1 Introduction

1.1 Example Application

1.1.1 Description

To illustrate the need for modeling uncertainty and the concepts, as well as tools, covered in this book, we start off with a virtual case study. “Virtual” meaning that the study concerns an actual situation in an actual area of the world; however, the data, geological studies and, most importantly, the practical outcomes of this example should not be taken as “truth,” which is understandably so after reading the application case.

Much of the world’s drinking water is supplied from groundwater sources. Over the past several decades, many aquifers have been compromised by surface-borne contaminants due to urban growth and farming activities. Further contamination will continue to be a threat until critical surface recharge locations are zoned as groundwater protection areas. This can only be successfully achieved if the hydraulically complex connections between the contaminant sources at the surface and the underlying aquifers are understood.

Denmark is one example of this type of scenario. Since 1999, in an effort to identify crucial recharge zones (zones where water enters the groundwater system to replenish the system), extensive geophysical data sets were collected over the Danish countryside—the areas designated as particularly valuable due to their high rate of water extraction. The data were collected with the intention of making more informed decisions regarding the designation of recharge protection zones. The magnitude of these decisions is considerable, as it could involve the relocation of farms, industry, city development and waterworks together with related large compensations. Consequently, incorrectly identifying a vulnerable area can lead to a costly error. In fact, the Danish Government set out a 10-point program (Figure 1.1) that sets certain objectives and formulates certain desired preferences, some of which may be in conflict with keeping the farming industry alive and ensuring economic health next to ecological health for this area.

The subsurface in Denmark consists of so-called buried valleys, which are considered the informal term for Pleistocene (Quaternary) subglacial channels. They have also been
Danish Government’s 10-point program (1994)

<table>
<thead>
<tr>
<th>Objective</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pesticides dangerous to health and environment shall be removed from the market</td>
<td>Pesticides dangerous to health and environment shall be removed from the market</td>
</tr>
<tr>
<td>Pesticide tax – the consumption of pesticides shall be halved</td>
<td>Pesticide tax – the consumption of pesticides shall be halved</td>
</tr>
<tr>
<td>Nitrate pollution shall be halved before 2000</td>
<td>Nitrate pollution shall be halved before 2000</td>
</tr>
<tr>
<td>Organic farming shall be encouraged</td>
<td>Organic farming shall be encouraged</td>
</tr>
<tr>
<td>Protection of areas of special interest for drinking water</td>
<td>Protection of areas of special interest for drinking water</td>
</tr>
<tr>
<td>New Soil Contamination Act – waste deposits shall be cleaned up</td>
<td>New Soil Contamination Act – waste deposits shall be cleaned up</td>
</tr>
<tr>
<td>Increased afforestation and restoration of nature to protect groundwater</td>
<td>Increased afforestation and restoration of nature to protect groundwater</td>
</tr>
<tr>
<td>Strengthening of the EU achievements</td>
<td>Strengthening of the EU achievements</td>
</tr>
<tr>
<td>Increased control of groundwater and drinking water quality</td>
<td>Increased control of groundwater and drinking water quality</td>
</tr>
<tr>
<td>Dialogue with the farmers and their organisations</td>
<td>Dialogue with the farmers and their organisations</td>
</tr>
</tbody>
</table>

Source: http://www.geus.dk/program-areas/water/denmark/case_groundwaterprotection_print.pdf

**Figure 1.1** Objectives of the Danish Government.

described as the result of waxing and waning of Pleistocene ice sheets. The primary method by which these valleys are formed is subglacial meltwater erosion under the ice or in front of the ice margin. Thus, the valley formation is directly related to the morphology and erodability of the geological strata. The secondary method is through direct glacial erosion by ice sheets.

Several of the processes that created and filled buried valleys are important for understanding the complexity of the Danish aquifer systems and their vulnerability to surface-borne pollutants. In Denmark, the superposition of three different generations of glaciations has been observed. Thus, multigeneration glacial valleys cross-cut each other and can also appear to abruptly end (as seen in Figure 1.2). The existence and location of these glacial valleys can be thought of as the primary level of Denmark’s aquifer system structure. If largely filled with sand, the buried valley has potential for being a high volume aquifer (reservoir). However, these buried valleys can be “re-used,” as revealed by the observed cut-and-fill structures. This describes the secondary level of uncertainty of heterogeneity in Danish aquifer systems.

Most cut-and-fill structures are narrower than the overall buried valley, but in some places very wide structures that span the entire valley width can be seen. The complex internal structure can be observed in seismic surveys, electromagnetic surveys and occasionally in borehole data.

—Sandersen and Jorgensen (2006)

Figure 1.2 shows a few different possible internal heterogeneities and varying extent of overlying strata, which deems the valley as actually “buried.”

Due to the generally complex internal structure of the valleys, potentially protective clay layers above the aquifers are likely to be discontinuous. The aquifers inside the valley will thus have a varying degree of natural protection. Even if laterally extensive clay layers are present, the protective effect will only have local importance if the surrounding
1.1 EXAMPLE APPLICATION

sediments are sand-dominated. The valleys may therefore create short-circuits between the aquifers in the valley and the aquifers in the surrounding strata.

1.1.2 3D Modeling

In this case study, the incompleteness of the information about the subsurface strata makes making specific decisions such as relocating farms difficult. A geologist may be tempted to study in great detail the process by which these glacial valleys were created and come up with a (deterministic) description of these systems based on such understanding, possibly a computer program to simulate the process that created these systems according the physical understanding of what is understood to occur. However, such description alone will fall short in addressing the uncertainty issue that has considerable impact on the decisions made. Indeed, even if full insight into the glaciation process exists (a considerable assumption), then that would not necessarily provide a deterministic rendering of the exact location of these valleys, let alone the detailed spatial distribution of the lithologies (shale, sand, gravel, clay) inside such valleys. This does not mean that the study of the geological processes is useless. On the contrary, such study provides additional information about the possible spatial variation of such channels next to the data gathered (drilling, geophysical surveys). Therefore, additional tools are needed that allow the building of a model of the subsurface glaciations as well as quantifying the uncertainty about the spatial distribution of valley/non-valley and the various lithologies within a valley. Such a model would ideally include the physical understanding as well as reflecting the lack of knowledge, either through limited data or limited geological understanding.

Data play a crucial role in building models and constraining any model of uncertainty, whether simple or complex. In the Danish case, two types of data are present: data obtained through drilling and data obtained through a geophysical method termed time-domain electromagnetic surveys (TEM surveys). Figure 1.3 shows the interpretation of the thickness of the valleys from such surveys, which are basically a collection of 1D (vertical) soundings. The data collected are typical of many Earth modeling situations: some detailed small scale information is gathered through sampling (in this case drilling a well) and some larger scale indirect measurement(s) collected either through geophysical

Figure 1.2 Geological interpretation of subsurface glacial channels cross-cutting each other (left). Conceptual view of the inner structure of the glacial channels (right).
1.2 Modeling Uncertainty

From this case study of modeling the subsurface, several elements in modeling uncertainty that are typical to many similar applications can be identified:

1 Decision making: modeling uncertainty is not a goal on its own, it is usually needed because a particular decision question is raised. In fact, this decision question is usually framed in a larger context, such as done by the 10-point program, specifying objectives and preferences. Two example decisions are in this case: (1) in which areas do we relocate pollution sources and (2) do we consider taking more geophysical data to narrow the uncertainty on locating vulnerable areas, hence increasing the probability of a good decision? This latter question is termed a “Value of Information” question. Clearly, we
1.2 MODELING UNCERTAINTY

need to make decisions without perfect information. These narrower decision questions should not be considered as independent of the larger objective outlined in Figure 1.1.

2 Importance of the geological setting: a critical parameter influencing the decision is the heterogeneity of the subsurface medium (fluids and soils/rocks). Rarely do we have perfect information to deterministically model the geological variability of the subsurface. Hence there is a need to model all aspects of uncertainty as related to the subsurface heterogeneity. While Figures 1.2 and 1.3 may provide one such interpretation of the system, often many alternative and competing interpretations are formed.

3 Data: several sources of data are available to constrain the models of uncertainty built. These data sources can be very diverse, from wells (driller’s logs, well-log, cores, etc.) to geophysical (TEM data in the Danish case) or remote sensing measurements. Tying all this data into a single model of uncertainty without making too many assumptions about the relationships between various data sources is challenging.

From this case study, it is clear that some of the tools for modeling random phenomena through traditional probability models are too rigid to handle all these complexities. The nature of modeling uncertainty in the Earth Science has various challenge and issues that need to be addressed.

1 Modeling uncertainty is often application tailored. If the application changes then the type of modeling and the approach to modeling uncertainty will be different, hence the model of uncertainty will be different. Building a model of uncertainty that includes all possible aspects of what is uncertain is too difficult and often not needed in the first place. Modeling uncertainty for the sake of uncertainty is basically irrelevant as well as an impossible task. For example, if one is looking to quantify the global reserves of an oil reservoir, then the focus should be on the structural model and global parameters such as net-to-gross, while if the question is about drilling the next well, than the analysis should focus on local reservoir heterogeneity and connectivity of flow units.

2 Several sources of uncertainty exist for this case study:

   a Uncertainty related to the measurement errors and processing of the raw measurements.

   b Uncertainty related to the fact that processed data can be interpreted in many ways and, in fact, that data interpretation and processing require a model on their own.

   c Uncertainty related to the type of geological setting used, which is interpreted from data or based on physical models which themselves are uncertain.

   d Spatial uncertainty: even if data were perfectly measured, they are still sparse with respect to the resolution at which we want to build models. This means that various models with different spatial distributions of properties or layering structures can be generated matching equally well the same data.
CH 1 INTRODUCTION

Response uncertainty: this includes uncertainty related to how geological uncertainty translates into modeling of processes such as flow, transport, wave, heat equations or even decisions made based on such models. There may be uncertainty related to the physics of these processes or other parameters that need to be specified to specify these processes. For example, solving partial differential equations requires boundary and initial conditions that may be uncertain.

Uncertainty assessment is subjective: while a “true” Earth exists with all of its true, but unknown properties, there is no “true uncertainty.” The existence of a true uncertainty would call for knowing the truth, which would erase the need for uncertainty assessment. Uncertainty can never be objectively measured. Any assessment of uncertainty will need to be based on a model. Any model, whether statistically or physically defined, based on probability theory or fuzzy logic, requires implicit or explicit model assumptions (because of lack of knowledge or data), hence is necessarily subjective. There is no true uncertainty; there are only models of uncertainty, hence the title of this book.

High dimensional/spatial aspect: we are dealing with complex Earth systems that require a large amount of variables to describe them. Typically, we will work with gridded models to represent all aspects of the natural system. If each grid cell in a model contains a few variables, then easily we have millions of variables for even relatively small models. Standard approaches of probability become difficult to apply, since probability theory and statistical techniques common to most introductory text books has not been developed with these complex situations in mind. Often, it is necessary to perform some sensitivity analysis to determine which factors impact our decision most. Traditional statistical methods for sensitivity analysis are difficult to apply in this high dimensional and spatial context.

Several data sources informing various scales of variability: we will need to deal with a variety of data or information to constrain models of uncertainty. Without any data, there would be no modeling. Such data can be detailed information obtained from wells or more indirect information obtained from geophysical or remote sensing surveys. Each data source (such as wells) informs what we are modeling at a certain “volume support” (such as the size of a soil sample) and measures what we are targeting directly or indirectly, for example, electromagnetic (EM) waves for measuring water saturation.

Following this introductory chapter, this book covers many of these issues in the following chapters:

Chapter 2 Probability, Statistics and Exploratory Data Analysis: basically an overview of basic statistics and probability theory that is required to understand the material in subsequent chapters. The aim is not to provide a thorough review of these fields, but to provide a summary of what is relevant to the kind of modeling in this book.
Chapter 3 Modeling Uncertainty: Concepts and Philosophies: uncertainty is a misunderstood concept in many areas of science, so the various pitfalls in assessing uncertainty are discussed; also, a more conceptual discussion on how to think about uncertainty is provided. Uncertainty is not a mere mathematical concept, it deals with our state of knowledge, or lack thereof, as the world can be perceived by human beings. Therefore, it also has some interesting links with philosophy.

Chapter 4 Engineering the Earth, Making Decisions Under Uncertainty: the basic ideas of decision analysis are covered without going too much into detail. The language of decision analysis is introduced, structuring decision problems is discussed and some basic tools such as decision trees are introduced. The concept of sensitivity analysis is introduced; this will play an important role through many chapters in the book.

Chapter 5 Modeling Spatial Continuity: the chapter covers the various techniques for modeling spatial variability, whether dealing with modeling a rock type in the subsurface, the porosity of these rocks, soil types, clay content, thickness variations and so on. The models most used in practice for capturing spatial continuity are covered; these models are (i) the variogram/covariance model, (ii) the Boolean or object model and (iii) the 3D training image model.

Chapter 6 Modeling Spatial Uncertainty: once a model of spatial continuity is established, we can “simulate the Earth” in 2D, 3D or 4D (including time, for example) based on that continuity model. The goal of such a simulation exercise, termed stochastic simulation, is to create multiple Earth representations, termed Earth models, that reflect the spatial continuity modeled. This set of Earth models is the most common representation of a “model of uncertainty” used in this book. In accordance with Chapter 5, three families of techniques are discussed: a variogram based, object based and 3D training image based.

Chapter 7 Constraining Spatial Uncertainty with Data: this chapter is an extension of the previous chapter and discusses ways for constraining the various Earth representations or models with data. Two types of data are discussed: hard data and soft data. Hard data are (almost) direct measurements of what we are modeling, while soft data are everything else. Typically, hard data are samples taken from the Earth, while typical soft data are geophysical measurements. Two ways of including soft data are discussed: through a probabilistic approach or through an inverse modeling approach.

Chapter 8 Modeling Structural Uncertainty: the Earth also consists of discrete planar structures such as a topography, faults and layers. To model these we often use a modeling approach tailored specifically for such structures that is not easily captured with a variogram, object or 3D training image approach. In this chapter the basic modeling approach to defining individual faults and layers and methods of combining them into a consistent structural model are discussed. Since structures are often interpreted
from geophysical data, the various sources of uncertainty for structural models and how they can be consistently constructed are discussed.

**Chapter 9 Visualizing Uncertainty:** because of the large uncertainty in Earth modeling and the many sources of uncertainty present, as well as the large amount of possible alternative Earth models that can be created, there is a need to get better insight into the integrated model of uncertainty through graphical representation. Some recently developed techniques based on distances to represent complex models uncertainty in simple 2D plots are discussed.

**Chapter 10 Modeling Response Uncertainty:** modeling uncertainty of the Earth by itself has little relevance in terms of the practical decision it is necessary to make. Instead, these models are used to evaluate certain scenarios or, in general, certain response functions, such as the total amount of contaminant, the best location for the sampling, the total amount of carbon dioxide that can be injected without risk, the best place to store nuclear waste and so on. This calls often for evaluating Earth models through CPU-expensive transfer functions, such as flow simulators, optimization codes, climate models and so on, that can take hours or days to run. A few techniques are presented for assessing response uncertainty that can deal, through model selection, with the issue of CPU cost in mind.

**Chapter 11 Value of Information:** before taking any data, such as costly drilling, sampling surveys or geophysical and remote sensing data, it can be quite useful to assess the value of taking such data. Such value will necessarily depend on the given model of uncertainty. Often, the more uncertain one is prior to taking any data, the more valuable the data may be. In this chapter, techniques are discussed for assessing such value of information in a formal decision analysis framework with a spatial context in mind.

**Chapter 12 Case Study:** the book concludes with a case study in value of information regarding a groundwater contamination problem. The aim of this case study is to illustrate how the various elements in this book come together: decision analysis, 3D modeling, physical modeling and sensitivity analysis.

**Further Reading**

