Author's Response

RC1 (Anonymous Referee #1)

This is an interesting paper, I think that the use of loss functions in this setting is a useful development and one with considerable potential, and I think that the paper is to be welcomed because of this.

- We thank the reviewer for this motivating content on the topic itself.

That said, the account of the statistical formulation of the author's model is very unclear and needs considerable improvement.
In section 2.2 the authors present their synthetic setting. This can be modelled, by an implicit modelling algorithm, from some set of input values. I assume that these input values are notional values of the z coordinate for a particular contact at each of a set of locations $\{x, y\}$ in addition to notional values of dip direction and angle.The authors state that they can sample 'deviation values' (I assume they mean errors) for these observations which are then added to the 'known' values. With one major caveat (below), the application of the implicit modelling algorithm to each of a series of 'observations' with stochastic error would generate a series of models of the underlying state of affairs. The caveat is that the authors give no indication of how they specify the joint distribution of errors, i.e. the joint distribution for the errors in z at locations x_1, y_1 and x_2, y_2. One can assume only that they treat these errors as independent, but that is most implausible in any geological setting, and could have dramatic consequences for the resulting model output. There are several studies, including some cited in this paper, which have considered the spatial dependence of model errors, and the authors should give this careful consideration in revising their work.

- The reviewer is right in his criticism of the description of our probabilistic model. We did not focus our attention too much on this aspect, as the probabilistic model is not the central point of the paper (see below).
- In our conceptual case study, errors are indeed treated as independent. The model is primarily aimed at providing an adequate base for generating distributions to apply our loss function to. This is, however, not a limitation - any knowledge about parameter correlations could be included in a concrete application (as the ones mentioned by the reviewer).
    - We clarified this aspect on P11 L18-20

The authors then attempt to set their work in a Bayesian context. However, they do not make this at all explicit, and it is very difficult to understand or assess what they have done. A Bayesian analysis requires some model, for which some set of parameters are unknown and are treated as random variables to reflect this uncertainty. Prior probability distributions express the modeller's subjective view, prior to examining the data under consideration, about the values of these parameters. Likelihood values can be computed for proposed values of the parameters, and this, possibly in conjunction with some sampling method such as MCMC, is then used to characterize the posterior distribution of the parameters, given the data. Posterior distributions, or samples from them, could then be used, for example, to estimate the expected value of a loss function under a particular decision rule.

While the paper starts with some very generalized statements about Bayesian methods (Section 2.4), we are nowhere given a sufficiently clear account of their model system and how it is expressed in Bayesian form. The jump from 2.4 to 2.4.1 is much too large. It is not possible for the reader to extract from section 2.4.1 a sufficiently clear account of what the authors have done to be able to reproduce it in a comparable setting, or to be confident that it has been done correctly. For example, the authors refer to 'layer thickness likelihoods'. I assume that the thickness of a layer is a model parameter (assumed to be the same everywhere in space?). It is not made clear how the likelihood is defined, given a set of notional observations. Are the observations treated as independent? Is this reasonable? I would certainly be concerned if the thicknesses of the separate layers are treated as independent random variables. This does not strike me as implausible and, as assumptions go, it is likely to be a sensitive one in this particular context.

- We didn't go into too much detail about Bayesian inference, as it is not the central aspect of this work. The loss function approach should equally be applicable in a general probabilistic context, using results from common Monte Carlo simulation.
- We now clearly emphasize the focus on the loss function approach. The manuscript was restructured accordingly: In the methods section, we first describe the concept of loss, loss functions and customization (P3 onwards), with a more in-depth explanation of the customization steps. In addition, we now included a more detailed figure explaining these aspects in a simple 1-D model (P8).
- The significance of Bayesian inference as a method is a secondary aspect of this work. It was used to generate comparable distributions in our synthetic

case study. We use it to update the model with additional information (e.g. combine prior information on interface depth with additional data on layer thickness). We clarified this aspect in the manuscript and provide references to similar previous work (P13-14). However, Bayesian inference is no strict prerequisite for the loss function application.

In short, I think that the authors need to start again, aiming at the simplest and clearest account they can give of their modelling framework so as to exemplify the loss function concept to best advantage. At present this gets lost in vague and hand-waving description.

- We took this comment very seriously, as we realized that the main focus of the paper was apparently not clear. As suggested, we streamlined the entire description and focused on the most important methods, with a clearer definition of the terms. Most importantly, we now included a simple 2-input-parameter "1-D model" in the Figure explanation on P8 and we are confident that this figure now helps clarify the main concept and relevance of the paper.

A couple of minor points. On page 6 line 10 you refer to a negatively skew normal distribution of mean zero used to generate errors of the throw for your fault. One would expect such a distribution to have a median value larger than the mean, and there statement that most of the values produced are negative is curious. Please clarify.

- Yes, the statement was erroneous and is now corrected (P11 L15).

Second, the statement of the loss functions from page 11 onward should be clarified. As stated Equations (4)–(7) are identical upto a constant value. I assume that you actually introduce the constants in (5) onward to introduce different degrees of asymmetry. This should be expressed in a rigorous way, indicating the conditions over which the constant is introduced. With some care a parsimonious but rigorous notation could be developed.

- As mentioned above, the loss function customization is now explained in further detail.

RC2 (Anonymous Referee #2)

This is an interesting and original contribution on simulating uncertainties and their impact on decision making in a structural hydrocarbon trap. Monte Carlo error

propagations and Markov Chain Monte Carlo sampling are used to consider the probability of different trap volume models based on stated uncertainties. Loss functions are used to explore potential decision making scenarios for high-risk and low-risk users. The technique has great potential to be used, with addition of more parameters, for hydrocarbon exploration and other geoscience applications.

- We thank the reviewer for this positive and motivating comment!

The main strengths of the paper are that it is tightly focused on a specific and important problem, it builds up the methods and results in logical step-wise sections, and it contains useful and clear figures to clarify key inputs/findings. The main weaknesses are that it is written in a fairly inaccessible way, with key elements and some assumptions undefined, and that it does not look beyond the broader applications beyond the specific one used.

- We thank the reviewer for pointing this aspect out, as well. As described above, we took this criticism very seriously and restructured the entire manuscript, including now a conceptualized step-by-step description of the customization with explanations in a 2-input-parameter "1-D model" in the Figure on page 8.
- We also extended the discussion with additional potential applications of the method.

There a few ways that it could be made more accessible to a non-specialist in decision analysis (discussed below) and some general editing would shorten the text and increase its clarity.
General: For reviewing, it would be much better to have continuous line numbers to refer to rather than numbering resetting back to 1 on each page, also every line rather than every 5th line should be numbered.

- We agree that this numbering is a bit cumbersome, but we used the standard Copernicus template for our submission and did not find an option for continuous line numbering.

Lots of instances of unnecessary modifiers e.g. line 10 (P1), line 3-4 (P4), convoluted sentences (e.g. line 2-4, P9) and redundant words (e.g. line 20 P12, line 4, P19) make the text difficult to access. Many of these could be drastically simplified, making the text shorter and more accessible.

● Thank you very much for your constructive remarks, and especially for highlighting the strengths and weaknesses you identified. You pointed out numerous valid concerns that we considered and tried to improve upon.

The word 'actor' is used throughout and in the title. It would be good to define the term in the context used early on. I see that it is commonly used in the decision sciences as a synonym for the more widely-understood term 'decision maker' or 'user', and seems to represent the human element of the process. However, in line 2 (P9) the phrase 'the case of an actor or decision maker' is used, suggesting that the two are different –perhaps actor has a very specific meaning here, which should be made clear. In line11 (P9), for example, the 'actor' would like to know the trap volume. Use of actor in this context seems obscure, and would be better replaced with a common word representing the human element in decision analysis like worker/geologist/company/user. Line15 (P17) reports observations of the behaviour of the modelled actors – reflecting that actors are a non-human element of the modelling (e.g. Fig. 3). Please clarify.

- We now use the word "decision maker" preferably, to avoid confusion. But we also define early on that an estimate-based decision can be referred to as an action a, and thereby a decision maker can also be referred to as actor (P3 L27-28).
- We overall still prefer talking about "decision makers" and "actors", due to the quite abstract nature of our approach. Our example application is not defined specifically enough, to name the exact type of individual, the decision maker could be. Therefore we prefer to clarify it in the introduction (P3 L4), as well as again on P4 L25.
- We also talk about companies as an example on P5 L18-21.
- As the sentence on P17 L15 was easily misunderstood, we dropped it.

Several other key concepts, terms and abbreviations are used throughout, but not all are clearly defined. It would be beneficial to place the work in the framework of such definitions.
The use of the hydrocarbon sector as an application for geological modelling is sensible. However, no other options are mentioned. What might other sectors be, and what modelling problems could be solved in the same way (perhaps things like nuclear waste disposal, landslide susceptibility)?

- Relevance for other fields is now referred to at the end of the discussion.

Throughout, volumetrics are discussed in the sense of 'trap volume' – which is appropriate and supported by the parameters used. However, there is one instance of the term 'reservoir volume', and there is consideration, in section 2.3, of OOIP. This indicates that actual hydrocarbon reserves are a key outcome of the analysis. But converting from a trap volume to a reserve volume of course involves extra parameters such as porosity, net:gross ratios, water saturation etc. These factors and their uncertainties may be as important or more important than the overall trap volume. While it is reasonable that other factors aren't included in the present modelling, it would be beneficial to mention them, justify their omission, and perhaps consider how they might be integrated in a future iteration of the model. I think they are alluded to in Fig 7 (reservoir and recovery parameters) but these should be made explicit, given the geologic focus of the topic.

- Very good point. The OOIP context is now established on P11 23 - P13 L9. OOIP is also mentioned before in the introduction, together with the NPV.

Lines 6-8 (P6). Is there actually an independent uncertainty related to fault offset? Since that parameter can only be inferred indirectly via stratigraphic surface picks, I would argue that there is no additional uncertainty on either the hanging wall or footwall beyond what has already been accounted for by the surface uncertainty (which is the sole observational basis for fault offset). It would be useful to see a short descriptive justification for including this additional uncertainty. The significance of this parameter is clear in Fig. 5, where the smaller probabilities of hanging wall seal and reservoir result (I think) from the additional uncertainty applied. If this extra uncertainty is not justified (I think it is not), then it places the subsequent results in doubt.

- *We have now added an according statement that our model is not aimed at complete geological plausibility* was added on P11 L18-20.

Line 14 (P6) OOIP/OOIG – presumably this should be OOIP/OGIP instead?
Line 25 (P6). SSF is not defined. I suspect it may be 'shale smear factor' (used without the abbreviation in line 19 (P18)). Please define accordingly.
Line 18 (P13). 'Low but positive volumes' – is a negative volume possible/meaningful? If not, simply use 'low volumes'.
A few minor typos throughout (e.g. line 24 (P18) should read: an individual's . . ., line 29 (P18) should read: to what extent . . .). Please check and amend generally.

- We overall found your comments and corrections highly useful. We named the Shale Smear Factor on P13 L14 and P15 L5, fixed the "low volumes" statement and corrected the typos.

Actors, actions and uncertainties: Optimizing decision making based on 3-D structural geological models

Fabian Antonio Stamm[1], Miguel de la Varga[1,2], and Florian Wellmann[1]

[1]Computational Geoscience and Reservoir Engineering (CGRE), RWTH Aachen University, Germany
[2]Aachen Institute for Advanced Study in Computational Engineering Science (AICES), RWTH Aachen University, Germany

Correspondence: Fabian Antonio Stamm (fabian.stamm@rwth-aachen.de)

Abstract. Uncertainties are common in geological models and have a considerable impact on model interpretations and subsequent decision making. This is of particular significance for high-risk, high-reward sectors~~, such as hydrocarbon exploration and production~~. Recent advances allows us to view geological modeling as a statistical problem that we can address with probabilistic methods. Using stochastic simulations and Bayesian inference, uncertainties can be quantified and reduced by incorporating additional geological information. In this work, we propose custom loss functions as a decision-making tool that builds upon such probabilistic approaches.

As an example, we devise a case in which the decision problem is one of estimating the uncertain ~~volume of a structural hydrocarbon trap. We construct a synthetic 3-D model to represent a potential hydrocarbon system and develop algorithms for automatic trap volume calculation. Various volume probability distributions for different information scenarios are attained via Monte Carlo error propagation and Markov chain Monte Carlo sampling.~~ economic value of a potential fluid reservoir. For subsequent true value estimation, we design a case-specific loss function to reflect not only the decision-making environment, but also the preferences of differently risk-affine ~~actors~~decision makers. Based on this function, optimizing for expected loss returns an actor's best estimate to base decision making on~~.~~, given a probability distribution for the uncertain parameter of interest. We apply the customized loss function in the context of a case study featuring a synthetic 3-D structural geological model. A set of probability distributions for the maximum trap volume as the parameter of interest are generated via stochastic simulations. These represent different information scenarios to test the loss function approach for decision making.

Our results show that the optimizing estimators shift according to the characteristics of the underlying ~~value~~ distribution. While overall ~~spread~~ variation leads to separation, risk-averse and risk-friendly decisions converge in the decision space and decrease in expected loss given narrower distributions. We thus consider the degree of decision convergence to be a measure for the state of knowledge and its inherent uncertainty at the moment of decision making. This decisive uncertainty does not change in alignment with model uncertainty but depends on alterations of critical parameters and respective interdependencies, in particular relating to seal reliability. Additionally, actors are affected differently by ~~one set of information~~ adding new information to the model, depending on their risk affinity. It is therefore important to identify the model parameters which are most influential for the final decision in order to optimize the decision-making process.

1 Introduction

In studies of the subsurface, data availability is often limited and characterized by high possibilities of error ~~, as well as epistemic uncertainty,~~ due to signal noise or inaccuracies . This, together with the inherent epistemic uncertainty of the modes,

5 leads to the inevitable presence of significant uncertainty in geological models, which in turn may affect interpretations and conclusions drawn from a model (Wellmann et al., 2018; de la Varga and Wellmann, 2016; de la Varga et al., 2019; Wellmann et al., 2010a; Bardossy and Fodor, 2004; Randle et al., 2019; Lark et al., 2013; Caers, 2011; Chatfield, 1995). Uncertainties are thus of particular importance for making responsible and good decisions in related economic settings, such as in hydrocarbon exploration and production (Thore et al., 2002; McLane et al., 2008; Smalley et al., 2008). Quantification and visualization of

10 such uncertainties and their consequences is ~~a currently~~ currently an active field of research. Recent developments allow us to view geological modeling as a statistical problem ~~(see Wellmann and Caumon (2018))~~(see Wellmann and Caumon, 2018). We particularly regard approaches to couple implicit geological modeling with probabilistic methods, as presented by de la Varga et al. (2019) with the Python library GemPy.

~~Following their example, we regard implicit geological modeling as a mathematical function of not only deterministic but~~
15 ~~also stochastic, i.e. uncertain, model input parameters. An essential aspect of this approach to modeling is that it allows for~~ ~~full automation, so that the consequences of a change in an input parameter are realized directly without the need for any~~ ~~further manual inspection or interaction (Wellmann et al., 2017; de la Varga and Wellmann, 2016). This not only enables us to~~ ~~quantify model uncertainties using Monte Carlo error propagation, but also allows us to incorporate additional information~~ ~~via Bayesian inference to reduce uncertainties.~~Building on this probabilistic perspective, ~~in this work,~~ we propose the use of

20 custom loss functions as a decision-making tool when dealing with uncertain geological models. In many applications, we are interested in some decisive model output value, for example reservoir volume. Given that such a ~~value~~ parameter is the result of a deterministic function of uncertain variables in our model, the ~~value~~ parameter of interest is likewise uncertain and can be represented by a probability distribution attained from stochastic simulations. A loss function can be applied to such a distribution to return a case-specific best estimate to base decision making on.

25 We consider hydrocarbon exploration and production as an exemplary high-risk, high-reward sector, in which good decision making is crucial. However, the described methods are potentially equally applicable to other types of fluid reservoirs (e.g. groundwater geothermal or CO_2-sequestration) and in the raw materials sector. Monte Carlo simulation for reservoir estimation and risk assessment has become common in this sector and is often used in combination with decision trees ~~(see Murtha et al. (1997), Mudford et al. (2000), Wim et al. (2001) and Bratvold and Begg (2010))~~(see Murtha et al., 1997; Mudford et al., 200

30 . However, it seems to us that distributions resulting from probabilistic modeling are mostly only considered to attain best estimates in the form of means. Most likely and extreme outcomes are identified as percentiles, typically P50 (the median), P10 and P90. We believe that this practice does not harness the full potential of such a probabilistic distribution and that much of the inherent information is discarded. Contrary to that, customized loss functions, as a Bayesian method, take into account

the full probability distribution and enable the inclusion of various conditions in the process of finding an optimal estimate. While used in statistical decision theory and other scientific fields, loss functions have, to the best of our knowledge, found no significant application in the ~~field of petroleum exploration and production~~ context of structural geological modeling. Thus, we intend to provide a new perspective with our methodology.

5 To illustrate our approach ~~, we apply it to a simple synthetic 3-D geological model that comprises deformed geological units and~~ of using custom loss functions for decision making, we first illustrate what such customization might look like step-by-step: starting off with a standard symmetrical loss function, incorporating scenario-specific conditions and assumptions, and lastly implementing a factor to represent varying risk affinities of different decision makers. As we assume a ~~normal fault which together form a potential structural hydrocarbon trap. Assuming a~~ petroleum exploration and production ~~case, we define the~~

10 ~~maximum trap volume as our value of interest.~~ decision-making scenario, our parameter of interest should be one that indicates the economic value of a potential hydrocarbon accumulation. In a larger context, including various geological and economic factors such as operational expenditures, this could be the net present value (NPV) of a project. In pre-production stages, original oil-in-place (OOIP) is commonly used for early assessments (Dean, 2007; Morton-Thompson et al., 1993). Decision makers would want to best estimate ~~this volume~~ the relevant parameter of interest to derive recoverable reserves, economic

15 value and subsequently allocate development resources accordingly~~.~~

2 ~~Methods~~

~~Design of our 3-D structural geological model. A 2-D cross section through the middle of the model ($y = 500$ m), perpendicular to the normal fault (parallel to the $x - z$ plane), is shown in (a). A 3-D voxel representation of the model, highlighting the reservoir and seal formations, is visualized in (b). In (d) and (d), the inclusion of parameter uncertainties is presented. Colors~~

20 ~~indicate certain layer bottoms (i.e. boundaries) which are assigned shared z-positional uncertainties (c). All points in the hanging wall are additionally assigned a fault offset uncertainty (d). Thicknesses of the three middle layers are defined by the distances of boundary points (e) and are thus directly dependent on (c).~~

1.1 ~~Computational implementation~~

~~Computationally and numerically, we implement all out methods in a Python programming environment, relying in particular~~

25 ~~on the combination of two crucial open source libraries: (1) GemPy (version 1.0) for implicit geological modeling and (2) PyMC (version 2.3.6) for conducting probabilistic simulations. Gempy is able to generate and visualize complex 3-D structural geological models based on a potential-field interpolation method originally introduced by Lajaunie et al. (1997) and further elaborated by Calcagno et al. (2008). GemPy was specifically developed to enable the embedding of geological modeling in probabilistic machine-learning frameworks, in particular by coupling it with PyMC (de la Varga et al., 2019)~~

30 ~~. PyMC was devised for conducting Bayesian inference and prediction problems in an open-source probabilistic programming environment (Davidson-Pilon, 2015; Salvatier et al., 2016). Different model-fitting techniques are provided in this library, such as various Markov chain Monte Carlo (MCMC) sampling methods. For our purpose we make use of Adaptive Metropolis~~

sampling by Haario et al. (2001) and check for MCMC convergence via a time-series method approach by Geweke et al. (1991) . Components of a statistical model are represented by deterministic functions and stochastic variables in PyMC (Salvatier et al., 2016) . We can thus use the latter to represent uncertain model input parameters and link them to additional data via likelihood functions. Other parameters, such as the value of interest for decision making, can be determined over deterministic functions, as children of parent input parameters. To visually compare the states of geological unit probabilities after conducting stochastic simulations, we consider the normalized frequency of lithologies in every single voxel and visualize the results in probability fields (see Wellmann and Regenauer-Lieb (2012)). Illustration of the process of trap recognition in 2-D, i.e. the conditions that have to be met by a model voxel, to be accepted as belonging to a valid trap. A voxel has to be labeled as part of the target reservoir formation (a) and positioned in the footwall (b). Trap closure is defined by the seal shape and the normal fault (c). Consequently, the maximum trap fill is defined by either the anticlinal spill point (S) or a point of leakage across the fault, depending on juxtapositions with layers underlying (L1) or overlying the seal (L2). The latter is only relevant if the critical Shale Smear Factor is exceeded, as determined over D and T in (d). In this example, assuming sealing of the fault due to clay smearing, the fill horizon is determined by the spill point in (d). Subsequently, only trap section 1 is isolated from the model borders in (d) and can thus be considered a closed trap. Voxels included in this section are counted to calculate the maximum trap volume.

1.1 Synthetic geological model

Our geological example model is designed to represent a potential hydrocarbon trap system. Stratigraphically, it includes one main reservoir unit (sandstone), one main seal unit (shale), an underlying basement and two overlying formations that are assumed to be permeable, so that hydrocarbons could have migrated upwards. Structurally, it is constructed to feature an anticlinal fold that is displaced by a normal fault. All layers are tilted and dip in the opposite direction of the fault plane dip. A potential hydrocarbon trap is thus found in the reservoir rock enclosed by the deformed seal and the normal fault. Using GemPy, we construct the geological model as follows: In principle, it is defined as a cubic block with an extent of 2000 m in x-, y- and z-directions. The basic input data for the interpolation of the geological features is composed of 3-D point coordinates for layer interfaces and fault surfaces, as well as orientation measurements that indicate respective dip directions and angles. From this data, GemPy is able to interpolate surfaces and compute a voxel-based 3-D model (see Fig. 3). which includes the possibility of walking away from a prospect. In this case, the decision maker might refer to an individual geological expert, but also to an exploration company as a whole.

We include uncertainties by assigning them to the z-positions of points which mark layer interfaces in the Once we have set up a loss function customized to this decision problem, we can apply it to probability density functions which represent our knowledge about the true value of the parameter of interest. As mentioned above, such distributions can result from geological modeling in a probabilistic context. To illustrate this, we include a synthetic 3-D space. This is achieved via probability distributions (PyMC stochastic variables) from which deviation values are drawn. These are then added to the original input data z-value. As the z-position is the most sensible parameter for predominantly horizontal layers, we can hereby not only implement uncertainties regarding layer surface positions in depth, but also layer thicknesses, geometrical shapes and degree

of fault offset. Input parameter uncertainties defined by distributions with respective means μ, standard deviations σ and shape factor α.**mu sigma alpha** 0 40 0 0 60 0 0 80 0 0 100 0 0 -150 -2 Such probability distributions can also be allocated as homogeneous sets to point and feature groups which are to share a common degree of uncertainty (see Table 1). We assign the same base uncertainty to groups of points belonging to the same layer bottom surface by referring them to one shared

5   distribution each. Assuming an increase of uncertainty with depth, standard deviations for the shared distributions are increased for deeper formations. Furthermore, uncertainty regarding the magnitude of fault offset is incorporated by adding a skew normal probability distribution that is shared by all layer interface points in the hanging wall. A left skewed normal distribution is chosen to reflect the nature of throw on a normal fault, in particular the slip motion of the hanging wall block. Mainly negative values are returned by this distribution. This way, the offset nature of the normal fault is maintained and inversion to a reverse

10   fault is avoided.

1.1   The value of interest for decision making

We define the trap volume V_t as the central feature of economic interest . For conducting straightforward volumetric calculations, we assume that found closed traps are always filled to spill, i.e. we only consider structural features as controlling mechanisms. This value is of central importance for calculating original oil or gas in place (OOIP/OOIG) and consequently, recoverable

15   reserves. This type of estimation is also the only approach to assess the amount of hydrocarbons in a reservoir before production has started (Dean, 2007; Morton-Thompson et al., 1993).By declaring these connections, we have given our model an economic significance. We can assume that the hydrocarbon trap volume is directly linked to project development decisions, i.e. investment and allocation of resources is represented by bidding on a volume estimate. In the course of this work, we developed a set of algorithms to enable the automatic recognition and calculation of trap volumes in geological models computed by GemPy.

20   The volume is determined on a voxel-counted basis via four conditions illustrated in Fig. 4 and further explained in Appendix A.Following these conditions, we can define four major mechanisms which control the maximum trap volume: (1) the anticlinal spill point of the seal cap, (2) the cross-fault leak point at a juxtaposition of the reservoir formation with itself, (3) leakage due to juxtaposition with overlying layers and SSF failure, and (4) stratigraphical breach of the seal, when its voxels are not continuously connected above the trap. Due to the nature of our model, (3) and (4) will always result in complete trap failure.

25   The trap volume V_t is a result from GemPy's implicit geological model computation. It is an output parameter dependent on deterministic and stochastic input parameters. With every model realization, input uncertainties will respectively propagate to structural geological model as a case study. In this context, we define the structurally determined maximum trap volume V_t as our parameter of interest and indicator for economic value. We generate different probability distributions via stochastic simulations and based on various information scenarios. It is important to note that these are always based on the same primary

30   input parameters. We attain altered states of information by updating the reference case (prior) with secondary information. In doing so, we make sure that the resulting distributions and the volume, which in turn is uncertain. We can thereby evaluate it using probabilistic methods and realizations of loss function applications can be directly compared. These case studies are synthetic and chosen here to exemplify the application of Bayesian decision theory in particular, as explained in the following.

and to show how additional information affects the optimality of decision.

2 Methods

2.1 Bayesian decision theory

We view the statistical analysis of ~~our model from a Bayesian~~ geological models from a probabilistic perspective, which is most importantly characterized by its preservation of uncertainty. Its principles have been presented and discussed extensively in literature (see Jaynes (2003), Box and Tiao (2011), Harney (2013), Gelman et al. (2014) and Davidson-Pilon (2015)). The Bayesian approach is widely seen as intuitive and inherent in the natural human perspective. It regards probability as a measure of belief about a true state of nature. ~~Such beliefs can be assigned to individuals. Thus, different and even contradicting beliefs about a true state of nature might be held by different individuals, based on variations and disparities in the information available to each one individual.~~In this work, the decision problem is one of estimating the true state of our value of interest, which we denote θ. Estimations are based on probability distributions attained from: (1) simple Monte Carlo error propagation, and (2) Bayesian inference using Markov chain Monte Carlo simulation. For (1), we are dealing with a prior probability distribution $p(\theta)$ that results from the deterministic function of the uncertain model input parameters~~

In many cases, decisions are made on the basis of summary parameters such as mean or standard deviation. This approximation works for well defined probability distributions but it may fail when the distribution does not have a defined structure, which is the usual case of distribution generated as a result of a Bayesian inference. ~~For (2), the prior distribution is revalued using Bayesian inference (see Appendix ??) given the presence of additional statistical information y, and using likelihood functions $p(y|\theta)$. Decision making is then based on the resulting posterior probability $p(\theta|y)$. In general, Bayesian inference is about updating a belief and reaching an estimate that is less wrong.~~In this work, we aim to tackle the decision problem associated to probabilistic inferences. By applying Bayesian decision theory concepts, we are capable to transform an arbitrary complex set of distributions onto a more adequate dimension for decision making, usually loss or score.

2.1.1 ~~Likelihoods~~

~~For the application of Bayesian inference, we implement two types of likelihoods:~~

1. ~~**Layer thickness likelihoods:** With every model realization, we extract the z-distance between layer boundary input points at a central $x - y$ position ($x = 1100$ m, $y = 1000$ m) in our input interpolation data. Resulting thicknesses can then be passed on to stochastic functions in which we define thickness likelihoods via normal distributions.~~

2. ~~**SSF likelihood**: SSF values are realized over more complex parameter compositions. We base this likelihood on a normal distribution which we link to the geological model output.~~

~~The inclusion of these likelihood is based on purely hypothetical assumptions and is intended to provide the opportunity to explore the effects different types and scenarios of additional information might have. While the thickness likelihood functions~~

~~are dependent on input parameters directly, the implementation of the SSF likelihood function requires a full computation of the model and extended algorithms of structural analysis.~~

2.1.1 Loss, expected loss and loss functions

Common point estimates, such as the mean and the median of a distribution, usually come with a measure for their accuracy (Berger, 2013). However, it has been argued by Davidson-Pilon (2015) that by using pure accuracy metrics, while this technique is objective, it ignores the original intention of conducting the statistical inference in cases, in which payoffs of decisions are valued more than their accuracies. A more appropriate approach can be seen in the use of loss functions (Davidson-Pilon, 2015).

Loss is a statistical measure of how "bad" an estimate is. Estimate-based ~~decision~~ decisions are also referred to as actions a. Therefore, we also refer to decision makers as actors. Loss is defined as $L(\theta, a)$, so $L(\theta_1, a_1)$ is the actual loss incurred when action a_1 is taken while the true state of nature is θ_1 (Berger, 2013). The magnitude of incurred loss related to an estimate is defined by a loss function, which is a function of the estimate and the true value of the parameter (Wald, 1950; Davidson-Pilon, 2015):

$$L(\theta, \hat{\theta}) = f(\theta, \hat{\theta}). \tag{1}$$

So, how "bad" a current estimate is, depends on the way a loss function weights accuracy errors and returns respective losses. Two standard loss functions are the absolute-error and the squared-error loss function. Both are objective, symmetric, simple to understand and commonly used

~~Davidson-Pilon (2015) and Hennig and Kutlukaya (2007) have proposed that it might be useful to move on from standard objective loss functions to the design of customized loss functions that specifically reflect an individual's (i.e. the decision maker's) objectives, preferences and outcomes. Hennig and Kutlukaya (2007) argue that choosing and designing a loss function involves the translation of informal aims and interests into mathematical terms. This process naturally implies the integration of subjective decisions and subjective elements. According to them, this is not necessarily unfavorable or less objective, as it may better reflect an expert's perspective on the situation.~~ ~~Standard symmetric loss functions can easily be adapted to be asymmetric, for example by weighting errors on the negative side stronger than those on the positive side. Preference over estimates larger than the true value (i.e. overestimation) is thus incorporated in an uncomplicated way..~~ ~~Much more complicated designs of loss functions are possible, depending on purpose, objective and application.~~The presence of uncertainty during decision making implies that the true parameter value is unknown and thus the truly incurred loss $L(\theta, a)$ cannot be known at the time of making the decision. The Bayesian perspective considers unknown parameters as random variables and samples that are drawn from ~~the posterior~~ a probability distribution as possible realizations of the unknown parameter, i.e. all possible true values are represented by this distribution.

~~Given a posterior distribution $p(\theta|y)$~~Under uncertainty, the expected loss of choosing an estimate $\hat{\theta}$ over the true parameter value θ ~~(after evidence y has been observed)~~ is defined by (Davidson-Pilon, 2015):

$$l(\hat{\theta}) = E_\theta[L(\theta, \hat{\theta})]. \tag{2}$$

The expectation symbol E is subscripted with θ, by which it is indicated that θ is the respective unknown variable. This expected loss l is also referred to as the Bayes risk of estimate $\hat{\theta}$ (Berger, 2013; Davidson-Pilon, 2015).

By the Law of Large Numbers, the expected loss of $\hat{\theta}$ can be approximated drawing a large sample size N from the posterior distribution, respectively applying a loss function L and averaging over the number of samples (Davidson-Pilon, 2015):

$$\frac{1}{N}\sum_{i=1}^{N} L(\theta_i, \hat{\theta}) \approx E_\theta[L(\theta, \hat{\theta})] = l(\hat{\theta}). \tag{3}$$

5

Hereby, we can approximate the expected loss l for every possible estimate $\hat{\theta}$ (every decision we can make) according to the loss function in use. Minimization of a loss function returns a point estimate known as Bayes action or Bayesian estimator, which is the decision with the least expected loss according to the loss function, and the decision we are interested in in this work (Berger, 2013; Moyé, 2006).

10 **2.2** ~~Customization of our case-specific loss function~~

~~Assigning an economic notion to our model and assuming the case of an actor or decision makerin any field naturally necessitates the consideration of preferences , interests and the overall subjective perspective such an individual or for example a company might have. Further constraints and conditions can also be specific to the field, industry or generally to the problem at hand~~

15 **2.1.1** **Customization of our case-specific loss function**

Davidson-Pilon (2015) and Hennig and Kutlukaya (2007) have proposed that it might be useful to move on from standard objective loss functions to the design of customized loss functions that specifically reflect an individual's (i.e. the decision maker's) objectives and preferences regarding outcomes. Especially as we assign an economic notion to geological models and related estimation problems, we argue that it is necessary to consider the subjective perspectives of involved decision makers,
20 say for example exploration and production companies. Consequently, the design of a more specific non-standard and possibly asymmetric loss function might be required. One that includes subjective aspects and ~~difference~~ differences in weighting of particular risks, arising from ~~an actor~~a decision maker's inherent preferences and the environment in which ~~the~~ this actor has to make a decision. In the face of several uncertain parameters, ~~a perfectly true estimate~~which is a given in complex geological models, a perfect estimate, a perfect decision, is virtually unattainable. However, an attempt can be made to design a custom
25 loss function that returns a Bayesian estimator involving the least bad consequences for ~~an individual~~ a decision maker in a specific environment (Davidson-Pilon, 2015; Hennig and Kutlukaya, 2007).

Hennig and Kutlukaya (2007) argue that choosing and designing a loss function involves the translation of informal aims and interests into mathematical terms. This process naturally implies the integration of subjective elements. According to them, this is not necessarily unfavorable or less objective, as it may better reflect an expert's perspective on the situation.
30 Standard symmetric loss functions can easily be adapted to be asymmetric, for example by weighting errors to the negative side stronger than those to the positive side. Preference over estimates larger than the true value, i.e. overestimation, is thus

incorporated in an uncomplicated way. Much more complicated designs of loss functions are possible, depending on purpose, objective and application. We will describe potential design options in the following.

For our example ~~ease of trap volume estimation~~of estimating the economic value of a hydrocarbon prospect, which is represented by the maximum trap volume V_t, we develop a custom loss function in five steps. Ideally, ~~an actor~~a decision maker would like to know the exact ~~trap volume~~true value, so that resources can be allocated appropriately in order to acquire economic gains by developing a project and producing from a reservoir. This conscious and irrevocable allocation of resources is the decision to be made or action to be taken (Bratvold and Begg, 2010). Thus, we treat estimating as equivalent to making a decision. Deviations from the unknown true value in the form of over- and underestimation bring about an error and loss accordingly. ~~In steps I - IV we make assumptions about the significance of such deviations and how they differently contribute to risks in the general decision making environment.~~

It can be assumed that several ~~actors~~decision makers in one such environment or sector may have the same general loss function but different affinities concerning ~~the~~risks. This might be based for example on different psychological factors or economic philosophies followed by companies. It might also be based on budgets and options such actors have available. An intuitive example is the comparison of a small and a large company. A ~~certain false estimate or error~~false estimate and wrong decision might have a significantly stronger impact on a company which has a generally lower market share and only few projects, than on a larger company which might possess a higher financial flexibility and for which one project is only one of many development options in a ~~portfolio. We therefore~~wide portfolio.

In steps I - IV we make assumptions about the significance of such deviations and how they differently contribute to expected losses in the general decision-making environment and introduce the concept of varying risk affinities in the final step V. ~~Plotting of expected loss realizations after including the risk affinity factor r in the loss function (see Eq. 8) for actors with risk affinities ranging from risk-friendly ($r = 0.5$ and 0.75), over risk-neutral ($r = 1$), to risk-averse ($r = 1.25$ and $r = 1.5$). Dots mark the respective positions of minimizing actions. We applied the function to samples from a normal distribution ($\mu = 0, \sigma = 500$) that represents the probability of a hypothetical score which is to be estimated. Jumps at zero are caused by the implementation of fatal over- and underestimation. Taking a closer look at the realizations for $r = 1.5$ and $r = 0.5$, we can recognize how they expect different losses and come to a different optimal decision given the same information. As in this case, positive and negative score values are equally likely, and overestimation errors are weighted stronger, only the most risk-friendly actor will bid on a positive estimate. Also, expected losses are lower for more risk-friendly decision makers.~~

– **Step I - Choosing a standard loss function as starting point**: In our case, we assume that investments increase linearly with linear growth in the value of the ~~resource (i. e. the volume).~~prospect. For this reason, we choose the symmetric absolute-error loss function as a basis for further customization steps:

$$L(\theta, \hat{\theta}) = |\theta - \hat{\theta}|. \tag{4}$$

– **Step II - Simple overestimation**: Considering the development of a hydrocarbon reservoir, it can be assumed that over-investing is worse than under-investing. Overestimating the size of an accumulation might for example lead to the installation of equipment or ~~facilites~~facilities that are actually redundant or unnecessary. This would come with

additional unrecoverable expenditures. Consequences from underestimating $(0 < \hat{\theta} < \theta)$, however, may presumably be easier to resolve. Additional equipment can often be installed later on. Hence, simple overestimation $(0 < \theta < \hat{\theta})$ is weighted stronger in this loss function by multiplying the error with an overestimation factor a:

$$L(\theta, \hat{\theta}) = |(\theta - \hat{\theta})|\, a. \tag{5}$$

5   – **Step III - ~~Fatal~~ Critical overestimation**: The worst case for any project would be that its development is set into motion, expecting a gain, only to discover later that the value in the reservoir does not cover the costs of realizing the project, resulting in an overall loss. A petroleum system might also turn out to be a complete failure, containing no value $(V_t = 0$ in our 3-D case study) at all, although the actor's estimate indicated the opposite. Here, we refer to this as ~~fatal~~ critical overestimation. A positive value is estimated, but the true value is zero or negative $(\theta \leq 0 < \hat{\theta})$. This is worse than simple

10   overestimation, where both values are positive and a net gain is still achieved, which is only smaller then the best possible gain of expecting the true value. ~~Fatal~~ Critical overestimation is included in the loss function by using another weighting factor b that replaces a:

$$L(\theta, \hat{\theta}) = |(\theta - \hat{\theta})|\, b. \tag{6}$$

In other words: With $b = 2$, ~~fatal~~ critical overestimation is twice as bad as simple ~~underestimation~~overestimation.

15   – **Step IV - ~~Fatal~~ Critical underestimation**: We also derive ~~fatal~~ critical underestimation from the idea of estimating zero (or a negative value), when the true value is actually positive $(\hat{\theta} \leq 0 < \theta)$. This is assumed to be worse than simple overestimation, but clearly better than ~~fatal~~ critical overestimation. No already owned resources are wasted, it is only the potential value that is lost, i.e. opportunity costs that arise from completely discarding a profitable project. ~~Fatal~~ Critical underestimation is weighted using a third factor c:

20
$$L(\theta, \hat{\theta}) = |(\theta - \hat{\theta})|\, c. \tag{7}$$

   – **Step V - Including different risk affinities**: We now further adapt the loss function to consider varying risk affinities of different actors. We follow the approach of Davidson-Pilon (2015), who implemented different risk affinities by simply introducing a variable risk factor. Using different values for this factor, we can represented how comfortable an individual is with being wrong and furthermore which "side of wrong" is preferred by that decision maker (Davidson-

25   Pilon, 2015). In our case, bidding lower is considered the cautious, risk-averse option, as smaller losses can be expected from underestimating. Guessing higher is deemed riskier, as losses from overestimation are greater. However, bidding correctly on a higher value, will also return a greater gain. It is assumed that risk-friendly actors care less about ~~fatal~~ critical underestimation, i.e. they will rather develop a project than discard it. In our finalized loss function, we simply

include these considerations via a risk affinity factor *r* which alters the incurred losses respectively:

$$L(\theta, \hat{\theta}) = \begin{cases} |\theta - \hat{\theta}|\, r^{-0.5}, & \text{for } 0 < \hat{\theta} < \theta \\ |\theta - \hat{\theta}|\, a\, r, & \text{for } 0 < \theta < \hat{\theta} \\ |\theta - \hat{\theta}|\, b\, r, & \text{for } \theta \leq 0 < \hat{\theta} \\ |\theta - \hat{\theta}|\, c\, r^{-0.5}, & \text{for } \hat{\theta} \leq 0 < \theta \end{cases}, \text{ with } a, b, c, r \in \mathbb{Q}. \tag{8}$$

This equation shows that the final custom loss function is in essence a composite of four different functions for the over- and underestimation cases explained in steps II - IV. It is important to note that the weighting factors *a, b* and *c* can take basically

5   any numerical values but should be chosen in a way that they appropriately represent the framework conditions of the problem. Here, we assume that simple overestimation is 25 % (*a* = 1.25), ~~fatal~~ critical overestimation 100 % (*b* = 2) and ~~fatal~~ critical underestimation 50 % (*c* = 1.5) worse than simple underestimation.~~An example plot of actual incurred losses via this loss function can be found in Appendix ??~~

According to Eq. 8, the risk-neutral loss function is returned for *r* = 1, as no re-weighting takes place. For *r* < 1, the weight

10   on overestimating (*a, b*) is reduced and increased for ~~fatal~~ critical underestimation (*c*), as well as normal underestimation. This represents a risk-friendlier actor that is willing to bid on a higher estimate to attain a greater gain. For *r* > 1, the overestimation weight (*a, b*) is increased in the loss function, underestimation and ~~fatal~~ critical underestimation weight (*c*) are decreased and respectively more risk-averse actors are prompted to bid on lower estimates. Since risk neutrality is expressed by *r* = 1, we consider values 0 < *r* < 2 to be the most appropriate choices to represent both sides of risk affinity equally.~~Accordingly~~

15   ~~different loss function realizations are~~

In Fig. 1, we illustrate different aspects and steps of adapting and applying the custom loss function. For these simple examples, we assume that the economic value of our reservoir is represented by an abstract score parameter. Fig. 1 (a) depicts the plotting of the absolute-error loss function (customization step I) applied to a normal distribution. It can be seen, that for this standard symmetrical function, the minimal point of expected losses and Bayes action corresponds the median (and

20   mean for this symmetric distribution). Fig. 1 (b) summarizes customization steps II - IV and visualizes how four different functions for four cases of under- and overestimation are summed up to one combined loss function that comprises all of the assumptions made for the decision-making environment. A jump of expected losses on the negative side of possible estimates can be attributed to the way we defined the function for critical underestimation as dependent on zero.

In Fig. 1 (c), the risk factor of step V was implemented without steps II - IV, i.e. only for the standard absolute-error loss

25   function. It can be seen that risk-averse and risk-friendly decision makers are represented by respectively different realizations of expected losses based on one and the same normal distribution: The narrow shape of the risk-friendly function represents improved confidence in the decision, while the increased expected loss (Bayes risk) of the minimum indicates that this comes along with the acceptance of a higher risk. Inversely, the flat shape of the risk-averse function can be seen as reduced confidence in the decision. There is less of a difference in taking a different decision, than for the risk-friendly actor. At the same time,

30   expected loss of the minimum, and thus the accepted risk, is lower. However, although they differ in expected losses, both decision makers share the same individual best estimate, since the loss function in itself is still symmetric. This changes in (d),
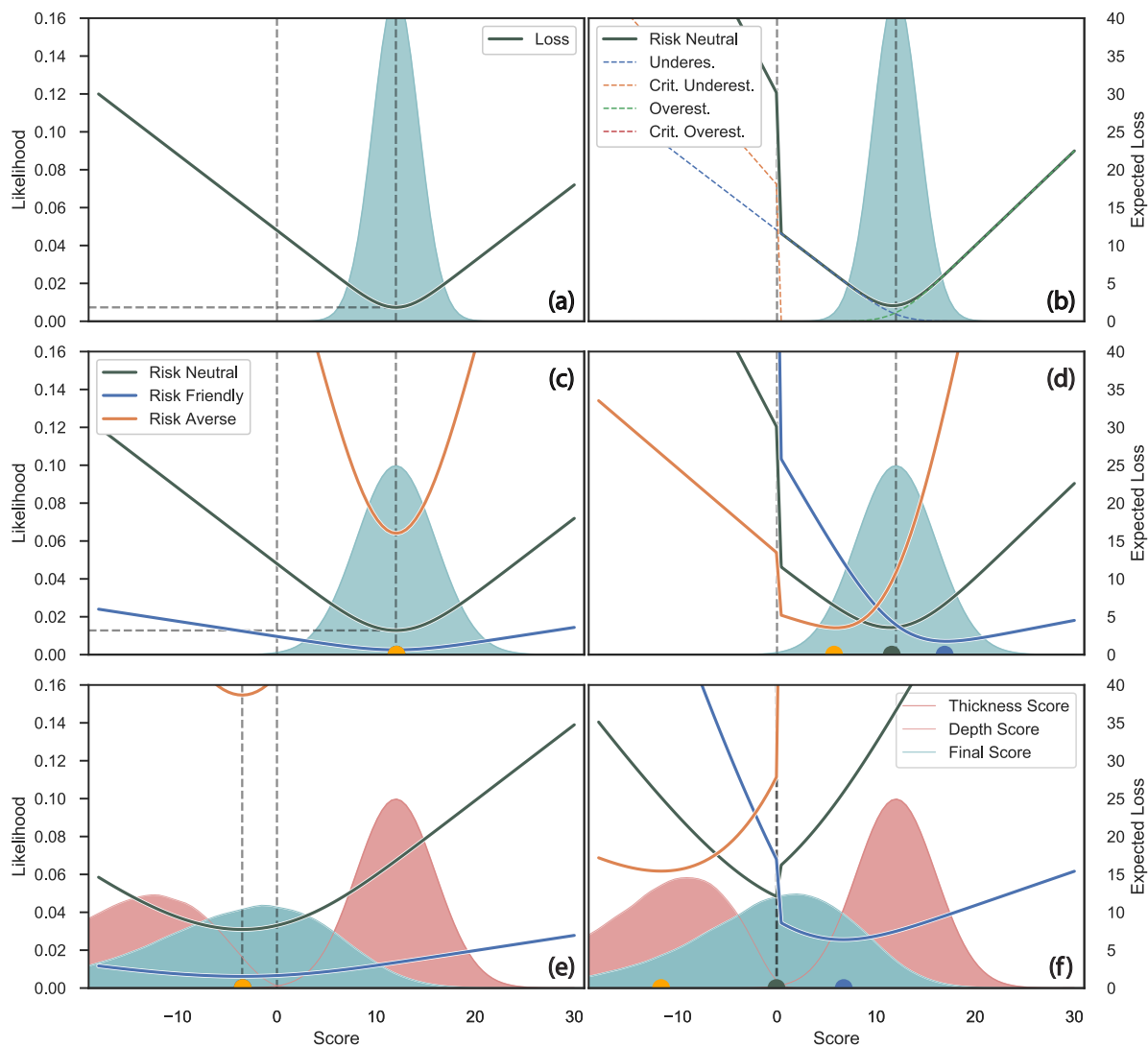
11

Figure 1. Illustration of different steps and aspects of our loss function customization. Functions are applied to an abstract score as the parameter of interest.

in which all customization steps were applied. Here, the risk factor re-weights the influence of the sub-functions shown in (b). Under- and overestimation cases are accordingly enhanced or reduced in impact, so that the resulting loss function becomes asymmetric and minima are found at different score estimates, given the same underlying information.

In Fig. 1 (e) and (f), the functions from (c) and (d) are applied on a score distribution resulting from the combination of two other uncertain parameters: Reservoir thickness and depth. This can be seen as an extremely simplified 1-D model with only two inputs that define one output as a parameter of interest, the final score. In this case, thickness is seen as the potential positive value in our reservoir, as it provides space for hydrocarbons to accumulate. Depth is subtracted from this, as it implies a cost of drilling. Thus, the final score is a very essential representation of the economic value given the information available. The respective final distribution is slightly skewed. Fig. 1 (e) depicts the respective application of the same functions used in (c). Symmetric, but including risk-affinity. Overall effects are the same as in (c). It can be additionally observed that since the underlying distribution is now asymmetric, all expected loss minima are found on the median estimate, lower than the mean. In (f), the complete custom loss function was applied as in (d). Based on the uncertain information about the final score, the three differently risk-affine loss functions plot differently, with minima in the negative space, at zero and in the positive space. This illustrates how the risk-averse decision maker tends to expect a possible negative outcome, while the risk-friendly actor bids on a positive value. This could be seen respectively as the decision to abandon versus the decision to invest in a prospect.

For better understanding of how our finalized custom loss function determines the incurrence of loss, actual losses for three fixed true values and risk-neutrality ($r = 1$) are plotted in Fig. ~~??~~ 2.
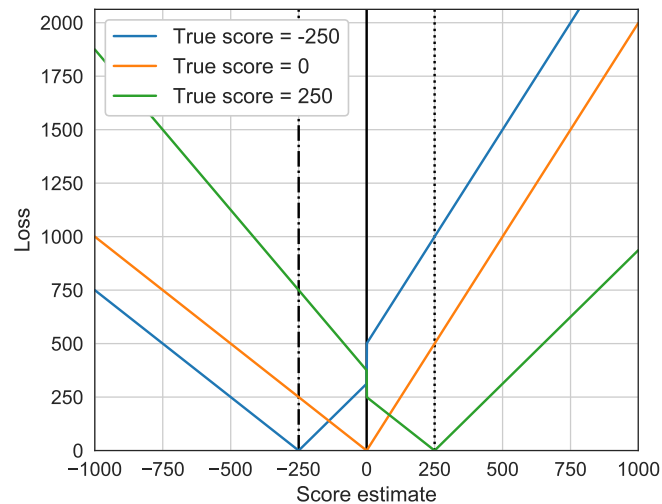

Figure 2. Loss based on the risk-neutral custom loss function (Eq. 8) for determined true scores of -250, 0 and 250. This plot is meant to clarify the way real losses are incurred for each estimate, relative to a given true value. The expected loss, as seen in Fig. 1 is acquired by arithmetically averaging over all deterministic loss realizations based on the score probability distribution by using Eq. 3.

It has to be emphasized that this is just one possible proposal for loss function customization. There exists not one perfect design for such a case (Hennig and Kutlukaya, 2007). Slight to strong changes can already be implemented by simply varying the values of the weighting factors *a, b* and *c*. Fundamentally different loss functions can also be based on a significantly different mathematical structure. ~~Loss~~ As loss functions are customized regarding the problem environment and according to

5   the subjective needs and objectives of the decision maker~~(Davidson-Pilon, 2015; Hennig and Kutlukaya, 2007). Thus~~, they are mostly defined by the actor expressing his perspective (Davidson-Pilon, 2015; Hennig and Kutlukaya, 2007). Changes in the individual's perception and attitude might lead to further customization needs at a future point in time, as was reported by Hennig and Kutlukaya (2007).

2.2   Case study: Synthetic 3-D structural geological model

10   Next we want to show that this loss function approach is not only applicable to simple probability distributions but is an equally useful tool to estimate the true value of a parameter of interest resulting from more complex geological models which encompass numerous uncertain input parameters. As a case study, we consider now a synthetic 3-D structural geological model which is placed in a probabilistic framework.

2.2.1   Computational implementation

15   Computationally, we implement all of our methods in a Python programming environment, relying in particular on the combination of two open-source libraries: (1) GemPy (version 1.0) for implicit geological modeling and (2) PyMC (version 2.3.6) for conducting probabilistic simulations.

Gempy is able to generate and visualize complex 3-D structural geological models based on a potential-field interpolation method originally introduced by Lajaunie et al. (1997) and further elaborated by Calcagno et al. (2008). GemPy was specifically

20   developed to enable the embedding of geological modeling in probabilistic machine-learning frameworks, in particular by coupling it with PyMC (de la Varga et al., 2019).

PyMC was devised for conducting Bayesian inference and prediction problems in an open-source probabilistic programming environment (Davidson-Pilon, 2015; Salvatier et al., 2016). Different model-fitting techniques are provided in this library, such as various Markov chain Monte Carlo (MCMC) sampling methods. For our purpose we make use of Adaptive Metropolis

25   sampling by Haario et al. (2001) and check MCMC convergence via a time-series method approach by Geweke et al. (1991). Components of a statistical model are represented by deterministic functions and stochastic variables in PyMC (Salvatier et al., 2016). We can thus use the latter to represent uncertain model input parameters and link them to additional data via likelihood functions. Other parameters, such as the value of interest for decision making, can be determined over deterministic functions, as children of parent input parameters.

30   To visually compare the states of geological unit probabilities after conducting stochastic simulations, we consider the normalized frequency of lithologies in every single voxel and visualize the results in probability fields (see Wellmann and Regenauer-Lieb (2).

14

Table 1. Input parameter uncertainties defined by distributions with respective means μ, standard deviations σ and shape factor α.

	mu	sigma	alpha
Overlying	0	40	0
Sandstone 2	0	60	0
Seal	0	80	0
Reservoir	0	100	0
Fault offset	0	-150	-2

2.2.2 Design of the 3-D structural geological model

Our geological example model is designed to represent a potential hydrocarbon trap system. Stratigraphically, it includes one main reservoir unit (sandstone), one main seal unit (shale), an underlying basement and two overlying formations that are assumed to be permeable, so that hydrocarbons could have migrated upwards. Structurally, it is constructed to feature an anticlinal fold that is displaced by a normal fault. All layers are tilted and dip in the opposite direction of the fault plane dip. A potential hydrocarbon trap is thus found in the reservoir rock enclosed by the deformed seal and the normal fault.

Using GemPy, we construct the geological model as follows: In principle, it is defined as a cubic block with an extent of 2000 m in x-, y- and z-directions. The basic input data for the interpolation of the geological features is composed of 3-D point coordinates for layer interfaces and fault surfaces, as well as orientation measurements that indicate respective dip directions and angles. From this data, GemPy is able to interpolate surfaces and compute a voxel-based 3-D model (see Fig. 3).

We include uncertainties by assigning them to the z-positions of points which mark layer interfaces in the 3-D space. This is achieved via probability distributions (PyMC stochastic variables) from which error values are drawn. These are then added to the original input data z-value. As the z-position is the most sensible parameter for predominantly horizontal layers, we can hereby not only implement uncertainties regarding layer surface positions in depth, but also layer thicknesses, geometrical shapes and degree of fault offset.

Such probability distributions can also be allocated as homogeneous sets to point and feature groups which are to share a common degree of uncertainty (see Table 1). We assign the same base uncertainty to groups of points belonging to the same layer bottom surface by referring them to one shared distribution each. Assuming an increase of uncertainty with depth, standard deviations for the shared distributions are increased for deeper formations. Furthermore, uncertainty regarding the magnitude of fault offset is incorporated by adding a skewed normal probability distribution that is shared by all layer interface points in the hanging wall. A left skewed normal distribution is chosen to reflect the nature of throw on a normal fault, in particular the slip motion of the hanging wall block. Skew to the negative side ensures that the offset nature of the normal fault is maintained and inversion to a reverse fault is avoided.

This model was designed for the primary purpose of testing our loss function method. All features, uncertainties and parameter relations were implemented in a way that they result in model variability and complexity that is adequate and
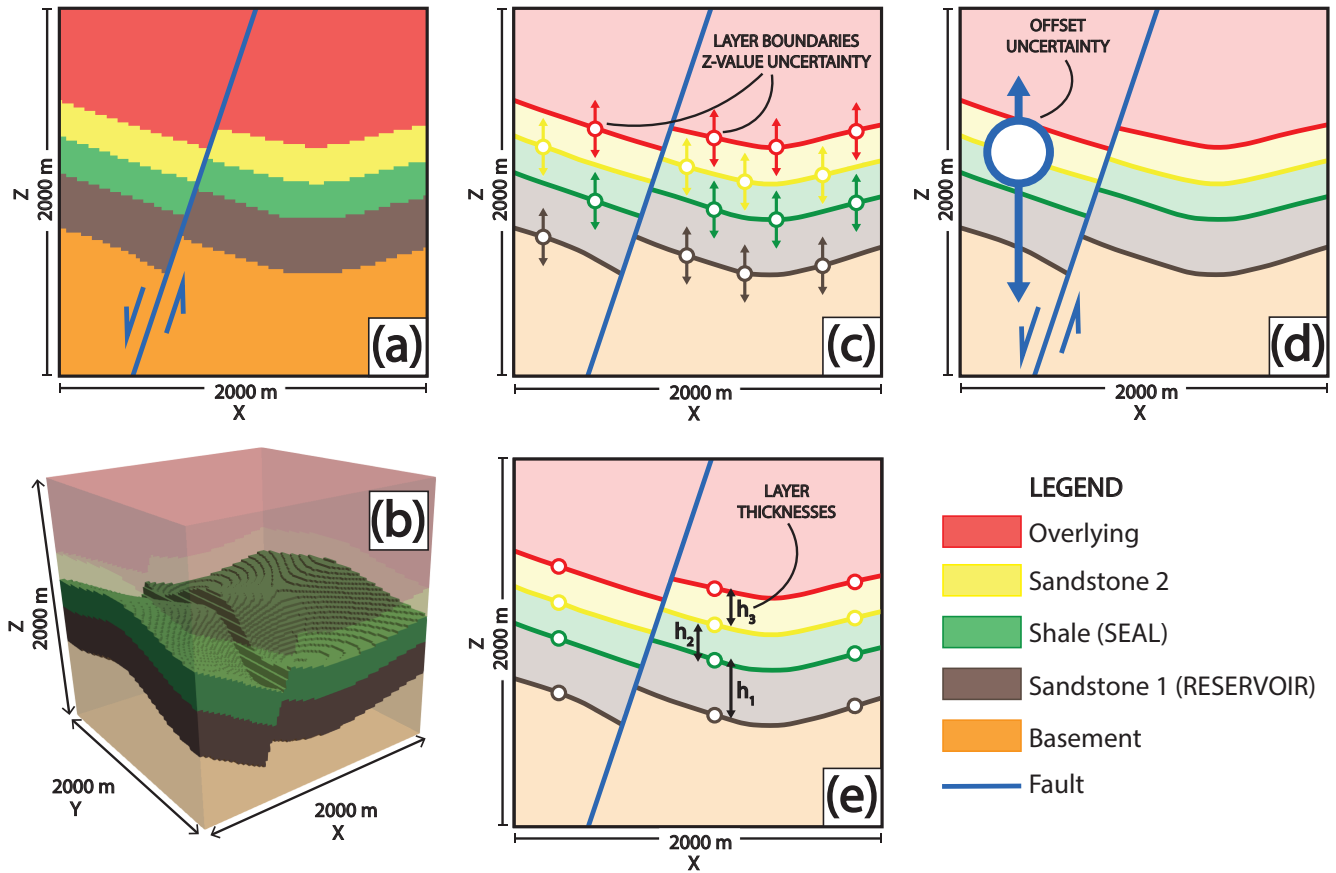
Figure 3. Design of the 3-D structural geological model. A 2-D cross section through the middle of the model ($y = 500$ m), perpendicular to the normal fault (parallel to the $x - z$ plane), is shown in (a). A 3-D voxel representation of the model, highlighting the reservoir and seal formations, is visualized in (b). In (d) and (d), the inclusion of parameter uncertainties is presented. Colors indicate certain layer bottoms (i.e. boundaries) which are assigned shared z-positional uncertainties (c). All points in the hanging wall are additionally assigned a fault offset uncertainty (d). Thicknesses of the three middle layers are defined by the distances of boundary points (e) and are thus directly dependent on (c).

significant to the decision problem in this work. The model is not aimed at representing a completely plausible or realistic geological setting.

2.2.3 V_t as the parameter of interest

Given full 3-D representation of geological structures, we can now define the trap volume V_t as the parameter of interest, a feature that indicates the economic value of the reservoir in this case. For conducting straightforward volumetric calculations,

5

16

we assume that closed traps are always filled to spill, i.e. we only consider structural features as controlling mechanisms and disregard other parameters in the OOIP equation (Eq.9).

We argue that V_t can be inserted for the hydrocarbon-filled rock volume $A * h$ in the OOIP equation (Dean, 2007; Morton-Thompson et al.):

5 $$\text{OOIP} = A * h * \phi * (1 - S_W) * 1/\text{FvF}, \qquad (9)$$

Where OOIP is returned in m^3, A is the drainage area, h the net pay thickness, ϕ the porosity, S_W the water saturation, and FvF the formation volume factor which determines the shrinkage of the oil volume brought to the surface.

By declaring these connections, we have given our model an economic significance. We can assume that the hydrocarbon trap volume is directly linked to project development decisions, i.e. investment and allocation of resources is represented by 10 bidding on a volume estimate.

In the course of this work, we developed a set of algorithms to enable the automatic recognition and calculation of trap volumes in geological models computed by GemPy. The volume is determined on a voxel-counted basis via four conditions illustrated in Fig. 4 and further explained in Appendix A.

Following these conditions, we can define four major mechanisms which control the maximum trap volume: (1) the anticlinal 15 spill point of the seal cap, (2) the cross-fault leak point at a juxtaposition of the reservoir formation with itself, (3) leakage due to juxtaposition with overlying layers and cross-fault seal breach (failure related to the Shale Smear Factor (SSF)), and (4) stratigraphical breach of the seal, when its voxels are not continuously connected above the trap. Due to the nature of our model, (3) and (4) will always result in complete trap failure. The occurrence of these trap control mechanisms can be tracked throughout stochastic simulations of the model.

20 **2.2.4 Generating different probability distributions for V_t**

The trap volume V_t is a result from GemPy's implicit geological model computation. It is an output parameter dependent on deterministic and stochastic input parameters. Conducting stochastic simulations, input uncertainties will respectively propagate to V_t, which is thereby represented by a respective probability distribution that our custom loss function can be applied to. Using simple Monte Carlo error propagation, with every iteration, we draw sample values for our uncertain primary 25 model input parameters defined in 2.2.2 and thus, with every iteration, we create one possible realization of our geological model, which in turn comes with one possible outcome for V_t. Results from all iterations together approximate the probability distribution for V_t according to the input parameters.

Furthermore, we consider the possibility to update our model by adding additional secondary information via Bayesian inference. We do this by introducing likelihood functions which constrain our primary parameters. We have to note that these 30 inputs remain unchanged, however, their prior probability distributions are revalued given the additional statistical information. We achieve this by conducting Markov chain Monte Carlo (MCMC) simulations. Decision making is then based on the resulting posterior probability. Using different likelihood functions, we can create generate different posterior probability distributions
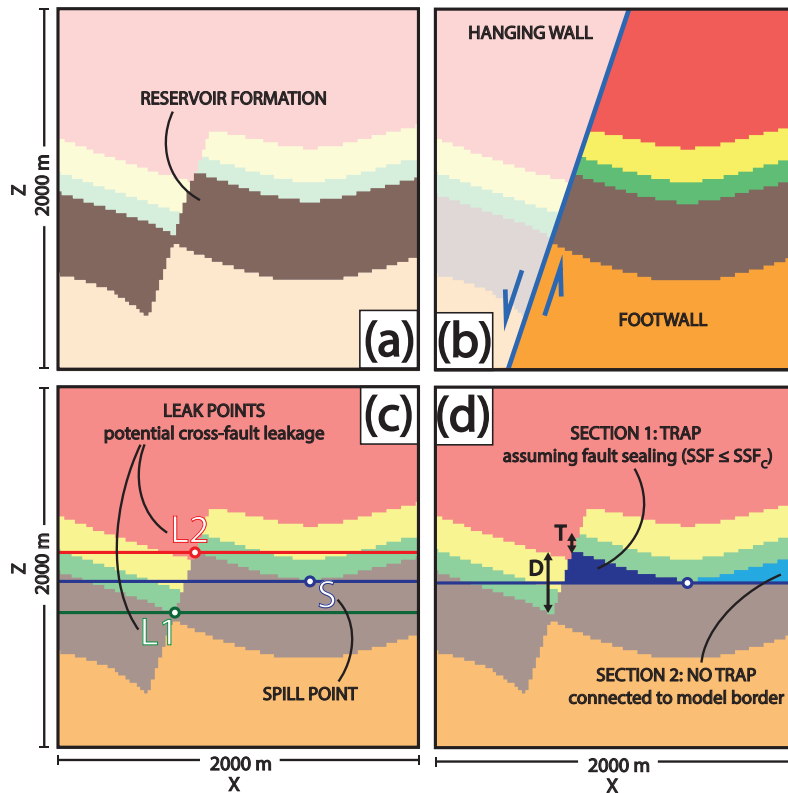
Figure 4. Illustration of the process of trap recognition in 2-D, i.e. the conditions that have to be met by a model voxel, to be accepted as belonging to a valid trap. A voxel has to be labeled as part of the target reservoir formation (a) and positioned in the footwall (b). Trap closure is defined by the seal shape and the normal fault (c). Consequently, the maximum trap fill is defined by either the anticlinal spill point (S) or a point of leakage across the fault, depending on juxtapositions with layers underlying (L1) or overlying the seal (L2). The latter is only relevant if the critical Shale Smear Factor is exceeded, as determined over D and T in (d). In this example, assuming sealing of the fault due to clay smearing, the fill horizon is determined by the spill point in (d). Subsequently, only trap section 1 is isolated from the model borders in (d) and can thus be considered a closed trap. Voxels included in this section are counted to calculate the maximum trap volume.

for V_t, which represent different information scenarios. Since we use Bayesian inference to revalue our original prior inputs, we can compare all outcomes and realizations of our custom loss function.

For the application of Bayesian inference, we implement two types of likelihoods:

1. **Layer thickness likelihoods:** With every model realization, we extract the z-distance between layer boundary input points at a central $x - y$ position ($x = 1100$ m, $y = 1000$ m) in our input interpolation data. Resulting thicknesses can then be passed on to stochastic functions in which we define thickness likelihoods via normal distributions.

2. **Shale Smear Factor (SSF) likelihood**: SSF values are realized over more complex parameter compositions. We base this likelihood on a normal distribution which we link to the geological model output.

The inclusion of these likelihood is based on purely hypothetical assumptions and is intended to provide the opportunity to explore the effects different types and scenarios of additional information might have. While the thickness likelihood functions are dependent on input parameters directly, the implementation of the SSF likelihood function requires a full computation of the model and extended algorithms of structural analysis.

Although Bayesian inference was utilized in this case study, it served primarily for the generation of these different but comparable distributions to base our decision making, i.e. the application of our custom loss function, on. For additional information on how implicit geological modeling can be embedded in a Bayesian framework, and how this can be used to reduce uncertainty, we refer to the work by Wellmann et al. (2010b), de la Varga and Wellmann (2016), de la Varga et al. (2019) and Wellmann et al. (2017).

3 Results

We applied our custom loss function to various different ~~volume~~ V_t probability distributions resulting from stochastic simulations. First, reference results were created using only ~~priors~~ primary inputs (priors) and simple Monte Carlo error propagation (10,000 sampling iterations, Scenario 1). Then we devised several scenarios of additional information and included these via likelihoods and Bayesian inference. For this, 10,000 MCMC sampling steps were conducted, with an additional burn-in phase of 1000 iterations. The prior parameter uncertainties were chosen to be identical for all simulations (see Table 1). Results of convergence diagnostics can be found in Appendix C.

We present the following information scenarios:

1. **Prior-only** model

2. Introducing **seal thickness likelihoods**

   (a) Likely thick seal

   (b) Likely thin seal

3. Introducing **reservoir thickness likelihoods**

19

Table 2. Normal distribution mean (μ) and standard deviations (σ) for the likelihoods implemented in the different scenarios.

	Seal thickness		Reservoir thickness		SSF	
	μ [m]	σ [m]	μ [m]	σ [m]	μ	σ
Scenario 1	-	-	-	-	-	-
Scenario 2a	300	30	-	-	-	-
Scenario 2b	50	30	-	-	-	-
Scenario 3a	350	30	-	-	-	-
Scenario 3b	300	30	300	30	-	-
Scenario 4a	-	-	-	-	5.1	0.3
Scenario 4b	300	30	-	-	2	0.3

    (a) Likely thick reservoir

    (b) Likely thick reservoir and thick seal

4. Introducing **SSF likelihoods**

    (a) SSF likely near its critical value

    (b) Likely reliable SSF and thick seal

The implemented likelihoods are listed in Table 2.

For the comparison of results, we consider in particular the following measures: (1) probability field visualization, (2) occurrence of trap control mechanisms, (3) resulting trap volume distributions, and (4) consequent realization of expected losses and related decisions.

3.1 Prior-only model (Scenario 1)

Probability field visualization illustrates well how the prior uncertainty is based on normal distributions (see Fig. B2). Trap control mechanisms are listed in Table B2. For this prior-only scenario, all four relevant mechanisms ~~occurred~~occur. The dominant factor is the anticlinal spill point with a 51.5 % rate of occurrence. It is followed by cross-fault leakage to the reservoir (25 %) and other permeable formations (12 %). Stratigraphical breaches of the seal were registered to be decisive in about 11 % of iterations. In only 0.5 % of iterations, the algorithm failed to recognize a mechanism, i.e. correct model realization failed.

Maximum trap volumes were calculated for each model iteration and plotted as a probability distribution in Fig. 5. In general, a wide range of volumes is possible, from zero to more than 3 million m^3. However, we can recognize a bimodal tendency: Low ~~but positive~~ volumes are less probable than significantly high volumes or complete failure ($V_t = 0$).

Consequently, applying our custom loss function to this distribution resulted in widely separated minimizing estimators for the differently risk-affine actors (see Fig. 5). Only the risk-friendliest estimates are found within the described highly positive

mode of the distribution. Risk-averse individuals bid on significantly lower estimates or even zero. The risk-neutral decision is found between both modes and presents the highest expected loss. Expected losses decrease towards the extreme decisions and closer to the modes.
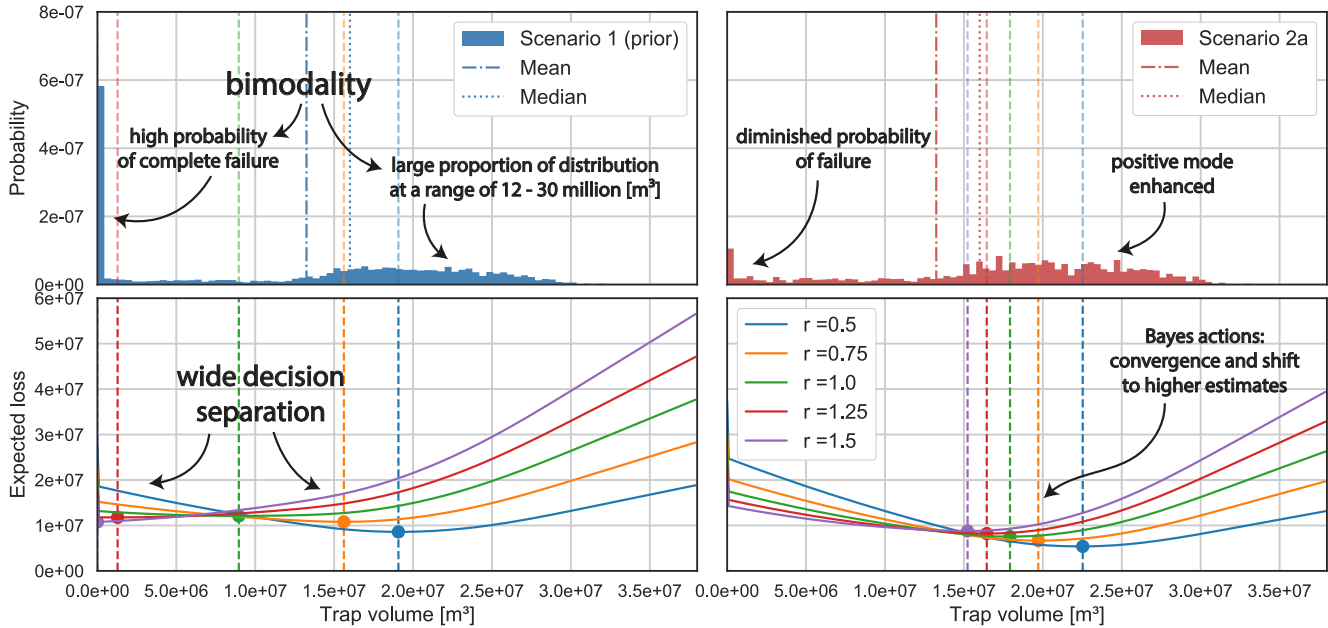

Figure 5. Trap volume distribution and resulting loss function realizations for Scenario 1 (prior) and Scenario 2a, where we introduced the likelihood of a thick seal. Comparing both, we can observe how the additional information reduced the bimodality in the posterior distribution (2a), particularly by reducing the probability of complete failure and enhancing positive probabilities. Consequently, Bayes actions converged and expected losses were reduced.

3.2 Introducing seal thickness likelihoods (Scenarios 2a and 2b)

We considered two scenarios of thickness likelihoods: The seal being (Scenario 2a) likely very thick, or (Scenario 2b) likely very thin (see Table 2).

In Scenario 2a, probability visualization illustrates that the presence of a thick seal is very probable (see Fig. B2). For Scenario 2b, the presence of a reliable seal is questionable.

A high likelihood of a reliable seal cap (2a) significantly reduced the probability of trap failure, while enhancing the mode of highly positive outcomes (see Fig. 5). This coincides with the predominance of the anticlinal spill point (63 %) and the leak point to the same reservoir (36%) as control mechanisms. The occurrence of other mechanisms was negligible (see Table B2). Inversely, a likely thin seal (2b) virtually eliminated the positive mode and focused almost the whole distribution on complete failure. Accordingly, seal-breach related control mechanisms gained importance (65.5 % occurrence rate for stratigraphical seal breach).

21

In both scenarios, Bayes actions shifted towards the respectively emphasized modes. This came with overall convergence of decisions and reduction of expected losses. In Scenario 2a, all decision makers bid on a positive outcome. Risk-averse individuals experienced the strongest shift, but also present the highest expected losses. In Scenario 2b, all individuals decide to not allocate resources. Even the risk-friendliest actor moved to a zero estimate, where the most risk-averse bid had already been placed in the prior Scenario 1. However, although all decision coincide, expected losses increase from risk-averse to risk-friendly (see Table B1).
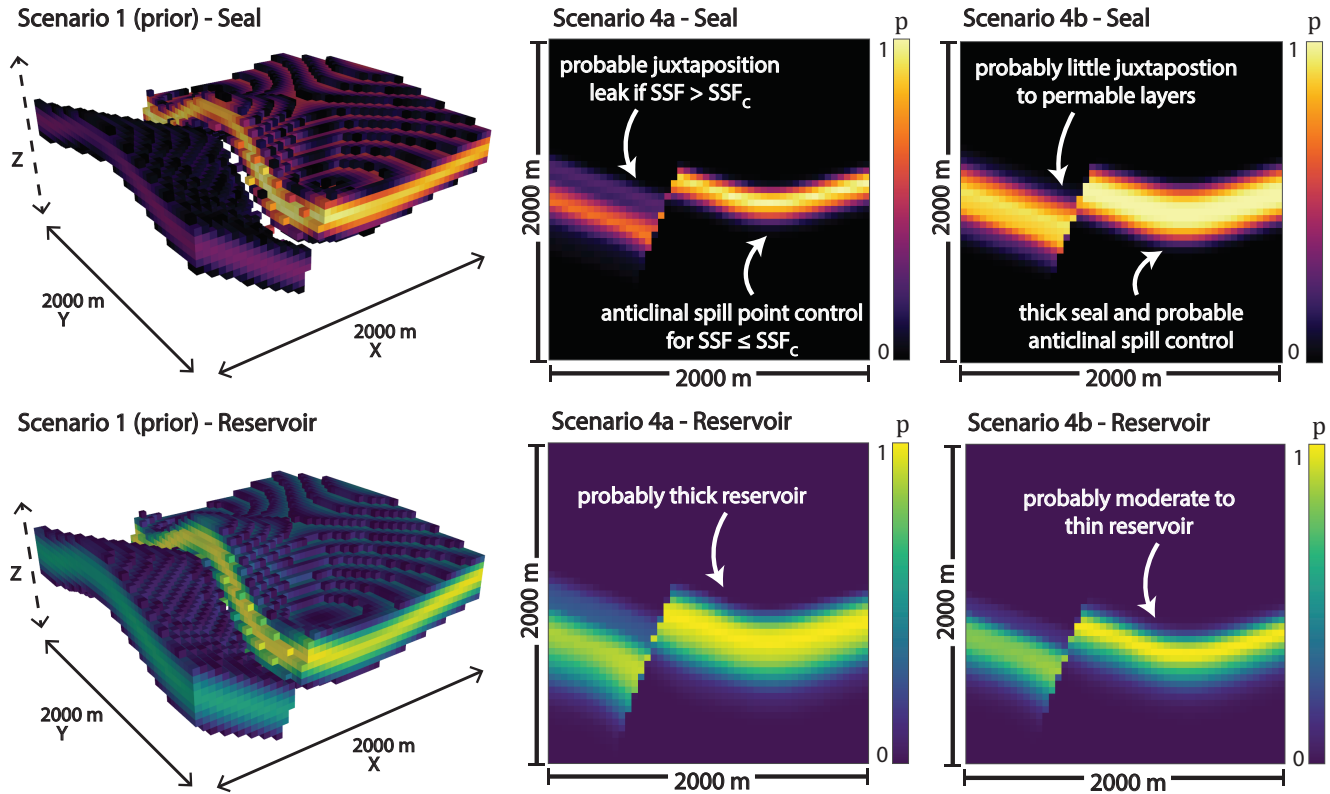
Figure 6. Probability field visualizations for seal and reservoir units in Scenarios 1 (prior), 4a and 4b. For Scenario 1, we used 3-D voxel visualizations and set a threshold at a probability of 0.5 (only voxels with a probability higher than 0.5 are shown). It can be recognized, that the seal is disrupted across the fault in more than 50 % of the prior model realizations. For the other scenarios, we show the full probability field for both units on a section through the middle of the model ($y = 500$ m), parallel to the $x - z$ plane.

3.3 Introducing reservoir thickness likelihoods (Scenarios 3a and 3b)

We also tested scenarios for a likelihoods of a thick reservoir formation alone (Scenario 3a) and in combination with the likelihood of a thick seal (Scenario 3b; see Table 2. The overall effect of using these reservoir-based likelihoods turned out to be minor compared to the seal-related scenarios.

In Scenario 3a, failure probabilities slightly increased, resulting in a decision shift towards lower values (see Fig. B1). Results for Scenario 3b are very similar to those of 2b, as can also be seen in Table B1. There was no significant reduction of expected losses or shift in decisions by adding the likelihood of a thick reservoir to the likelihood of a thick seal.
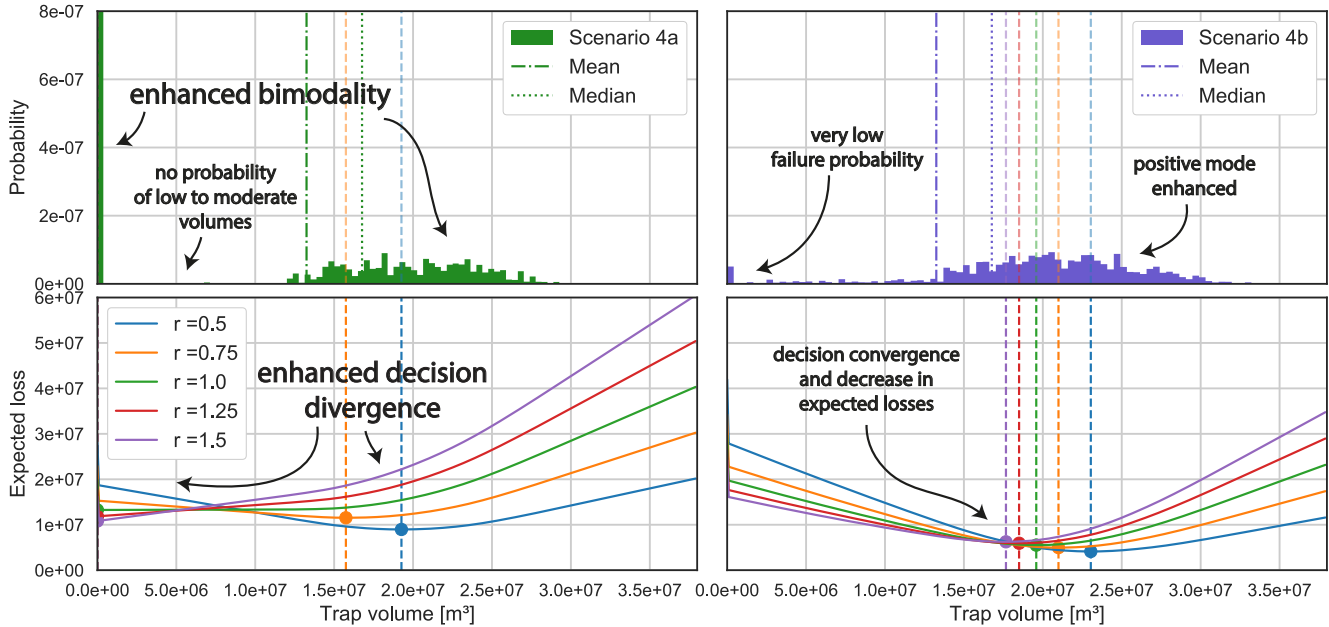

Figure 7. Trap volume distribution and resulting loss function realizations for Scenario 4a and Scenario 4b. Adding a likelihood of the SSF being around its critical value led to increased bimodality and an elimination of low to moderate volume probabilities. Bayes actions diverged accordingly in Scenario 4a. Implementing a reliable SSF value likelihood ($\mu = 2$, $\sigma = 0.3$) in combination with the thick seal likelihood from Scenario 2a, resulted in a emphasis on highly positive volumes. This, in turn, led to a stark convergence of decisions and reduction of expected losses.

3.4 Introducing SSF likelihoods

5 We considered two SSF-related likelihood scenarios. In Scenario 4a, we implemented solely a SSF likelihood that was based on a narrow normal distribution ($\mu = 5.1$, $\sigma = 0.3$) with a mean near the critical value $SSF_c = 5$. In Scenario 4b, we combined the likelihood of a thick seal (2a) with a likely moderate but reliable SSF value (SSF normal distribution with $mu = 2$ and $\sigma = 0.3$). Figure 6 illustrates the posterior situations well.

Scenario 4a resulted in increased bimodality of the posterior distribution (see Fig. 7). Accordingly, the Bayes action divergence and expected losses increased. Only two trap control mechanisms remained relevant for 4a (see Table B2): anticlinal

10 spill (66 %) and cross-fault leakage to overlying formations (34 %).

The results for 4b were comparable to those of 2a, but more pronounced. Entropies, particularly related to the seal thickness, were clearly reduced, also in the hanging wall. Probabilities of failure and low volumes were almost eliminated, further

enhancing the highly positive mode. This consequently resulted in an even higher convergence of Bayes actions, as well as reduction of expected losses compared to Scenario 2a. Anticlinal spill is the decisive control mechanism in 79.5 % of cases, otherwise only cross-fault leakage to the reservoir occurred (20.5 %).

4   Discussion

~~In this work, we build upon the recent advances presented by de la Varga et al. (2019) which enable us to view geological modeling as a probabilistic statistical problem. We expand on this by proposing custom loss functions as a useful decision-making tool when dealing with uncertain structural geological settings and to measure the effects of adding new information to a model. This is also aimed to illustrate the significance of the Bayesian perspective with regards to model interpretation in an economic context. As an explanatory example, we chose the hydrocarbon sector. This field is characterized by the necessity to make decisions in the face of high risks and potentially high rewards. These decisions are often closely linked to geological modeling and~~

Our results show that it is possible to apply Bayesian decision theory to geological models, as an approach to obtain an objective basis for decisions, considering uncertainties in these models. Even though the concept itself is not new, the application to the ~~estimation of reservoir-related values. By developing case-specific~~ *context of probabilistic geological modeling requires some adaptation and care when constructing appropriate loss functions. Our results highlight the potential use of* custom loss functions, ~~we intended to show that this estimation approach is suitable and useful to express the nature of complexities behind decision-making problems, decision environments and~~ *first for a simplified 1-D case, and then for a more complex full 3-D model. Even though these models are both conceptual, they highlight in our point of view the interesting potential of the method, as the optimal decision,* the ~~risk-related behavior of actors.~~ *Bayes action, is not always directly obvious when only considering posterior predictive distributions. Especially the addition of subjective risk affinity and the risk of critical overestimation lead to interesting changes in the optimal decision. Given these aspects, we consider the use of custom loss functions with probabilistic geological modeling as a very suitable combination in the framework of Bayesian decision theory.*

The case study considered here addressed a typical scenario of exploration for a fluid reservoir. We first discuss additional relevant points with regard to this specific case and provide then more general comments on extensions and the application in additional fields where geological models are commonly used.

4.1   ~~State of knowledge, decision uncertainty and consequent decision making~~

~~As we defined trap volume to be in essence a deterministic function of uncertain modelinput parameters, uncertainties propagate to this parameter of interest when conducting stochastic simulations. We consider the resulting volume probability distributions to be expressions of the respective state of knowledge (or information) on which the decision making is to be based. As this should include all parameters and conditions relevant for decision making, we furthermore propose that the overall uncertainty inherent in this probability distribution can be referred to as "decisionuncertainty" and that this entity should be viewed~~

separately from geological model uncertainty.By viewing decision making as a problem of optimizing a case-specific custom loss function applied to such a state of knowledge and decision uncertainty, we were able to observe clear differences in the respective behavior of distinctly risk-affine actors.The position and separation of their minimizing estimators, i.e. their decisions, manifested according to the properties of the value distributions. General spread and the occurrence of modes relative to the overall distribution and the relevant decision space appear to be particularly significant. High spread and bimodal tendencies, i.e. high overall uncertainty, resulted in a wider separation of different actions. Reduction of the distribution to one mode conversely led to their convergence. A decrease in decision uncertainty furthermore was accompanied by a reduction in expected loss for each Bayes estimator. Considering these observations, we derive that the degree of action convergence and respective expected losses can be considered measures for the state of knowledge and decision uncertainty at the moment of making a decision. The better these are, the more similar the decisions of differently risk-affine actors and the lower their loss expectations are . Given perfect information all actors would bid on the same estimate (the true value) and expect no loss, since no risk would be present. It furthermore follows from this that the relevance of risk affinity decreases with greater reduction of decision uncertainty.

4.1 ~~The~~ On the impact of additional information on decision making

We used these loss-function-related indicators to assess the significance ~~additional information might have for decisionmaking~~of additional information on a decision. We observed that the impact on decision uncertainty, induced by Bayesian inference, is not simply strictly aligned with the change in uncertainty regarding model parameters, but on those parameter combinations which are relevant for the outcome of the value of interest. It seems to be of central importance (1) "where" in the model uncertainty is reduced, i.e. in which spatial area or regarding which model parameters, and (2) which possible outcome is enhanced in terms of probability. An increased probability of a thick or thin seal in our model equally reduced decision uncertainty significantly, by raising the probability of a positive or negative outcome, respectively. Improved certainty about our reservoir thickness, however, had far lesser impact on decision making. This shows that some areas and parameter combinations have a much greater influence on the decision uncertainty than others, depending on the way they contribute to the outcome of the value of interest.

Some types of additional information could even lead to increased decision uncertainty. We observed this in Scenario 4a. The introduced SSF likelihood practically constrained our geological model to two possible situations: (1) a trap which is sealed off from juxtaposing layers and full-to-spill, and (2) complete failure of the trap due to a breached seal across the fault. This made the decision problem a predominantly binary one and split the outcome distribution into two narrowed but distant modes. The resulting increase in decision divergence and expected losses show that, in some cases, adding information might leave actors in greater disagreement than before.

However, we furthermore have to consider that actors weight possible outcomes of the value distribution differently. They consequently are affected differently by the same type of information. Risk-friendly actors were the most robust in their decision making in the face of possible trap failure. Eliminating this risk proved to be far less significant to the most risk-friendly, than for risk-averse actors. Accordingly, it should be of foremost importance for risk-averse actors to reduce the

uncertainty regarding critical factors, such as seal integrity, which might decide between the success and complete failure of a project. This is less relevant for risk-friendly decisions makers, who respectively might acquire a comparable benefit from knowing more about the probability of positive outcomes. They are less afraid of failure and than they are of missing out on opportunity.

Crucial risks might be easily assessed if they are dependent on only one or a few parameters, such as seal thickness. In other cases, they are derived from more complex parameter interrelations, as is the case for the Shale Smear Factor. To approach an effective mitigation of high risks, the complexities behind decisive factors need to be assessed thoroughly and respective parent parameters, as well as their interdependencies, need to be identified. This might enable a better understanding of which type of information is missing and where in the model additional data might be of use for improved decision making.

More of simply any type of information does not necessarily lead to better decisions. Instead, improved decision making is achieved by attaining the right kind of information that is able to shed light on uncertainties which are relevant to an ~~individuals~~ individual's own goals and preferences, as well as the general problem at hand. Bratvold and Begg (2010) stated that value is not generated by uncertainty quantification or reduction in itself, but is created to the extent that these processes have potential to change a decision. Such decision changes were clearly indicated by the shifting of actions in our different scenarios. According to Hammitt and Shlyakhter (1999), the difference in expected payoff between the prior and posterior optimal decision gives the expected value of information. This raises the question as to what ~~extend~~ extent a change in expected losses in itself might be an indicator for the value of information and if there is value in gaining confidence in a decision, even though it remains unchanged.

4.1.1 ~~The significance of our method for the hydrocarbon sector~~

While Monte Carlo simulation is by now common in the hydrocarbon sector, it does not make decisions, as Murtha et al. (1997) ~~emphasized. It~~ emphasized—it merely prepares for it. We believe that loss functions have the potential to go one step further. A hypothetical ideal loss function would consider all conditions in an economic environment, as well as perfectly represent preferences and goals of an actor and consequently be able to automatically find an optimal decision. While this is obviously unrealistic, we presume that an elaborate loss function might at least provide a very good preliminary decision recommendation. It might furthermore be able to weight risks that are not immediately apparent to an individual as a person. Furthermore, the influence of human biases and psychological behavioral challenges, as described by Bratvold and Begg (2010), could be mitigated.

Bayesian inference and MCMC methods have been applied for OOIP estimation and forecasting of reservoir productivity by Wadsley et al. (2005), Ma et al. (2006) and Liu et al. (2010). However, their research focused on history-matching simulations for already producing fields. Our approach of applying Bayesian inference for structural geological modeling and volumetric reservoir calculations is intended to support decision making in the earliest stages of a reservoir, when it has to be decided whether a project should be developed or not. Nevertheless, it was shown in the research conducted by Wadsley et al. (2005) that early volumetric OOIP estimates can be combined with later calculations from production data via MCMC methods.

Our continuous approach could be integrated into common discrete decision-making frameworks, such as decision trees. In real cases, normally only a limited number of options is given. In the context of hydrocarbon exploration and production, this would relate to fixed magnitudes of resource allocation, such as a certain number of required drilling wells or the size of a production platform. Based on such previously defined actual options, we could discretize our value probability distribution into sections, which represent each decision scenario accordingly. Our minimizing estimators would then indicate the best discrete option for a decision maker.
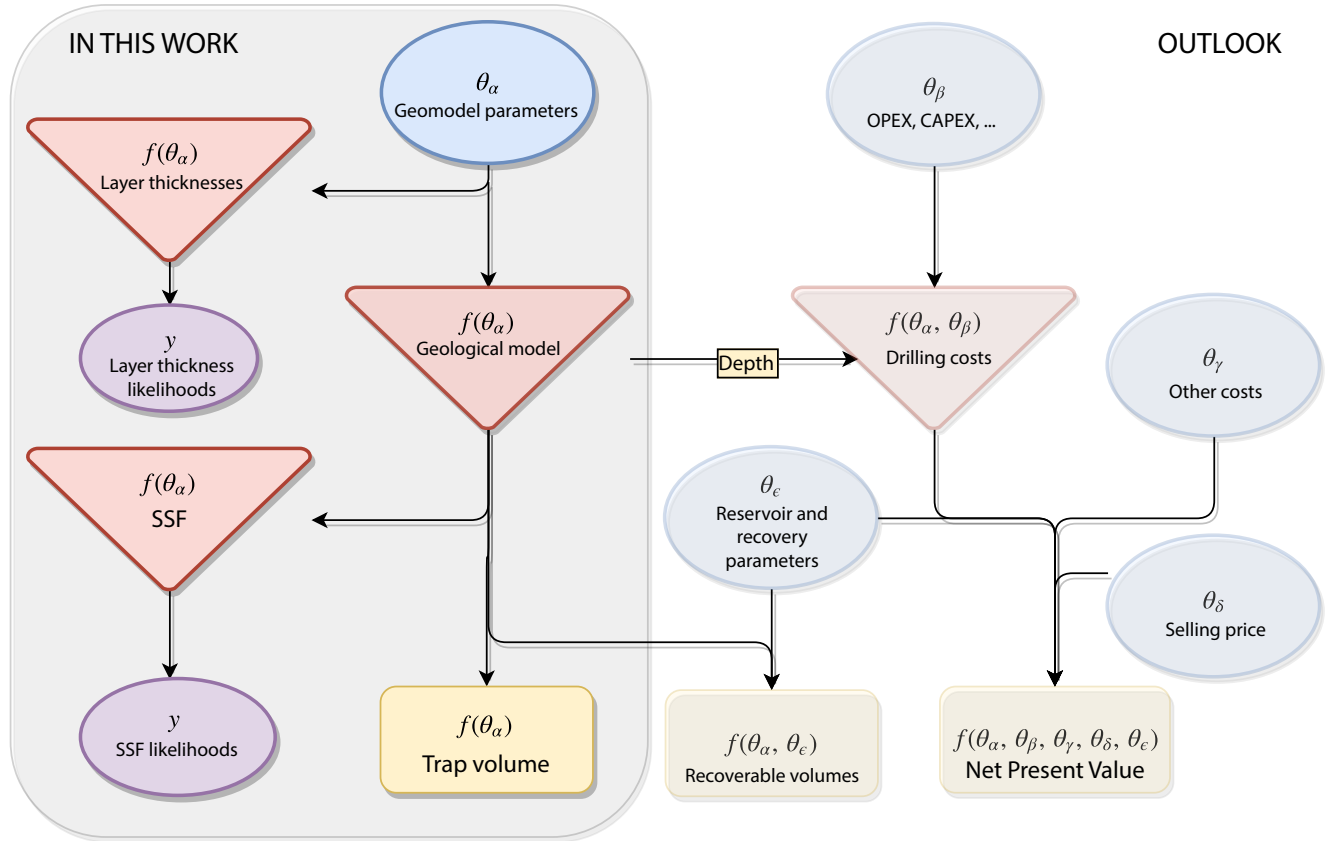
Figure 8. In this work, we applied our loss function approach to estimate a hydrocarbon trap volume. For this, we considered stochastic geomodeling parameters, defined deterministic functions to acquire volume, layer thicknesses and SSF values, and linked the latter two to respective likelihoods. Regarding the bigger picture, this methodology is expandable and could include other parameters and dependencies. By taking into account other reservoir parameters and recovery factors, we could for example base decision making on recoverable volumes. We could also take depth information from our model and combine this with other cost parameters to calculate drilling costs. Including additional costs, but also the selling price of hydrocarbons, we could attain the NPV as our final value of interest.

4.1.1 ~~Limitations and outlook~~

It has to be emphasized that we used

4.2 State of knowledge, decision uncertainty and consistent decision making

We applied the concept here to a synthetic geological model, not based on real data. Nevertheless, it was designed to include some typical structural characteristics related to hydrocarbon fluid reservoir systems. We developed algorithms aimed

5 to consider the most common conditions that define structural traps. However, these conditions needed to be simplified and were implemented on a very conceptual level. Furthermore, uncertainties employed in the 3-D model related to z-positional values only and were thus of primarily one-dimensional nature. It follows that no effective uncertainty concerning the overall structural shape was implemented, particularly regarding anticlinal features and the lateral position of the spill point.

As we defined trap volume to be in essence a deterministic function of uncertain model input parameters, uncertainties

10 propagate to this parameter of interest when conducting stochastic simulations. We consider the resulting volume probability distributions to be expressions of the respective state of knowledge (or information) on which the decision making is to be based. As this should include all parameters and conditions relevant for decision making, we furthermore propose that the overall uncertainty inherent in this probability distribution can be referred to as "decision uncertainty" and that this entity should be viewed separately from geological model uncertainty.

15 By viewing decision making as a problem of optimizing a case-specific custom loss function applied to such a state of knowledge and decision uncertainty, we were able to observe clear differences in the respective behavior of distinctly risk-affine actors.

The position and separation of their minimizing estimators, i.e. their decisions, manifested according to the properties of the value distributions. General spread and the occurrence of modes relative to the overall distribution and the relevant decision

20 space appear to be particularly significant. High spread and bimodal tendencies, i.e. high overall uncertainty, resulted in a wider separation of different actions. Reduction of the distribution to one mode conversely led to their convergence. A decrease in decision uncertainty furthermore was accompanied by a reduction in expected loss for each Bayes estimator.

Considering these observations, we derive that the degree of action convergence and respective expected losses can be considered measures for the state of knowledge and decision uncertainty at the moment of making a decision. The better these

25 are, the more similar the decisions of differently risk-affine actors and the lower their loss expectations are. Given perfect information all actors would bid on the same estimate (the true value) and expect no loss, since no risk would be present. It furthermore follows from this that the relevance of risk affinity decreases with greater reduction of decision uncertainty.

4.3 Extensions and Outlook

We applied the concept of decision theory here to an implicit geological modeling method (de la Varga et al., 2019). Depending

30 on the application, other types of geometric interpolations may be more suitable to represent the geological setting. More details on these methods, as well as the consideration of respective model uncertainties and the potential integration into probabilistic frameworks are described, for example, in Wellmann and Caumon (2018).

We defined risk affinity to be dependent on arbitrarily chosen risk factors which led to according reweighting. Davidson-Pilon (2015) used risk parameters determined by the maximal loss each actor could incur. Other approaches could be based on more tangible values, for example by making risk attitude dependent on a fixed budget.

There are still many points that could be expanded on in future research. It would be of interest to apply the same overall concept and methodology to an authentic case based on real datasets. Given a realistic economic scenario, including capital and operational expenditures of a project, possibly a full net-present-value (NPV) analysis could be conducted, applying a loss function to a NPV distribution (see Fig. 8). A more elaborate loss function could be customized on the base of surveys, acquiring the specific preferences of one or several companies and thus attaining a better profile of the economic environment, as well as the individuals acting in it.

~~Furthermore, additional and different uncertain parameters should be considered in the future. A respective next step regarding our model would be the incorporation of uncertainties which to a wider extent affect structural shapes in all three dimensions. Otherwise, non-structural reservoir parameters could be included as uncertain values, such as porosities and permeabilities of different layers. This might be of particular interest considering factors which are related to high risks in decision making.~~ We chose hydrocarbon systems and petroleum exploration as a sector for an exemplary application, ~~but many other settings can be found in which~~ as studies on risk related to geological modeling are most prominent in this field. However, geological modeling is of central ~~significance for decision making . One example would be subsurface storage of fluids in a reservoir, such as~~ importance to decision making in several other fields. Directly related are all other types of subsurface fluid reservoirs, for example in groundwater extraction or geothermal energy usage. Also closely related are applications of fluid storage in subsurface reservoirs, most prominently carbon capture and storage (CCS) applications. Questions regarding storage capacity and safety deal with similar conditions and geological problems as the ones presented this work. ~~Other examples might be mining or geotechnical projects. In general, it does not lack of geo-related areas in which a good understanding of uncertainties , risks and the general state of information is crucial for good decision making, and in which~~ The described concepts can similarly be applied to other types of geological features, for example ore bodies in mineral exploration or subsurface structures and materials in geotechnical applications. In all of these cases, the geological model can have significant uncertainties and, similar to the example described in this manuscript, further engineering and usage aspects carry high costs. We are therefore confident that a more detailed analysis of uncertainties and the definition and understanding of custom loss functions ~~could prove to be a beneficial tool~~ in the context of Bayesian decision theory are very interesting paths for more research with wide possible applications.

29

Appendix A: Determination of the maximum trap volume

The volume is calculated on a voxel-count basis. To assign model voxels to the trap feature, it is checked whether the following conditions (illustrated in Figure 4) are satisfied by each individual voxel:

1. **Labeled as reservoir formation:** The voxel has been assigned to the target reservoir formation (see Sandstone 1 in Fig. 4 (1)) in GemPy's lithology block model.

2. **Location above spill point horizon:** The voxel is located vertically above the final spill point of the trap. In the algorithm to find this final spill point, it is distinguished between a spill point defined by the folding structure, referred to as anticlinal spill point, and a cross-fault leak point, that depends on the magnitude of displacement and the resulting nature of juxtapositions. Once both of these points have been determined, the higher one is defined to be the final spill point used to determine the maximum fill capacity of the trap. Given a juxtapostion with layers overlying the seal, due to fault displacement, the respective section is checked for fault sealing by taking into account the Shale Smear Factor (SSF) value which is the ratio of fault throw magnitude D to displaced shale thickness T (Lindsay et al., 1993; Yielding et al., 1997; Yielding, 2012):

$$\text{SSF} = \frac{D}{T}. \tag{A1}$$

We attain both D and T by examining the contact between the seal lithology voxels and the fault surface.

For our model, we define the critical SSF to be $\text{SSF}_c = 5$. We assume that cross-fault sealing is breached when this threshold is surpassed. For simplicity, the fault is considered to be sealing along its plane.

3. **Location inside of closed system:** The voxel is part of a model section inside of the main anticlinal feature. All of the voxels inside this particular section are separated from the borders of the model by voxels that do not meet the first two conditions above, which primarily means that they are encapsulated by seal voxels upwards and laterally. This condition is relevant under the assumption, that connection to the borders of the model lead to leakage. A trap is thus defined as a closed system in this model and trap closure is assumed to be void outside of the space of information, i.e. the model space. In our example model, this also means that hydrocarbons escape in the hanging wall due to respective layer dipping upwards towards the model borders.

It has to be emphasized that these conditions have been fitted to our synthetic example model. For other models featuring different geological properties, structures and levels of complexities, these conditions and respective algorithms might not apply. Models of higher complexities will surely require the introduction of further conditions.

A1   Anticlinal spill point detection

Regarding anticlinal structures and traps, it can be observed that, geometrically and mathematically, a spill point is a saddle point of the reservoir top surface in 3-D. This was described by Collignon et al. (2015), who pointed out that the linkage of

folds is given by saddle points. These are thus a controlling factor for spill-related migration from respective structural traps. For anticlinal traps, closure can consequently be defined as the distance between the saddle point (i.e. spill point) and maximal point of the trap (Collignon et al., 2015).

Regarding a surface defined by $f(x, y)$, a local maximum at (x_0, y_0, z_0) would resemble a hill top (Guichard et al., 2013). Local maxima will be found looking at the cross-sections in the planes $y = y_0$ and $x = x_0$. Furthermore, the respective partial derivatives (i.e. gradients) $\frac{\delta z}{\delta x}$ and $\frac{\delta z}{\delta y}$ will equal zero at x_0 and y_0, i.e. that the extremum is a stationary point (Guichard et al., 2013; Weisstein, 2017). In the context of a geological reservoir system, such a hill can be regarded as a representation of an anticlinal structural trap. Local minima are defined analogously, presenting local minima in both planes at a stationary point (Guichard et al., 2013). A saddle point, however, is a stationary point, while not being an extremum (Weisstein, 2017). In general, saddle points can be distinguished from extrema by applying the second derivative test (Guichard et al., 2013; Weisstein, 2017): Considering a 2D function $f(x, y)$ with continuous partial derivatives at a point (x_0, y_0), so that $f_x(x_0, y_0) = 0$ and $f_x(x_0, y_0) = 0$, the following discriminant D can be introduced:

$$D(x_0, y_0) = f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2. \tag{A2}$$

Using this, the following holds for a point (x_0, y_0):

1. If $D > 0$ and $f_{xx}(x_0, y_0) < 0$, there is a local maximum.

2. If $D > 0$ and $f_{xx}(x_0, y_0) > 0$, there is a local minimum.

3. If $D < 0$, there is a saddle point at the point (x_0, y_0).

4. If $D = 0$, the test fails (Guichard et al., 2013).

According to Verschelde (2017), a saddle point in a matrix is maximal in its row and minimal in its column. This corresponds to the logical geometrical deduction, that a saddle point for a surface defined by $f(x, y)$ is marked by a local maximum in one plane, but a local minimum in the perpendicular plane. In our spill point detection algorithm, we make use of GemPy's ability to return layer boundary surfaces (simplices and vertices) as well as the gradients of the potential fields in discretized arrays:

1. We first look for vertices at which the surface of interest coincides with a gradient zero point.

2. Then, we check for the change in gradient sign at each such point in perpendicular directions. If they are opposite to one another, we can classify the vertex as a saddle point.

3. Lastly, we declare the highest saddle point to be our anticlinal spill point.

A2 Cross-fault leak point detection

For the potential point of leakage to formations underlying the seal across the normal fault (including the reservoir itself), we take the highest z-position of the reservoir units' contact (voxelized) with the fault in the hanging wall.

In the case of a juxtaposition with seal-overlying formations and a failed SSF-check, the maximum contact of the trap with the fault becomes the final spill point. Due to the shape of the trap in our model, we can then expect full leakage and set the maximum trap volume to zero.

A3   Calculating the maximum trap volume

5   Have all trap voxels been determined via the conditions defined in Section 2.2.3, the maximum trap volume V_t is calculated by simply counting the number of trap voxels and rescaling their cumulative volume depending on the resolution in which the model was computed:

$$V_t = n_v * (\frac{S_o}{R_m})^3, \tag{A3}$$

Where n_v is the number of trap voxels, S_o gives the original scale and R_m the used resolution for the model.

10   For the example of a cubic geological model with an original extent of 2000 m in three directions, computed using a resolution of 50 voxels in every direction, the scale factor is 40 m. Every voxel thus accounts for $40 \, \text{m} * 40 \, \text{m} * 40 \, \text{m} = 64{,}000 \, \text{m}^3$ in volume. It has to be noted, that this direct approach to rescaling and calculating the volume requires the model to be computed in cubic voxels.

Appendix B: Bayesian inference

15   Bayesian inference is defined by and conducted via the following equation, called the Bayes' Theorem (Jaynes, 2003; Gelman et al., 2014; I
÷

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

Appendix B: Further analysis of the custom loss function

For better understanding of how our custom loss function determines the incurrence of loss, actual losses for three fixed true
20   values and risk-neutrality ($r = 1$) are plotted in Fig. 2. Loss based on the risk-neutral custom loss function (Eq. 8) for determined true scores of -250, 0 and 250. This plot is meant to clarify the way real losses are incurred for each guess, relative to a given true value. The expected loss, as seen in Fig. ?? is acquired by arithmetically averaging such deterministic loss realizations based on the true score probability distribution by using Eq. 3.
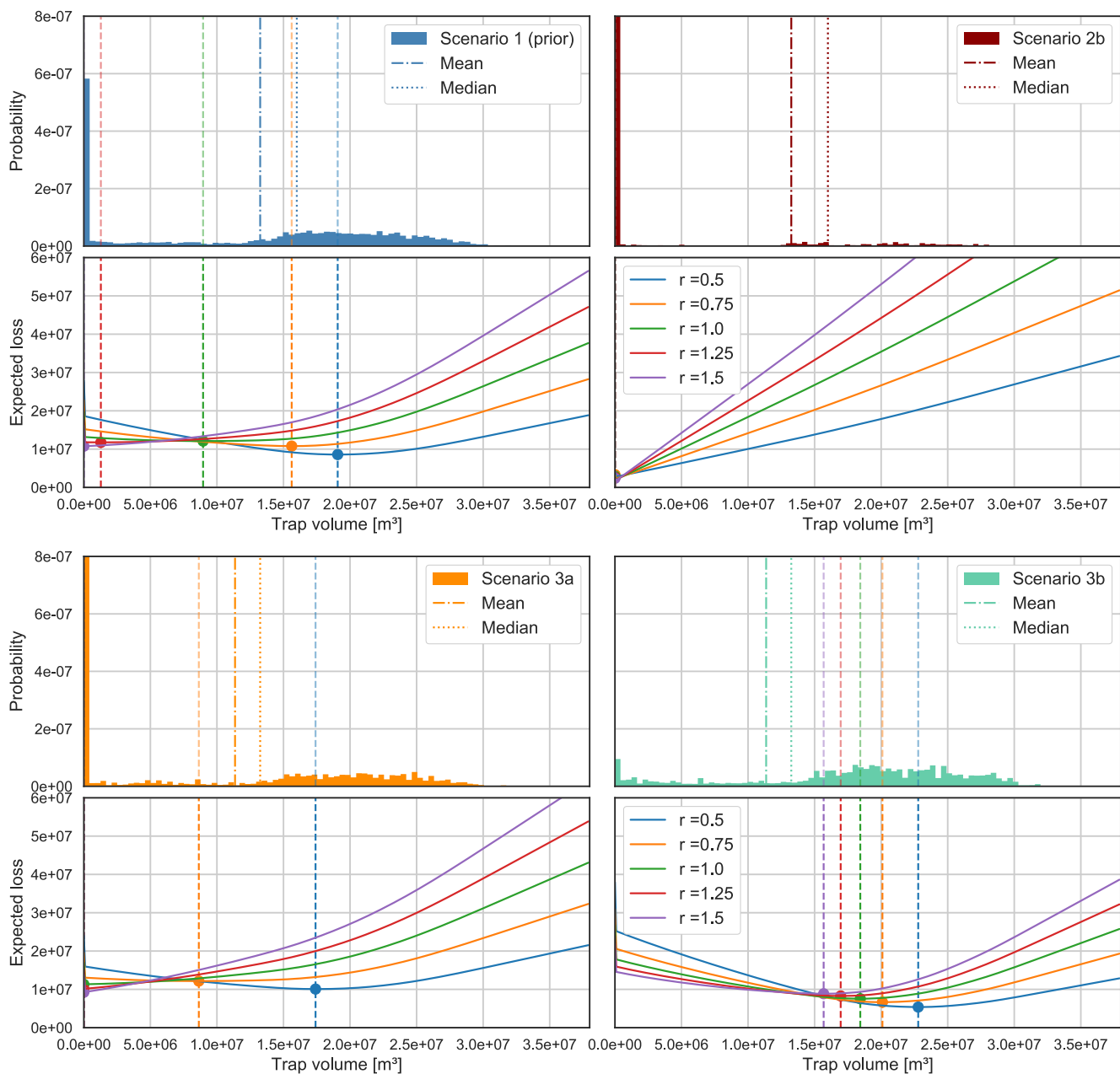
Appendix B: Results data

Figure B1. Posterior trap volume distributions and respective loss function realization plots for Scenarios 1 (prior), 2b, 3a and 3b.

Table B1. Decision results for all considered scenarios and each actor. Respective optimal estimates (decisions) are represented by $\hat{\theta}$, while $\Delta\hat{\theta}$ indicates posterior changes relative to the prior (Scenario 1) result. Expected losses are given by l, changes relative to the prior by Δl.

|		Decision makers				
		risk-friendly		**risk-neutral**		**risk-averse**
		r = 0.5	r = 0.75	r = 1.0	r = 1.25	r = 1.5
Scenario 1	$\hat{\theta}$	19072000.00	15616000.00	8960000.00	1280000.00	0.00
Prior	l	8582112.55	10785632.54	12100484.80	11759772.46	10763671.94
Scenario 2a	$\hat{\theta}$	22528000.00	19712000.00	17920000.00	16448000.00	15232000.00
Thick seal	$\Delta\hat{\theta}$	3456000.00	4096000.00	8960000.00	15168000.00	15232000.00
	l	5387582.96	6654239.73	7544384.00	8220155.30	8776678.80
	Δl	-3194529.59	-4131392.81	-4556100.80	-3539617.16	-1986993.14
Scenario 2b	$\hat{\theta}$	0.00	0.00	0.00	0.00	0.00
Thin seal	$\Delta\hat{\theta}$	-19072000.00	-15616000.00	-8960000.00	-1280000.00	0.00
	l	2743719.13	2240237.29	1940102.40	1735280.34	1584086.98
	Δl	-5838393.42	-8545395.25	-10160382.40	-10024492.12	-9179584.96
Scenario 3a	$\hat{\theta}$	17408000	8640000	0	0	0
Thick reservoir	$\Delta\hat{\theta}$	-1664000.00	-6976000.00	-8960000.00	-1280000.00	0.00
	l	10073515.53	12159993.48	11319609.6	10124566.62	9242422.54
	Δl	1491402.98	1374360.94	-780875.20	-1635205.84	-1521249.40
Scenario 3c	$\hat{\theta}$	22784000.00	20096000.00	18432000.00	16960000.00	15680000.00
Thick reservoir and seal	$\Delta\hat{\theta}$	3712000.00	4480000.00	9472000.00	15680000.00	15680000.00
	l	5380782.45	6658861.07	7551644.80	8278631.71	8857405.68
	Δl	-3201330.10	-4126771.47	-4548840.00	-3481140.75	-1906266.26
Scenario 4a	$\hat{\theta}$	19264000.00	15744000.00	0.00	0.00	0.00
Near critical SSF	$\Delta\hat{\theta}$	192000.00	128000.00	-8960000.00	-1280000.00	0.00
	l	8959284.13	11533073.67	13250828.80	11851901.58	10819256.41
	Δl	377171.58	747441.13	1150344.00	92129.12	55584.47
Scenario 4b	$\hat{\theta}$	23040000.00	20992000.00	19584000.00	18496000.00	17664000.00
Reliable SSF and thick seal	$\Delta\hat{\theta}$	3968000.00	5376000.00	10624000.00	17216000.00	17664000.00
	l	4112858.01	4964529.37	5513651.20	5929335.97	6245426.13
	Δl	-4469254.54	-5821103.17	-6586833.60	-5830436.49	-4518245.81
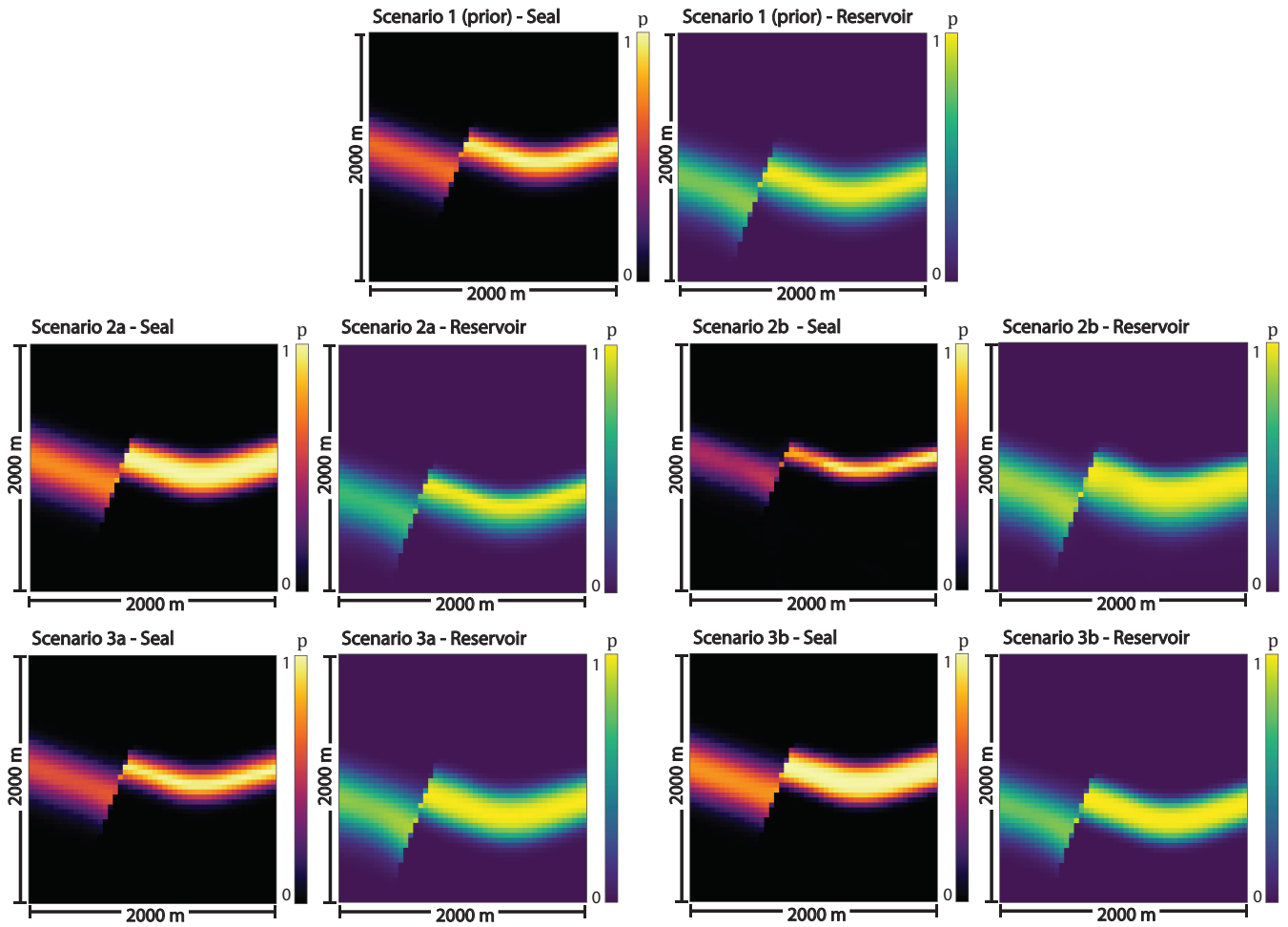
Figure B2. Probability field visualizations for Scenarios 1 to 3b.

**Table B2.** Occurrence rate of trap control mechanisms in % for each information scenario.

	1 - Anticlinal spill	2 - Leak to reservoir	3 - Leak to overlying	4 - Stratigraphic breach	5 - Unclear
Scenario 1	51.47	25.11	12.36	10.56	0.5
Scenario 2a	63.1	35.8	0.41	0.49	0.2
Scenario 2b	10.04	1.53	20.82	65.51	2.1
Scenario 3a	41.99	23.21	23.06	11.38	0.36
Scenario 3b	61.86	36.59	0.53	1.02	0
Scenario 4a	66.4	0.01	33.59	0	0
Scenario 4b	79.45	20.55	0	0	0

Appendix C: MCMC convergence
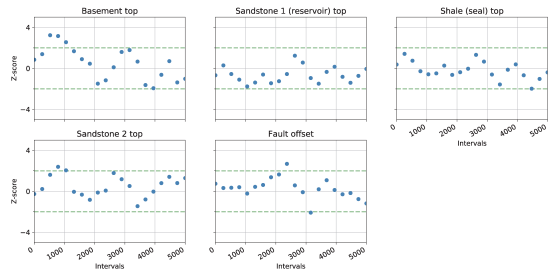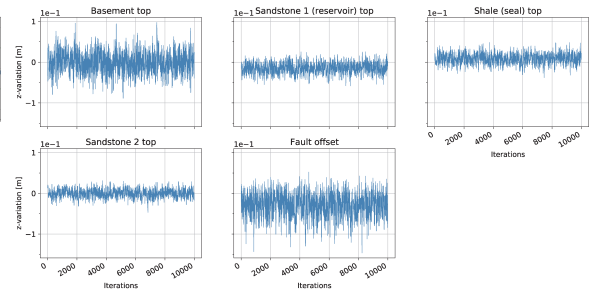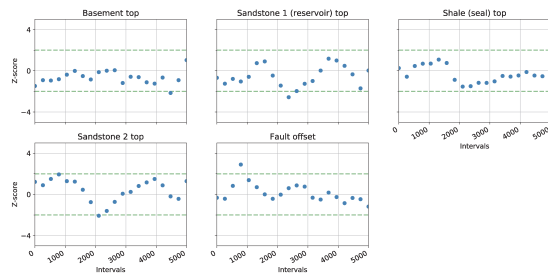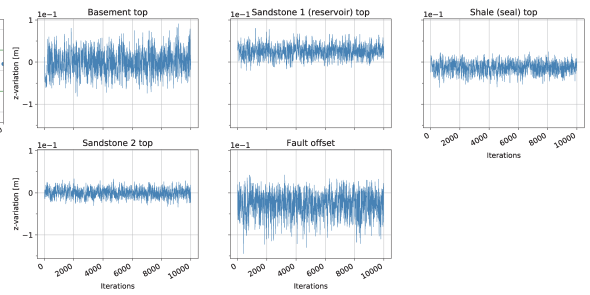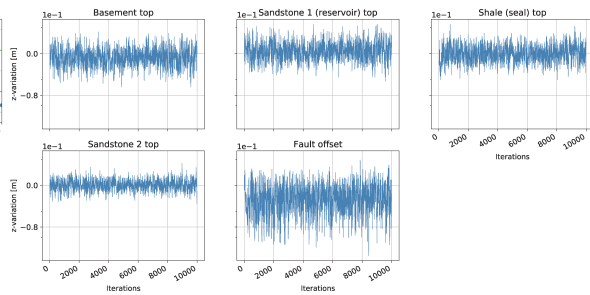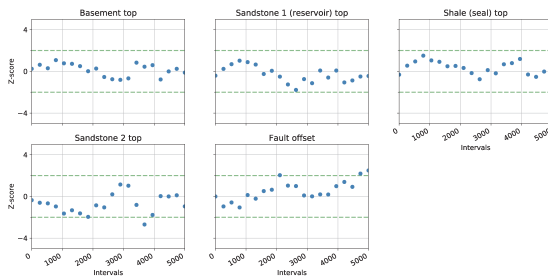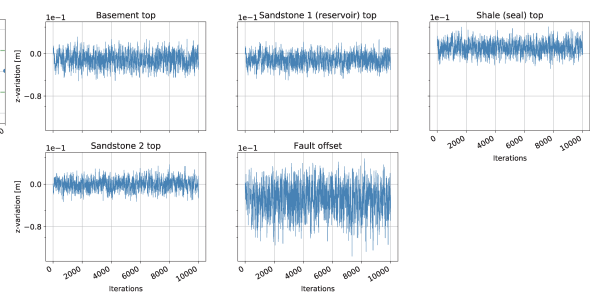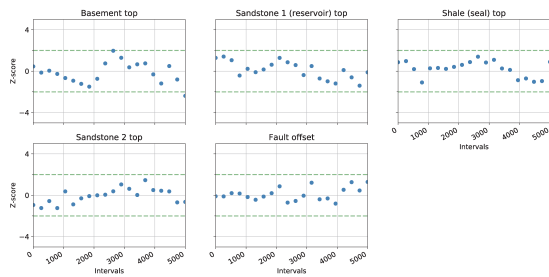
Figure C1. Geweke plots and traces for Scenarios 2a to 3b.
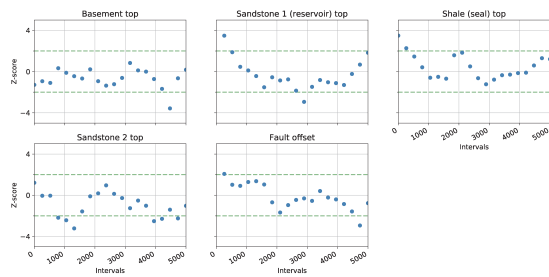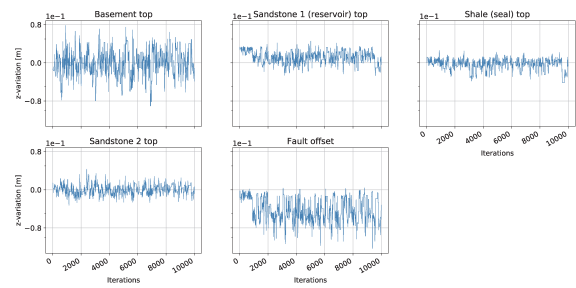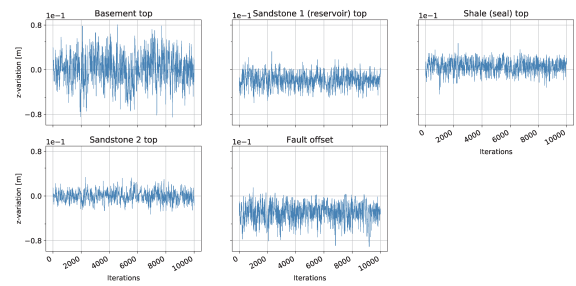
Figure C2. Geweke plots and traces for Scenarios 4a and 4b.

References

Bardossy, G. and Fodor, J.: Evaluation of Uncertainties and Risks in Geology: New Mathematical Approaches for their Handling, Springer, Berlin, Germany, 2004.

Berger, J. O.: Statistical decision theory and Bayesian analysis, Springer Science & Business Media, 2013.

5 Box, G. E. and Tiao, G. C.: Bayesian inference in statistical analysis, vol. 40, John Wiley & Sons, 2011.

Bratvold, R. and Begg, S.: Making good decisions, Society of Petroleum Engineers, 2010.

Caers, J.: Modeling Uncertainty in the Earth Sciences, John Wiley & Sons, Ltd, Chichester, UK, 2011.

Calcagno, P., Chilès, J.-P., Courrioux, G., and Guillen, A.: Geological modelling from field data and geological knowledge: Part I. Modelling method coupling 3D potential-field interpolation and geological rules, Physics of the Earth and Planetary Interiors, 171, 147–157, 2008.

10 Chatfield, C.: Model Uncertainty, Data Mining and Statistical Inference, Journal of the Royal Statistical Society. Series A (Statistics in Society), 158, 419–466, 1995.

Collignon, M., Fernandez, N., and Kaus, B.: Influence of surface processes and initial topography on lateral fold growth and fold linkage mode, Tectonics, 34, 1622–1645, 2015.

Davidson-Pilon, C.: Bayesian Methods for Hackers: Probabilistic Programming and Bayesian Inference, 2015.

15 de la Varga, M. and Wellmann, J. F.: Structural geologic modeling as an inference problem: A Bayesian perspective, Interpretation, 4, SM1–SM16, 2016.

de la Varga, M., Schaaf, A., and Wellmann, F.: GemPy 1.0: open-source stochastic geological modeling and inversion, Geoscientific Model Development, 12, 1–32, https://doi.org/10.5194/gmd-12-1-2019, https://www.geosci-model-dev.net/12/1/2019/, 2019.

Dean, L.: Reservoir Engineering for Geologists - Volumetric Estimation, The Monthly Magazine of the Canadian Society of Petroleum
20 Geologists, pp. 11–14, 2007.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B.: Bayesian data analysis, vol. 2, Taylor & Francis, 2014.

Geweke, J. et al.: Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments, vol. 196, Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, USA, 1991.

Guichard, D., Koblitz, N., and Keisler, H. J.: Calculus: Early Transcendentals, D. Guichard, 2013.

25 Haario, H., Saksman, E., and Tamminen, J.: An adaptive metropolis algorithm. Bernoulli 7 223–242, Mathematical Reviews (MathSciNet): MR1828504 Digital Object Identifier: doi, 10, 3318 737, 2001.

Hammitt, J. K. and Shlyakhter, A. I.: The expected value of information and the probability of surprise, Risk Analysis, 19, 135–152, 1999.

Harney, H. L.: Bayesian inference: Parameter estimation and decisions, Springer Science & Business Media, 2013.

Hennig, C. and Kutlukaya, M.: Some Thoughts About the Design of Loss Functions, REVSTAT - Statistical Journal, 5, 19–39, 2007.

30 Jaynes, E. T.: Probability theory: The logic of science, Cambridge university press, 2003.

Lajaunie, C., Courrioux, G., and Manuel, L.: Foliation fields and 3D cartography in geology: principles of a method based on potential interpolation, Mathematical Geology, 29, 571–584, 1997.

Lark, R., Mathers, S., Thorpe, S., Arkley, S., Morgan, D., and Lawrence, D.: A statistical assessment of the uncertainty in a 3-D geological framework model, Proceedings of the Geologists' Association, 124, 946–958, 2013.

35 Lindsay, N., Murphy, F., Walsh, J., and Watterson, J.: Outcrop studies of shale smears on fault surfaces, The Geological Modelling of Hydrocarbon Reservoirs and Outcrop Analogues, pp. 113–123, 1993.

Liu, C., McVay, D. A., et al.: Continuous reservoir-simulation-model updating and forecasting improves uncertainty quantification, SPE Reservoir Evaluation & Engineering, 13, 626–637, 2010.

Ma, X., Al-Harbi, M., Datta-Gupta, A., Efendiev, Y., et al.: A multistage sampling method for rapid quantification of uncertainty in history matching geological models, in: SPE Annual Technical Conference and Exhibition, Society of Petroleum Engineers, 2006.

5 McLane, M., Gouveia, J., Citron, G. P., MacKay, J., and Rose, P. R.: Responsible reporting of uncertain petroleum reserves, AAPG Bulletin, 92, 1431–1452, 2008.

Morton-Thompson, D., Woods, A. M., et al.: Development Geology Reference Manual: AAPG Methods in Exploration Series, No. 10, 10, AAPG, 1993.

Moyé, L. A.: Statistical reasoning in medicine: the intuitive P-value primer, Springer Science & Business Media, 2006.

10 Mudford, B. et al.: Valuing and comparing oil and gas opportunities: A comparison of decision tree and simulation methodologies, in: SPE Annual Technical Conference and Exhibition, Society of Petroleum Engineers, 2000.

Murtha, J. A. et al.: Monte Carlo simulation: its status and future, Journal of Petroleum Technology, 49, 361–373, 1997.

Randle, C. H., Bond, C. E., Lark, R. M., and Monaghan, A. A.: Uncertainty in geological interpretations: Effectiveness of expert elicitations, Geosphere, 15, 108–118, 2019.

15 Salvatier, J., Wiecki, T. V., and Fonnesbeck, C.: Probabilistic programming in Python using PyMC3, PeerJ Computer Science, 55, 2016.

Smalley, P. C., Begg, S. H., Naylor, M., Johnsen, S., and Godi, A.: Handling risk and uncertainty in petroleum exploration and asset management: An overview, AAPG Bulletin, 92, 1251–1261, 2008.

Thore, P., Shtuka, A., Lecour, M., Ait-Ettajer, T., and Cognot, R.: Structural uncertainties: Determination, management, and applications, Geophysics, 67, 840–852, 2002.

20 Verschelde, J.: Programming Tools and File Management - Part I: A Tour of Python, http://homepages.math.uic.edu/~jan/mcs275/mcs275notes/index.html, last visited on 18/10/2017, 2017.

Wadsley, A. W. et al.: Markov Chain Monte Carlo Methods for Reserves Estimation, in: International Petroleum Technology Conference, International Petroleum Technology Conference, 2005.

Wald, A.: Statistical decision functions., 1950.

25 Weisstein, E. W.: Saddle Point. From MathWorld—A Wolfram Web Resource, http://mathworld.wolfram.com/SaddlePoint.html, last visited on 18/10/2017, 2017.

Wellmann, F. and Caumon, G.: 3-D Structural geological models: Concepts, methods, and uncertainties, 2018.

Wellmann, J. F. and Regenauer-Lieb, K.: Uncertainties have a meaning: Information entropy as a quality measure for 3-D geological models, Tectonophysics, 526, 207–216, 2012.

30 Wellmann, J. F., Horowitz, F. G., Schill, E., and Regenauer-Lieb, K.: Towards incorporating uncertainty of structural data in 3D geological inversion, Tectonophysics, 490, 141–151, 2010a.

Wellmann, J. F., Horowitz, F. G., Schill, E., and Regenauer-Lieb, K.: Towards incorporating uncertainty of structural data in 3D geological inversion, Tectonophysics, 490, 141–151, 2010b.

Wellmann, J. F., de la Varga, M., Murdie, R. E., Gessner, K., and Jessell, M.: Uncertainty estimation for a geological model of the Sandstone
35 greenstone belt, Western Australia – insights from integrated geological and geophysical inversion in a Bayesian inference framework, in: Characterization of Ore-Forming Systems from Geological, Geochemical and Geophysical Studies, edited by Gessner, K., Glessinkop, T. G., and Sorjonen-Ward, P., Geological Society, London, Special Publications, 453, https://doi.org/https://doi.org/10.1144/SP453.12, 2017.

Wellmann, J. F., de la Varga, M., Murdie, R. E., Gessner, K., and Jessell, M.: Uncertainty estimation for a geological model of the Sandstone greenstone belt, Western Australia–insights from integrated geological and geophysical inversion in a Bayesian inference framework, Geological Society, London, Special Publications, 453, 41–56, 2018.

Wim, J., Swinkels, H., et al.: Guidelines for the evaluation of petroleum reserves and resources, Society of Petroleum Engineers, 2001.

5 Yielding, G.: Using probabilistic shale smear modelling to relate SGR predictions of column height to fault-zone heterogeneity, Petroleum Geoscience, 18, 33–42, 2012.

Yielding, G., Freeman, B., and Needham, D. T.: Quantitative fault seal prediction, AAPG bulletin, 81, 897–917, 1997.