Accounting for Forest Growth Uncertainty due to Climate Change in Harvest Scheduling

Martin B Bagaram, Sándor F Tóth, Andrés Weintraub

Accounting for Forest Growth Uncertainty due to Climate Change in Harvest Scheduling

Martin B. Bagaram[a,b], Sándor F. Tóth[a,1], Andrés Weintraub P.[1]

[a]*School of Environmental & Forest Sciences, University of Washington, Box 352100, Seattle, WA 98195, USA*
[b]*Department of Industrial & Systems Engineering, University of Washington, Seattle, WA 98195-2650, USA*
[c]*Department of Industrial Engineering, University of Chile, República 701, Santiago, Chile*

Abstract

Climate change affects forest growth and yield, making it difficult to make optimal harvest decisions without knowing the future climate. We propose a multi-stage stochastic programming framework to address this uncertainty. Our proposed framework requires an a priori construction of scenario trees where the branches represent non-anticipativities of future climate paths affecting forest growth. We illustrate our model with four climate paths using a case study in Washington State, United States. Using this case, we study as to whether it is optimal to expect a positive climate effect on yield or not. We find in this particular case that it is optimal to be optimistic. We conclude that multi-stage stochastic programs with non-anticipativity constraints can be used to analyze the effects of decision makers' perception of uncertain futures on the optimality of their decisions.

Keywords: data driven optimization, climate change, harvest planning, uncertainty, forest

1. Introduction

The sustainable provision of forest products and services is at the core of strategic (long-term) forest planning. Although there are a multitude of stewardship objectives that might guide decision makers in charge of forest management, such as habitat conservation [46, 41, 47] or wildfire hazard reduction [42],

timber investment companies maximize net present value for their shareholders. To manage forests as financial assets, these companies need to be able to forecast the future growth of their inventories with a certain level of accuracy. Using historical data from the past to predict the future has its limits because climate change increasingly affects forest growth via more frequent fires, slides, windstorms, drought or other "disturbances" [14]. Climate change is often cited as the primary concern for public forest sectors in the United States and beyond De Pellegrin Llorente et al. [11].

Growth and yield parameters are used to parameterize forest harvest scheduling models. These models optimize as to which forest units or stands should be prescribed which treatments, e.g. harvest- or thinning actions, and when to best meet management objectives such as maximizing NPV. When these input parameters are known with certainty, *deterministic* harvest scheduling models can be used (e.g., [45], [41], [36], and [5]). However, deterministic models are inadequate when forest growth is subject to considerable variation due to climate change. Multi-stage stochastic programming models allow analysts to make robust decisions in the present and have an optimal plan for future contingencies with limited information about future growth or productivity. The stochastic model considers the range of possible future growth and yield instead of a single deterministic one. These alternative futures, also known as scenarios, are simply discretizations of random parameters along their predicted but uncertain domains. The set of relevant scenarios can be structured as a *scenario tree* where each node is a possible future state of the forest in terms of volume and each branch is a possible trajectory of subsequent growth leading to the next state (node). Each "path" through this tree is a scenario with computed probability. This structured discretization of future states of interest is crucial in multi-stage stochastic programming. Too coarse of a discretization can lead to sub-optimal decisions [40], whereas too high of a resolution can render the program computationally intractable.

Previous studies that incorporated climate change uncertainty in forest harvest scheduling (e.g., [2, 16]) used scenarios generated by experts without taking

into account the degree of change in forest growth resulting from climate variations and whether the scenarios sufficiently captured the underlying growth uncertainty. The scenarios were identified as possible futures but without knowing the associated probabilities of these futures, it remains unclear how this information should be incorporated in the models. The unique challenge of building stochastic harvest scheduling models under climate change is that climate uncertainty occurs at three levels. First, climate change is forecast as climate paths or pathways each of which represents a set of assumptions about the future such as advancement in technology, human population growth, or CO_2 emissions. The problem is that we have no information about the probability of these climate paths and thus it is not possible to incorporate them in the stochastic program with conventional methods.

Second, there is uncertainty with regards to how forests would respond in terms of growth under each climate path. Thus, even if we knew which climate path would materialize, there is still uncertainty as to what degree actual growth would change under that path. A study conducted on how forest growth would change in the face of four different climate paths in the Pacific Northwest United States [28] illustrates this layer of uncertainty. The results showed that forest behavior under each one of those climate paths will be different depending on the location and the altitude of the forest. Similar results were obtained in Brazil [14]. Considering all possible climate paths, with all possible associated growth responses, could make the stochastic model too unwieldy and its solutions less than optimal.

Lastly, because of climate change mitigation efforts and the advancement in technology, there is no guarantee that if the climate of the next decade represents one of the climate paths, then the same climate will remain in the following decades. Indeed, because of mitigation efforts, it is possible to transition from one climate path to a different one in the same decade. It is important to highlight that the objective of generating the scenario tree is not to have a perfect scenario tree but a tree that adequately captures the underlying stochastic process that is suitable for strategic planning. Thus, it is important to determine

what the desired characteristics of a suitable tree should be and how one can determine whether or not a given given tree adequately represents the underlying stochastic space.

The objective of this paper is to devise a method for incorporating forest growth uncertainty associated with climate change into harvest scheduling models. The novelty of the approach stems from the unique characteristic of climate uncertainty and its layered effects on decision making in forestry as described above. In a nutshell, we propose that each climate path is considered individually in its own stochastic program first. In the absence of empirical evidence with regards to the probability distribution of forest growth responses to specific climate scenarios, we assign arbitrary distributions, such as normal and uniform theoretical distributions, to the growth response under each scenario. We then solve the stochastic program for each distribution and each climate scenario. If using the stochastic solution from one climate path in place of the stochastic solution provided by a different climate path yields a stochastic objective function value that is within the associated bounds, then we consolidate those two paths into one.

We develop a very comprehensive harvest scheduling model that is suitable for both tactical and strategical harvest planning which integrates all the constraints in the forest industry in particular the spatial constraints. To our knowledge, most papers limit themselves to simplified harvest models. We show as well that if the decision maker is more interested in maximizing their net present value, then the optimal attitude they should have is to expect an optimistic climate change which predicts an increase of forest growth.

The rest of this paper is organized as follows. We start with a review of the scenario generation literature for stochastic programming in Section 2. In Section 3, we present the scenario generation procedure we propose specifically for forest harvest scheduling under climate change. In Section 4, we apply the methodology to a case study; and finally in Section 5, we discuss the results and provide some conclusions and recommendation for future work.

2. Literature review

To begin, it is worth emphasizing that though this paper deals with stochastic programming, the uncertainty in mathematical models can be addressed in different ways such as robust optimization, stochastic programming, and chance constrained optimization [3, 10, 34]. The aim of robust optimization is to guarantee feasibility over the specified uncertainty set [44]. It is known that when the full set of uncertainty is considered, robust optimization results in the worst-case solution that will likely be very expensive in terms of foregone objective function value. This is because robust optimization produces solutions that must be feasible under every possible scenario no matter how unlikely. According to Apap and Grossmann [3], robust optimization is only suitable for short term planning where feasibility is the main concern. In contrast, forest harvest scheduling is a very long-term planning problem.

In stochastic programming, the decision maker must make a decision at the beginning of the planning horizon without knowing what values some uncertain parameter(s) affecting the model will take in subsequent planning periods. After a period in which the uncertainty might reveal itself, the decision maker can take a recourse action at a cost (to the objective function). Depending on whether the sequence of initial decisions and taking recourse actions occurs once or more than once, the model must be cast as a two- or multi-stage stochastic program. Because multistage stochastic programming does not fix all actions that should be taken in advance, it is a method suitable for long term planning such as harvest scheduling [3]. However, unlike robust optimization, stochastic programming relies on the discretization of the continuous uncertain parameters and the probability associated with each realization of the uncertainty. Each realization of the uncertain parameters is known as a scenario and the set of scenarios form the scenario tree. The scenario tree generation constitutes the first step for building multistage stochastic programming models.

There are many methods for building scenario trees. These methods include moment matching, sampling average approximation, clustering and conditional

sampling. Moment matching is a method aiming at matching statistical moments between the scenario tree and the distribution of the random parameter. The technique was developed by Fleishman [15] for univariates, however Høyland et al. [25] extended its use to multivariate cases. The principle of this method is to generate a scenario tree with certain properties (with moments such as average, standard deviation, skewness and kurtosis) matching the ones of the theoretical distribution of the random parameter. A description of the steps to undertake for implementing moment matching is described in [25]. A shortcoming of the method is that the scenario generation step itself requires solving an optimization problem that is not linear. Furthermore, as shown by [21] and highlighted by [32] there could be many theoretical distributions having the same first moments such as the one listed. The method has been mainly successfully applied for scenario generation in portfolio management [26, 25, 37]. Nevertheless, [40] tested the performance of the method for forest harvest scheduling with price uncertainty with limited success.

The sampling average approximation (SAA) technique is relatively simple to implement. It consists of drawing repetitively many samples from the distribution of the random parameter and solving the optimization problem for each of these samples. If the sample size is large enough, the average solution will approximate the true solution [4]. To that end, Mak et al. [33] showed that the expectation of individual solutions corresponds to the lower bound on the true solution of the stochastic model and that the bound monotonically increases as the sample size increases (for minimization problems). For a formal description, the reader is referred to Löhndorf [32]. Sample average approximation (SAA) has been applied in several fields such as portfolio selection [49], supply chain design and supply chain network, transportation [43, 9], personnel assignment [38] and forest planning under climate change uncertainty [4]. One limitation of the method is that it may require solving hundreds or thousands of optimization problems in order to achieve stability. Although the method performs well for two-stage problems, it does not yield the level of flexibility a decision maker may need in multi-stage problems.

The idea of scenario clustering is to generate a set of data paths known as a 'fan' [7] that represents possible futures and then to proceed into grouping these paths into a scenario tree. This technique is sometimes referred to as distribution free scenario generation since it relies on generating paths that correspond to past scenarios [13] or experts view of the future [21, 20]. The advantage of scenario clustering is to reduce the computational burden a fully rendered scenario tree might impose on the stochastic program [44]. Nevertheless, the technique inherits the drawbacks of the methods used for generating the initial scenario tree [50]. The technique has been extensively employed in the field where the future behavior of the stochastic process is deemed to be identical to previous observations. Particular domains of applications of the method include portfolio management [18, 21, 7], interest rate management in investments [39], hydroelectric power management [12, 17, 50], and reservoir management [30].

None of the aforementioned methods is suitable without modifications for harvest scheduling with climate uncertainty. For instance, in the case of [2, 16], forest growth scenarios resulting from climate change were provided by an independent research. Therefore, the scenario tree might not be tailored to address the level of uncertainty in forest growth. In addition, because the scenario tree was given in these studies, they did not explicitly consider the two levels of uncertainty that are considered in this research. In the following section, we describe our methodological approach in overcoming the challenges that harvest scheduling under climate uncertainty poses especially when we consider forest growth prediction from statistical models.

3. Methodology

To formally introduce the scenario generation procedure, we present the general form of the stochastic problem we intend to solve. Let us consider that the stochastic problem of interest is presented as (1) where ξ is the continuous random vector which does not depend on x; the expectation is taken with respect to ξ. X is the set of constraints that the decision variable vector x needs to

satisfy, and z^* is the true objective function value of the stochastic program.

$$z^* = \min_{x \in X} \mathbb{E}\, f(x, \xi) \tag{1}$$

However, we cannot solve directly problem (1) because of the presence of the continuous random vector ξ, we can use the approximation of (1) by

$$\tilde{z} = \min_{x \in X} \mathbb{E}\, f(x, \tilde{\xi}), \tag{2}$$

where $\tilde{\xi}$ is the discretization of the random continuous vector into a scenario tree (see Section 3.1 for formal definition of the scenario tree). Since $\tilde{\xi}$ is discrete realizations of the random vector ξ, we could rewrite (2) using the summation. However, we leave this definition for later once we have defined the structure of $\tilde{\xi}$.

3.1. Scenario trees generation

3.1.1. Scenario tree structure

Before diving into how to build scenario trees, first let's briefly describe the scenario tree structure. The random process in multistage stochastic programming can be represented as a "scenario tree" which has the following structure. Let $\mathcal{T}$ denote the set of periods in the planning horizon with $T = |\mathcal{T}|$ being the number of stages at which decisions can be made. A node of the scenario tree represents the realization of the uncertain parameters and variables at a given time period. It is a possible state of the forest at a given time $t \in \mathcal{T}$. Let n and $\mathcal{N}$ describe the node and the lexicographically numbered set of nodes $\{1, ..., |\mathcal{N}|\}$ in the tree, respectively. From each node n, for $t \in \mathcal{T} \setminus \{T\}$ there is at least one branch leading to another node m with probability P_m. Let Ω represent the finite set of representative scenarios in the tree. A scenario $\omega \in \Omega$ is a particular realization of the uncertain parameter represented as a path from the root-node to a leaf-node (terminal nodes). Each scenario ω has a probability or weight denoted by w^ω. Note that $\sum_{\omega \in \Omega} w^\omega = 1$. We can now rewrite (2) as:

$$\tilde{z} = \min_{x \in X} \sum_{\omega=1}^{|\Omega|} w^\omega f(x, \omega), \tag{3}$$

where $f(\cdot, \omega)$ is the optimization function evaluated for the scenario ω.

Decisions are made at each stage and implemented in the subsequent period. For instance, in Figure 1, $t = 1$ represents the first stage (stage 1). At the first stage, the decision maker needs to decide which forest units will be harvested in the first period without knowing which scenario will occur. The period is the time between two consecutive stages. Hence, $t = 1$ marks the beginning of the first period and $t = 2$ marks its end. In the example of Figure 1, we have three stages and two periods. The value of the random parameter is only revealed in periods while the decision needs to be taken at the stage before the uncertainty is revealed. The natural question becomes: What is the procedure to generate the scenarios necessary for solving (3)? and what is the appropriate number of scenarios $|\Omega|$? We address the two questions in the next sections.

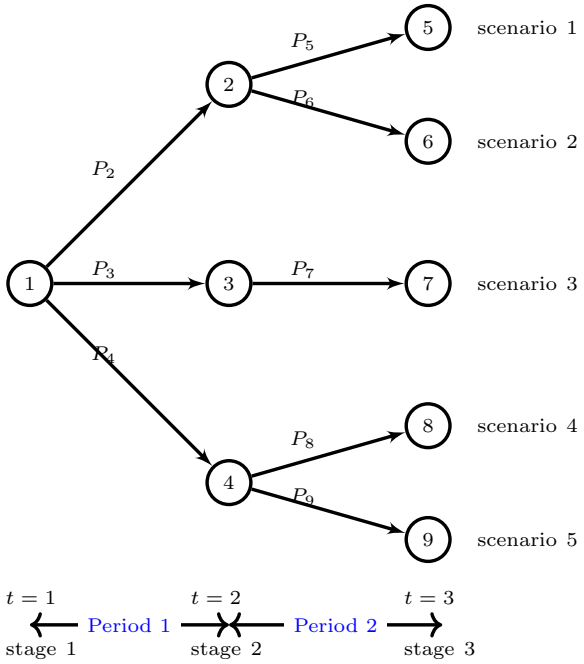

Figure 1: *Scenario tree representation of stochastic programming. The scenario tree shows five scenarios with three stages*

3.1.2. Scenario generation procedure

Scenario generation is more than a science, it is an art [8] that gives the modeler the flexibility to decide on the structure of the scenario tree through discretization of the random vector. Fine discretization of the continuous random vector leads to a computationally intractable stochastic program, while a coarse discretization leads to a tree that may completely alter the structure of the underlying stochastic process the scenario tree ought to represent. Although many researchers focused on scenarios' generation for stochastic programming, the scenario tree is not an end but a wherewithal to solve stochastic programs. The scenario tree, therefore, ought to have some properties like stability, unbiasedness and minimal stochastic optimality gap[1]. These properties will be discussed in Section 3.2.

The framework used for building scenario trees in this research is based on conditional sampling. The method consists of fitting at each node of the scenario tree a conditional probability density function and sampling from it the values that successor nodes will have. Hence, except the root-node which has no predecessor node, the value of each node depends on its predecessor's value. This method has the advantage of controlling the range of values that each node may take depending on the process that led to it. In addition, the modeler could specify edge cases that should be represented by the scenario tree.

The procedure implemented for scenario generation is inspired from [1], however, with many differences. It consists of dividing the sampling space of forest growth change into an equal number of parts corresponding to number of branches the scenario tree should have at the given stage. To illustrate the method, let's suppose the random parameter is normally distributed as illustrated in Figure 2. In the figure, $t = \{1, 2, 3\}$ represents stages at which decisions are taken. Let L and U be the lower and the upper bounds, respectively, of the support of the random parameter ξ. The node a is the root-node and there

[1]This is different from MIP optimality gap

is no growth change associated with it. However, the values of nodes b, c and d are drawn from a normal distribution with mean μ_a, standard deviation σ_a and support Ξ_a. We denote by Φ_a the associated cumulative density function. For all nodes j in the scenario tree, we require that $\Xi_j \subseteq [L, U]$. Notice that for normally distributed random variables with mean μ_a and variance σ_a^2, 99.73% of the cumulative density is within the interval $\Xi'_a = [\mu_a - 3\sigma_a, \mu_a + 3\sigma_a]$. Let β designate the number of branches that should originate from node a (β is the number of successor nodes). Let's divide Ξ'_a into equal intervals of width $w = 6\sigma_a/\beta$. The probability associated with each one of the successor nodes is given using the cumulative density for the interval in which the successor node is uniformly drawn from. For instance, the value of $b \in [L, \mu_a - 3\sigma_a + w]$ with a probability of P_b which is the cumulative density of that interval. For example, from Figure 2, $\beta = 3$, $w = 2\sigma_a$. Hence, nodes b and c, for instance, are uniformly drawn from interval $[L, \mu_a - \sigma_a]$ with a probability of 0.16 and $[\mu_a - \sigma_a, \mu_a + \sigma_a]$ with a probability of 0.68, respectively, In other words, $P_b = \Phi_a(\mu_a - \sigma_a) = 0.16$ and $P_c = \Phi_a(\mu_a + \sigma_a) - \Phi_a(\mu_a - \sigma_a) = 0.68$. To build the scenario tree, we repeat the process for each node until the leaf nodes corresponding to the terminal stage. As shown, in Figure 2, we arbitrarily chose the number of branches $\beta = 3$. One may wonder whether such a value of β, dictating the number of scenarios in the scenario tree, is justified. The question is what represents the appropriate number of branches suitable? To answer to this question, we need to assess the properties of the generated scenario tree.

3.2. Properties of a good scenario tree

Since the process of generating scenario trees described in Section 3.1 is stochastic, two runs of the scenario generation procedure might lead to two different scenario trees. However, this difference should not be significant as to affect the optimal solution of the stochastic programming. If the difference is substantial, then the scenario tree is not stable. This stability in scenario generation is measured through in-sample stability and out-of-sample stability

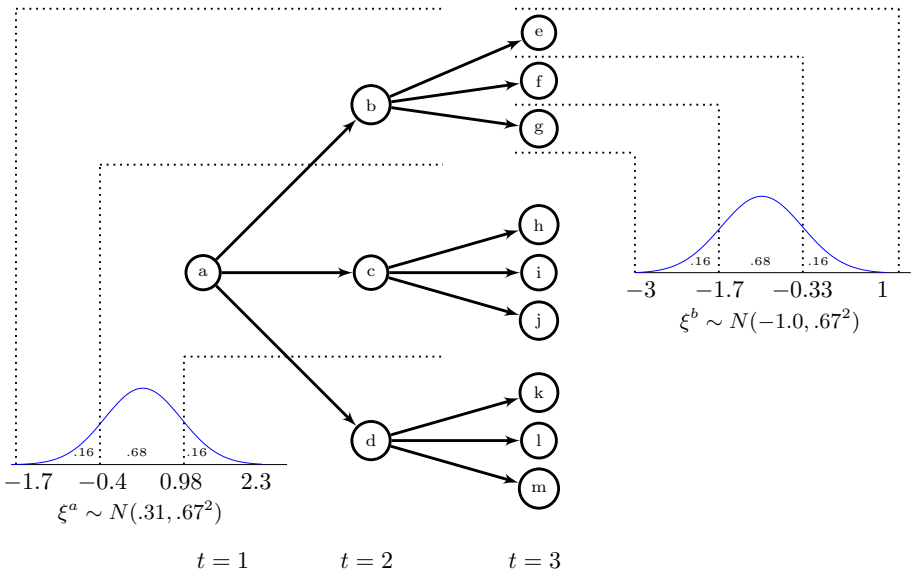

Figure 2: *Growth change scenario generation process*

[27].

3.2.1. In-sample stability

In-sample stability measures whether the difference in the solution using two different scenario trees generated from the same process is just due to the randomness in the process and not to the structure of the scenario tree. Let's suppose we generate two scenario trees $\tilde{\xi_1}$ and $\tilde{\xi_2}$ with the same structure (same number of scenarios and same number of branches at each stage). By in-sample stability, the objective function values of the two scenario trees are approximately equal as:

$$\min_{x \in X} \mathbb{E}\, f(x, \tilde{\xi_1}) \approx \min_{x \in X} \mathbb{E}\, f(x, \tilde{\xi_2}). \tag{4}$$

A different way of expressing in-sample stability is to consider $\tilde{\xi}_1, \tilde{\xi}_2, \ldots, \tilde{\xi}_k$ scenario trees with increasing size, such that $|\tilde{\xi}_1| < |\tilde{\xi}_2| < \cdots < |\tilde{\xi}_{k-1}| < |\tilde{\xi}_k|$ where $|\cdot|$ is the number of scenarios. If there is in-sample stability then

$$\min_{x \in X} \mathbb{E}\, f(x, \tilde{\xi}_{k-1}) \approx \min_{x \in X} \mathbb{E}\, f(x, \tilde{\xi}_k). \tag{5}$$

In other words, increasing the size of the scenario tree does not alter the solution and therefore, the scenario tree reached in-sample stability.

3.2.2. Out-of-sample stability

Out-of-sample stability is guaranteed if for two scenario trees $\tilde{\xi}_1$ and $\tilde{\xi}_2$ with the same structure, we can write:

$$\mathbb{E}\left[f\left(\arg\min_{x\in X}\mathbb{E}\left[f(x,\tilde{\xi}_1)\right],\xi\right)\right] \approx \mathbb{E}\left[f\left(\arg\min_{x\in X}\mathbb{E}\left[f(x,\tilde{\xi}_2)\right],\xi\right)\right]. \quad (6)$$

In practice, (6) is impossible to evaluate because we cannot evaluate the value of the approximated solution on each value of the continuous random vector. After all, we would not need the scenario tree if we could solve directly the continuous random process expressed by ξ.

Since (6) is impossible to verify, we can check out-of-sample stability by implementing a Monte-Carlo like approach [33, 40] and assuming out-of-sample stability if we can state that

$$\frac{1}{n}\sum_{i=1}^{n} f(\tilde{x}^1,\tilde{\xi}_i) \approx \frac{1}{n}\sum_{i=1}^{n} f(\tilde{x}^2,\tilde{\xi}_i), \quad (7)$$

where $\tilde{\xi}_1,...,\tilde{\xi}_n$ are i.i.d scenario tree samples of the same structure like $\tilde{\xi}$ and $\tilde{x}^k = \arg\min_{x\in X}\mathbb{E}\left[f(x,\tilde{\xi}_k)\right], \quad k=1,2.$

3.2.3. Alternative stability measurement

A different way we measured the scenario tree stability since we are confident the scenario generation procedure possesses sufficient randomness, was through relative stability measurement [19]. We generated a given number of scenarios of the same structure and computed the objective function for each one of them. We can find the largest (z^+) and the smallest (z^-) objective function values and compute the relative variability as $(z^+ - z^-)/z^+$. Then, we picked the scenario structure with a minimum number of scenarios that leads to a relative variability lower than a threshold fixed, say 1% or 0.5% (low variability means high stability).

3.2.4. Unbiasedness

Even when in-sample and out-of-sample stability are guaranteed, the discretization could still lead to a biased solution. The scenario tree leads to unbiased solutions if the expected solution and the true solution are approximately equal such as:

$$\mathbb{E}\left[f\left(\arg\min_{x\in X}\left[\mathbb{E}\, f(x,\tilde{\xi})\right],\xi\right)\right] \approx \min_{x\in X}\mathbb{E}\, f(x,\xi). \tag{8}$$

Although (8) is important for assuring that the scenario generation leads to a solution diverging little from the true solution, it cannot be evaluated since the original problem cannot be solved because of the presence of the continuous random vector ξ. Kaut and Wallace [27] recommend to approximate the continuous variable through a 'reference tree'. Such a tree is the biggest possible tree that can be solved and it must be generated from a process known to be unbiased. We can assume an unbiassed scenario tree if on one hand, the process that generates the tree is random and on the other hand, the scenario tree fulfills the aforementioned properties.

3.3. Confidence interval on stochastic optimality gap

We have stated that it is impossible to solve the original problem with the continuous random vector ξ, as consequence, its approximation with $\tilde{\xi}$ was required. Even if the approximation leads to a scenario tree with the aforementioned properties (in-sample and out-of-sample stability, unbiasedness), we still need to know to what extent the objective function value of the approximation diverges from the true optimal value. The stochastic optimality gap $e(\tilde{\xi},\xi)$ evaluates the error (negative bias in the case of minimization problems) of approximating ξ by $\tilde{\xi}$. Hence, $e(\tilde{\xi},\xi)$ is computed as:

$$e(\tilde{\xi},\xi) = \mathbb{E}\, f(\tilde{x},\xi) - \min_{x\in X}\mathbb{E}\, f(x,\xi) \tag{9}$$

$$= \mathbb{E}\, f(\tilde{x},\xi) - z^* \tag{10}$$

$$= \mu_{\tilde{x}}\ ; \tag{11}$$

where $\mu_{\tilde{x}}$ is the stochastic optimality gap of the solution $\tilde{x}$ obtained from the scenario tree $\tilde{\xi}$. Since we cannot evaluate the stochastic optimality gap directly, we can assess its bounds. The computation of those bounds are provided in Bayraksan and Morton [6]. We can estimate the stochastic optimality gap by computing its upper bound. We readily know that its lower bound is zero, corresponding to the solution from the scenario tree approximation which is as good as the true solution. The stochastic optimality gap of a candidate solution $\hat{x}$ is given by:

$$G_n(\hat{x}) = \frac{1}{n}\sum_{i=1}^{n} f(\hat{x}, \tilde{\xi}_i) - \min_{x \in X} \frac{1}{n}\sum_{i=1}^{n} f(x, \tilde{\xi}_i) \; ; \tag{12}$$

where $\tilde{\xi}_i$ is a scenario tree from the discretization of the random parameter ξ. Note that the term $G_n(\hat{x})$ is always positive. In-sample stability increases with the number of scenarios (the number of branches at each stage) for the four climate paths (Figure 3). The relative stability presented in Table 4 shows that except for the climate path B1, when the random parameter is uniformly distributed, which has a relative stability less than 1% when considering three branches at Consequently, it is necessary to define its upper bound. There are many methods used to that end. Those methods include multiple replications procedure (MRP), two replications procedure (TRP) and single replication procedure (SRP) [33, 6]. However, in this research, we focus on MRP presented in Algorithm 1 because it allows to compute more robust bounds.

Algorithm 1 MRP

Input: Candidate solution $\hat{x}$, the number of samples n, the number of replications n_G and $\alpha \in [0, 1]$

1: **for** $i = 1, 2, ..., n_G$ **do**

2: Sample i.i.d observations $\tilde{\xi}_{i1}, \tilde{\xi}_{i2}, ..., \tilde{\xi}_{in}$ from the distribution of ξ

3: Calculate $G_n^i(\hat{x})$ using the (12)

4: **end for**

5: Calculate gap estimate $\overline{G}_{n_G}$ and variance $s_G^2(n_G)$

$$\overline{G}_{n_G} = \frac{1}{n_G} \sum_{i=1}^{n_G} G_n^i(\hat{x}) \qquad s_G^2(n_G) = \frac{1}{n_G - 1} \sum_{i=1}^{n_G} \left(G_n^i(\hat{x}) - \overline{G}_{n_G} \right)^2$$

6: Output the one-sided confidence interval on $\mu_{\hat{x}}$

$$\mu_{\hat{x}} = \left[0, \overline{G}_{n_G} + \frac{t_{n_G - 1, \alpha} s_G(n_G)}{\sqrt{n_G}} \right]$$

As highlighted by Bayraksan and Morton [6], although the stochastic optimality gap is not normally distributed, from the central limit theorem, since $\overline{G}_{n_G}$ is the mean of i.i.d random variables it can be approximated to a normal distribution.

3.4. Convergence of two climate paths

We have so far covered how to generate good scenario trees for the purpose of stochastic programming; we have computed the bounds on the stochastic optimality gap that arise from approximating the continuous random vector by a scenario tree. However, we still have an issue of how to deal with the different climate scenarios (climate paths) that are forecast by climate scientists. We present here a method for reducing the number of climate paths that are worthy of consideration. Let's suppose two climate paths i and j ($i \neq j$). We can solve (2) for the two climate paths to obtain the tuple of solutions and objective function values $(\hat{x}_i, \hat{z}_i)$ and $(\hat{x}_j, \hat{z}_j)$ corresponding to climate paths i and j, respectively. We claim that the two climate paths are not different if they have

the same solution. However, because it is possible to have multiple optimal solutions or to have two different solutions that lead to the same objective function value, we use the objective function to evaluate the similarity between the two climate paths.

Hence, we conclude convergence if using the solution $\hat{x}_i$ from climate path i in the function of climate path j leads to the same objective function as $\hat{z}_i$. Put differently, we claim convergence of two climate paths i and j if

$$\hat{z}_i \approx \mathbb{E}\left[f_i\Big(\arg\min_{x\in X}\mathbb{E}[f_j(x,\tilde{\xi})],\tilde{\xi}\Big)\right]; \tag{13}$$

where f_i and f_j are the objective functions of the optimization problem for climate paths i and j, respectively. If there is convergence, then one of the climate paths is sufficient to capture the underlying random process and there is no need to consider both climate paths. In practice, we can conclude that the two scenario paths lead to the same solution if the right hand side term of (13) belongs to the confidence interval of z_i^*. From the previous sections, we have all the material to compute the confidence interval on the true objective function value z_i^* of the climate path i using a candidate solution $\hat{x}_i$ as follows:

$$z_i^* \in \left[\hat{z}_i,\ \hat{z}_i + \Big(\overline{G}_{n_G}(\hat{x}_i) + \frac{t_{n-1,\alpha}s_G(n_G)}{\sqrt{n_G}}\Big)\right]. \tag{14}$$

The confidence interval in (14) is one sided. It is computed taking into account the stochastic optimality gap from Algorithm 1. For minimization problems, $\hat{z}$ is negatively biased; which means that $\mathbb{E}\,\hat{z} \leq z^*$ [33, Theorem 1]. This stems from the fact that the solution from the discretization is more optimistic because it only considers a finite number of scenarios we optimize against.

3.5. Practical considerations

In practice, it is difficult to compute the stochastic optimality gap ($G_n(\hat{x})$ arising from the discretization of the random parameter), as defined in (12) . The difficulty stems from the impossibility to solve certain mixed integer

programs to the full optimality in a reasonable clock time. Here, we need to highlight that the optimality gap discussed is related to the mixed integer program (MIP) and the solver used to solve the MIP. That gap (g) is computed by comparing the objective value of the incumbent solution (x) to the best lower bound $\mathcal{L}$ (for minimization problems) as follows:

$$g = \frac{f(x) - \mathcal{L}}{f(x) + \varepsilon}\ , \tag{15}$$

where ε is a small quantity that prevents from dividing by zero. $g = 0$ means that the incumbent solution x is the optimal solution of the MIP. However, for harvest scheduling MIPs, it is difficult to achieve such a solution in a reasonable time. The common practice is to set an acceptable stopping optimality gap, $g > 0$. Unfortunately, this simplification affects (12). We cannot guarantee anymore that each $G_n(\hat{x})$ is positive because the solution to $\min_{x \in X} \sum_i^n f(x, \tilde{\xi}_i)$ might not be the true optimal solution. We solve this issue by computing the lower bound on the stochastic optimality gap as computed in (12) by only using positive $G_i(\hat{x})$. We compute both the lower bound ($G_n^l(\hat{x})$) and upper bound ($G_n^u(\hat{x})$) on the stochastic optimality gap using (16)

$$G_n(\hat{x}) = \begin{cases} G_n^l(\hat{x}) = \frac{1}{n}\sum_{i=1}^n f(\hat{x}, \tilde{\xi}_i) - \min_{x \in X} \frac{1}{n}\sum_{i=1}^n f(x, \tilde{\xi}_i) & \text{if (12) is positive} \\ \\ G_n^u(\hat{x}) = \frac{1}{n}\sum_{i=1}^n f(\hat{x}, \tilde{\xi}_i) - \frac{1}{n}\sum_{i=1}^n \mathcal{L}_i, & \text{Otherwise} \end{cases} \tag{16}$$

where $\hat{x}$ is a candidate solution of the stochastic program, $\tilde{\xi}_i$ is a scenario tree from the discretization of the uncertain parameter, and $\mathcal{L}_i$ is the lower bound obtained while solving the stochastic mixed integer program with an optimality gap g. As a consequence, from the Algorithm 1, we have to compute the mean stochastic optimality gap by computing its lower and upper values as $\overline{G}_{n_G}^l$ and $\overline{G}_{n_G}^u$, respectively.

4. Case study

4.1. Climate change data

Although climate experts define several potential climate paths known as representative concentration pathways [48], in this research, we are more interested in how climate change influences forest growth rather than the actual climate change parameters such as precipitation, temperature, etc. There exist two paradigms to translate climate data into forest growth. The two paradigms are known as process-based modeling and statistical or empirical modeling. Process-based models are known for offering the flexibility to integrate different interactions that explain forest growth. However, these models are mostly suitable for short rotation forests such as eucalyptus plantations [35, 2, 16] which are not the kind of forests in the Pacific Northwest. It goes without saying that the empirical modeling is more suitable for this study. Hence, using empirical modeling, Latta et al. [28] showed that in the Pacific Northwest, the four climate paths forecast by climate experts will affect forest growth disparately. The authors predicted also what the change of potential forest growth will be in 100 years. As results, they provided the potential mean annual increment (pMAI) of forests for the year 2100.

In Table 1, *pMAI* refers to the potential mean annual increment change, which is the average forest growth change in one year. The value in the table represents the potential mean annual increment change that will be observed in 100 years. It is assumed that the change will be linear from now (year 2020) up to that year. The four climate paths, A1B, A2, B1, and Commit (hereafter referred to as C) correspond to different climate forecast in response to human activities, technological advancement, population growth, etc. The values in Table 1, represents the expected pMAI, however, the values were calculated from a spatial auto-regressive model developed in Latta et al. [29]. The model used as input environmental parameters such as the slope, air moisture, temperature, precipitation, and predicted forest growth. As results, there is a large uncertainty on the predicted pMAI and the prediction interval is quite large

[29]. To account for this uncertainty, we use the prediction interval instead of the expected pMAI in harvest scheduling models.

Table 1: *Potential mean annual increment change*

	A1B	A2	B1	Commit
pMAI ($m^3/ha/year$)	2.5	3.1	1.3	0.2

4.2. Scenario trees

To generate scenario trees, we use the conditional sampling method described in Section 3.1. Forest growth change used is the one presented in Table 1. The statistical model predicting the forest growth change had a root mean squared error δ. We use the error term δ to build the 99% prediction interval of the growth change associated with each climate path. For practicality, we used the predicted growth change $\pm 3\delta$ ($pMAI \pm 3\delta$) to build the lower and upper bounds of the predicted growth change. We generated scenario trees for each climate path by sampling within the lower and upper bounds of the prediction interval (see Section 3.1 for more details).

To generate the scenarios for each climate path, we propose supposing forest growth change ξ, within each prediction interval follows either a normal or uniform distribution. The normal distribution assigns high probabilities to scenarios that are closer to the expected predicted forest growth change, whereas uniform distribution assigns the same probabilities to all scenarios within the prediction interval. The objective of having the two distributions is to test the sensitivity of the solutions to different probability schemes.

The planning horizon adopted in this experiment is 50 years divided into five periods (with five decision stages as well). For each distribution, we generated scenario trees by using $\beta = \{2, 3, 4, 5\}$ corresponding to two, three, four, and five branches at each stage and leading to 16, 81, 256, and 625 scenarios, respectively. For each branching scheme, we proceeded into generating ten replications of the

scenario tree with the same structure. These replications served into computing in-sample, out-of-sample stability and the convergence of two climate paths.

There are a few assumptions this research relies on. First, we assume the forest of interest is small enough in size that our management decisions do not significantly affect climate change. Consequently, the stochastic programming is the one with exogenous uncertainty [22, 3]. If we were interested in managing forests at a global level, then we would have had to consider a case where our decisions may affect back climate change. In this latter case, the stochastic programming is known as stochastic programming with endogenous uncertainty [24, 23]. Second, the growth change is linear from one period to another. Third, although forest growth changes, the forest site will still be suitable for the species of interest and therefore there is no need to worry about species shifting from one site to another. This assumption is warranted by the first assumption that the study area is relatively small.

4.3. Optimization model

In this section, we present the optimization model that served as harvest scheduling model. The main objective of harvest scheduling in commercial forests is the maximization of the net present value from the harvest actions during the planning horizon. The list of parameters, variables and sets is reported in Table 2.

4.3.1. Objective function

The objective function which aims at maximizing the expected net present value considering all the scenarios can be written in the form of:

$$\max \sum_{\omega \in \Omega} w^{\omega} \left[\sum_{t \in \mathcal{T}} \left(p_t H_t^{\omega} - \sum_{s \in \mathcal{S}} c_{st} x_{st}^{\omega} \right) + \sum_{s \in \mathcal{S}} p_0 n_s^{\omega} \right]. \tag{17}$$

We make sure to subtract from the revenue we get from the harvest of each forest unit, the cost of replanting the same unit. Hence, we balance between the revenue from harvesting a unit and the cost of replanting it. The objective function (17) maximizes the net present value while meeting various sustainability and logical restrictions.

Table 2: *Nomenclature*

Indices

s	Stand
t	Time of harvest or the year
ω, ω'	Scenario

Variables

H_t^ω	Volume harvested in year t under scenario ω (m^3)
n_s^ω	Binary variable: 1 if stand s should not be harvested during the whole planning horizon under scenario ω; and 0 otherwise
x_{st}^ω	Binary variable: 1 if stand s is scheduled to be harvested in year t under scenario ω; and 0 otherwise
z_{st}	Binary: 1 if stand s has been harvested before the current management such that the stand is not green-up yet at time t; and 0 otherwise. It is not an actual variable since it is defined while building the model

Parameters

A_{max}	Maximum contiguous area that should not be exceeded during harvest for green-up (120 acres in Washington state)
a_s	Area of stand s (ha)
age_{st}	Age of stand s at the end of the planning horizon if harvested in year t (yr)
$age_{s.}$	Current age of the stand s (yr)
age_{s0}	Age of stand s at the end of the planning if not harvested during the planning horizon (yr)
b	Minimum age that a stand can have before it is considered green-up or old enough to not be considered as an opening
c_{st}	Discounted cost of regenerating stand s in year t ($)
f_{max}	Allowable upper bound of percentage of fluctuation of volume of wood
f_{min}	Allowable lower bound of percentage of fluctuation of volume of wood
p_0	Discounted value of the forest at the end of the planning horizon if it is not harvested ($)
p_t	Discounted price of wood in year t under scenario ω ($)
v_{st}^ω	Productivity of stand s if harvested in year t according to scenario ω (m^3/ha)
w^ω	Weight or probability of scenario ω

Sets

$\mathcal{B}$	Set of stands that are big such that their area exceeds A_{max}
$\mathcal{C}$	Set of stands forming a minimum infeasible cluster. (They don't include large units which size exceeds A_{max})
$\mathcal{K}_s$	Set of stands neighbor to the stand s
Λ	Set of minimally infeasible clusters. (They don't include large units which area exceed the maximum opening size)
$\mathcal{S}$	Set of stands
$\mathcal{T}$	Set of years for the planning horizon
Ω	Set of scenarios in the scenario tree

4.3.2. Constraints

(a) A stand can be harvested only once during the planning horizon

$$n_s^\omega + \sum_{t \in \mathcal{T}} x_{st}^\omega = 1 \quad \forall s \in \mathcal{S} \quad \forall \omega \in \Omega \tag{18}$$

(b) Volume of wood harvested each period

$$\sum_{s \in \mathcal{S}} a_s v_{st}^\omega x_{st}^\omega = H_t^\omega \quad \forall t \in \mathcal{T} \quad \forall \omega \in \Omega \tag{19}$$

(c) Wood flow constraints

$$H_t^\omega - (1 - f_{min}) H_{t+1}^\omega \geq 0 \quad \forall t \leq |\mathcal{T}| - 1, \quad \forall \omega \in \Omega \tag{20}$$

$$H_t^\omega - (1 + f_{max}) H_{t+1}^\omega \leq 0 \quad \forall t \leq |\mathcal{T}| - 1, \quad \forall \omega \in \Omega \tag{21}$$

(d) Spatial configuration restrictions (Adjacency rules)

$$\sum_{s \in \mathcal{C}} \left(z_{st} + \sum_{\substack{q=t-b \\ q>0}}^{t} x_{sq}^\omega \right) \leq |\mathcal{C}| - 1 \quad \forall t \in \mathcal{T}, \quad \forall \mathcal{C} \in \Lambda, \quad \forall \omega \in \Omega \tag{22}$$

$$x_{st}^\omega + z_{it} + \sum_{\substack{q=t-b \\ q>0}}^{t} x_{iq}^\omega \leq 1 \quad \forall t \in \mathcal{T}, \quad \forall s \in \mathcal{B}, \quad \forall i \in \mathcal{K}_s, \quad \forall \omega \in \Omega \tag{23}$$

$$\sum_{i \in \mathcal{K}_s} \sum_{\substack{q=t-b \\ q>0}}^{t} x_{iq}^\omega \leq M(1 - x_{st}^\omega - z_{st}) \quad \forall t \in \mathcal{T}, \quad \forall s \in \mathcal{B}, \quad \forall \omega \in \Omega \tag{24}$$

$$z_{st} = 1 \quad \text{if } s \text{ harvested in } t' \text{ and } t < b + t' \text{ with } t' \in [-b, 0] \tag{25}$$

(e) Ending inventory (ending age)

$$\sum_{s \in \mathcal{S}} a_s \left[\sum_{t \in \mathcal{T}} age_{st} x_{st}^\omega + age_{s0} n_s^\omega \right] \geq \sum_{s} a_s age_{s.} \quad \forall \omega \in \Omega \tag{26}$$

(f) Non anticipativity

$$x_{st}^\omega = x_{st}^{\omega'} \quad \omega \neq \omega' \quad \forall s, t \tag{27}$$

If two scenarios ω and ω' are indistinguishable at time t, then the decision in the two scenarios should be the same at that time t.

(g) Nature of variables

$$H_t^{\omega} \in \mathbb{R}^+,\ x_{st}^{\omega} \in \{0,1\},\ n_{st}^{\omega} \in \{0,1\} \quad \forall t \in \mathcal{T}, \quad \forall s \in \mathcal{S}, \quad \forall \omega \in \Omega \tag{28}$$

4.3.3. Meaning of the constraints

Constraint set (18) states that each forest unit can only be harvested once for each scenario. We use the variable n_s^{ω} as a counter of the stands that are not scheduled for harvest during the whole planning horizon. The set of constraints (19) computes the volume of wood harvested in each period for each scenario. As we can see, the parameter v_{st}^{ω} depends on the scenario ω. Constraint sets (20) and (21) are volume flow restrictions and they ensure that the volume of wood harvested in period t is within f_{min} and f_{max} percentage of the one harvested in the period $t-1$. The set of constraints (26) states that the age of the forest at the end of the planning horizon should be greater or equal to the current age of the forest. This set of constraints is a proxy for sustainability; it ensures that resources are not depleted during the planning horizon. Constraint set (27) imposes non-anticipativity for scenario ω and ω'. It states that if there are two scenarios ω and ω' that are indistinguishable in time t, then the decision should be the same for the two scenarios up to that time t.

Constraints (22), (23), (24) and (25) refer to the green-up constraints. Green-up constraints or green-up rules are a set of regulations that aim at limiting the size of the openings and the length of time before adjacent forest units can be harvested. A_{max} is the maximum opening area that contiguous forest units harvested can create. A forest unit is considered as an opening if the forest in that unit is not older than b. In practice we do not need the variable z_{st} since its values will be defined while building the model. However, for easiness of the model and its readability, that variable was necessary. Constraints (22) say only a feasible cluster is allowed. It includes the fact that there might be stands that have not yet reached the green-up requirement at the start of the planning because those stands were harvested in the previous planning. Constraints (23) impose that a large unit cannot be scheduled for harvest if it is adjacent to any other unit that is not green-up yet. The units that are adjacent and not

green-up yet could be of two sources. They could originate from the previous harvest planning in which case the stands on that unit are less than b years old prior to this planning. The second case is that the stand is harvested either in an anterior or the current year. Constraint set (24) says that if there is a large stand harvested then no neighbor to that stand can be eligible for harvest. Notice that for any t, $x^{\omega}_{st} + z_{st} \leq 1, \forall \omega, s$. This equation is deactivated if the large stand is unscheduled for harvest (M is a big number). Constraint set (25) just informs that the values of z are defined while building the model. $z_{st} = 1$ if the stand has been harvested in the prior management such that it is still considered not green-up yet in time t.

During implementation, Constraint set (22) is written slightly differently as follows:

$$\sum_{s \in \mathcal{C}} \sum_{\substack{q=t-b \\ q>0}}^{t} x^{\omega}_{sq} \leq \max\left(0, |\mathcal{C}| - 1 - \sum_{s \in \mathcal{C}} z_{st}\right) \quad t \in \mathcal{T}, \quad \forall \mathcal{C} \in \Lambda, \quad \forall \omega \in \Omega \quad (29)$$

This reformulation is important because the value of $\sum_{s \in \mathcal{C}} z_{st}$ can be greater than $|\mathcal{C}| - 1$. Let's for example have a cluster of two stands ($|\mathcal{C}| = 2$) that were harvested in the previous planning such that the two stands are not green-up yet at t^*, thus $\sum_{s \in \mathcal{C}} z_{st^*} = 2$. With the new formulation, the right-hand side will be zero; as a result, the cluster cannot be harvested until the stands have reached the harvestable age. This situation arises because we do not have control over previous managements and how minimally infeasible clusters were defined. Equations (28) define the domain of the variables.

4.3.4. Recourse constraints

Although the above model is sufficient to define the harvest scheduling, we still need to define the structure of the recourse variables in case the decision maker assumes a climate path, say A2, will occur and actually climate path C, for instance, materializes. Notice that except the wood flow constraints (Constraint sets (20) and (21)), all other constraints will remain satisfied. Hence, to satisfy the wood flow constraints, we suppose that in case of shortage of wood in any time period, because we were expecting one climate path and a different

one occurred, we have to purchase wood on the market from a competitor to fulfill the demand. It goes without saying that the price at which we buy the wood exceeds the price at which we would sell ours. The rational behind this reasoning is that the competitor has no incentive to sell to us their wood. In sum, we incur a high cost for not having enough inventory to fulfill the implicit demand. Similarly, in case of excess of wood, we incur a cost of holding excessive inventory. Wood flow constraints are one of the most important policy in forest management planning. It ensures employment and the stability of the involved local communities [31]. To implement these policies, we need to define a set of variables and parameters.

First, let π_t^+ be the cost of holding inventory $e_t^{+\omega}$ in case the wood we produce in period t exceeds the maximum allowable harvest in period t ($\overline{H}_t^{\omega}$) from the wood flow constraints of the scenario ω. Second, let π_t^- be the price at which we buy wood on the market in case we have a shortage of production if the wood we produce in period t is bellow the minimum volume, $\underline{\mathrm{H}}_t^{\omega}$, required by the wood flow constraints by a value of $e_t^{-\omega}$. We can compute the shortage volume and the excess of inventory by using (30) and (31), respectively.

$$e_t^{-\omega} = \max(\underline{\mathrm{H}}_t^{\omega} - H_t^{\omega}, 0) \tag{30}$$

$$e_t^{+\omega} = \max(H_t^{\omega} - \overline{H}_t^{\omega}, 0) \tag{31}$$

Note that $e_t^{+\omega} * e_t^{-\omega} = 0,\ \forall t, \omega$ because we cannot have at the same time shortage and excess. The new objective function becomes

$$\max \sum_{\omega \in \Omega} w^{\omega} \left[\sum_{t \in \mathcal{T}} \left(p_t H_t^{\omega} - \pi_t^+ e_t^{+\omega} - \pi_t^- e_t^{-\omega} - \sum_{s \in \mathcal{S}} c_{st} x_{st}^{\omega} \right) \right]. \tag{32}$$

This objective is necessary in computing the net present value if the decision maker plans for one climate path and actually a different climate path occurs. Since the decision variables are fixed, this is just computing the new objective function value of wrongly assuming the future climate path.

4.3.5. Values of the model's parameters

The methodology was tested on Phyllis Perry forest with 89 stands at different maturity age. The planning horizon was 50 years divided into five planning periods of ten years each.

Table 3: *Optimization model parameters*

Parameter	Value
b	5 years
f_{min}	15%
f_{max}	15%
π_t^+	$0.8p_t$
π_t^-	$2p_t$
Discount rate	3.5%
Planning horizon	50 years
Length of a period	10 years

We solved the models for all data sets using IBM ILOG CPLEX 64-bit 12.9.0 on a Dell Power Edge 510 Server with an Intel Xeon(R) CPU, E5-2680 v3 @2.50 GHz (two processors) with 256 GB RAM and the Windows Server 2012R2 64-bit operating system. The optimization model was implemented on Python. We ran CPLEX using the default settings except limiting the run time to 24h (wall clock time) and setting the optimality gap (g) to 2% for each model. We excluded from the analysis all models that terminated because of time limit. Table 3 summarises the values of several parameters used in the optimization model. To check for stability, we generated ten replications of each scenario tree structure.

4.4. Experimental results

4.4.1. In-sample stability

In-sample stability increases with the number of scenarios (the number of branches at each stage) for the four climate paths (Figure 3). The relative

stability presented in Table 4 shows that except for the climate path B1, when the random parameter is uniformly distributed, which has a relative stability less than 1% when considering three branches at each stage ($\beta = 3$), all other climate paths reach that stability level when four branching is considered ($\beta = 4$). The four branching at each stage corresponds to 256 growth scenarios.

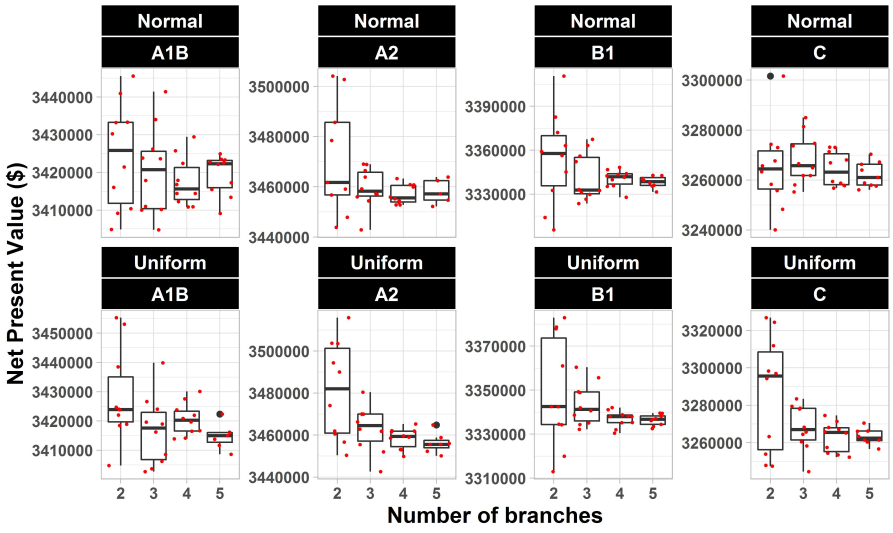

Figure 3: *In-sample stability analysis*

Table 4: *Percentage of stability using maximum and minimum values for both normal and uniform distributions (relative stability)*

	Normal				Uniform			
Branch (# scenarios)	A1B	A2	B1	C	A1B	A2	B1	C
2 (16)	1.21	2.15	3.06	1.65	1.63	1.88	2.49	2.34
3 (81)	1.20	1.37	1.34	1.11	1.05	1.09	0.86	1.46
4 (256)	0.69	0.44	0.66	0.73	0.43	0.40	0.31	0.59
5 (625)	0.45	0.34	0.35	0.50	0.39	0.38	0.37	0.43

4.4.2. Out-of-sample stability

We get high out-of-sample stability when considering four branching (or 256 scenarios) compared to five branching ($\beta = 5$ or 625 scenarios). However, these stabilities do not appear to be statistically different. The trade-off between the time required to solve the problem with 625 scenarios and the increase of out-of-sample stability does not motivate the adoption of the scheme with five branching. Hence, the best sampling schema is four branching at each stage leading to 256 scenarios (Figure 4).

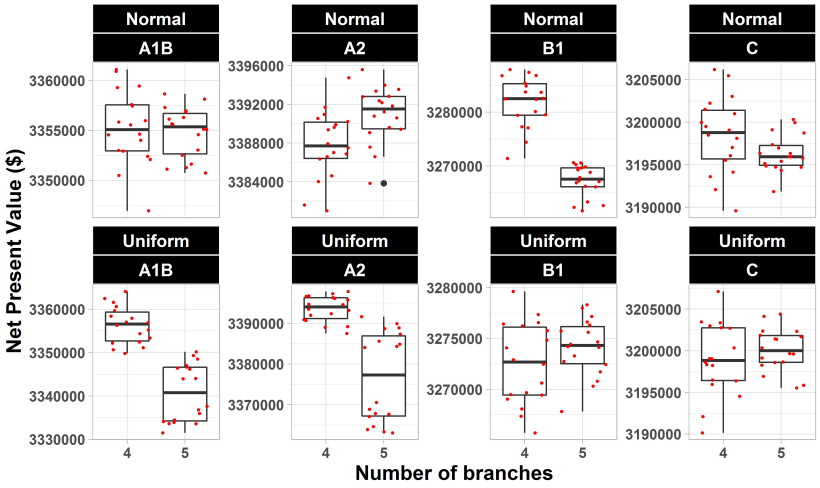

Figure 4: *Out-of-sample stability analysis*

4.4.3. Stochastic optimality gap

The results of the upper and lower bounds on stochastic optimality gap for the four climate paths are reported in Table 5. The lower bound of the stochastic optimality gap ranges from 0.09% to 0.18% corresponding to the climate path A1B and C, respectively, when the growth is normally distributed. However, the upper bound on the stochastic optimality is less than 2%.

Table 5: *Stochastic optimality gap on each climate path. The percentage of the bound is relative to the mean NPV*

Distribution	Climate path	Mean NPV($)	$\overline{G}^{l}_{30}$($)	$\overline{G}^{l}_{30}$(%)	$\overline{G}^{u}_{30}$($)	$\overline{G}^{u}_{30}$(%)
Normal	A1B	3,354,555	2,980	0.09	62,593	1.87
	A2	3,387,797	6,063	0.18	65,974	1.95
	B1	3,278,346	3,914	0.12	60,970	1.86
	C	3,208,195	5,743	0.18	58,427	1.82
Uniform	A1B	3,355,126	4,677	0.14	63,121	1.88
	A2	3,391,432	4,401	0.13	65,150	1.92
	B1	3,272,744	5,609	0.17	63,356	1.94
	C	3,200,770	4,444	0.14	60,942	1.90

4.4.4. Convergence of climate paths

Table 6 displays the expected net present value that we get if we commit to each of the climate paths. The expected NPV is high when we manage the forest expecting climate path A2 regardless of if that climate path actually materializes or not. Managing the forest expecting climate path C leads to the lowest NPV. The results of two-way ANOVA shows that the expected NPV we get depends on the climate path we commit to (Table 7). However, the distribution of the random parameter for each climate path has no significant effect. In other words, the probabilities of the scenarios did not have any significance on the value of the NPV.

Figure 5 presents the summary of the net present value (NPV) of managing the forest expecting one climate path while the materialized climate path may be the same or a different one. For instance, if we manage the forest expecting climate path A1B and climate path A2 occurs, the NPV is expected to increase, although it remains in the uncertainty margin. However, if climate path B1 or C occurs, then the obtained NPV will be much lower than the one we would get if actually the climate A1B had occurred. Similarly, if we expect climate A2, for instance, and climate path C occurs, the NPV obtained is much higher

Table 6: *Expected NPV when the forest is managed expecting a specific climate path regardless of which climate path actually materializes for both the normal and the uniform distributions. SD = standard deviation*

Distribution	Climate	NPV($)	SD(NPV)
Normal	A1B	3,335,126	30,825.77
	A2	3,352,117	33,563.86
	B1	3,288,788	27,300.93
	C	3,244,643	26,940.26
Uniform	A1B	3,333,478	33,689.93
	A2	3,350,849	38,496.11
	B1	3,285,235	29,311.52
	C	3,238,941	28,359.64

Table 7: *Two-way ANOVA of the NPV distribution and the climate path (climate) as factors*

	df	Sum Sq	Mean Sq	F Value	Pr(>F)
Distribution	1	3.70E+08	3.70E+08	0.3784	0.5394
Climate	3	2.93E+11	9.77E+10	99.888	<2e-16
Distribution × Climate	3	1.24E+08	4.14E+07	0.0423	0.9884
Residuals	152	4.87E+10	9.79E+08		

than the NPV we would get if we managed the forest expecting climate path C and it actually materialized. This second analysis may be counter intuitive. It seems like if we knew which climate path would occur, we would make the best decision here and now, and therefore the NPV should be higher for such a good foresee. The second thing one may wonder is whether we are not over-harvesting if we plan foreseeing an optimistic forest growth climate path, say climate path A2, and actually climate path C materializes. The following paragraphs address these two points.

First, when we plan for an optimistic climate path, the best decision is to

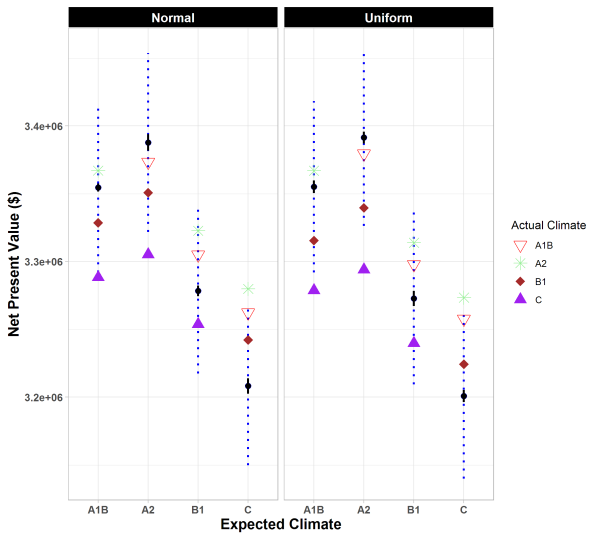

Figure 5: *Net present value from a management expecting one climate path versus the materialization of a different climate path. Dashed blue lines correspond to the confidence of the NPV from a correct prediction of the climate path.*

increase the volume of forest harvested here and now (harvested in the first period) because the future forest growth will increase and compensate the volume we may have over-harvested here and now. Second, the discounting affects revenues and costs that we incur in the future, hence the future actions are less significant. It is clear, therefore, that if we only care about the ending age inventory requirement (constraint set (26)) and the wood flow, the best decision is to plan for optimistic future forest growth and if we lack forest material in the future, we can still purchase wood even at a higher cost as long as the ending age constraint is satisfied.

Second, we can see as illustrated on Figure 6, that the volume of wood harvested is the same by the end of the planning horizon, regardless of if the management anticipated correctly the climate path to occur or not. Indeed, if we expect an optimistic climate path, say A2, we tend to increase the volume of wood harvested in the first period. However, if a less optimistic climate path, say C, occurs, we will harvest less volume in the future compared to the volume

we would have harvested if we knew from the beginning that the climate path C would occur.

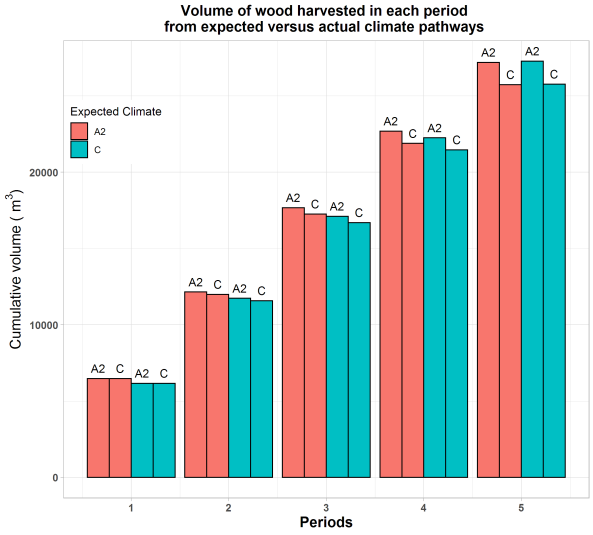

Figure 6: *Cumulative volume of wood harvested in each period when planning for either climate paths A2 and C and actually any of the two materializes for normally distributed growth change. The letters on the bars designate the actual climate path that occurred*

5. Conclusions and discussion

In this study, we have developed a framework for incorporating forest growth and yield uncertainty due to climate change in harvest scheduling models. We considered forest growth change with different climate paths separately. We generate growth scenarios within the prediction interval of each climate path. Increasing the number of scenarios allows to capture the variability of the forest growth that affects optimal management decisions. However, there is no real benefit at increasing that number beyond a given point since it increases the problem computational complexity without significantly increasing the quality of the scenario tree. The analysis showed that out of the four climate paths considered, one may just focus on the two extreme climate paths, namely climate

path A2 and C since the decisions one makes considering the two others (climate path A1B and B2) are included in the decision if these two climate paths.

Subsequently, at the high level, it does not really matter which climate path will materialize; the optimal decision is to plan forest harvests expecting an optimistic climate, which in this study is climate path A2. Whether the uncertainty of the forest response is normally or uniformly distributed around its expected value, the results have shown that there is no significant impact considering one or the other.

Although, we advocate that the decision maker ought to have an optimistic attitude expecting a future increase of the forest growth in the context of Pacific northwest, this recommendation can be viewed with skepticism especially for a large forestry company. The idea of purchasing forest products in case of shortage makes sense for a small company but may not be appealing to a large company that is the leading in the industry. Similarly, we would like to stress that this recommendation has to be taken with caution since it is not a sustainable practice if the whole forest industry adopt it. One way this caution can be integrated in the model is to reinforce the wood flow constraints to be feasible for all climate paths.

One of the limitations of this research is that we supposed that climate change will affect forest growth without any feedback from our management decisions. Although this assumption was valid for our case, it is a limitation of this research. Similarly, we supposed that if there is a shortage of wood supply because of our management decisions, then we could procure some wood from competitors. This might not be true if all decision makers adopt this attitude. Moreover, supposing just two distributions of the random variable may be limiting. In the continuation of this research, we intend to use distributionally robust optimization to integrate more distributions in the harvest scheduling and at the same time consider all the climate paths at once. Finally, we supposed in this research that the price of wood is independent from climate change. Although we do not have enough evidence to refute this assumption, we do think it might not be the case. Wood can become a prime commodity if climate

change leads to a substantial reduction in forest growth. However, this change can be attenuated by the technological advancement that can reduce the need of wood for construction.

[1] Alonso-Ayuso, A., Escudero, L. F., Guignard, M., and Weintraub, A. (2018). Risk management for forestry planning under uncertainty in demand and prices. *European Journal of Operational Research*, 267(3):1051–1074.

[2] Álvarez-Miranda, E., Garcia-Gonzalo, J., Ulloa-Fierro, F., Weintraub, A., and Barreiro, S. (2018). A multicriteria optimization model for sustainable forest management under climate change uncertainty: An application in Portugal. *European Journal of Operational Research*, 269(1):79–98.

[3] Apap, R. M. and Grossmann, I. E. (2017). Models and computational strategies for multistage stochastic programming under endogenous and exogenous uncertainties. *Computers and Chemical Engineering*, 103:233–274.

[4] Bagaram, M. B. and Tóth, S. F. (2020). Multistage sample average approximation for harvest scheduling under climate uncertainty. *Forests*, 11(11):1230.

[5] Bagaram, M. B., Tóth, S. F., Jaross, W. S., and Weintraub, A. (2020). A parallelized variable fixing process for solving multistage stochastic programs with progressive hedging. *Advances in Operations Research*, 2020:117.

[6] Bayraksan, G. and Morton, D. P. (2006). Assessing solution quality in stochastic programs. *Mathematical Programming*, 108(2-3):495–514.

[7] Beraldi, P. and Bruni, M. E. (2013). A clustering approach for scenario tree reduction: An application to a stochastic programming portfolio optimization problem. *Top*, 22(3):1–16.

[8] Casey, M. S. and Sen, S. (2005). The Scenario Generation Algorithm for Multistage Stochastic Linear Programming. *Mathematics of Operations Research*, 30(3):615–631.

[9] Chunlin, D. and Liu, Y. (2012). Sample Average Approximation Method for Chance Constrained Stochastic Programming in Transportation Model of Emergency Management. *Systems Engineering Procedia*, 5:137–143.

[10] Crespi, G. P., Kuroiwa, D., and Rocca, M. (2018). Robust optimization: Sensitivity to uncertainty in scalar and vector cases, with applications. *Operations Research Perspectives*, 5:113–119.

[11] De Pellegrin Llorente, I., Eyvindson, K., Mazziotta, A., Lms, T., Eggers, J., and hman, K. (2023). Perceptions of uncertainty in forest planning: Contrasting forest professionals perspectives with the latest research. *Canadian Journal of Forest Research*.

[12] Dembo, R. S., Chiarri, A., Martin, J. G., and Paradinas, L. (1990). Managing Hidroeléctrica Española's Hydroelectric Power System. *Interfaces*, 20(1):115–135.

[13] Dupacová, J., Consigli, G., and Wallace, S. W. (2000). Scenarios for Multistage Stochastic Programs. *Annals of Operations Research*, 100:25–53.

[14] Elli, E. F., Sentelhas, P. C., and Bender, F. D. (2020). Impacts and uncertainties of climate change projections on Eucalyptus plantations productivity across Brazil. *Forest Ecology and Management*, 474(March):118365.

[15] Fleishman, A. I. (1978). A method for simulating non-normal distributions. *Psychometrika*, 43(4):521–532.

[16] Garcia-Gonzalo, J., Pais, C., Bachmatiuk, J., Barreiro, S., and Weintraub, A. (2020). A Progressive Hedging Approach to Solve Harvest Scheduling Problem under Climate Change. *Forests*, 11(2):224.

[17] Gröwe-Kuska, N., Heitsch, H., and Römisch, W. (2003). Scenario reduction and scenario tree construction for power management problems. *2003 IEEE Bologna PowerTech - Conference Proceedings*, 3:152–158.

[18] Gülpinar, N., Rustem, B., and Settergren, R. (2004). Simulation and optimization approaches to scenario tree generation. *Journal of Economic Dynamics and Control*, 28(7):1291–1315.

[19] Guo, Z., Wallace, S. W., and Kaut, M. (2019). Vehicle Routing with Space- and Time-Correlated Stochastic Travel Times: Evaluating the Objective Function. *INFORMS Journal on Computing*, Articles i(September):1–17.

[20] Heitsch, H. and Römisch, W. (2009). Scenario tree modeling for multistage stochastic programs. *Mathematical Programming*, 118(2):371–406.

[21] Hochreiter, R. and Pflug, G. C. (2007). Financial scenario generation for stochastic multi-stage decision processes as facility location problems. *Annals of Operations Research*, 152(1):257–272.

[22] Hooshmand, F. and MirHassani, S. (2016). Efficient constraint reduction in multistage stochastic programming problems with endogenous uncertainty. *Optimization Methods and Software*, 31(2):359–376.

[23] Hooshmand, F. and MirHassani, S. A. (2018). Reduction of nonanticipativity constraints in multistage stochastic programming problems with endogenous and exogenous uncertainty. *Mathematical Methods of Operations Research*, 87(1).

[24] Hooshmand Khaligh, F. and MirHassani, S. (2016). A mathematical model for vehicle routing problem under endogenous uncertainty. *International Journal of Production Research*, 54(2):579–590.

[25] Høyland, K., Kaut, M., and Wallace, S. W. (2003). A Heuristic for Moment-Matching. *Computational Optimization and Applications*, 24:169–185.

[26] Høyland, K. and Wallace, S. W. (2001). Generating Scenario Trees for Multistage Decision Problems. *Management Science*, 47(2):295–307.

[27] Kaut, M. and Wallace, S. W. (2007). Evaluation of scenario-generation methods for stochastic programming. *Pacific Journal of Optimization*, 3(2):257–271.

[28] Latta, G., Temesgen, H., Adams, D., and Barrett, T. (2010). Analysis of potential impacts of climate change on forests of the United States Pacific Northwest. *Forest Ecology and Management*, 259(4):720–729.

[29] Latta, G., Temesgen, H., and Barrett, T. M. (2009). Mapping and imputing potential productivity of Pacific Northwest forests using climate variables. *Canadian Journal of Forest Research*, 39(6):1197–1207.

[30] Li, Q. Q., Li, Y. P., Huang, G. H., and Wang, C. X. (2018). Risk aversion based interval stochastic programming approach for agricultural water management under uncertainty. *Stochastic Environmental Research and Risk Assessment*, 32(3):715–732.

[31] Liu, C. M. (2002). A Primal-dual Steepest-edge Method for Even-flow Harvest Scheduling Problems. *International Transactions in Operational Research*, 9(1):33–50.

[32] Löhndorf, N. (2016). An empirical analysis of scenario generation methods for stochastic optimization. *European Journal of Operational Research*, 255(1):121–132.

[33] Mak, W. K., Morton, D. P., and Wood, R. K. (1999). Monte Carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters*, 24(1):47–56.

[34] Marla, L., Rikun, A., Stauffer, G., and Pratsini, E. (2020). Often, the RO Approach and the CCPRobust Modeling and Planning: Insights from Three Industrial Applications. *Operations Research Perspectives*, page 100150.

[35] Miehle, P., Battaglia, M., Sands, P. J., Forrester, D. I., Feikema, P. M., Livesley, S. J., Morris, J. D., and Arndt, S. K. (2009). A comparison of four process-based models and a statistical regression model to predict growth of Eucalyptus globulus plantations. *Ecological Modelling*, 220(5):734–746.

[36] Moriguchi, K., Ueki, T., and Saito, M. (2020). Establishing optimal forest harvesting regulation with continuous approximation. *Operations Research Perspectives*, 7(April):100158.

[37] Ponomareva, K., Roman, D., and Date, P. (2015). An algorithm for moment-matching scenario generation with application to financial portfolio optimisation. *European Journal of Operational Research*, 240(3):678–687.

[38] Pour, A. G., Naji-Azimi, Z., and Salari, M. (2017). Sample average approximation method for a new stochastic personnel assignment problem. *Computers and Industrial Engineering*, 113:135–143.

[39] Pranevicius, H. and Sutiene, K. (2007). Scenario tree generation by clustering the simulated data paths. *Proceedings of 21st European Conference on Modeling and Simulation*, page 203208.

[40] Rios, I., Weintraub, A., and Wets, R. J. (2016). Building a stochastic programming model from scratch: a harvesting management example. *Quantitative Finance*, 16(2):189–199.

[41] Ross, K. L. and Tóth, S. F. (2016). A model for managing edge effects in harvest scheduling using spatial optimization. *Scandinavian Journal of Forest Research*, 31(7):646–654.

[42] Schroder, S. A., Tóth, S. F., Deal, R. L., and Ettl, G. J. (2016). Multiobjective optimization to evaluate tradeoffs among forest ecosystem services following fire hazard reduction in the Deschutes National Forest, USA. *Ecosystem Services*, 22:328–347.

[43] Schütz, P., Tomasgard, A., and Ahmed, S. (2009). Supply chain design under uncertainty using sample average approximation and dual decomposition. *European Journal of Operational Research*, 199(2):409–419.

[44] Séguin, S., Fleten, S. E., Côté, P., Pichler, A., and Audet, C. (2017). Stochastic short-term hydropower planning with inflow scenario trees. *European Journal of Operational Research*, 259(3):1156–1168.

[45] St. John, R. and Tóth, S. F. (2015). Spatially explicit forest harvest scheduling with difference equations. *Annals of Operations Research*, 232:235257.

[46] St John, R., Tóth, S. F., and Zabinsky, Z. B. (2018). Optimizing the geometry of wildlife corridors in conservation reserve design. *Operations Research*, 66(6):1471–1485.

[47] Tóth, S. F., Haight, R. G., Snyder, S. A., George, S., Miller, J. R., Gregory, M. S., and Skibbe, A. M. (2009). Reserve selection with minimum contiguous area restrictions: An application to open space protection planning in suburban Chicago. *Biological Conservation*, 142(8):1617–1627.

[48] van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J. F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., and Rose, S. K. (2011). The representative concentration pathways: An overview. *Climatic Change*, 109(1):5–31.

[49] Wang, W. and Ahmed, S. (2008). Sample average approximation of expected value constrained stochastic programs. *Operations Research Letters*, 36(5):515–519.

[50] Xu, B., Zhong, P.-A., Zambon, R. C., Zhao, Y., and Yeh, W. W.-G. (2015). Scenario tree reduction in stochastic programming with recourse for hydropower operations. *Water Resources Research*, 51(8):6359–6380.