylenburg, M., & van der Plas, R. (2010). Implications of an Uncertain Future for Port Planning. *Maritime Policy & Management,* 37(3), 221-245.

To70. (2007). *Eisen voor gelijkwaardige bescherming bij wijziging van de Luchthavenbesluiten*. Retrieved from www.evaluatieschipholbeleid.nl/download/3005/06b.pdf

U.S. Department of Healt and Human Services - Food and Drug Administration. (1998). *Guidance for Industry - Providing Clinical Evidence of Effectiveness for Human Drugs and Biological Products*: U.S. Department of Healt and Human Services - Food and Drug Administration.

V&W. (2007). *Nieuwsbrief Schipholbeleid nr 12 [newsletter Schiphol policy, nr 12]*: Ministry of Transport.

V&W, & VROM. (2007). *Lange Termijn Verkenning Schiphol, probleemanalyse [Long term exploration Schiphol: problem exploration]*: Ministry of Transport, Ministry of Housing, Spatial Planning and the Environment.

Walker, W. E., Harremoës, J., Rotmans, J. P., Van der Sluijs, J. P., Van Asselt, M. B. A., Janssen, P. H. M., et al. (2003). Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment,* 4(1), 5-17.

Walker, W. E., Rahman, S. A., & Cave, J. (2001). Adaptive Policies, Policy Analysis, and Policymaking. *European Journal of Operational Research,* 128(2), 282-289.

Williamson, O. E. (1997). Transaction cost economics: the governance of contractual relations. *Journal of Law and Economics,* 22(October), 3-61.

Summary: The treatment of uncertainty in airport strategic planning

J.H. Kwakkel

The air transport industry operates in a fast changing environment. At the end of the 1970's, the industry was liberalized and privatized in the U.S.A. Europe followed in the 1990's. As a result of this privatization and liberalization, the industry has undergone unprecedented changes, exemplified by the rise of airline alliances and low cost carriers. Parallel to this, the industry has witnessed increasing environmental awareness, which has resulted in more attention being paid to the negative external effects of aviation, such as noise and emissions, and, since 9/11, safety and security are also of more concern. It is likely that the aviation industry will become even more dynamic in the coming years, for example because of the recently signed U.S.A.-Europe Open Skies treaty. In addition to these changes in the air transport industry itself, outside influences, such as the oil price, flu epidemics, and financial and economic woes, further add to the volatility of aviation demand development. All these changes together pose a major challenge for airports. They have to plan the various investments that will shape the future of the airport for many years to come, taking into consideration the many uncertainties that are present.

The current dominant approach for the long-term planning of an airport is Airport Master Planning (AMP). The goal of AMP is to provide a blueprint that will determine the future development of the airport. AMP is a formalized, structured planning process that results in a Master Plan that presents the planner's conception of the ultimate development of a specific airport. In the United States, the Federal Aviation Authority has set up strict guidelines for an AMP study. Internationally, reference manuals of the International Air Transport Association (IATA), the International Civil Aviation Organization (ICAO), and books about airport planning by leading scholars heavily influence AMP practices. There are many cases that that clearly show that AMP has proven to be ineffective. This inefficacy of AMP is also well recognized in the scientific literature. The central question this thesis addressed is "how can uncertainties in the long-term planning of an airport be treated in a better way than is currently done in AMP?"

Uncertainty is an ambiguous term. It has different meanings depending on the context in which it is used. Before being able to address the central research question, we therefore, in chapter 2, first elucidate the various meanings and connotations associated with the term uncertainty. In order to address this problem, we analyzed the language of scientists as found in the abstracts of papers when they use words such as uncertain, uncertainty, and uncertainties. The aim of this analysis was to reveal the various meanings that uncertainty has in the sciences and where these meanings are used. The analysis revealed that there are at least nine different answers to the question what is meant by uncertainty. The answer given by someone working in the medical sciences will differ substantially from the answer given by a scientist working in the natural sciences. It also became clear that the various sciences that are involved in offering decision support for the long-term planning of airports are not internally consistent in their usage of the term.

In chapter 3 we set out to develop a framework that offered a common conceptual basis for the various scientific disciplines that are involved in the long-term planning of airports. In general, these various disciplines all make use of models. The framework therefore focuses on

offering a common conceptual basis for discussing uncertainty in model-based decision support. According to this framework, uncertainty is a three dimensional concept:

- *Location*: where is the uncertainty located? Uncertainty can be located in any of the various aspects of a policy problem: (a) the future world, (b) the model of the relevant system for that future world, (c) the outcomes from the system, and (d) the weights that the various stakeholders place on the various outcomes of the system
- *Nature*: is the uncertainty due to a lack of knowledge, due to inherent variability in the real world system, or due to different ways of framing?
- *Level*: how severe is or what is the degree of the uncertainty? Broadly speaking, the level of uncertainty is the assignment of likelihood to things or events. So, uncertainty for model-based policy analysis means that there are multiple plausible alternative future worlds, system models, outcomes, and/or weights. In some cases the likelihood or plausibility of these alternatives can be expressed using numbers, but in other cases more imprecise labels are used, such as more likely, less likely, or equally likely. We define four levels of uncertainty: We speak of Level 1 uncertainty, or shallow uncertainty, when one is able to enumerate multiple alternatives and is able to provide probabilities. We speak of Level 2 uncertainty, or medium uncertainty, when one is able to enumerate multiple alternatives and able to rank order the alternatives in terms of perceived likelihood. However, how much more likely or unlikely one alternative is compared to another cannot be specified. We speak of Level 3 uncertainty, or deep uncertainty, when one is able to enumerate multiple alternatives without being able or willing to rank order the alternatives in terms of how likely or plausible they are judged to be. Finally, we speak of Level 4 uncertainty, or recognized ignorance when one is unable to enumerate multiple alternatives, while admitting the possibility of being surprised.

Having provided clarity on how uncertainty is understood in this research, we turn to addressing the central research question: How can the uncertainties currently prevalent in the long-term planning of airports be treated better? Chapter 4 presents an approach that is designed to better handle uncertainty than AMP. Figure 1 shows the approach that we call Adaptive Airport Strategic Planning (AASP). It is a synthesis of various ideas that were put forward in the (airport) planning literature and literature on adaptive policymaking. In short, in Step I, the existing conditions of an airport are analyzed and the goals for future development are specified. In Step II, the way in which this is to be achieved is specified. This basic plan is made more robust through four types of actions specified in Step III: mitigating actions are actions to reduce the *certain* adverse effects of a plan; hedging actions are actions to spread or reduce the risk of *uncertain* adverse effects of a plan; seizing actions are actions taken to seize certain available opportunities; and shaping actions are actions taken to reduce the chance that an external condition or event that could make the plan fail will occur, or to increase the chance that an external condition or event that could make the plan succeed will occur. Even with the actions taken in Step III, there is still the need to monitor the performance of the plan and take action if necessary. This is called contingency planning and it is specified in Step IV. Signposts specify information that should be tracked in order to determine whether the plan is achieving its conditions for success. Critical values of signpost variables (triggers) are specified, beyond which actions should be implemented to ensure that the plan keeps moving the system in the right direction and at a proper speed. There are four different types of actions that can be triggered by a signpost: defensive actions are taken to clarify the basic plan, preserve its benefits, or meet outside challenges in response to specific triggers that leave the basic plan remains unchanged; corrective actions are adjustments to the basic plan; capitalizing actions are actions trigger to take advantage of opportunities that improve the performance of the basic plan; and a reassessment of the plan is initiated when the analysis and assumptions critical to the plan's success have clearly lost validity. Step V is the actual implementation. In this step, the actions to be taken immediately (from Step II and Step III) are implemented and a monitoring system (from Step IV) is established. Then time starts running, signpost information related to the triggers is collected, and actions are started, altered, stopped, or expanded in response to this. After implementation of the initial actions, the implementation of other actions is suspended until a trigger event occurs. The central claim of AASP is that a more successful coherent development of infrastructure can be achieved by planning for adaptation instead of having to make ad-hoc adaptations and modifications to existing plans

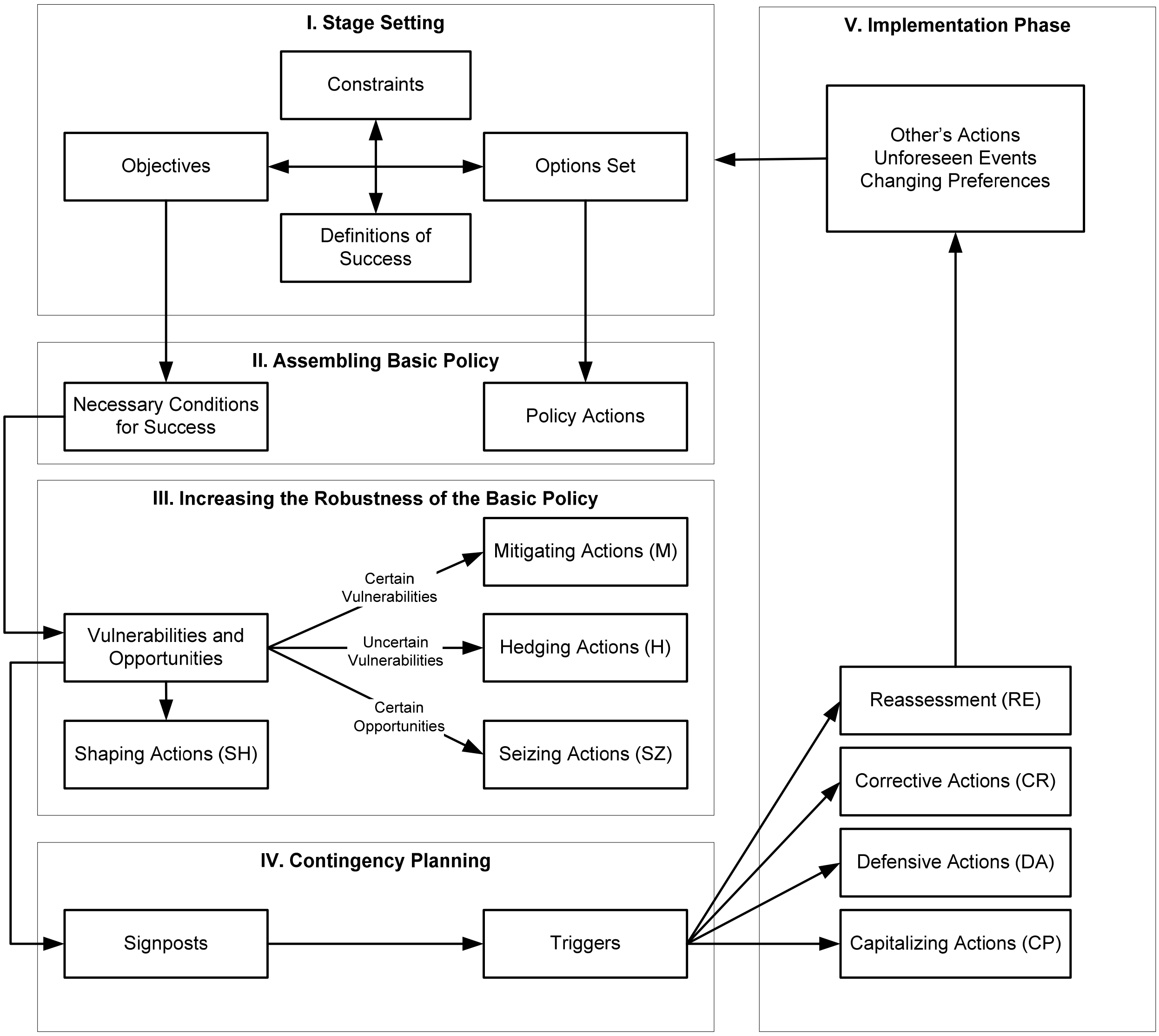


Figure 1: Adaptive Airport Strategic Planning

AASP covers the four levels of uncertainty as identified in chapter 3. Level 1 and 2 uncertainty are considered jointly in Steps I and II. These two steps are comparable to AMP. Step III focuses on increasing the robustness of the plan through static actions taken at the start of the implementation. This corresponds to the use of for example scenario planning, which can be used to handle Level III uncertainty about the future world. The fourth step in AASP focuses on contingency planning. It prepares the basic plan for surprises and builds in mechanisms to handle these surprises. This fits with flexible and adaptive approaches that have been advocated for handling Level 4 uncertainty. AASP thus is capable of handling all the levels of uncertainty

The next step is to test whether AASP is indeed better than AMP. However, this is not an easy question. How can this be shown? The monitoring, over time, of a single instance of a plan generated by AASP, can only prove the efficacy of the planning approach if there is a competing Master Plan. This way of gathering evidence for the efficacy of AASP is practically impossible. It is next to impossible to find two airports that are comparable enough to implement two rival plans. Moreover, it would take a long time before the evidence comes in. In order to identify an alternative way of gathering evidence, we draw an analogy with medicine. Medicine has a well-established approach for gathering evidence about the efficacy of new treatments. Evidence for the efficacy of medical treatments is derived from five sources. *Theoretical developments* in disciplines such as biology, biochemistry, and the life sciences suggest new treatments. *Animal models* are used as a surrogate for the human body in early tests of new treatments. Human tissue is used *in vitro* to assess the efficacy of the treatment, without risking harm to human subjects. *Natural experiments* are utilized when controlled experimentation is not possible (e.g. to study to long-term health effects of a treatment). *Clinical trials* are the final test for new treatments and are subject to strict rules to minimize various threats to validity. In Chapter 5, we argue that these same five sources of evidence can also be used to provide evidence for the efficacy of new planning approaches such as AASP.

In Chapter 6 we provide evidence for the efficacy of AASP through series of computational experiments. This use of simulation models as a stand in for the real system corresponds to the use of animal models in medicine. Animal models in medical research and development only bear a partial resemblance to the real, human, subject. As such they are comparable to simulation models frequently employed in engineering. By comparing the performance of an Adaptive Plan and a Master Plan as calculated by the computer models, one can reason about how these plans would behave in the real world. We developed a model of an airport that calculates various airport performance metrics, such as the size of the noise contour, the capacity of the airport, the cumulative emissions of the aircraft operations, and metrics related to external safety. Next, we developed a scenario generator that can generate a wide range of different plausible futures. This created the possibility to assess the performance of a Master Plan and an Adaptive Plan across a wide range of plausible futures. This showed that an Adaptive Plan has less variance in its outcomes then a Master Plan. Thus, an Adaptive Plan exposes an airport to less risk than a Master Plan. Furthermore, AMP is preferable to AASP only if the future demand were to fall into a narrow bandwidth. Conversely, if there is even small uncertainty about future demand, the Adaptive Plan is to be preferred. The series of computational experiments thus provides corroborating evidence for the efficacy of AASP.

The aim of this thesis was to design and test and approach for Airport Strategic Planning that was capable of handling the various uncertainties airports are facing. In line with this, the thesis' main contributions are to the literature on uncertainty and to the adaptive policymaking literature. With respect to the uncertainty literature, the contribution of this thesis is twofold. First, it provides empirical insight into the wide variety of meanings that uncertainty has in the scientific literature. Moreover, it not only shows the various meanings that exist, but also reveals which scientific fields make use of a particular meaning. Second, this thesis contributes to the literature on uncertainty analysis and classification by synthesizing the literature on model-based decision support in a conceptual framework. With respect to the adaptive policymaking literature, this thesis makes three contributions. First, it expands the adaptive policymaking framework in order to bring out more clearly how opportunities fit into this framework. The resulting adaptive policymaking framework is now also consistent with the theoretical literature on uncertainties in that the various steps specified in the framework map to the four levels of uncertainty that have been recognized in the uncertainty literature. Second, this thesis outlines a methodology for providing evidence for the efficacy of adaptive policymaking. Third, utilizing this methodology, this thesis provides evidence for the efficacy of adaptive policymaking based on series of computational evidence.

Samenvatting: De behandeling van onzekerheid in de strategische planning van luchthavens

J.H. Kwakkel

De luchttransportsector opereert in een snel veranderende omgeving. Aan het einde van de jaren zeventig werden in de Verenigde Staten delen van de sector geprivatiseerd. Europa volgde in de jaren negentig. Als gevolg van deze ontwikkeling heeft de sector een enorme verandering ondergaan, dat zichtbaar wordt in het ontstaan van allianties en de opkomst van low-cost carriers. Parallel aan deze ontwikkelingen is er meer aandacht gekomen voor de negatieve milieu effecten van luchtvaart, zoals geluidsoverlast en de uitstoot van vervuilende gassen. Sinds 11 september 2001 is veiligheid ook veel belangrijker geworden. De verwachting is dat deze ontwikkelingen zich verder zullen doorzetten waardoor de sector nog dynamischer zal worden. Mede ook vanwege het recente 'Open Skies' verdrag tussen Europa en de Verenigde Staten. Naast deze veranderingen binnen de luchtvaartsector zijn er ook nog allerlei invloeden van buitenaf, zoals de ontwikkeling van de olieprijs, griep epidemieën, en financiële en economische problemen. Al deze veranderingen gezamenlijk vormen een enorme uitdaging voor de langetermijnplanning van luchthavens, waarbij met al deze onzekerheden rekening gehouden dient te worden.

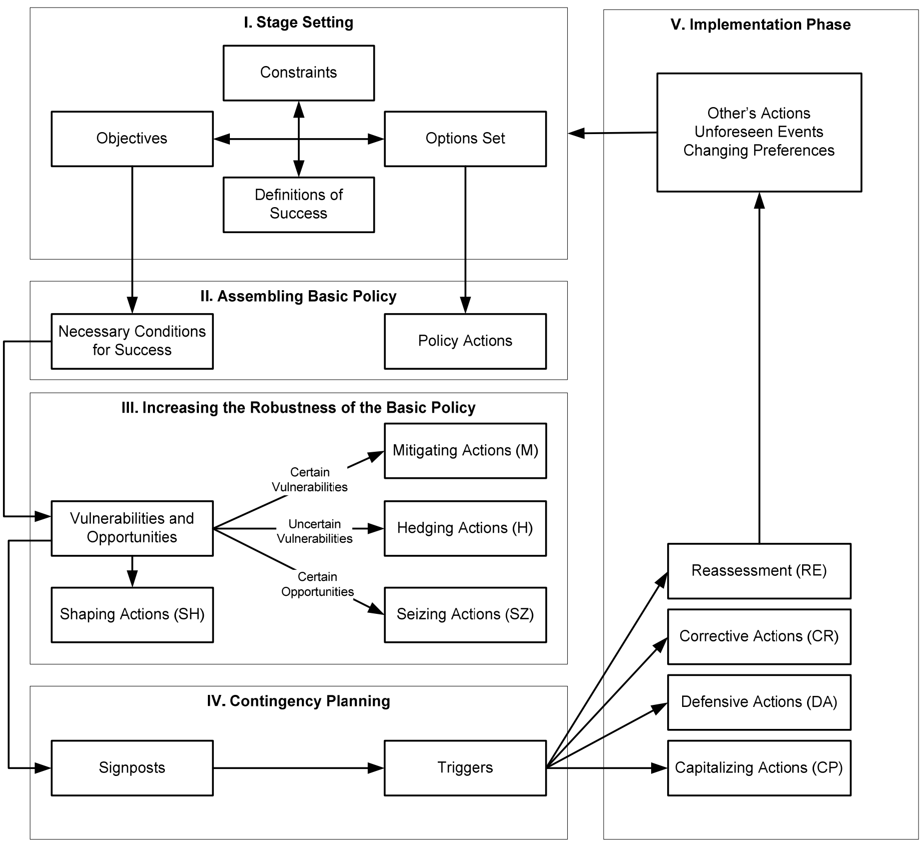
Momenteel wordt bij de langetermijnplanning van luchthavens veelal gebruik gemaakt van luchthavenmasterplanning (LMP). Het doel van LMP is om een blauwdruk te ontwikkelen die de toekomstige ontwikkeling van de luchthaven zal bepalen. LMP is een formeel gestructureerd planningsproces dat resulteert in een visie op de ultieme ontwikkeling van een specifieke luchthaven. In de Verenigde Staten heeft de Federal Aviation Authority strikte regels opgezet voor een LMP studie. Internationaal zijn naslagwerken van International Air Transport Association (IATA), de International Civil Aviation Organization (ICAO), en boeken van wetenschappers van grote invloed op de planningspraktijk. Er zijn veel voorbeelden te vinden waaruit blijkt dat masterplanning niet effectief is. Dit wordt ook onderkent in de wetenschappelijke literatuur. De centrale vraag die behandeld wordt in deze dissertaties is "hoe kunnen onzekerheden in de langetermijnplanning van een luchthaven beter behandeld worden dan momenteel gebeurt in LMP?"

De term onzekerheid is ambigu. Het heeft verschillende betekenissen, afhankelijk van de context waarbinnen het gebruikt wordt. Voordat de centrale vraag beantwoord kan worden is het nodig om eerst helderheid te verschaffen over de verschillende betekenissen en connotaties van de term onzekerheid. Dit gebeurt in hoofdstuk 2. Daar wordt het taalgebruik van wetenschappers, zoals dat aangetroffen wordt in de abstracts van papers over onzekerheid geanalyseerd. Het doel van deze analyse was om de verschillende betekenissen van de term onzekerheid in de wetenschappen en waar deze verschillende betekenissen gebruikt worden, helder te krijgen. De analyse liet zien dat er minstens negen verschillende antwoorden mogelijk zijn op de vraag naar de betekenis van de term onzekerheid. Het antwoord van een medicus zal duidelijk anders zijn dan dat van een natuurwetenschapper. Daarnaast werd ook duidelijk dat de verschillende wetenschappelijke disciplines die betrokken zijn bij de langetermijnontwikkeling van luchthavens niet intern consistent zijn in hun gebruik van de term.

In hoofdstuk 3 wordt een framework ontwikkeld dat kan dienen als een gedeelde conceptuele basis voor de verschillende wetenschappelijke disciplines die betrokken zijn bij de langetermijnontwikkeling van luchthavens. Aangezien nagenoeg al deze disciplines gebruik maken van modellen, richt het framework zich op het bieden van een gedeelde conceptuele basis voor het bespreken van onzekerheden in modelgebaseerde besluitvormingsondersteuning. Volgens dit framework kent onzekerheid drie dimensies:

- Locatie, waar is de onzekerheid gelokaliseerd? Onzekerheid kan zich bevinden in elk van de aspecten van een beleidsprobleem (a) de toekomstige wereld, (b) het model van het systeem in die toekomstige wereld, (c) de uitkomsten van het systeem, en (d) de waardering die verschillende partijen toekennen aan de verschillende uitkomsten.
- Aard, is de onzekerheid een gevolg van een gebrek aan kennis, inherent aan de natuurlijke variabiliteit in de werkelijke wereld, of het gevolg van verschillende manieren van framen?
- Niveau, hoe ernstig is de onzekerheid? In het algemeen is het niveau van de onzekerheid het toekennen van een waarschijnlijkheid aan een gebeurtenis of ding. Onzekerheid voor modelgebaseerde beleidsondersteuning betekent dus dat er meerdere alternatieve toekomstige werelden, systeemmodellen, uitkomsten en/of waarderingen zijn. In sommige gevallen wordt de waarschijnlijkheid van deze verschillende alternatieven uitgedrukt middels getallen, maar soms worden er ook andere minder precieze labels gebruikt, zoals waarschijnlijker, minder waarschijnlijk, of even waarschijnlijk. We definiëren vier niveaus van onzekerheid. We spreken van niveau 1 onzekerheid, of oppervlakkige onzekerheid, als men in staat is om verschillende alternatieven op te sommen en in staat is om hieraan kansen toe te kennen. We spreken van niveau 2 onzekerheid, of matige onzekerheid, wanneer men in staat is om meerdere alternatieven op te sommen en in staat is om deze te ordenen in termen van gepercipieerde waarschijnlijkheid. Hoeveel meer of minder waarschijnlijk het ene alternatief is ten opzichte van de ander kan of wil men niet aangeven. We spreken van niveau 3 onzekerheid, of diepe onzekerheid, wanneer men in staat is om meerdere alternatieven op te sommen, maar men is niet in staat of wil deze alternatieven niet ordenen in termen van gepercipieerde waarschijnlijkheid. Tenslotte spreken we van niveau 4 onzekerheid, of erkende onwetendheid, wanneer men niet in staat is om verschillende alternatieven op te sommen, maar wel onderkent dat men verrast kan worden.

Nu het verheldert is wat er onder onzekerheid verstaan wordt in dit onderzoek, kunnen we terug keren naar de centrale vraag van deze dissertatie. Hoe kunnen onzekerheden in de langetermijnplanning van een luchthaven beter behandeld worden dan momenteel gebeurt in LMP? Hoofdstuk 4 presenteert een aanpak die ontworpen is om beter om te gaan met onzekerheid dan LMP. Figuur 1 toont deze aanpak die we Adaptive Airport Strategic Planning noemen. Het is een synthese van verschillende ideeën die in de literatuur over (luchthaven)planning en adaptief beleid naar voren zijn gekomen. In het kort komt het er op neer dat in stap 1 de bestaande situatie van een luchthaven in kaart wordt gebracht en de doelen voor toekomstige ontwikkeling worden vastgesteld. In stap 2 wordt de manier waarop deze doelen bereikt gaan worden vastgesteld. Dit basisplan wordt in stap 3 robuuster gemaakt via 4 soorten acties: mitigerende acties zijn acties die de zekere negatieve effecten van een plan verminderen, indekkende acties zijn acties die de onzekere negatieve effecten verminderen of verspreiden, aangrijpende zijn acties die genomen worden om de beschikbare kansen te grijpen, vormende acties zijn acties die de mogelijkheid dat een externe conditie of gebeurtenis die het plan zou doen mislukken verminderd of juist de mogelijkheid dat een externe conditie of gebeurtenis die het succes van het plan groter wordt doet toenemen. Zelfs met de acties die gespecificeerd zijn in stap 3 is het nog steeds noodzakelijk om de prestaties van het plan te monitoren en eventuele acties voor te bereiden. Dit wordt plannen voor onvoorziene gebeurtenissen genoemd en dit wordt gedaan in stap 4. Een signpost specificeert welke informatie in de gaten gehouden moet worden om na te kunnen gaan of het plan zijn gestelde doelen bereikt. Kritieke waarden, de zogenaamde triggers, geven aan welke acties genomen moeten worden om er voor te zorgen dat het plan het systeem in de gewenste richting blijft sturen met de gewenste snelheid. Er zijn vier verschillende typen acties die door een trigger geactiveerd kunnen worden: (i) defensieve acties worden genomen om het bestaande plan de verhelderen, zijn voordelen vast te houden, of te kunnen reageren op uitdagingen van buitenaf, ze laten het bestaande plan ongewijzigd; (ii) corrigerende acties zijn wijzigingen van het bestaande plan; (iii) kapitaliserende acties zijn acties die voordeel proberen te halen uit geboden kansen, (iv) een herziening van het hele plan vindt plaats als blijkt dat de cruciale aannames die ten grondslag lagen aan het plan niet langer blijken te kloppen. Stap 5 is de feitelijke implementatie. In deze stap wordt het plan van stap 2 samen met de acties van stap 3 geïmplementeerd en het monitoringsysteem van stap 4 wordt ingesteld. Dan begint de tijd te lopen, de informatie van de signposts wordt verzameld en acties worden in gang gezet of gestopt al naar gelang de informatie van de signposts. Nieuwe acties worden pas geïnitieerd als deze getriggered worden. De centrale claim van deze aanpak is dat een meer succesvolle en coherente ontwikkeling van een infrastructuur mogelijk is door te plannen met adaptatie in gedachte, in plaats van op een ad-hoc manier adaptaties en modificaties van het bestaande plan door te voeren.



Figuur 1: Adaptive Airport Strategic Planning

AASP bestrijkt de vier niveaus van onzekerheid zoals die geïdentificeerd zijn in hoofdstuk 3. Niveau 1 en 2 worden gezamenlijk behandeld in stap 1 en 2. Deze twee stappen zijn vergelijkbaar met het huidige LMP. Stap 3 richt zich op het verhogen van de robuustheid van het plan door het nemen van statische acties bij het begin van de implementatie. Dit is vergelijkbaar met bijvoorbeeld scenarioplanning, wat een methode is die primair bedoeld is voor het omgaan met niveau 3 onzekerheid over de toekomstige wereld. De vierde stap in AASP richt zich op plannen voor onvoorzienbare gebeurtenissen. Het bereidt het basisplan voor op verassingen en bouwt mechanismen in om met deze verassingen om te gaan. Dit komt overeen met flexibele en adaptieve aanpakken die voorgesteld zijn als manier om om te gaan met niveau 4 onzekerheid. AASP dekt dus alle vier niveaus van onzekerheid expliciet af.

De volgende stap in het onderzoek is om te testen of AASP inderdaad beter is dan LMP. Dit is echter niet eenvoudig want hoe kan dit aangetoond worden? Het volgen van één adaptief plan kan alleen de werkzaamheid van AASP aantonen als er ook een alternatief traditioneel masterplan is. Maar dit vereist twee nagenoeg identieke luchthavens, wat praktisch niet haalbaar is. Daarnaast zou een dergelijke test veel tijd kosten voordat het bewijs beschikbaar komt. Immers de prestaties van beide plannen zou voor langere tijd gevolgd moeten worden. Om te komen tot een alternatieve methode voor het testen van AASP wordt er een analogie gemaakt met hoe medicijnen getest worden. Bewijs voor de werkzaamheid van medicijnen komt uit vijf verschillende bronnen. *Theoretische ontwikkelingen* in bijvoorbeeld de biologie en de biochemie kunnen nieuwe medicijnen suggereren. *Dierproeven* worden gebruikt als surrogaat voor mensen in de vroege tests van nieuwe medicijnen. Menselijk materiaal wordt gebruikt *in vitro* om het effect van een medicijn vast te stellen zonder daarbij het risico te lopen mensen schade te berokkenen. *Natuurlijke experimenten* worden gebruikt als gecontroleerde experimenten niet mogelijk zijn, zoals bijvoorbeeld in het geval van onderzoek naar de langetermijngezondheidseffecten van een medicijn. *Klinische experimenten* zijn de laatste test voor nieuwe medicijnen en zijn onderhevig aan een breed scala aan strikte regels en eisen om allerlei bedreigingen voor de validiteit uit te sluiten. In hoofdstuk 5 wordt betoogd dat deze vijf bronnen van bewijs ook gebruikt kunnen worden om bewijs te leveren voor de werkzaamheid van nieuwe planningsaanpakken zoals AASP.

In hoofdstuk 6 wordt er bewijs geleverd voor de werkzaamheid van AASP via een serie van computationele experimenten. Het gebruik van simulatiemodellen als vervanger van het werkelijke systeem is analoog aan het gebruik van dieren in de medische wereld. Dieren zijn slechts gedeeltelijk te vergelijken met mensen. Dit is vergelijkbaar met de manier waarop simulatiemodellen vaak gebruikt worden in de ingenieurswereld. Door de prestatie van een adaptief plan en een masterplan, zoals uitgerekend door een model, te vergelijken, kan er inzicht gekregen worden in hoe deze plannen zouden kunnen presteren in de werkelijke wereld. Daarom is er een model ontwikkeld dat een aantal belangrijke prestatie indicatoren voor een luchthaven uitrekent, zoals de omvang van een geluidscontour, de capaciteit van de luchthaven, de emissies van de vliegtuigen die starten en landen op de luchthaven, en indicatoren met betrekking tot de externe veiligheid. Vervolgens hebben we dit model gekoppeld aan een scenariogenerator die een enorme variëteit aan mogelijke toekomsten kan genereren. Dit geeft de mogelijkheid om de prestaties van een masterplan en een adaptief plan te vergelijken voor een enorme variëteit aan mogelijke toekomsten. Deze vergelijking liet zien dat er minder variantie is in de uitkomsten van een adaptief plan. Anders gezegd, een adaptief plan vermindert het risico van een luchthaven over de te verwachten toekomstige prestaties. Daarnaast werd ook duidelijk dat LMP alleen de voorkeur verdient als de vraag naar luchtvervoer binnen een beperkte bandbreedte viel. Buiten deze bandbreedte heeft adaptief plannen de voorkeur. Deze serie van experimenten levert hiermee bewijs voor de werkzaamheid van AASP.

Het doel van deze dissertatie was om een planningsaanpak te ontwerpen en te testen voor de langetermijnontwikkeling van luchthavens, die goed kan om gaan met de verschillende onzekerheden waarmee luchthavens te maken hebben. Deze dissertatie draagt dan ook primair bij aan de literatuur over onzekerheid en adaptief plannen. Wat betreft de onzekerheidsliteratuur zijn er twee bijdragen. Allereerst geeft hoofdstuk 2 empirisch inzicht in de variëteit aan betekenissen die de term onzekerheid heeft in de wetenschappen. Ten tweede levert deze dissertatie een bijdrage aan de onzekerheidsliteratuur door de synthese van verschillende papers in een framework voor de analyse en classificatie van onzekerheden in modelgebaseerde beleidsondersteuning. Wat betreft de literatuur over adaptief plannen levert deze dissertatie drie bijdragen. Ten eerste breidt het een bestaand framework, dat de nadruk legt op hoe er om gegaan moet worden met bedreigingen uit, door expliciet in te gaan op hoe er om gegaan kan worden met kansen. Daarnaast zorgt deze uitbreiding ervoor dat het framework voor adaptief plannen consistent is met de vier onderkende niveaus van onzekerheid. Ten tweede reikt deze dissertatie een methodologie aan die gebruikt kan worden voor onderzoek naar de werkzaamheid van adaptief plannen. Ten derde, gebruikmakend van deze methodologie levert deze dissertatie bewijs voor de werkzaamheid van adaptief plannen via een serie computationele experimenten.

About the author

Jan Kwakkel was born in Biddinghuizen on 9 July 1982. He studied Systems Engineering, Policy Analysis and Management at Delft University of Technology. He graduated in 2006 on a thesis about the failure of decision support systems. He also studied Philosophy at Leiden University. He finished his bachelor degree there in 2010 with a paper applying Hussler's phenomenological theory of knowledge to the problem of uncertainty for scientific decision support. From 2006 till 2010 he was a PhD. Student at Delft University of Technology. He was located at the Policy Analysis section of the Faculty of Technology and Management, although his research was co-funded by the Transport Policy and Logistics' Organization section of the same faculty.

During his PhD research, he co-authored various papers related to his research interest: uncertainty, model-based decision support, airport planning, and content analysis. He also worked on a European fifth and sixth framework project called SPADE, which aimed at developing a decision support system for airport strategic and operational planning. He has been involved in teaching various policy analysis and system dynamics courses in the different programs offered by the faculty of Technology, Policy and Management. Most notably, this resulted in a book on Policy Analysis of Multi-Actor systems.

Overview of Scientific Publications

Journal papers

WIJNEN, R. A. A., WALKER, W. E. & KWAKKEL, J. H. (2008) *Decision Support for Airport Strategic Planning*. Transportation Planning and Technology, vol. 31, no. 1, pp. 11-34.

KWAKKEL, J. H. & CUNNINGHAM, S. W. (2009) *Managing polysemy and synonymy in science mapping using the mixtures of factor analyzers model*. Journal of the American Society for Information Science and Technology, vol. 60, no. 10, pp. 2064-2078.

KWAKKEL, J. H., WALKER, W. E. & MARCHAU, V. A. W. J. (2010) *Classifying and communicating uncertainties in model-based policy analysis*. International Journal of Technology, Policy and Management, 10(4), 299-315

KWAKKEL, J. H., WALKER, W. E. & MARCHAU, V. A. W. J. (2010) *Adaptive Airport Strategic Planning*. European Journal of Transportation and Infrastructure Research, 10(3), 227-250.

CUNNINGHAM, S.W. & KWAKKEL J.H. (in press) *Innovation Forecasting: A Case Study of the Management of Engineering and Technology Literature*, Technological Forecasting and Social Change

Conference papers

KWAKKEL, J. H., WALKER, W. E. & MARCHAU, V. A. W. J. (2007) *Dealing with Uncertainty in Airport Strategic Planning*. IN HANSSEN, M. (Ed.) 11th Air Transport Research Society World Conference. Berkeley, USA.

PRUYT, E., KWAKKEL, J.H., (2007) *Combining System Dynamics and Ethics: Towards more Science*? Annual System Dynamics Conference, Boston, United States

CUNNINGHAM, S. W. & KWAKKEL, J. H. (2008) *Mixed Representations of Science and technology Data for Use in the Management of Technology*. Portland International Conference on Management of Engineering and Technology. Cape Town, South Africa.

KWAKKEL, J. H. & CUNNINGHAM, S. W. (2008) *The Semantics of the Uncertainty Literature*. Portland International Conference on Management of Engineering and Technology. Cape Town, South Africa.

KWAKKEL, J. H. & CUNNINGHAM, S. W. (2008) *Techniques and Methods for Uncertainty Management*. Portland International Conference on Management of Engineering and Technology. Cape Town, South Africa.

KWAKKEL, J. H., WALKER, W. E. & MARCHAU, V. A. W. J. (2008) *Adaptive Airport Strategic Planning*. IN ZUYLEN, H. J. V. & BINSBERGEN, A. J. V. (Eds.) 10th International Trail Congress. Rotterdam, The Netherlands.

KWAKKEL, J. H., WALKER, W. E. & MARCHAU, V. A. W. J. (2008) *Adaptive Airport Strategic Planning*. IN ZUYLEN, H. J. V. & BINSBERGEN, A. J. V. (Eds.) Proceedings 2008, 10th International Trail Congress, Selected Papers. Rotterdam, The Netherlands.

KWAKKEL, J. H., WALKER, W. E. & MARCHAU, V. A. W. J. (2008) *Adaptive Policy Making for Airport Strategic Planning*. NGInfra/IEEE conference on next generation infrastructures "Building Networks for a Brighter Future". Rotterdam, The Netherlands.

KWAKKEL, J. H., WALKER, W. E. & WIJNEN, R. A. A. (2008) *The Treatment of Uncertainty in Airport Strategic Planning: The Case of Schiphol Airport's Long-Term Vision*. IN PAPATHEODOROU, A. & IATROU, K. (Eds.) 12th Air Transport Research Society World Conference. Athens, Greece.

SLINGER, J. KWAKKEL, J.H., VAN DER NIET, M. (2009) *Does learning to reflect make better modelers?* Annual System Dynamics Conference, Athens, Greece

WIJNEN, R. A. A., CHIN, R. T. H., WALKER, W. E. & KWAKKEL, J. H. (2008) *Decision Support for Mainport Strategic Planning*. IN ZARATE, P., BELAUD, J. P., CAMILLERI, G. & RAVAT, F. (Eds.) Proceedings of the International Conference on Collaborative Decisionmaking. Toulouse, France, IOSPress publishers.

CUNNINGHAM, S. W. & KWAKKEL, J. H. (2009) *Innovation Forecasting: A Case Study of the Management of Engineering and Technology Literature*. Portland International Conference on Management of Engineering and Technology. Portland, United States.

CUNNINGHAM, S. W. & KWAKKEL, J. H. (2009) *A Theory of Infrastructure Provision*. 2nd Next Generation Infrastructure Conference: 'Developing 21st Century Infrastructure Networks'. Chennai, India.

KWAKKEL, J. H. & CUNNINGHAM, S. W. (2009) *Remining Picmet: 1997-2008*. Portland International Conference on Management of Engineering and Technology. Portland, United States.

KWAKKEL, J. H., CUNNINGHAM, S. W. & PAS, J. W. G. M. V. D. (2009) *Evaluation of Infrastructure Planning Approaches: An Analogy with Medicine*. 2nd Next Generation Infrastructure Conference: 'Developing 21st Century Infrastructure Networks'. Chennai, India.

KWAKKEL, J. H., WIJNEN, R. A. A., WALKER, W. E. & MARCHAU, V. A. W. J. (2009) *A Fast and Simple Model for a Quick Scan of Airport Performance*. IN HOOPER, P. (Ed.) 13th Air Transport Research Society World Conference. Abu Dhabi, United Arab Emirates.

CUNNINGHAM, S. W. & KWAKKEL, J. H. (2010) *Anticipating New Knowledge Transfer in Nanotechnology*. Atlanta Transatlantic Workshop on Nanotechnology Innovation and Policy Georgia Institute of Technology, Atlanta, United States, 24-26 March 2010.

DE HAAN, J., KWAKKEL, J.H. (2010) *Uncertainty Framed*. 25th Mini-EURO Conference on "Uncertainty and Robustness in Planning and Decisionmaking" (URPDM2010) University of Coimbra, Portugal, 15-17 April 2010.

KWAKKEL, J. H., WALKER, W. E. & MARCHAU, V. A. W. J. (2010) *From Predictive Modeling to Exploratory Modeling: How to use Non- Predictive Models for Decisionmaking under Deep Uncertainty*. 25th Mini-EURO Conference on "Uncertainty and Robustness in Planning and Decisionmaking" (URPDM2010) University of Coimbra, Portugal, 15-17 April 2010.

KWAKKEL, J. H. & CUNNINGHAM, S. W. (2010) *Nanotechnology Research: An analysis of past trends and expectations for future developments*. Atlanta Transatlantic Workshop on Nanotechnology Innovation and Policy. Georgia Institute of Technology, Atlanta, United States, 24-26 March 2010.

KWAKKEL, J. H., WALKER, W. E. & MARCHAU, V. A. W. J. (2010) *Assessing the Efficacy of Adaptive Airport Strategic Planning: Results from Computational Experiments*. World Conference of the Transportation Research Society, Lisbon, Portugal.

KWAKKEL, J.H., VREUGENHIL, H. & SLINGER, J. (2010) *Paradigms in Sustainability Research*, ERSCP-EMSU 2010 Conference, Delft, the Netherlands, 25-29 October 2010.

KWAKKEL, J.H & WALKER, W.E. (2010) *Grappling with Uncertainty in the Long-Term Development of Infrastructure Systems*, 3nd Next Generation Infrastructure Conference: ‘Infrastructures for Ecocities'. Shenzhen, China, 11-13 November 2010.

Books

ENSERINK, B., HERMANS, L., KWAKKEL, J. H., THISSEN, W., KOPPENJAN, J. F. M. & BOTS, P. W. G. (2010) *Policy Analysis of Multi-Actor Systems*, Utrecht, Lemma.

Part in edited volume

BANKES, S., WALKER, W.E., & KWAKKEL J.H. *Exploratory Modeling and Analysis*, in Encyclopedia of OR, Gass, S. (ed), upcoming 3th edition

Popular scientific publications

KWAKKEL, J.H. (2008) *The Problem of Uncertainty in Airport Master Planning*, Aerlines, 39, p. 1-4

KWAKKEL, J.H. (2008) *Alternative Approaches to Traditional Airport Master Planning*, Aerlines, 41, p. 1-5

TRAIL Thesis Series

The following list contains the most recent dissertations in the TRAIL Thesis Series. For a complete overview of more than 100 titles see the TRAIL website: www.rsTRAIL.nl.

The TRAIL Thesis Series is a series of the Netherlands TRAIL Research School on transport, infrastructure and logistics.

Kwakkel, J.H., *The Treatment of Uncertainty in Airport Strategic Planning*, T2010/13, December 2010, TRAIL Thesis Series, the Netherlands

Pang, Y., *Intelligent Belt Conveyor Monitoring and Control,* T2010/12, December 2010, TRAIL Thesis Series, the Netherlands

Kim, N.S., *Intermodal Freight Transport on the Right Track? Environmental and economic performances and their trade-off,* T2010/11, December 2010, TRAIL Thesis Series, the Netherlands

Snelder, M., *Designing Robust Road Networks: a general design method applied to the Netherlands,* T2010/10, December 2010, TRAIL Thesis Series, the Netherlands

Hinsbergen, C.P.IJ. van, *Bayesian Data Assimilation for Improved Modeling of Road Traffic*, T2010/9, November 2010, TRAIL Thesis Series, the Netherlands

Zuurbier, F.S., *Intelligent Route Guidance*, T2010/8, November 2010, TRAIL Thesis Series, the Netherlands

Larco Martinelli, J.A., *Incorporating Worker-Specific Factors in Operations Management Models*, T2010/7, November 2010, TRAIL Thesis Series, the Netherlands

Ham, J.C. van, Zeehavenontwikkeling in Nederland: naar een beter beleidsvormingsproces, T2010/6, August 2010, TRAIL Thesis Series, the Netherlands

Boer, E. de, *School Concentration and School Travel*, T2010/5, June 2010, TRAIL Thesis Series, the Netherlands

Berg, M. van den, *Integrated Control of Mixed Traffic Networks using Model Predictive Control,* T2010/4, April 2010, TRAIL Thesis Series, the Netherlands

Top, J. van den, *Modelling Risk Control Measures in Railways*, T2010/3, April 2010, TRAIL Thesis Series, the Netherlands

Craen, S. de, *The X-factor: A longitudinal study of calibration in young novice drivers*, T2010/2, March 2010, TRAIL Thesis Series, the Netherlands

Tarau, A.N., *Model-based Control for Postal Automation and Baggage Handling,* T2010/1, January 2010, TRAIL Thesis Series, the Netherlands

Knoop, V.L., *Road Incidents and Network Dynamics: Effects on driving behaviour and traffic congestion*, T2009/13, December 2009, TRAIL Thesis Series, the Netherlands

Baskar, L.D., *Traffic Control and Management with Intelligent Vehicle Highway Systems,* T2009/12, November 2009, TRAIL Thesis Series, the Netherlands

Konings, J.W., *Intermodal Barge Transport: Network Design, Nodes and Competitiveness,* T2009/11, November 2009, TRAIL Thesis Series, the Netherlands

Kusumaningtyas, I., *Mind Your Step: Exploring aspects in the application of long accelerating moving walkways,* T2009/10, October 2009, TRAIL Thesis Series, the Netherlands

Gong, Y., *Stochastic Modelling and Analysis of Warehouse Operations*, T2009/9, September 2009, TRAIL Thesis Series, the Netherlands

Eddia, S., *Transport Policy Implementation and Outcomes*: *the Case of Yaounde in the 1990s*, T2009/8, September 2009, TRAIL Thesis Series, the Netherlands

Platz, T.E., *The Efficient Integration of Inland Shipping into Continental Intermodal Transport Chains: Measures and decisive factors,* T2009/7, August 2009, TRAIL Thesis Series, the Netherlands

Tahmasseby, S., *Reliability in Urban Public Transport Network Assessment and Design*, T2009/6, June 2009, TRAIL Thesis Series, the Netherlands

Bogers, E.A.I., *Traffic Information and Learning in Day-to-day Route Choice,* T2009/5, June 2009, TRAIL Thesis Series, the Netherlands

Amelsfort, D.H. van, *Behavioural Responses and Network Effects of Time-varying Road Pricing,* T2009/4, May 2009, TRAIL Thesis Series, the Netherlands

Li, H., *Reliability-based Dynamic Network Design with Stochastic Networks,* T2009/3, May 2009, TRAIL Thesis Series, the Netherlands

Stankova, K., *On Stackelberg and Inverse Stackelberg Games & their Applications in the Optimal Toll Design Problem, the Energy Markets Liberalization Problem, and in the Theory of Incentives*, T2009/2, February 2009, TRAIL Thesis Series, the Netherlands

Li, T., *Informedness and Customer-Centric Revenue*, T2009/1, January 2009, TRAIL Thesis Series, the Netherlands

Agusdinata, D.B., *Exploratory Modeling and Analysis: A promising method to deal with deep uncertainty,* T2008/17, December 2008, TRAIL Thesis Series, the Netherlands

Kreutzberger, E., *The Innovation of Intermodal Rail Freight Bundling Networks in Europe: Concepts, Developments, Performances,* T2008/16, December 2008, TRAIL Thesis Series, the Netherlands