Bayesian Evidential Learning applied to Mineral Resource Modelling to reduce uncertainties

Aafje Houben
4732677

Bachelor Thesis
Delft University of Technology
Department of Geoscience and Engineering

Delft, 27 October 2022
Supervisors: M. S. Shishvan &
A. D. Daniilidis & G. Rongier

1 Abstract

Mineral Resource Modeling (MRM) is used to predict the properties of an orebody, however it does not come without uncertainties. Multiple approaches can be used to reduce the latter. Due to limited knowledge about the subsurface, predictions are difficult to be made. In this thesis the application of Bayesian Evidential Learning (BEL), in order to reduce uncertainties on MRM, will be researched. The uncertainties present in MRM will be linked to the knowledge obtained from case studies in different geological domains where BEL has been applied successfully and reduced certain parameter uncertainties. The gap in todays industry is the knowledge and proof that using BEL for MRM will reduce uncertainties and risks, works effectively and will consume less money.

The aim of this study is to show that BEL is a useful approach to reduce uncertainty of the predictions in MRM. The success of a project is supported by the accuracy of the model utilized and the geological interpretation. BEL is a framework based on statistical relationships between data and prediction variables. It will predict the posterior distribution of the prediction variable.

The various case studies that have been discussed are (Hermans et. al, 2019), (Hermans et. al, 2018), (Thibaut et. al, 2021) and (Tadjer and Bratvold, 2021). BEL has successfully been able to reduce uncertainties related to geological problems of the following:

- The temperature in an alluvial aquifer
- The efficiency of the thermal energy storage capacity in an alluvial aquifer
- The wellhead protection areas surrounding the pumping well using tracing experiments as predictors
- The prediction of leakages of CO_2 and the storage of CO_2

Thus, BEL can be seen as a potential approach to also reduce uncertainties in a mineral resource domain; the research of this thesis.

In attempt to prove this, a descriptive case study on Tropicana Gold Mine has been executed. The aim is to show that the use of BEL will, based on the geological and geochemical properties such as lithology, grade and mineral type, reduce the uncertainty in the prediction of a geometallurgy property, namely the hardness of a rock. The six steps of BEL's framework will be followed consisting of Monte Carlo simulations, Principal Component Analysis (PCA) and Canonical Correlation Analysis (CCA) in order to obtain relations between the data and predictor variables. Where the data variables are from exploration drillhole data and prediction variable the hardness of the rock. It shows that BEL is able to reduce the uncertainty of a geometallurgy property by using geological and geochemical properties. Meaning BEL can be applied to MRM to reduce uncertainties.

Table of contents

2 Introduction

In todays world the use of cell phones is a very straight forward assumption, considered weird when not having one. Never standing still by the fact that to get such a phone, mineral resources are needed. The exploitation of the latter can have huge impacts and disasters are not always excluded. Why is it needed to have a small uncertainty and how can you reduce such an uncertainty?

The main hypothesis of this study is whether Bayesian Evidential Learning (BEL) can be applied to reduce the uncertainty in Mineral Resource Modeling (MRM). In order to test this hypothesis the following research questions are defined.

- What are the uncertainties present in MRM
- How is BEL applied in todays industry
- How can BEL be used in the MRM industry to reduce the uncertainty of the parameters

BEL will be explained within a geostatistical framework. The subsurface is so complex that any prediction is subject to large uncertainties. Prediction is not sufficient, but an entire uncertainty quantification (UQ) is needed for a proper risk analysis and subsequent decision making.

This study contains a literature review of both MRM and BEL, concluding with a literature synthesis in Chapter 3. Then the link is explained between MRM and BEL in Chapter 4 followed by a discussion and recommendations for further research in Chapter 5, conclusion in Chapter 6.

3 Literature Review

This chapter introduces MRM, its uses, how it is applied and its main uncertainties; next BEL will be explained and shown how it is applied in four different case studies. The chapter concludes with a synthesis of literature review with a short summary of the main findings.

3.1 Mineral Resource Modeling

3.1.1 Why is Mineral Resource Modeling performed

Mineral resources from the Earths subsurface are certainly part of today's industry. Mineral resources are crucial for industrial processes, contemporary transportation system construction and social growth. Without these minerals, the worlds living standards would crash because they are essential to ones daily life. All of the shallow, easy accessible deposits seem to have been found and exploited. The remaining deposit domains containing mineral resources create serious difficulties since they could be marginal deposits, happening at considerable depths or having other access limitations [Leuangthong, 2021].

The following defines what a mineral resource is: mineral resources are concentrations or occurrences of naturally occurring, solid, inorganic, or fossilized organic material within or on the crust of the Earth that are of such quantity, grade, or quality to have a good chance of being economically extracted. From specific geological evidence and information, the location, quantity, grade, geological properties, and continuity of a Mineral Resource are known, estimated, or interpreted[Dohm, 2005].

MRM is done to get an accurate estimation of the Mineral in a certain domain. Mineral resource estimation entails applying a variety of estimating techniques to determine the grade and quantity of a Mineral deposit based on its geological features [Dumakor-Dupey and Arya, 2021]. In the fields of science, industry, and government, there is an increasing amount of data available. Modeling plays a part in the opportunities and challenges that are being created by this. Regardless of the project's size, deposit type, or product, extracting these mineral resources will always be difficult. In order to ensure that the data points are correlated to the actual resource, the exploitation of mineral resources requires a certain level of trust in the persistence of materialization[Dumakor-Dupey and Arya, 2021]. In order for the project to be successful, the estimation must be trustworthy and to provide the project with technical support for the extraction of geoscience and the quantitative estimation of mineral resources, MRM is therefore required[Wang and Huang, 2012].

Modeling is used to gain a better understanding of the deposit's mineral system. Models of mineral resources provide the framework and the underlying support for the mineral investigation. More information about geological uncertainty, grade severity, and quantification and variability errors in predicted geotechnical parameters will be made available[Dumakor-Dupey and Arya, 2021]. MRM provides more trust in geoscientific data and the resource estimation that goes along with a grade estimate[Dohm, 2005]. This confidence is needed since all risk analysis, technical and financial, is based on the knowledge and confidence of the project and that is why MRM is needed to create the best possible understanding of the project domain [Abzalov, 2016].

Dealing with large data sets, such as seismic data transfer or weather forecasting, can be difficult in the field of Earth sciences. In this geospatial setting, modeling is frequently carried out under the premise of conventional statistical learning. A sample of something, like soil, can only be taken once from a certain area. Therefore, the basis of the truth is a single random variable realization. Therefore, it would be possible to take further measurements across the study region to obtain a spatial distribution of samples. However, this is not the same because they can vary due to varied structural shaping because it is connected to geological processes. Therefore, some assumptions must be made in order to use these samples [Bogrash, 2020]. MRM increases the capacity for comprehension and trust in geoscience [Dohm, 2005].

3.1.2 What are the uncertainties regarding Mineral Resource Modeling

The success of a project is supported by the accuracy of the model utilized and the geological interpretation. It is essential to explicitly define the characteristics and the geological parameters that have influence on the estimation on the mineral resources. The influence of these parameters is based on the minerals present, and its materialization style [Abzalov, 2016], the spatial variability and the uncertainty of the geological formations[Dumakor-Dupey and Arya, 2021]. Major failures can occur as a result from an incorrect degree of the uncertainty estimation in the geological models.

Mineral resources are divided into three different groups. When the grade, mineral and tonnage content can be estimated with low confidence it is called inferred mineral resources. Indicated mineral resources is when the level of confidence is reasonable to estimate its tonnage, shape, physical characteristics, grade, densities and mineral content. Lastly, measured mineral resources is when all the content can be estimated with a high level of confidence. This is shown in Figure 1.

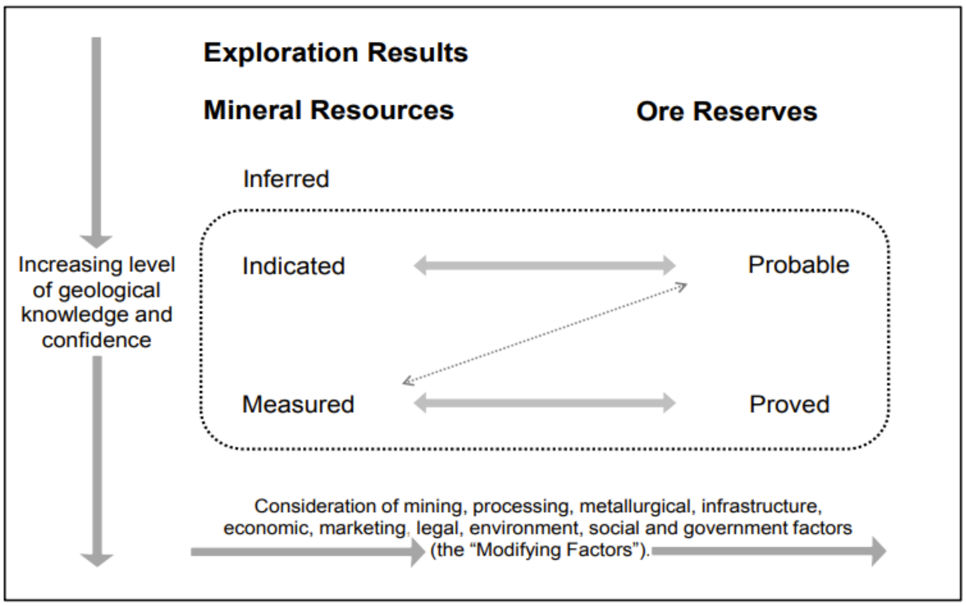

Figure 1: General definition and relationship between the results, resources and reserves[Committee, 2012].

The parameters influencing the uncertainty regarding the MRM process are the following:

- Geometry of the ore body
  The first parameter influencing the uncertainty is the geometry of the ore body. A too big uncertainty on the geometry of the ore body leads to an inaccurate resource domain. Shape and volume can be influenced purely based on a geological interpretation leading to errors in the resource tonnages that are estimated [Abzalov, 2016]. Not including the shapes uncertainty of the domain causes an underestimated uncertainty in the resource estimates resulting that a bad overview is obtained of where the project will take place. Therefore it is important to quantify the degree of uncertainty of this particular parameter. When the geometry of an ore body changes, the tonnage and grade reserves will change as well. The depth of an open pit mine is depending on the geometry of the ore body. Shown in Figure 2 are the kinematic controls on the geometry of the orebody, a lot of factors affect the geometry. Foliations are typical in all grades of metamorphic rocks and are systematically linked to tectonic deformation. Axis plane foliations are thought to be primarily caused by ductile smoothing and the parallel alignment of platy minerals. It might have been created by the parallel alignment

of pebbles, minerals, or fossils during deformation [Jpb, 2020]. It can provide information on regional stress and plate tectonic analyses as well as the direction of increased strain. Different degrees of temperature and pressure can be deduced from the sorts of minerals that are present. Any recurring, typically penetrative, and parallel alignment of linear features within a rock is referred to as lineation. These two controls can be caused by differential stress, it shows the response on the stress and strain because of variations in layer thickness, viscosity, the contrast between the viscosities of the matrix and the layers, mechanical anisotropy and more [Nabavi and Fossen, 2021].

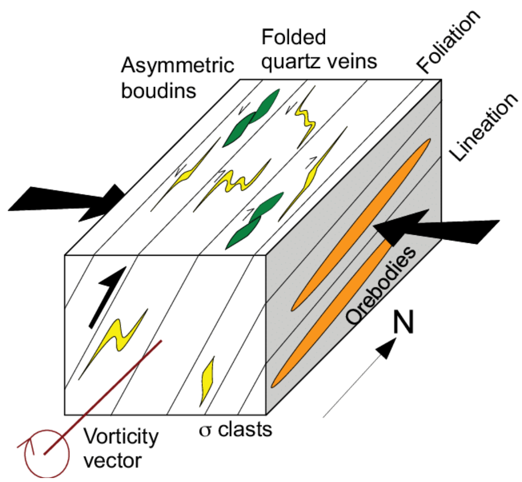

Figure 2: Kinematic controls on the orebody geometry at the Golden Pig mine[Blenkinsop et al., 2020].

- Characterization of the contacts, sharp or gradational, straight or irregular
  Another parameter is the characteristics of the geological contacts, the vertical profile, the sequence of the material present and the changes in chemistry are looked at. The grade distribution profiles will give the information about its type of contact; either sharp or gradational shown in Figure 3.

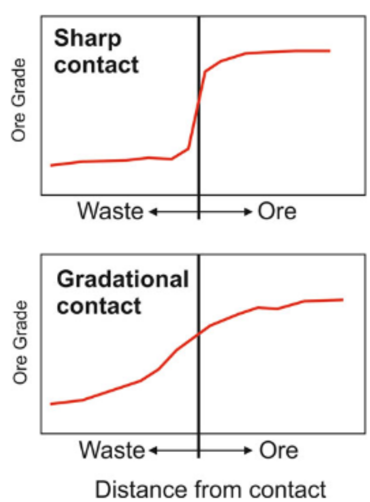

Figure 3: Shown what the difference is between a sharp contact and a gradational contact[Abzalov, 2016].

The deposit's grade distribution profile can vary depending on the domain, the drill hole and the section of the deposit. A gradational contact is a gradual, slow change in the deposition. A sharp

contact is an abrupt change in the sediment type and a minor depositional break is present. It gives information about where the resource is located. A sharp contact can mean that the distribution of the resource can change fast from no resource present to a very high grade value present. An accurate type of contact is needed for the estimation of the estimated resources' location. A small inaccuracy in the model can lead to a big error. When sharp contacts are present, it is needed to have a hard-boundary approach to prevent an inaccurate location [Abzalov, 2016]. When a sharp contact is present, the geometry is relatively more simple than a gradational contact however this does not mean that the uncertainty is greatly reduced. It can be caused by lack of information [Dominy et al., 2002]. The layers change slowly when the contact type is gradational. Having a gradational contact means that the tonnage is depending on the chosen cut-off grade so this means indirectly on the economic parameters[Dominy et al., 2002].
Another difference that need to be made in characteristics of the contacts is that a contact can be straight or irregular shown in Figure 4. It does not depend on it's contacts type. It is visible, presented on a topographic map. The higher the degree of irregularity, the more subjective is the approach which can result in different correlations found by different geologists [Dominy et al., 2002].

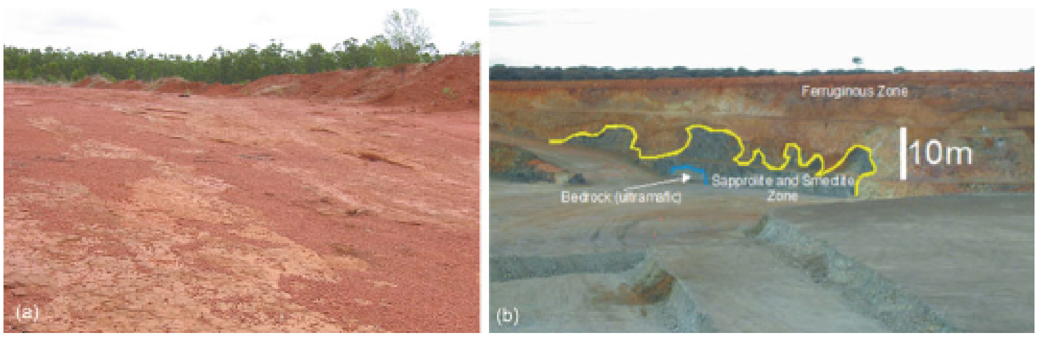

Figure 4: Shown in (a) a straight footwall, a smooth transition and in (b) an open pit wall with highly irregular contact, marked by the yellow line[Abzalov, 2016].

This parameter has impact on the estimation of excessive mining dilution and/or unforeseen mining losses.

- Internal structure of the ore body, geochemical and mineralogical zoning and layering
  For MRM, its first steps are based on data from for example drill holes data, an example is shown in Figure 5. To get an accurate model based on this might not be feasible leading to an incorrect representation of the internal structure of the ore body. Knowing the relationship between the rocks and the structure of the domain provides information about the failure geometry and the kinematics. A minerals structure depends on its block size which is depending on the spatial distribution patters of the samples and the continuity of the grade [Abzalov, 2016] or doing it the geostatistical way through simple kriging by looking at the weight of the mean.

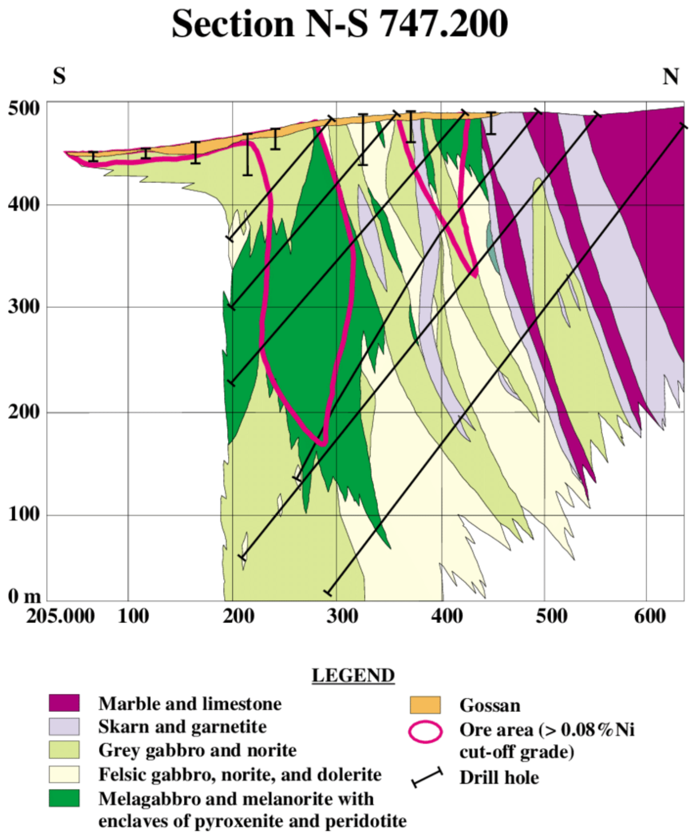

Figure 5: The internal structure based on drillhole data of the norther border of the Aguablanca orebody and host rocks[Ortega et al., 2004].

The mineralogical and geochemical data influences the metallurgical characterisation of the mineralization. The chosen process is of great importance for further procedure of the exploitation of the mineral resource. A wrong defined metallurgical zone can cause estimation inaccuracies for the relative proportion of various materials [Dominy et al., 2002]. Is is needed to characterize the process of exploring a material's intrinsic structure and qualities using methods form outside the material. Variables that have affect on this process is rock weathering, textures, grindability [Abzalov, 2016]. It influences to decision for the boundaries for the objects that are situated far below the surface [Wang and Huang, 2012].
When changes in the geochemical and mineralogical zoning is falsely identified it can result in an overvaluation of intersections because such changes may have an effect on metallurgical recoveries and the concentrations of harmful elements [Dominy et al., 2002]. This parameter is mostly based on prior geological knowledge and interpretation. When an inaccurate uncertainty is used it results in a wrong diagnosis of the ore body. This has a great influence during the estimation of local reserves before production. It is critical for the development of the project [Dominy et al., 2002].

- The presence of the multiple generations of mineralization and differences of their structural controls
  The structural controls knowledge is used for the characteristics of the origin of the deposited material. A incorrect uncertainty can lead to mistakes concerning the mineral exploration, the evaluation of the ore deposits and planning of the project. It can be passive, meaning tectonic structures were developed before mineralization took place or active, meaning tectonic structures are developing together with mineralization [Funedda et al., 2018]. An example of multiple generations of mineralization is shown

in Figure 6.

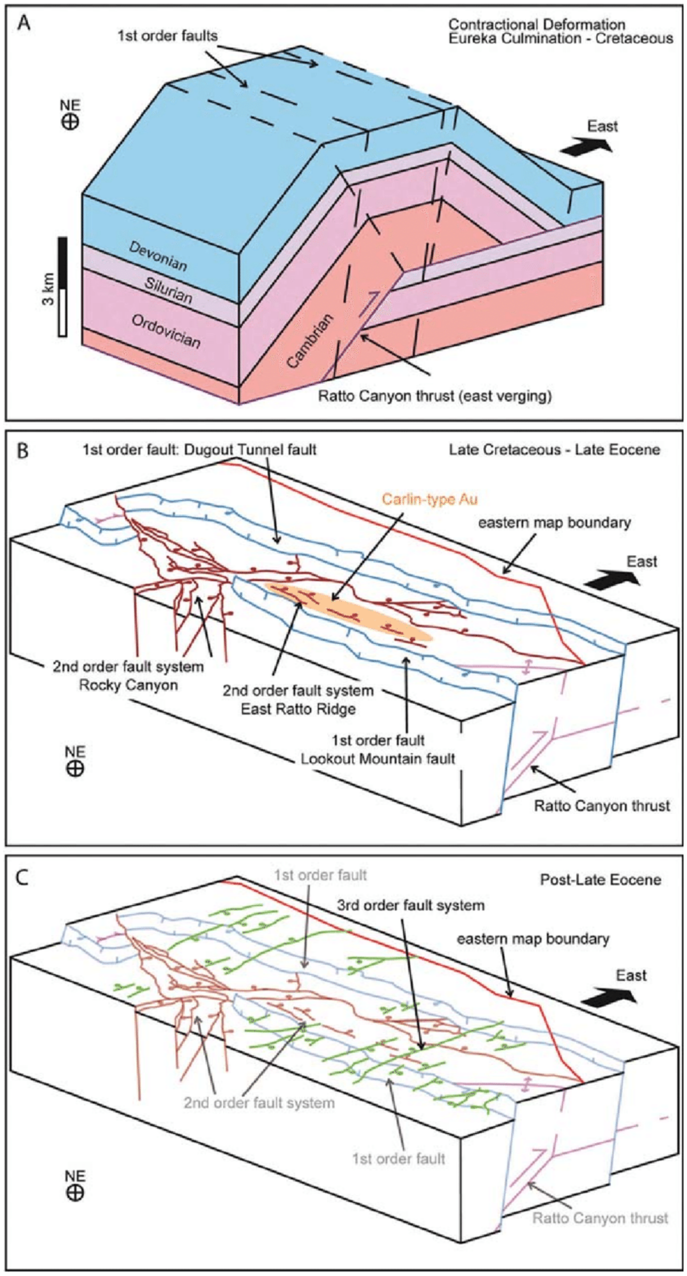

Figure 6: Showing the evolution of the structural geometry and mineralization[Di Fiori et al., 2015].

- The presence of the internal waste within the ore body-intrusion in the ore body
  The next parameter is the presence of the internal waste within the ore body. The internal waste that is intruded into the ore body can result in too much dilution of the ore [Abzalov, 2016]. Internal waste can be characterised by delineated contours. To determine whether there is waste rock, the soil's coefficient of permeability must be taken into account and its leachability [Ambrosone et al., 2019].

- Distribution of the weathered and oxidized rocks affecting the stabilities of the pit walls or underground

workings
The distribution of the rocks that are weathered and oxidized affect the stability of the for example underground workings or pit walls. Weathered rocks have a low core recovery when data is drilled [Abzalov, 2016]. The strength of a rock is highly influenced by weathering thus its importance defining its level of rock weathering. Oxidized rocks affect the permeability of the host rocks.
It is important that the lithology is measured well, however this has it's own parameters such as the oxidized state of the rocks present and the grain sizes. Oxidized rocks are more prone to weathering, its distribution will influence the stability. Weathering and oxidation is affected by exogenic forces. It has influence on the slope and pillar stability and it can provide slip surfaces for failures[Dominy et al., 2002]. It is influenced by the distance, the rate of thickness, the grade diminution and the irregularity that goes towards the limits[Dominy et al., 2002]. The variability of the bulk density depends on the degree of weathering and oxidation. The reported reserves and resources will be biased when the bulk density is wrongly assumed. This can even be a small error but it has great influence on the economic variability of the project [Dominy et al., 2002] since the tonnage that is computed, depends on the ore bulk density.

- Faults, dykes and pegmatite veins representing the geotechnical hazards and also internal waste within the pre-body
The presence of geotechnical hazards such as faults, dykes and pegmatite veins have influence on the estimation. Faulting can make a significant impact on the definition of the ore body for mineable minerals [Dominy et al., 2002] shown in Figure 7.

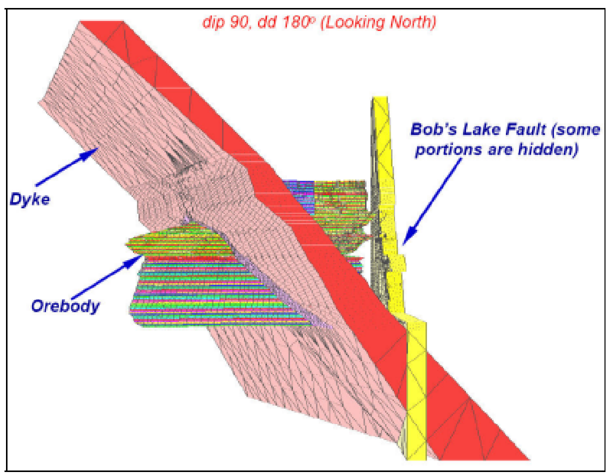

Figure 7: The effect of a fault and a dyke present in the orebody domain[Hosseini and Lewis, 2009].

Information on these geotechnical hazards aids in making the decision to determine how the geology will affect the modeling of the mineralized zone. To knowledge of folding being present gives information about having a duplication of the mineralized zone or an incomplete intersection. To avoid grade dilution and marginal smearing throughout the block modeling process, reliable information is required. The characteristics of geotechnical hazards are the hardness, rock density and moisture [Abzalov, 2016].

- Potholes
Another parameter is the presence of potholes, shown in Figure 8. The material in potholes and their surroundings is fractured and fragmented, which causes issues for mining activities. These cause major safety risks and ground instabilities, particularly in the hanging wall. It is possible to not find the existence of a pothole by logged data from the well providing the data. It can be difficult to detect the presence of a pothole. Potholes have a shape such as a ellipse. It is a disturbance for mining and it

shows alterations in the mineralogy. When potholes are present it is affecting the distribution of the grade in the gravels [Robb, 2005].

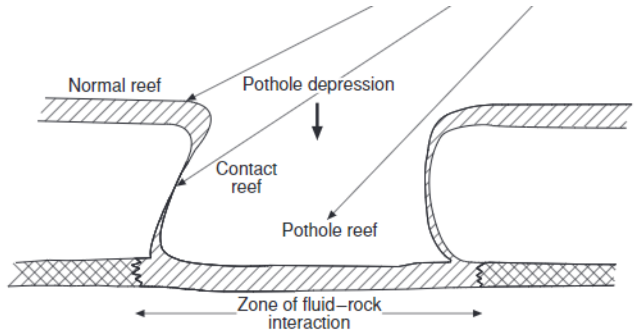

Figure 8: Schematic illustration of a pothole[Robb, 2005].

- Grade and its physical limits
  A great source for error in resource estimations is the grade and its physical limits. It is one that is very difficult to improve [Dominy et al., 2002]. It is depending on the volume-variance effect in the grade interpolations. This is based on the appropriate block size. The latter is used to predict the grade and tonnage relationship [Glacken and Snowden, 2001]. In Figure 9 is shown what an example is of a block that is chosen. The material that is extracted during the mining process from outside the ore body is referred to as external dilution. There is internal dilution within the ore body. This waste cannot be prevented by either circumstances through selective mining, hence it is always mined alongside the ore [GmbH, 2020].

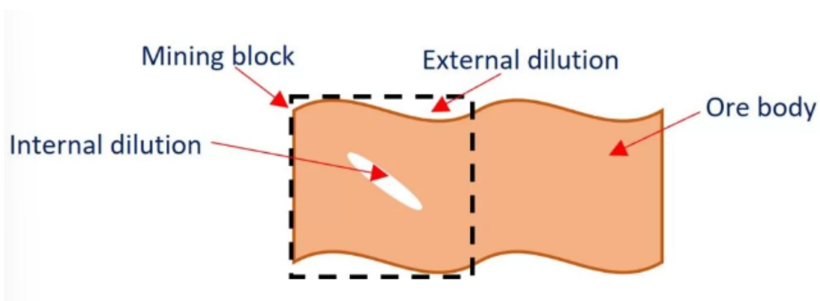

Figure 9: Example of a block size depending on internal and external dilution[GmbH, 2020].

When extraction of the orebody is done, planned dilution and unplanned dilution can occur. In stope limits, the rock having a lower mineralization concentration than the cut-off grade is referred to as planned dilution. Unplanned dilution is the term for rock that comes from outside the intended stope limits, such as through poor rock control, and has a lower mineralization concentration than the cut-off grade [GmbH, 2020].

- Continuity and variability in grade and geology
  Grade and geological continuity and variability are scale sensitive. There is a difference between the local and global continuity of a certain data set. These parameters are related to the spacing and the density of the drill holes. When the uncertainty is big, the estimation can be far off[Dominy et al., 2002]. It will help to decide on the best spacing.

3.1.3 What and how is the current practice to reduce uncertainties in the industry today

MRM requires enough geological information to define the geological model. It must be consistent, display the lithological and mineralogical domains and depict how the mineralization was distributed during the sampling. Only then, the resource estimation can be generated. The latter entails defining the geological or mineralization limitations, analyzing the sample data statistically and/or geostatistically, and using an appropriate grade interpolation approach. Each approach must be founded on a properly specified geological model, and each requires considerable validation by the practitioners, among other commonalities [Glacken and Snowden, 2001]. There are different techniques in the industry today to reduce uncertainty in MRM. All techniques respect the statistical and geostatistical properties of the constraining data, although they vary in the specifics and the approach[Abzalov, 2016] and they all rely on the stationarity assumption, which is not a testable hypothesis but rather the choice to collect data from a certain area or domain. The dimensions are very important for all methods, a wrong size such as too small results in over-smoothing of the sample data [Glacken and Snowden, 2001] and this results in conditional bias. The latter can show in a underestimation of the high grade blocks. The following techniques are mentioned in (Glacken and Snowden, 2001):

- Geostatistical technique
  Geostatistical techniques use semi-variogram measurements of the spatial relationship between samples to calculate weights of the estimation of the unknown point or block. The most common techniques used are variations of ordinary kriging, belonging to the linear kriging techniques. It determines how the grades are distributed across each point or block, the spatial variations. The latter gives a sense of regional uncertainty. The workflow of this approach is shown in Figure 10.
  It will aid in selecting an appropriate grade interpolation method. It should make clear the geographic continuity of the structure of the data and confirm the geostatistically predicted geological trends. Additionally, this study will look to confirm if soft or hard domain boundaries were chosen. The selection of the block size to be utilized in block modeling approaches will aid in the development and modeling of variograms [Glacken and Snowden, 2001].

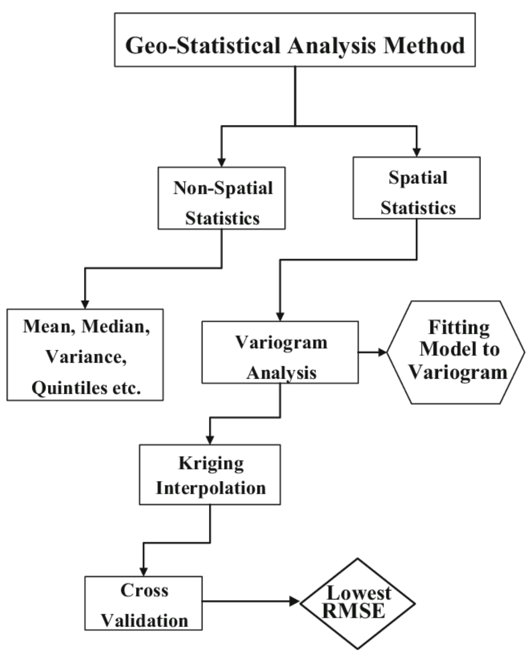

Figure 10: Flow diagram of the geostatistical analysis method[Javed et al., 2017].

- Conditional simulation
  Conditional simulation is based on kriging and Monte Carlo sampling[Glacken and Snowden, 2001]. The name of this approach is based on the simulated models that are adjusted, conditioned, to the actual data samples[Abzalov, 2016]. It can replicate the degree of variability in the samples. The greater the degree of the variations between realisations, the more data variability there is and the fewer samples are available to constrain the models[Abzalov, 2016]. It is particularly useful when data are densely spaced and could be viewed as supplementing rather than replacing other estimating techniques. The big difference that simulations include in their model is that risks are considered. Conditional simulation can establish the ideal block sizes and the amount of resource smoothing that exists. It utilizes stochastic random sampling techniques in addition to kriging. This provides a complete assessment of uncertainty. This method addresses many of the drawbacks that are present in the kriging methods, however it is very time consuming and not simple. The models in this method are conditioned to the actual data samples. It is based on the Monte Carlo stochastic algorithm. The uncertainty can be attained by statistically analysing the differences between the simulated realisations [Abzalov, 2016]. in Figure 11 it is visible the random sampling what the prediction is based on.

  Machine learning techniques are mostly applied when the training data is generously present, however in the geological field this is not always the case and must it be done by indirect measurements and in shorter supply.

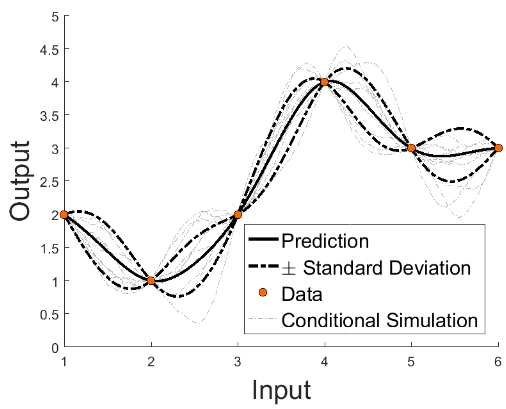

Figure 11: Conditional simulations shown based on prediction and data[Freier, 2016].

The method can be divided into three groups based on the characteristics of the variable that needs to be simulated. Simulations based on continuous variables, categorical variables and objects. The quantification of the uncertainty in the estimated grade is the conditional simulation's most typical application. This is used for the quantification of the production risks for mining. It is also used to classify the mineral resources into the measured, indicated and inferred category. The method is based on the creation of various equiprobable models of the grade under study's spatial distribution. To quantify the estimated grade uncertainty, the findings are contrasted and their differences statistically analyzed. These estimations of the grade uncertainty can be done by different techniques such as turning bands which is shown in (Battalgazy,2019;Paravarzar,2015), sequential Gaussian simulation (SGSim) which is shown in (Albuquerque,2014;Paravarzar,2015), or sequential indicator simulation shown in (Sojdehee,2015). All three need a transformation of the data. By comparing the estimations, independent from each other, performed by different techniques can show if if the results support each other and contribute to the validity[Abzalov, 2016].

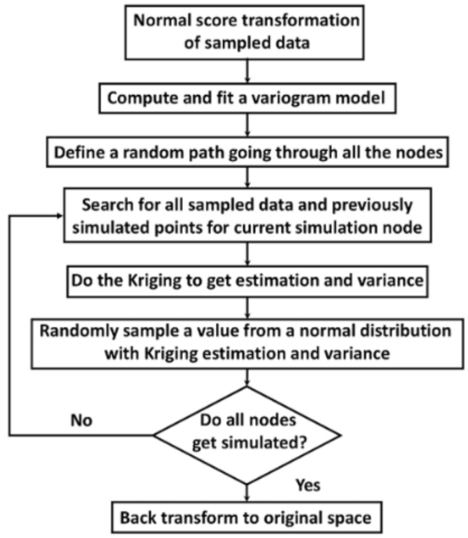

Figure 12: How to get to a stochastic realization[Bai and Tahmasebi, 2022].

- Stochastic framework
  In a stochastic framework shown in Figure 12, the SGSims are designed to address problems with underestimation of high values and overestimation of small values caused by smoothing. It is also used to get through limitations like the size of the problems or irregular spacing between the data and the target areas. However when there is too much data, cokriging becomes computationally expensive. This occurs when there are a lot of pre-existing data or many places that are being targeted for simulation[Par, 2015].
  Due to its ease of use and clearness in a wide range of situations, this strategy is frequently utilized in operation. In order for SGSims to function, a grid of randomly assigned values from a typical normal distribution, where the mean is equal to zero and the variance equal to one, must first be created. Then, the grid is subjected to the covariance model, based on the semivariogram in the Simple Kriging layer which is a must as input for the simulation. This makes sure that the grid values follow the spatial layout of the input data set. One unconditional realization is represented by the resulting grid, and many more can be created by repeatedly utilizing a different grid of values with a normally distributed distribution[Kasa, 2018]. The accuracy of this method is not always guaranteed, and applying it to scenarios involving multiple variables may be difficult and call for simplicity [Par, 2015].
  It is used to asses the uncertainty in the unknown values of coregionalized variables, such as the grades of elements of interest, petrophysical properties of the subsurface, or geometallurgical properties. It requires a stochastic model that describes the spatial distribution of the coregionalized variables and an algorithm that constructs realizations of the prescribed model.

- Geological model
  When the geological model is based on the prior knowledge and the available data, it is needed to create a domain where the characteristics are representing the materialization [Glacken and Snowden, 2001]. For the resource model, the domain is based on the reflection of the geology. When the latter is not thorough enough it should be based on another domain boundary. This can be a grade boundary that is defined by the cut-off grade, which will consider a relation to the economics of the deposit. So the domain can be chosen based on statistical and geostatistical and a cut-off grade.
  The domains can be defined by soft and hard boundaries shown in Figure 13. A soft boundary domain can have variations, this is not permitted in the hard boundary domain. In the same region it is

possible to have different types of domains. It is needed to define the resource that is present in this region. To decide what kind of boundaries for the domain are used, the geological characteristics of the materialization needs to be considered [Abzalov, 2016].

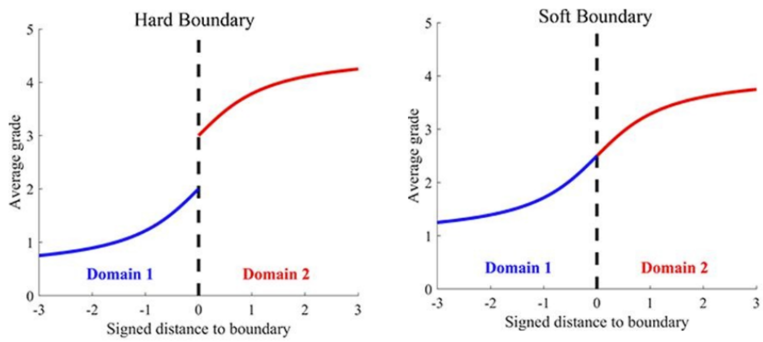

Figure 13: Difference between a hard and soft geological boundary domain[Maleki and Emery, 2020].

- Statistical analysis
  Statistical analysis helps with deciding on the nature of the boundaries. The analysis looks at how the grades are changing at the boundaries. The numerical properties of the mineralization should be specified in order to aid in the selection of a grade interpolation method and highlight any unique data treatments. It will also identify any patterns or correlations between the minerals or relevant variables. All these statistical analyses take place in the defined domain. The samples that are used in the data analysis need to have the same volume. This can be by looking at equal lengths of drilled data[Glacken and Snowden, 2001]. Clustering of the data is required before the data analysis can be carried out. Both the statistical analysis and the bias of the variography are affected by this. By using declustering techniques, it may be ensured to reflect an equivalent volume. Data clustering may produce biased findings. Compositing is further justified by the possibility of extremely variable small-scale experiments, which can be reduced. By composing to an acceptable length, the variability will be reduced, resulting in a more reliable geostatistical analysis that includes variography. It is crucial to the resource model's overall quality [Geologist, 2020]. It also can indicate that more domains are needed to be separated or to use indicator techniques for kriging and variography. The bulk density should subsequently be used in the actual estimation procedure when it demonstrates a positive correlation between the bulk density, the ore, and the grade of the minerals of interest [Glacken and Snowden, 2001].

3.2 Bayesian Evidential Learning

3.2.1 What is Bayesian Evidential Learning

Bayesian :
Based on the fundamental principle of information acquired previously. An inductive-deductive approach to science reasoning[Caers, 2018].
Evidential :
Both field observations and modeling data are used as evidence[Caers, 2018].
Learning :
Based on Monte Carlo machine learning[Caers, 2018].

BEL enables to model a posterior distribution in the prior model space, using parameters. It provides an indication as to how future data might appear, given the data and model. By separating the quantity of interest's range into several distinct bins of equal length, it will give a prior probability that represents what is initially assumed to be true before new evidence is presented, and posterior probability accounts for this new knowledge. BEL uses statistical relations between the predictor variables and data from the sampled realizations from the prior distribution. The establishment of statistical correlations between these variables is a prerequisite for the BEL's interpretation component. These connections result from geophysical problem models and simulations that take into account the prior knowledge that is currently accessible. Projects need uncertainty to be considered because it gives decision-makers reliable information. It relies on linearized approximation-based data error propagation. Numerous inversion or Monte Carlo techniques are used in stochastic methods to compute many sets[Thibaut et al., 2021]. Unlike unified or sequential formulations of inversion, BEL is capable of demonstrating the statistical relationship between the variables of relevance and measured data[Bogrash, 2020]. The BEL framework can involve falsification, global sensitivity analysis and direct forecasting [Tadjer and Bratvold, 2021]. The BEL approach is subdivided in six different stages to get to the final goal of decision making. These six stages will be explained in the next section with case studies that show how to get to the final result step by step. This approach will follow a certain logic such that in the end there is a certain uncertainty quantification so that the qualification makes scientific sense and that can be used for decision making.

1. Formulating the decision question and statement of prediction variables

   This first step is very important for the process, to formulate a decision question and state the important prediction variables. Objectives need to be formed, risk and return objectives are usually the case and these are the prediction variables. The objectives need to be evaluated against the alternatives done with modeling. What is different in the BEL approach in comparison to other approaches is that it will only look at the field data, observed data, during regression and not in the beginning of creating the model. First the a-priori information will be considered[Scheidt et al., 2018]. Building an appropriate prior distribution for the model's parameters is the base of BEL[Thibaut et al., 2021].

2. Statement of model complexity and prior uncertainty

   In this step the Earth needs to be conceptualized and the global knowledge will be taken up in the process. This can be research about the area of the problem before or the geological history that is known. So the combination of what the aim is to predict and the knowledge already present will provide the model complexity and the prior uncertainty. This stage is using an initial hypothesis, it does not need to be accurate. No observations or dimension reductions are included yet. This needs to be done for all of the fields jointly, which can lead to many models.
   An example of present information about the area is geophysical information which can have information about the variations of the density in a particular area.
   You have structural, rocks and fluid components of parameters. The distributions are not to be used with standard deviations, the wide range is used and later on these will be reduced. The data variable is obtained by talking a model and applying the data for model. Using this data, the prediction variable is obtained.
   Prior sampling is done to get the possible range of the model parameterization and the probability distribution of the geological parameters [Tadjer and Bratvold, 2021].

3. Monte Carlo and falsification of prior uncertainty using data

   Each variable obtained in the previous step will have one realization or sample to get one model leading to many realizations. This prior information will be used for random sampling and model realizations will be produced. This is mathematically shown as:

   $$m^{(1)}, m^{(2)}, ..., m^{(L)} \tag{1}$$

The prediction variable will be shown as h and the data observable variable as d. So d is first based on prior knowledge and global information, which does not need to be inverted with data, and further in the process d_{obs} is used, which is the actual field, measured data. A model $m^{(l)}$is based on prior information and not the field data. Giving $d^{(l)}$ and $h^{(l)}$ by forward modeling. The used $d^{(l)}$ will most likely not match d_{obs}. So the model is based on information that is independent of the process and it will create or produce values that can be used for the first prediction to then get a statistical relationship between these data values $d^{(l)}$ and prediction $h^{(l)}$. When this relationship is created, another prediction can be made by using d_{obs} [Scheidt et al., 2018]. Learning the link between the data variables and the decision factors is the goal. So a big difference from other methods is that the model is not based on the evidence and then a prediction is made but the measured data is directly influencing the decisions.
Applying Monte Carlo means making a table, shown as the box in the middle in Figure 14 and then to apply many realizations of the model, obtained in the prior step to continue with forward simulations on the model resulting in data variables and the future that is simulated. The actual observation come in place only after this step.

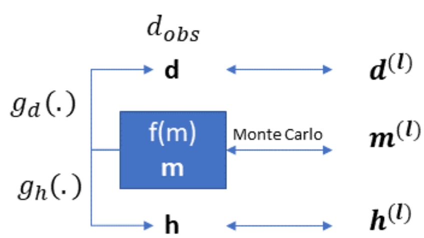

Figure 14: Visualization of what the application Monte Carlo is. g_d is the data forward model, g_h is the prediction forward model, d_{obs} is the actual measurements, **d** the data variables and **h** the prediction variables. $l = 1, ..., L$ is for the models obtained as shown in Formula 1[Caers, 2018].

The subsurface model parameters' prior distribution results in a statement of uncertainty regarding prior prediction or decision hypothesis. Shown in 2 what the *Bayesian* part is of this approach. These are based on limited samples and the estimated decision hypothesis, $\hat{P}(h)$ comes from the same models as the estimation of the data variables $\hat{P}(d)$. The estimation for the statistical relationship can be done by a direct estimate $\hat{P}(h|d_{obs})$ or by estimating the likelihood of the evidence to be occurring given a certain hypothesis, $\hat{P}(d_{obs}|h)$. This is all coming from Bayes' theorem A.1[Scheidt et al., 2018]:

$$\hat{P}(h|d_{obs}) = \frac{\hat{P}(d_{obs}|h)}{\hat{P}(d_{obs})}\hat{P}(h) \tag{2}$$

So the h is based on the model parameters, the d is data variables that are observed. Now there is the ability to calculate the needed prediction for each posterior distribution model and do a risk analysis relying on the quality of these predictions[Hermans et al., 2018]. The statistical model called m describes uncertain subsurface properties. The objective of statistical modeling is to describe potential outcomes of a random event or quantity and determine the likelihood that each outcome will occur. However, it becomes significant only after obtaining certain data[Scheidt et al., 2018].

To visualize all these different models can be difficult and unclear. Dimension reduction can be a solution for this. Because the number of samples required increases exponentially and it will consume way more time, there is a need for a decreased dimensional space. Multicollinearity is another potential outcome of high dimensionality in data. The variables are correlated at this point, although they may originate from an unobserved, lower-dimensional source[Scheidt et al., 2018]. Many vectors such as an object, a model, a data variable, a prediction variable, anything can be taken and written as a linear combination of non-linear functions. These combinations can be scalars or fixed shapes. Often Principal Component Analysis (PCA) is performed. Other techniques are that can be used are kernel density estimation[Hermans et al., 2019], multidimensional scaling or functional data analysis[Scheidt et al., 2018]. What happens when applying PCA, is that the non-linear functions are rewritten into a model giving linear combinations of these vectors.

Principal Component Analysis (PCA) gives a series of the original high-dimensional variable's best linear approximations. It can be applied to matrices, a set of vectors, and geographic maps as well and it can combine data variables. For example when there are two different well locations, it is possible to combine them and get one variable from it.
The size of dimension of the data an predictor variables is important for how much their need to be reduced. By examining the generated eigenvalues allows the determination of the minimum number of dimensions required to capture the majority of the variability. However these variables are reduced independently from each other. So the d and the h become d^f and h^f, this means that the variables are in their reduced dimension space[Hermans et al., 2019]. Because PCA is bijective, it is possible to recover the original high-dimensional variable by reversing the projection. As a result, after any predictions are made in the low-dimensional space, the original space may be easily and uniquely recreated[Scheidt et al., 2018]. This allows to apply other analyses on the variables in the reduced dimension space and they can still be rebuild to the original space.
Additionally, PCA can be used to merge data from various sources, find overlaps, and produce a reduced dimensional projection of the combined data. Using mixed PCA, data from numerous well locations are integrated into a single variable. The procedure consists of three steps shown in 15. A standard PCA is initially performed on each of the data sources to determine which value is the most distinctive. Second, each data source is normalized using the first singular value in order to take into account any scale variations among the data sources. The third step is the combination of the normalized data inputs, followed by a typical PCA on the resulting matrix[Scheidt et al., 2018]. This process is shown in Figure 15.

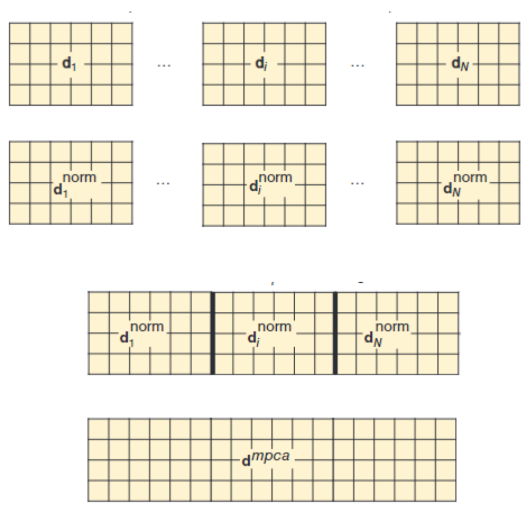

Figure 15: The steps to perform mixed PCA[Scheidt et al., 2018].

The dimension reduction is based on the formula given below.

$$x = \Sigma_{k=1}^{K} \alpha_k \phi_k \tag{3}$$

A model is written as a linear combination of the basis vectors. The α in this formula is the variance, α_1 will be a larger variance than α_2 and so on. These α's can be plotted against each other which results in a 2D space which makes it easier to visualize the data.
The model can be depending on reaction rates, concentrations of the mineral, the surface area, geochemical and geological uncertainties are obtained from this. When the case is that the prior distribution is unreliable it will need a falsification of the prior. This is the case when the observed data is outside the data variable samples. Then the prior hypothesis will be falsified, it will require extrapolation when the observed data is outside the sample range. Because there might not be any outliers in the training sample, the statistical models for classifying outliers could be difficult. (Schiedt et al, 2018) describes a possible outlier identification algorithm which is a support vector machine (SVM).
The fact that the data variable informs on the prediction variable should result in any reduction in uncertainty between the prior and posterior probability density functions(pdfs). In the BELs case it is based on the samples of h and d, obtained from sampling the prior distribution. To find out whether it is informative or uninformative, concerning the statistical significance, the p-value is looked at. When the p-value is lower than 0.05 it is indicated that informative data variables are used [Scheidt et al., 2018]. However the uncertainty significance is not necessarily correlated to the statistical significance. Another method can be testing the uncertainty significance using a hypothesis test. It tests whether d^*_{obs} is informative, shown when there is a big difference between the prior and the posterior distribution. The test takes the angle for it to be $false$, so the test will look if the data is **not** informative. Meaning the wanted outcome is that there is no difference between the prior and posterior distribution. The null hypothesis is as follows:

$$H_0 : f(h^*) = f(h^*|d^*_{obs}) \tag{4}$$

This method is based on limited number of samples and not an exact distribution, which can be impractical [Scheidt et al., 2018]. The distance between the two distribution needs to be measured, which can be a difficult task since there are many variables concerning the problem. However, to use the first component can be a solution, h^*_1. This is not a complete test, that would include all the variables. The goal is to get a low achieved significance level (ASL), this means that with this null hypothesis that d^*_{obs} is informative of h^*. When the ASL is high, it means that null hypothesis is true. In this case this would mean that any predicted posterior uncertainties might not be accurate, therefore any judgments based on them should be carefully considered [Scheidt et al., 2018]. After a falsification on the prior uncertainty the model complexity and the model uncertainty needs to be increased.

4. Sensitivity analysis on both data and prediction variables

   When the prior model uncertainty has passed the falsification, the sensitivity analysis is performed. The method used is global sensitivity analysis (GSA). Meaning that the Monte Carlo results will be used again, the sensitivity analysis will be ran for these. These analyses are performed to find out what variables have the most influence on the models predictions. It will lead to a simplification of the model because the valuable information and variables will be prioritized in the reduced dimensions. Data giving most information about the prediction is what is needed to have an efficient working model. An example of such a GSA is shown in 16.

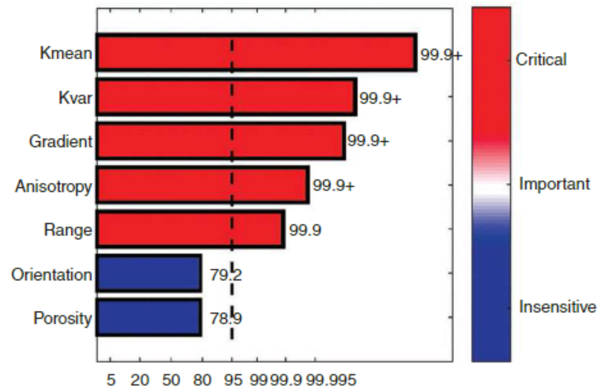

Figure 16: Example and indication of sensitivity of the parameters[Scheidt et al., 2018].

It is good for the UQ since it evaluates sensitivity measures by assessing the impact of inputs throughout the full parameter space. The response of the parameters on one collective something is compared.

5. Design of uncertainty reduction on prediction variables based on data

   At this stage, uncertainty reduction is done since the models have gone through the falsification process and the sensitivity analysis is performed. In this stage, observed data variables are used to reduce the uncertainty in the variables. Instead of model inversions, direct forecasting, shown in Figure 17, will be done on the multiple realizations of the data and prediction variables already obtained through Monte Carlo. Machine learning will learn the data from the prediction variable. This will give a regression relationship and then the observed data can be used to directly reduce the uncertainty on the prediction variables. The answer is obtained directly based on what the predictions are.

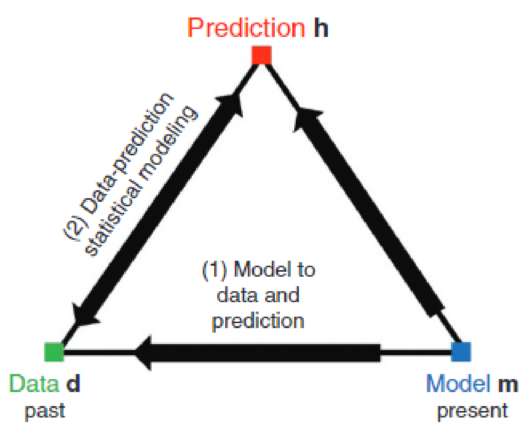

Figure 17: Shows what direct forecasting is based on[Caers, 2018].

In this step the aim is to estimate the posterior probability distribution, by regression analysis, of the prediction variable h and d_{obs} used. In this step Bayes' rule will come up for expression the posterior as a product between the prior distribution and the likelihood function[Scheidt et al., 2018]. So the variables have all kind of transformations and will be noted as d^* and h^*. The posterior distribution formula will be as follows:

$$f(h^*|d^*_{obs}) = const \times f(d^*_{obs}|h^*)f(h^*) \tag{5}$$

The constant is there to normalize both functions. Choosing the appropriate regression techniques will influence the method of estimating these distributions. It depends on the underlying variables' characteristics as well as how they are related. The distribution of potential unobserved values depending on the observed values is known as the posterior predictive distribution. This is based on h, given that the d is equal to the observed data d_{obs}. Getting this relationship, regression analysis is performed. There is parametric regression and non-parametric regression[Scheidt et al., 2018]. In this thesis, it will be focused on the parametric regression technique. This is applicable when the the process generates a linear relationship [Hermans et al., 2019].
Canonical Correlation Analysis (CCA) is a method to explore and quantify the relationships between the two data matrices[Scheidt et al., 2018]. It relies on projections in the lower dimensions. The goal is to find the maximum correlation. It depends on the eigenvalues of the matrices. It is applied to make the models less complex and to reduce the dimensions. By applying CCA it will find the optimal linear combination of the correlations obtained by PCA. When the data variable contains of three Principal Components (PCs) and the prediction variable of two PCs it leads to a maximum of two Canonical Components (CCs) can be obtained. Shown in 18 is what the impact on CCA can have to make the visualization better.

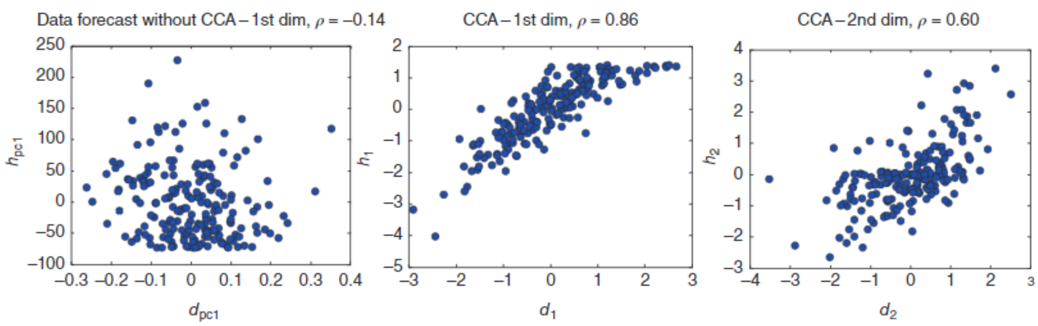

Figure 18: The first graph shows the correlation between two PCs without any CCA, the second shows the correlation between the first dimension PCs after CCA is applied and the third the second dimension of PCs[Scheidt et al., 2018].

The in depth formulas of regression are shown in A.2.1. The main advantage of this method is that posterior sampling, $f(h^*|d^*_{obs})$ is straightforward. Each sample taken from the posterior can be projected into the original space, $f(h|d_{obs})$ by undoing any transformations and dimension reductions.

6. Posterior falsification and sensitivity, decision making

   Once knowing the uncertainties, it is time to make a decision. Evaluating all different numbers on return and benefit and risk scores. Here is where the reduced predicted target with a reduced uncertainty will be taken into account for further considerations.

3.2.2 Applications of Bayesian Evidential Learning in other industries

In this chapter multiple case studies will be shown and studied to learn how BEL has been used already in certain cases to obtain a reduction of uncertainty on the prediction parameter and decision making processes. Each case will follow the six steps as explained before.

Case 1: Bayesian Evidential Learning: a field validation using push-pull tests [Hermans et al., 2019]
This case shows that BEL can be used to predict the range of uncertainty regarding a given prediction in the field of Earth sciences. The following information is from (Hermans et al.,2019). The site is checked to

not have any abrupt, physical not plausible reactions during the experiments considering the temperature changes.

1. Formulating the decision question and statement of prediction variables
   In this case the objective is to predict the posterior temperature distribution during a cyclic heat tracer push-pull test in an alluvial aquifer [Hermans et al., 2019]. The change in temperature in the aquifer during the different phases will give information to predict. A narrow range of uncertainty is preferred to get from using BEL. The over time evolution of the temperature in the well is the prediction variable.

2. Statement of model complexity and prior uncertainty
   The prior model is based on the knowledge of the site, which is a lot in this case, the three layers are defined. Its orientation is along the flow direction. For the model, 500 independent realizations are used. The uncertainty variable parameters and its range for this case are shown in Table 1.

   Table 1: Parameters of the uncertainty variables of case 1[Hermans et al., 2019].

Parameter	**Range of uncertainty**
Mean of $\log_{10} K(m/s)$	U[-4 to -1]
Variance of $\log_{10} K(m/s)$	U[0.05 to 2]
Range (m)	U[1 to 10]
Anisotropy ratio	U[0.1 to 0.5]
Orientation	U[0 to π]
Porosity	U[0.05 to 0.30]
Gradient (%)	U[0.083 to 0.167]

   These prior uncertainty parameters are all randomly and independently sampled. The hydraulic conductivity affects the aquifer's flow dynamics in general and the advection velocity in particular. A high variance shows that the heterogeneity is high. The gradient will affect the advective fluxes, all based on prior knowledge.

3. Monte Carlo and falsification of prior uncertainty using data
   The past two steps are field data independent. The prior model that was used with the data must be equally consistent in order for the results to be accepted. If it is inconsistent, it is falsified, which then influences the entire process. For both data and prediction, the prior model's consistency is confirmed. The prediction would not have been feasible yet using other techniques. Both in this instance involve the injection well's temperature over time. No evidence or prediction can be used to invalidate the prior model. With a rapid or gradual fall in temperature during the pumping and storage phases, the preceding model can produce a wide range of potential outputs. Field data and forecasts show similar temporal characteristics and fall within the same range as that seen in past samples' responses as shown in Figure 19. This is done in a reduced dimension space, 500 temperature curves and field curves are considered. The reduced dimension space is obtained using PCA. The visualization is to know whether the field curves lay in the span of the temperature curves.

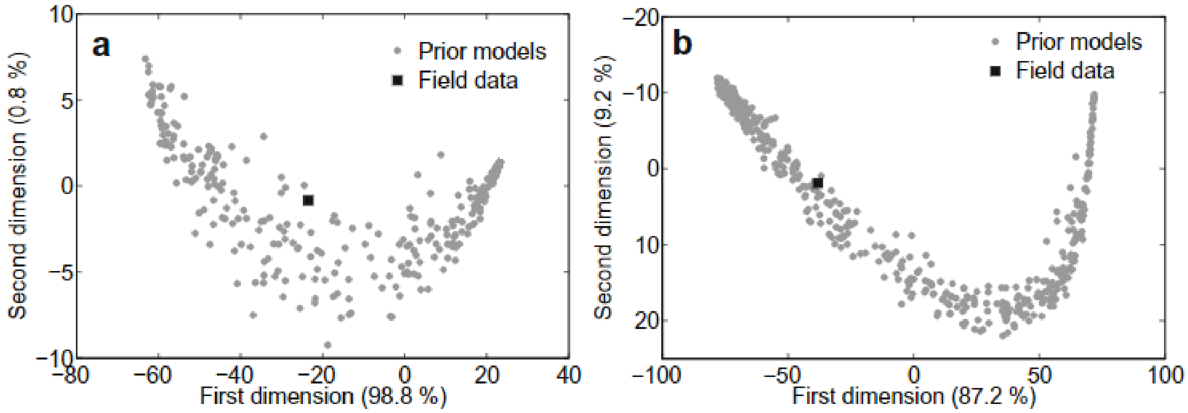

Figure 19: The prior model falsifications in the reduced dimension space that shows that the prior model is not falsified[Hermans et al., 2019].

4. Sensitivity analysis on both data and prediction variables
   The sensitivity analysis is based on the distance between the responses from pairs of models within the 500 models. The ranges of the anisotropy ratio and of the variogram, relating to the spatial distribution of the hydraulic conductivity, and the variance will control the degree of the heterogeneity in the well area and affect the temperature curves. Expected that the mean and the variance of the hydraulic conductivity will control the results of the aquifer considering its sensitivity as is shown in Figure 20.

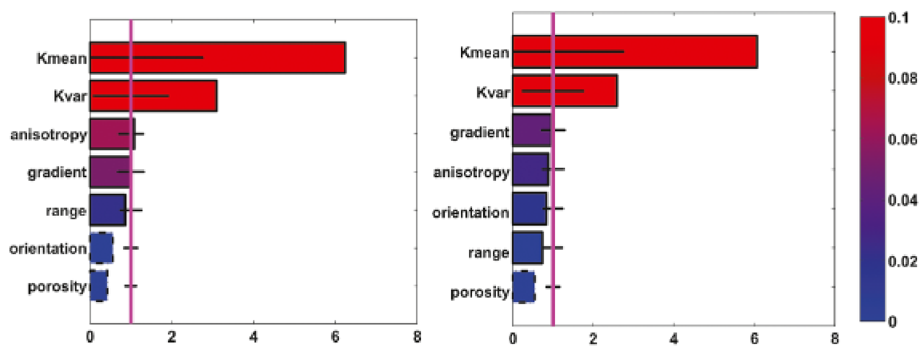

Figure 20: The most sensitive parameters are shown for both experiments in this case. The mean and the variance of the hydraulic conductivity is in both experiments of most influence[Hermans et al., 2019].

5. Design of uncertainty reduction on prediction variables based on data
   CCA is applied to the reduced data and the prediction variables to get a linear relationship that is independent. The first dimension space is not linear and nor has the second a unique linear relationship. Kernel Density Estimation (KDE) will be used, it is based on how previous samples in the CCA space were distributed. CCA is helpful even to obtain the most linear relationship, given the observed data. The posterior distribution of the prediction in the CCA space is determined. The computed posterior distribution of the prediction is shown in Figure 21, which is in a reduced dimension space and in a CCA space. This distribution can be back transformed to its original space.

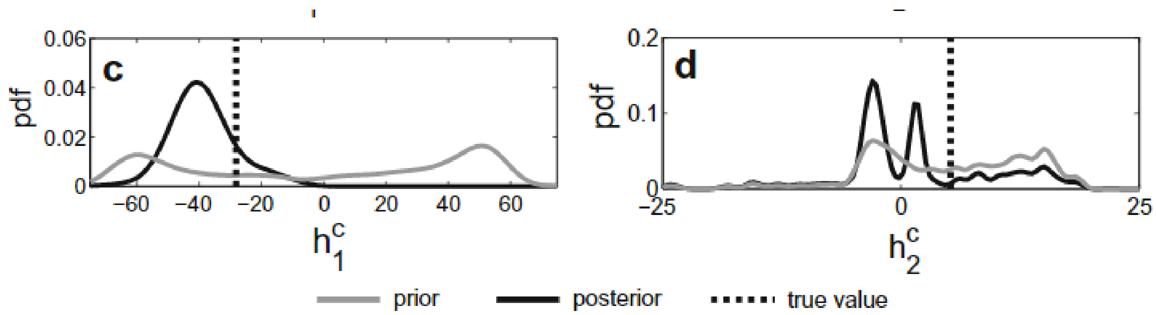

Figure 21: Posterior distribution of the prediction in CCA space and reduced dimension space through PCA [Hermans et al., 2019].

The peaks shown in the second graph shown in Figure 21 are because of the density of the points in the CCA space, the samples where laying very close to the place where the data variables of the second dimension in the CCA space are almost equal to zero. Figure 21 shows that by using BEL it is succeeded to forecast a posterior distribution of the prediction not getting falsified.

6. Posterior falsification and sensitivity, decision making
The posterior distribution of the prediction has been achieved and is back transformed to its originals space, shown in Figure 22. Showing that BEL most definitely has been successfully applied to reduce the uncertainties. The difference between the range of blue lines and grey lines is the uncertainty reduction.

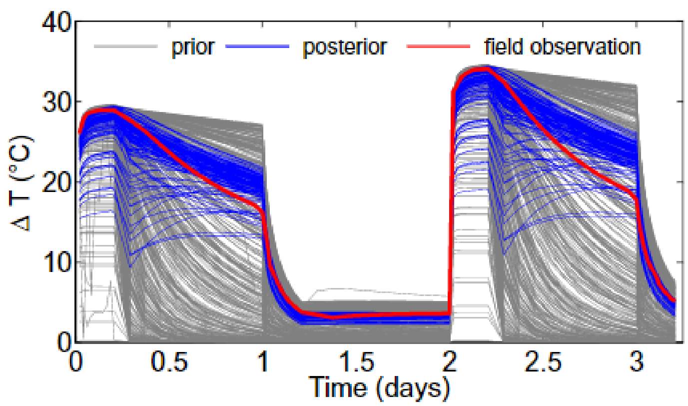

Figure 22: Posterior distribution of the prediction including the field observation to forecast the response of the aquifer for other solicitations [Hermans et al., 2019].

It can be more precise if the unexpected behavior of the posterior models that is physical not plausible are filtered out. 300 models are removed from the prior realizations. The distribution of the model parameters in the removed samples and in the reduced prior model are shown in Figure 23. What is visible is that the mean and the variance of the hydraulic conductivity are different. This is showing that the sensitivity analysis is done correct since there was mentioned that the main factors influencing the models response are the mean and the variance of the hydraulic conductivity.

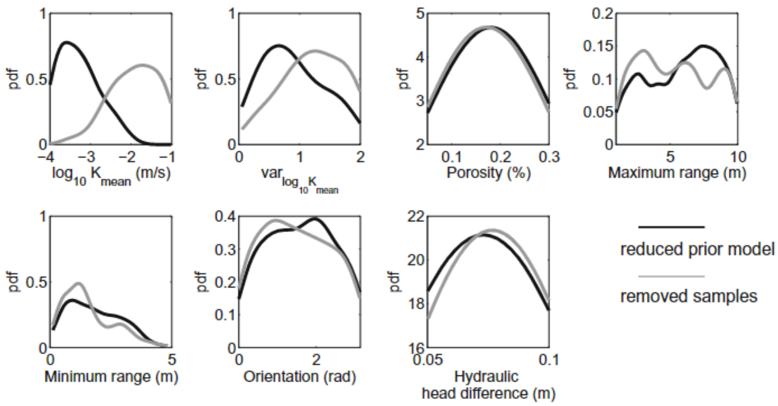

Figure 23: The most influential parameters on the models response are shown [Hermans et al., 2019].

And the final posterior distribution of the prediction is shown in Figure 24 after having eliminated the models that did not have much influence on the posterior distribution of the prediction.

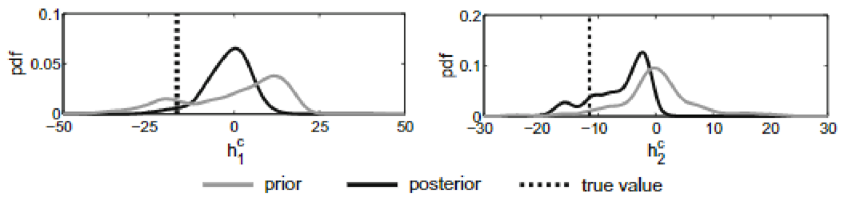

Figure 24: After eliminating 300 prior models this is the posterior distribution of the prediction in CCA space and reduced dimension space through PCA obtained [Hermans et al., 2019].

By removing these samples it improves the relationship that is derived in the CCA space between the data and prediction variables. This is shown in Figure 25, however the effect on uncertainty reduction is not immense.

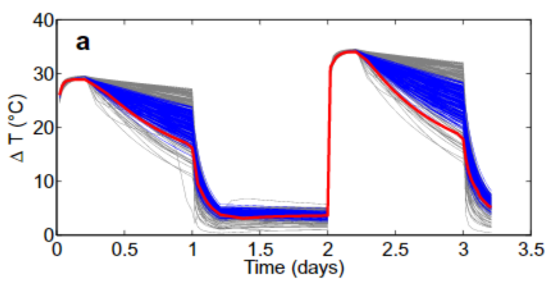

Figure 25: After eliminating 300 prior models this is the posterior distribution of the prediction in CCA space and reduced dimension space through PCA obtained [Hermans et al., 2019].

The prediction is still in the extremes of the distribution. It is challenging to forecast the true prediction since it is situated at the edge of the prior distribution. Additionally, as a result of this, it is also towards the edge of the posterior distribution.
Few samples op the prior and the real observation lay outside the main trend. Samples that are more densely distributed will get more samples generated which gives a crooked temperature prediction. If

the real prediction was included into the predicted probability density function it would be as shown in Figure 26 in the first dimension. Compared to Fig 25, it will get less high temperatures with the corrected mean of the first dimension prediction. This distribution is the final result using the BEL approach and it is visible that the uncertainty range is reduced. BEL has been applied successfully.

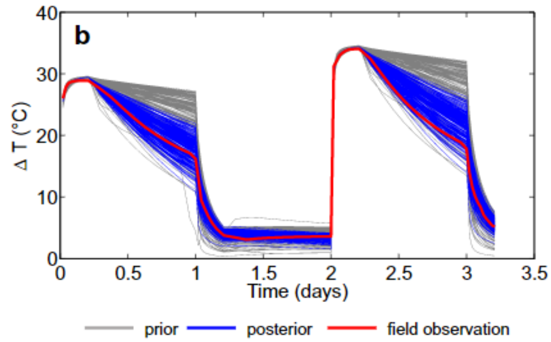

Figure 26: Posterior distribution of the prediction after correcting the mean of the first dimension of the prediction[Hermans et al., 2019].

Case 2: Uncertainty quantification of medium-term heat storage from short-term geophysical experiments using Bayesian Evidential Learning[Hermans et al., 2018] This case will show that by using BEL a full quantification of the prediction posterior distribution will be achieved purely based on observed data without needing inversion models. The aim is to design an ATES system that considers the uncertainties concerning the spatial heterogeneity of hydraulic properties or non-favorable hydrogeological conditions. The following information is from (Hermans et al.,2018).

1. Formulating the decision question and statement of prediction variables
   This case is about how aquifer thermal energy storage (ATES) systems can recover in winter the heat stored in the aquifer during the summer to increase the energy efficiency of the system. This however is often lower than expected.
   BEL will estimate the heat storage capacity of an alluvial aquifer using a heat tracing experiment, it will predict the efficiency of the thermal energy storage in a real aquifer.

2. Statement of model complexity and prior uncertainty
   The prior seeks to identify the potential range of parameter change based on current knowledge. The geological data used for the model's discretization was collected using existing boreholes. Looking at heat tracing will help since it is affected by hydraulic and thermal properties of the aquifer so it will give a lot of information about the parameters present. The three types of uncertainty that are being considered are hydrogeological conditions and properties and the model of the spatial continuity that gives information about the spatial heterogeneity. The parameters that are used in the simulations are shown in Table 2

Table 2: Parameters of the uncertainty variables of case 2[Hermans et al., 2018].

Parameter	**Range of uncertainty**
Mean of $\log_{10}K(m/s)$	U[-4 to -1]
Variance of $\log_{10}K(m/s)$	U[0.05 to 1.5]
Range(m)	U[1 to 10]
Anisotropy ratio	U[0.5 to 10]
Orientation	U[$\frac{-\pi}{4}$ to $\frac{-\pi}{4}$]
Porosity	U[0.05 to 0.40]
Gradient (%)	U[0 to 0.167]

3. Monte Carlo and falsification of prior uncertainty using data
   Simple Monte Carlo simulations are used to create a collection of 500 realizations of the prior model's parameters. The temperature curve is generated and shown in Figure 27, during the pumping phase. The prior distribution of the prediction is shown.

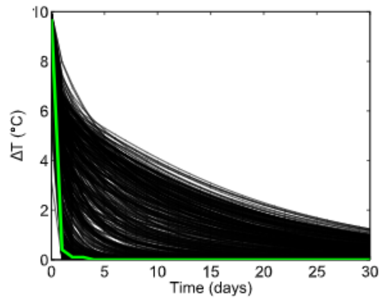

Figure 27: The prior distribution of the prediction during the heat storage. These are temperature curves[Hermans et al., 2018].

One curve is one prediction of one model of the prior. The green curve is not the true data but obtained through a calibration of the hydraulic conductivity distribution.
An experiment is performed with heat tracing to reduce the uncertainty of the prediction on the 500 models of the prior set. The change in the electrical resistance is the data set for the subsurface models. The experiments are done in different time steps, for one-day and for five-days. The data set is reduced in its dimension space. The prior gets falsified. After removing noise and clear outliers the curve from the observed data is expected to be within the distribution of the prior curves. In this case the overall change in resistance in this instance is consistent with the preceding for all electrode designs. This makes sure that the temporal behavior is constant and that the prior can, on the whole, record the amplitude of the change. This is shown in the first graph in Figure 28.
While the temporal behavior is very similar for the individual electrode design, some are at the edge of the prior distribution or even located outside the distribution domain, shown in the third graph in Figure 28. It will not cause a falsification of the prior since the difference between the observed curve and the model curves is not bigger than the difference between the extreme curves in the distribution of the the first and second figure shown in Figure 28.

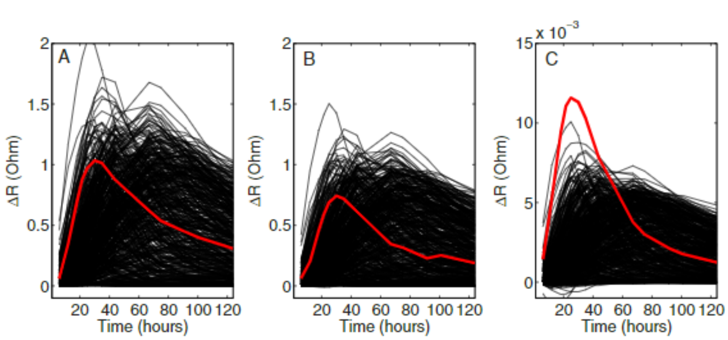

Figure 28: The prior falsification using observed data, shown as the red curve[Hermans et al., 2018]. THe first graph is the average change in resistance and the second individual change. The range showing in these graphs cause that the outliers visible in the third graph will not influence the falsification.

4. Sensitivity analysis on both data and prediction variables

The sensitivity analysis is done in a reduced dimension space to get robust samples. There is a difference in the global sensitivity between the two experiments. This is mainly because the beginning of the experiment provides mostly information about the arrival time of the tracer and the latter is related to the mean of the hydraulic conductivity. After these time steps, following, will the heterogeneity influence most. This can be concluded by seeing that the variance of the hydraulic conductivity is most sensitive in the experiment over five-days. This is shown in Figure 29.

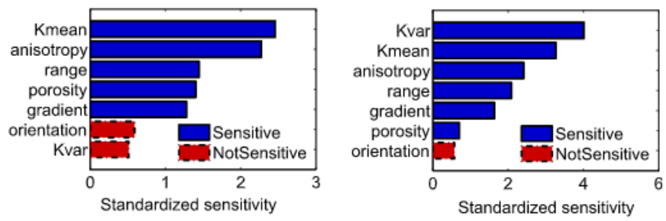

Figure 29: The global sensitivity of the parameters for the one-day experiment and the five-day experiment[Hermans et al., 2018].

Figure 29 also shows that the natural gradient has a smaller sensitivity for the heat storage.

To get a more precise result for the falsification of the prior, dimension reduction is applied using PCA. Both experiments have the observed data laying in the domain of the prior distribution. However it does show in Figure 30 that for both experiments the observed data is located at the edge and a very poor sampled domain.

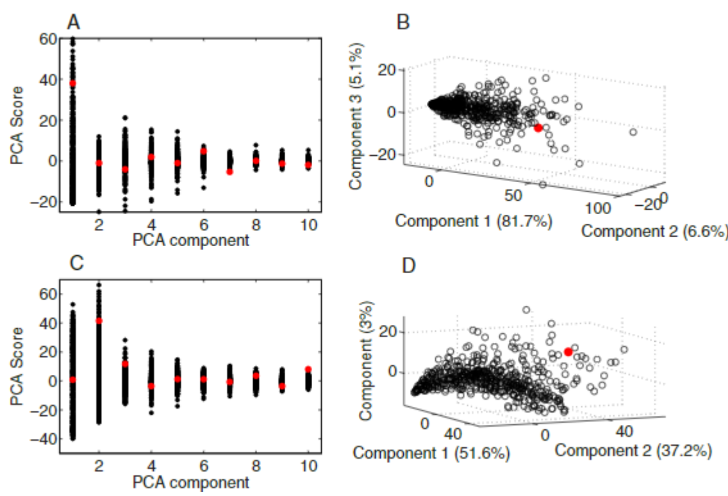

Figure 30: The prior falsification in the reduced dimension space for the one-day experiment and the five-days experiment. The red dot is the observed data, showing that it is not an outlier however poorly sampled around it[Hermans et al., 2018].

Shown in Figure 30 are the distributions of all PCA components and the the distribution of the three most influential components.

5. Design of uncertainty reduction on prediction variables based on data
   In this case direct measurements of the temperature are performed. This observed, measured data is in the domain of the distribution of the prior. In the reduced dimension space the observed data is present on a well sampled place. This is visible in Figure 31.

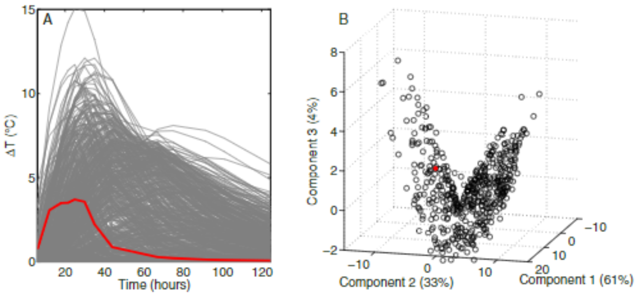

Figure 31: Direct measurements of the temperature changes are shown. The red line and dot are the observed data[Hermans et al., 2018].

The prior is not falsified and the predicting of the temperature at the well during the experiment of the heat storage can take place. In this case, the first ten dimensions of the data variables and the first two dimensions of the prediction variables are used for the one-day experiment, representing 99% of the variance, which is good. Shown in Figure 32, the temperature change is decreasing very fast once the pumping starts. This indicates that the capacity of the storage of heat is in proportion low. The uncertainty reduction is high comparing the posterior and the prior. In the posterior distribution the visible percentiles are almost similar however this is not the case with the prior distribution. This implicates that the thermal energy recovery of the 50% is very low.
The five-days experiment has more information to predict the heat storage. However the uncertainty is a bit bigger than the one for the one-day experiment. The 90% percentile shows that for the five-days experiment it is more likely to have a higher heat storage capacity. The five-days experiment is more valid however it does influence the computational costs due to more data acquisition [Hermans et al., 2018].

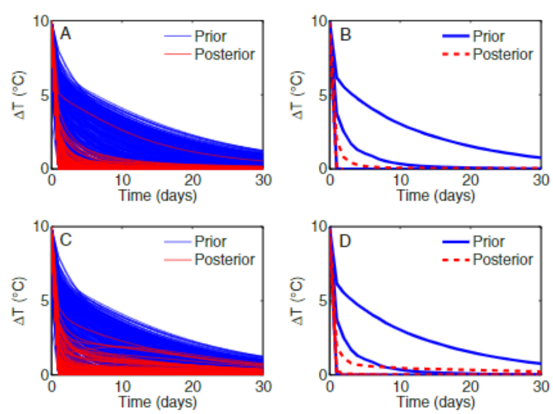

Figure 32: The prior and posterior distributions of the prediction are given. The top is information about the one-day and the bottom about the five-day experiment. The dotted red lines are the 10, 50 and 90 % percentiles of the posterior distribution of the prediction[Hermans et al., 2018].

The different time steps cause variation in the posterior distribution. The uncertainty is depending on the temperature change and the temperature distribution.

6. Posterior falsification and sensitivity, decision making
The direct measurements of the temperature changes that are performed validate the predictions at

those certain locations shown in Figure 33. This shows that BEL has achieved to reduce the uncertainty. The uncertainty reduction in comparison to the prior distribution is high, however the variations of the temperature are still present.

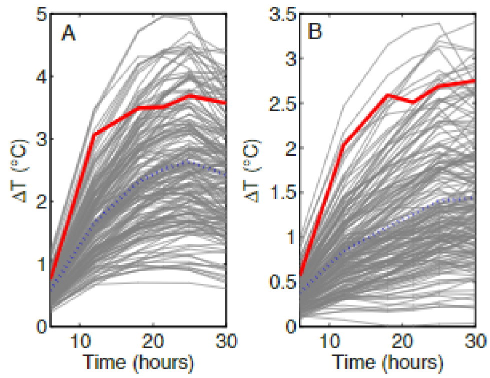

Figure 33: The direct measurements at the two well locations are validating the posterior distribution [Hermans et al., 2018].

Case 3: A new framework for experimental design using Bayesian Evidential Learning: The case of wellhead protection area [Thibaut et al., 2021]. This case will show that BEL is able to reduce the uncertainty of tracing experiments to predict the wellhead protection area. Meaning that the training set of N models will be used as data variables and observed data variables. BEL has the ability for the posterior uncertainty to correspond to new input and directly be used for the posterior distrubution.

1. Formulating the decision question and statement of prediction variables This case is about preventing the groundwater to be contaminated and to extract it in a sustainable way. Contaminants can be reaching the pumping well, this will be determined by the flow velocity around the well which can be calculated by using particle tracking, transport simulation or tracer tests. The targets, wellhead protection areas (WHPAs) surrounding the pumping well, are stochastically predicted using breakthrough curves (BCs) from tracing experiments as predictors.
   The aim is to find the best location for the injection well considering the logistic and economic issues. The problem set up is shown in Table 3, the pumping well surrounded by injection wells. The injection wells will have individual tracers that will model their transport and obtain the BCs at the pumping well location.

2. Statement of model complexity and prior uncertainty
   The experiment in this case is depending on the unknown hydraulic conductivity **K** field. The parameters of the variogram and the parameters of the structural uncertainty, shown in Table 3 are considered.

Table 3: Parameters of the uncertainty variables of case 3[Thibaut et al., 2021].

Parameter	**Range of uncertainty**
K Mean ($\frac{m}{d}$)	U[25 to 100]
Range(m)[min,max]	U[25 to 100]
$log_{10}K$ Standard deviation ($\frac{m}{d}$)	U[0.1 to 0.3]
Angle around vertical axis (degrees)	U[-30 to 30]
Range max (m)	U[200 to 400]

3. Monte Carlo and falsification of prior uncertainty using data

First the direct relationship will be detected between predictor and target. This is done by using 400 training models, from the prior distribution of the hydraulic conductivity, K. The collected field data will be directly used to predict the posterior distribution of WHPA. The number and location of the injection wells influence the uncertainty range of the posterior WHPA distribution. In this case a data set of $n = 500$ is used, where 80% of the models is used as the training set and 20% as the models to validate as a test. The WHPA gets simulated using particle backtracking and the BCs by using solute transport modeling. This gives you a 2D grid.
This case has used 1000 samples as training set and 250 samples as test set. One of the predictions is shown in Figure 34. The test variables are in the domain of the training set variables. To conclude from these different predictions is that all the samples test set are in the domain of the training set samples so it does not need to get falsified.
PCA is performed to the data set, the size is reduced by four times and it is remaining 99.87% of its variance. The remaining variance of 0.13% will not be taken into account for the target prediction. These transformations will be back-transformed later on.

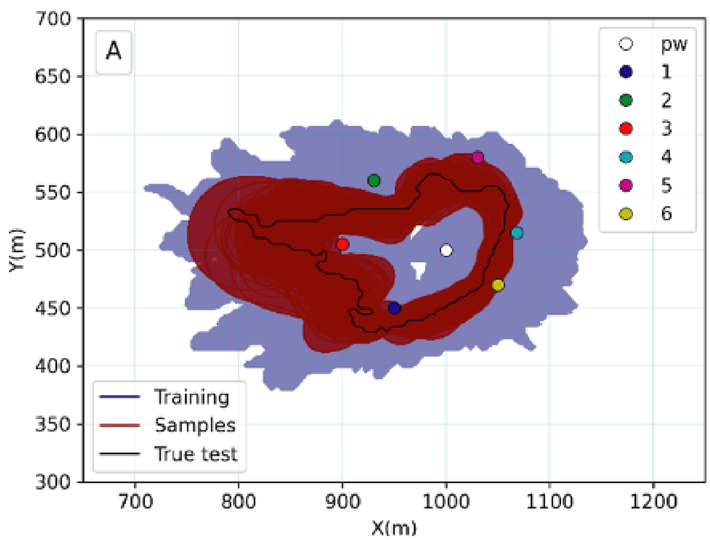

Figure 34: WHPA prediction showing the training data set and the sample data set and the true values[Thibaut et al., 2021]. The pumping well and the six other wells are shown as well.

Shown in Figure 34 is the inferred posterior distribution. The uncertainty reduction can be seen by comparing the prediction and the prior prediction range.

4. Sensitivity analysis on both data and prediction variables
   This case is not depending on many variables. The last three shown in Table 3 are considering the structural uncertainty and the two first the variogram uncertainty. This leads to that the paper did not perform a explicit sensitivity analysis.

5. Design of uncertainty reduction on prediction variables based on data
   CCA is applied to the reduced dimension variables. The correlations are very linear obtaining a great visualization of the joint probability distribution comparing them with the prior prediction range. By increasing the number of sampled models in the prior, the relationship will be learned better through BEL. The number is determined by the target's complexity. The ultimate goal is to get a stable predicted range. The test data variables not falsified.

6. Posterior falsification and sensitivity, decision making
   Some PCA components of the target are ignored in CCA so part of the variability cannot be explained. CCA is not always capturing properly the non-linear components, this can lead to bias and result in an overestimation of the posterior uncertainty. An option is to apply KDE instead of linear regression [Michel et al., 2020]. BEL has a risk of producing unrealistic posterior distribution. The prediction uncertainty may be slightly overestimated, due to the PCA and CCA which simplifies the problem. In this case, linear CCA combined whit linear regression was good enough to predict the WHPA. The size of the training set is case dependent, mainly on the complexity of the target. A large prior would reduce the risk of misestimating the predictor.

Case 4: Managing Uncertainty in Geological CO2 Storage Using Bayesian Evidential Learning [Tadjer and Bratvold, 2021] This case will that BEL has the ability to predict to prevent storage reservoir leakage and drinking water to be contaminated. Carbon capture and storage (CCS) is used to reduce the CO_2 emissions. This can cause storage reservoir leakage and it can contaminate the drinking water in the groundwater aquifers. The aim of this case is to look at the distribution of the CO_2 mass and leakages in the top layer in the future.

1. Formulating the decision question and statement of prediction variables
   This aim of this case for the Utsira saline aquifer, located west of Norway, is that BEL predicts the amount of CO_2 storage room and possible leakages and it will help improve to make a decision concerning CO_2 storage. This will be done by looking at mass and leakage of CO_2.

2. Statement of model complexity and prior uncertainty
   The model m contains the uncertain parameters using the historical knowledge, prior data d variables, which is the CO_2 saturation that is near the well-bore region. The prediction h variables are the amount of CO_2 mass and leakage.
   The parameters that will be looked at to reduce its uncertainty of this case are shown in Table 4.

   Table 4: Parameters of the uncertainty variables of case 4[Tadjer and Bratvold, 2021].

Parameter	**Range of uncertainty**
Depth (m)	U[300 to 1400]
Storage capacity (Gt)	U[0.5 to 1.5]
Permeability (D)	U[0.5 to 2.5]

3. Monte Carlo and falsification of prior uncertainty using data
   The set up of this case is that 200 prior geological realizations are produced, generated by Monte Carlo applied to the prior distributions. A reference model is used, which is the black line in Figure 35 and the red line is the prior probability density function, these graphs show the prior distribution of the prediction variables.

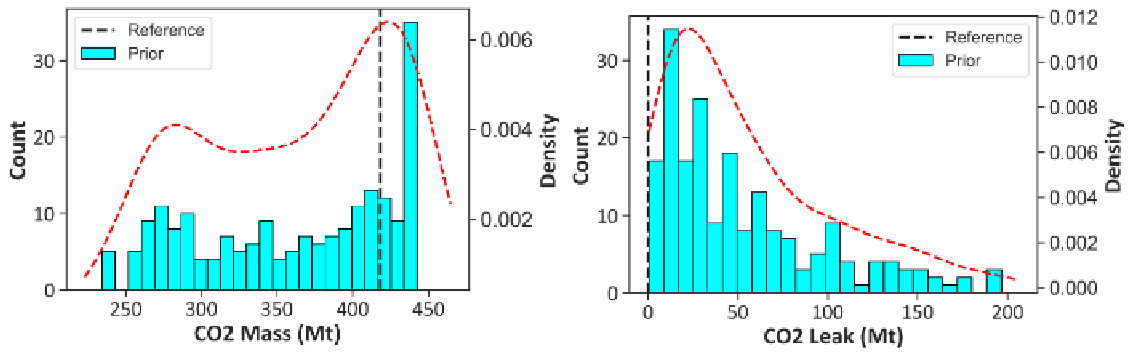

Figure 35: Prior distribution of the prediction data variables[Tadjer and Bratvold, 2021].

The falsification is done by looking at the quality of the injection well of the 200 prior models and d_{obs}, whether outliers can be detected. The Mahalanobis Distance (MD), explained in A.2.2 and shown in Figure 42 is below the 95-percentile threshold, so the model is not falsified, shown in Figure 36

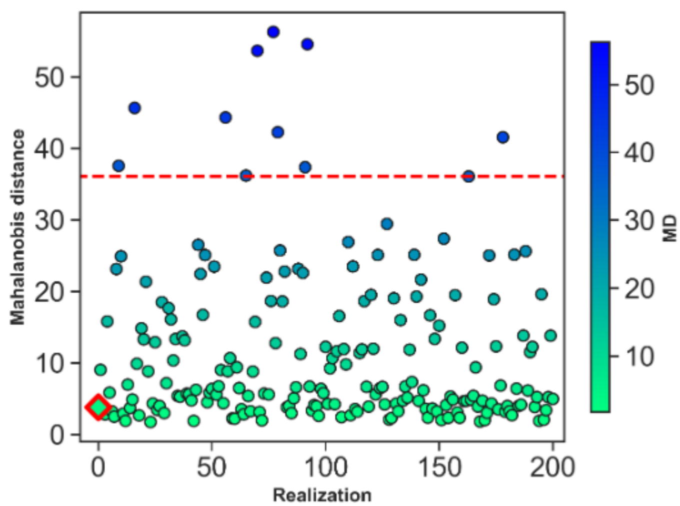

Figure 36: Prior falsification using MD, the red line shows the 95-percentile[Tadjer and Bratvold, 2021].

The aim is to get a statistical relationship between the data and forecast variables. PCA is performed, three dimensions are kept to have a high enough variance (90 %) and to reduce the complexity.
CCA is applied afterwards on the dimension reducted variables to maximize the linearity, however the relation is not unique linear. The PCA and CCA applications for the CO_2 mass are shown in Figure 37 and for the CO_2 leakage in Figure 38.

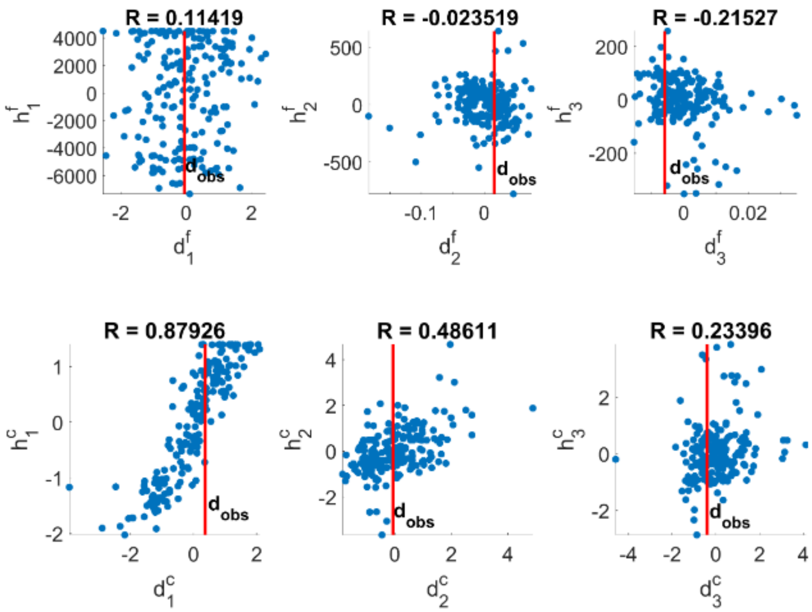

Figure 37: PCA correlation analysis on top and in a CCA domain below for the CO_2 mass. The red line is d_{obs}[Tadjer and Bratvold, 2021].

What is visible that for the CO_2 mass, the CCA works good for the first dimension however for the second and third there is little difference.

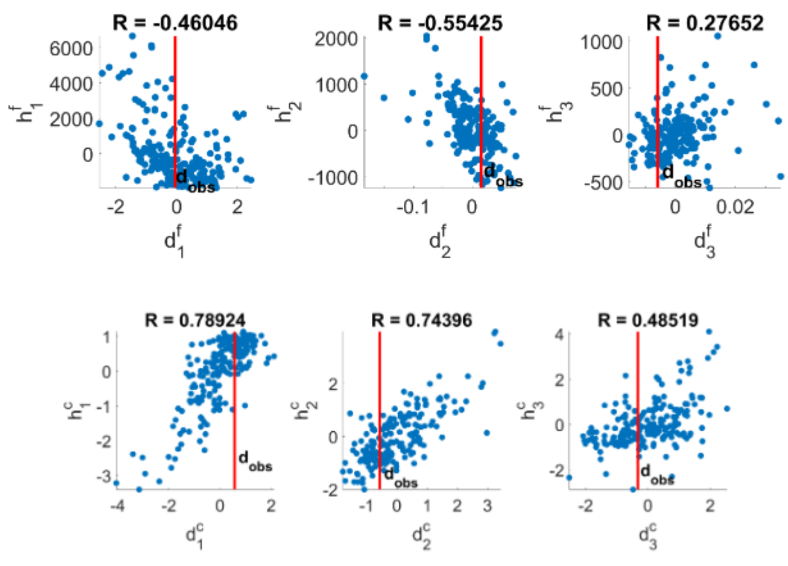

Figure 38: PCA correlation analysis on top and in a CCA domain below for the CO_2 leakage. The red line is d_{obs}[Tadjer and Bratvold, 2021].

For the CO_2 leakage the CCA is improving the linearity of the first and second dimension.

4. Sensitivity analysis on both data and prediction variables
The sensitivity analysis has been based on prior knowledge, (Nilsen et al., 2015) shows that the variance

of the porosity influences the total rock volume that the plume connects with and the permeability affects the CO_2 plume flows behavior. Pressure and temperature variations impact the density which influences the storage capacity of CO_2, shown in Figure 39.

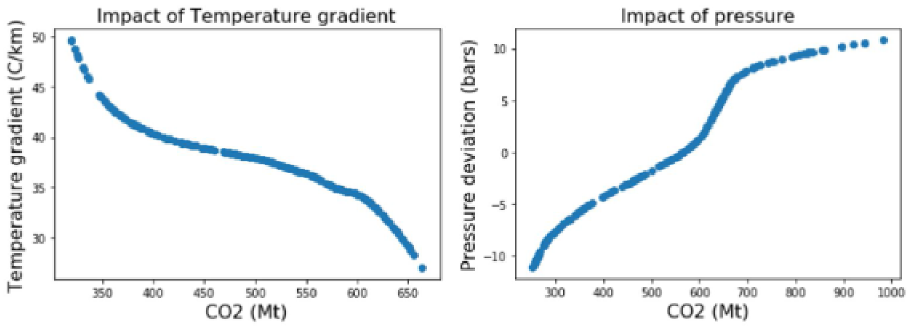

Figure 39: The impact of the pressure and the temperature gradient on the storage capacity of CO_2 [Tadjer and Bratvold, 2021].

5. Design of uncertainty reduction on prediction variables based on data
   Since the linear relationship has been maximized, the linear Gaussian regression equation has been applied, since it is linear enough, to calculate the posterior distribution of the prediction variables. After the latter is obtained, it will be back transformed to the original space. The result of the posterior CO_2 mass and leakage are shown in Figure 40. The uncertainty range has been reduced by applying BEL's framework.

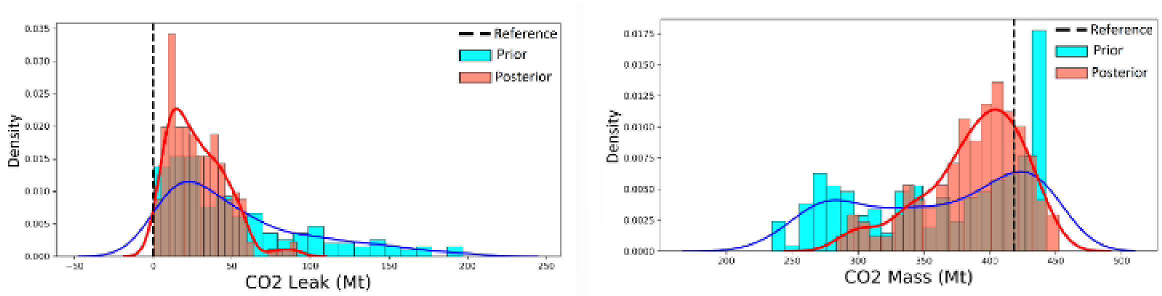

Figure 40: Posterior prediction in original space of the CO_2 mass and leakage by using BEL[Tadjer and Bratvold, 2021].

6. Posterior falsification and sensitivity, decision making
   A statistical forecast of the interest quantities, CO_2 mass and CO_2 leakage is what is achieved. This is reached without experiencing problems with history matching or iterative data inversion, it offers an outcome of CO_2 leakage and CO_2 mass forecasts. The decision can be made on how much CO_2 will be stored without leakage.

3.2.3 Advantages and pitfalls of Bayesian Evidential Learning

In this section the advantages and the pitfalls on the BEL approach will be shown. It is based on the different cases and theories.

The advantages of using BEL as approach for the uncertainty quantification of a process are the following. A case where a correlation between the model parameters is present and is defined in the prior, without changing the framework, it is possible for BEL to perform. No modification or inversion is needed [Hermans et al., 2018]. BEL is able to show the statistical link between the relevant variables and measured data, in contrast to unified or sequential formulations of inversion [Bogrash, 2020]. BEL can work with a narrow number of models which results in fast computations and the need of less data [Thibaut et al., 2021] meaning it can tackle problems where there is not many samples to use [Glacken and Snowden, 2001]. It does not demand advanced computational capabilities and only a small bit of forward modeling is necessary, which is time wise a great advantage. Another great difference between BEL and inversion models is that there is no assurance that inverted models will hold true when new information is gathered in the future. Every time new data becomes available, the model needs to be updated and rebuilt. BEL's model can immediately be updated during the development of the process and when new data occurs[Hof, 2022]. Additionally, moving forward is costly and challenging. Consequently, over the course of a project, this may result in fluctuations or contradictory forecasts. Inverse problems need often ad-hoc modifications that can lead to a bias[Scheidt et al., 2018].

Any collection of data that is consistent with the earlier distribution can be subjected to BEL. In contrast to an inversion, it provides a much larger and more comprehensive approach [Thibaut et al., 2021].The inconsistency between the prior and the data is pretty quickly indicated by BEL. BEL provides a straightforward and distinctive framework to incorporate any kind of uncertainty (conceptual decision, parameter distribution, spatial uncertainty). The uncertainty quantification is simpler and quicker. BEL makes use of multiple forward model runs that works faster than data inversions [Hof, 2022]. These forward simulations can if need be completely parallelized which reduces the costs. A requirement for proper risk analysis and judgment [Hermans et al., 2018].
In parametric regression, a mathematical expression can be obtained from the posterior distribution. By doing this, computationally potentially costly numerical interactions to estimate the posterior are avoided [Scheidt et al., 2018]. The use of BEL in experimental design enables the identification of data sources that maximize the information richness of any measurement data while adhering to financial restrictions and minimizing computing expenses[Thibaut et al., 2021].
BEL can be used for unidimensional interpretations of data since it works with linear and non-linear relationships and all different techniques that can do dimension reduction and transform the realizations back to the original space [Michel et al., 2020].

The BEL approach is known to have pitfalls, the following are the most significant. The predictions are based on prior knowledge but the posterior models can show behavior that is physical not plausible. The sampled values are never guaranteed to be observed inside the prior because the forecast is produced statistically. This can result in implausible answers[Hermans et al., 2019]. It can also be due to simplifying the problem by using PCA and CCA causing an overestimation in the prediction uncertainty. It is unable to produce samples that are clearly different from the prior. As a result, classifying outliers can be difficult because the training set might not have any outliers present [Scheidt et al., 2018].
More simulations are required when the prior knowledge is scanty or when parameters are unclear. The observed data determines the number of models. Inferring that intensively sampled regions of the data space are simpler to predict. So it is influenced by its prior knowledge which is not unlimited in the subsurface[Hermans et al., 2018].

When a case is based on complex priors, the size of the training set is influenced. A large prior reduces the risk of misestimating the relationship between predictor and target [Michel et al., 2020]. BEL has been successfully used for complex priors shown in [Satija and Caers, 2015] and [Hermans et al., 2016].

3.3 Synthesis of literature review

Mineral Resource Modeling has ranges of uncertainties influencing the process of decision making based on risk analysis. In todays industry these uncertainties are reduced by inversion techniques such as conditional simulation and geostatitical modeling. These techniques are based on various versions of kriging. These methods are conditioned to actual data for models to be obtained. Stochastic random sampling techniques are introduced to obtain a complete uncertainty quantification however these techniques are time consuming and complex. The inversion models are very expensive due to the many models and the result are not assured to be true when new data is applied to the model. The conventional deterministic calibration methods for computing the dimensions of an particular area may not be suitable because MRM can be very complex and the methods used in todays industry do not account for the uncertainty present in such prediction problems, which is caused by a limited understanding of the subsurface's heterogeneity.

BEL is an approach where the model is based on assumptions, prior and global knowledge. This is a convenient trade for the subsurface. It enables to model a posterior distribution in the prior model space. Based on statistical relationships between predictor and data variables a posterior prediction can be made. Based on one measured data variable, the model can provide a posterior distribution prediction. BEL can still successfully be used when the problem is very complex, by dimension reduction and regression analysis. The original space can easily be back transformed creating the availability for other analyses to be performed on the reduced space. BEL does not need many models and gives a quick and straightforward result.

The cases of (Hermans et. al, 2019), (Hermans et. al, 2018), (Thibaut et. al, 2021) and (Tadjer and Bratvold, 2021) are showing that BEL can reduce the uncertainties. The temperature in an alluvial aquifer, the efficiency of the thermal energy storage capacity in an alluvial aquifer, the wellhead protection areas surrounding the pumping well using tracing experiments as predictors and the prediction of leakages of CO_2 and the storage of CO_2 are all shown to be reduced in uncertainty. These are all related to geology but not yet in a Mineral Resource domain. Considering these four cases have shown by using BEL the uncertainties are successfully reduced, the uncertainties in MRM will be compared and linked to these in the case studies. It will be shown in a recommended case study for the future.

4 How BEL can be applied to reduce uncertainty in MRM

This section will propose a case study to show how BEL and MRM can be linked. This will be done by using geological and geochemical properties in order to predict geometallurgical properties (e.g. rock hardness). The hardness is falling under the variability within the block [Dominy et al., 2018]. The aim of this case study is to show that the use of BEL will reduce the uncertainty in the prediction of the hardness based on the above mentioned properties. This is done by reducing the uncertainty range in the proxy variables such as penetration rate, Leeb measurements and bound index. The proxy variables are used to estimate the hardness of the rock. For example, the penetration rate can be measured while drilling. This variable is an indication of the hardness of the rock i.e. the softer the rock the higher the penetration rate. A cross section of such a drilling data is shown in Figure 41. Steps that have to be taken in order to apply BEL will only be explained in a descriptive way.

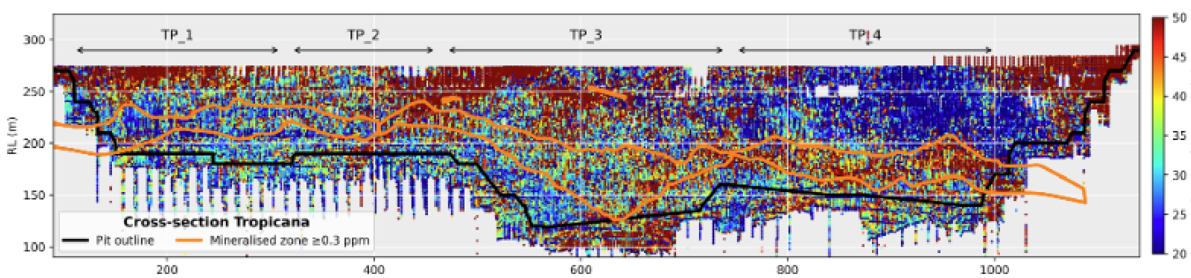

Figure 41: A cross section of penetration data through Tropicana Gold Mine[van Duijvenbode et al., 2022].

1. Formulating the decision question and statement of prediction variables
   The objective is to reduce the uncertainty in the estimation of the hardness of the orebody using the data obtained from the exploration drillholes. The hardness of the rock present in the orebody will be the prediction variable.

2. Statement of model complexity and prior uncertainty
   The model will be based on geological and geochemical properties such as the lithology, the grade and the mineral type. These properties are obtained from samples or sample points that are spatially correlated [Dominy et al., 2018]. These samples will be obtained from exploration data such as a drillhole. It can also be obtained by geophysical methods such as the seismicity or blasted drillholes. These will have a certain range of uncertainty. The uncertainty of these parameters is assessed by using the same model to the range of the simulated parameter realizations. In this open pit faults and dykes are present [van Duijvenbode et al., 2022] which fall under the MRM uncertainties described in the prior.

3. Monte Carlo and falsification of prior uncertainty using data
   Monte Carlo realizations are obtained based on the exploration drillholes and assumed uncertainty range. These realizations give all kinds of results to estimate the hardness from hard to soft. After obtaining these models, a relationship needs to be obtained between lithology, grade and mineral type variables and the hardness variables. PCA should be applied to get a better visualization. In this case a 2D domain is used. The variables with most impact on the hardness are kept and looked at. Then the observed data from the drill holes which has the hardness of the rock present will be used to compare and see if the prior model needs to get falsified.

4. Sensitivity analysis on both data and prediction variables
   This will be done on all variables, looking which have the most influence on the prediction.

5. Design of uncertainty reduction on prediction variables based on data

   To apply the regression analysis, in this case assumed for the variables to be linear. CCA is then applied to find the combination of the variables that has the maximized linearity between the penetration data and the prediction data. In this domain space will the prediction be made and then back transformed to its original space.

6. Posterior falsification and sensitivity, decision making
   When the hardness predictions are in its original space and it is not falsified by the observed data from the drill holes the decisions can be made. By the predictions of the hardness, one can now decide what speed will be used to start exploiting the orebody.

5 Discussion and recommendations

The aim of this research is to substantiate the hypothesis that BEL can be used to reduce uncertainty in MRM. To prove this hypothesis, a literature review has been done on MRM's uncertainties and the techniques used in today's industry. Several papers were reviewed, it became clear that not all have been updated which can lead to wrong assumptions.

The significance to reduce uncertainty in MRM is that exploiting mineral resources can cause serious difficulties since they could be marginal deposits, happening at considerable depths or having other access limitations. BEL should be considered to reduce the uncertainties present in MRM since these difficulties can have huge impacts and cause disasters.

BEL has a risk of producing unrealistic posterior distribution. The prediction uncertainty may be slightly overestimated, due to the PCA and CCA which simplifies the problem. However in the cases studied where BEL is applied, the variance was able to be kept above 90%. The size of the training set is case dependent, mainly on the complexity of the target. So a large prior reduces the risk of misestimating the predictor and target relationship shown in (Michel et al. 2020). BEL has been successfully used for complex priors shown in (Satija et al. 2015) and (Hermans et al. 2016).

The next step is to demonstrate BELs utility in other domains as mentioned in the case studies. These show that BEL indeed reduces uncertainty. However these cases are not based on mineral resources so it does not proof immediately that BEL can be used for MRM.

This research provides the background knowledge for further analyses regarding the use of BEL for MRM. The research performed was to prove that BEL can be used for UQ in MRM. This is shown in a descriptive case study. A next step could be to perform a case study with actual data and model it. The results obtained can then be compared against existing data.
This case is focusing on a geometallurgical property of MRM and many other uncertainties have been mentioned. In further research these should be looked at as well for a UQ.

6 Conclusion

This thesis proposes to use a different framework than is used in the industry today, namely BEL, to estimate the uncertainty of the prediction for MRM. By reviewing the uncertainties in MRM and how BEL is applied in different case studies a new case study is proposed.

The uncertainties in MRM lie in geological, geochemical and geometallurgical properties. In the industry of today the uncertainties are reduced through complex models, such as inversion, and time consuming methods. Every time new data becomes available, the model needs to be updated and rebuilt.

BEL constitutes statistic based relationships between simulated data and associated model parameters, estimated from data. The model requires this data when making prediction. The predictions are estimated or learned from data. BEL depends on prior knowledge consisting of uncertainty. By applying a sensitivity analysis the level of influence from the parameters on the prediction are obtained. The consistency of the prior data is checked through falsification.
Next up, the predictions can be made. This technique makes it possible to calculate the prediction's level of uncertainty, providing a foundation for risk analysis and fully informed decision-making. The advantage of only needing limited amount of forward simulations and no model calibration through data inversion, it will cost less money and time. The various case studies that have been discussed show that BEL has successfully reduced the range of uncertainty. In the (Hermans et. al, 2019) the uncertainty of the temperature in an alluvial aquifer, in (Hermans et. al, 2018) the uncertainty of the efficiency of the thermal energy storage capacity in an alluvial aquifer, in (Thibaut et. al, 2021) the uncertainty of the wellhead protection areas surrounding the pumping well using tracing experiments as predictors and in (Tadjer and Bratvold, 2021) the uncertainty of the prediction of leakages of CO_2 and the storage of CO_2.

The link between BEL and MRM is shown through a case study. This will be done by using geological and geochemical properties in order to predict geometallurgical properties (e.g. rock hardness). Reducing the uncertainty is done by reducing the uncertainty range in the proxy variables such as penetration rate, Leeb measurements and bound index. The proxy variables are used to estimate the hardness of the rock. For example, the penetration rate can be measured while drilling. This variable is an indication of the hardness of the rock i.e. the softer the rock the higher the penetration rate. This is showing that BEL can be used for MRM to reduce uncertainty.

References

[Par, 2015] (2015). Comparing sequential gaussian and turning bands algorithms for cosimulating grades in multi-element deposits. *Comptes Rendus - Geoscience*, 347:84–93.

[Hof, 2022] (2022). Modeling geospatial uncertainty of geometallurgical variables with bayesian models and hilbert–kriging. *Mathematical Geosciences*.

[Abzalov, 2016] Abzalov, M. (2016). *Applied Mining Geology*.

[Ambrosone et al., 2019] Ambrosone, M., Giuliani, G., Chatenoux, B., Rodila, D., and Lacroix, P. (2019). Definition of candidate essential variables for the monitoring of mineral resource exploitation. *Geo-Spatial Information Science*, 22:265–278.

[Bai and Tahmasebi, 2022] Bai, T. and Tahmasebi, P. (2022). Sequential gaussian simulation for geosystems modeling: A machine learning approach. *Geoscience Frontiers*, 13.

[Blenkinsop et al., 2020] Blenkinsop, T., Oliver, N., Dirks, P., Nugus, M., Tripp, G., and Sanislav, I. (2020). Chapter 1 structural geology applied to the evaluation of hydrothermal gold deposits. *Economic Geology*, 21:1–23.

[Bogrash, 2020] Bogrash, A. (2020). Bayesian evidential learning as an alternative to spatial inverse problems. *CCG Annual Report*.

[Caers, 2018] Caers, J. (2018). 02 bayesian evidential learning.

[Committee, 2012] Committee, J. O. R. (2012). The jorc code 2012 edition ee ective 20 december 2012 and mandatory from 1 december 2013 australasian code for reporting of exploration results, mineral resources and ore reserves prepared by the joint ore reserves committee of the australasian institute of mining and metallurgy, australian institute of geoscientists and minerals council of australia (jorc).

[Di Fiori et al., 2015] Di Fiori, R., Long, S., Muntean, J., and Edmondo, G. (2015). Structural analysis of gold mineralization in the southern eureka mining district, eureka county, nevada: A predictive structural setting for carlin-type mineralization. *Geological Society of Nevada Symposium Proceedings*.

[Dohm, 2005] Dohm, C. (2005). Quantifiable mineral resource classification: A logical approach.

[Dominy et al., 2002] Dominy, S. C., Noppé, M. A., and Annels, A. E. (2002). Errors and uncertainty in mineral resource and ore reserve estimation: The importance of getting it right.

[Dominy et al., 2018] Dominy, S. C., O'connor, L., Parbhakar-Fox, A., Glass, H. J., and Purevgerel, S. (2018). Geometallurgy—a route to more resilient mine operations.

[Dumakor-Dupey and Arya, 2021] Dumakor-Dupey, N. K. and Arya, S. (2021). Machine learning—a review of applications in mineral resource estimation. *Energies*, 14.

[Freier, 2016] Freier, L. (2016). Where can i find a good derivation for conditional simulation?

[Funedda et al., 2018] Funedda, A., Naitza, S., Buttau, C., Cocco, F., and Dini, A. (2018). Structural controls of ore mineralization in a polydeformed basement: Field examples from the variscan baccu locci shear zone (se sardinia, italy). *Minerals*, 8.

[Geologist, 2020] Geologist, M. (2020). Why drillhole compositing before resource estimation is important?

[Glacken and Snowden, 2001] Glacken, I. M. and Snowden, D. V. (2001). Mineral resource estimation, in mineral resource and ore reserve estimation-the ausimm guide to good practice mineral resource estimation.

[GmbH, 2020] GmbH, S.-S. (2020). The concept of dilution is introduced here. there are four different types of dilution in mining: planned, unplanned, external and internal dilution.

[Hermans et al., 2019] Hermans, T., Lesparre, N., Schepper, G. D., and Robert, T. (2019). Bayesian evidential learning: a field validation using push-pull tests. *Hydrogeology Journal*, 27:1661–1672.

[Hermans et al., 2018] Hermans, T., Nguyen, F., Klepikova, M., Dassargues, A., and Caers, J. (2018). Uncertainty quantification of medium-term heat storage from short-term geophysical experiments using bayesian evidential learning. *Water Resources Research*, 54:2931–2948.

[Hermans et al., 2016] Hermans, T., Oware, E., and Caers, J. (2016). Direct prediction of spatially and temporally varying physical properties from time-lapse electrical resistance data. *Water Resources Research*, 52:7262–7283.

[Hosseini and Lewis, 2009] Hosseini, Z. and Lewis, R. (2009). Difficult mining conditions in 153 orebody at a depth of 4550 ft at vale inco's coleman mine.

[Javed et al., 2017] Javed, S., Ali, A., and Ullah, S. (2017). Spatial assessment of water quality parameters in jhelum city (pakistan). *Environmental Monitoring and Assessment*, 189.

[Jpb, 2020] Jpb (2020). Secondary planar structural elements.

[Kasa, 2018] Kasa, F. (2018). Geostatistics-sgems:sequential gaussian simulations.

[Lee et al., 2020] Lee, W. J., Mendis, G., Triebe, M., and Sutherland, J. (2020). Monitoring of a machining process using kernel principal component analysis and kernel density estimation. *Journal of Intelligent Manufacturing*, 31.

[Leuangthong, 2021] Leuangthong, O. (2021). The future of mineral resource modelling. *SRK News*, 63.

[Maleki and Emery, 2020] Maleki, M. and Emery, X. (2020). Geostatistics in the presence of geological boundaries: Exploratory tools for contact analysis. *Ore Geology Reviews*, 120.

[Michel et al., 2020] Michel, H., Nguyen, F., Kremer, T., Elen, A., and Hermans, T. (2020). 1d geological imaging of the subsurface from geophysical data with bayesian evidential learning. *Computers and Geosciences*, 138.

[Nabavi and Fossen, 2021] Nabavi, S. T. and Fossen, H. (2021). Fold geometry and folding – a review.

[Ortega et al., 2004] Ortega, L., Lunar, R., Moreno, T., Prichard, H., and Fisher, P. (2004). The aguablanca ni-cu-pge deposit, southwestern iberia: Magmatic ore-forming processes and retrograde evolution. *The Canadian Mineralogist*, 42:325–350.

[Robb, 2005] Robb, L. (2005). Introduction to ore-forming processes.

[Satija and Caers, 2015] Satija, A. and Caers, J. (2015). Direct forecasting of subsurface flow response from non-linear dynamic data by linear least-squares in canonical functional principal component space. *Advances in Water Resources*, 77:69–81.

[Scheidt et al., 2018] Scheidt, C., Li, L., and Caers, J. (2018). *Quantifying uncertainty in subsurface systems.*

[Tadjer and Bratvold, 2021] Tadjer, A. and Bratvold, R. B. (2021). Managing uncertainty in geological co 2 storage using bayesian evidential learning.

[Thibaut et al., 2021] Thibaut, R., Laloy, E., and Hermans, T. (2021). A new framework for experimental design using bayesian evidential learning: The case of wellhead protection area. *Journal of Hydrology*, 603.

[van Duijvenbode et al., 2022] van Duijvenbode, J. R., Cloete, L. M., Shishvan, M. S., and Buxton, M. W. (2022). Material fingerprinting as a tool to investigate between and within material type variability with a focus on material hardness. *Minerals Engineering*, 189:107885.

[Wang and Huang, 2012] Wang, G. and Huang, L. (2012). 3d geological modeling for mineral resource assessment of the tongshan cu deposit, heilongjiang province, china. *Geoscience Frontiers*, 3:483–491.

A Appendix

A.1 Bayes' Theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{6}$$

Bayes' Theorem states that the conditional probability of an event, based on the occurrence of another event, is equal to the likelihood of the second event given the first event multiplied by the probability of the first event.
Bayes Theorem provides a useful method for thinking about the relationship between a data set and a probability. In other words, the theorem says that the probability of a given hypothesis being true based on specific observed data can be stated as finding the probability of observing the data given the hypothesis multiplied by the probability of the hypothesis being true regardless of the data, divided by the probability of observing the data regardless of the hypothesis.

A.2 Explanation BEL

A.2.1 Regression Analysis BEL

$$f(h^*) = const \times exp(-\frac{1}{2}(h^* - \bar{h^*})^T C^{-1}_{h^*h^*}(h^* - \bar{h^*})) \tag{7}$$

Giving the next formula:

$$d^* = Ah^* + \epsilon^* \tag{8}$$

Where A consists of unknown coefficients that map h^* to d^*.
To estimate A, the training set is used that was generated from the prior models. This will give the likelihood as a Gaussian centered around d^*_{obs} and the covariance $C_{d^*d^*}$. Get the minimum A that minimizes the sum of the squared errors. ϵ^* is assumed to be Gaussian with zero mean and covariance C^*_e [Scheidt et al., 2018]. This process starts of by having $f(h^*)$ expressed as the Gaussian prior. This will give the mean and the covariance, Formula shown in A.2.1. The relation between h^* and d^*_{obs} is the likelihood function. How well will this prediction variable link to the observed data variable. So how well is the model at generating h^* and d^*_{obs}. Assumed, as mentioned before, a linear relationship between the variables, Formula shown in **??**. The goal is to get the minimized sum of the squared errors [Scheidt et al., 2018]. This gives the likelihood and the covariance. The latter has also undergone dimension reduction and this is, in the case of a linear process, done by a linear operator. The covariance is a source of error that can be appear due to imperfect fitting or by measuring the observed data. The first can be computed using residuals from training data fitting and the last is the effect of independent factors such as noise.
So to have the Gaussian prior distribution and the likelihood function, the mean $\bar{h^*}$ and the covariance $C_{\hat{h^*h^*}}$ of the posterior distribution of the prediction variable is expressed as the following[Scheidt et al., 2018].

$$\bar{h^*} = \bar{h^*} + C_{\hat{h^*h^*}}\hat{A}^T(\hat{A}C_{\hat{h^*h^*}}\hat{A}^T + \hat{C^*_{dd}} + \hat{C}_e)^{-1}(d^*_{obs} - \hat{A}\bar{h^*}) \tag{9}$$

$$C_{\hat{h^*h^*}} = C_{\hat{h^*h^*}} - C_{\hat{h^*h^*}}\hat{A}^T(\hat{A}C_{\hat{h^*h^*}}\hat{A}^T + \hat{C^*_{dd}} + hatC_e)^{-1}\hat{A}C_{\hat{h^*h^*}} \tag{10}$$

A.2.2 Mahalanobis Distance

The Mahalanobis distance (MD) is a measure of the distance between a point P and a distribution D. It is a multi-dimensional generalization of the idea of measuring how many standard deviations away P is from the mean of D. This distance is zero for P at the mean of D and grows as P moves away from the mean along each principal component axis. If each of these axes is re-scaled to have unit variance, then the MD corresponds to standard Euclidean distance in the transformed space. The MD is thus unitless, scale-invariant, and takes into account the correlations of the data set shown in 42.

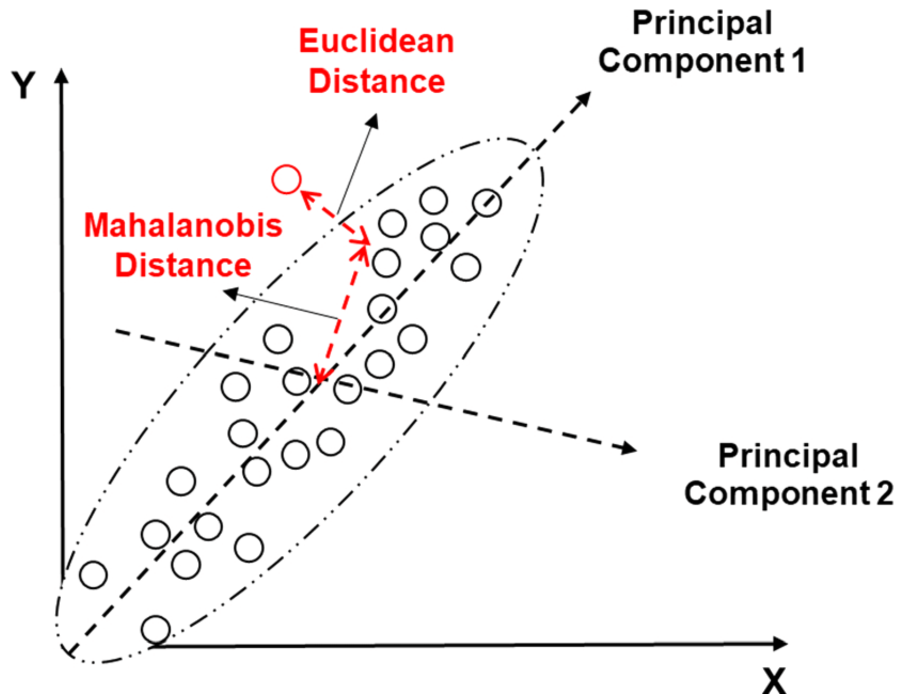

Figure 42: Shown MD and Euclidean Distance [Lee et al., 2020].