

Woodruff, Joshua; Dimitrov, Nedialko B.

**Article**

## Optimal discretization for decision analysis

Operations Research Perspectives

**Provided in Cooperation with:**

Elsevier

*Suggested Citation:* Woodruff, Joshua; Dimitrov, Nedialko B. (2018) : Optimal discretization for decision analysis, Operations Research Perspectives, ISSN 2214-7160, Elsevier, Amsterdam, Vol. 5, pp. 288-305,  
<https://doi.org/10.1016/j.orp.2018.09.002>

This Version is available at:

<https://hdl.handle.net/10419/246356>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>



## Optimal discretization for decision analysis

Joshua Woodruff\*, Nediialko B. Dimitrov

The University of Texas at Austin, Austin, TX, United States



### ARTICLE INFO

**Keywords:**  
Optimization  
Discretization  
Joint discretization

### ABSTRACT

The use of discretization in decision analysis allows practitioners to use only a few assessments to estimate the certain equivalent (CE) or expected value of a decision without knowing the functional form of the distribution of each uncertainty. The discretization shortcuts are fast, but are created with a specific distribution, or families of distributions in mind. The discretizations are not formulated with the decision problem in mind. Each discretization is specific to one uncertainty distribution, or is even more generalized. In this article, we introduce a novel mathematical formulation for selecting an optimal discretization for a specific problem. With optimal discretization, a decision analyst can use the newly-created shortcuts in repeated decisions and improve the expected accuracy of the CE calculations.

### 1. Introduction

A common task in Decision Analysis is *discretization*, which involves reducing probability distributions of uncertainties to just a few point masses [13]. For many uncertainties in a Decision Analysis model, determining the value of multiple points in a distribution, let alone, the true distribution may be impossible or costly. Discretization is intended to reduce the cost of calculating certain equivalents (CEs) and simplify both communication and computation because it substitutes otherwise complex and computationally intensive integrations to the evaluations of just a handful of utilities. Discretizations significantly improve a decision analyst's ability to communicate with clients [16]. Even with an increase in computing power, discretizations allow for human-understandable assessment and evaluation of decisions.

The discretization process can be thought of as having two distinct components. The first component is to select the points that will be used for discretization. Often, these points are chosen to be percentiles of the original distribution. The second component of discretization is to assign a probability mass to each point. An example of a discretization applied to two common distributions is shown in Fig. 1.

Because of the usefulness of discretization, many methods for discretization exist [13]. One discretizations method divides the probability distribution into intervals based on values or cumulative distributions. Each interval is given a percentile equal to its median or its mean. Finally, each interval is assigned a probability based on the size of the interval. Another way of discretizing is to choose percentiles and probabilities that match the moments of the original distribution.

Typically, the underlying distributions of the uncertainties are unknown. As a result, these methods have given way to shortcut methods that are easy to implement and work across a broad range of distributions. Decision analysts seek discretization methods that produce near-correct certain equivalents across the entire class of decision problems of interest. Typically the number of percentiles that a decision analyst uses in a discretization is three, which provide a low, high, and most-likely value.

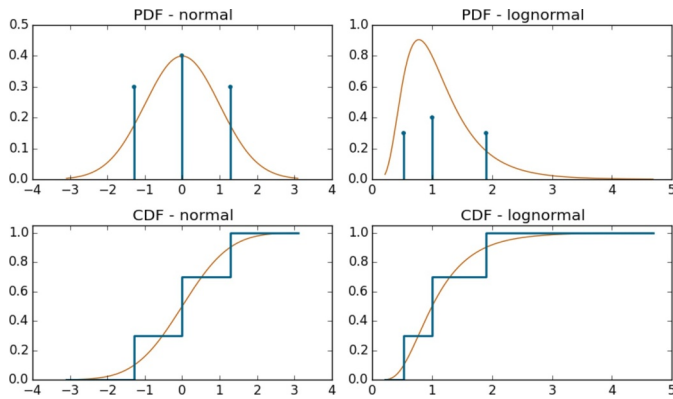
Though shortcut methods are easy to use, they are not created with the decision problem in mind. This can lead to reduced accuracy and situations where the results of different strategies are within a few percent of each other, sub-optimal decisions. Our method for choosing a discretization differs from these methods in that we seek a set of discretizations that provide the lowest certain equivalent error over a large set of potential distributions for a specific problem. The error particularly matters when the certain equivalent of a decision is close to zero. In a go, no-go decision, an error in the calculated certain equivalent would result in the wrong decision. We generate problem-specific discretizations by solving an optimization problem that chooses the percentiles and the corresponding probabilities that minimize an error metric over a large set of potential versions of a problem.

The main contributions of this article are: 1) We mathematically formulate the problem of selecting an optimal discretization for a set of decision problems. 2) We are able to derive tractable instances of this optimization problem. These tractable instances allow us to compute optimal discretizations for a specified set of decision problems. 3) Prior discretization methods produce discretizations for each uncertainty in

\* Corresponding author.

E-mail addresses: [joshua.woodruff@utexas.edu](mailto:joshua.woodruff@utexas.edu) (J. Woodruff), [ned.dimitrov@gmail.com](mailto:ned.dimitrov@gmail.com) (N.B. Dimitrov).

URLs: <http://www.optimizedfinancialsystems.com/> (J. Woodruff), <http://neddimitrov.org/> (N.B. Dimitrov).



**Fig. 1.** These are two examples of the same three-point discretization applied to a standard normal and a log-normal distribution. The placement of the points on the independent axis are determined by the percentiles of the discretization, and the height is based on the probabilities assigned.

the decision problem. We introduce a *joint discretization* where the probabilities of the percentiles of a discretization are not independent from each other. The mathematical formulation for computing optimized discretizations opens a new area of computing joint discretizations. Finally, we show that our methodology for computing independent and joint discretizations outperforms prior methods across a set of computational examples.

This article is structured as follows. We begin with related work in Section 2, describing some popular and novel methods of discretization. Next we provide a general formulation for optimal discretization in Section 3. The general formulation is followed by the a tractable instance in Section 4.1 and modifications required to create a joint discretization in Section 4.2. In Section 5, we analyze two examples from the literature to determine the benefits of optimal and joint discretization, and the effects of sampling and pre-determining percentiles. We conclude with a discussion and future work in Section 6.

**2. Related work**

Discretizations are typically divided into distribution-specific methods and shortcuts. Distribution-specific methods require knowledge of the probability density function (PDF) of an uncertainty’s distribution prior to discretization. These methods choose the discretization based on some criteria of the original distribution that the decision analyst is trying to match. Discretization shortcuts require experts to assess a few (usually three) percentiles of the uncertainty’s distribution. They do not require knowledge of the shape or moments and are easily applied. A third type of method is a hybrid approach. This method assumes limited knowledge of the underlying distribution and provides a discretization based on this knowledge. Prior methods focus on computing discretizations independently for each uncertainty. In areas such as stochastic optimization there are more examples of multivariate discretizations [20]. We will review a few of these discretization methods and compare them to our approach.

In a decision analysis project, if the client, decision analyst, or some other expert knows the true distribution of each uncertainty, then the decision analyst can determine the CE of even the largest problem using Monte Carlo sampling or some other technique. With the true CE, the decision analyst can recommend the strategy with the highest CE with the certainty that this is the best recommendation. In reality, the form of each uncertainty is unknown. The decision analyst will elicit assessments for the uncertainties and apply a discretization to these uncertainties to create an estimate of the CE.

Two common distribution-specific discretization methods are bracket mean and Gaussian quadrature. Bracket mean, described by McNamee and Celona [15], is also known as the equal areas

discretization. In bracket mean discretization, the PDF is partitioned into three sections, with probabilities 0.25, 0.50, and 0.25. The percentiles assigned to each region represent the mean value within the probability region. This method matches bounded means of the uncertainty’s PDF. Though the method matches these bounded means for each uncertainty, there is no guarantee that the discretization will match the certain equivalent (CE) in the decision problem.

Miller III and Rice [16] and later Smith [19] proposed techniques based on Gaussian quadrature (GQ). With a discretization of  $N$  points, it is possible to match the first  $2N - 1$  moments of an uncertainty’s distribution. The idea behind matching the moments comes from considering expectations of low degree polynomials. A discretization that matches an uncertainty in the first  $2N - 1$  moments, will produce the same expectation for any  $2N - 1$  order polynomial. Smith [19] improved the method for GQ by making it more efficient and provided examples of how GQ matched the moments of the input distributions ([19], Table 2, P.345). GQ requires knowledge of the distributions being discretized – at least its first  $2N - 1$  moments. If these are not known in the literature, one may require complex numerical integrations. An application of GQ also requires solving a multivariate system of polynomial equations, which is easily computed with matrix manipulation software. The polynomial matching argument misses cross-terms of the uncertainties when the utility depends on several uncertainties. From a client perspective, this discretization may ask for percentiles that are not easily assessed, such as the value of the uncertainty at the 99.50th percentile. In contrast, our method is able to limit the discretization’s percentiles to easily assessed values, directly targets computing CEs, and the discretization is dependent on a set of decision problems, as opposed to a single specific distribution.

Shortcuts are easy to use and generally perform well. Shortcut discretizations do not require knowledge of the distributions of the uncertainties and apply the same discretization percentiles and probabilities to all uncertainties. Two common shortcut methods are the McNamee–Celona Shortcut (MCS) and the Extended Swanson–Megill (ESM) method. Both of these discretization methods use the 10th, 50th, and 90th percentiles. The Extended Swanson–Megill (ESM) method for discretization is described by Hurst et al. [10] and is commonly used in the oil and gas industries [2]. MCS assigns probabilities of (0.25,0.50,0.25) while ESM assigns probabilities of (0.3,0.4,0.3) to each of the respective percentiles. In examining several discretization methods, Keefer and Bodily [13] proposed the Extended Pearson–Tukey method. This method proposed using the 5th, 50th, and 95th percentile with probabilities of (0.185,0.630,0.185). Keefer and Bodily [13] found EPT outperformed several other methods. Additionally Hammond and Bickel [8] found EPT to have the best performance among MCS, EPT, and ESM when calculating average absolute error, average absolute percent error, maximum error, and maximum percent error. While easy to apply, these methods do not take into consideration specific knowledge about the decision problem. In contrast, our method computes an optimized discretization for a specific set of decision problems.

It is often known if the distributions of the uncertainties have specific shapes or are bounded. For example, when drilling for oil it can be assumed that the percentage of oil to be recovered will be between 0 and 100%. To that extent, there are newer shortcuts that are based on shape-specific assumptions. Hammond and Bickel [9] and Hammond and Bickel [8] offer new shortcuts based on the Pearson and Johnson families of distributions. Based on the specific zone of the assumed distribution of the uncertainty, Hammond and Bickel [8] provide a specific discretization to use. For example, a normal distribution is symmetric and unbounded on either side. A log-normal distribution is bounded from below, might be right-skewed, and unbounded from above. These methods are similar to the shortcut methods, but address a specific family of distributions. There are shortcuts for bounded, semi-bounded, and unbounded distributions. The purpose is to provide a better discretization that uses potentially available information while still allowing for an unknown distribution. Distribution-specific

shortcuts are similar to the other discretization methods in that these methods only take into account individual uncertainties as opposed to the decision problem as a whole. In contrast, our method considers a set of similar decision problems and discretizes all the uncertainties simultaneously.

Wallace and Hoyland [20] developed a technique based on non-linear programming that can be used to generate a finite number of discrete events that satisfy specific statistical properties. The main difference between Wallace and Hoyland [20] and our method is that they consider a single distribution, and attempt to match potentially multiple statistical properties of that distribution using a small discrete set of scenarios. On the other hand, we have potentially many distributions and decision problems and we are looking for a single set of discrete percentiles that return approximately correct CEs for all the decision problems.

Most research has focused on matching aspects of a single uncertainty’s distribution – for example, its moments or bounded means. Brockett and Kahane [3] found that moment matching does not produce accurate CEs, partly because it does not take into account other aspects of the decision problem such as the value function and the utility function. Keefer [12] also tested the accuracy of three point discretizations and found “substantial” (P.763) errors in the CEs they produce. Three point discretizations can match the first five moments, which may have contributed to the failure of moment matching in this instance. While the uncertainty’s moments may match, the uncertainties combine and are passed through a value function, a utility function, and a certain equivalent calculation. In certain problems this may result in CE errors because the distribution of the CE may be completely different than any of the uncertainty distributions. Keefer [12] further tested six shortcut methods to determine their CE errors, with EPT identified as the best performer overall. Smith [19] found that bracket median and bracket mean had higher errors when calculating the CE than did EPT and GQ. In contrast to past methodologies, our method focuses on matching the CEs of a set of decision problems. In this way, the methodology can both capture lack of knowledge – because the discretization is based on a set of decision problems – and specific knowledge about the decision problem structure – which is incorporated through similarities in the set of decision problems.

### 3. General formulation for discretization

Optimal discretization finds a discretization for the uncertainties of a problem that is going to be revisited frequently. In the problem there are uncertainties whose distributions will change over time. For example, each oil field will have different potential reservoir and recovery ratio. In the consumer packaged goods industry every new product will have a different potential market share and market size. The realizations of the uncertainties are fed into a value function, which computes the net present value. Finally, there may be a utility function to convert project values to certain equivalents.

The general discretization problem assumes independence among the percentile values of the uncertainties. In the independent discretization, the percentiles chosen from one uncertainty are independent from one another. We still allow for dependence between the values of uncertainties. In the examples we use in this article, the values of one or more uncertainties depend on one or more other uncertainties. We use various methods to determine those values. In an assessment framework, the decision analyst would still elicit dependent (correlated) assessments from the independent percentiles.

In our decision problem, the functional form of the uncertainties is known. What we do not know is which decision problem we are facing and consequently are unsure of the functional forms we will face when needing to make a decision. The combination of all the potential uncertainty distributions when applied to a value and utility function come to define our set of decision problems,  $\mathbb{D}$ . To determine the CE of one instance of  $\mathbb{D}$ , it is possible to compute the CE by means of Monte

Carlo sampling. When interacting with a client, a decision analyst does not know the exact problem in  $\mathbb{D}$  they are addressing. We seek a discretization that works over all cases of  $\mathbb{D}$ . Let  $\mathcal{D}$  define a prior probability distribution over the potential decision problems  $\mathbb{D}$ . This is a probability assignment on each problem  $d \in \mathbb{D}$ . This allows the decision analyst to specify that some decision problems are more likely than others.

For a specific decision problem  $d \in \mathbb{D}$ , let  $CE_d$  denote the certain equivalent of decision problem  $d$ . Let  $\mathbb{P}$  be the set of allowable probability mass functions of the uncertainties. Let  $CE_d(p)$  for  $p \in \mathbb{P}$  denote the certain equivalent of decision problem  $d$  when the distribution  $p$  is used for the uncertainties instead of the true distribution. In other words,  $CE_d(p)$  is the certain equivalent when we use the discretized distribution instead of the true distribution of the uncertainties. We want to find a  $p \in \mathbb{P}$  that matches  $CE_d(p) \approx CE_d$  for all decision problems  $d \in \mathbb{D}$ . A discretization with a perfect fit will have  $CE_d = CE_d(p)$  for all  $d \in \mathbb{D}$ .

We begin by formulating discretization as an optimization problem. This is a novel contribution to the area of discretization, and leads us to the results in the rest of the paper. The discretization problem can be formulated as an optimization as follows:

$$\arg \min_{p \in \mathbb{P}} \left( \lambda \left[ \max_{d \in \mathbb{D}} \text{Err}(d, p) \right] + (1 - \lambda) [E_{d \in \mathcal{D}} \text{Err}(d, p)] \right), \tag{1}$$

where

$$\text{Err}(d, p) = \left| \frac{CE_d - CE_d(p)}{CE_d} \right| \text{ or} \tag{2}$$

$$\text{Err}(d, p) = |CE_d - CE_d(p)|. \tag{3}$$

Given a parameter  $\lambda \in [0, 1]$ , optimization (1) defines the discretization problem. The problem seeks to find a distribution  $p \in \mathbb{P}$  that yields the minimum convex combination of worst case absolute error and expected absolute error,  $E_{d \in \mathcal{D}} \text{Err}(d, p)$ , with respect to the distribution  $\mathcal{D}$ . We use  $\mathcal{D}$  as the distribution of decision problems to indicate there is a probability associated with each decision problem. Eq. (2) defines the absolute percentage error in the CE when using the discretized distribution of uncertainties instead of the true distribution. We also include (3) as an alternative formulation to the error function when  $CE_d$  has values that are orders of magnitude different, such as when  $CE_d$  may be positive or negative. When  $\lambda$  is one, we seek a discretized distribution that yields the minimum worst case error. When  $\lambda$  is zero, we seek a discretized distribution that yields the minimum average error over the distribution of decision problems  $\mathcal{D}$ .

### 4. A tractable discretization instance

In order to determine an optimal discretization, we need to formulate a model that is solvable by available software in a reasonable amount of time. In this section, we give a specific and tractable instance of the more general discretization problem described in the previous section. We do this by defining a specific set of discretized probability distributions  $\mathbb{P}$ , a specific set of problems  $\mathbb{D}$ , and a probability distribution  $\mathcal{D}$  over the problems. The optimal choice of  $p \in \mathbb{P}$  defines the optimal discretization. These definitions allow us to formulate the discretization problem as a tractable, though non-linear, integer program (NLIP).

In this section we discuss our process for creating a tractable model. In our process we have a challenge that prevents us from formulating the model as we envision in Section 3. We discuss how we solve this challenge. Finally, we find there are some assumptions we can relax and provide a second tractable formulation. Solving a tractable discretization instance requires defining the objective value, decisions, and constraints with data such that the engines that solve the model are able to solve it within a given time.

### 4.1. NLIP formulation of discretization problem

The successful implementation of our NLIP requires we pre-calculate  $CE_d$  for each  $d \in \mathbb{D}$ . We estimate the  $CE_d$  values through simulation. Specifically, we sample the uncertainties from the distribution combinations defined by  $d \in \mathbb{D}$ . From the simulation we compute uncertainty values, a project value distribution, expected utility, and a certain equivalent.

To formulate optimization (1) as a tractable NLIP, one key obstacle to overcome is that the objective  $Err(d, p)$  is a non-linear function in  $p$  as shown in (2). In this formula,  $CE_d$  is already a constant we obtained from our formulation, but  $CE_d(p)$  is our calculated CE. The formula for  $CE_d(p)$  is given by

$$\text{CertainEquivalent: } CE_d(p) = -\rho_d \cdot \ln \left( - \sum_{p \in \mathbb{P}} prob_p \cdot u_d(p) \right), \quad (4)$$

$$\text{with exponential utility: } u_d(x) = -\exp(-x/\rho_d), \quad (5)$$

$$\text{and } P(X \leq x) = p \quad (6)$$

where  $\rho_d$  is the risk tolerance,  $prob_p$  is the probability assigned to percentile combination  $p$  and  $u_d(p)$  is calculated expected utility for decision problem  $d$  with percentile combination  $p$ . We use the exponential utility function in this example. Different utility functions will change the formulations for both  $u_d(p)$  and  $CE_d$ . Normally utility is expressed in terms of the value of the project/decision,  $x$ . We are searching for the optimal percentiles, so we used (6), the definition of the CDF, to relate  $p$  to  $x$ .

When we formulate a non-linear IP by using formula for CE given by (4) directly, our non-linear solver, Bonmin 1.8.4, we found it to be intractable. To create a tractable formulation we linearize the objective function. We choose to linearize  $Err(d, p)$  around the expected utility,  $E[u_d(x)]$ . We first substitute the equation for certain equivalent, (17), into the definition of  $Err(d, p)$  to get

$$Err(d, p) = \left| 1 - \frac{-\rho_d \ln(-E_p[u_d(x)])}{CE_d} \right|, \quad (7)$$

where  $E_p[u_d(x)]$  is the expected utility of the decision problem under the new discretized distribution  $p \in \mathbb{P}$ . Given a  $d \in \mathbb{D}$ , the only variable in the above formula is  $E_p[u_d(x)]$ , and everything else is a constant. We choose to linearize this equation, viewed as an equation in the variable  $E_p[u_d(x)]$  around the true expected utility  $E[u_d(x)]$ . We choose to linearize because this is easier for most solvers to solve. An extra Taylor expansion term can be added to improve the accuracy of the approximation at a cost of additional solve time. For brevity let  $T_d$ , our target utility, denote  $E[u_d(x)]$ .

To compute a linearization, we first drop the absolute value sign, assuming the second summand of (7) is smaller than one. This gives

$$f(w) = 1 - \frac{-\rho_d \ln(-w)}{CE_d},$$

which is now a continuous function of  $w$ . We create  $w$  as a shorthand for the variable  $E_p[u_d(x)]$ . We can now do a first order Taylor expansion of this function around  $T_d$  to get

$$\begin{aligned} f(w) &\approx 0 + f'(T_d)(w - T_d) \\ &= \frac{-\rho_d}{CE_d \cdot T_d} (w - T_d). \end{aligned}$$

This approximation is valid when the second summand of (7) is smaller than one. This happens when  $w$  is greater than or equal to  $T_d$ . A similar argument, assuming that the second summand is greater than one, yields the approximation  $\frac{\rho_d}{CE_d \cdot T_d} (w - T_d)$ , which is valid when  $w$  is smaller than or equal to  $T_d$ . Together, these two linearizations can be summarized as

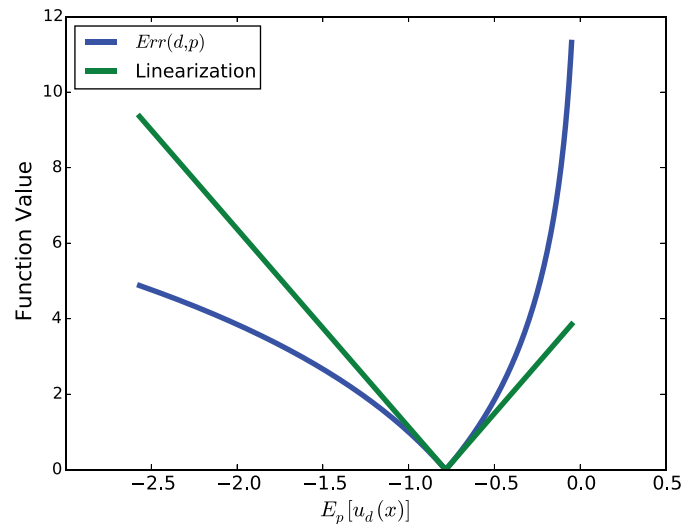


Fig. 2. This is a sample of a linearization of the absolute percentage error and the true absolute percentage error as a function of the expected utility. If the results of the discretization are close enough to the true value, the linearization of the error function is sufficient.

$$\delta_d = \frac{-\rho_d}{CE_d \cdot T_d} \quad (8)$$

$$Err(d, p) \approx \delta_d \cdot |E_p[u_d(x)] - T_d|, \quad (9)$$

which is a linearization of Eq. (7). Fig. 2 plots an example of the true error function and corresponding linearization. In the case where the decision maker is risk neutral, we can skip the calculation of  $\delta_d$  (8) and just use  $\delta_d = \frac{1}{CE_d}$ .

With this linearized objective function, we can now write an IP for computing an optimal discretization as follows.

#### Indices and sets

- $i \in I$  :the set of uncertainties
- $v^i \in V^i$  :the set of percentile discretization of uncertainty  $i$ . These are candidate percentiles for the uncertainty, such as {5, 10, 15, 45, 50, 55, 90, 95}.
- $\mathbf{v} \in \otimes V_i$  :a percentile combination for all uncertainties.  $\mathbf{v}$  is a vector of length  $|I|$ .
- $d \in \mathbb{D}$  :a finite, discrete set of decision problems
- $j \in J$  :the indexes for each of the  $|J|$  incompatible sets discretizations

#### Parameters

- $\lambda$  :used to compute a convex combination of average and maximum error
- $T_d$  :the true expected utility for decision problem  $d$
- $\delta_d$  :a shorthand for  $\frac{-\rho_d}{CE_d \cdot T_d}$ , a constant used in linearization
- $N_i$  :the maximum number of percentiles per uncertainty for the output discretization
- $U_d(\mathbf{v})$  :the utility of the project value for decision problem  $d$  and at percentile combination  $\mathbf{v}$
- $\delta_j$  :the incompatible discretizations,  $\mathbf{v}$ , in set  $j$
- $P_d$  :the probability assigned to decision problem,  $d$

#### Decision variables

- $P_{\mathbf{v}}$  :the probability assigned to a combination of percentiles  $\mathbf{v}$ .
- $o_d$  :the over-estimation in approximating  $T_d$  with a discretized probability distribution
- $u_d$  :the under-estimation in approximating  $T_d$  with a discretized probability distribution
- $z$  :the estimated  $\max_{d \in \mathbb{D}} Err(d, p)$

$x_{v^i}$  :the probability assigned to candidate percentile  $v^i$  for uncertainty  $i$   
 $y_{v^i}$  :1 if percentile  $v^i$  is used for uncertainty  $i$  and 0 otherwise

**Formulation**

$$\min \lambda \cdot z + (1 - \lambda) P_d \sum_{d \in \mathcal{D}} P_d \cdot \delta_d \cdot (u_d + o_d) \tag{10a}$$

$$\text{s.t.} \sum_{\mathbf{v} \in \otimes V_i} U_d(\mathbf{v}) p_{\mathbf{v}} - o_d + u_d = T_d \quad \forall d \in \mathbb{D} \tag{10b}$$

$$z \geq \delta_d \cdot (o_d + u_d) \quad \forall d \in \mathbb{D} \tag{10c}$$

$$\sum_{v^i \in V^i} x_{v^i} = 1.0 \quad \forall i \in I \tag{10d}$$

$$\sum_{v^i \in V^i} y_{v^i} \leq N_i \quad \forall i \in I \tag{10e}$$

$$\sum_{v^i \in \delta_j} y_{v^i} \leq 1 \quad \forall i \in I, \forall j \in J \tag{10f}$$

$$x_{v^i} \leq y_{v^i} \quad \forall i \in I; v^i \in V^i \tag{10g}$$

$$p_{\mathbf{v}} = \prod_{v^i \in \mathbf{v}} x_{v^i} \quad \mathbf{v} \in \otimes V_i \tag{10h}$$

$$0 \leq x_{v^i} \leq 1 \quad \forall i \in I; v^i \in V^i \tag{10i}$$

$$y_{v^i} \in \{0, 1\} \quad \forall i \in I; v^i \in V^i \tag{10j}$$

$$o_d, u_d \geq 0 \quad \forall d \in \mathbb{D} \tag{10k}$$

Objective (10a) of the optimization model is to minimize a convex combination of the largest error  $z$  and the average error. In this formulation we show the generalized distribution on  $\mathcal{D}$ . The second term of the objective function is the average error. In order to compute the linearized error (9), we should compute the absolute value of the difference between  $T_d$  and  $E_p[u_d(x)]$ . The formula for  $T_d$  is given for formula (18) in Appendix A.1. Constraint (10b) computes the difference between the target and expected utility. Constraint (10c) computes the maximum error,  $z$ . Constraint (10d) forces the sum of the probabilities for each uncertainty to sum to one. Constraint (10e) limits the number of percentiles allowed for each uncertainty. Constraint (10f) forces only a single low, a single medium, and a single high percentile in our discretizations. This helps the optimization engine find a solution faster. For example, our low-percentile candidates are P5 and P10. Only one may be selected for the discretization. Constraint (10g) forces the assigned probability to zero if the percentile is not used in the discretization. Constraint (10h) computes the probability assigned to a percentile combination as a function of the probabilities of each of the uncertainties. This is the only non-linear constraint in the formulation and it enforces that the output distribution  $p \in \mathbb{P}$  is independent over the uncertainties. The remaining constraints bound the probability values between 0 and 1, make the indicator variables binary, and make the underage and overage non-negative.

4.2. Joint discretization problem

A joint discretization differs from an independent discretization in that the probability of a percentile combination is assigned individually, and is not the product of the probability assigned to each percentile in the percentile combination. For a joint distribution, we assign the probability to each percentile combination separately. This relaxation increases the flexibility of the values assigned to a percentile combination and it linearizes the model formulation. Previous discretization techniques only considered uncertainty discretizations independently. Because we consider a set of decision problems and

compute the best discretization for that set of problems, it is possible to compute this joint discretization.

The feasible solutions for a joint discretization include the feasible solutions for independent discretizations. As a result, computing an optimal joint discretization is guaranteed to produce at least as good a result as computing an optimal independent discretization. Furthermore, computing an optimal joint discretization has the potential to decrease solution times, as it removes the non-linearity in Model (10). By removing the non-linearity, we could also switch solvers from Bonmin to CPLEX. In this section, we compute such optimal joint discretizations.

We alter Model (10) as follows to compute optimal joint discretizations. The formulation drops variables  $x_{v^i}$  and any constraints where they appear. These are Constraints (10d) and (10h). We also add the following constraints:

$$\sum_{\mathbf{v} \in \otimes V_i} p_{\mathbf{v}} = 1 \tag{11a}$$

$$p_{\mathbf{v}} \leq y_{v^i} \quad \forall \mathbf{v} \in \otimes V_i, v^i \in \mathbf{v} \tag{11b}$$

$$p_{\mathbf{v}} \geq 0. \tag{11c}$$

Constraints (11a) and (11c) ensure the variables  $p_{\mathbf{v}}$  compute a joint probability. Constraint (11b) ensures the support of that joint probability is limited to the  $N_i$  percentiles for each uncertainty  $i$ . The result of Constraint (11b) is that experts make the same number of assessments as before.

5. Analysis

In this section we solve Model (10) for a sample problem given by Smith [19]. We briefly describe the example here and more in depth in Appendix A.1. We also apply the methodology to a second problem originally given by Clemen [4] and expanded by Reilly [18] and further described in Section A.2. We begin with the Smith [19] wildcatter problem.

A wildcatter is a person that drills for oil in a previously undeveloped field, and will not know exactly what he or she will find. In this simple model there are four uncertainties that impact the value of the well: the oil price at which the wildcatter can sell, the amount of oil in the well, the amount of oil the wildcatter can recover, and the cost of the recovery. The exact distribution of the each uncertainty is unknown. Rather than solving the problem with the functional forms used by Smith [19], we use several candidate distributions as shown in Fig. 3. Any one of the candidate distributions could be the true distribution. We use nine distributions per uncertainty resulting in a total of 6561 decision problems, any of which is equally likely to be the true problem. In this formulation we assume the risk tolerance is known at the time of the problem definition. The objective of optimal discretization is to find the discretization that minimizes that error over all the problems.

For each of the 6651 decision problems we use Latin hypercube sampling as originally described by McKay et al. [14]. We generate 4,000,000 values for each uncertainty to generate a set of present values,  $\mathbb{X}$ . For each  $x \in (X)$  we generate a utility and determine  $CE_d$  using (17) with a risk tolerance value,  $\rho = \$16,000,000$ . The distribution of the  $CE_d$  is found in Fig. 4. From each  $CE_d$  we are also able to obtain a target utility,  $T_d$ , using (18). For each decision problem we also calculate  $\delta_d$  using (8).

The set of percentile combinations is drawn from a Cartesian product of the candidate percentile for each uncertainty. We define  $\otimes V_i$  as the set of potential percentile combinations. We allow each of the four uncertainty percentiles to be in the set {5, 10, 45, 50, 55, 90, 95}. This represents the set of potential assessment values we might ask an expert to give for each uncertainty. This set is neither inclusive, nor exhaustive. These are illustrative choices, and a decision analyst may add

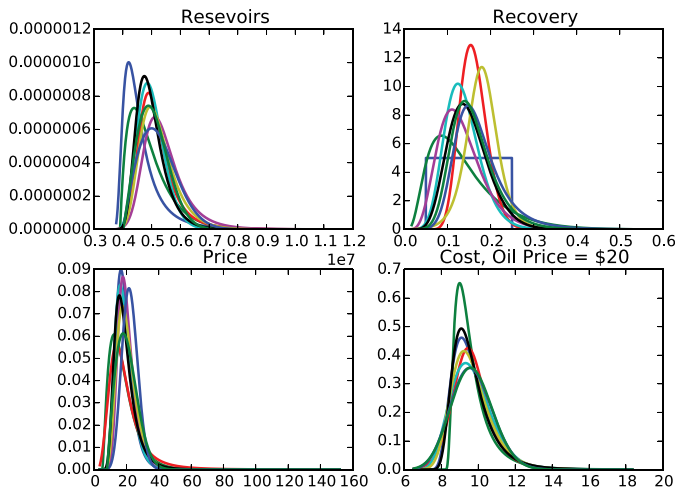


Fig. 3. Each uncertainty has nine candidate distributions. The reservoir, price, and cost distributions are bounded from below at zero. The recovery distribution is bounded by 0 and 100%. Though most distributions are similar in shape, we also included a uniform distribution in as potential distribution for the fraction of the reserves that may be recovered.

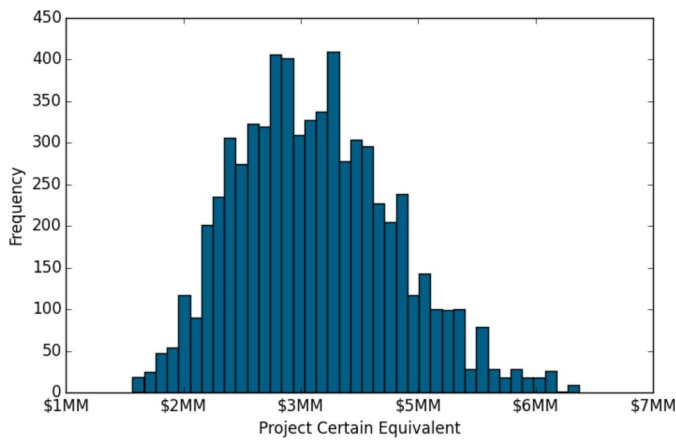


Fig. 4. The distribution of CE values for the decision problems created from all the combinations of potential distributions for each uncertainty and a single risk tolerance.

or remove percentiles. An increase in candidate percentiles may improve accuracy, and it will increase computational complexity. For each decision problem we have 2401 potential percentile combinations. we

will choose  $3^4 = 81$  of those percentile combinations and assign them probabilities to create our optimal discretization. For this discretization instance, we are defining the distributions in  $\otimes V_i$  as independent over the uncertainties. For each of the  $2401 p \in \mathbb{P}$  we calculate the utility for each decision problem  $d \in \mathbb{D}$  using (16) to calculate  $U_d(v)$  for each  $v \in \otimes V_i$ . This provides the data we need to populate our optimization model.

We begin our comparison of optimal discretization to four incumbent distributions. These are the MCS, ESM, EPT, and HB methods. These discretizations each use three percentiles and assign probabilities. With four uncertainties in the problems, this yields 81 potential outcomes for each decision problem. We use the percentile from each discretization to get a value from the decision problem’s uncertainty distributions inverse CDF. We compute the project value and utility based on the samples. Finally, we compute the CE using the probabilities assigned to each percentile. This gives us an estimated CE for each decision problem. We compare the estimated CE using the discretization to the CE we obtained by using the simulation for the same problem using the equation

$$100 * \frac{CE_d - CE_d(p)}{CE_d} \tag{12}$$

We create a distribution of errors for each discretization method and present them in Fig. 5. The HB and EPT methods use more extreme percentiles like the 5th and 95th percentiles. The MCS and EPT discretizations use the 10th and 90th percentiles. The accuracy of the discretizations with more extreme values is visible in Fig. 5. We use two measures of accuracy. The first is the worst-case error. This is the largest absolute value of a percent error from the true CE across all decision problems. The other error metric is the average of the absolute errors. When we measure the worst case error, HB has a worst-case of 1.75% and EPT has a worst case of 1.47%. ESM has an absolute worst case error of 2.02% and MCS has a worst case of 6.34%. The mean absolute errors of HB and EPT are both 0.25%. HB has a slightly better performance in terms of absolute error, but when rounded to the nearest hundredth of a percent, they are the same. MCS and ESM have average absolute errors of 1.30 and 0.41%, respectively. In the wild-catter example, the standard deviation of  $Err(d, p)$  is larger and the mean CE is further away from 0 when the discretizations using the extreme (5th and 95th) percentiles is used.

In this section, we solve Model (10) using two parameters for  $\lambda$ . When we set  $\lambda = 1$ , we minimize the worst case error, when we set  $\lambda = 0$ , we minimize the average error.

5.1. Independent discretization

In creating optimal discretizations we have two goals in mind. The

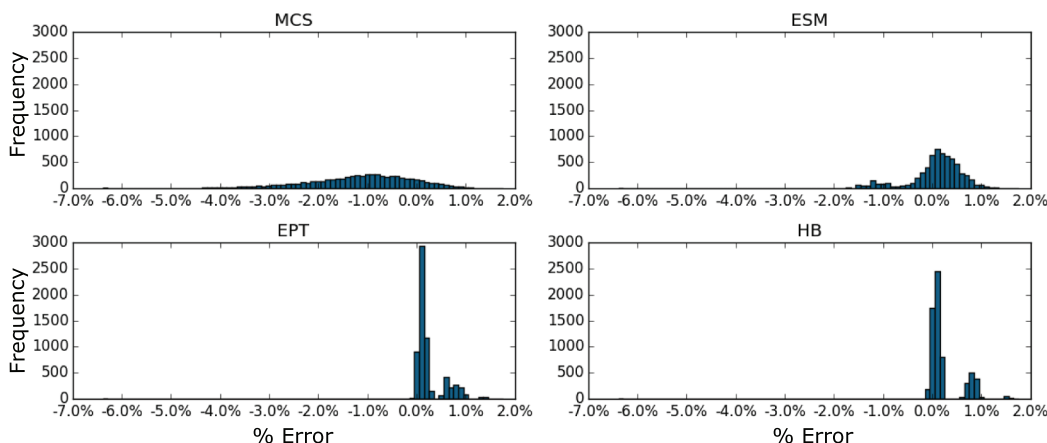


Fig. 5. The distribution of percent errors for four shortcut methods.

first goal is to find discretizations that minimize (10a). The second goal is to find this solution quickly. We define quickly rather loosely. If this is being done for an ongoing project, we want to be able to generate an optimal discretization for the client between sessions; Otherwise, we want to have the discretizations computed for the next time a decision problem comes up.

We solved the complete model, with all the candidate percentiles for each decision problem. We also solved different versions of problem (1) using subsets of the candidate discretizations or using a subset of the decision problems. By limiting the candidate discretizations, we are able to reduce the number of variables. Specifically, when we reduce the number of candidate percentiles to three, we are able to solve the model as a continuous problem instead of as a non-linear mixed-integer problem. Our candidate percentiles for the full problem are  $P_5, P_{10}, P_{45}, P_{50}, P_{55}, P_{90}, P_{95}$ . Some of the most popular discretization methods use either  $P_{10}, P_{50}, P_{90}$  or  $P_5, P_{50}, P_{95}$ . We try these two sets of three plus the full set. Our second way of reducing the computational time is to reduce the number of decision problems by sampling them. We test how the results differ when we chose to minimize the worst-case discretization error and when we try to minimize the average discretization error.

5.2. How much benefit do we get from optimizing an independent discretization?

The calculation of the certain equivalent is given by multiplying the

probability of each percentile of each uncertainty to determine the probability of an outcome. There may be a covariance among the resulting values, but the percentiles are treated as independent. In the case of the four uncertainties in our sample problem, the probability of any one outcome is the product of the probability of each of the individual uncertainties. The drawback of the nonlinear approach is that there are few available solvers, and large problems generally take too long to solve. For example, in our test problem, solving the full problem with  $\lambda = 0.0$  using Bonmin 1.8.4 using an Intel 6-core I7 processor running at 2.6 GHz, the average time to generate the model and solve the problem was 129,937.72 s (1.5 days). This problem has 6651 decision problems and over 15,000,000 nonzeros.

The results from of the optimization are shown in Fig. 6. When comparing to HB, which has the best results in Fig. 5, independent discretization improves the worst case mean error and the standard deviation of error.

5.3. How much do we lose by solving a smaller sample of decision problems?

Given the intractability of solving for 6651 decision problems with 20,736 possible combinations of percentiles for the uncertainties, we tested the effects of sampling the decision problems. In sampling the decision problems, we randomly select a subset of the decision problems and then applied the optimal discretization for that problem to the entire set of problems. The results are displayed in Fig. 7. We test the sampling with 1, 5, 10, and 20% of the 6651 decision problems. We

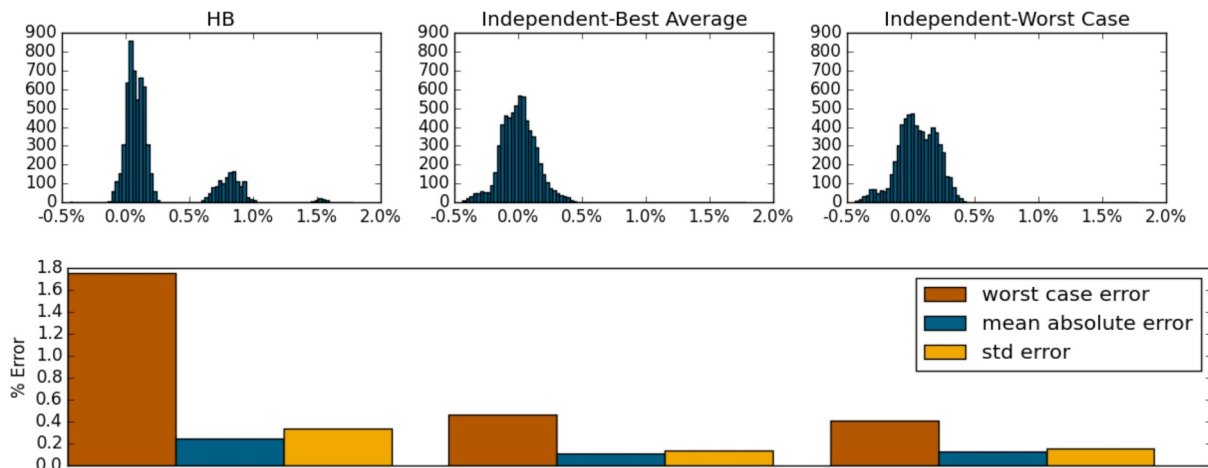


Fig. 6. The histogram of results compares HB, which has the best average error of the incumbent methods with the optimized discretizations using both the optimal average and optimal worst-case preferences. The optimized results show a reduction in average absolute error of 56% and a reduction of worst-case error of 74%.

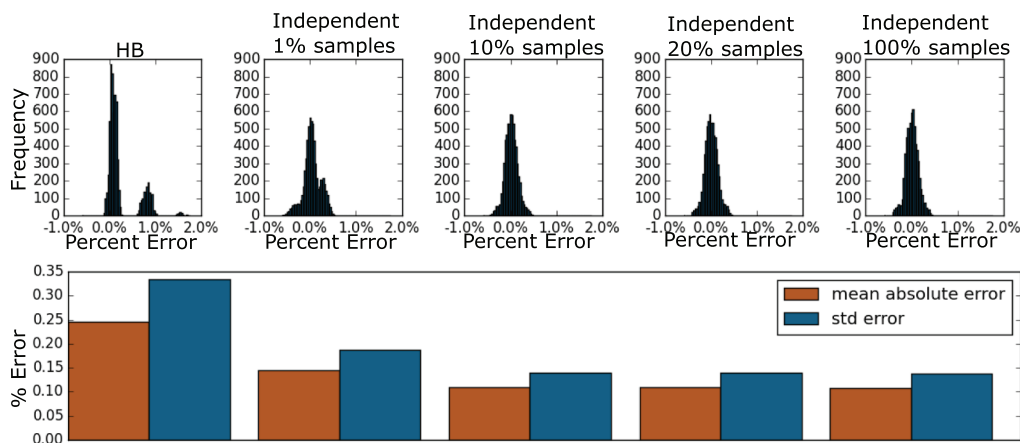


Fig. 7. As the number of samples increases in percentage, the overall accuracy of the discretization improves. After 10% of the samples, the average error is 1% worse than the optimum using the entire set of decision problems.



test this method using the worst case and the best average objectives. The solve time increases roughly linearly with the number of decision problems that we sample. Sampling with 1% took 24 min, sampling with 10% took just under 3 h, and sampling with 20% took just under 6 h. Solution quality, as measured as the increase in objective value from the optimal value with 100% sampling improves with the number of samples, but not linearly. Sampling with 1% results in a 34% increase in the average error, sampling with 10% results in an increase of 1% in the average error, and sampling with 20% results in an increase of 1.4% in the average error. The increase in average error in the sampling is likely due to the randomness of the sampling. When we looked at the solution quality, there is a noticeable difference between choosing  $\lambda = 1$  and  $\lambda = 0$ . For the smaller samples ( $< 20\%$ ), minimizing the worst case led to varying degrees of over fitting, with increases in worst case error of 25, 15.7, and 6% for the 1, 10, and 20% samples respectively. Sampling the decision problems results in roughly linear speedups in performance with a small loss of accuracy.

5.4. How much do we lose by restricting the candidate percentiles?

Some of the most common discretizations use either P10, P50, P90 or P5, P50, P95. In comparing both the shortcut methods and the discretization results, it seems the most accurate discretizations come from using the more extreme percentiles. If P5, P50, P95 are more accurate, it may be better to only consider these percentiles and optimize the probabilities. Given the P10, P50, P90 discretizations are also popular, we also want to know what improvement in accuracy we can expect when considering the more extreme percentiles. We solve the problem

using either the maximum error or the average error objectives. The first improvement is the rapid speedup in solution time. The range of reduction is from 99.5 to 99.9% reduction in the time required to generate a discretization. The solution times were in the 100–300 s range. When limiting the candidate discretizations to P10, P50, P90, the mean absolute error is 4.8% lower than the mean absolute error using the HB shortcut. It should be pointed out this slight improvement comes using less-extreme values than those required by HB. In comparison to using all the candidate percentiles from a full optimization, the mean absolute error is still 117% worse when using P10, P50, P90. These results are shown in Fig. 8. Using the P5, P50, P95 percentiles improves the accuracy of the optimal discretization while solving quickly. The optimized discretization increases worst case error by just 0.25% over the optimal results obtained from considering all the percentiles. This result is shown in Fig. 9.

5.5. What benefit do we derive when we remove the independence of uncertainties?

The results in solving the problem with a nonlinear solver have mixed results. Though the improvement in accuracy is substantial, some instances take days to solve. In a large business problem with 12 or 15 uncertainties, the size of the problem could become intractable. Previous discretization methods focused on individual uncertainties, which were combined to create a distribution of the decision problem values. We propose a new approach which relaxes the independence of uncertainty percentiles and creates a joint distribution.

Joint discretization improves both performance time and the

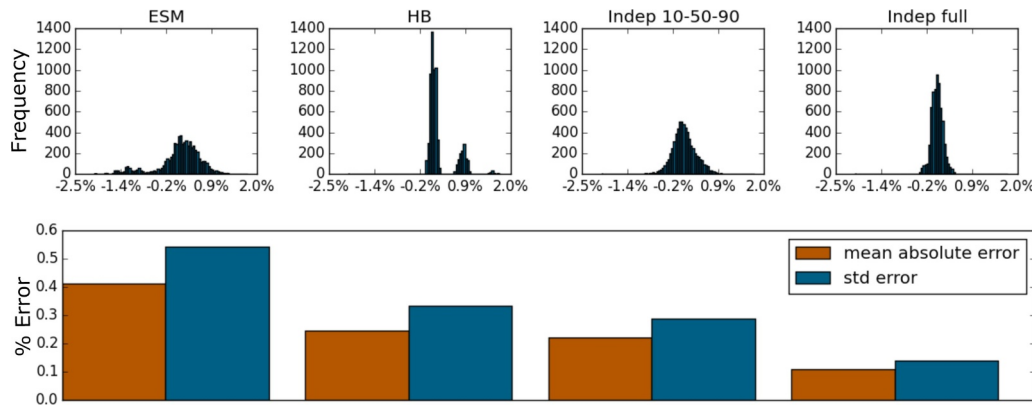


Fig. 8. Limiting the candidate percentiles to 10-50-90 reduces the error in comparison to the shortcut methods, and is worse than when considering a larger assortment of candidate percentiles.

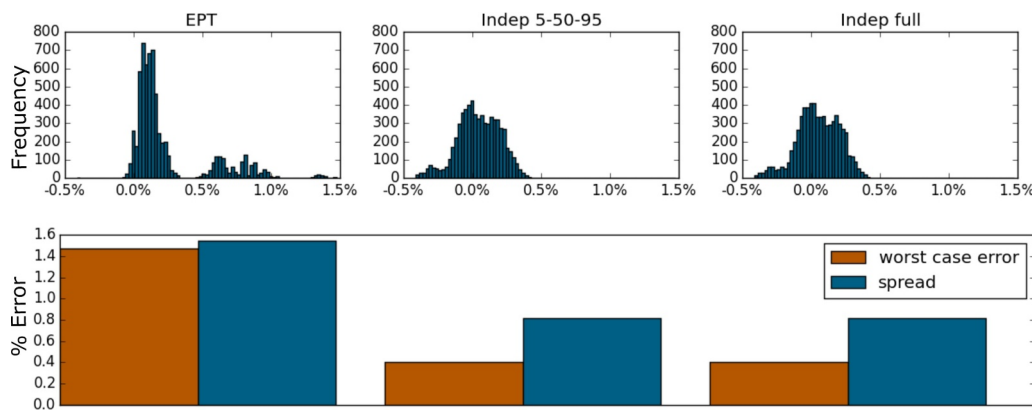


Fig. 9. Limiting the candidate percentiles to 5-50-95 reduces the error in comparison to the shortcut methods, and is worse than when considering a larger assortment of candidate percentiles.



Fig. 10. A comparison of the independent and joint discretization method results.

accuracy of the discretizations. The solution time using CPLEX 12.5 is just under 2 h for the mean absolute error, and about 1 h and 40 min for the worst case error. This compares favorably to the 1.5 and 1.2 day solution times for the independent discretizations. The joint discretization also reduces the mean absolute error by 34% in comparison to the independent discretization. In comparison to the shortcut methods, this is a 71% reduction in mean absolute error of the best-performing shortcut (HB). For the worst case error, the joint discretization reduces the error by 41% when compared to the independent discretization, and it reduces the worst case error by 86% when compared to the best shortcut method (EPT). These results are visible in Figs. 10 and 11.

Joint discretization has another benefit over independent discretization. As seen in Appendix A.3 a joint discretization does not use every possible percentile combination. For both a practitioner and a client, this means there are fewer assessments required. For example, if there are two correlated uncertainties for a decision problem and three discretized values for each uncertainty, then 12 assessments are required. First, the first uncertainty must be assessed at its low, base, and high values, and then the correlated uncertainty must also be assessed for its low, base, and high values for each value of the first uncertainty. In contrast, joint discretization often omits certain percentile combinations, reducing the number of assessments.

### 5.6. What is the value provided by optimal discretization?

In order to determine the effectiveness of optimal discretization, we determine how much of a boost in CE do we expect to get from using

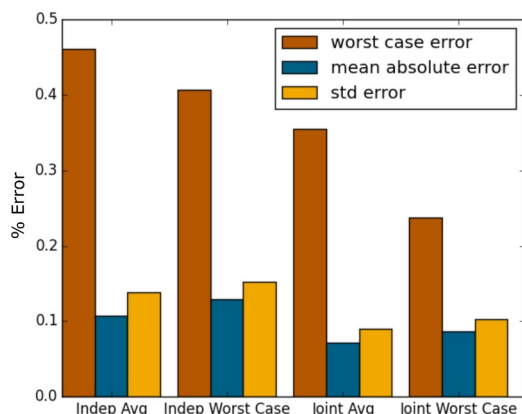


Fig. 11. A comparison of the independent and joint discretization method results.

optimal discretization instead of shortcut discretizations. We define our value based on whether the decision changes based on the results of the discretization. If the results of two discretizations both indicate that the company should initiate a project, their value is the same because the decision is the same. In this section, we modify the problem in order to induce an increase in different decisions and compare the results.

We adjust the initial capital required in Eq. (15) so that the median CE is now zero. In half the decision problems, the best decision is now to pass on the project. The histogram of the project values is the same as in Fig. 4 but shifted lower by \$3.50MM. With several decision problems having CEs near zero, we modify the equation for  $Err(d, p)$  to be Eq. (3). This changes the formula for the error approximation equation, Eq. (9), to

$$Err(d, p) \approx \frac{-\rho_d}{T_d} |E_p[u_d(x)] - T_d|. \tag{13}$$

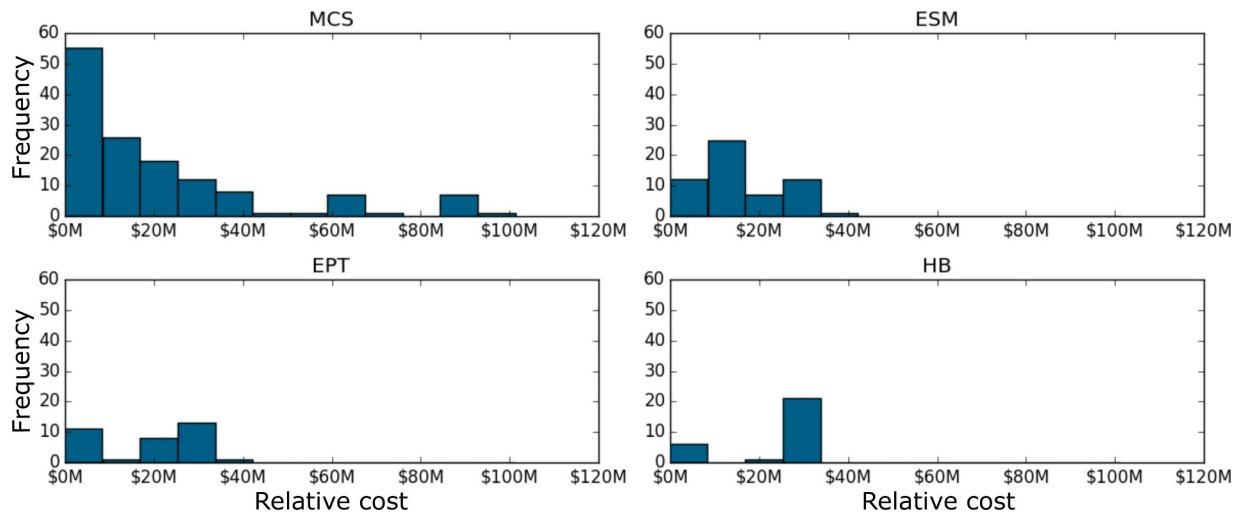
Using a new values for  $\delta_d$  in Model (10), we solve the same set of models again to obtain optimal discretizations. For each discretization we determine the additional value derived from knowing the true distributions of the uncertainties as opposed to using the discretizations of the uncertainties. From our initial Monte Carlo integration, we determine the  $CE_d$  of each decision problem. We compare  $CE_d$  to the  $CE_d(p)$  given by the discretization. In our sample problem the two strategic options are to initiate the project, or to not initiate the project. The outcomes from the discretizations are to correctly initiate or pass on the project, or to incorrectly initiate or pass on the project. We define relative cost (RC) as the expected additional cost of using a discretization instead of knowing the functional form of the uncertainties. For each decision problem  $d \in D$ ,  $RC_d$  is the mean absolute value of the  $CE_d$  when the wrong decision is made due to the discretization and zero otherwise. For example, when the true CE is 100, and the discretized CE is negative, the value of having the true CE is 100. When the true CE is 10,000, and the discretized CE is 1, both CE values will recommend initiating the project. In this case, the value of knowing the true CE is 0. The relative cost of the discretization for a decision problem,  $d$  is as follows:

$$RC_d = \begin{cases} CE_d & \text{if } CE_d > 0 \text{ and } CE_d(p) < 0 \\ -CE_d & \text{if } CE_d < 0 \text{ and } CE_d(p) > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

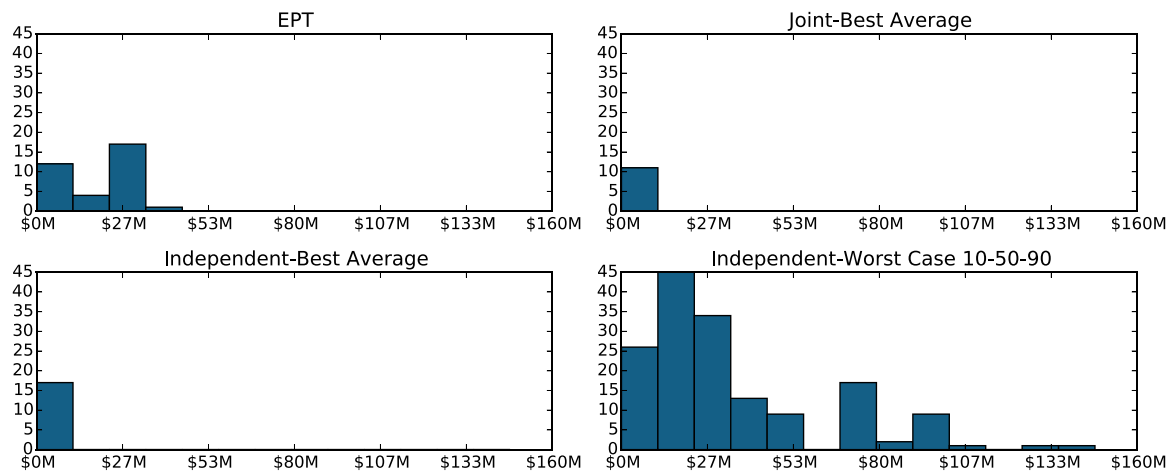
over all the decision problems. The discretization with the lowest relative cost is the discretization where the decision from using the discretization matches the decision that would come from knowing the functional forms of the uncertainties the most. The higher the RC, the worse a discretization is in terms of value. We can compare the average RC for the different discretizations to determine how much additional value one method has over another.

We begin by comparing the RC for the shortcuts. Fig. 5 indicates the MCS shortcut tends to have the largest RC values. Among the shortcuts, this produces the largest RC. Again, EPT and HB perform better than MCS and ESM. This time, the average RC for ESM is only 35.28% worse than EPT. This compares to the average error being about 67.08% worse than HB. In absolute terms, the additional value provided by EPT over ESM is \$35.48, which for a project with an average CE of \$52, 642.89 is only 0.07%. It is likely different decision problems with a more strategic options and a larger range of project values will result in larger RC differences (Fig. 12).

The optimized discretizations showed a surprising range of RC as seen in Fig. 13. At one extreme, the optimizations using average error had RC values of \$1.62 and \$4.04 for the joint discretization and independent discretization respectively. This means knowing the functional forms of the uncertainty distributions provides almost no value above using a discretization (as long as the assessments are accurate). At the other extreme, the optimizations using worst case error performed substantially worse than the shortcuts. The RC for worst case



**Fig. 12.** The distribution of the relative cost (of not knowing the true distributions of the uncertainties) for four shortcut methods. The value of knowing the true distribution for most of the decision problems is \$0. This means most of the time, the discretization is on the right side of 0. In some cases, as with MCS, the relative cost can be as high as \$100, 000. Note: there were a large number of observations at zero, which were removed to better visualize the remaining observations.



**Fig. 13.** The distribution of the relative cost comparing EPT with the results from optimal average error for joint and independent discretizations and optimal worst-case independent discretization using only the 10th, 50th, and 90th percentiles. The worst case optimization has the most decision problems and highest RC of any discretization we test. For this discretization, the RC of one of the decision problems is over \$140, 000. Note: Each distribution has a large frequency of values at zero that we have removed to better show the scale of the non-zeros.

errors were \$695.30 and \$201.09 for the joint discretization and independent discretization respectively. The reason behind this complete flip in performance is that minimizing worst case error tends to focus on the most extreme-valued decision problems. For the joint discretization, the true CEs of the decision problems where the optimal discretization leads to the wrong decision, has a range between  $-\$72, 903$  and  $\$79, 812$ . The independent discretization has a range between  $-\$72, 903$  and  $\$5884$ .

Sampling the decision problems and limiting the percentiles yields similar results as compared to the original decision problems and error function. The general exception is that worst-case optimization underperformed its best average counterpart (in terms of samples and allowed percentiles). In six out of 30 runs minimizing the worst case had a lower RC than minimizing the average error. The best-performing methods used the more extreme percentiles. Using more samples typically resulted in better results, but not always. For instance, the best RC came from solving the joint discretization optimization using 20% of the decision problems and the best average. It yields a RC of only \$0.52. This is a result of serendipitous sampling. The worst result comes from optimizing for the worst case, maintaining independence of

uncertainties, and using 10th, 50th, and 90th percentiles. This discretization had a RC of \$840.59.

*5.7. How well do the discretizations work with new uncertainty distributions when applied to the original problem?*

So far this method has performed extremely well when optimized against a training set of distributions. We call it a training set as it is used in the same way as a training set is used in predictive analytics. In predictive analytics set of data is used to generate model parameters; in our case these are the discretizations. The results of the first model are tested against another set of data in order to determine if the model works for the entire set of data. It can also be noted that if the problem to be solved is from one of a potential set of distributions, the decision analyst can estimate the CE of every uncertainty distribution combination and come up with a distribution of the CE. In a situation like this, the process of optimization does not help the decision-making process. In practice, all potential distributions for every uncertainty should be used in the optimization model.

To test the performance of the optimal discretization we change the

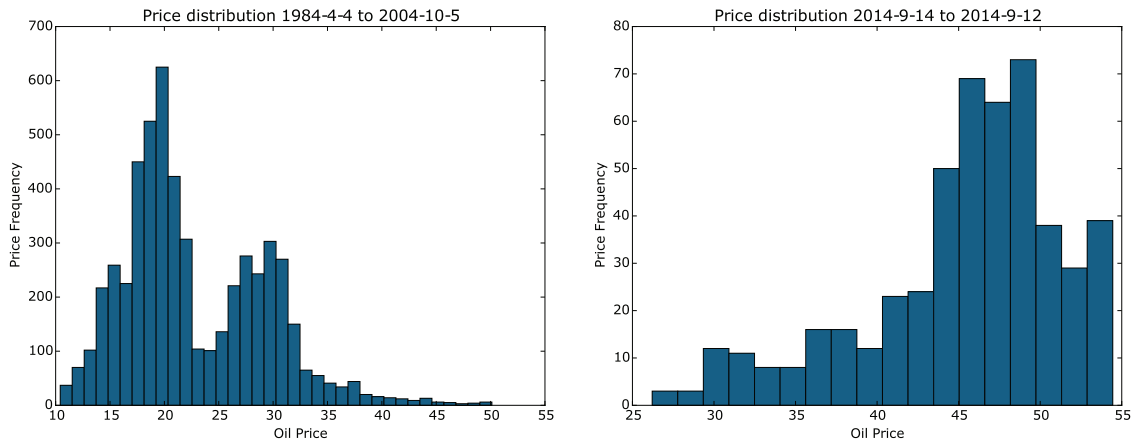


Fig. 14. The oil price distributions, when drawn from historical data, do not resemble any of the distributions we have used to train the model.

functional form of all the uncertainties. We begin by using the historical pricing of the West Texas Intermediate benchmark. We downloaded the prices from the United States Energy Information Administration for the front month Cushing, OK Crude Oil Future Contract available on the US Energy Information Administration web site. We used the reservoir and cost data distributions from the original Smith [19] paper, and we used a beta(3,27) distribution for the recoverable oil percentage. We chose this number to have a mean of 10% and would range between 1.5% and 26%. When comparing to the distributions in Fig. 3, this tends to be on the low side, but within the realm of the feasible.

Examining the price distribution in the original Smith [19] paper and in Fig. 3, we determined the oil price was somewhere between \$10 and \$50. The WTI price data begins on April 4, 1983, with a price of 29.44 and remains below \$50 until October 5, 2004. We use the daily closing price to populate our price distribution in our first example.

In a second test, we wanted to see if the methodology might also be applicable to shale drillers. In this test case, we used recent prices. We used the two years of price history, from September 14, 2015 until September 12, 2017. We also doubled the capital cost of drilling a well, and we doubled the production rate. Because we used historical data, our distributions as seen in Fig. 14 have their own shapes. Even the data pulled from a 20 year seems to be multi-modal.

In both examples, we estimate the CE using the Latin hypercube technique. We applied the discretization percentiles and probabilities that we generated previously. These are available for reference in

Appendix A.3. We chose the best average and worst case discretizations for all candidate percentiles as well as the 5, 50, 95, and 10,50, 90 optimized discretizations. It should be noted none of the uncertainty distributions in our new problem (the test set) were any of the distributions used to calculate the optimal discretizations (the training set). We find that without having the new distributions in the training set, some of the very best performing optimal discretizations from Sections 5.2 and 5.5 underperformed the shortcuts. We also find that the consistently best-performing discretization is an optimized discretization.

The results from this example indicate that simplicity is best. The results are shown in Fig. 15. The best-performing discretization is the independent discretization that discretizes using the 10 – 50 – 90 percentiles. In general, the optimized 10 – 50 – 90 shortcuts performed better using the new distributions in the example problems, while in the training sets, the optimized 5 – 50 – 95 discretizations performed better. In the first example, the mean is much further away from zero, so differences in percent error tend to be closer. In the second example, the mean is much closer to zero, and differences are greater. It should be noted that just as with the optimal discretizations, the shortcuts also vary in their performance between the two examples. In the first example, ESM has the best performance of all the shortcuts we test. MCS, which is also a 10 – 50 – 90 shortcut performs better than HB and EPT, which use more extreme values.

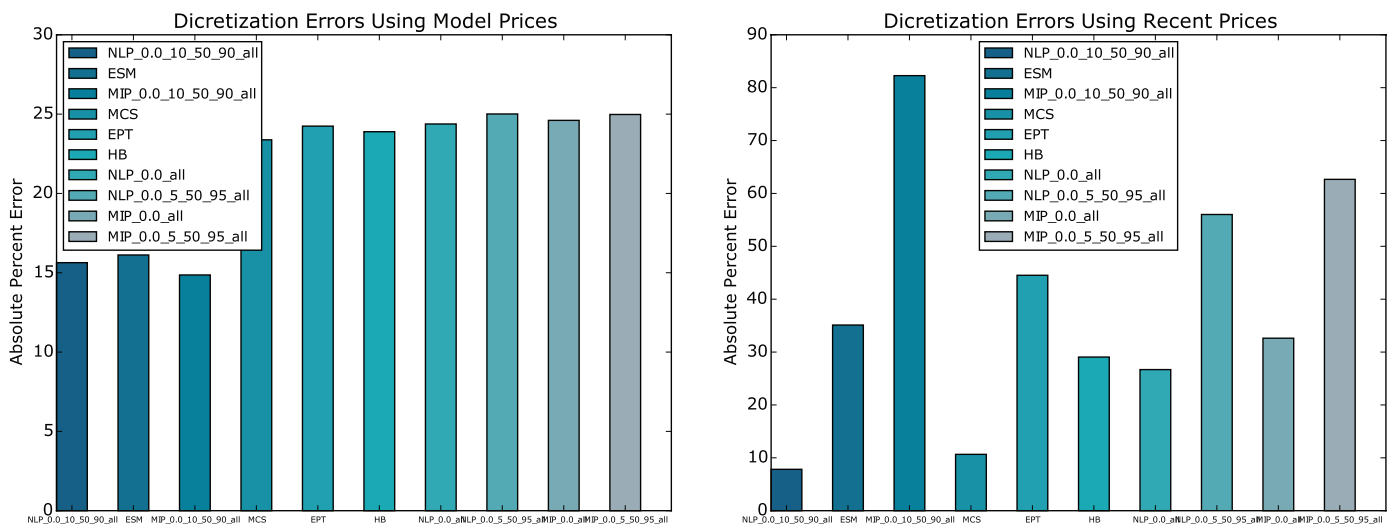


Fig. 15. West Texas Intermediate oil prices representing a twenty year history and oil prices representative of the oil prices during the fracking boom in the united states.

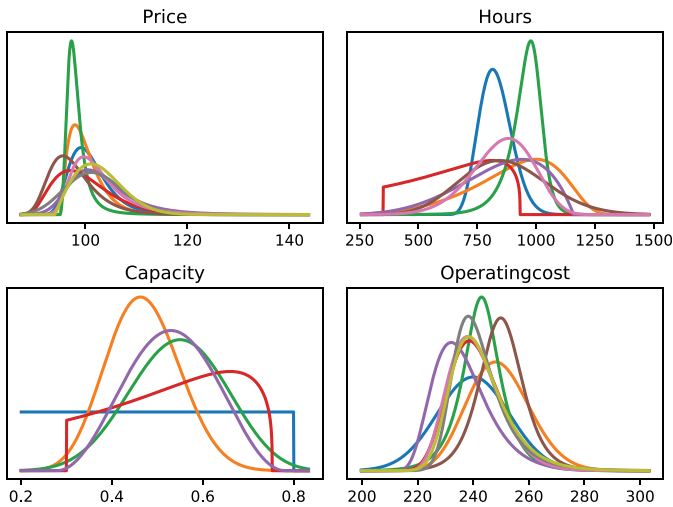


Fig. 16. The potential distributions for the four uncertainties of Eagle Airlines. None of these distributions is the true distribution of the given problem.

5.8. How well do the discretizations work with other problems?

So far the numerical analysis has focused on the Wildcatter problem introduced by Smith [19]. We now present a shorter analysis of Eagle Airlines, first introduced by Clemen and Reilly [5]. A short description of Eagle Airlines is given in Appendix A.2. For this problem we also have four important correlated uncertainties, price, hours, capacity, and operational cost that affect the value of purchasing an airplane for passenger and charter trips. For each of these uncertainties we created a set of potential uncertainties. These are shown in Fig. 16. With the Cartesian combination of each of these uncertainties we determined the expected value (risk neutral) of the purchase decision. The distribution of the expected value of the purchase is given by Fig. 17.

We solve for the independent discretizations using  $P_{10} - P_{50} - P_{90}$ ,  $P_5 - P_{50} - P_{95}$ , and free discretizations. We apply the resulting discretizations to the correlated uncertainties of the true distributions to determine the error of the optimized discretizations and the shortcut methods. These results are shown in Fig. 18, and we present the discretizations in Appendix A.3.

In this example, the best discretization for the training set that uses some potential distributions to generate the discretization uses  $P_5 - P_{50} - P_{95}$  for each uncertainty, with values similar to EPT. When we determine the error using various discretizations and using the true

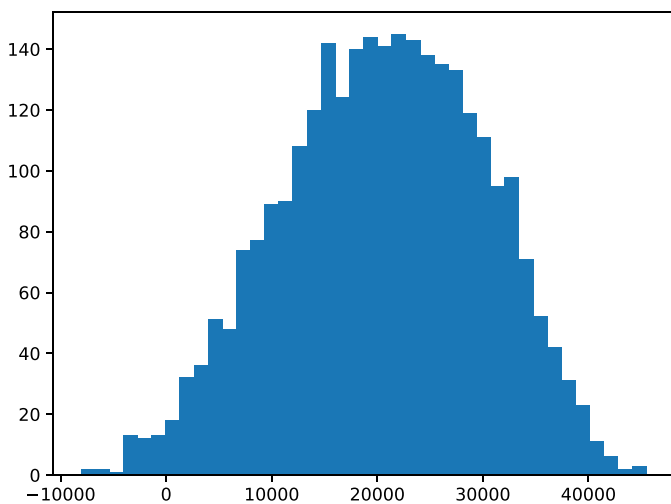


Fig. 17. The histogram of the expected value of the various uncertainty distribution combinations.

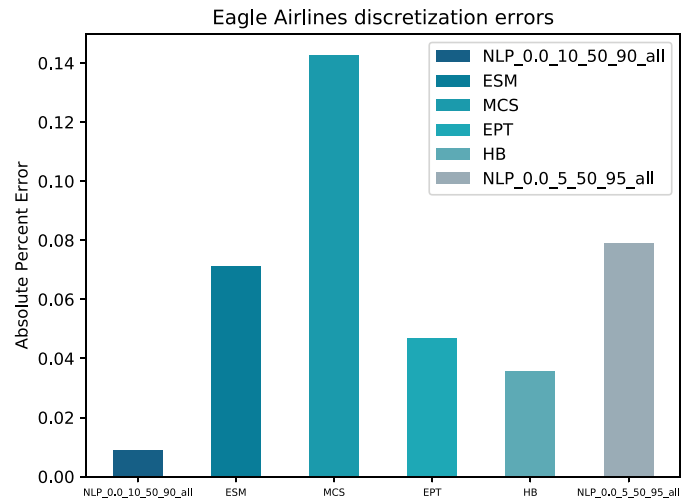


Fig. 18. A summary of selected discretizations using the example of Eagle Airlines given by Clemen and Reilly [5]. All errors are less than 0.14% with the  $P_{10} = P_{50} - P_{90}$  turning out the best.

distributions given by Montiel and Bickel [17]. The discretization using the  $P_{10} - P_{50} - P_{90}$  percentiles has the least absolute error from the true expected value. Though this example does not provide absolute proof, the Eagle Airlines example shows that a less extreme set of percentiles is more robust for determining the CE or expected value of a project when the true distributions are unknown and may not be part of the training set.

These examples should not be taken as conclusive. They illustrate that the improved performance of an optimized discretization or a shortcut is dependent on the distributions of the uncertainties. The results show optimal discretizations can have robust results for a specific type of decision problem that is repeatable. A result we do not show in Fig. 15 is that worst-case optimization consistently underperforms the average-case optimization. We also found independent discretizations outperform joint discretizations. Finally, we find that discretizations that only use a sample of the data still perform within a few percentage points of the best one. A ten percent sample performs the best over the examples we test. We believe the degradation in performance in both the shortcut and optimized discretizations in comparison to the  $10 - 50 - 90$  discretizations is due to over-fitting and the use of extreme results to provide an initially better fit.

6. Discussion, conclusion, and future work

In the computational experiments we performe, optimized discretization of Model (1) improves on existing discretization methods. This improvement can be over worst-case  $Err(d, p)$ , average  $Err(d, p)$ , or both. Furthermore, the methodology provides a large amount of flexibility. First, one could exchange the function  $Err(d, p)$  for something other than absolute percent CE error – potentially repeating the linearization process described in Section 4.1 as we did in Section 5.6. Second, the methodology is able to compute both independent and joint discretizations – a significant improvement over past methods that focus solely on independent discretizations.

Based on our findings, we make recommendations to the practitioner who would like to improve the accuracy of their discretizations and value of their recommendations. We believe more testing is necessary before choosing between joint and independent discretization methods. While a joint discretization generally provides more accurate discretizations over the training problems, they may over-fit. When using a joint discretization, it is important to have a large number of training problems.

In our tests, the 10% sample size results offer significant error

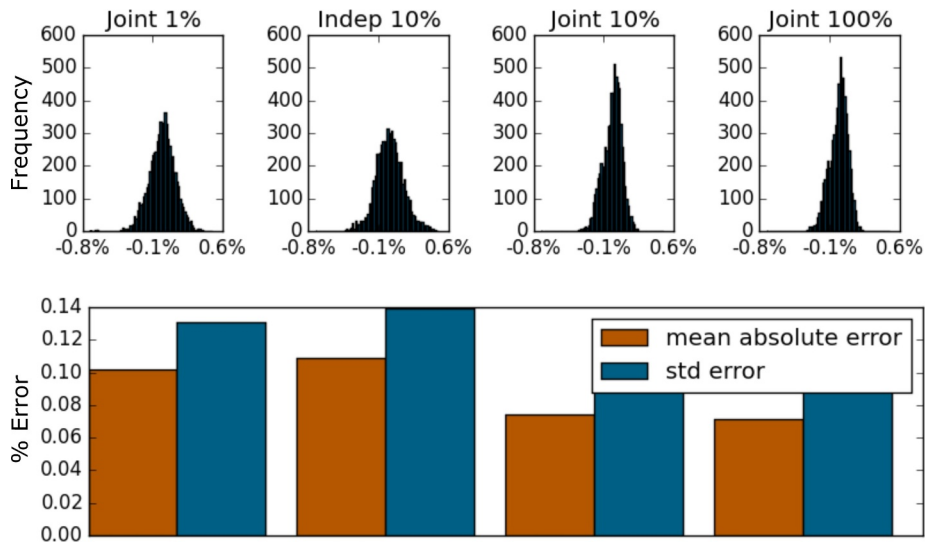


Fig. 19. A joint discretization that samples 10% of the decision problems yields results that are better than shortcuts, solve quickly, and are close in terms of resulting error to the use of 100% sampling.

reduction over shortcuts (70%) and reduce the time for the non-linear optimization by 90%. We do not know if this improvement in performance while maintaining and edge in accuracy will hold with other problems. The time required for non-linear optimization solvers to generate solutions can take days, and it is worth experimenting with smaller sample sizes to generate results that are better than shortcuts in a reasonable amount of time.

We recommend using an average error method over a worst-case error method. The analysis of the relative cost of the worst-case analysis tends to show that optimizing to the worst case provides the least value of any discretization method. Optimizing over the average error provided the highest value discretizations. The results when using the recommended problem size and discretization method are shown in Fig. 19.

From our observations, the computation time required for finding an optimized discretization increases linearly in  $|\mathcal{D}|$  due to increases in the number of constraints. Computing optimized joint discretizations depends on solving a mixed integer linear program which is generally faster than computing optimized independent discretizations. In both independent and joint discretization the computation time increases exponentially with the number of uncertainties, and the number of candidate percentiles.

We found optimized discretizations make a greater use of the 5th and 95th percentiles relative to the use of the 10th and 90th percentiles. Alpert and Raiffa [1] noted that assessing more extreme values is also more prone to error, and the results from Hammond and Bickel [8, 9] also make use of more extreme percentiles. An expert that has twice the experience is likely to have seen twice the number of extreme events, and is likely to be able to better assess the value of those extreme events. The result is that someone who is assessing the 95th percentile, may only be assessing a value at the 90th percentile. This adds an additional uncertainty to the set of decision problems. In future work we plan to determine whether using more extreme assessments results in better estimates of the true CE given different assumptions regarding the accuracy of the assessments.

## Appendix A

### A1. Wildcatter problem description

This appendix provides the details of the decision analysis problem described by Smith [19]. We cover the uncertainties and their PDFs, the valuation model, the utility and CE functions, and the sources of risk. We take variants of this basic problem to construct our decision problem sets in the article.

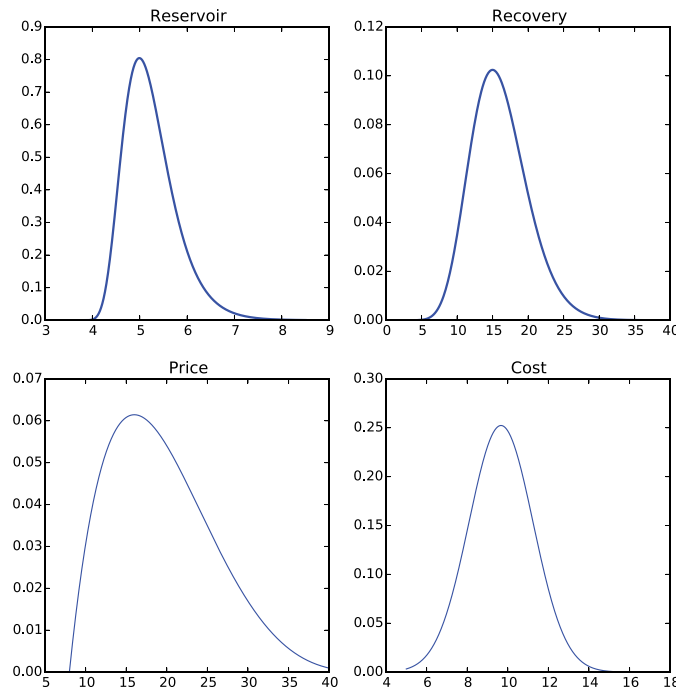
There is a potential for introducing novel algorithms for computing optimized discretizations. For example, the number of variables for both independent and joint discretizations grows exponentially with the number of possible percentiles. It may be possible to construct column generation algorithms [6] that iteratively introduce these variables as needed. This process is likely easier with the linear mixed integer program for computing optimized joint discretizations.

The computation time increases linearly with the size of the decision problem set  $\mathcal{D}$ . However, it may also be possible to reduce the size of  $\mathcal{D}$  while maintaining the quality of the resulting optimized discretization. One could view the expected error portion of the objective function of Model (1) as a type of stochastic optimization, where the scenarios are the decision problems in  $\mathcal{D}$ . There is then a potential of using scenario reduction techniques [7] to reduce the size of  $\mathcal{D}$ .

In our analysis we used a linearization of absolute percentage error. Optimal discretization is flexible in its ability to use multiple objective functions. Other objectives we have considered are measuring deviation from expected value or adding additional terms to the Taylor expansion of the error function. It is our recommendation that the Taylor expansion of the objective function be linear if possible so as to keep solution times as short as possible. In conclusion, optimized discretization can help decision analysis practitioners create discretizations that are specific to their current projects and perform better than alternate methodologies. Intuitively, the key difference between optimized discretizations and other discretization methods is that optimized discretizations take as input an entire decision problem set  $\mathcal{D}$  and a valuation function like (15). This allows optimized discretizations to focus on producing lower CE errors than using traditional discretization.

### Acknowledgements

We would like to thank our peer reviewers, Dr. Robert Hammand, Dr. J. Eric Bickel, and Dr. John Hasenbein for their recommendations and feedback.



**Fig. 20.** The original distributions are similar to the candidate distributions given in Fig. 3. As a point of reference, when using the optimal discretizations, the independent worst-case had an error of  $-0.0369\%$ , the best average discretization had an error of  $-0.0613\%$ . The joint best average discretization yields an error of  $0.0199\%$  and joint worst case discretization yields an error of  $0.0823\%$ . HB, EPT, ESM, MCS had errors of  $0.0357$ ,  $0.0411$ ,  $0.0337$ , and  $-0.6197\%$  respectively.

The problem described by Smith [19] is a wildcatting decision problem. There are four uncertainties that determine the project value. These are the oil price, reservoir volume, recovery rate, and production cost. We refer readers to the [19] for a visualization of the influence diagram. The present value of the project given realizations for the four variables is

$$Value = \begin{cases} \frac{1}{\delta} \cdot (p - c) \cdot k \cdot (1 - \exp(-\delta \cdot T)) - C & \text{if } p > c \\ -C & \text{if } p \leq c, \end{cases} \tag{15}$$

where  $v$  is the reservoir volume;  $r$  is the recovery rate;  $p$  is the oil price;  $c$  is the production cost  $k$  is a fixed production rate of 100,000 barrels per year;  $T = r \cdot v / k$  is the years of production;  $\delta$  is a fixed discount rate of 5% per year; and  $C$  are the initial capital expenditures of \$2.5 million. In the project valuation using Formula (15) the wildcatter will lose money for each barrel pumped if  $p \leq c$ . Even if the wildcatter has already made the decision to expend capital costs  $C$ , he or she may choose to not drill when each additional unit of production is not profitable.

The PDFs of the uncertainties defined by Smith [19] are:

$$\begin{aligned} \text{ReservoirVolume: } f\left(\frac{x}{10^6}\right) &= \frac{1}{\sqrt{2\pi}(x-3.5)} \cdot \exp\left(-\frac{(\ln(x-3.5)-0.5)^2}{0.2}\right) & \frac{x}{10^6} \geq 3.5 \\ \text{Recovery: } f(x \cdot 10^2) &= \frac{1}{15!} \cdot x^{15} \cdot \exp(-x) & x \geq 0 \\ \text{OilPrice: } f(x) &= \frac{1}{40} \cdot \frac{6!}{114!} \cdot \left(\frac{x-8}{40}\right)^1 \cdot \left(\frac{48-x}{40}\right)^4 & 8 \leq x \leq 48 \\ \text{ProductionCost|OilPrice: } f(x, p) &= \frac{\sqrt{8}}{\sqrt{\pi}} \cdot \exp\left(-\frac{8 \cdot (x-p/3-3)^2}{p}\right) \end{aligned}$$

The PDFs of the uncertainties and are shown in Fig. 20. Smith [19] has additional visualizations of the cumulative distribution functions for the uncertainties and the project value. These make for an interesting problem because the distributions take on many shapes, some are non-symmetrical, and the cost is dependent on the oil price. The reservoir uncertainty follows a lognormal distribution and is bounded from below at 3.5M barrels. The recovery uncertainty is a gamma distributions bounded from below at 0%. It is not bounded from above, though in practical terms it should be 100%. The oil price uncertainty follows a beta distribution with bounds at 8 and 48 dollars per barrel. Finally, the cost follows a normal distribution with a mean and standard deviation that are functions on the price.

From the project values, we are able to generate the project utility. From the expected utility we are able to generate a CE. Smith [19] uses an exponential utility function to convert the project value,  $x$  to a utility,  $u(x)$ , with a risk tolerance parameter  $\rho$ . The expected utilities are converted to a CE. The combination of the uncertainty PDFs, the valuation model, the utility function, and the risk tolerance value  $\rho$  combine to make one problem instance  $d \in \mathcal{D}$ . For all problem instances in our decision problem sets, the functions for the utility and CE are defined as:

$$\text{Utility: } u_d(x) = -\exp(-x/\rho_d) \tag{16}$$

$$\text{CertainEquivalent: } CE_d = -\rho_d \cdot \ln(-E[u_d(x)]) \tag{17}$$

$$\text{Target: } T_d = -\exp(-CE_d/\rho_d). \tag{18}$$

With this exponential utility function, utility values are between  $-\infty$  and 0, and projects with a CE of zero have a utility of  $-1$ . Lower risk

tolerances penalize losses more. An infinite risk tolerance makes the CE to be equal to the expected value of the project. Given a  $d \in \mathbb{D}$ , we can compute a CE numerically by sampling project values based on the uncertainty distributions, converting values to expected utility, and expected utility to a CE.

### A2. Eagle Airlines

Eagle Airlines is a problem described by Clemen [4] and further refined by Clemen and Reilly [5] and Reilly [18] in the area of fleet expansion. Here we used the information from [18] and the distributions given by Montiel and Bickel [17]. In this problem the owner of Eagle Airlines must decide whether or not to expand his fleet with the addition of one plane. The alternative is to invest the money in a money market earning a certain return. The problem has several uncertainties, and the ones determined to be of significance are price ( $P$ ), hours flown ( $H$ ), capacity ( $C$ ), and operational cost ( $O$ ). The owner is risk neutral and will make the decision based on comparing the expected profit to the risk-free return of the money market.

In addition to the uncertainties, the owner uses the following parameters in the profit calculation:

Parameter	Value	Description
$CR$	0.5	Charter ratio
$PF$	40%	Percentage financed
$I$	11.5%	Risk-free interest rate
$PU$	\$87, 500	Purchase price
$IN$	20,000	Insurance cost
$CP$	$3.25 \cdot P$	Charter price
$N$	5	Number of seats

The true distributions for the uncertainties are:

Uncertainty	Distribution	Parameters	Range
$P$	Beta	$\alpha = 9, \beta = 15$	[\$81.94, \$133.96]
$H$	Beta	$\alpha = 9, \beta = 15$	[66.91, 1, 136.26]
$C$	Beta	$\alpha = 9, \beta = 15$	[0, 1]
$O$	Normal	$\mu = 245, \sigma = 11.72$	$(-\infty, \infty)$

The formulas for revenue, costs and profits are:

$$Cost = H \cdot O + IN + PU \cdot PF \cdot I \tag{19}$$

$$Revenue = CR \cdot H \cdot CP + (1 - CR) \cdot H \cdot C \cdot N \cdot P \tag{20}$$

$$Profit = Revenue - Cost \tag{21}$$

Furthermore, the uncertainties are have a Spearman rank correlation with each other:

Uncertainty	Spearman correlation			
	$P$	$H$	$C$	$O$
$P$	1			
$H$	-0.5	1		
$C$	-0.25	0.5	1	
$O$	0	0	0.25	1

In order to calculate the expected mean for this problem we use the simulation methods described in [11] to generate values for the uncertainties. For the discretizations, we generate the correlated uniform values from the percentile discretizations which we then use to generate the Pearson rank correlated uncertainty values. When using rank correlation, the values we get for each percentile for an uncertainty, say  $H$ , will be different when we use different percentiles for  $P$ .

### A3. Discretization values

We present selected joint and independent discretization values for the Wildcatter problem and then Eagle Airlines. The shortcut names are the same as used in the article. The optimized discretization names are coded. The first code is either “NLP” or “MIP”. Independent discretizations are solved with a non-linear programming solver and hence have the code “NLP”. Joint discretizations are solved with a mixed-integer programming solver and hence have the code “MIP”. The next code is either a zero or a one. Zero indicates the discretization is solved to minimize the average error. A one indicates the discretization is solved to minimize the worst case error. The next three numbers are optional. These numbers indicate whether specific percentiles are used. For example, “5\_50\_95” indicates the 5th, 50th, and 95th were the only percentiles allowed in the discretization. The final code indicates how many sample decision problems are used to create the discretization.

Wildcatter Independent Discretizations

Discretization	Uncertainty	Percentiles			Probabilities		
		Q1	Q2	Q3	P1	P2	P3
MCS	Reservoir	0.10	0.50	0.90	0.25	0.50	0.25
MCS	Recovery	0.10	0.50	0.90	0.25	0.50	0.25
MCS	Price	0.10	0.50	0.90	0.25	0.50	0.25
MCS	Cost	0.10	0.50	0.90	0.25	0.50	0.25



ESM	Reservoir	0.10	0.50	0.90	0.30	0.40	0.30
ESM	Recovery	0.10	0.50	0.90	0.30	0.40	0.30
ESM	Price	0.10	0.50	0.90	0.30	0.40	0.30
ESM	Cost	0.10	0.50	0.90	0.30	0.40	0.30
HB	Reservoir	0.04	0.50	0.96	0.16	0.67	0.16
HB	Recovery	0.05	0.50	0.95	0.18	0.63	0.18
HB	Price	0.04	0.50	0.96	0.16	0.67	0.16
HB	Cost	0.04	0.50	0.96	0.16	0.67	0.16
EPT	Reservoir	0.05	0.50	0.95	0.18	0.63	0.18
EPT	Recovery	0.05	0.50	0.95	0.18	0.63	0.18
EPT	Price	0.05	0.50	0.95	0.18	0.63	0.18
EPT	Cost	0.05	0.50	0.95	0.18	0.63	0.18
NLP_1.0_5_50_95_all	Reservoir	0.05	0.50	0.95	0.25	0.61	0.14
NLP_1.0_5_50_95_all	Recovery	0.05	0.50	0.95	0.27	0.50	0.23
NLP_1.0_5_50_95_all	Price	0.05	0.50	0.95	0.03	0.86	0.11
NLP_1.0_5_50_95_all	Cost	0.05	0.50	0.95	0.21	0.63	0.16

Wildcatter Independent Discretizations

Discretization	Uncertainty	Percentiles			Probabilities		
		Q1	Q2	Q3	P1	P2	P3
NLP_0.0_5_50_95_all	Reservoir	0.05	0.50	0.95	0.15	0.66	0.18
NLP_0.0_5_50_95_all	Recovery	0.05	0.50	0.95	0.20	0.60	0.20
NLP_0.0_5_50_95_all	Price	0.05	0.50	0.95	0.19	0.63	0.18
NLP_0.0_5_50_95_all	Cost	0.05	0.50	0.95	0.19	0.62	0.19
NLP_0.0_10_50_90_all	Reservoir	0.10	0.50	0.90	0.35	0.34	0.31
NLP_0.0_10_50_90_all	Recovery	0.10	0.50	0.90	0.27	0.46	0.27
NLP_1.0_10_50_90_all	Price	0.10	0.50	0.90	0.16	0.62	0.22
NLP_1.0_10_50_90_all	Cost	0.10	0.50	0.90	0.35	0.41	0.24
NLP_0.0_all	Reservoir	0.05	0.50	0.95	0.20	0.61	0.19
NLP_0.0_all	Recovery	0.10	0.55	0.95	0.26	0.57	0.17
NLP_0.0_all	Price	0.05	0.50	0.95	0.19	0.63	0.18
NLP_0.0_all	Cost	0.10	0.55	0.95	0.29	0.53	0.18
NLP_1.0_all	Reservoir	0.05	0.55	0.90	0.22	0.52	0.27
NLP_1.0_all	Recovery	0.05	0.55	0.90	0.32	0.33	0.36
NLP_1.0_all	Price	0.05	0.50	0.90	0.25	0.44	0.31
NLP_1.0_all	Cost	0.10	0.55	0.95	0.28	0.52	0.19

Wildcatter Independent Discretizations

Discretization	Uncertainty	Percentiles			Probabilities		
		Q1	Q2	Q3	P1	P2	P3
NLP_0.0_66	Reservoir	0.05	0.45	0.95	0.13	0.67	0.20
NLP_0.0_66	Recovery	0.05	0.50	0.95	0.20	0.61	0.19
NLP_0.0_66	Price	0.05	0.50	0.95	0.18	0.64	0.18
NLP_0.0_66	Cost	0.05	0.50	0.95	0.19	0.62	0.19
NLP_0.0_328	Reservoir	0.10	0.50	0.95	0.18	0.64	0.19
NLP_0.0_328	Recovery	0.05	0.50	0.95	0.20	0.60	0.20
NLP_0.0_328	Price	0.05	0.50	0.95	0.18	0.64	0.18
NLP_0.0_328	Cost	0.05	0.45	0.95	0.15	0.64	0.20
NLP_0.0_656	Reservoir	0.10	0.50	0.95	0.20	0.61	0.19
NLP_0.0_656	Recovery	0.05	0.50	0.95	0.20	0.61	0.19
NLP_0.0_656	Price	0.05	0.50	0.95	0.18	0.64	0.18
NLP_0.0_656	Cost	0.10	0.55	0.95	0.29	0.52	0.19
NLP_0.0_1312	Reservoir	0.10	0.55	0.95	0.24	0.59	0.18
NLP_0.0_1312	Recovery	0.05	0.50	0.95	0.20	0.60	0.20
NLP_0.0_1312	Price	0.05	0.50	0.95	0.19	0.63	0.18
NLP_0.0_1312	Cost	0.10	0.55	0.95	0.29	0.53	0.18
NLP_0.0_all	Reservoir	0.05	0.50	0.95	0.20	0.61	0.19
NLP_0.0_all	Recovery	0.10	0.55	0.95	0.26	0.57	0.17
NLP_0.0_all	Price	0.05	0.50	0.95	0.19	0.63	0.18
NLP_0.0_all	Cost	0.10	0.55	0.95	0.29	0.53	0.18

Wildcatter Joint Discretizations

Reservoir Quantile	Recovery Quantile	Price Quantile	Cost Quantile	MIP_1.0_10_50_90_all	MIP_0.0_10_50_90_all
0.10	0.10	0.90	0.90		0.088
0.10	0.50	0.10	0.10	0.086	

0.10	0.50	0.50	0.10	0.133	0.026
0.10	0.50	0.50	0.90	0.047	
0.10	0.50	0.90	0.10		0.156
0.10	0.90	0.50	0.50	0.065	
0.10	0.90	0.90	0.10	0.049	
0.10	0.90	0.90	0.50	0.023	
0.50	0.10	0.10	0.90		0.135
0.50	0.10	0.50	0.90	0.254	0.109
0.50	0.50	0.50	0.10		0.048
0.50	0.50	0.50	0.50	0.052	
0.50	0.50	0.90	0.50		0.051
0.50	0.90	0.50	0.90		0.030
0.50	0.90	0.90	0.50	0.045	0.022
0.90	0.10	0.10	0.50	0.031	
0.90	0.10	0.10	0.90		0.001
0.90	0.10	0.50	0.50	0.057	
0.90	0.50	0.10	0.10	0.057	0.102
0.90	0.90	0.10	0.50		0.061
0.90	0.90	0.50	0.50		0.172
0.90	0.90	0.90	0.50	0.100	

Wildcatter Joint Discretizations

Reservoir Quantile	Recovery Quantile	Price Quantile	Cost Quantile	MIP_1.0_all
0.10	0.05	0.55	0.90	0.033
0.10	0.50	0.55	0.45	0.064
0.10	0.50	0.55	0.90	0.176
0.10	0.95	0.55	0.45	0.060
0.45	0.05	0.55	0.45	0.040
0.45	0.05	0.55	0.90	0.100
0.45	0.50	0.05	0.10	0.147
0.45	0.50	0.55	0.10	0.035
0.45	0.50	0.55	0.45	0.158
0.45	0.95	0.55	0.45	0.006
0.90	0.05	0.55	0.10	0.008
0.90	0.05	0.55	0.45	0.047
0.90	0.95	0.95	0.10	0.111
0.90	0.95	0.95	0.45	0.016

The next joint discretization is for the average error minimization. The number of non-zero percentile combinations is much larger here, which gives a more robust answer when computing out-of-sample percent errors.

Wildcatter Joint Discretizations

Reservoir Quantile	Recovery Quantile	Price Quantile	Cost Quantile	MIP_0.0_all
0.05	0.10	0.05	0.10	0.045
0.05	0.10	0.50	0.10	0.010
0.05	0.10	0.50	0.55	0.018
0.05	0.55	0.05	0.10	0.001
0.05	0.55	0.05	0.95	0.013
0.05	0.55	0.50	0.10	0.049
0.05	0.55	0.50	0.95	0.035
0.05	0.55	0.95	0.10	0.004
0.05	0.90	0.05	0.55	0.013
0.05	0.90	0.05	0.95	0.001
0.05	0.90	0.95	0.10	0.026
0.50	0.10	0.05	0.95	0.020
0.50	0.10	0.50	0.55	0.111
0.50	0.55	0.50	0.10	0.046
0.50	0.55	0.50	0.55	0.210
0.50	0.55	0.50	0.95	0.060
0.50	0.55	0.95	0.55	0.007
0.50	0.90	0.50	0.10	0.060
0.50	0.90	0.95	0.55	0.047
0.50	0.90	0.95	0.95	0.028
0.95	0.10	0.05	0.95	0.007
0.95	0.10	0.50	0.55	0.028
0.95	0.10	0.95	0.10	0.060

0.95	0.55	0.50	0.55	0.002
0.95	0.55	0.50	0.95	0.015
0.95	0.55	0.95	0.95	0.003
0.95	0.90	0.05	0.55	0.074
0.95	0.90	0.50	0.10	0.000
0.95	0.90	0.50	0.55	0.001
0.95	0.90	0.50	0.95	0.004

Eagle Airlines Independent Discretizations

Discretization	Uncertainty	Percentiles			Probabilities		
		Q1	Q2	Q3	P1	P2	P3
NLP_0.0_5_50_95_all	Price	0.05	0.50	0.95	0.186	0.629	0.185
NLP_0.0_5_50_95_all	Hours	0.05	0.50	0.95	0.184	0.632	0.184
NLP_0.0_5_50_95_all	Capacity	0.05	0.50	0.95	0.190	0.620	0.190
NLP_0.0_5_50_95_all	Operating Cost	0.05	0.50	0.95	0.185	0.630	0.185
NLP_0.0_10_50_90_all	Price	0.10	0.50	0.90	0.308	0.386	0.306
NLP_0.0_10_50_90_all	Hours	0.10	0.50	0.90	0.293	0.415	0.292
NLP_0.0_10_50_90_all	Capacity	0.10	0.50	0.90	0.270	0.461	0.269
NLP_0.0_10_50_90_all	Operating Cost	0.10	0.50	0.90	0.318	0.373	0.309

Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.orp.2018.09.002](https://doi.org/10.1016/j.orp.2018.09.002).

References

[1] Alpert A, Raiffa H. Judgement under uncertainty: heuristics and biases. A progress report on the training of probability assessors. Cambridge, England: Cambridge University Press; 1982. p. 294–305.

[2] Bickel JE, Lake LW, Lehman J, et al. Discretization, simulation, and Swanson’s (inaccurate) mean. Soc Petrol Eng Econ Manage 2011;3(03):128–40.

[3] Brockett PL, Kahane Y. Risk, return, skewness and preference. Manage Sci 1992;38(6):851–66.

[4] Clemen RT. Making hard decisions: an introduction to decision analysis. 2nd ed. Duxbury Press; 1996.

[5] Clemen RT, Reilly T. Correlations and copulas for decision and risk analysis. Manage Sci 1999;45(2):208–24.

[6] Dantzig GB, Wolfe P. Decomposition principle for linear programs. Oper Res 1960;8(1):101–11.

[7] Dupačová J, Gröwe-Kuska N, Römisch W. Scenario reduction in stochastic programming. Math Program 2003;95(3):493–511.

[8] Hammond RK, Bickel JE. Reexamining discrete approximations to continuous distributions. Decis Anal 2013;10(1):6–25.

[9] Hammond RK, Bickel JE. Approximating continuous probability distributions using the 10th, 50th, and 90th percentiles. Eng Econ 2013;58(3):189–208.

[10] Hurst A, Brown G, Swanson R. Swanson’s 30-40-30 rule. Am Assoc Pet Geol Bull 2000;84(12):1883–91.

[11] Iman RL, Conover W-J. A distribution-free approach to inducing rank correlation among input variables. Commun Stat-Simul Comput 1982;11(3):311–34.

[12] Keefer DL. Certainty equivalents for three-point discrete-distribution approximations. Manage Sci 1994;40(6):760–73.

[13] Keefer DL, Bodily SE. Three-point approximations for continuous random variables. Manage Sci 1983;29(5):595–609.

[14] McKay M, Beckman R, Conover W. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 1979;21(2):239–45.

[15] McNamee P, Celona J. Decision analysis with supertree. San Francisco, CA: Scientific Press; 1990.

[16] Miller III AC, Rice TR. Discrete approximations of probability distributions. Manage Sci 1983;29(3):352–62.

[17] Montiel LV, Bickel JE. A simulation-based approach to decision making with partial information. Decis Anal 2012;9(4):329–47.

[18] Reilly T. Sensitivity analysis for dependent variables. Decis Sci 2000;31(3):551–72.

[19] Smith JE. Moment methods for decision analysis. Manage Sci 1993;39(3):340–58.

[20] Wallace SW, Hoyland K. Generating scenario trees for multistage decision problems. Manage Sci 2001;47(2):295–307.