On the number of stages in multistage stochastic programs

Giovanni Pantuso,
Department of Mathematical Sciences, University of Copenhagen,
Universitetsparken 5, 2100 Copenhagen, Denmark
pantuso@math.ku.dk

January 11, 2016

Abstract

Multistage stochastic programs are a viable modeling tool for sequential decisions conditional on information revealed at different points in time (stages). However, as the number of stages increases their applicability is soon halted by the curse of dimensionality. A typical, sometimes forced, alternative is to approximate stages by their expected values thus considering fewer stages in the resulting model. This paper shows how concepts in the available literature, such as the value of the stochastic solution, can be slightly extended to evaluate the benefit from solving a multistage stochastic program rather than an approximation obtained by reducing the number of stages. A numerical procedure for the calculation of this benefit, as well as bounds, are presented. The procedure is explanatorily applied to the well know investor problem.

1 Introduction

Multistage stochastic programs (MSPs) are a well recognized technique for handling sequential decisions conditional on information revealed at different points in time, so called *stages*. Problems with such a structure arise, for instance, in asset/liability management Cariño et al. (1994); Mulvey and Shetty (2004), capacity expansion Ahmed et al. (2003), energy systems Wallace and Fleten (2003); Fleten and Kristoffersen (2008); Kristoffersen and Fleten (2010), water resources management Li et al. (2006), air transportation Alonso et al. (2000), and maritime transportation Pantuso et al. (2015); Bakkehaug et al. (2014).

An MSP can be formulated in different ways. For the sake of compactness we will adopt the following formulation:

$$\min f(x,\xi) = \left\{ E\left[\sum_{t=1}^{T} c_t^T(\xi_t)x_t\right] \;\middle|\; \begin{array}{l} \sum_{\tau=1}^{t} A_{t\tau}(\xi_t)x_\tau = b_t(\xi_t),\ t = 1\ldots,T, \\ x_t \text{ is } \mathcal{F}_t(\xi_{[1,t]})\text{-measurable},\ t = 1,\ldots,T, \\ x_t \in X_t,\ t = 1\ldots,T \end{array} \right\} \tag{1}$$

where $1,\ldots,T$ are stages, $\xi := (\xi_1,\ldots,\xi_T)$ is a stochastic process defined on some probability space $(\omega,\mathcal{F},P)$, and $x := (x_1,\ldots,x_T)$ is the collection of decisions. There is no uncertainty in ξ_1. MSP (1) requires that decisions x_t are functions of the history of decisions until $t-1$, and that decisions are chosen in a feasibility set X_t which may, in general, require integrality on some (all) of the decision variables. Finally, the requirement that x_t is $\mathcal{F}_t(\xi_{[1,t]})$-measurable, where $\mathcal{F}_t(\xi_{[1,t]}) \subseteq \mathcal{F}_{t+1}(\xi_{[1,t+1]}) \subseteq \mathcal{F}$ is the sigma-algebra generated by $\xi_{[1,t]} := (\xi_1,\ldots,\xi_t)$, ensures that x_t depends only on available information, i.e., is *nonanticipative*. In almost all practical applications ξ is discrete, possibly by assumption. Also, typically, recourse matrices are fixed, i.e., $A_{tt}(\xi_t) = A_{tt}, t = 1,\ldots,T$.

Despite solution methods for MSPs have been proposed (see, e.g., Birge (1985), Løkketangen and Woodruff (1996), Lulli and Sen (2004), Escudero et al. (2009), and Pantuso et al. (2015)) solving a stochastic program which considers all the stages in the real-life problem it attempts to model is often prohibitive. Many problems have an infinite planning horizon and, consequently, infinite stages. In these cases one must rely on a significant portion of the planning horizon – and consequently on a representative number of stages. Also in many finite-horizon problems considering all the stages may pose serious tractability issues. In any case, a typical way out is to solve simpler problems by reducing the number of stages, i.e., by considering the future deterministic after a given stage. In these cases it is relevant for decision makers to know what is the benefit from solving the original problem rather than an approximation.

Metrics in the available literature help evaluating the benefit from solving the original problem. Given a two-stage stochastic program, the *Value of the Stochastic Solution* (VSS), described in Birge (1982), provides the benefit from solving the stochastic program rather than its mean value problem (MVP), i.e., the problem obtained by replacing the random variables in the second stage with their expected values. The procedure consists of solving the original two-stage program and the MVP separately. The expected return from using the first-stage component of the solution to the MVP is then calculated. The difference between this value and the optimal objective to the original two-stage program provides the VSS. The VSS can be calculated also for multistage programs as shown by Escudero et al. (2007) using the *rolling-horizon value of the stochastic solution* also discussed in Maggioni et al. (2013). It provides the benefit from solving a MSP rather than using the MVP every time a decision must be made. The procedure is slightly more complex than for two-stage problems. The MVP is solved in a rolling-horizon framework in order to account for the fact that in a multistage problem decision are made every time new information is obtained. Additional types of deterministic approximations to a MSP are available and are discussed in Maggioni et al. (2013).

In this paper we show how the available metrics can be extended to account for approximations obtained by reducing the number of stages, though without reverting to a deterministic problem. That is, we show how to evaluate the benefit from solving a T-stage stochastic program rather than approximating it by a T'-stage stochastic program with $1 \leq T' \leq T-1$. We refer to this value as the *Marginal Stage Value*. A procedure to numerically evaluate the marginal stage value, as well as bounds, are proposed.

The remainder of the paper is organized as follows. In Section 2 we introduce the marginal stage value while in Section 3 we provide bounds on its value. In Section 4 the concepts presented are explanatory applied to the well know investment problem, see Birge and Louveaux (1997, p. 20). Finally, conclusions are drawn in Section 5.

2 The Marginal Stage Value

We want to evaluate the benefit from solving a T-stage stochastic program rather than a T'-stage approximation, with $1 \leq T' \leq T-1$, obtained by replacing stages with their expected values. This quantity will be referred to as the *Marginal Value of the T-th Stage with respect to the T'-th*, or simply as the Marginal Stage Value, ($MSV_{T,T'}$).

In order to calculate the $MSV_{T,T'}$ initially we solve the original T-stage problem (1). Let z^T be its optimal objective value. When using a T'-stage approximation we make first-stage decisions by replacing the T-stage stochastic process $\xi = (\xi_1, \ldots, \xi_T)$ with $\xi^{1,T'} := (\xi_1^{1,T'}, \ldots, \xi_T^{1,T'}) = (\xi_1, \ldots, \xi_{T'}, E[\xi_{T'+1}], \ldots, E[\xi_T])$ and solving $\min\ f(x, \xi^{1,T'})$. Notice that in $\xi^{1,T'}$ there is no uncertainty after T', rendering it a T'-stage stochastic process. Let $x_1^{T'}$ be the first-stage component of $x^{T'} = \arg\min\ f(x, \xi^{1,T'})$ – this is the only component of the solution which would be actually implemented. Upon reaching the t-th stage the history $\xi_{[1,t]}$ becomes known and new conditional decisions are made. Again, these decisions are made based on a T'-stage program[1]. This means that the remainder stochastic process $\xi_t, \ldots, \xi_T$, is replaced by $\xi^{t,T'} := (\xi_t, \ldots, \xi_{t+T'-1}, E[\xi_{t+T'}], \ldots, E[\xi_T])$. The history of previous T'-stage-based decisions, $x_{[1,t-1]}^{T'} := (x_1^{T'}, \ldots, x_{t-1}^{T'})$ is data. This corresponds to solving

[1]Unless fewer stages than T' remain until the end of the planning horizon.

problem

$$\min f(x_t,\dots,x_T,\xi^{t,T'}) = \left\{ E\left[\sum_{\tau=t}^{T} c_\tau^T(\xi_\tau^{t,T'})x_\tau \middle| \xi_{[1,t]}\right]\right.$$

$$\left.\begin{array}{|l} \sum_{\tau=1}^{t-1} A_{t',\tau}(\xi_{t'}^{t,T'})x_\tau^{T'} + \sum_{\tau=t}^{t'} A_{t',\tau}(\xi_{t'}^{t,T'})x_\tau = b_{t'}(\xi_{t'}^{t,T'}),\ t'=t,\dots,T, \\ x_{t'} \text{ is } \mathcal{F}_{t'}(\xi_{[1,t']})\text{-measurable},\ t'=t,\dots,T, \\ x_{t'} \in X_{t'},\ t'=t,\dots,T \end{array}\right\}$$

Let $x_t^{T'}$ be the component of its optimal solution for stage t. We collect such solution component for stages $1,\dots,T-T'$. The reason is that for stages $T-T'+1,\dots,T$ solving a T'-stage problem would not be an approximation anymore. As an example, consider a five-stage problem, approximated by a two-stage problem. When reaching stage $T-T'=5-2=3$, the original problem has still three stages to go (i.e., the current stage, the fourth and the fifth), therefore a two-stage approximation would replace the fifth stage with its expected values. When reaching stage $T-T'+1=4$, there are two stages left, the current and the fifth. A two-stage problem would not be an approximation anymore.

Once T'-stage based decisions $x_{[1,T-T']}^{T'}$ have been collected, we need to evaluate their performance, i.e., the expected return from solving a T'-stage approximation. This can be done by solving min $f((x_1^{T'},\dots,x_{T-T'}^{T'},x_{T-T'+1},\dots,x_T),\xi)$, where $x_{[1,T-T']}^{T'}$ is input data, and we solve for $x_{T-T'+1}$, $\dots,x_T$. Let $E_{T'}$ be its optimal objective value representing the *expected return from using a T'-stage approximation*. We can then define the *marginal value of the T-th stage with respect to the T'-th* as:

$$MSV_{T,T'} = E_{T'} - z^T$$

Notice that $MSV_{2,1}$ corresponds to the "Value of the Stochastic Solution" (VSS) for two-stage stochastic programs (see Birge (1982)) and $MSV_{T,1}$ corresponds to the "Dynamic Value of the Stochastic Solution" (see Escudero et al. (2007)) and to the "Rolling Horizon Value of the Stochastic Solution" whit the mean value problem used as the reference scenario (see Maggioni et al. (2013)).

Proposition 1. *For any multistage stochastic program $MSV_{T,T'} \geq 0$.*

Proof. For the proof it is enough to observe that $E_{T'}$ and z^T are obtained by solving the same MSP, with the exception that the variables for stages $1,\dots,T-T'$ are fixed in the case of $E_{T'}$ producing, in general, a suboptimal solution. □

3 Computation and Bounds

In practice stochastic processes are represented by scenario trees. Consider a four-stage stochastic program. We want to evaluate what is the benefit from solving the original four-stage problem rather than a two-stage approximation. Let us follow the procedure illustrated in Fig. 1. In Step 1 we solve the two-stage approximation and store its first-stage decision, $x_1^{T'} = x_1^2$. In step 2, for each of the possible nodes in the second stage (i.e., outcomes of the second-stage random variable) we make a decision based on a two-stage program. That is, we generate a two-stage scenario sub-tree rooted in each of the two nodes at the second stage. These sub-trees are conditional on the information obtained in their respective second-stage nodes and on their history. We solve the resulting stochastic programs, constrained by x_1^2, and store the respective conditional decisions for the second stage, x_{2n}^2 for nodes $n=1,2$. Notice that the decision we make for each second stage node are, in general, different from the decision that the original four-stage program (constrained by x_1^2) would make for them. Finally, in Step 3 we pick up the original four-stage program and we fix its first- and second-stage decisions to the corresponding decisions we stored in the previous steps, namely x_1^2, x_{21}^2 and x_{22}^2. We solve the resulting four-stage stochastic program, whose optimal objective value represents the *expected return from using a two-stage approximation*, E_2. This quantity can be compared to the optimal objective value to the original four-stage program z^4 and we can calculate $MSV_{4,2} = E_2 - z^4$. Notice that, when calculating $MSV_{T,T-1}$, in order to calculate $E_{T'}$ one needs to solve only one T'-stage problem.

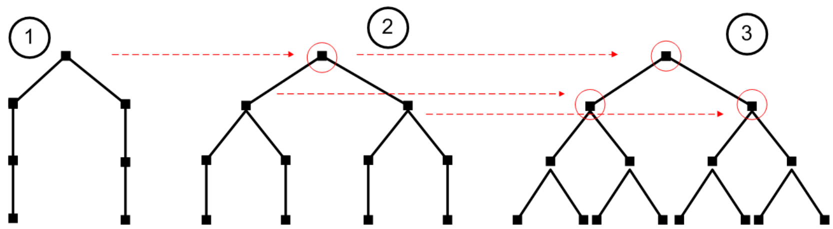


Figure 1: Procedure for the calculation of the expected return from using two stages instead of four

The computation of $MSV_{T,T'}$ requires solving, in general, many stochastic programs. While the computation of $E_{T'}$ only requires solving T'-stage problems, obviously the real difficulty lies in the solution of the original T-stage stochastic program. In cases when this can be solved, though paying a price in terms of complexity and solution time, $MSV_{T,T'}$ can be directly calculated. When the original T-stage problem cannot be solved bounds on $MSV_{T,T'}$ become useful in practice.

Let the *wait-and-see* value be the following quantity:

$$WS = E_{\xi}[\min f(x, \xi)]$$

WS represents the expected value of the anticipative solutions, i.e., of the optimal solution for each possible scenario.

Proposition 2. *For any T-stage stochastic program and $1 \leq T' \leq T$ the following inequalities hold:*

$$WS \leq z^T \leq E_{T'} \tag{2}$$

Proof. For the first inequality we observe that, for any realization $\hat{\xi}$ of ξ we have:

$$\min_x z(x, \hat{\xi}) \leq z(x^*_{\hat{\xi}}, \hat{\xi})$$

where $x^*_{\hat{\xi}}$ represents the components of the optimal solution to the multistage stochastic program for scenario $\hat{\xi}$. For the scenario-problem in fact $x^*_{\hat{\xi}}$ is just a feasible solution. The expectation of both sides provides the first inequality in (2). The second inequality follows from Proposition 1 and from the definition of $MSV_{T,T'}$. □

An immediate consequence of Proposition 2 is the following:

Proposition 3. *For any T-stage stochastic program and $1 \leq T' \leq T$ the following inequality holds:*

$$MSV_{T,T'} \leq E_{T'} - WS$$

Proof. Follows directly from Proposition 2. □

Notice that the calculation of these bounds require at most the solution of T'-stage problems, which are assumed to be easier than the original T-stage problem.

4 Example: The Investor Problem

In this section, the marginal stage value is explanatorily calculated for the "investor problem" (see Birge and Louveaux (1997)).

An investor is to decide how to invest \$55 in order to have, after 15 years, an amount at least equal to \$80 - their children's tuition fee. There are two possible investment types, namely stocks and bonds, and the investment strategy can be revisited every five years in light of new information. With this setting, the problem has four stages. At the end of the planning horizon, exceeding the planned \$80 tuition fee will generate an income of 1% of the excess, while not meeting the goal will require borrowing at 4%. The return of stocks and bonds is uncertain. After each five-year period, two equally likely scenarios may occur, having returns of 1.25 for stocks and 1.14 for bonds the first and 1.06 for stocks and 1.12 for bonds the second. This generates a total of eight scenarios over the four stages. Mathematical model (3) formalizes the problem. Notice that the problem is expressed as a minimization problem in order to make the discussion consistent with the previous sections.

$$
\begin{aligned}
\min & \sum_{n \in N_{\bar{T}}} p_n(-q^+ y_n^+ + q^- y_n^-) && \\
\text{s.t.} & \sum_{i \in I} x_{i1} = 55, && \\
& \sum_{i \in I} x_{in} - \sum_{i \in I} r_{in} x_{i,a(n)} = 0, && n \in N_t, t = 2, \ldots, \bar{T} - 1, \\
& \sum_{i \in I} r_{in} x_{i,a(n)} - y_n^+ + y_n^- = 80, && n \in N_{\bar{T}}, \\
& x_{in} \geq 0, && i \in I, n \in N, \\
& y_n^+, y_n^- \geq 0, && n \in N_{\bar{T}}
\end{aligned}
\tag{3}
$$

Here, $T = \{1, \ldots, \bar{T}\}$ is the set of stages, with $\bar{T} = 4$ in our case. N represents the set of nodes in the scenario tree, N_t the set of nodes at stage t and I the set of investment options. Variables x_{in} represent the amount of money invested in investment type i at node n, while variables y_n^+ and y_n^- represent the surplus and deficit, respectively a nodes $n \in N_{\bar{T}}$. Finally, r_{in} represents the return of investment type i at node n and $a(n)$ represents the predecessor of node n.

Let us initially calculate $MSV_{4,2}$, while $MSV_{4,3}$ will be calculated afterwards. Fig. 2 shows the solution to the original four-stage problem while its optimal objective value is $z^4 = \$1.51$. In order to calculate $MSV_{4,2}$, initially we solve a two-stage approximation of the problem, where the random returns for stages $t = 2, \ldots, 4$ are replaced by their expectations. Fig. 3a shows the two-stage approximation and its optimal first-stage solution. The procedure continues by moving onto the second stage. We take the solution just found as fixed, and for each node in the second stage we solve a two-stage stochastic program. Here the random variables in the last stage ($\bar{T}$) are replaced by their expectation, while the realizations of the random variables for the nodes at the second and third stage are exactly those seen by the original problem. Fig. 3b shows this procedure. Notice that the first-stage solution is shown in a box to indicate that it fixes first-stage decision variables. Once the problems at the second stage have been solved we collect the decisions made at each node in the second stage, as shown in Fig. 3b. We can now calculate E_2 as shown in Fig. 4. The expected return of using a two-stage problem is $E_2 = \$2.71$. Consequently, $MSV_{4,2} = 2.71 - 1.51 = \$1.2$, stating that solving a two-stage approximation would cause a loss of \$1.2.

The WS value is $\$-10.50$, therefore inequalities (2) hold, as $-10.50 \leq 1.51 \leq 2.71$. Consequently a valid bound is $MSV(4, 2) \leq 2.71 - (-10.50) = 13.21$.

In order to calculate $MSV_{4,3}$ the procedure is similar to that used for calculating $MSV_{4,2}$ and is not show here for the sake of brevity. The only difference is that the only decision component which is fixed when calculating E_3 is the first-stage decision obtained by solving a three-stage approximation

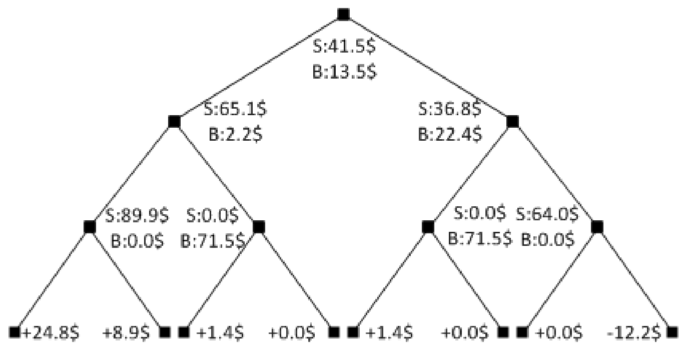
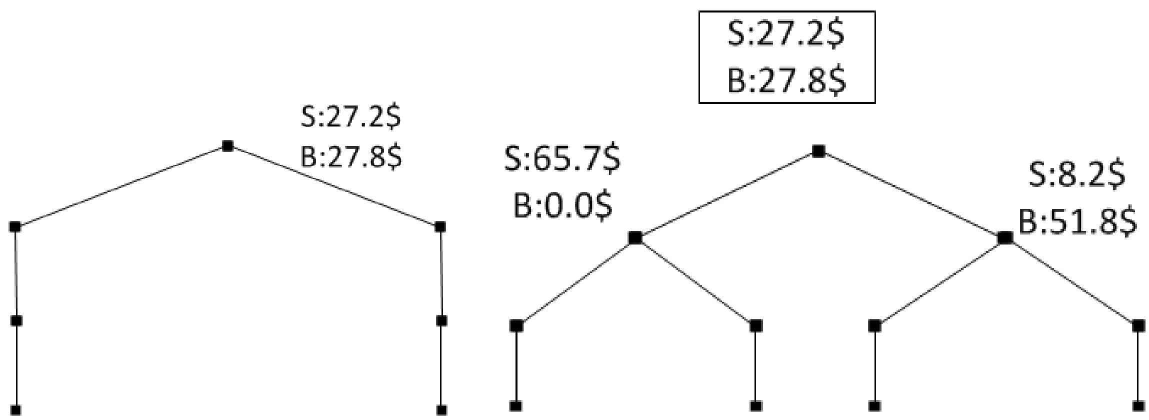


Figure 2: Solution to the original four-stage problem



(a) Two-stage approximation of the original four-stage problem

(b) Two-stage problem approximations solved at the second stage

Figure 3: Constructing solution components based on two-stage problems

problem. This solution suggests investing \$14.0 in stocks and \$40.95 in bonds. We obtain $MSV_{4,3} = E_3 - z^4 = 1.92 - 1.51 = \$0.41$. As intuition suggests, a three-stage approximation is preferable to a two-stage approximation. A valid bound is $MSV_{4,3} \leq 1.92 - (-10.50) = 12.42$, which is tighter than for $MSV_{4,2}$.

5 Conclusions

Solving multistage stochastic programs is often challenging, and approximations obtained by reducing the number of stages are typically used. We showed how concepts in the available literature can be extended to evaluate the benefit from solving the original problem rather than an approximation. This metric, referred to as the marginal stage value, offers an appreciation of how god (or how bad) an approximation is. A numerical procedure for the calculation of the marginal stage value – as well as bounds – have been proposed, and the concepts introduced have been explanatorily applied to the investor problem.

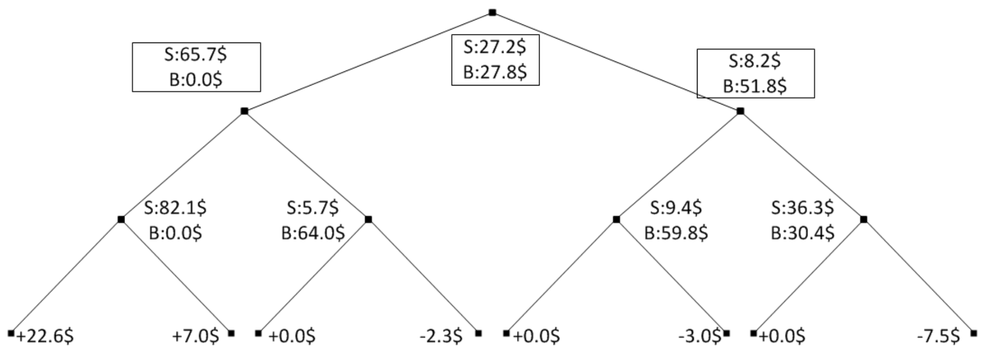


Figure 4: Calculation of the expected return from using a two-stage approximation

References

S. Ahmed, A. King, and G. Parija. A multi-stage stochastic integer programming approach for capacity expansion under uncertainty. *Journal of Global Optimization*, 26:3–24, 2003.

A. Alonso, L. F. Escudero, and M. T. Ortuño. A stochastic 0-1 program based approach for the air traffic flow management problem. *European Journal of Operational Research*, 120:47 – 62, 2000.

R. Bakkehaug, E. S. Eidem, K. Fagerholt, and L. M. Hvattum. A stochastic programming formulation for strategic fleet renewal in shipping. *Transportation Research Part E: Logistics and Transportation Review*, 72:60 – 76, 2014.

J. R. Birge. The value of the stochastic solution in stochastic linear programs with fixed recourse. *Mathematical Programming*, 24:314–325, 1982.

J. R. Birge. Decomposition and partitioning methods for multistage stochastic linear programs. *Operations Research*, 33:989–1007, 1985.

J. R. Birge and F. Louveaux. *Introduction to stochastic programming*. Springer, New York, 1997.

D. R. Cariño, T. Kent, D. H. Myers, C. Stacy, M. Sylvanus, A. L. Turner, K. Watanabe, and W. T. Ziemba. The Russell-Yasuda Kasai Model: An Asset/Liability Model for a Japanese Insurance Company Using Multistage Stochastic Programming. *Interfaces*, 24:29–49, 1994.

L. Escudero, A. Garín, M. Merino, and G. Pérez. The value of the stochastic solution in multistage problems. *TOP*, 15:48–64, 2007.

L. F. Escudero, A. Garín, M. Merino, and G. Pérez. BFC-MSMIP: an exact branch-and-fix coordination approach for solving multistage stochastic mixed 0–1 problems. *Top*, 17:96–122, 2009.

S.-E. Fleten and T. K. Kristoffersen. Short-term hydropower production planning by stochastic programming. *Computers & Operations Research*, 35:2656 – 2671, 2008.

T. K. Kristoffersen and S.-E. Fleten. Stochastic programming models for short-term power generation scheduling and bidding. In E. Bjørndal, M. Bjørndal, P. M. Pardalos, and M. Rönnqvist, editors, *Energy, Natural Resources and Environmental Economics*, Energy Systems, pages 187–200. Springer Berlin Heidelberg, 2010.

Y. Li, G. Huang, and S. Nie. An interval-parameter multi-stage stochastic programming model for water resources management under uncertainty. *Advances in Water Resources*, 29:776 – 789, 2006.

A. Løkketangen and D. L. Woodruff. Progressive hedging and tabu search applied to mixed integer (0, 1) multistage stochastic programming. *Journal of Heuristics*, 2:111–128, 1996.

G. Lulli and S. Sen. A branch-and-price algorithm for multistage stochastic integer programming with application to stochastic batch-sizing problems. *Management Science*, 50:786–796, 2004.

F. Maggioni, E. Allevi, and M. Bertocchi. Bounds in multistage linear stochastic programming. *Journal of Optimization Theory and Applications*, pages 1–30, 2013.

J. M. Mulvey and B. Shetty. Financial planning via multi-stage stochastic optimization. *Computers & Operations Research*, 31:1 – 20, 2004.

G. Pantuso, K. Fagerholt, and S. W. Wallace. Solving hierarchical stochastic programs: application to the maritime fleet renewal problem. *INFORMS Journal of Computing*, 27:89–102, 2015.

S. W. Wallace and S.-E. Fleten. Stochastic programming models in energy. In A. Ruszczynski and A. Shapiro, editors, *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*, pages 637 – 677. Elsevier, 2003.