University of Stavanger

FACULTY OF SCIENCE AND TECHNOLOGY

MASTER'S THESIS

Study programme/specialisation: **Petroleum Engineering /** **Reservoir Engineering**	Spring semester, 2021 Open
Author: **Stacia Elvaretta**	
Supervisor(s): **Reidar B. Bratvold, Aojie Hong**	
Title of master's thesis: **Value of Information Analysis in CO_2 Sequestration Projects**	
Credits (ECTS): 30	
Keywords: Value of Information CO_2 Storage CCS Decision Analysis Monte Carlo Simulation Simulation-Regression	Number of pages: 62 + supplemental material/other: 0 Stavanger, 15th June 2021

Acknowledgements

This thesis concludes my master's study in Petroleum Engineering – Reservoir Engineering program at University of Stavanger. I would like to express my gratitude to those who have helped and supported me during this thesis project.

First and foremost, I would like to thank my supervisors, Professor Reidar B. Bratvold and Professor Aojie Hong, for their impeccable advice, vast knowledge, kindness, and encouragement during this project and my master's study. It has been a pleasure and wonderful experience to work together with the experts in the field of decision analysis and programming.

My sincere thanks to Amine Tadjer, for his valuable contribution in sharing knowledge and advice which was very helpful in the initial process of my thesis.

I owe my deepest gratitude to my parents and family for providing me the best educations and giving me endless support throughout my studies and life. I would also give my special thanks to Finn Haakonsen and my friends, I am truly grateful for their help and encouragement throughout the entire process.

Abstract

As one of the main risks related to CO_2 storage is leakage, it is important to consider the impact of the reservoir uncertainties on the long-term CO_2 migration and leakage which affecting the decisions in CO_2 storage project. Value of information (VOI) is an effective means to assess the maximum a decision maker should pay for gathering and assessing data and information in a given decision context, in order to achieve a high-quality decision.

This study introduces, illustrates, and discusses the role of VOI analysis in deciding whether to accept a contract for a CO_2 storage project. A decision problem is framed where a company that provides CO_2 storage as commercial services plans to sign a contract to store CO_2 at the Utsira Formation on the Norwegian Continental Shelf. The evaluated reservoir uncertainties include permeability, porosity, caprock elevation, aquifer conditions (pressure and temperature). A key contribution of this study is to introduce and discuss a workflow that is useful for a preliminary evaluation of information gathering schemes in CO_2 storage projects carried out before injection commences, with the purpose of achieving cost-effectiveness and high-quality decisions. Moreover, it demonstrates the VOI application with multiple objective values: to maximize the Net Present Value (NPV) and to minimize the long-term leakage. The proposed workflow involves the determination of the material variables, the use of swing weighting, CO_2 injection simulation, and a simulation-regression approach to approximate the posterior value that is used to calculate the VOI. Finally, the study also investigates how sensitive the decision model and the objective values are to variations in certain parameters. For the decision situation framed in this study, the VOI analysis can be used to inform the company's decision of whether to gather and assess more data and information before signing a CO_2 storage contract.

Table of Contents

List of Figures

List of Tables

Nomenclature

CCS	-	Carbon Capture and Storage
CDF	-	Cumulative Distribution Function
CO_2	-	Carbon Dioxide
GHG	-	Greenhouse Gases
GRF	-	Gaussian Random Field
IID	-	Independent and Identically Distributed
KNN	-	***k*** – Nearest Neighbors
MCS	-	Monte Carlo Simulation
MRST	-	Matlab Reservoir Simulation Toolbox
MSE	-	Mean Squares Error
NPD	-	Norwegian Petroleum Directorate
NPV	-	Net Present Value
OLS	-	Ordinary Least-Squares
PoV	-	Posterior Value
PV	-	Prior Value
RF	-	Random Forest
VOI	-	Value of Information

Chapter 1 – Introduction

The carbon dioxide (CO_2) concentration in the world is one of the causes of global warming and climate change. The global emissions trend of Greenhouse Gases (GHGs) has steadily increased since the beginning of the 21st century. This is mainly due to human activities such as burning fossil fuels and deforestation. According to the Emission Database for Global Atmospheric Research (Crippa et al., 2020), the global CO_2 emissions were 38,016.6 Mton in 2019, an increase of 68% from 1990. Without climate change mitigation policies to significantly reduce the emissions of GHGs by 2030, we are likely to see endless crises and irreversible losses for the most vulnerable people, societies, and ecosystems (IPCC, 2018).

This issue has motivated global effort of the Paris Agreement by setting a global framework for avoiding dangerous climate change by limiting global warming to below 2°C and pursuing efforts to limit it to 1.5°C. The carbon capture and storage (CCS) technology is considered an essential element in meeting the target set by the Paris Agreement (United Nations, 2015) on reducing GHG emissions by at least 40% by 2030 compared to 1990. CCS can reduce CO_2 emissions from the extraction of resources, the production and use of fuels, and the generation of electricity by capturing produced CO_2, transporting and storing it in a safe place from the atmosphere. Possible storage places are geological storages such as oil and gas reservoirs, deep saline formations, and un-minable coal beds. The Utsira Formation is one of the deep saline-formations that has a large potential capacity to store CO_2 as a part of the large-scale deployment of CCS in the North Sea region (Norwegian Petroleum Directorate, 2019).

One of the main risks related to CO_2 storage is leakage and it is therefore important to consider the effects of reservoir uncertainties on the long-term migration and leakage. Acquiring new information may reduce uncertainties in the reservoir parameters. However, reducing uncertainties does not necessarily alter any decisions in a CO_2 storage project; furthermore, information-gathering is likely to be expensive and time-consuming. It is accordingly important to justify the costs of the information-gathering by assessing its impact on the decisions that need to be made. Value of information (VOI) is very effective and suited for this purpose. Bratvold et al. (2009) determined the VOI as the most the decision maker should pay for the additional information; it is equal to the difference between the certain equivalent with information and the certain equivalent without information.

Eidsvik et al. (2015) evaluated VOI analysis integrated with spatial modeling and decision analysis. Several examples are used to illustrate the applicability of VOI for domains such as energy, geophysics, geology, mining, and environmental science. Puerta-Ortega et al (2013) introduced and used a decision-analytic framework to quantify VOI of permeability data in a United States based CCS project. Sato (2011) illustrated the application of VOI analysis in decision-making problems with discrete probabilities and continuous probability distributions. The study also demonstrated the improvement in the information accuracy increases the VOI.

To address the computational tasks of VOI analysis which can be daunting in spatial decision situations, a regression-based approximation approach was introduced for VOI assessment (Eidsvik et al., 2015, 2017; Dutta et al., 2019). The method involves Monte Carlo simulation (MCS) followed by linear regression to fit the conditional expectation expression that is required in VOI analysis. A recent study by Anyosa et al. (2021) discussed monitoring schemes for leakage detection and assessed the VOI of the seismic monitoring schemes. Their workflow involved the simulation of CO_2 saturations and amplitude versus offset (AVO) attributes, as well as the use of statistical modeling (regression-based approximation approach) to inform the decision about the best monitoring time and to assess the VOI.

Allen et al. (2018) investigated how uncertainty in geological models affects CO_2 storage capacities in large-scale saline aquifers. Their focus was on uncertainties in top-surface elevation, rock properties (porosity, permeability), fault transmissibility, and aquifer conditions (pressure and temperature). In this thesis study, the approach used by Allen et al. (2018) is extended by applying VOI analysis on both individual and combined uncertain variables such as permeability, porosity, caprock elevation, aquifer conditions (pressure and temperature) that can affect the decision in a CO_2 storage project. The objective of this study is to use VOI analysis to determine whether the company should invest in additional information to reduce reservoir uncertainties prior to signing a CO_2 storage contract. The reservoir model is based on the Utsira Formation on the Norwegian Continental Shelf. A decision problem is framed where the decision maker has the option to gather new information regarding reservoir uncertainties before choosing whether to sign a CO_2 storage contract. The company has identified two objectives: maximize the Net Present Value (NPV) that the company will get over 60 years of CO_2 storage injection and minimize the long-term CO_2 leakage, where each objective has a different weight. The material variables are determined by evaluating the Tornado Diagram of each uncertainty variable on the objective value. The Matlab

Reservoir Simulation Toolbox (MRST) is used to vary the material variables and simulate 100 realizations of CO_2 injection. The outputs from the simulation are used to estimate the VOI using the simulation-regression approach. Furthermore, three regression methods are used for the assessment: linear regression, random forest, and ***k***-nearest neighbors. A sensitivity analysis is performed to evaluate the impact of carbon price uncertainty, leakage value constraint, and different injection periods on the VOI estimations.

A key contribution of this study is to introduce and discuss a workflow that is useful for an initial evaluation of information gathering schemes in CO_2 storage projects, with the purpose of achieving cost-effectiveness and high-quality decisions. Moreover, it demonstrates the VOI application with multiple objective values and weights which can be adjusted according to the company requirements.

This thesis consists of 5 chapters. Following this introduction, Chapter 2 explains the key uncertainties in the CO_2 storage migration, the basics of MCS, and CO_2 injection simulation. Chapter 3 describes the theory of VOI analysis and its examples. Chapter 4 contains a case study of VOI analysis in a CO_2 storage project at the Utsira Formation. Finally, Chapter 5 provides the conclusions of this thesis and future recommendations.

Chapter 2 – CO_2 Storage Uncertainties and Simulation

Zhang and Song (2014) described that CO_2 can be stored as compressed gas, liquid, or supercritical phase via four major geological carbon sequestration mechanisms. When injection CO_2 into a reservoir, most of the injected CO_2 will be initially be sequestered in a mobile phase, free to move laterally or migrate vertically towards the caprock of the aquifer (structural or stratigraphy trap). Second, a residual gas trap can be formed when formation water encroaches or invades the CO_2 plume. Third, dissolving CO_2 partially into the aqueous phase, leading to solubility trapping, and finally, it can also react with native minerals which resulting in mineral trapping.

Chapter 2 explains the impact of key uncertainties on long-term CO_2 storage migration, the basics of MCS, and provides an introduction to CO_2 injection simulation which will be used in the VOI analysis.

2.1 CO_2 Storage Uncertainties

Saline aquifers have been identified as having the largest storage capacity amongst all geologic formations (Temitope et al., 2016). The storage capacity and plume dynamics are particularly important when selecting aquifers as potential storage sites. There are several reservoir parameters that affect capacity estimation and long-term plume migration including top-surface elevation, rock properties (porosity, permeability), and aquifer initial conditions (pressure and temperature).

Top-Surface Elevation. Allen et al. (2018) observed that CO_2 migrates like a thin plume underneath the undulating caprock of the aquifer due to the large density difference between CO_2 and brine. Consequently, the aquifer's surface topography is the main parameter dictating the migration direction and the overall plume shape. The caprock topography is often measured using seismic surveys, therefore the top-surface elevation uncertainty is likely related to aquifer depth.

Rock Properties. Allen et al. (2018) conducted research that shows the impact of porosity on the total rock volume that occupied by the plume; the higher porosity reduces the total rock volume which slows down the plume migration so that it does not travel as far upslope and also impacts the amount of CO_2 trapped within the pore grains. Furthermore, permeability also influences the plume migration by changing its speed which results in a thinner plume that reaches farther upslope.

Aquifer Initial Conditions. The influence of pressure and temperature on CO_2 density

which further impacts the plume migration and storage capacity has been discussed previously in Allen et al. (2018). They concluded that with increasing pressure from a hydrostatic condition, CO_2 density increases resulting in an increased CO_2 volume in the structural traps. On the other hand, a warmer aquifer decreases CO_2 density resulting in lower CO_2 storage volume.

2.2 Monte Carlo Simulation

Monte Carlo simulation (MCS) is a methodology for analyzing uncertainty which is very popular among petroleum engineers and geoscientists (Bratvold and Begg, 2010). Bonate (2001) illustrated that MCS differs from deterministic simulation where the variability of the model parameters is included in the model and its long-term impact is examined. Bratvold and Begg (2010) summarized the procedure of the MCS in Figure 2.1. First, to perform the MCS, the objectives of the model need to be appropriately defined. Second, the decision maker's beliefs about the uncertainty variables are represented by the probability distributions and defined as inputs in MCS. For instance, a normal distribution with mean and standard deviations values. The variables are assumed to be independent in MCS, therefore dependencies need to be included in the model if two or more variables are dependent on each other. Third, MCS uses a random number to sample each input probability distribution and computes the output variables. The result of each trial will be stored. And finally, this process is repeated many times and the stored results will be used to generate histogram and statistics of the output variables. The distributions of the output variables represent the probabilities of their occurrence. The greater number of samples taken, the more representative the output distributions.

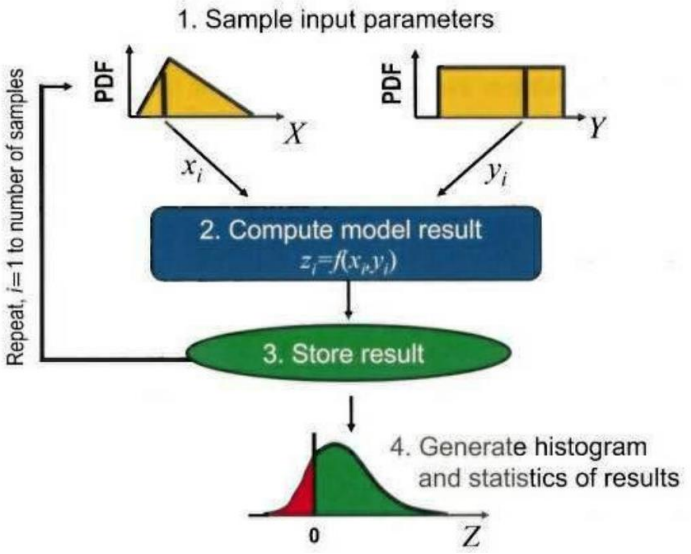


Figure 2.1 Schematic of Monte Carlo simulation procedure (Bratvold and Begg, 2010)

Bratvold and Begg (2010) also described several advantages of MCS as follows:

- The technique does not require any approximation of the input variable distributions;
- Correlation and dependencies can be modeled easily;
- It provides solutions where analytical solutions are not available and it is easier to understand than complex mathematical equations;
- Compared to the analytical approach, MCS has a lower chance of making errors in solving the problem;
- Commercial software is available for the tasks involved in the simulation;
- Complex and nonlinear mathematics can be included with no additional difficulty;
- The results of MCS are likely to be accepted by both analysts and decision-makers as it is recognized as a valid technique;
- It is easy to investigate the behavior of the model and make quick changes to the model.

2.3 Gaussian Random Field

Random fields are widely used to model uncertainty. Generally, statistical measures, e.g. mean, standard deviation, and probability distribution, are necessary to describe in the random field. Liu et al. (2019) described that among the commonly used probability functions, the Gaussian distribution is the most fundamental distribution in statistics which has the advantage of flexible shape in data fitting by controlling the mean and variance. In addition, the Gaussian distribution has well-established tools for computer simulations, e.g. Matlab, R packages, and other commercial software, that can easily generate Gaussian random variables and fields.

A random field $X = \{X(t), t \epsilon T\}$ is a family of random variables with values in state space S, where T is the parameter set. If $T \subseteq \mathbb{R}^N$ and $S = \mathbb{R}^d (d \geq 1)$, then X is called an (N, d) random field. Gaussian random field (GRF) is a random field involving Gaussian probability distribution of the variables. A GRF, $X(s)$, is defined by a mean function $\mu(s) = E(X(s))$ and covariance function $C(s,t) = Cov(X(s), X(t))$. The mean value μ is a constant over the field, while the covariance function only depends on the distance of two spatial locations in an isotropic case (Liu et al., 2019). Furthermore, there are several choices for the covariance functions such as Matérn covariance, linear covariance, circular covariance, spherical covariance, Whittle covariance.

In this study, the geological model realizations of reservoir uncertainty are generated using approximate GRF from SINTEF (2016b), which convolving a field of independent normally

distributed variables with a Gaussian kernel. Approximation GRF is one of simplified methods for generating geostatistical realizations.

2.4 CO_2 Injection Simulation

Lie (2019) provides an introduction and several examples using the Matlab based Reservoir Simulation Toolbox (MRST), an open-source software for reservoir simulator and modeling. SINTEF (2016a) developed an add-on module of MRST, i.e. MRST-CO2lab which offers a set of open-source simulators and workflow tools that were designed for modeling geological CO_2 storage. The module provides simplified access to publicly available datasets from the Norwegian CO_2 Storage Atlas and a unified toolchain that enables the user to visualize migration paths and compute upper theoretical bounds on structural, residual, and solubility trapping. The CO2lab simulator is using a vertical-equilibrium formulation that can be used to analyze pressure build-up and plume migration. The vertical-equilibrium method improves the simulation time and consumes significantly less memory than conventional 3D simulators.

In this study, the CO_2 injection will be simulated using one injection well, adjusting injection rate and period, and framing several uncertainties. The outputs of the simulations are a forecast of CO_2 injection, CO_2 leakage amount, and average variable value in the geological model for each realization.

Chapter 3 – Value of Information Analysis

In this chapter, the theory of value of information will be introduced with examples. Value of information can be computed using analytical methods (conjugate priors) and approximate methods (e.g. gridding, Markov Chain Monte Carlo, or regression).

3.1 Value of Information

Any information-gathering activity must meet the following three criteria if it is to be valuable: relevant, material, and economic. In Eidsvik et al. (2015), these criteria are illustrated by the "pyramid of conditions" in Figure 3.1.

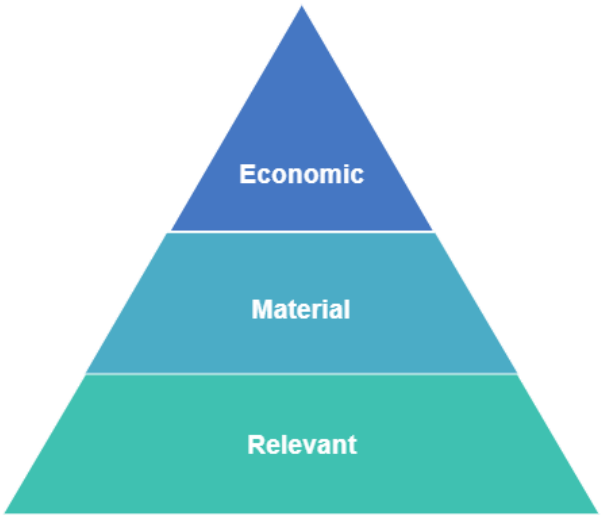


Figure 3.1 The pyramid of conditions information-gathering schemes (Eidsvik et al., 2015)

The first criteria: ***relevant***, the information must have an impact on the decision maker's beliefs about the distinction of interest. For example, while making a decision about whether a company should drill a well in the field, information about reservoir properties from logging data along the well location will be very useful in determining the hydrocarbons present at the well location. The second criteria: ***material***, the information-gathering outcome must have the possibility of changing a decision the decision maker might otherwise make. For instance, the probability of high permeability given that the test indicates "high permeability" is not equal to the probability of high permeability given that the test indicates "low permeability". Finally, the last criteria: ***economic***, the cost of obtaining the information must be less than its value. The three criteria are presented as a pyramid since higher conditions cannot be satisfied until the conditions at the bottom of the pyramid are met.

The VOI for uncertainty $\boldsymbol{x}$ and a decision $\boldsymbol{a}$ is the value at which the decision maker is indifferent between acquiring the information making the decision with the current (prior) information only (Eidsvik et al., 2015). If the decision maker is risk-neutral, the certain equivalents are equal to expected values and the VOI is given by the difference between the expected value with information and the expected value without information (Bratvold et al., 2009).

Assume a corporation is facing a decision making situation with some alternative $\boldsymbol{a} \in \boldsymbol{A}$, where $\boldsymbol{A}$ denotes the set of possible alternatives and there is an uncertainty $\boldsymbol{x}$ associated with the decision situation. After $\boldsymbol{x}$ is observed and alternative $\boldsymbol{a}$ is chosen, the decision maker acquires value $\boldsymbol{v}(\boldsymbol{x},\boldsymbol{a})$. The prior probability distribution $\boldsymbol{p(x)}$ represents the decision maker's beliefs about the uncertainty $\boldsymbol{x}$.

Motivating problem example:

To illustrate the basic ideas, consider a simplified example of a decision situation inspired by Puerta-Ortega et al (2013) where a company that sells CO_2 storage services is considering a CO_2 storage contract to inject 4 Mt CO_2 per year for three years. The decision maker has two alternatives, $\boldsymbol{a} \in \{0,1\}$, i.e. sign the contract ($\boldsymbol{a} = 1$) or turn down the contract ($\boldsymbol{a} = 0$). There are two possible scenarios faced by the operator, $\boldsymbol{x} \in \{0,1\}$, where:

- S1 = The targeted formation can store the required CO_2 amount of 4 Mt ($\boldsymbol{x} = 0$), and
- S2 = The targeted formation fails to store the required CO_2 amount of 4 Mt due to leakage ($\boldsymbol{x} = 1$).

The prior belief about the uncertainty is $\boldsymbol{p}(\boldsymbol{x} = 0) = 0.3$, $\boldsymbol{p}(\boldsymbol{x} = 1) = 0.7$. The decision maker obtains no value if the contract is turned down, $\boldsymbol{v}(\boldsymbol{x},\boldsymbol{a} = 0) = 0$, while the value of each possible scenario depends on the profit for inject the required CO_2 amount which is 60.7 in some monetary unit and the penalty fee for the CO_2 leakage of 30 monetary units. The following value function $\boldsymbol{v}(\boldsymbol{x},\boldsymbol{a})$ is calculated:

- $\boldsymbol{v}(0,1) = 60.7$
- $\boldsymbol{v}(1,1) = -30$
- $\boldsymbol{v}(\boldsymbol{x},0) = 0$

3.1.1 Prior Value

Assuming the decision maker is risk-neutral who choose alternatives based on expected value maximization, the prior value is given by:

$$PV = \max_{a \in A}\{E|v(x,a)\} = \max_{a \in A}\left\{\int_x v(x,a)\, p(x)\, dx\right\}. \quad (3.1)$$

Using the previous CO_2 storage problem, the expected values for each alternative are:

$$E[v(x, a = 1)] = 0.3 \cdot 60.7 + 0.7 \cdot -30 = -2.79,$$

$$E[v(x, a = 0)] = 0.$$

Assuming a discrete sample for the data, the prior value is calculated using equation (3.1) with:

$$PV = max\{E|v(x, a = 1), E|v(x, a = 0)\} = \max\{-2.79, 0\} = 0.$$

From the above analysis, the company should not sign the contract as the expected value is 0. Figure 3.2 illustrates the decision tree of the decision situation above. The rectangular node represents a decision node and the circular node represents a probability node. Branches from decision nodes indicate the decision maker's alternatives, whilst branches from probability nodes indicate possible uncertainty outcomes.

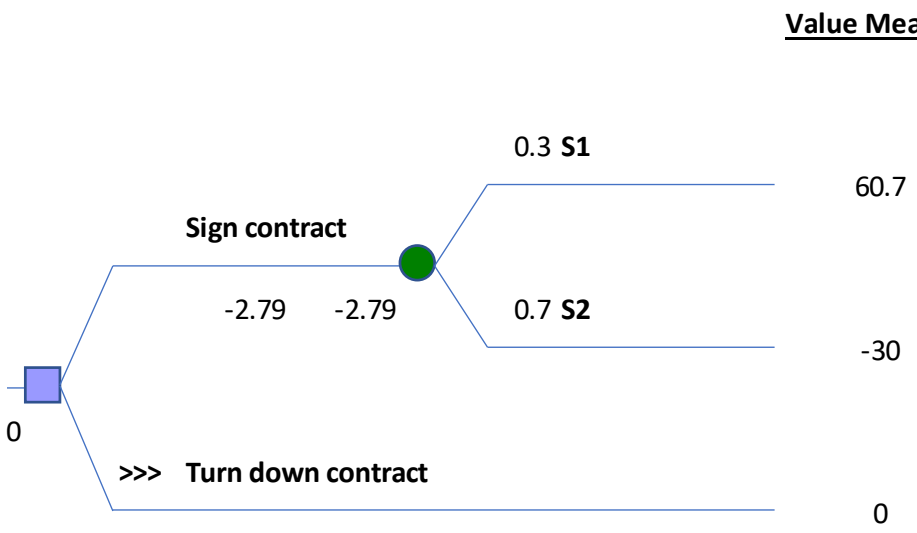


Figure 3.2 The CO_2 storage problem's decision tree without additional information

3.1.2 Posterior Value

Suppose the decision maker can obtain additional information prior to making the contract signing decision. The data are represented by a likelihood probability $\boldsymbol{p(y|x)}$, the conditional relationship between the information $\boldsymbol{y}$ and the uncertain variable $\boldsymbol{x}$. The information-gathering scheme will affect the likelihood function.

Perfect Information.

A predictor's information is perfect if it is always correct. In the hypothetical case of perfect information gathering, the posterior value (***PoV***) for a risk-neutral decision maker is defined as:

$$PoV(x) = \int_x \max_{a \in A}\{E[v(x,a)|y]\}\, p(x)\, dx. \tag{3.2}$$

The VOI is computed as the difference between posterior and prior value:

$$VOI(x) = PoV(x) - PV. \tag{3.3}$$

Although perfect information is rare to be obtained, the concept of perfect information can be very useful as a theoretical model when evaluating the information because it provides the upper value bound of any imperfect information.

Imperfect Information.

Most information sources in Earth sciences are arguably imperfect sources of information (Eidsvik et al., 2015). Information gathering schemes are often noisy measurements due to human error, instrument inaccuracy, etc. Therefore, the noise should be taken into account in the VOI analysis. The posterior value (***PoV***) can be defined by:

$$PoV(y) = \int_y \max_{a \in A}\{E[v(x,a)|y]\}\, p(y)\, dy. \tag{3.4}$$

The pre-posterior, or total probability ***p(y)*** in Eq. (3.4) is given by:

$$p(y) = \int_x p(y|x)p(x)dx, \tag{3.5}$$

and the expectation by

$$E[v(x,a)|y] = \int_x v(x,a)p(x|y)dx, \tag{3.6}$$

where the posterior p(x|y) is calculated using Bayes' rule:

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}.$$

For the risk-neutral decision maker, the VOI is the difference between posterior and prior value:

$$VOI(y) = PoV(y) - PV. \tag{3.7}$$

Consider again the previous CO_2 storage problem, the posterior value of the decision situation with perfect information is calculated by:

$$PoV(x) = \sum_x \max_{a \in A}\{E[v(x,a)|y]\}\, p(x)$$
$$= max\{v(0,1), v(0,0)\}p(x=0) + max\{v(1,1), v(1,0)\}p(x=1)$$
$$= max\{60.7,0\}(0.3) + max\{-30,0\}(0.7) = 18.21.$$

Then the value of the perfect information can be calculated:

$$VOI(x) = PoV(x) - PV = 18.21 - 0 = 18.21.$$

From the above analysis, the company can conclude that they should not pay more than 18.21 monetary units to obtain any additional information about the given uncertainty ***x***.

More realistically, the information gathered will be imperfect but can still help determine which information gathering will be more valuable. Assume that the company can perform a test indicating whether there is CO_2 leakage in the reservoir (***y*** = 1) or not (***y*** = 0). Assume that the test has an accuracy of 80%, the likelihood is given by ***p***(***y*** = 0|***x*** = 0) = ***p***(***y*** = 1|***x*** = 1) = 0.8.

First, the probability of each of the outcomes from the test are calculated:

$$p(y=0) = \sum_x p(y|x)p(x) = \mathrm{p(y=0|x=0)p(x=0)} + \mathrm{p(y=0|x=1)p(x=1)}$$
$$= 0.8 \cdot 0.3 + 0.2 \cdot 0.7 = 0.38,$$
$$p(y=1) = \sum_x p(y|x)p(x) = \mathrm{p(y=1|x=0)p(x=0)} + \mathrm{p(y=1|x=1)p(x=1)}$$
$$= 0.2 \cdot 0.3 + 0.8 \cdot 0.7 = 0.62.$$

Second, the posterior for different outcomes of ***x*** given the different outcomes of ***y*** are calculated:

$$p(x=0|y=0) = \frac{p(y=0|x=0)p(x=0)}{p(y=0)} = \frac{0.8 \cdot 0.3}{0.38} = 0.63$$
$$p(x=1|y=0) = 1 - p(x=0|y=0) = 1 - 0.63 = 0.37$$
$$p(x=1|y=1) = \frac{p(y=1|x=1)p(x=1)}{p(y=1)} = \frac{0.8 \cdot 0.7}{0.62} = 0.90$$
$$p(x=0|y=1) = 1 - p(x=1|y=1) = 1 - 0.90 = 0.10.$$

The probability trees of above calculations are illustrated in Figure 3.3. The next step is to calculate the expected values given the possible test outcomes:

- $E[v(x,a=0)|y=0] = \sum_x v(x,a=0)p(x|y=0)$
$$= v(x=0,a=0)p(x=0|y=0) + v(x=1,a=0)p(x=1|y=0)$$
$$= 0 \cdot 0.63 + 0 \cdot 0.37 = 0$$

- $E[v(x,a=1)|y=0] = \sum_x v(x,a=1)p(x|y=0)$

$$= v(x=0,a=1)p(x=0|y=0) + v(x=1,a=1)p(x=1|y=0)$$

$$= 60.7 \cdot 0.63 + (-30) \cdot 0.37 = 27.28$$

- $E[v(x,a=0)|y=1] = \sum_x v(x,a=0)p(x|y=1)$

$$= v(x=0,a=0)p(x=0|y=1) + v(x=1,a=0)p(x=1|y=1)$$

$$= 0 \cdot 0.1 + 0 \cdot 0.9 = 0$$

- $E[v(x,a=1)|y=1] = \sum_x v(x,a=1)p(x|y=1)$

$$= v(x=0,a=1)p(x=0|y=1) + v(x=1,a=1)p(x=1|y=1)$$

$$= 60.7 \cdot 0.1 + (-30) \cdot 0.9 = -21.22.$$

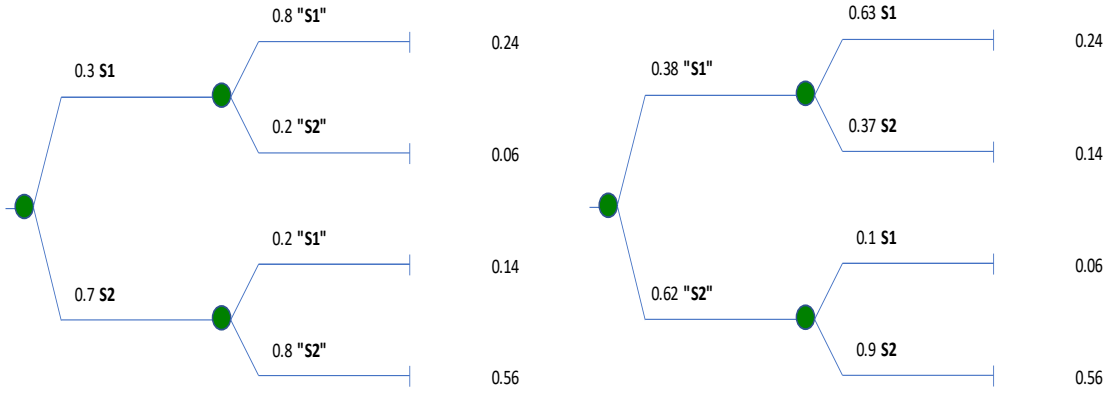


Figure 3.3 Probability trees in the assessed and inferential forms for the information-gathering

Then the posterior value of the decision situation with imperfect information is calculated by:

$$PoV(y) = \sum_y \max_{a \in A}\{E[v(x,a)|y]\}\, p(y)$$

$$= max\{E[v(x,a)|y=0]\}p(y=0) + max\{E[v(x,a)|y=1]\}p(y=1)$$

$$= max\{0,27.28\}(0.38) + max\{0,-21.22\}(0.62) = 10.37.$$

Finally, the value of the imperfect information is calculated:

$$VOI(y) = PoV(y) - PV = 10.37 - 0 = 10.37.$$

In conclusion, the company should conduct the test with 80% accuracy if the cost of the test is less than 10.37 monetary units. Figure 3.4 illustrates the decision situation with the imperfect information. The optimal decision is conditional on the outcome of the test (marked arrows in Figure 3.4) where the probability of alternative 2 given test "S2" is higher than the probability of alternative 1 given test "S1".

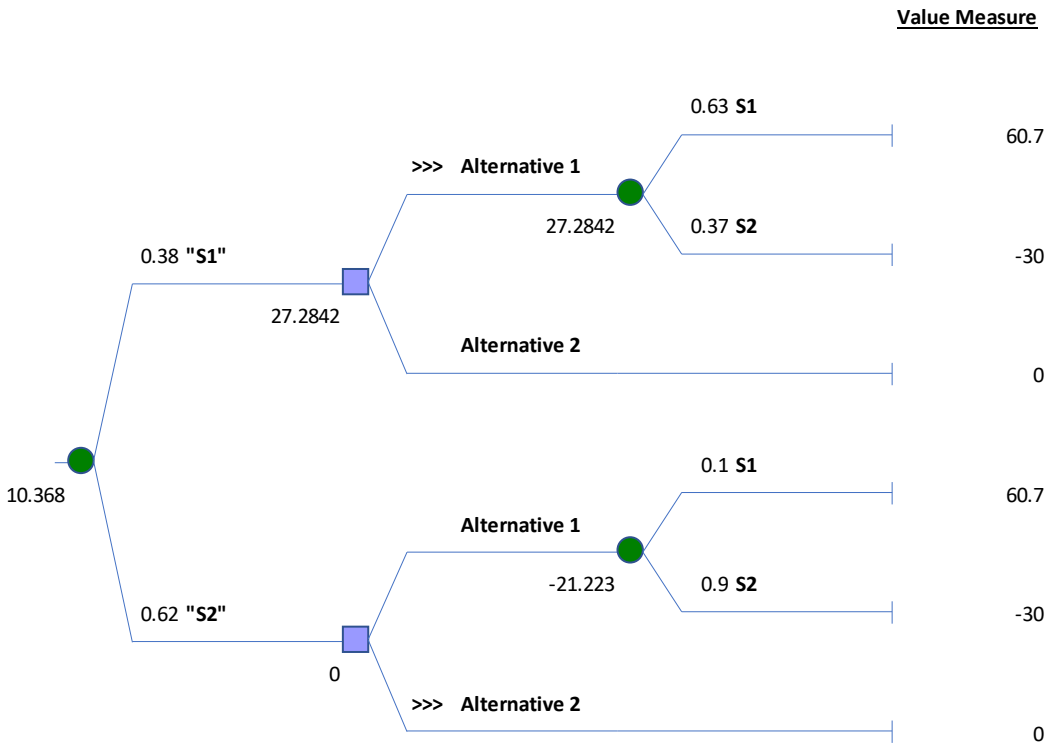


Figure 3.4 The decision tree for the CO_2 storage problem with imperfect information

A sensitivity analysis is conducted to show the relationship between the VOI and the accuracy of the test (Figure 3.5). It shows that the VOI decreases as the test accuracy decreases and that any test with an accuracy below 50% cannot add value.

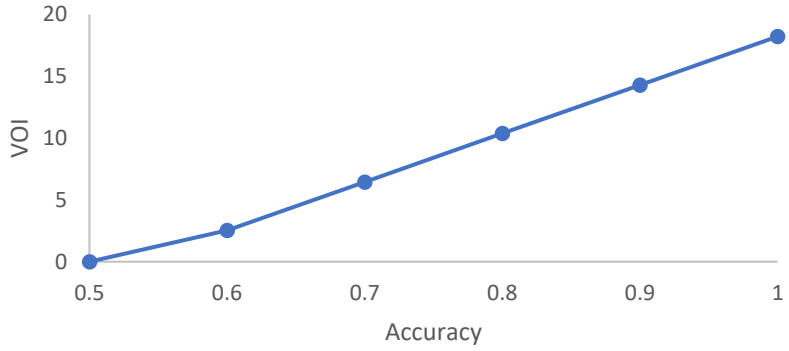


Figure 3.5 The sensitivity analysis accuracy vs VOI for the CO_2 storage problem

3.2 Simulation-Regression Approach

In equation (3.4), the posterior value is calculated from maximum expected values conditional on the data outcomes (inner expression) and then weighted over all possible data (outer expression). The general applicable computational approach for the posterior value is double

Monte Carlo sampling (Eidsvik et al., 2017) that involves the use of MCS sampling of data $\boldsymbol{y}$ for the outer expression and another round of MCS sampling for the inner expression, although it is time-consuming. Simulation-regression approach is considered to be one solution to overcome the computational issue by constructing efficient approximations for the VOI.

In this study, VOI will be estimated using the simulation-regression approach. In the simulation-regression approach, VOI is computed by simulating the model parameters, the data and prospect values, then regressing the prospect values on the data (Eidsvik et al., 2015, 2017; Dutta et al., 2019). The steps involved in this method are as follows:

(1) From the prior probability $\boldsymbol{p(x)}$, generate $\boldsymbol{B}$ realizations of parameter $\boldsymbol{x}$, given by $\boldsymbol{x^1}$,…, $\boldsymbol{x^B}$, where the parameter $\boldsymbol{x}$ in this study are porosity, permeability, caprock elevations, pressure, and temperature. The realization method of each uncertainty $\boldsymbol{x}$ will be explained in the next chapter;

(2) For each realization of $\boldsymbol{x}$, $\boldsymbol{x^b}$, and for each alternative $\boldsymbol{a}$, conduct the forward modeling using MRST to obtain the corresponding samples of data ($\boldsymbol{y^b}$) and prospect values ($\boldsymbol{v_a^b}$), such as the average value of parameter $\boldsymbol{x^b}$ in the geological model and the objective value in our case, respectively. The objective value is a result of the uncertainty in the prior model, it includes the leakage amount which is directly obtained from MRST results, and the NPV which is calculated with the NPV functions explained in the next chapter;

(3) For each alternative $\boldsymbol{a}$, fit a model regressing the values on the dataset using the selected regression model:

$$\hat{v}_a^b = F_a(y^b), \tag{3.8}$$

which approximates the conditional expectation $E[v(x,a)|y^b]$. The optimal inputs for the regression model are defined using 10-fold cross-validation;

(4) According to Eidsvik et al. (2017), the prior value is approximated to ensure that the VOI is always non-negative. The prior value is given by:

$$PV = \max_{\boldsymbol{a} \in \boldsymbol{A}} \left[\frac{1}{B} \sum_{b=1}^{B} v_a^b \right]; \tag{3.9}$$

(5) Approximate the posterior value with information by:

$$\begin{aligned} PoV(y) &= \int_y \max_{\boldsymbol{a} \in \boldsymbol{A}} \{E[v(x,a)|y]\}\, p(y)\ dy \\ &\approx \frac{1}{B} \sum_{b=1}^{B} \max_{\boldsymbol{a} \in \boldsymbol{A}} E[v(x,a)|y^b] \approx \frac{1}{B} \sum_{b=1}^{B} \max_{\boldsymbol{a} \in \boldsymbol{A}} \hat{v}_a^b ; \end{aligned} \tag{3.10}$$

(6) Finally, VOI is approximated by:

$$VOI = PoV(y) - PV. \tag{3.11}$$

The approximation can be improved by increasing the sample size $\boldsymbol{B}$ although it also depends on the choice of the regression model. Furthermore, the VOI result should be compared with the price of the information-gathering, the new information will be worthwhile for decision making if the price is lower than its VOI.

The value of imperfect information is reflected by adding some measurement error as a noise value that changes the value $\boldsymbol{y^b}$ before creating the regression model, while the value of perfect information is calculated without noise, i.e. noise = 0%. In this study, the noise value is represented by a normal distribution MCS sample with a mean of $\boldsymbol{y^b}$ and standard deviation of average variable realization value multiplied by the percentage error, $\left(\frac{1}{B}\sum_{b=1}^{B} \boldsymbol{y^b}\right) \cdot \frac{\%\, noise}{100}$. The use of MCS sampling can create spurious correlation effects and biased results due to different results in each repetition. Thus, the calculations of VOI with imperfect information are repeated 1000 times and averaged to reduce bias and provide accurate approximations.

3.3 Regression Methods

The regression methods used in this study are ordinary least-squares linear regression (OLS), random forest (RF), and $\boldsymbol{k}$-nearest neighbors (KNN) from scikit-learn (Pedregosa et al., 2011). OLS is chosen to fit the linear relationships, while RF and KNN are considered for fitting non-linear relationships. Machine learning algorithms can be controlled by some parameters that are set before the learning process begins, which are commonly named hyperparameters. For optimizing the hyperparameters of the regression model, $\boldsymbol{k}$-fold cross-validation is used by minimizing mean squares error (MSE) of the regression model scores aiming for good accuracy of an estimator. Cross-validation is primarily used to estimate the skill of machine learning model on unseen data, that is, using a limited sample to assess how the expected performance of the model in general when used to make predictions on unused data during the model training (Brownlee, 2018). The scores from $\boldsymbol{k}$-fold cross validation represent the overfitting or underfitting estimate of a machine learning model by randomly dividing the set of observations into $\boldsymbol{k}$ groups (James et al., 2013). In this study, the analysis use $\boldsymbol{k}$=10 becoming 10-fold cross-validation, which is commonly used in the field of applied machine learning as it shows a balance between computational complexity and validation accuracy (Grootendorst, 2019).

Ordinary Least-Squares Linear Regression (OLS).

Linear regression is a standard statistics and machine learning modeling method, it is considered to be the most frequently used statistical technique in practice (Mahaboob et al., 2018). The least-squares method used in Ordinary Least-Squares (OLS) linear regression is straightforward, it uses gradient descent to minimize the residual sum of the squared differences between observed targets in the dataset and estimated output values by the linear approximation.

The input data is the independent variable denoted as ***x*** with ***n*** samples, while the output is the dependent variable denoted as ***y*** with one output for each input. The prediction of the regression method is denoted as $\hat{y}$,

$$\hat{y}_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \varepsilon_i, i = 1,2, \dots, n. \quad (3.12)$$

where $\beta_j, j = 0, 1, 2, \dots, k$ are regression coefficients, *x* is the independent variables, and ε_i is a statistical error, i.e., a random variable that accounts for the failure of the model to fit the data exactly.

Random Forest (RF).

Breiman (2001) defined random forest (RF) as regressions that are formed by growing trees depending on independent and identically distributed (IID) random vectors (Θ), taking the average over ***k*** of the trees $\{\boldsymbol{h}(\boldsymbol{x},\Theta_k)\}$ to improve the predictive accuracy and control the over-fitting condition. The study also described the advantages of the random forest which are that it keeps the benefits achieved by the decision trees and improves the results by using the bagging concept on samples. In this study, RF regression is implemented using scikit-learn (Pedregosa et al., 2011) with the maximum depth of tree adjusted by 10-fold cross-validation. A different number of maximum depth of tree values in the range 1 to 80 is assigned to 10-fold cross-validation.

K-Nearest Neighbors (KNN).

KNN regressor from scikit-learn (Pedregosa et al., 2011) computes each query point based on the ***k*** nearest neighbors. The algorithm of the model predicts the dependent values based on the average responses of the ***k*** closest independent data points in the training set (Hastie et al., 2009). The variable ***k*** is an integer value optimized using 10-fold cross-validation in the application of this study. A different number of neighbors (***k***) values in the range 1 to 80 is considered to perform 10-fold cross-validation and achieve the minimum MSE.

Chapter 4 – Case Study at Utsira Formation

Chapter 4 demonstrates a case study of VOI analysis in a CO_2 storage project at the Utsira Formation on the Norwegian Continental Shelf. The reservoir simulation is done using MRST and the VOI analysis is done using the methodology discussed in chapter 3. This chapter introduces and discusses the Utsira Formation reservoir model, the decision frame and the workflow used in the case study, reservoir simulation and VOI analysis results, sensitivity analysis, and discussions of the overall analysis.

4.1 Utsira Formation Reservoir Model

In this study, the proposed methodology is examined in the Utsira Formation, a saline formation located in the North Sea. This upper Miocene to Lower Pliocene age formation has an average top-surface depth of almost 800 meters (ranging from 300 to 1400 meters) and a critical point of 31°C and 73.8 bars (Allen et al., 2018). Boundaries of the aquifers are considered open which indicates there is communication between the aquifer and another aquifer or the sea bottom that lies adjacent to it. The average porosity and permeability of Utsira Formation are reportedly 0.2112 and 1000 mD in the CO_2 Storage Atlas, Norwegian Petroleum Directorate (2019). Since the heterogeneous rock properties are not available, Allen et al. (2018) created a model using the porosity-depth and porosity-permeability functions from the Stø aquifer model which is given as follows:

$$\emptyset = -7.0461 \cdot 10^{-5} z + 0.32343, \tag{4.1}$$

$$\log(k) = 34.9718\emptyset - 7.246. \tag{4.2}$$

The geological realizations of the reservoir rock properties (porosity and permeability) and caprock elevations are generated using GRF. The porosity realizations used a random field interval value of ±0.1 and standard deviation of 0.65, while the permeability realizations were generated from the logarithmic relationship between permeability and porosity from the same uncertainty model as porosity. The caprock elevation realizations used a random field interval of ±5 m and standard deviation of 0.65.

Pressure and temperature conditions of the reservoir model were computed with the functions of the depth:

$$P = (\rho_w g z + P_s), P_0 = P + \frac{\sum_i^n pv_i P_i}{\sum_i^n pv_i} d/100, \tag{4.3}$$

$$T = T_b + \nabla T(z - z_b), \tag{4.4}$$

Where P_0 is initial hydrostatic pressure, ρ_w is water density, g is gravitational acceleration, z is caprock depth, P_s is surface pressure, pv_i is pore volume of cell $\boldsymbol{i}$, d is the deviation in percent, T is initial caprock temperature, T_b is seafloor temperature, $(z - z_b)$ is depth below seafloor, and ∇T is the thermal gradient in the vertical direction. We are uncertain about the deviation (d) in the initial hydrostatic pressure, equation (4.3), and the thermal gradient (∇T) in the initial caprock temperature, equation (4.4). Both pressure and temperature realizations are generated using normal distribution MCS. The pressure realizations are formed by using a mean and standard deviation of [0, 5]% deviation (d), which corresponds to a standard deviation of roughly ± 13 bars as the model reference pressure is 80 bars. Furthermore, the temperature realizations are generated by sampling the thermal gradients (∇T) with mean and standard deviation of [37.5, 3.36]°C/km.

4.2 Decision Frame and Workflow

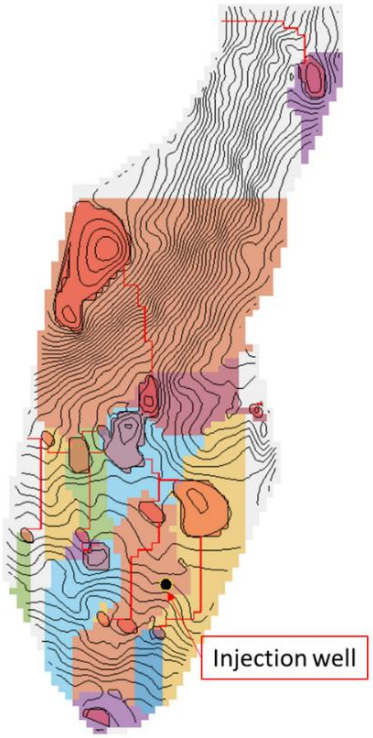


Figure 4.1 The location of injection well in the south of Utsira Formation

Assume a company that offers CO_2 storage as commercial services is going to sign a contract to store CO_2 at the Utsira Formation. The contract requires the company to store 5 Mt CO_2 per year for 60 years with 1 injection well located in the south (Figure 4.1). The company has two alternatives $\boldsymbol{a} \in \{0,1\}$, sign the contract ($\boldsymbol{a}$=1) and turn down the contract ($\boldsymbol{a}$=0). The decision of the company is based on 2 objectives, to maximize the Net Present Value (NPV) over the injection period and to minimize the long-term leakage from the CO_2 storage (post-injection). NPV is calculated from the profit that represents revenue of the company, considering that the company offers CO_2 storage as commercial services in this study case. In other cases where a company that performs CO_2 storage projects also produces CO_2 emissions and has to pay the carbon tax, the profit can be considered as a carbon tax reduction. A carbon tax is a form of carbon pricing and is a charge that was established by governments on emitters for every metric ton of CO_2 they produce to incentivize the investment in CCS and other CO2-emission reducing measures. Furthermore, both objectives are formed in an objective value for each realization. If the company decides to sign the contract, the objective value from the project is a function of several uncertainties related to CO_2 leakage, i.e., top-surface elevation, rock properties (porosity, permeability), and aquifer initial conditions (pressure and temperature). Figure 4.2 illustrates the influence diagram, including dependencies, for this analysis.

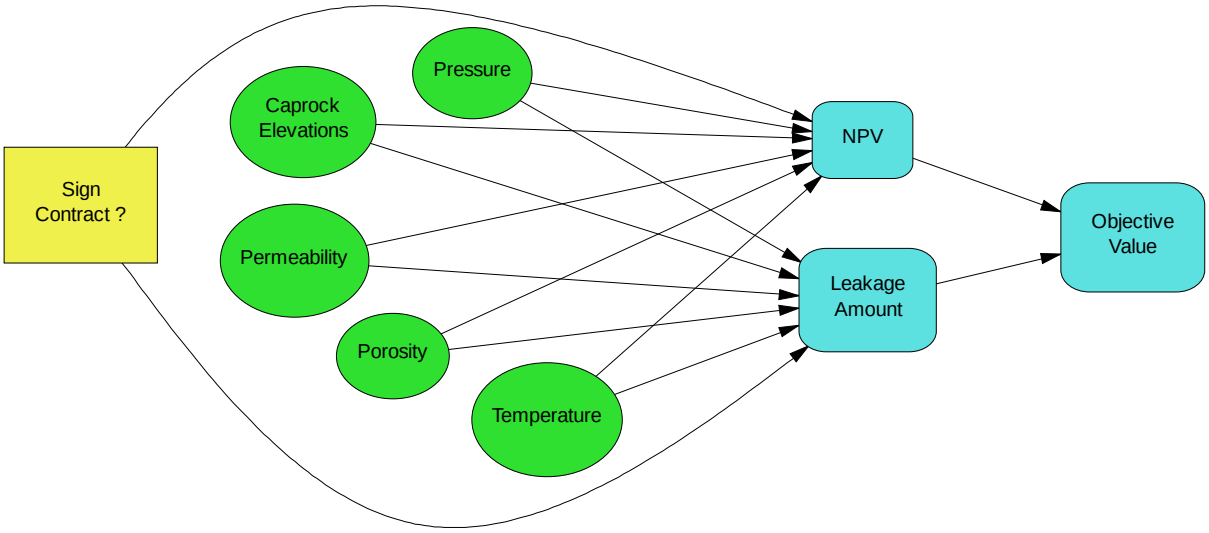


Figure 4.2 Influence diagram of the model analysis

Figure 4.3 represents the workflow of the case study. The analysis starts with uncertain variables that were included in the model one at a time to investigate their individual impact on the objective value and determine the material variables. The objective value is represented as an objective function that contains two attributes and weights. MRST is used to run 100 realizations of CO_2 injection simulations into the Utsira Formation aquifer. The forecast of CO_2 injection, CO_2 leakage amount, the average variable value in the geological model from each realization are used as inputs to estimate VOI. We consider the linear regression, random forest, and ***k***-nearest neighbors as the options of regression method in the simulation-regression approach. The 10-fold cross-validation of machine learning is used to determine which methods give the best fit from each dataset. Finally, VOIs are estimated for perfect and imperfect information and three scenarios of sensitivity analysis will be carried out.

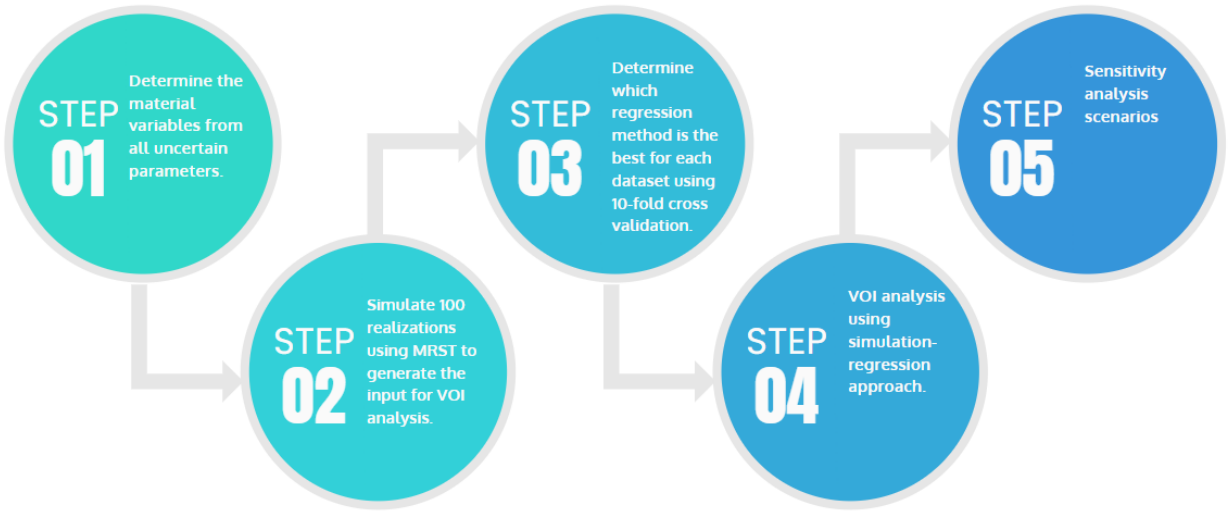


Figure 4.3 Workflow of the case study

The company will gain profit from CO_2 storage based on the carbon price at March 2021 of \$40/tCO_2 (Quandl, 2021) minus the total cost of \$38.2/tCO_2 which includes the following: \$21.5/tCO_2 cost for capturing CO_2 from gas processing and bio-ethanol production (Irlam, 2017), \$13/tCO_2 cost of CO_2 offshore storage (Zero Emissions Platform, 2011), \$3.5/tCO_2 costs of construction, operation, and maintenance (Bock et al., 2003), \$0.2/tCO_2 cost for CO_2 monitoring (Anyosa et al., 2021). The NPV is calculated with a discount rate of 8% per year using input from the reservoir simulations, i.e. total CO_2 injection ($\boldsymbol{M_{inj}}$) and total CO_2 leakage ($\boldsymbol{M_{leak}}$), the NPV($\boldsymbol{x}$,$\boldsymbol{a}$) and is given by:

$$NPV(x, a = 0) = 0 \tag{4.5}$$

$$NPV(x, a = 1) = \frac{(profit - cost)}{(1 + i)^n} \tag{4.6}$$

where,

$$profit = \$40\ /tCO_2 \cdot (M_{inj} - M_{leak}),$$

$$cost = \$38.2\ /tCO_2 \cdot (M_{inj}),$$

$$i = interest\ rate, n = the\ discounted\ period$$

The value of the CO_2 storage contract is reflected by the objective to maximize the NPV and minimize the CO_2 leakage. To measure the achievement of the objective values with the assumption that the company is risk-neutral, value functions of each attribute are defined as follows:

$$v_A(NPV) = a - b \cdot (NPV - x) \tag{4.7}$$

$$v_B(Leak) = a - b \cdot Leak \tag{4.8}$$

Where x is the cost of additional information, note that this value will be increased until the objective value is equal to or close to zero to calculate the maximum price of additional information that the company should pay. The variables a and b are calculated with the upper and lower limit values from the reservoir simulation in the probabilistic model:

$0 \leq v_A(NPV) \leq 1$, for $-200\ \text{M} \leq NPV \leq 111\ \text{M}$, $v_A(111\ \text{M}) = 1, v_A(-200\ \text{M}) = 0$

$0 \leq v_B(Leak) \leq 1$, for $110\ \text{M} \leq Leak \leq 0$, $v_B(0) = 1, v_B(110\ \text{M}) = 0$

Therefore, the value functions become:

$$v_A(NPV) = 0.64 + 3.21e - 09 \cdot (NPV - x) \tag{4.9}$$

$$v_B(Leak) = 1 - 9.072\text{e} - 12 \cdot Leak \tag{4.10}$$

According to Bratvold and Begg (2010), the objective values should be ranked based on their importance in distinguishing between alternatives by using the swing weighting method which takes the relative magnitudes of the payoffs into account. Using the values from the probabilistic evaluation, the swing weighting method is applied in Table 4.1. The attribute with the greater percentage gain over its worst score is given a higher rank than the other. The result shows that minimizing the leakage amount is ranked as more important than maximizing NPV.

Table 4.1 Swing weighting results

Attributes	Actual Alternatives				Hypothetical Alternatives		Swing Rank
	Sign			Not Sign	Worst	Best	
	Min	Mean	Max				
NPV	-2.00E+08	6.13E+07	1.11E+08	0	-2.00E+08	1.11E+08	2
Leak	8.08E+02	1.96E+10	1.10E+11	0	1.10E+11	0	1

The next step is to combine the weighted values of each objective to determine an overall value for each alternative which can be seen in Table 4.2. Finally, the objective function can be defined from the swing weights as:

$$O_f(Leak, NPV) = w_1 v_B(Leak) + w_2 v_A(NPV), \tag{4.11}$$

where the weights are $w_1 = 0.56$ and $w_2 = 0.44$, and $0 \leq O_f(Leak, NPV) \leq 1$. In the prior model (without information), the alternative to turn down the CO_2 storage contract has an expected objective value of 0.84 which is higher than the alternative to sign it that has an expected objective value of 0.83.

Table 4.2 Expected objective values and weights of each objective

Attributes				Expected Objective Values	
Name	Swing Rank	Weights		Sign	Not Sign
		abs	norm		
Leak	1	100	0.56	0.82	1.00
NPV	2	80	0.44	0.84	0.64
		180	**1.00**	**0.83**	**0.84**

4.3 Reservoir Simulation in MRST

The CO_2 injection forecast is simulated using MRST with a top surface grid and thickness maps of the Utsira Formation. The setup consists of one injection well placed in the south of the Utsira Formation with an injection rate of 5 Mt per year over a period of 60 years (given a reference density of 760 kg/m3) and a simulation time of 3000 years to investigate the long-term migration and leakage (post-injection). Each uncertainty variable is included in the model one at a time with 100 realizations to evaluate their individual impact on the objective value. A sensitivity value analysis is conducted using the Tornado technique to evaluate the impact of the minimum and maximum values of each variable on the expected objective value (Figure 4.4).

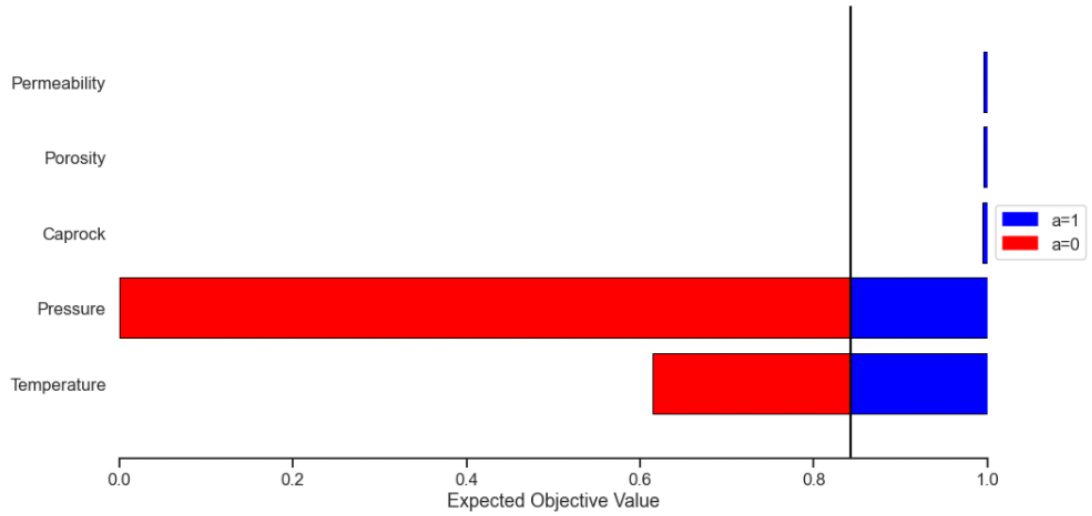


Figure 4.4 Tornado diagrams of uncertainty variables

The results in Figure 4.4 show that only pressure and temperature can alter the decision between signing the contract (***a***=1) and turning down the contract (***a***=0). The pressure and temperature diagrams show how the expected objective values vary and could lead to different optimal decisions. On the other hand, permeability, porosity, and caprock elevation will not alter the decision since they always result in higher expected objective values compared to the alternative to turn down the contract. The impact of pressure and temperature on the objective values can be seen in Figure 4.5. As explained in Chapter 2, the decreasing pressure from a hydrostatic condition and a warmer aquifer will cause a decrease in CO_2 density and resulting in a lower CO_2 storage capacity. The plateau in Figure 4.5 represents the constraint of the maximum objective value, considering the minimum leakage amount of zero and the maximum NPV of $111 million that the company can get for injecting 5 Mt CO_2 per year for 60 years. Within each variable interval, the objective value decreases with decreasing average pressure when it is below the critical point and with increasing average temperature when it is above the critical point. Note that the variable values in Figure 4.5 are averaged, however, the changes in objective values due to the critical point occur in each grid cell simulation. This indicates that if the standard deviation is reduced to a value that is not close to the critical point, the impact on the objective value will not be as big as it is. Consequently, the material variables for this decision context are pressure and temperature only.

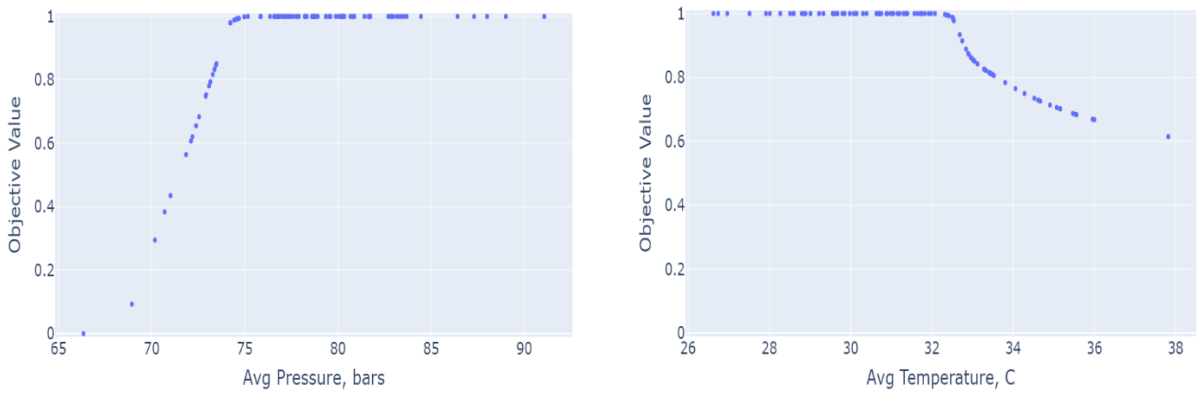


Figure 4.5 Impact of pressure and temperature on the objective value

All material variables will be assessed as uncertainties in the next step, while the non-material variables are considered fixed at the base value. A hundred realizations of CO_2 injection simulations in MRST are run to generate a forecast of CO_2 injection, CO_2 leakage amount, and the average variable value in the geological model from each realization. These values are then used to calculate the expected objective value using the value functions and the objective function previously described in chapter 4.2. Based on the probabilistic evaluation of the prior model (without information), the company should turn down the contract because it has an expected objective value of 0.911, higher than the alternative to sign the contract which has an expected value of 0.833. The risk profiles and expected objective value of each alternative are illustrated in Figure 4.6. The blue line shows the CDF or the possible outcomes of the alternative to sign the contract, while there is no CDF of the alternative to turn down the contract because it has a certain value (consider the leakage amount and NPV of zero).

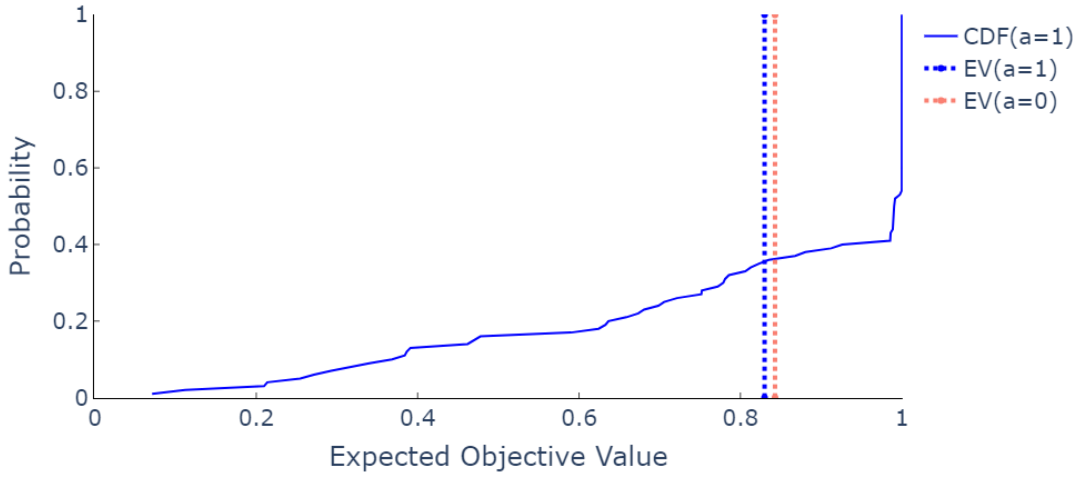


Figure 4.6 Risk profiles of each alternative in the prior model

4.4 VOI calculation

Applying the methodology presented in Chapter 3, VOI is estimated using the simulation-regression approach with regression methods OLS, RF, and KNN. The analysis begins with regression method selections for each additional information alternative, i.e., pressure, temperature, and combined information. The 10-fold cross-validation is applied to determine which method gives the best fit for each dataset. Table 4.3 indicates that the KNN method has the lowest MSE value and as a result, it was chosen to estimate VOI in the next step.

Table 4.3 Cross-validation scores

Uncertain Variables	Cross-validation scores		
	OLS	Random Forest	KNN
pressure	1.019	1.018	1.012
temperature	1.029	1.033	1.028
combine	1.0136	1.0032	1.0026

The VOI analysis will be done for perfect information and imperfect information with noise= {5%, 8%, 10%, 50%} for each information gathering scheme. Hence, there are 15 data sets and regression models. The noises are represented by a normal distribution MCS sample of each realization and the calculations of VOI with imperfect information are repeated 1000 times. The values for ***k*** in the KNN method are determined using 10-fold cross-validation where it has the lowest MSE value. Thereafter, the VOIs are predicted from the fitted KNN model and shown in Figure 4.7. Overall, the VOIs drop as soon as there is noise in the information, i.e. the information is imperfect.

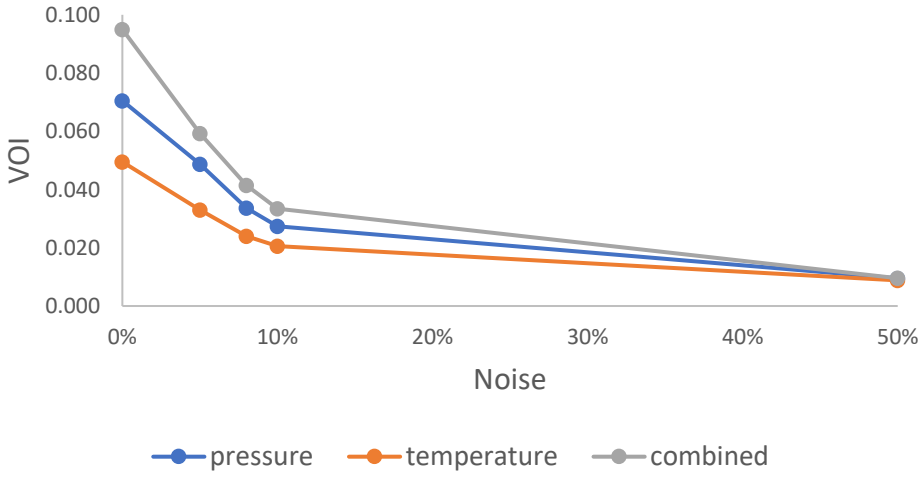


Figure 4.7 Summary of the VOI estimations

The risk profiles based on perfect information are plotted in Figure 4.8. The cumulative distribution functions (CDFs) represent how spread-out the possible outcomes are in each information-gathering alternative, it shows that the combined information scheme first-order stochastically dominates the pressure information which also dominates the temperature information. Therefore, the combined information has the highest expected objective value (and VOI), followed by the pressure and temperature, respectively.

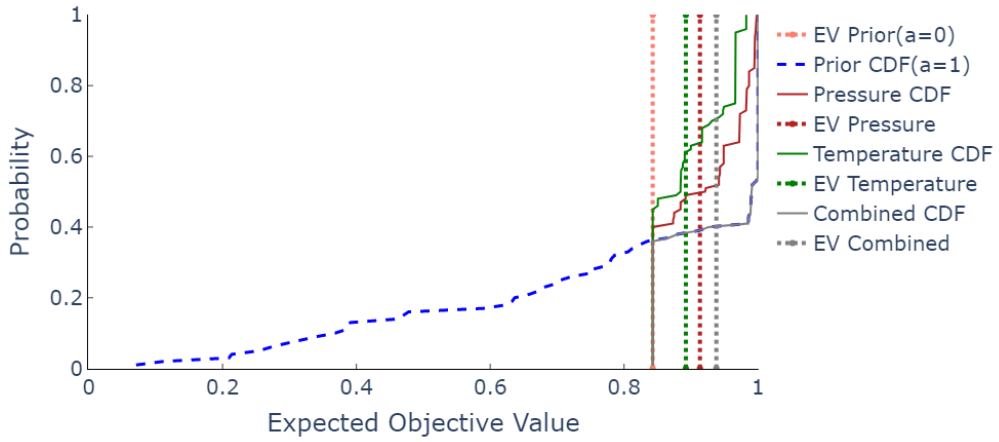


Figure 4.8 Risk profiles of each information-gathering alternative

The frequency distribution of the optimal decisions for each information gathering alternative with perfect information is shown in Figure 4.9. Given the results from the perfect information gathering schemes, either individual uncertain parameters or combined show that the frequency of choosing the "Sign" is greater than the "Don't Sign" alternative. However, the optimal decision relies on the result from the information-gathering that the company will choose. The relationship between the optimal decision and the perfect information outcomes is plotted in Figure 4.10. It indicates that pressure information has a more dominant impact on the decision.

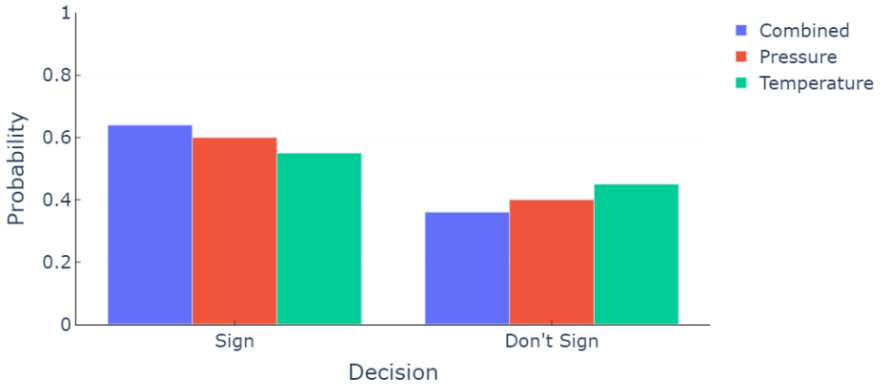


Figure 4.9 Optimal decisions in each uncertainty model with perfect information

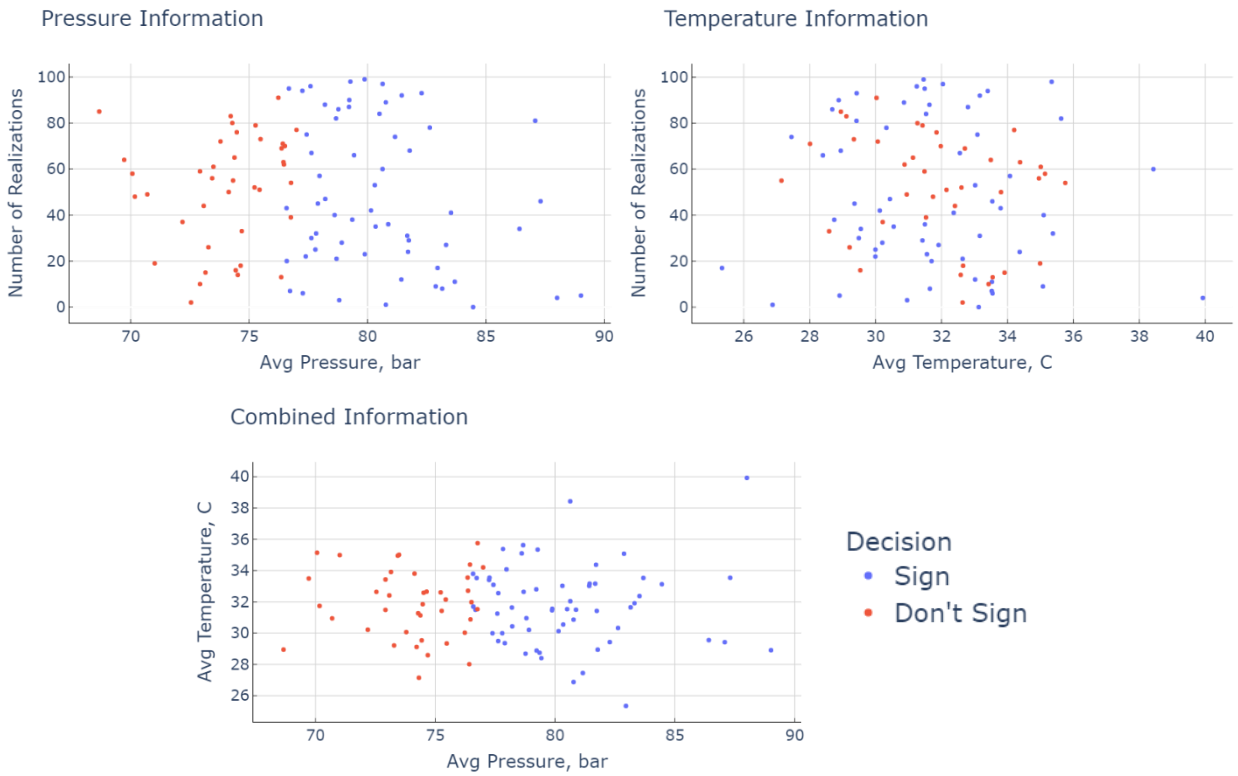


Figure 4.10 The optimal decisions based on the information outcomes

Figure 4.11 shows that the maximum price of the additional information that company should pay varies, depending on the magnitude of the information on the objective value. For the perfect information, the price of the combined information, pressure, and temperature are $110.9 million, $107.9 million, $94.6 million, respectively. When the noise of 50% is added to the model, the price of the information is reduced to $43.2 million for combined information, $41.4 million for pressure, and $39.2 million for temperature. Furthermore, the correlation between the objective value and the maximum price of the additional information that the company should pay is displayed in Figure 4.12 from all the data points in Figure 4.11. Overall, it indicates a logarithmic trend, the price remains stable if the objective value is more than 0.06 and falls dramatically when the objective value is less than 0.06.

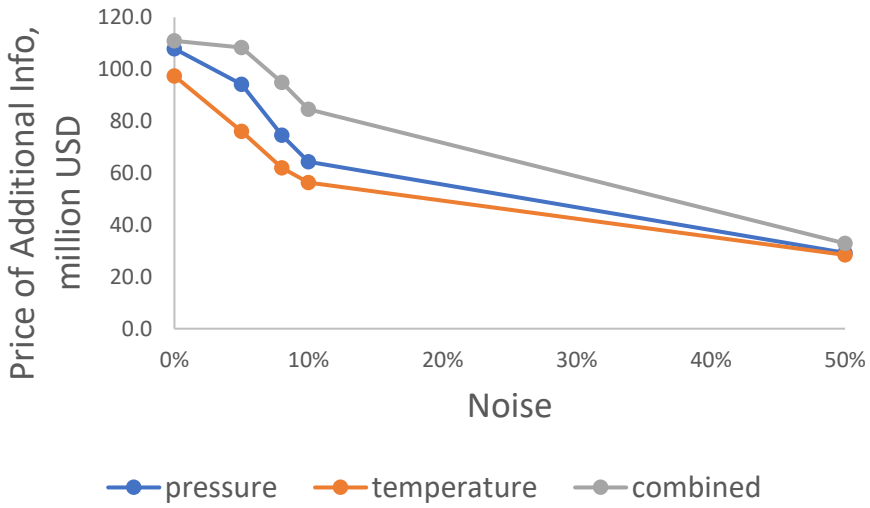


Figure 4.11 Maximum price of the additional information

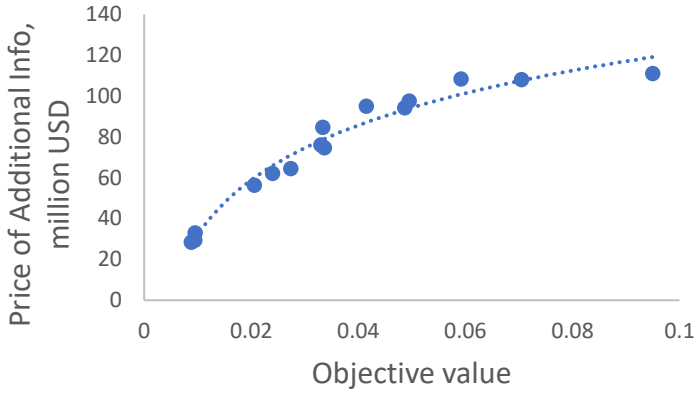


Figure 4.12 Correlation between objective value and price of the additional information

Noise not only changes the expected objective value of the information-gathering schemes but also the frequency of the optimal decisions as can be seen in Figure 4.13. Compared to the perfect information situation, the probability distribution of optimal decision given imperfect combined information switches when 50% noise is added, i.e., the frequency of signing the contract cases is lower than the frequency of not signing it. Meanwhile, the distributions of the optimal decisions based on pressure and temperature information change after the incremental noise reaches 10% and 5%, respectively. For this reason, the company is suggested to choose the information-gathering scheme with noise less than 50% for combined information, 10% for pressure, and 5% for temperature to get a satisfying result.

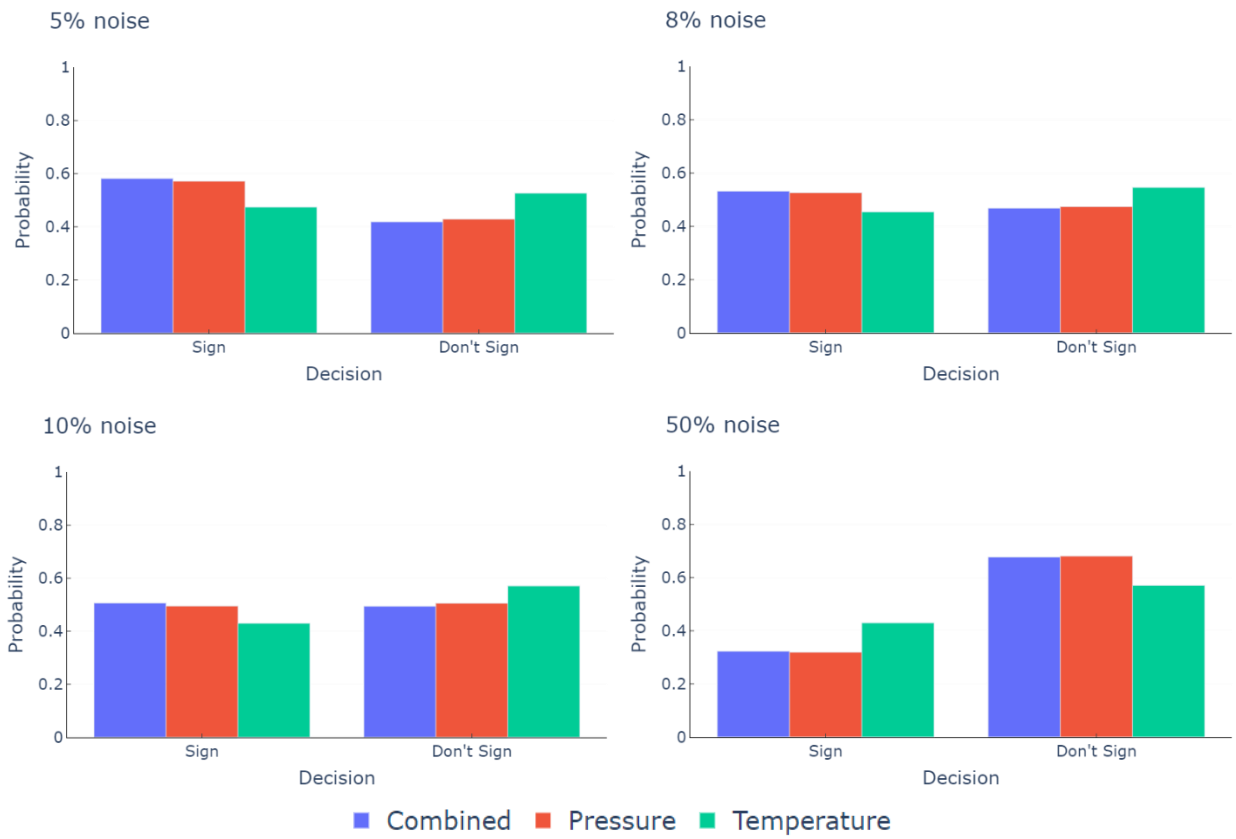


Figure 4.13 Probability distributions of optimal decision with imperfect information

4.5 Sensitivity analysis

In this chapter, three scenarios were added to examine the effect of certain parameters on the decision and the VOI estimations. The base model for the analysis is taken from the previous chapter.

4.5.1 Scenario I: Add carbon price uncertainty to the model

As explained in the previous chapter, the carbon price can affect the NPV value. The carbon price uncertainty is modeled using the geometric Ornstein-Uhlenbeck process. The stochastic differential equation for the Ornstein-Uhlenbeck process is given by:

$$dS_t = \lambda(\mu - S_t)dt + \sigma dW_t, \tag{4.12}$$

where S_t is the logarithm of the carbon price, λ is the speed of mean reversion, μ is the long-term mean, σ is the volatility of the process, and W_t denotes the Wiener process (a Brownian motion). S_t is normally distributed with mean and variation of:

$$\mathbb{E}(S_t) = S_0 e^{-\lambda \Delta t} + \mu(1 - e^{-\lambda \Delta t}) \tag{4.13}$$

$$\mathbb{V}ar(S_t) = \frac{\sigma^2}{2\lambda}(1 - e^{-2\lambda \Delta t}) \tag{4.14}$$

Equation (4.8) can be discretized by:

$$S_{t+1} = e^{-\lambda \Delta t} S_t + \left(1 - e^{-\lambda \Delta t}\right)\mu + \sigma \sqrt{\frac{(1 - e^{-2\lambda \Delta t})}{2\lambda}} \Delta W_t \tag{4.15}$$

where ΔW_t is a standard normal Wiener process. The least-squares polynomial regression method is used to calibrate the input parameters for modeling the carbon price resulting in a mean reversion speed of 0.115, volatility of 0.503, and long-term mean of 3.075 (logarithm scale). The historical data of carbon price is taken from ECX EUA Futures (Quandl, 2021), converted from Euro to USD (1 Euro = 1.2 USD), and plotted in Figure 4.14. The calibration of the input parameters used the historical data from the year 2011 to 2021.

Figure 4.14 Historical data of carbon price

The carbon price is simulated for 100 realizations as input to the VOI analysis. The price simulation is shown in Figure 4.15. The mean carbon price is \$45.1/tCO₂. As a consequence of value changes in the carbon price, the minimum and maximum values of NPV are also changed:

$0 \leq v_A(NPV) \leq 1$, for -1.92 B $\leq NPV \leq 6.92$ B , $v_A(6.92 \text{ B}) = 1, v_A(-1.92 \text{ B}) = 0,$

Therefore, the NPV value function becomes:

$$v_A(NPV) = 0.217 + 1.13e - 10 \cdot (NPV - x), \tag{4.16}$$

while the leakage value function and the weights remain the same.

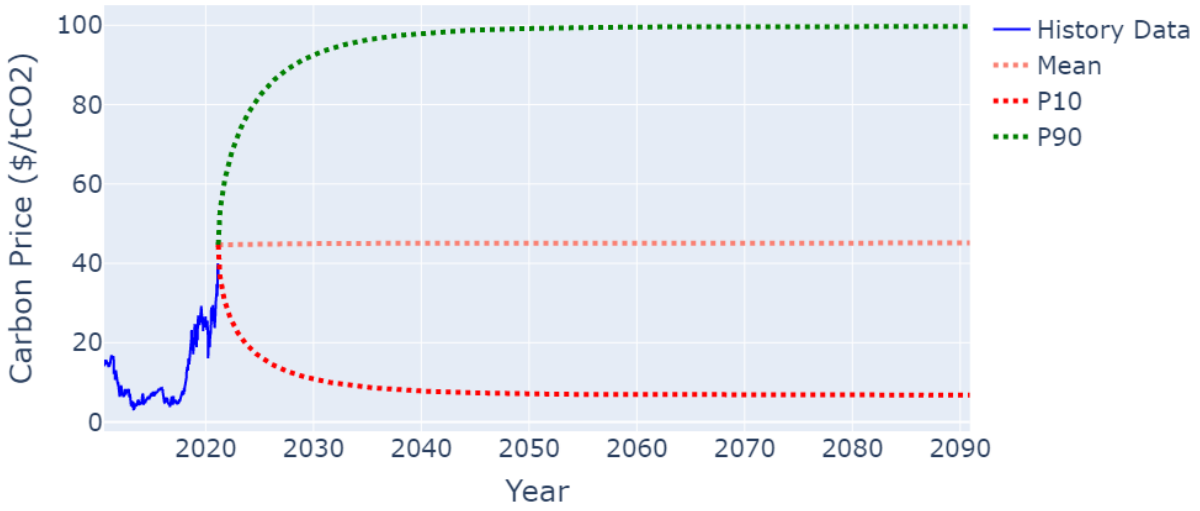


Figure 4.15 Ornstein-Uhlenbeck model for carbon price

First, carbon price variable is examined using the Tornado technique in Figure 4.16, where it indicates that carbon price variable is material. Then, the VOI estimations are done as in the base model. The summary of VOI estimation and its maximum price of additional information is plotted in Figure 4.17. It shows that the expected objective values of individual pressure and temperature information are becoming similar. In comparison with the base model, the maximum price of additional information that the company should pay escalates significantly up to 20-30 times the base model values, although the decline is steeper than the base model when the noise increases. The company should obtain new information if the price of the additional information is lower than the values in Figure 4.17.

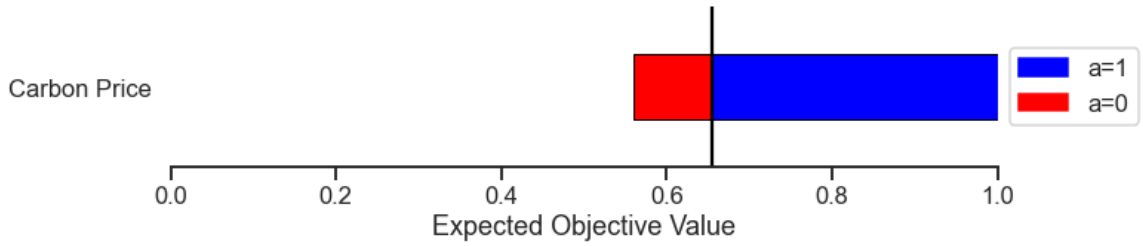


Figure 4.16 Tornado diagram of carbon price

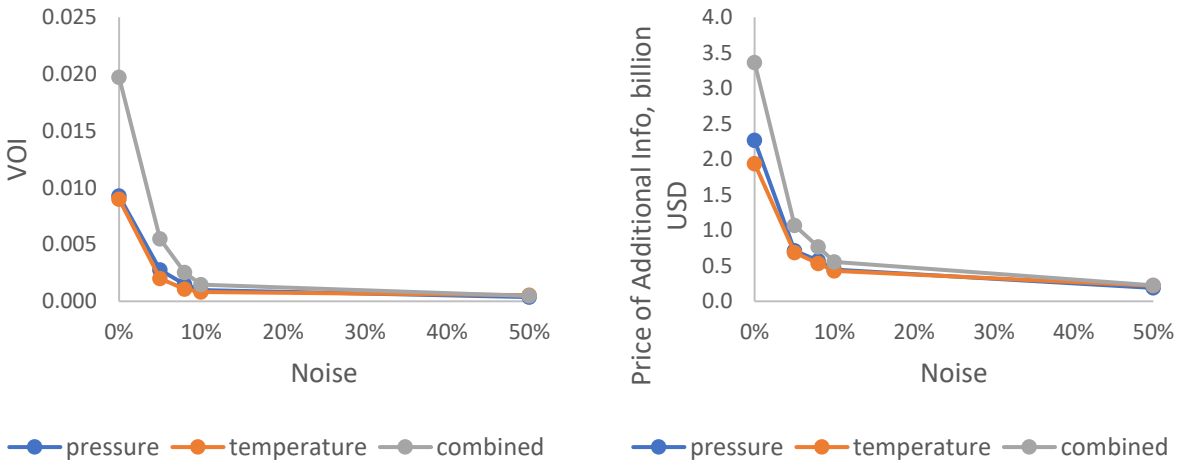


Figure 4.17 VOI estimations and the maximum price of additional information in scenario I

The risk profiles of scenario I without noise are plotted in Figure 4.18. The prior model remained the same, indicating that the company should not sign the contract. Compared to the base model, the increments of expected value after getting additional information are smaller since the carbon price remains uncertain in this model. According to the optimal decisions in the posterior models with perfect information shown in Figure 4.19, the frequency of not signing the contract is higher than the frequency of signing it given the result from any information-gathering schemes. The optimal decision strategy for different information outcomes can be seen in Figure 4.20. There is no clear pattern on the decision impact given the variations in the input variables.

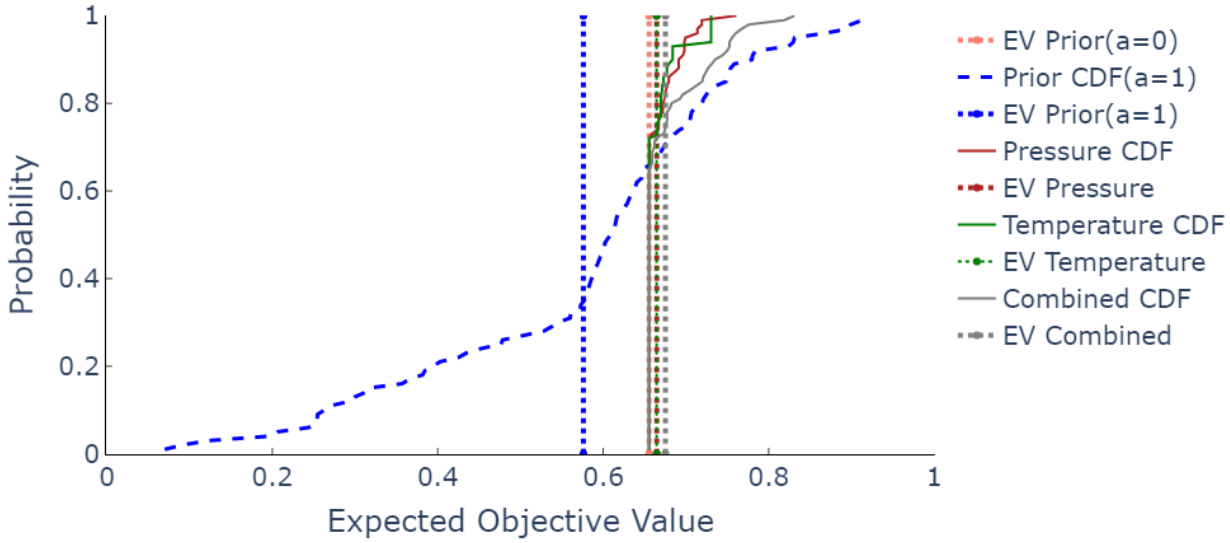


Figure 4.18 Risk profiles in scenario I

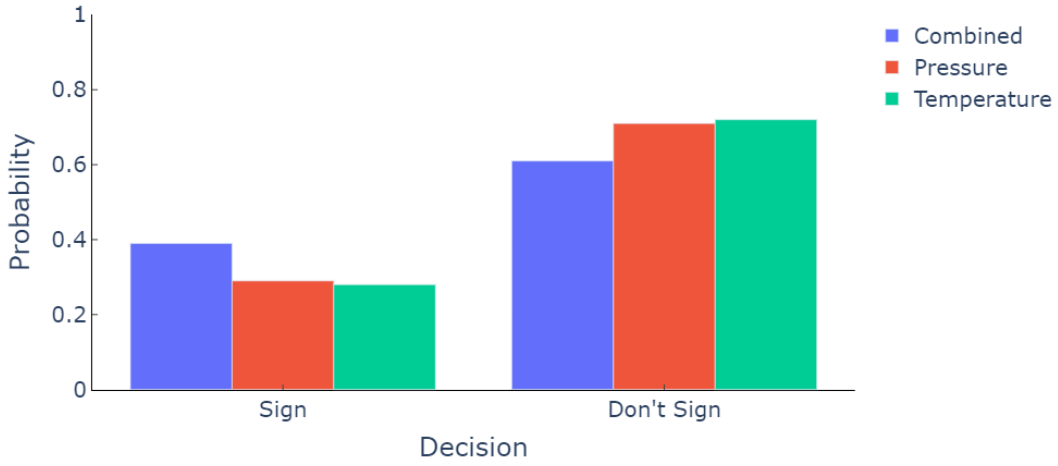


Figure 4.19 Probability distributions of optimal decision in scenario I

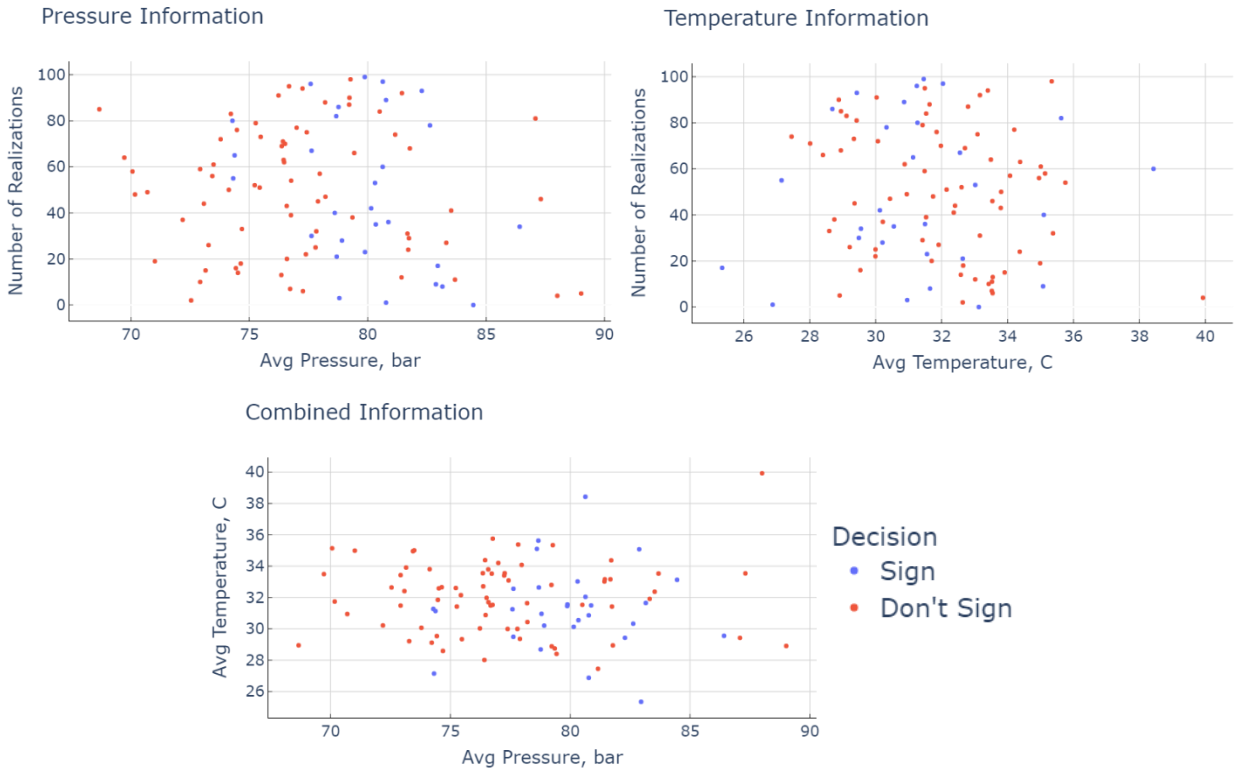


Figure 4.20 The optimal decisions based on information outcomes in scenario I

A sensitivity analysis of the carbon price on the objective value is shown in Figure 4.21. The mean carbon price represents the mean value of each carbon price realization. When the mean carbon price model is increased, the objective value increases gradually with a larger variation, leading to higher uncertainty in the objective value.

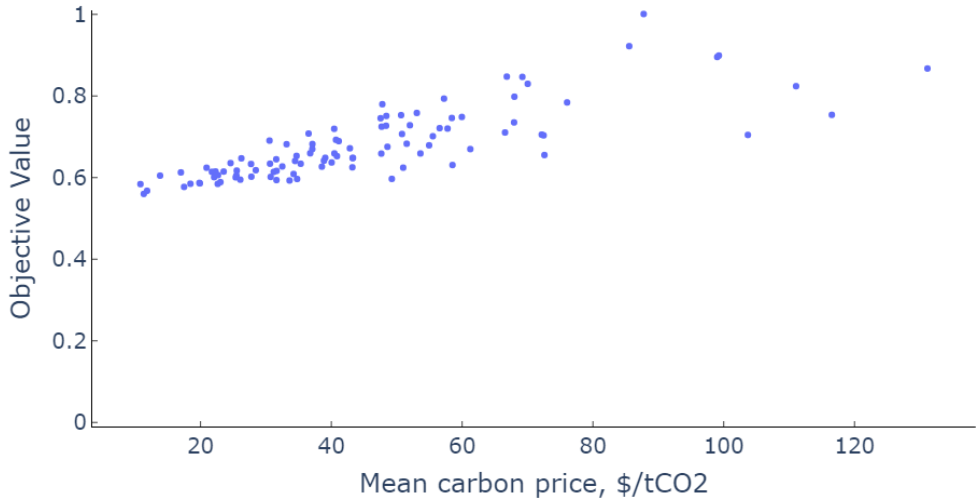


Figure 4.21 Sensitivity analysis of carbon price on the objective value

4.5.2 Scenario II: The leakage volume constraint

Van der Zwaan and Smekens (2009) recommended a maximum acceptable value for the leakage rate below 0.5% annually which implies leakage of 50% over a century, whilst Bielicki et al. (2015) and Vinca et al. (2018) suggested a lower maximum leakage rate. In scenario II, the leakage is limited to i = {5%, 10%, 20%} of the total CO_2 injection, indicating that it is not acceptable to have a leakage greater than these values. Due to the leakage constraint, the value of total leakage more than i is truncated to zero. These maximum leakage values are applied to the leak function (4.6) by changing the lower value to the constraint values:

$0 \leq v_B(Leak) \leq 1$, for i $\leq Leak \leq 0$, $v_B(0) = 1, v_B(\text{i}) = 0$,

Where i = {5%, 10%, 20%}, therefore, the value function is changed to:

$v_B(Leak) = 1 - 6.659\text{e} - 11 \cdot Leak$ for i = 5%, (4.17)

$v_B(Leak) = 1 - 3.329\text{e} - 11 \cdot Leak$ for i = 10%, (4.18)

$v_B(Leak) = 1 - 1.665\text{e} - 11 \cdot Leak$ for i = 20%, (4.19)

while the NPV value function and the weights remain the same.

The VOI estimations are done as in the base model using the same material variables. The VOI estimations using different leakage constraints are displayed in Figure 4.22. Note that each case has a different value function, therefore the expected objective value comparison between the base model and this scenario shall not be done.

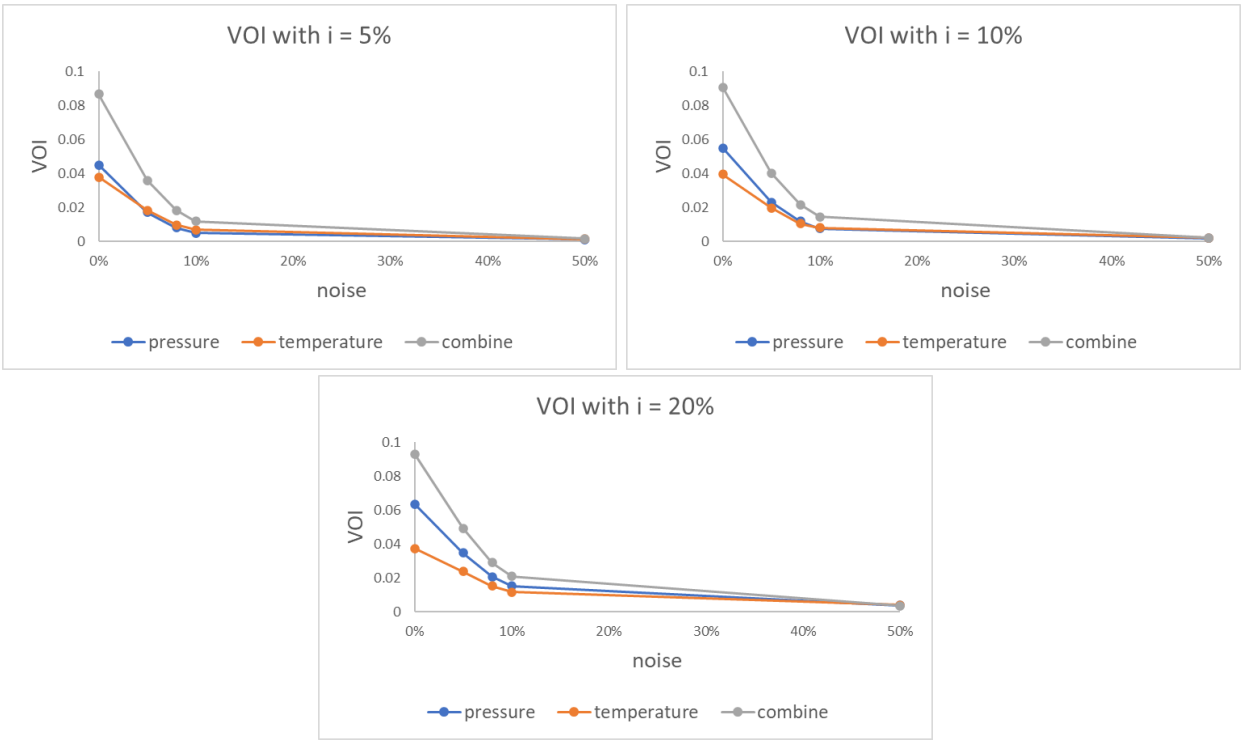


Figure 4.22 VOI estimations for different leakage constraint

The VOI relationship between pressure-temperature-combined information can be inspected from Figure 4.22, and the corresponding maximum price of information can be seen in Figure 4.23. It can be observed that as the leakage constraint becomes smaller, the difference between VOIs of individual pressure and temperature information decreases, resulting in approximately the same expected objective value. This is a result of some realization values being truncated to zero which causes the maximum, the minimum, and the average values of the model to change. A sensitivity analysis of individual pressure and temperature information with different leakage functions (leakage constraint) was done to verify the changes on the maximum, minimum, and average values. Table 4.4 indicates a reduction in both pressure and temperature mean values due to the decreased leakage constraint. Notice that the temperature expected objective value (mean value) becomes lower than the expected objective value of the alternative to turn down the contract (***a***=0) when the leakage constraint is equal to or lower than 10%, while the difference between the minimum and maximum values of pressure is always greater than that of the temperature. Evidently, both conditions give approximately the same impact on the pressure and temperature VOIs in the model Figure 4.22.

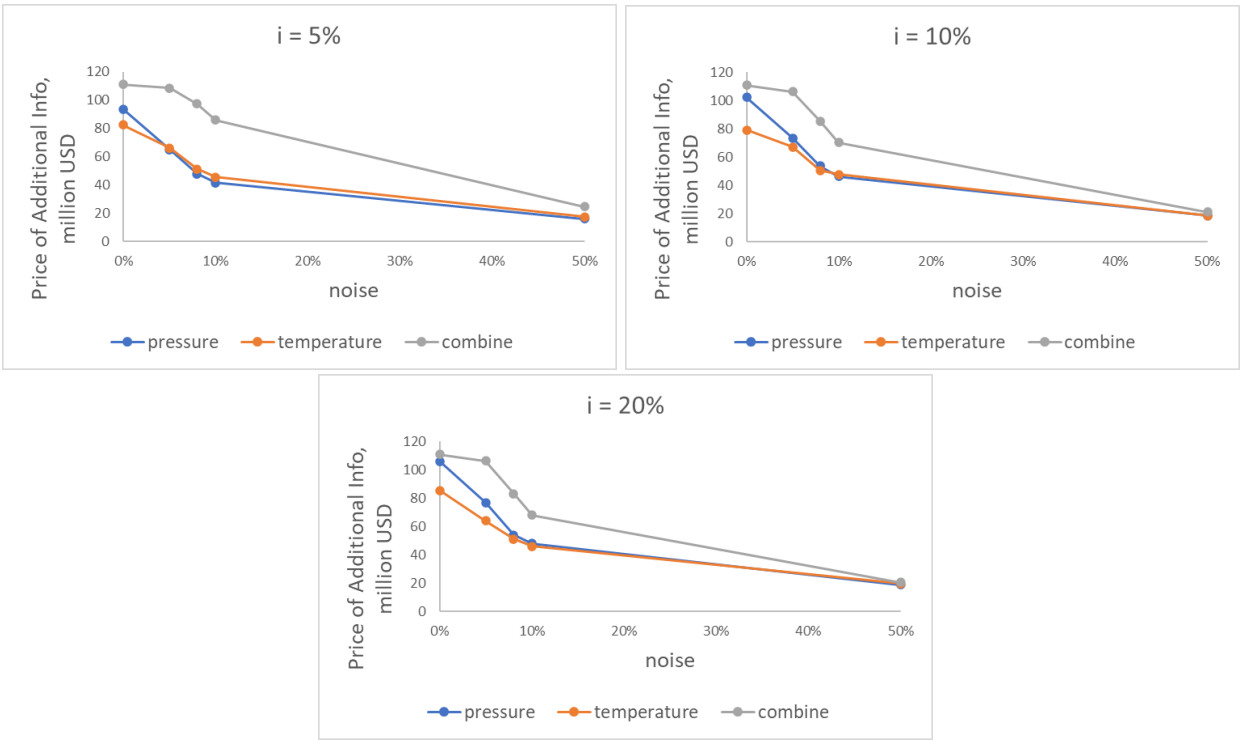


Figure 4.23 The maximum price of information comparison for different leakage constraints

Table 4.4 Sensitivity analysis with different leakage constraints

Max. Leakage Constraint	Expected Objective Value (a=0)	Pressure Objective Value			Temperature Objective Value		
		Min.	Mean	Max.	Min.	Mean	Max.
5%	0.843	0	0.868	0.999993	0.162	0.748	1
10%		0	0.874	0.999996	0.329	0.802	1
20%		0	0.898	0.999998	0.525	0.881	1

Leakage constraints not only change the VOI estimation but also the frequency distributions of the optimal decisions. The optimal decision in the prior model remained the same as the base model, indicating that the company should not sign the contract. Optimal decision summaries of all information-gathering schemes with perfect information can be seen in Figure 4.24. The perfect information with leakage constraints of 5% and 10% show that both individual pressure and temperature test results have a higher frequency in the "Don't Sign" alternative, although the combined information indicates that the "Sign" alternative has a higher frequency. On the other hand, the model with a 20% maximum leakage constraint shows that the "Sign" alternative has a higher frequency than the "Don't Sign" alternative given all information-gathering schemes with perfect information.

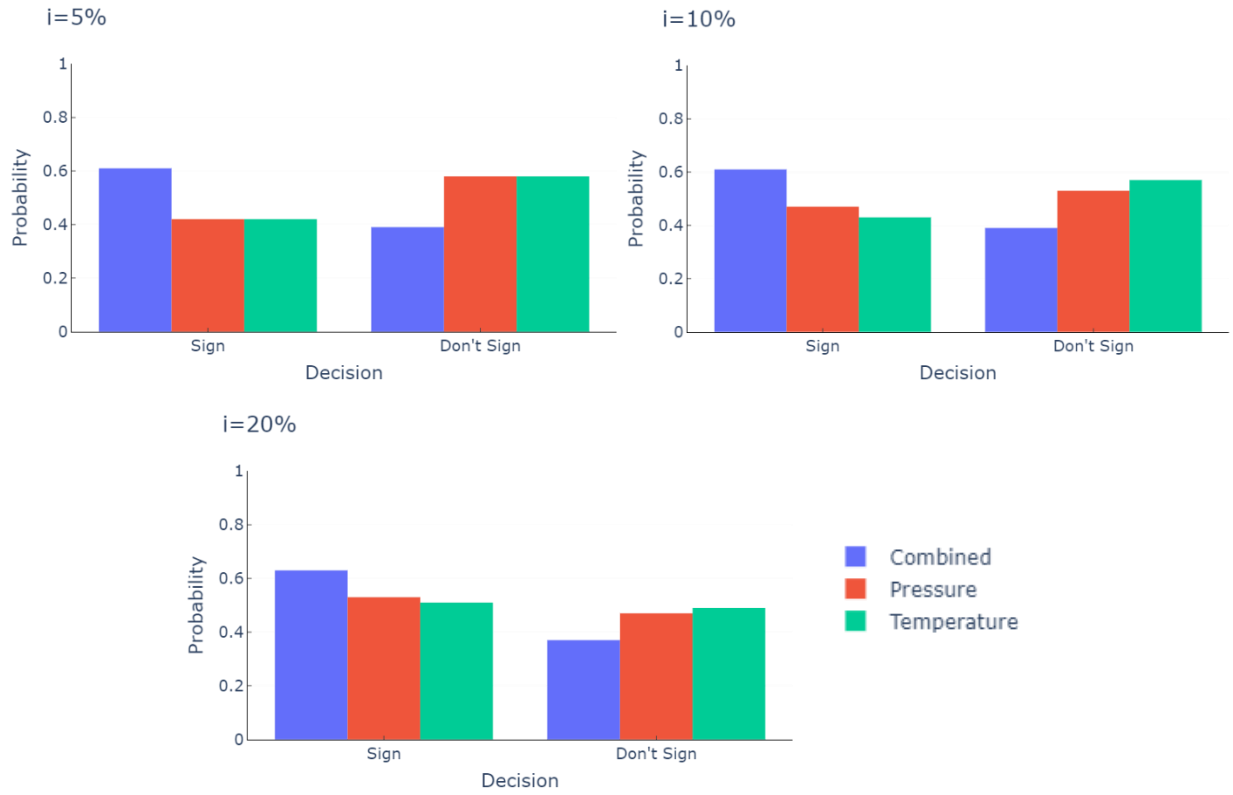


Figure 4.24 Probability distributions of optimal decision for different leakage constraints

4.5.3 Scenario III: Changes in the injection period

In scenario III, the effect of the injection period on VOI estimation is analyzed. The injection periods of {40 years, 80 years} are assigned while the CO_2 injection rate of 5 Mt per year remains the same with the base model. Therefore, the total CO_2 injections for 40 years and 80 years are 200 Mt and 400 Mt, respectively. Due to changes in the minimum and maximum values of NPV and long-term leakage, the value functions (4.5) and (4.6) are rearranged:

For 40 years CO_2 injection,

$$0 \le v_A(NPV) \le 1 \quad \text{, for } -151\text{ M} \le NPV \le 106\text{ M} \quad \text{, } v_A(106\text{ M}) = 1, v_A(-151\text{ M}) = 0$$

$$0 \le v_B(Leak) \le 1 \quad \text{, for } 57\text{ M} \le Leak \le 0 \quad \text{, } v_B(0) = 1, v_B(57\text{ M}) = 0$$

$$v_A(NPV) = 0.59 + 3.87e-09 \cdot (NPV - x), \tag{4.20}$$

$$v_B(Leak) = 1 - 1.75\text{e} - 11 \cdot Leak, \tag{4.21}$$

and for 80 years CO_2 injection,

$$0 \le v_A(NPV) \le 1 \quad \text{, for } -207 \text{ M} \le NPV \le 112 \text{ M} \quad \text{, } v_A(112 \text{ M}) = 1, v_A(-207 \text{ M}) = 0$$

$$0 \le v_B(Leak) \le 1 \quad \text{, for } 184 \text{ M} \le Leak \le 0 \quad \text{, } v_B(0) = 1, v_B(184 \text{ M}) = 0$$

$$v_A(NPV) = 0.65 + 3.13e - 09 \cdot (NPV - x), \tag{4.22}$$

$$v_B(Leak) = 1 - 5.43\text{e} - 12 \cdot Leak, \tag{4.23}$$

while the weights remain the same.

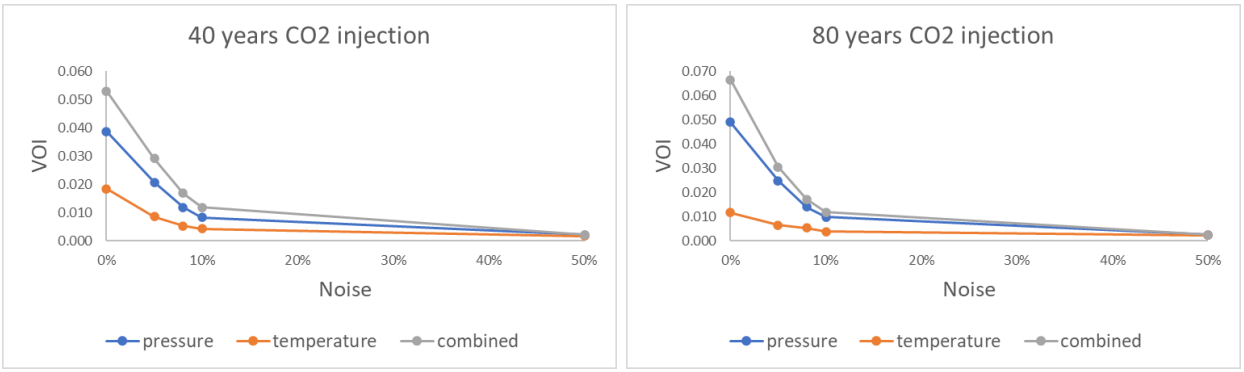


Figure 4.25 VOI estimations for different CO_2 injection periods

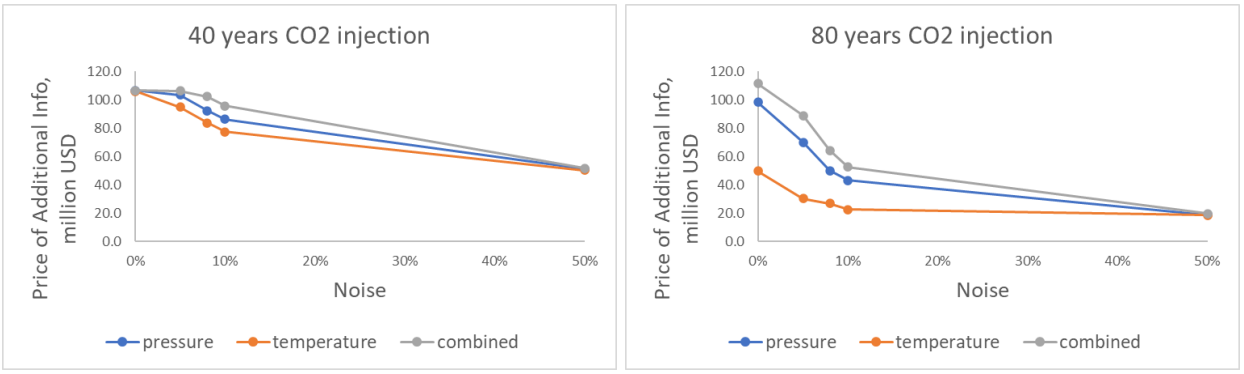


Figure 4.26 The maximum price of information comparison for different CO_2 injection periods

Using the same methodology as in the base model, the VOI results for different injection periods are displayed in Figure 4.25, followed by the maximum price of the additional information in Figure 4.26. There is a slight increase in the price of the combined perfect information when the injection period increases, from \$107 million to \$112 million, although the difference in prices of information between information-gathering schemes becomes larger. The decreases in VOI and its price are also sharper with the increasing noise when CO_2 injection period is longer.

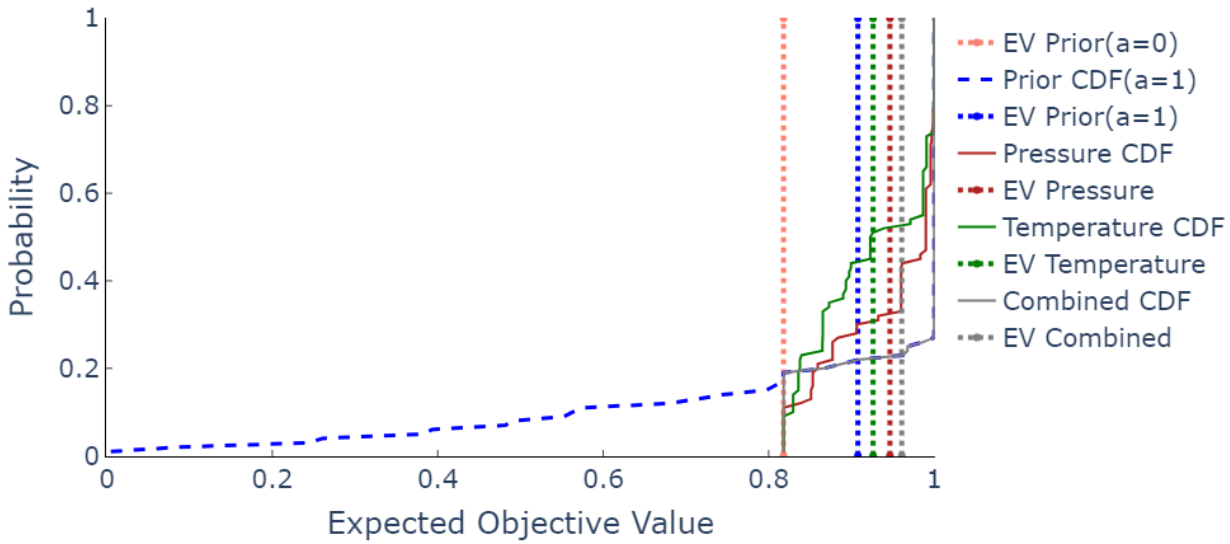


Figure 4.27 Risk profiles with 40 years CO_2 injection

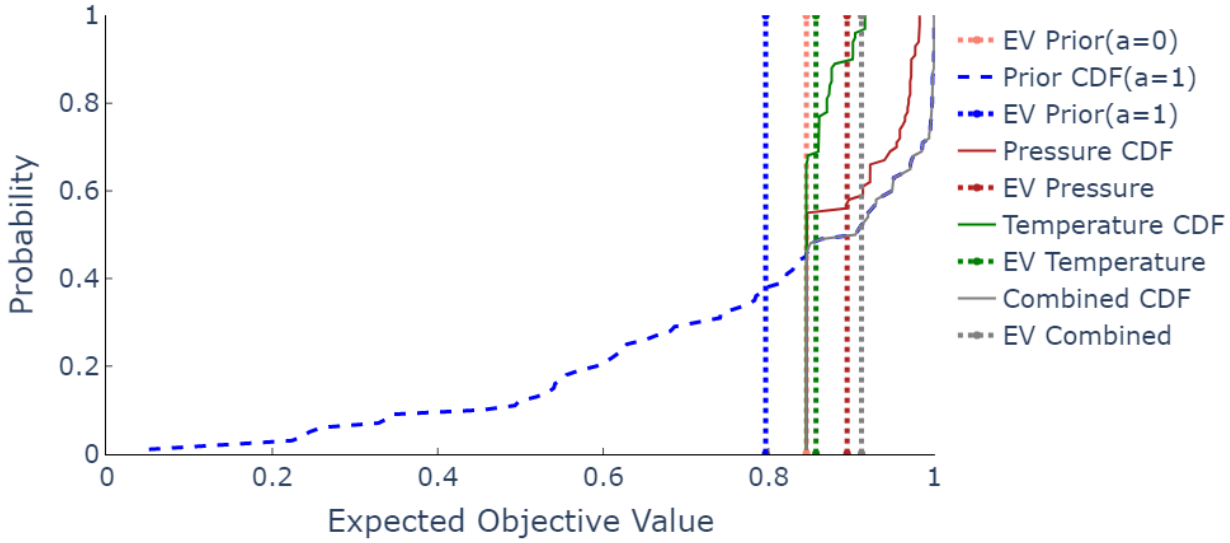


Figure 4.28 Risk profiles with 80 years CO_2 injection

The risk profiles of 40 years and 80 years injection periods can be seen in Figure 4.27 and Figure 4.28, respectively. The optimal decisions of each model with perfect information in scenario III can be seen in Figure 4.29. In the model with 40 years CO_2 injection, the prior model suggests the company to sign the contract. It also showed that there is no stochastic dominance between the information-gathering schemes. If they obtain any additional information scheme with perfect information, the "Sign" alternative has a higher frequency than the "Don't Sign" alternative. On the contrary, the prior model with 80 years CO_2 injection advises the company to turn down the contract. The risk profiles indicate that the combined information scheme first-order

stochastically dominates pressure information which also dominates temperature information. Furthermore, the outcomes from both individual information schemes with perfect information show that the "Don't Sign" alternative has a higher frequency. However, the outcome from the combined information indicates that the distribution of the optimal decision is switched, i.e. the "Sign" alternative has a higher frequency.

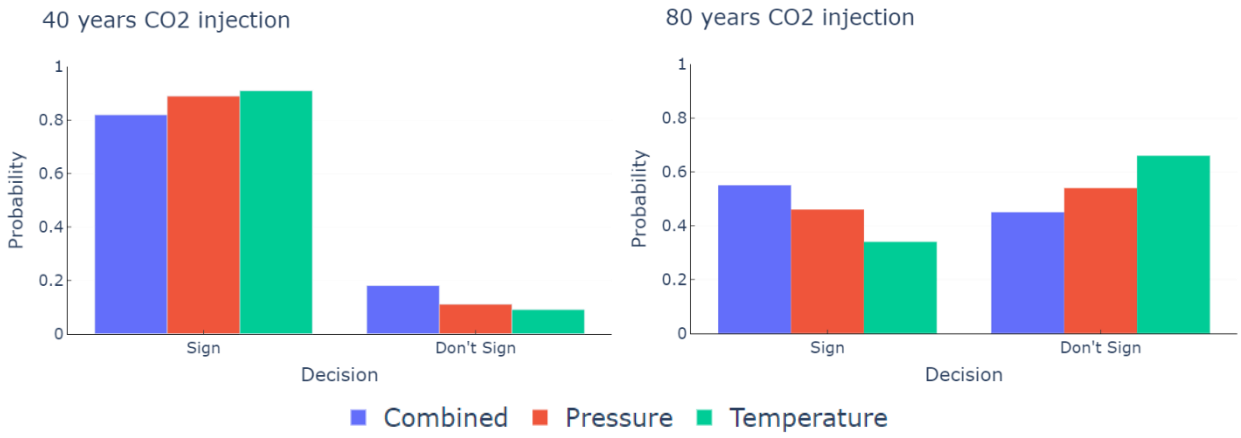


Figure 4.29 The optimal decisions of each model in scenario III

4.6 Discussion Summary

In Chapter 4, we illustrate and discuss the simulation-regression approach to estimating the VOI for individual and combined information sources. The VOI in this study is measured by an objective value that consists of 2 attributes: (i) maximize the NPV over 60 years of CO_2 storage injection and (ii) minimize the long-term leakage from the CO_2 storage. Three regression methods were tested using 10-fold cross-validation, and the KNN method was selected as the best fit for the model regression. The estimated values from the regression models were used to estimate prior and posterior values, which were applied in the VOI estimations for perfect and imperfect information. Several MCS noise values, {5%, 8%, 10%, 50%}, were added to demonstrate the effect of decreasing information accuracy in VOI estimations. The VOI calculations of imperfect information were repeated 1000 times and averaged to reduce bias from the result.

In the base model, the combined information has the highest VOI, followed by information on pressure and temperature, respectively. The VOI and the maximum price that the company should pay show a decreasing trend with increasing noise. The model identified the "Don't Sign" alternative as the optimal decision based on the prior information. Given the perfect information from all information-gathering schemes, the "Sign" alternative has a higher decision frequency

than the "Don't Sign" alternative. The frequency distributions of the optimal decisions change as the accuracy of the information decreases. The VOI assessments provide relevant and material information supporting the company's decision of whether to sign the contract or not, as well as the maximum price that the company should pay for the information. This demonstrates the role of VOI analysis in selecting the best decisions for CO_2 storage projects.

Three sensitivity scenarios were constructed to investigate the effect of carbon price uncertainty, leakage constraints, and varying injection periods. In scenario I, where a carbon price uncertainty with a mean value of \$45.1/tCO_2 is added, the prices of the information increase significantly to 20-30 times compared to the base model, although the decline is steeper than the base model when the noise increases. The frequency of not signing the contract is higher than the frequency of signing it given the result from all perfect information-gathering schemes. In Scenario II, the VOIs of individual pressure and temperature information become similar as the leakage constraint decreases. From the sensitivity analysis with different leakage constraints, we see that when the leakage constraint is equal to or lower than 10%, the expected objective value on temperature information is smaller than the expected objective value of the alternative to turn down the contract ($\boldsymbol{a}$=0). This condition will lead not only to a smaller difference between VOIs of individual pressure and temperature but also changes in the frequency distribution of the optimal decisions. Scenario III illustrates that the upper and lower limit values of NPV and leakage change as the injection period is changed from 60 to 40 and 80 years. The injection period of the 80 years model leads to a higher variance of NPV and leakage amount, compared to the 40 years injection period. Therefore, when the injection period is longer, the frequency of the "Sign" alternative is decreased while the frequency of the "Don't Sign" alternative is increased. It also shows that the price of the combined perfect information is slightly increased when the injection period increases, although the difference in information prices between each information-gathering scheme becomes larger.

The choice of regression model for simulation-regression approach is essential in the analysis. Estimates using different regression methods which has a higher MSE score of cross-validation may give higher uncertainty in VOI estimation due to over-fitted or under-fitted regression models. This also applies to the hyperparameter of the regression model, e.g., the $\boldsymbol{k}$ value in the KNN method. The VOI estimate is very sensitive to the $\boldsymbol{k}$ value and therefore it had to be controlled by cross-validation.

Chapter 5 – Conclusions & Recommendations

VOI analysis is an effective way to optimize information gathering decisions for a project with uncertain elements. In this study, we have proposed a workflow for supporting the decision of whether to enter into a CO_2 storage contract by implementing a VOI analysis. The contract decision is multi-objective as it is conditional on 2 objectives; maximize NPV and minimize long-term leakage. The workflow includes the identification of material uncertainties, swing weighting, CO_2 injection simulation, and a simulation-regression machine learning approach to approximating the posterior values that are required to estimate the VOI. The approach and results are relevant and suitable for a preliminary study of CO_2 storage projects carried out before injection commences as it includes leakage risk as a criterion. Studies on monitoring ongoing CO_2 injection programs and VOI analysis associated with leakage risk after the injection starts are discussed in Sato (2011) and Anyosa et al. (2021). For the decision situation framed in this study, the results of the VOI analysis can help the management team to determine whether the company should gather additional information, and if so which information and how much the price that the company should pay for it, to increase the quality of the injection decision.

According to Fenwick et al. (2020), one of the limitations of VOI analysis is its estimation quality depends on the model and the specification of parameter uncertainty. In this study, we have evaluated the use of three possible regression methods. However, other regression methods such as the Gaussian process, convolutional neural networks, principal components, or moving windows could also be implemented to find the optimal model for approximating the VOI. Our analysis was limited by the reservoir uncertainty, i.e. top-surface elevation, porosity, permeability, and aquifer initial conditions (pressure and temperature); while the model could also include other reservoir or valuation uncertainties that have an impact on CO_2 storage and its migration, for instance, faults and compartmentalization.

The analysis can be extended for use on actual projects with detailed revenues, costs, and taxes as most companies involved in CO_2 storage projects also produce CO_2 emissions and have to pay carbon taxes. Moreover, the methodology can also be expanded to improve the quality of decisions in choosing the best alternative among several projects.

References

Allen, R., H. M. Nilsen, K.-A. Lie, O. Møyner, and O. Andersen, 2018, Using simplified methods to explore the impact of parameter uncertainty on CO2 storage estimates with application to the Norwegian Continental Shelf: International Journal of Greenhouse Gas Control, v. 75, p. 198–213, doi:10.1016/j.ijggc.2018.05.017.

Anyosa, S., S. Bunting, J. Eidsvik, A. Romdhane, and P. Bergmo, 2021, Assessing the value of seismic monitoring of CO2 storage using simulations and statistical analysis: International Journal of Greenhouse Gas Control, v. 105, p. 103219, doi:10.1016/j.ijggc.2020.103219.

Arild, Ø., H. P. Lohne, and R. Bratvold, 2008, A Monte Carlo Approach to Value of Information Evaluation: p. 12.

Bielicki, J. M., C. A. Peters, J. P. Fitts, and E. J. Wilson, 2015, An examination of geologic carbon sequestration policies in the context of leakage potential: International Journal of Greenhouse Gas Control, v. 37, p. 61–75, doi:10.1016/j.ijggc.2015.02.023.

Bonate, P. L., 2001, A Brief Introduction to Monte Carlo Simulation: Clinical Pharmacokinetics, v. 40, no. 1, p. 15–22, doi:10.2165/00003088-200140010-00002.

Bock, B., R. Rhudy, H. Herzog, M. Klett, J. Davison, D. G. D. L. T. Ugarte, and D. Simbeck, 2003, Economic Evaluation of CO2 Storage and Sink Enhancement Options, Final Technical Report: TVA Public Power Institute, doi:10.2172/826435.

Bratvold, R. B., J. E. Bickel, and H. P. Lohne, 2009, Value of Information in the Oil and Gas Industry: Past, Present, and Future: Society of Petroleum Engineers, p. 9, doi:10.2118/110378-PA.

Bratvold, R. B., and S. H. Begg, 2010, Making Good Decisions: Richardson, TX : Society of Petroleum Engineers.

Breiman, L., 2001, Random Forests: Machine Learning, v. 45, no. 1, p. 5–32, doi:10.1023/A:1010933404324.

Brownlee, J., 2018, A Gentle Introduction to k-fold Cross-Validation. Available at: https://machinelearningmastery.com/k-fold-cross-validation/ [Accessed 1 June 2021].

Cavanagh, A. J., and R. S. Haszeldine, 2014, The Sleipner storage site: Capillary flow modeling of a layered CO2 plume requires fractured shale barriers within the Utsira Formation: International Journal of Greenhouse Gas Control, v. 21, p. 101–112, doi:10.1016/j.ijggc.2013.11.017.

Chaiyapo, N., and N. Phewchean, 2017, An application of Ornstein-Uhlenbeck process to commodity pricing in Thailand: Advances in Difference Equations, v. 2017, no. 1, p. 179, doi:10.1186/s13662-017-1234-y.

Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Solazzo, E., Monforti-Ferrario, F., Olivier, J.G.J., Vignati, E., 2020, Fossil CO2 emissions of all world countries - 2020 Report, EUR 30358 EN: Publications Office of the European Union, Luxembourg, ISBN 978-92-76-21515-8, doi:10.2760/143674, JRC121460.

Dutta, G., T. Mukerji, and J. Eidsvik, 2019, Value of information analysis for subsurface energy resources applications: Applied Energy, v. 252, p. 113436, doi:10.1016/j.apenergy.2019.113436.

Eidsvik, J., T. Mukerji, and D. Bhattacharjya, 2015, Value of Information in the Earth Sciences: Cambridge University Press.

Eidsvik, J., G. Dutta, T. Mukerji, and D. Bhattacharjya, 2017, Simulation–Regression Approximations for Value of Information Analysis of Geophysical Data: Mathematical Geosciences, v. 49, no. 4, p. 467–491, doi:10.1007/s11004-017-9679-9.

Fenwick, E., L. Steuten, S. Knies, S. Ghabri, A. Basu, J. F. Murray, H. (Erik) Koffijberg, M. Strong, G. D. Sanders Schmidler, and C. Rothery, 2020, Value of Information Analysis for Research Decisions—An Introduction: Report 1 of the ISPOR Value of Information Analysis Emerging Good Practices Task Force: Value in Health, v. 23, no. 2, p. 139–150, doi:10.1016/j.jval.2020.01.001.

Grootendorst, M., 2019, Validating your Machine Learning Model. Available at: https://towardsdatascience.com/validating-your-machine-learning-model-25b4c8643fb7 [Accessed 1 June 2021].

Hastie, T., R. Tibshirani, and J. H. Friedman, 2001, The elements of statistical learning: data mining, inference, and prediction. New York: Springer.

IEA (2021), Is carbon capture too expensive?, IEA, Paris https://www.iea.org/commentaries/is-carbon-capture-too-expensive

Irlam, L., 2017, Global Costs of Carbon Capture and Storage: Global CCS Institute Publications.

IPCC, 2018: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty [Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (eds.)]. In Press.

James, G., D. Witten, T. Hastie, and R. Tibshirani, 2013, An introduction to statistical learning: with applications in R: Springer-Verlag New York, Springer texts in statistics., XIV, 426 p.

Lie, K.-A., 2019, An Introduction to Reservoir Simulation Using MATLAB/GNU Octave: SINTEF, Norway, Cambridge University Press.

Liu, Y., J. Li, S. Sun, and B. Yu, 2019, Advances in Gaussian random field generation: a review: Computational Geosciences, v. 23, no. 5, p. 1011–1047, doi:10.1007/s10596-019-09867-y.

Mahaboob, B., B. Venkateswarlu, C. Narayana, J. Ravi sankar, and P. Balasiddamuni, 2018, A Treatise on Ordinary Least Squares Estimation of Parameters of Linear Model: International Journal of Engineering & Technology, v. 7, no. 4.10, p. 518, doi:10.14419/ijet.v7i4.10.21216.

Norwegian Petroleum Directorate, 2019, CO2 Storage Atlas Norwegian North Sea. Available at: https://www.npd.no/en/facts/publications/co2-atlases/co2-storage-atlas-norwegian-north-sea/ [Accessed 1 April 2021].

Pedregosa, F. et al., 2011, Scikit-learn: Machine Learning in Python: Machine Learning in Python, p. 6.

Puerta-Ortega, C., J. E. Bickel, and S. Hovorka, 2013, Assessing the value of permeability data in a carbon capture and storage project: International Journal of Greenhouse Gas Control, v. 17, p. 523–533, doi:10.1016/j.ijggc.2013.06.003.

Quandl, 2021, ECX EUA Futures. Available at: https://www.quandl.com/data/CHRIS/ICE_C1 [Accessed 25 March 2021].

SINTEF, 2016a, MRST-CO2lab. Available at: https://www.sintef.no/projectweb/mrst/modules/CO2lab/ [Accessed 5 April 2021].

SINTEF, 2016b, An Introduction to Reservoir Simulation Using MATLAB: SINTEF ICT, Norway, Department of Applied Mathematics.

Sato, K., 2011, Value of information analysis for adequate monitoring of carbon dioxide storage in geological reservoirs under uncertainty: International Journal of Greenhouse Gas Control, v. 5, no. 5, p. 1294–1302, doi:10.1016/j.ijggc.2011.07.010.

Temitope, A., N. Sulemana, J. S. Gomes, and R. Oppong, 2016, Statistical Uncertainty Analysis and Design Optimization of CO2 Trapping Mechanisms in Saline Aquifers, *in* Day 1 Mon, November 14, 2016, International Petroleum Technology Conference, Bangkok, Thailand: IPTC, p. D012S055R002, doi:10.2523/IPTC-18837-MS.

United Nations, 2015, Paris Agreement: United Nations Framework Convention on Climate Change.

Van der Zwaan, B., and K. Smekens, 2009, CO2 Capture and Storage with Leakage in an Energy-Climate Model: Environmental Modeling & Assessment, v. 14, no. 2, p. 135–148, doi:10.1007/s10666-007-9125-3.

Vinca, A., J. Emmerling, and M. Tavoni, 2018, Bearing the Cost of Stored Carbon Leakage: Frontiers in Energy Research, v. 6, p. 40, doi:10.3389/fenrg.2018.00040.

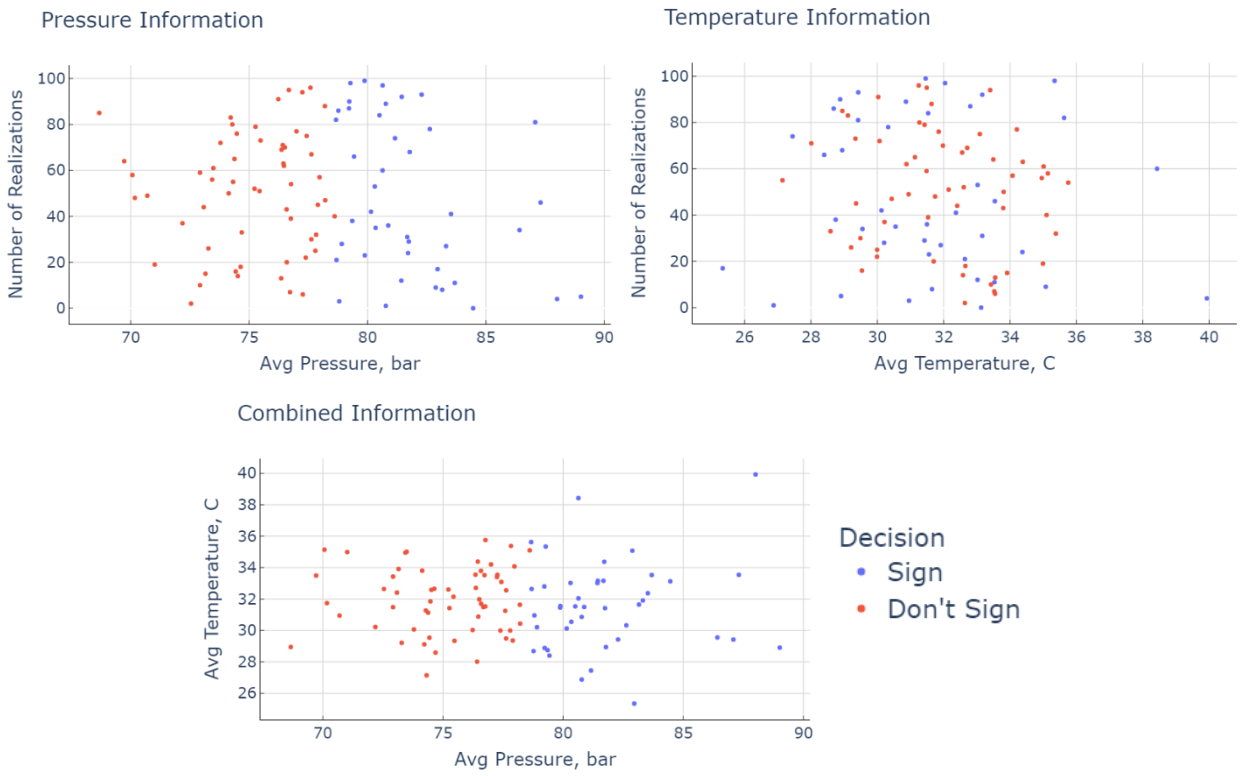
Zero Emissions Platform, 2011, The Costs of CO2 Storage: Global CCS Institute Publications.

Zhang, D., and J. Song, 2014, Mechanisms for Geological Carbon Sequestration: Procedia IUTAM, v. 10, p. 319–327, doi:10.1016/j.piutam.2014.01.027.

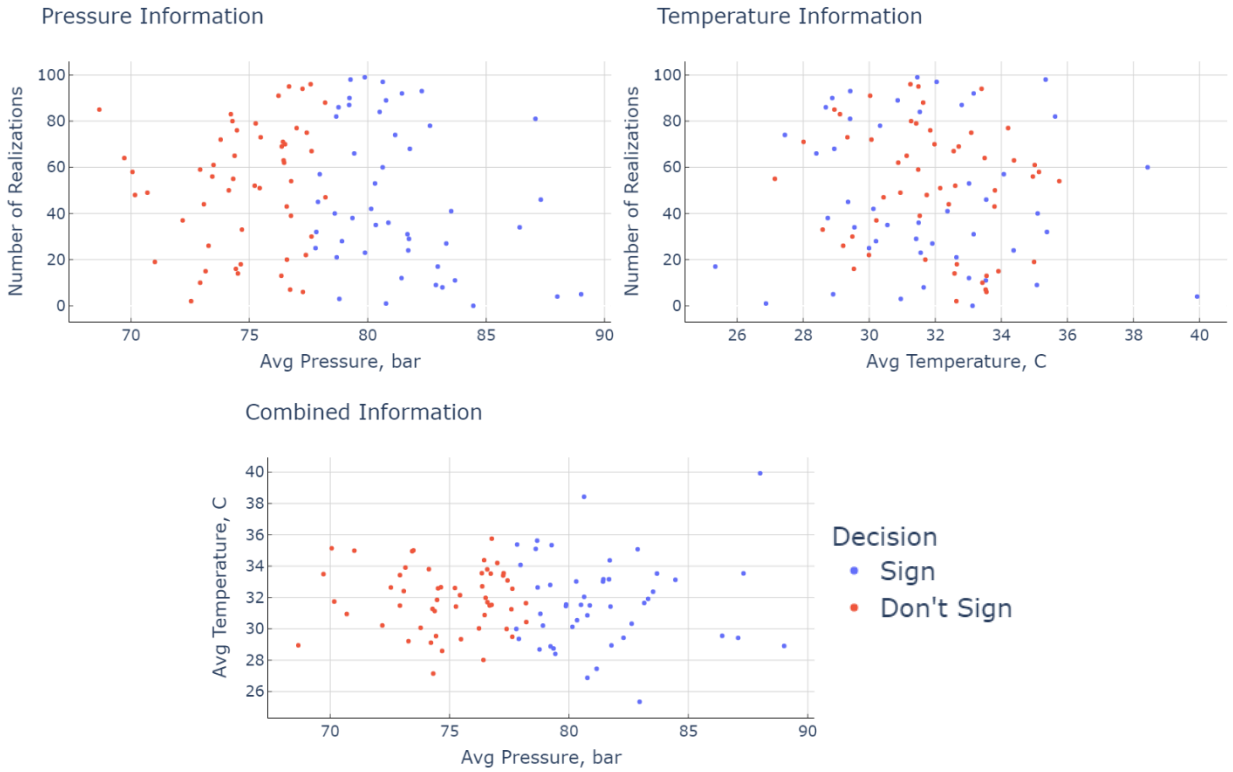
Zhang, B., L. A. Grzelak, and C. W. Oosterlee, 2012, Efficient pricing of commodity options with early-exercise under the Ornstein–Uhlenbeck process: Applied Numerical Mathematics, v. 62, no. 2, p. 91–111, doi:10.1016/j.apnum.2011.10.005.

Appendix 1 – Optimal Decisions based on Perfect Information Outcomes

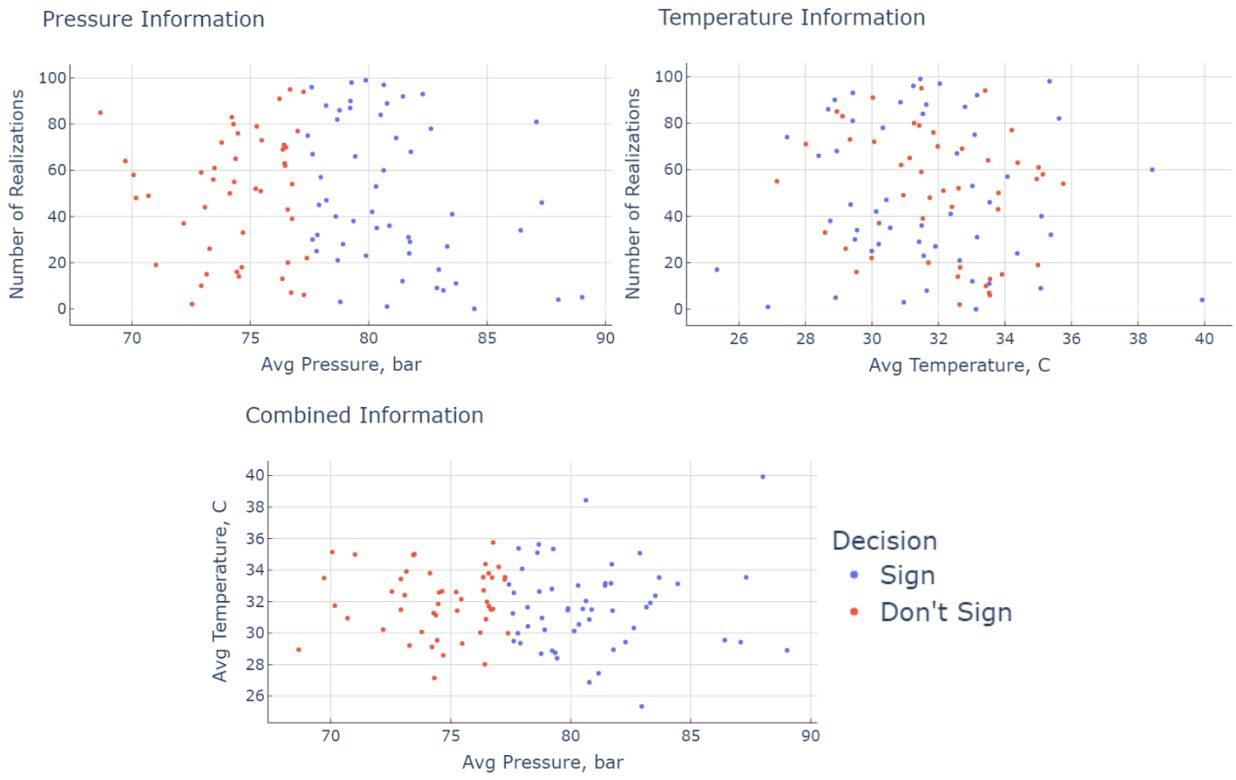
Scenario II – 5% maximum leakage constraint



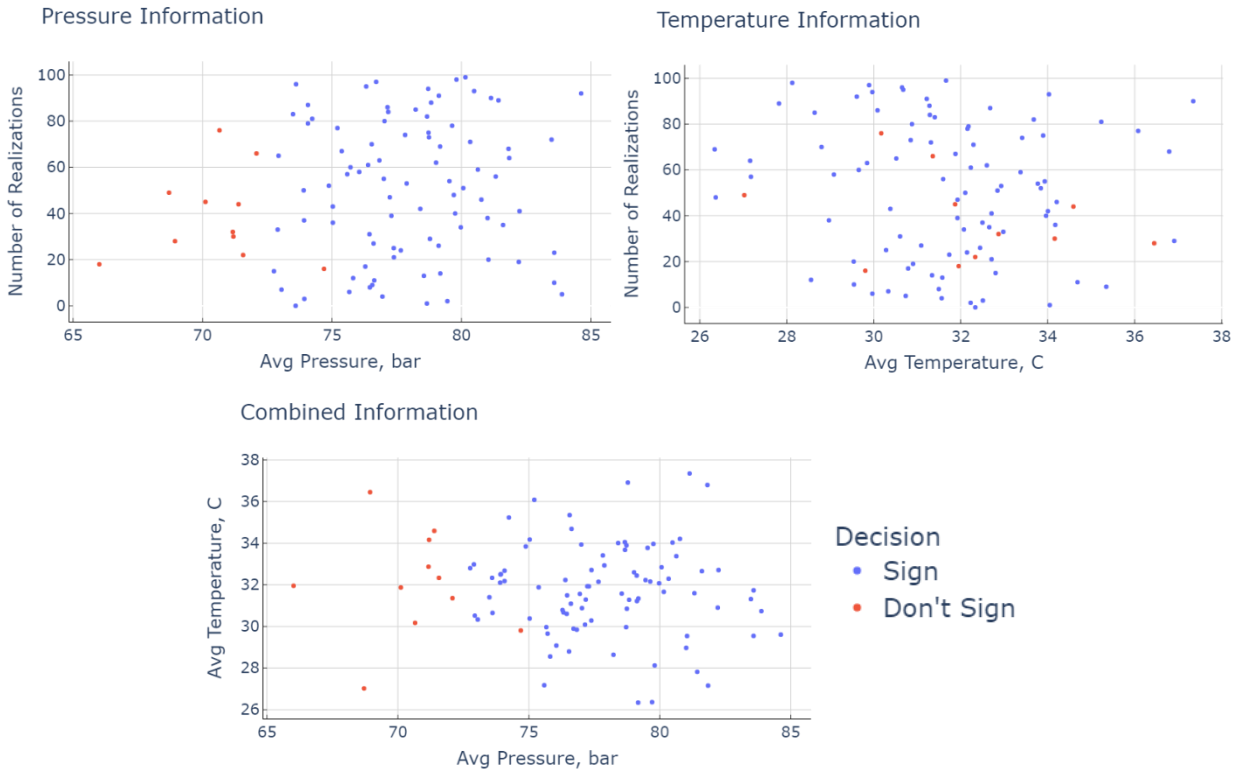
Scenario II – 10% maximum leakage constraint



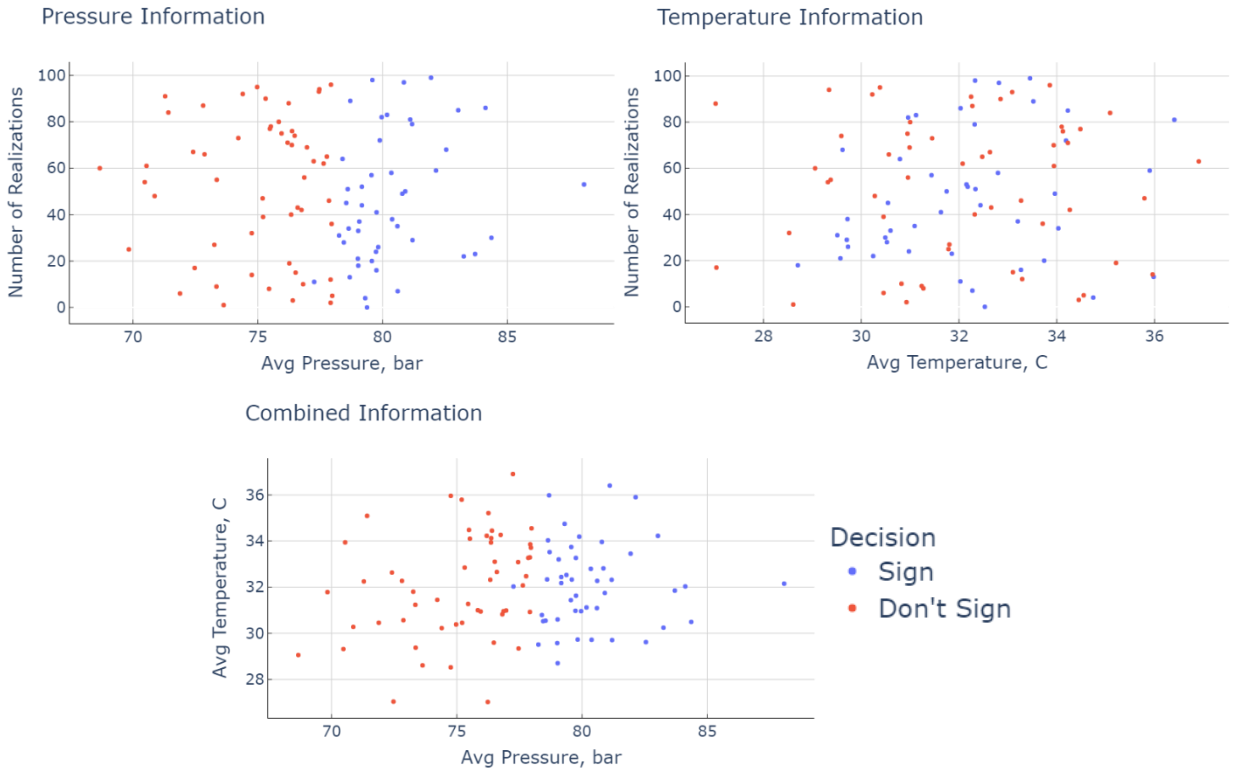
Scenario II – 20% maximum leakage constraint



Scenario III – 40 years injection period

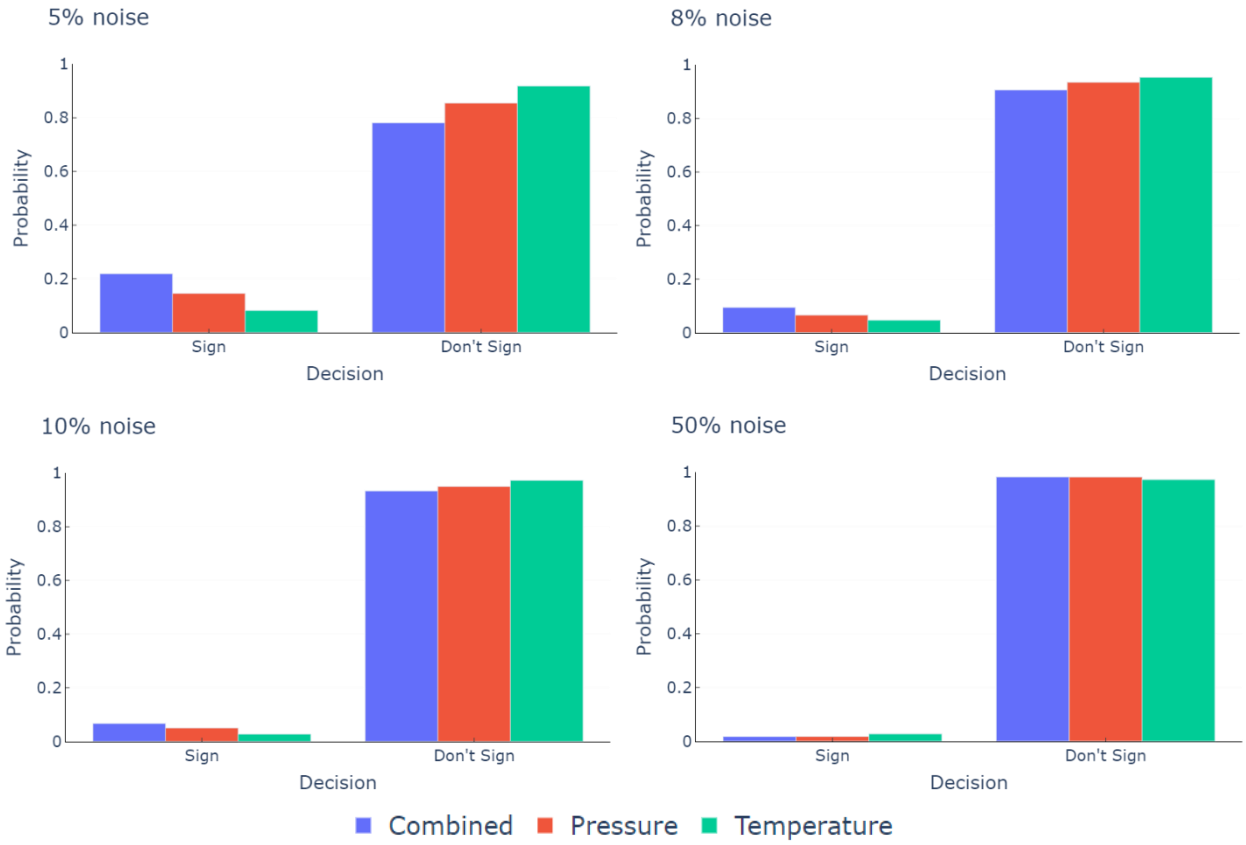


Scenario III – 80 years injection period

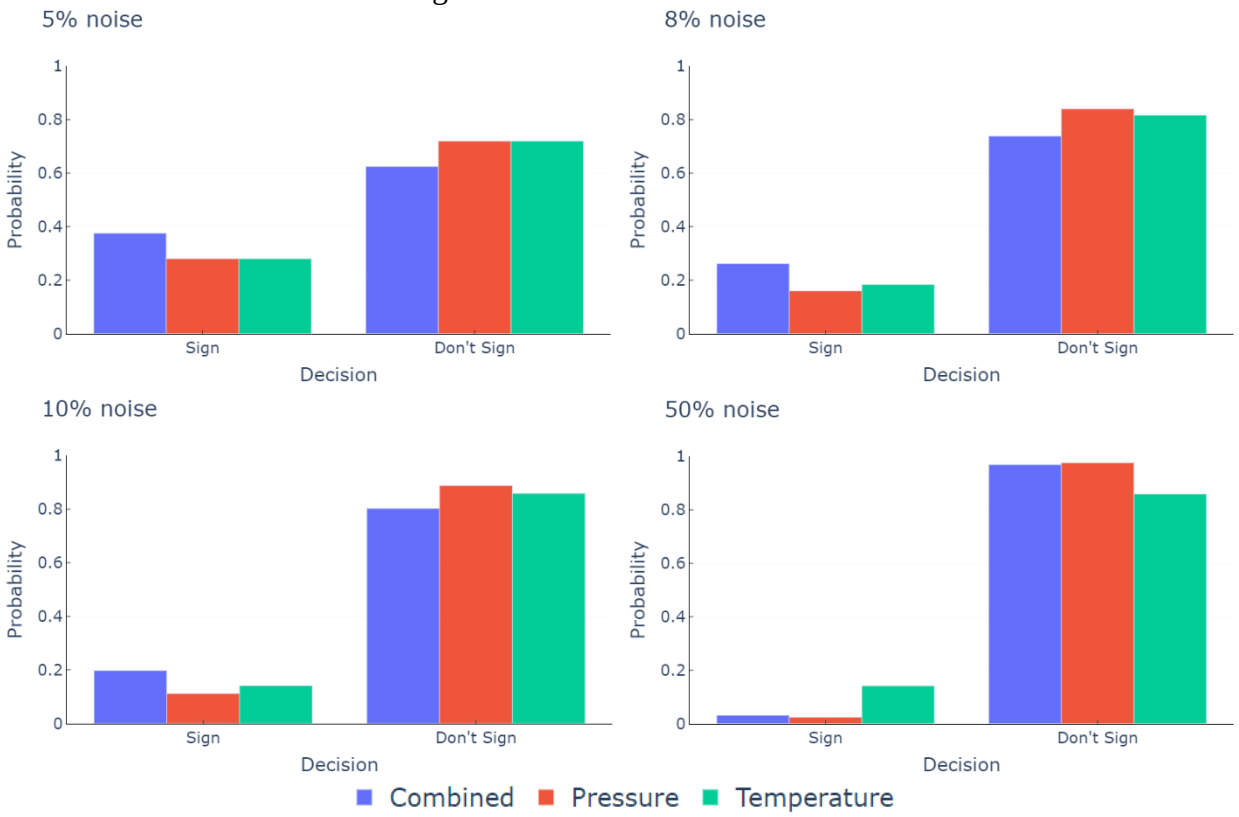


Appendix 2 – Optimal Decisions with Noises

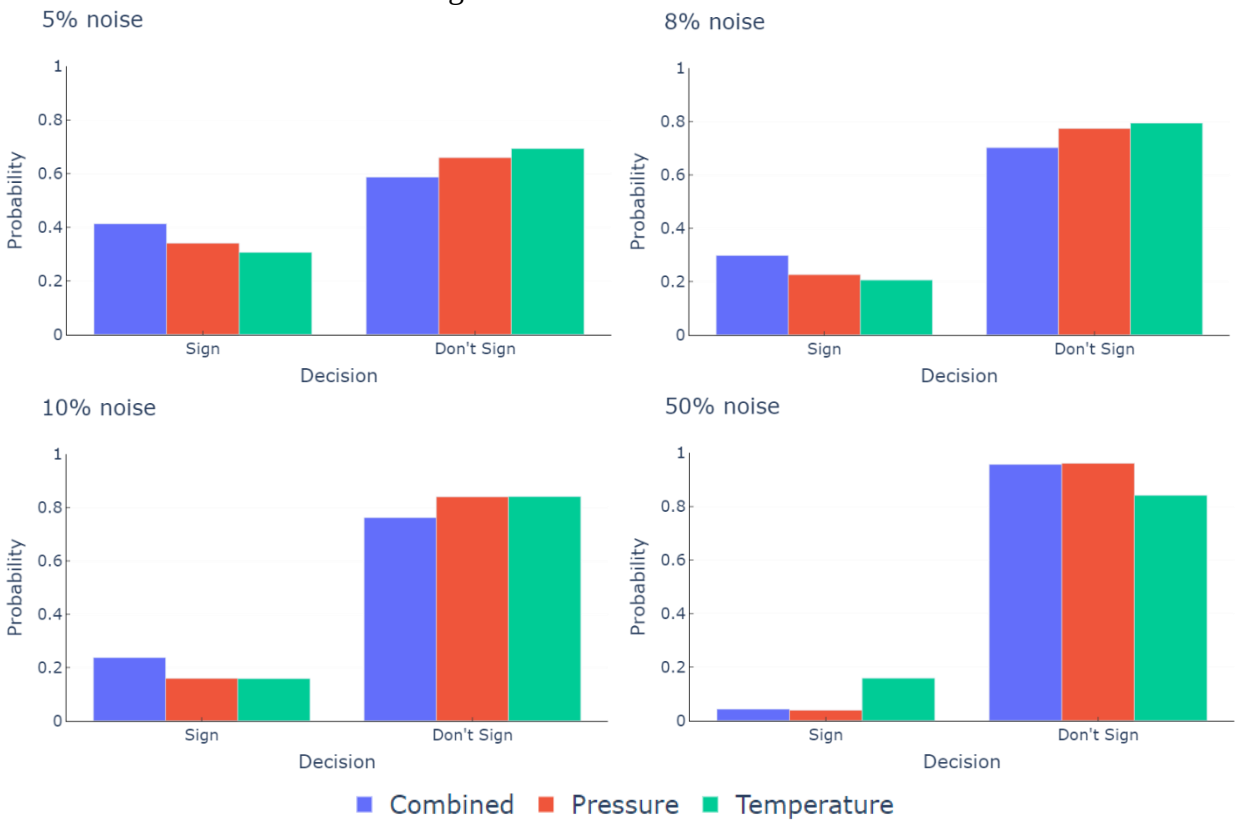
Scenario I



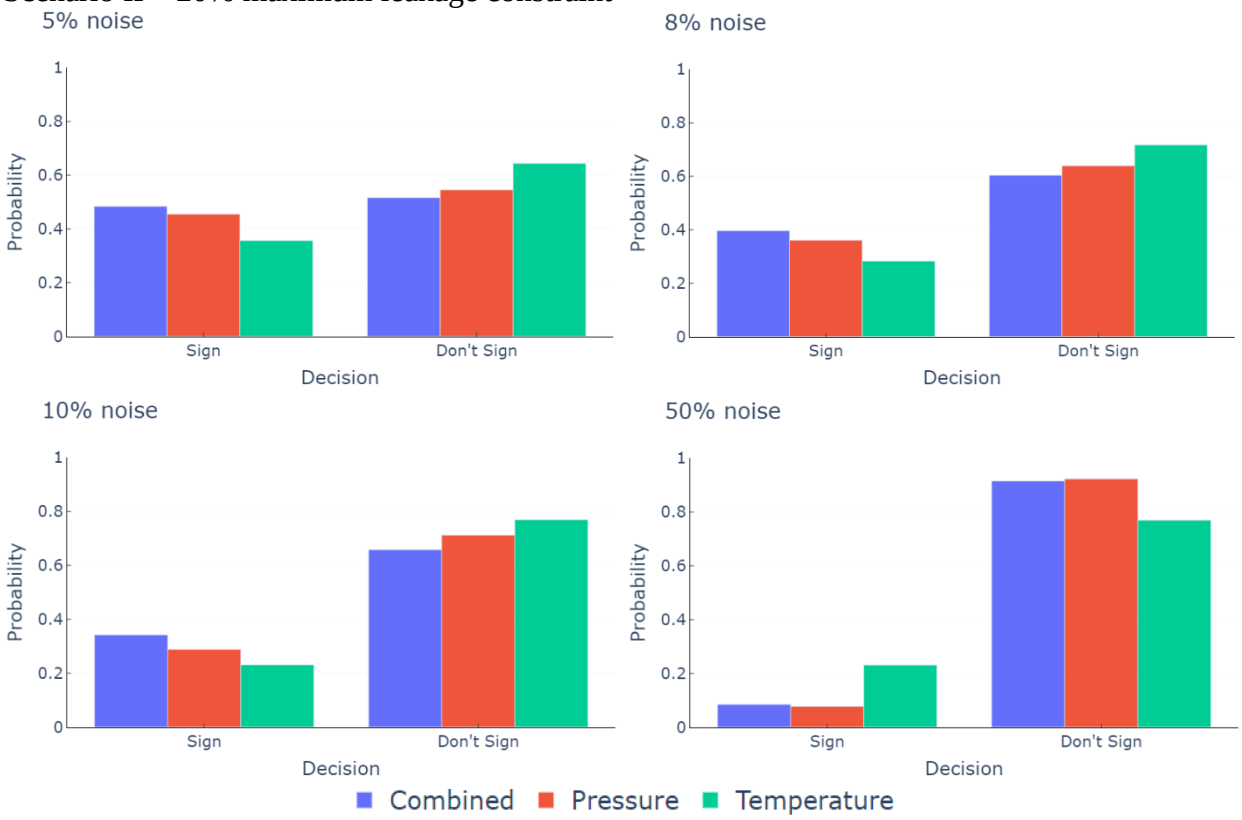
Scenario II – 5% maximum leakage constraint



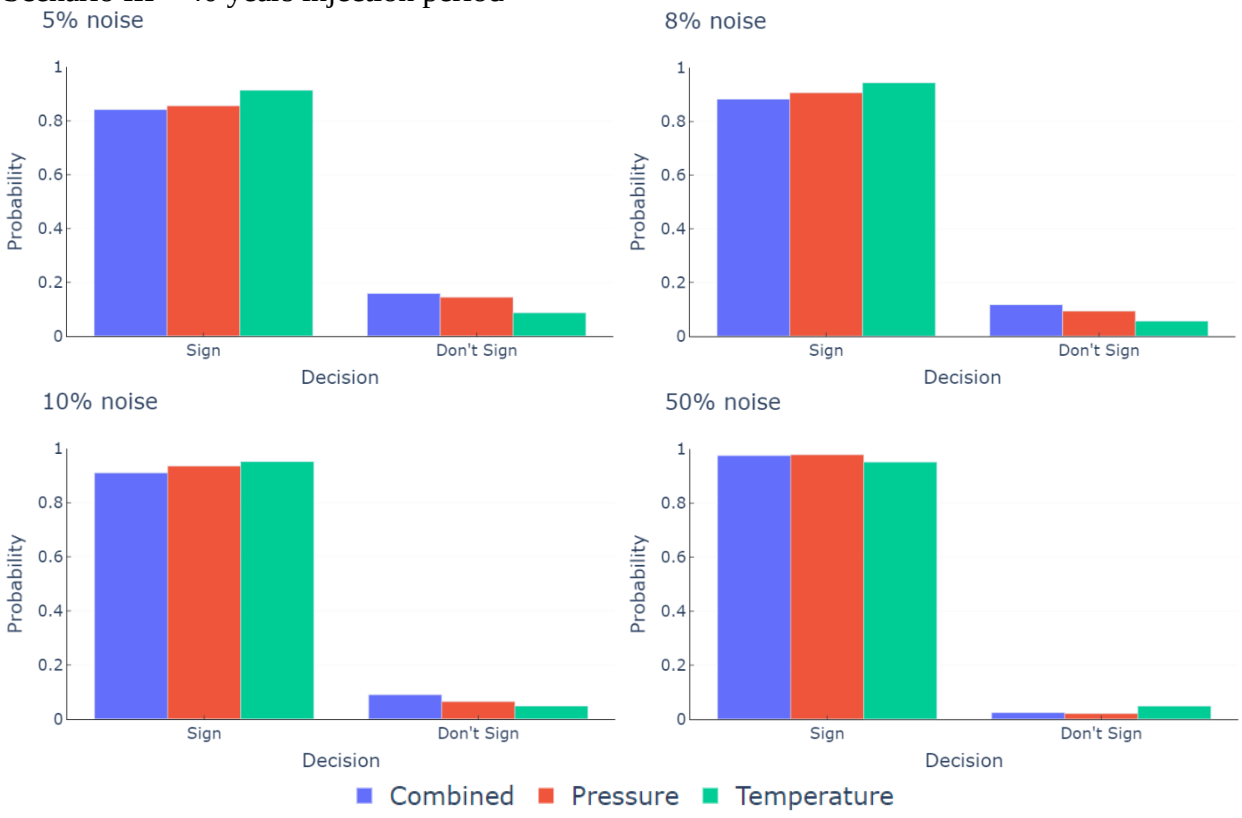
Scenario II – 10% maximum leakage constraint



Scenario II – 20% maximum leakage constraint



Scenario III – 40 years injection period



Scenario III – 80 years injection period