Improved Inflow Modeling in Stochastic Dual Dynamic Programming

Hamed Poorsepahy-Samian; Vahid Espanmanesh; and Banafsheh Zabraie

Abstract: Stochastic dual dynamic programming (SDDP) is a widely used technique for operation optimization of large-scale hydropower systems in which reservoir inflow uncertainty is modeled with discrete scenarios produced by statistical time series models, such as the family of periodic auto-regressive (PAR) models. It is a common practice in statistical modeling of hydrologic time series to fit a well-known probability distribution (usually normal distribution) to the data by applying proper transformation. Box-Cox transformation is a commonly used transformation in the case of normal distribution fitting. The convexity requirement of SDDP means that nonlinearly transformed time series cannot be used for statistical inflow model calibration. In this paper, a linear approximation is proposed to estimate the expected value of the next stage inflow. In the proposed approach, next-stage inflows are estimated by a model that uses transformed time series. Furthermore, using the proposed linear approximation, it is shown that it is possible to utilize the time series transformed by Box-Cox transformation for scenario generation in SDDP. The Karoon multireservoir system in Iran has been used as a case study in order to show the effectiveness of the proposed method. Some concluding remarks have also been provided by comparing the results of the two SDDP models, with and without the proposed linear approximation. DOI: 10.1061/(ASCE)WR.1943-5452.0000713. © 2016 American Society of Civil Engineers.

Author keywords: Multireservoir systems; Stochastic dual dynamic programming (SDDP); Optimization; Box-Cox transformation.

Introduction

Stochastic dual dynamic programming (SDDP) is an extension of stochastic dynamic programming (SDP). The main difficulty with SDP is the exponential increase in the computational time and memory needed to solve the optimization problem by increasing the size of the problem, which is the well-known “curse of dimensionality,” first introduced by Bellman (1961). The curse of dimensionality restricts the applicability of SDP for the cases of large-scale reservoir systems with 3–4 reservoirs (Pereira and Pinto 1991). Different approaches have been proposed to circumvent this problem. One approach is to aggregate the reservoirs, which leads to difficulties in disaggregating the results (Saad et al. 1994). Another approach is to construct either an interpolation (e.g., Johnson et al. 1993) or an extrapolation of the expected benefit-to-go function (e.g., Castelletti et al. 2007).

SDDP is one of the most successful approaches to overcome the computational problems of SDP, by constructing an extrapolation of the expected benefit-to-go function. Pereira and Pinto (1985, 1991) proposed SDDP for the problem of hydrothermal scheduling. SDDP uses the dual information of one-stage optimization to approximate the expected benefit-to-go function in each stage of the problem. In this sense, SDDP is an extension of Benders decomposition, and each approximation of the benefit-to-go function is a Benders cut. The accuracy of this approximation is increased through an iterative simulation-optimization process.

SDDP has been applied to a number of hydrothermal scheduling problems (e.g., Homem-de-Mello et al. 2011; Rebennack et al. 2012). Tilmant and Kelman (2007) utilized SDDP in order to analyze trade-offs between agricultural development scenarios and hydropower benefits in a big water resources system. Tilmant et al. (2008) proposed a method for incorporating irrigation benefits in the objective function of SDDP in order to assess the marginal value of water. Goor et al. (2011) proposed a method based on the convex hull approximation of the true hydropower function in order to take into account the variable productivity of hydropower plants. Recently, Maceira et al. (2015) have applied a direct conditional value-at-risk (CVaR) approach in a SDDP-based model for long-term hydrothermal coordination.

A difficulty with SDDP has been the incorporation of interstage dependency of inflows. In SDDP, inflow uncertainty is modeled with discrete scenarios synthetically generated by statistical models, such as auto-regressive (AR) models. Infanger and Morton (1996) provided a methodology for sharing cuts in multistage programming when the random variable satisfies certain classes of interstage dependency. Maceira and Damazio (2004) applied the methodology of Infanger and Morton (1996), to take into account the interstage dependencies of inflows in the operation planning of Brazilian hydropower system. A similar formulation was used by Tilmant and Kelman (2007), Tilmant et al. (2008), and Goor et al. (2011).

However, in some cases, the linear interstage dependent models are not well suited for inflow modeling. In most of the parameter estimation methods, it is assumed that the time series probability distribution is normal, but in many cases, the hydrologic time series do not follow normal distribution and are not symmetrically distributed. Therefore, it is often necessary to transform those time series to better fit normal distribution before performing statistical modeling (Karamouz et al. 2003). Box-Cox transformation (Box and Cox 1964) is a commonly used transformation in the case of normal distribution fitting.

© ASCE
Besides, in some cases, linear interstage dependent models generate negative inflow values. Eliminating the negative inflow values by, for instance, setting a lower bound on the inflow could be an option, but it results in convergence issues in SDDP. Because setting a lower bound for the inflow changes the properties of the probability distribution, it causes problems in SDDP convergence. Gjelsvik et al. (2010) discussed the application of logarithmic transformation in order to circumvent this problem and concluded that such a transformation results in a nonconvex optimization model, which is not solvable by SDDP. A similar argument has been made by de Matos and Finardi (2012).

In this paper, a linear approximation is proposed for the expected values of the inflow from a Box-Cox transformed time series model. The results have shown that the proposed linear approximation is very accurate for a wide range of Box-Cox transformation parameters. Then, based on the proposed linear approximation, a methodology is developed to include Box-Cox transformation in the inflow modeling in SDDP. The resulting policies of the SDDP model with Box-Cox transformation are then compared to a SDDP model without transformed inflows. Finally, some concluding remarks are provided.

The rest of the paper is organized as follows. The optimization problem formulation and the SDDP algorithm are introduced in “SDDP Formulation.” In “Hydrological Model,” the Box-Cox transformation and the hydrological model are introduced. “Proposed Method for Cut Estimation” explains the methodology for developing a closed form equation of cut parameters when Box-Cox transformation is used in the inflow model. Karoon multireservoir system is then introduced in “Results,” which is followed by the model results for this case study. Finally, discussion and concluding remarks are presented in the last section.

SDDP Formulation

In SDDP, the value of the expected benefit-to-go function in each stage is approximated utilizing some, say \( L \), linear constraints. Taking the reservoir storages as the state variables, in a system with \( J \) reservoirs, the \( j \)th approximation (Benders cut) has the form of

\[
F^*_t \leq \alpha^t s_t + \delta^t_t \tag{1}
\]

where \( F^*_t \) (\$/) is approximated value of the expected benefit-to-go function in stage \( t \); \( \alpha^t \) (\$/m³) = \( 1 \times J \) matrix of the slopes of the \( j \)th Benders cuts with respect to storages; \( s_t \) (m³) = \( J \times 1 \) matrix of reservoir storages; and \( \delta^t_t \) (\$/) is offset value corresponding to the \( j \)th Benders cut.

The value of the objective function also depends on the random variables. Therefore, if a cut of the form (1) is generated for a certain scenario of the random variable, it cannot be used for other scenarios of the random variable. Cut sharing is the ability to share cuts generated for a certain scenario of the random variable in the stage \( t \) with any other scenario in that stage. Infanger and Morton (1996) have developed formulations for cut sharing when the random variable satisfies certain interstage dependency. In brief, they have shown that the current state of the random variable should be taken into account in the Benders cut. It should be noted that the main random variable in reservoir operation problem is the inflow to the reservoir. If an autoregressive model of order \( p \) is used for modeling the inflows, the inflow of the each stage will depend on the inflows of the previous \( p \) stages. Therefore, the state of the inflow at the start of each stage will consist of the inflows of the previous \( p \) stages, \( q_{t-1}, \ldots, q_{t-p} \), and the Benders cut will be of the following form:

\[
F^*_t \leq \alpha^t s_t + \sum_{j=1}^{J} \sum_{i=1}^{p} \beta_{t,j}^i q_{t-i,j} + \delta^t_t \tag{2}
\]

where \( \beta_{t,j}^i \) (\$/m³) = coefficient associated with the inflow of the reservoir \( j \) in the \( i \)th previous stage, in the \( i \)th Benders cut at stage \( t \).

Considering a reservoir that supplies water to agricultural lands downstream, which is also the case of this research, benefits from supplying irrigation demands can also be included in the model formulation. This adds another variable into the Benders cuts based on the method proposed by Tilmant et al. (2008)

\[
F^*_t \leq \alpha^t s_t + \sum_{j=1}^{J} \sum_{i=1}^{p} \beta_{t,j}^i q_{t-i,j} + y_t^j + \delta^t_t \tag{3}
\]

where \( y_t = (D \times 1) \) matrix of agricultural state variable at the start of stage \( t \) (m³); \( y_t^j = (S/m³) = (1 \times D) \) matrix of the coefficients associated to the agricultural state variable matrix in the \( i \)th Benders cut in the stage \( t \); and \( D \) = total number of agricultural demand sites.

The variable \( y_t \) represents the cumulative water allocated to each agricultural demand site from the beginning of irrigation season until the current stage \( t \).

The one-stage optimization in SDDP has the following general form:

\[
\begin{align*}
\max & \quad F_t = f_t + F^*_{t+1} \\
\text{Subject to} & \quad s_{t+1} - C^t r_t - r^\text{null} - C^t w_t = s_t + q_t, \\
& \quad r_t \leq r_t \leq \bar{r}_t, \\
& \quad s_{t+1} \leq s_{t+1} \leq \bar{s}_{t+1} \quad t \neq T: \quad s_{t+1} + x_{s,t} = s_{\text{target}} \quad t = T, \\
& \quad w_t \leq w_t \leq \bar{w}_t, \\
& \quad y_{t+1} - \alpha w_t = y_t, \\
& \quad \sum_{j=1}^{J} G_{s,j} + x_{G,s} \geq G_{s}
\end{align*}
\]

where \( f_t = \bar{p}_t \sum_{j=1}^{J} G_{s,j} + \sum_{d=1}^{D} A R_{s,d} - (p f_G x_{G,s} + pf_e \sum_j x_{s,j}) \) and \( t \neq T \)

\begin{align*}
& t = T \\
& f_t = \bar{p}_t \sum_{j=1}^{J} G_{s,j} + \sum_{d=1}^{D} A R_{s,d} - (p f_G x_{G,s} + pf_e \sum_j x_{s,j} + pf_e \sum_j x_{s,j})
\end{align*}
where \( F_t \) is expected benefit-to-go from the beginning of the stage \( t (t = 1, \ldots , T) \); \( R_t \) is immediate revenue at the stage \( t \); \( \text{AR}_{t+1} \) is revenue from the agricultural demand point \( d_t \) in the stage \( t \); \( w_t \) is upper and lower bounds of agricultural allocations, respectively \( (m^3) \); \( a \) is \( (D \times 1) \) matrix of irrigation efficiency in different demand sites; \( r_t \) is \( (J \times 1) \) matrix of the reservoir releases in the stage \( t \); \( q_t \) is \( (J \times 1) \) matrix of the reservoir inflows in stage \( t \); \( s_t \) is \( (1 \times J) \) matrix of the spills in the stage \( t \); \( \text{S}_{t+1} \) and \( \text{S}_t \) are upper and lower bound matrices of reservoir storages at the end of stage \( t \), respectively \( (m^3) \); \( C_t \) is minimum required energy generation of the system in the stage \( t \) (MWh); \( P_t \) is average power price in time stage \( t \) (S/MWh); \( s_{\text{target}} \) is end of horizon target storage (which is generally determined using the long-term models or is assumed to be the same as the initial storage) \( (m^3) \); \( \text{df}_t \) is amount of hydropower demand deficit in the stage \( t \) (MWh); \( \text{df}_t \) is \( (1 \times J) \) matrix of the violation from the end of horizon storage target \( (m^3) \); \( x_{t,j} \) is \( (J \times 1) \) matrix of the instream flow demand deficit downstream of the reservoirs in the stage \( t \); \( \text{pf}_a \) is penalty factor for deficit in supplying hydropower demand; \( \text{pf}_e \) is \( (1 \times J) \) matrix of the penalty factors for violation from target storage at the end of planning horizon; \( \text{oe}_t \) is \( (1 \times J) \) matrix of penalty factors for deficit in supplying instream flow demands.

\( \text{C}^0 \) is a \((J \times J)\) matrix, which shows how the reservoirs are connected in a multireservoir system. If reservoir \( j \) receives the released water from reservoir \( j' \), then \( C_{j,j'} = 1 \) and \( C_{j,j'} = -1 \), \forall j, j' = 1, \ldots, J. \text{C}^0 \) is a \((J \times J)\) matrix that shows the connectivity between agricultural demand sites and reservoirs. If reservoir \( j \) receives return flows from agricultural demand site \( d \), then \( \text{C}(j,d) = \omega_{jd} \), where \( \omega_{jd} \) is the percent of water allocation to the demand site \( d \) that drains back to the system.

Solving the optimization model in SDDP includes two main phases. The first phase is backward recursion, which includes a backward pass through all stages and estimating new cuts for each stage based on some, say \( L \), trial values for the state variables at the start of that stage. For each trial value, \( K \) different scenarios (also called backward openings) are generated for the uncertain parameter in that stage, and cuts are estimated by averaging over those scenarios. The second phase is forward simulation, which includes sampling some, say \( M \), different scenarios for the uncertain parameter and simulating the system based on those scenarios. This two phases are iteratively done until reaching the stopping criterion. Further details about the SDDP algorithm are available in Pereira and Pinto (1991) and Tilmant and Kelman (2007).

The classic stopping criterion, which was proposed by Pereira and Pinto (1991), is based on statistically comparing the observed expected benefit at the first stage (which is a lower bound of the optimal expected benefit) with the approximation of the expected benefit at the first stage, which is obtained by averaging the expected benefits defined by Benders cuts over \( M \) scenarios (which is an upper bound of the optimal expected benefit). Shapiro (2011) argued that this criterion is too optimistic and proposed that the SDDP algorithm should be stopped when the value of the upper bound has started to stabilize. He also proposed to statistically test whether the further runs of SDDP algorithm can significantly improve the accuracy of the solution. Another statistical test was discussed by Homem-de-Mello et al. (2011). In this paper, the classic stopping criteria is controlled as an indicator of the accuracy of the Benders cuts. Meanwhile, in order to prevent the premature convergence of the algorithm, the algorithm is further continued until reaching a maximum number of iterations. It is common in SDDP applications to let the model continue up to a maximum number of iterations (Philpott and Guan 2013).

### Hydrological Model

The one parameter Box-Cox transformation (Box and Cox 1964) is defined by

\[
\hat{q}_t = \begin{cases} 
\frac{q_{t-1}^{\lambda}}{\lambda} & \lambda \neq 0 \\
\ln(q_t) & \lambda = 0
\end{cases}
\]

where \( \hat{q}_t \) is transformed values of the original time series, \( q_t \); and \( \lambda \) is transformation parameter. It is worth noting that Box-Cox transformation is continuous for \( \lambda = 0 \), and the logarithmic transformation is included in Box-Cox transformation as a special case of \( \lambda = 0 \).

The periodic autoregressive model of order \( p \) for \( \hat{q}_t \), PAR\((p)\), can be written as

\[
\frac{\hat{q}_t - \hat{\mu}_t}{\hat{\sigma}_t} = \sum_{i=1}^{p} \phi_i (\hat{q}_{t-i} - \hat{\mu}_{t-i}) + \varepsilon_t
\]

where \( \phi_i = \text{AR} \) coefficient corresponding to the season \( z \) with lag \( i \). Note that for a weekly model, \( z = 1, \ldots, 52 \), whereas \( t = 1, \ldots, T \). Also, \( \hat{\mu}_t \) and \( \hat{\sigma}_t \) are the mean and the standard deviation of the transformed time series corresponding to the season \( z \). \( \varepsilon_t \) is the Gaussian white noise with mean zero and variance \( \sigma_z^2 \). In order to generate inflow scenarios, first, \( \hat{q}_t \) is estimated from Eq. (14), and then, \( q_t \) is estimated by applying the inverse Box-Cox transformation. The same process is used both for generating an inflow scenario for forward simulation and generating the backward openings for backward optimization (Hipel and McLeod 1994).

Based on Eq. (14), the expected value of the transformed inflow (or the forecast in the transformed domain) to each reservoir \( j \) in the stage \( t \) is obtained using the following equation:

\[
E[\hat{q}_t(j)] = \left[ \sum_{i=1}^{p} \phi_i (\hat{q}_{t-i}(j) - \hat{\mu}_{t-i}(j)) \right] \hat{\sigma}_t + \hat{\mu}_t(j)
\]

According to Hipel and McLeod (1994), the minimum mean square error of the expected value of inflow (or forecasted inflow) in the untransformed domain is as follows:

\[
E[q_t(j)] = \frac{1}{\sqrt{2\pi\sigma_z^2}} \int_{-\infty}^{\infty} (\lambda q + 1)^{1/\lambda} e^{-1/2((q-E[\hat{q}_t(j)])^2/\sigma_z^2)} dq
\]

The integral in Eq. (16) should be determined using numerical integration (Hipel and McLeod 1994).

The white noises corresponding to the reservoirs that are in the same basin are dependent in space, which means that if one reservoir is experiencing a dry condition, other reservoirs in the same basin would experience somewhat the same condition. Following the work of Tilmant and Kelman (2007), this kind of dependency is taken into account in this study by using the Cholesky factorization of the lag-zero cross-correlation matrices between the historical inflows into different reservoirs (Hipel and McLeod 1994).

### Proposed Method for Cut Estimation

An important step in the SDDP algorithm is to estimate the cut parameters \( \alpha, \beta, \chi, \) and \( \delta \). In this paper, the main concern is the estimation of \( \beta \) for a SDDP model with Box-Cox transformed
inflow model. Hence, only the formulas for estimating $\beta$ are considered. Interested readers could find the formulas for estimating other cut parameters in Tilmant and Kelman (2007).

As mentioned previously, the state of inflow at the start of each stage should be taken into account in the Benders cuts. In the case of an autoregressive model of order $p$, which does not include a nonlinear transformation, the inflow of each stage has a linear relation with the inflows of the previous $p$ stages. However, using the Box-Cox Transformation in the inflow modeling results in a nonlinear relationship between the inflow of each stage and the inflows of the previous stages.

In this paper, it is proposed to apply a linear regression in order to estimate a linear approximation of the relationship between $E(q_t(j))$ [Eq. (16)] and the inflows of the previous stages. Because $E(q_t(j))$ is nonlinearly related to both the transformed and untransformed inflows of the $p$ previous stages, both $q_{t-p,j}$, ..., $q_{t-1,j}$ and $\hat{q}_{t-p,j}$, ..., $\hat{q}_{t-1,j}$ are taken into account in the proposed methodology.

In order to estimate such a linear regression, a sufficiently large set of predictors ($q_{t-p,j}$, ..., $q_{t-1,j}$ and $\hat{q}_{t-p,j}$, ..., $\hat{q}_{t-1,j}$) and the corresponding values of dependent variable, $E(q_t(j))$, for each time stage $t$ are needed. To this end, $N$ different synthetic scenarios are generated, each with a length of $T$. Hence, $N$ different combinations of predictors and dependent variables for each stage $t$ can be obtained on which a linear regression of the following form could be fitted:

$$E(q_t(j)) \approx \sum_{i=1}^{p} u^i_{t,j} q_{t-i,j} + \sum_{i=1}^{p} \tilde{u}^i_{t,j} \hat{q}_{t-i,j} + \psi_{t,j} \quad (17)$$

where $u^i_{t,j}$ and $\tilde{u}^i_{t,j}$ are regression coefficients associated to $q_{t-i,j}$ and $\hat{q}_{t-i,j}$, respectively; and $\psi_{t,j}$ is corresponding offset term of the regression. This regression will only be used in estimating the Benders cuts and is not utilized in scenario generation or generation of backward openings.

It should be noted that the process of generating $N$ different inflow scenarios and estimating the regression coefficients does not impose any significant computational burden to the iterative process of SDDP because this process takes place only once before the main iteration of SDDP. Another point worth noting here is that it is also possible to consider only the transformed (or untransformed) values in the linear approximation. However, the results have shown that including both the transformed and untransformed values increases the accuracy of the linear approximation.

Based on the estimated linear approximation, the expected inflow in each stage $t$ depends linearly on $q_{t-p,j}$, ..., $q_{t-1,j}$ and $\hat{q}_{t-p,j}$, ..., $\hat{q}_{t-1,j}$. Hence, the state of the inflow at the start of each stage $t$ includes both $q_{t-p,j}$, ..., $q_{t-1,j}$ and $\hat{q}_{t-p,j}$, ..., $\hat{q}_{t-1,j}$. Consequently, the Benders cut will have the following form:

$$F_{t+1} \leq \alpha_{t+1} s_{t+1} + \sum_{j=1}^{J} \sum_{i=1}^{p} \beta_{t,i,j}^L q_{t-i,j} + \sum_{j=1}^{J} \sum_{i=1}^{p} \beta_{t,i,j}^U \hat{q}_{t-i,j} + \lambda_{t+1}^L \eta_{t+1} + \kappa_{t+1}^U$$

$$F_{t+1} \leq \alpha_{t+1} s_{t+1} + \sum_{j=1}^{J} \sum_{i=1}^{p} \beta_{t,i,j}^L q_{t-i,j} + \sum_{j=1}^{J} \sum_{i=1}^{p} \beta_{t,i,j}^U \hat{q}_{t-i,j} + \lambda_{t+1}^L \eta_{t+1} + \kappa_{t+1}^U$$

This equation replaces Eq. (11) in the one-stage optimization problem.

The cut parameters $\beta_{t,i,j}^L$ and $\beta_{t,i,j}^U$ are the gradient of the expected benefit-to-go function with respect to $q_{t-i,j}$ and $\hat{q}_{t-i,j}$, respectively. These parameters are estimated as follows:

$$\beta_{t,i,j}^L = \frac{\partial F^L}{\partial q_{t-i,j}} = \frac{\partial F^L}{\partial q_{t-j}} \left|_{q_{t-i,j}} \right. + \sum_{i=1}^{p} \left( \frac{\partial F_i}{\partial q_{t-i,j}} \frac{\partial q_{t-i,j}}{\partial q_{t-i,j}} + \frac{\partial F_i}{\partial \hat{q}_{t-i,j}} \frac{\partial \hat{q}_{t-i,j}}{\partial q_{t-i,j}} \right)$$

$$\beta_{t,i,j}^U = \frac{\partial F^U}{\partial \hat{q}_{t-i,j}} = \frac{\partial F^U}{\partial \hat{q}_{t-j}} \left|_{\hat{q}_{t-i,j}} \right. + \sum_{i=1}^{p} \left( \frac{\partial F_i}{\partial \hat{q}_{t-i,j}} \frac{\partial \hat{q}_{t-i,j}}{\partial \hat{q}_{t-i,j}} + \frac{\partial F_i}{\partial q_{t-i,j}} \frac{\partial q_{t-i,j}}{\partial \hat{q}_{t-i,j}} \right)$$

where $\pi_{t,i,j}^L$ = dual variable associated with constraint (5) for the $k$th backward opening; and $\lambda_{t,i,j}^L$ = Lagrange multiplier associated to the $l$th linear approximation in the one-stage optimization in stage $f$ for the $k$th backward opening.

As shown in the Eqs. (19) and (20), $\beta_{t,i,j}^L$ and $\beta_{t,i,j}^U$ have two main terms. The first term is related to $\pi_{t,i,j}^L$, which represents how $q_{t-i,j}$ and $\hat{q}_{t-i,j}$ affect the expected value of the objective function by contributing to the right hand side of the continuity constraints. The second term represents how $q_{t-i,j}$ and $\hat{q}_{t-i,j}$ affect $F_t$ by contributing to the right-hand side of the Benders cuts in Eq. (18). It should be noted that in the first term, which relates to constraint (5), $\partial F_i/\partial q_{t-i,j}$ is used to estimate the cut parameter. If $\partial F_i/\partial q_{t-i,j}$

was used instead of $\partial F_i/\partial q_{t,i}$, the problem would have become nonconvex and that would have led the whole cut estimating process to fail. In other words, $F_i(q_t, \hat{q}_t)$ is estimated rather than $F_i(\hat{q}_t)$.

From Eq. (17), the following derivatives can be estimated:

$$\frac{\partial q_{t,i}}{\partial q_{i-1,j}} = \frac{\partial E(q_{t,i})}{\partial q_{i-1,j}} = u_{i,j}^t$$

$$\frac{\partial q_{i-1,j}}{\partial q_{t,i}} = -\frac{\partial E(q_{i-1,j})}{\partial q_{t,i}} = \hat{u}_{i,j}^t, \quad i = 1, \ldots, p$$

(21)

Eq. (15) shows that $\hat{q}_{t,j}$ has a linear relationship with $q_{t-1,j}$, $\ldots$, $q_{1,j}$. Hence

$$\frac{\partial \hat{q}_{t,i-1,j}}{\partial q_{i-1,j}} = 0$$

(22)

Now, in Eqs. (19) and (20), the derivatives $\partial q_{t,i}/\partial q_{i,j}$, $\partial q_{t,i}/\partial q_{i,j}$, and $\partial q_{t,i}/\partial q_{i,j}$ remain to be estimated, all of which could be separated to the two following different cases:

1. $i = 1$: In this case, $\partial q_{t,j}/\partial q_{i,j}$, $\partial q_{t,j}/\partial q_{i,j}$, and $\partial q_{t,j}/\partial q_{i,j}$ have to be estimated. The first two derivatives are already given by Eq. (21).

2. $i \neq 1$: In this case, if $j = i + 1$, the derivatives $\partial q_{t,i-1,j}/\partial q_{i-1,j}$, $\partial q_{t,i-1,j}/\partial q_{i-1,j}$, and $\partial q_{t,i-1,j}/\partial q_{i-1,j}$ would become $\partial q_{t,i-1,j}/\partial q_{i-1,j}$, $\partial q_{t,i-1,j}/\partial q_{i-1,j}$, and $\partial q_{t,i-1,j}/\partial q_{i-1,j}$. Hence, the following equations can be written for this case:

$$\frac{\partial q_{t,i-1,j}}{\partial q_{i-1,j}} = \begin{cases} 1 & i = i + 1 \\ 0 & \text{else} \end{cases}$$

$$\frac{\partial q_{t,i-1,j}}{\partial q_{i-1,j}} = \begin{cases} 1 & i = i + 1 \\ 0 & \text{else} \end{cases}$$

$$\frac{\partial q_{t,i-1,j}}{\partial q_{i-1,j}} = 0$$

As a summary of the previous

$$\frac{\partial \hat{q}_{t,i-1,j}}{\partial \hat{q}_{i-1,j}} = \begin{cases} \text{Eq. (23)} & i = 1 \\ 1 & i = i + 1 \\ 0 & \text{else} \end{cases}$$

$$\frac{\partial q_{t,i-1,j}}{\partial q_{i-1,j}} = \begin{cases} \text{Eq. (21)} & i = 1 \\ 1 & i = i + 1 \\ 0 & \text{else} \end{cases}$$

$$\frac{\partial q_{t,i-1,j}}{\partial q_{i-1,j}} = \begin{cases} \text{Eq. (21)} & i = 1 \\ 0 & \text{else} \end{cases}$$

$$\frac{\partial \hat{q}_{t,i-1,j}}{\partial \hat{q}_{i-1,j}} = 0$$

By replacing Eqs. (21) and (27)–(30) into Eqs. (19) and (20), a closed-form equation for $\beta_{t,j}$ and $\beta_{t,j}$, $i = 1, \ldots, p$ results.

As an example, the case of $p = 1$ can be considered. In this case, $\beta_{t,j}$ and $\beta_{t,j}$ are given as follows:

$$\beta_{t,j} = \frac{1}{K} \sum_{k=1}^{K} \left[ n_{t,k,j} u_{i,j}^t + \sum_{l=1}^{L} n_{c,t,l} \left( \beta_{t+1,j} \sigma_{\hat{q}_{t+1,j}} \right) \right]$$

(31)

$$\beta_{t,j} = \frac{1}{K} \sum_{k=1}^{K} \left[ n_{t,k,j} \hat{u}_{i,j}^t + \sum_{l=1}^{L} n_{c,t,l} \left( \beta_{t+1,j} \sigma_{\hat{q}_{t+1,j}} \right) \right]$$

(32)

Results

The proposed methodology has been applied to the Karoon Reservoir system in Iran, which provides 14% of the total national power generation (Tavanir Holding Company 2012). This system is the largest and most important hydropower system in Iran and consists of five hydropower reservoirs in series. The location map and the schematic of the system are shown in Fig. 1. The annual irrigation demands downstream of the Karoon 1, Godarlandar and Gotvand dams, are 689, 1,216; and 2,165 ($10^6$ m$^3$), respectively. It is expected that a minimum annual power generation equal to 3,900 GWh is supplied by the Karoon Reservoir system. All the reservoirs are assumed to be 40% filled initially (at the start of Stage 1), and end-of-horizon target storage is also assumed to be the same as the initial storage for each reservoir.

The main inflows to this system are supplied from Karoon River, which enters the system as inflow to the most upstream reservoir, the Karoon 4 Reservoir and Kherans River, which joins the main stream from the subbasin between Karoon 4 and Karoon 3 Dams. Table 1 shows some specifications of the reservoirs and their corresponding hydroelectric power plants in the system. Observed inflow time series between the years 1970 and 2004 have been used for calibrating statistical inflow models, and the observed inflow series between 2005 and 2014 have been utilized for simulation. Based on the maximum likelihood method, Box-Cox transformations with $\lambda = -0.248$ and $\lambda = -0.236$ have been found to be appropriate for transforming the time series of the inflows from Karoon and Khersans Rivers, respectively. The normality plots of the Khersans River’s inflow time series before and after applying the Box-Cox transformation are represented in Fig. 2(a) as an example.

 Afterwards, a PAR(4) model has been fitted to the transformed time series in each of the inflow branches. To show how applying Box-Cox transformation affects the residuals of the models, the normality plot of residuals in Season (Week) 21 in the Khersan River are shown as an example in Fig. 2(b) for the models calibrated for the untransformed and transformed time series. As it can be seen in the figure, the residuals are highly deviated from normal distribution in the case of the model without Box-Cox transformation, whereas the residuals are close to normal distribution in the model calibrated for transformed series. A Kolmogorov-Smirnov test shows that the normality assumption holds only for the residuals of the model with transformed inflow time series. Further investigation has shown that the time independence assumption holds true for model residuals of both inflow models (Karoon and Khersans Rivers) with and without Box-Cox transformation.

It has been found that the accuracy of the linear regression shown in Eq. (17) in approximating the expected value of inflow depends mainly on Box-Cox transformation parameter $\lambda$. Fig. 3 represents the relationship between $\lambda$ and the multiple correlation coefficient (average of correlation coefficients over time stages) of
the linear regression for the K hersan River. As shown in this figure, the proposed linear regression is accurate enough for a wide range of Box-Cox transformation parameters, which also includes logarithmic transformation as a special case.

Two optimization models have been developed, namely Model 1 and Model 2. Model 1 is an SDDP model in which the inflow modeling has been carried out with the transformed time series. Model 2 is also a SDDP model in which the original time series (without applying Box-Cox transformation) have been used for calibrating the inflow model. The only difference between the inflow models in Model 1 and Model 2 is the inclusion of the Box-Cox transformation in the inflow modeling of Model 1.

Midterm operation optimization models generally have weekly or monthly time stages with a planning horizon of one to five years. A two-year planning horizon has been considered with weekly time stages in this study. SDDP parameters $L$, $K$, and $M$ have been set as 15, 30, and 100, respectively. There are no specific guidelines available for the selection of these parameters. However, there are some considerations. $L$ controls the number of the Benders cuts generated in each iteration. Although a small value for $L$ ensures faster iterations, on the other hand, it increases the number of iterations needed for the model to converge. Based on the authors' previous experience with the model, a value of 15 seems a good trade-off between fast iterations and fast convergence. The value of $K$ defines the number of backward openings, and the cuts are estimated by averaging the dual variables corresponding to each of the backward openings. Hence, the value of $K$ should be chosen such that the sample can be considered as a good representation of the true probability distribution, and as such, a value of 30 is chosen for this parameter. $M$ is the number of inflow scenarios generated for the forward simulation. de Matos et al. (2015) have proposed a tree traversing strategy in which the number of scenarios visited in each forward simulation increases with each iteration of the SDDP algorithm. However, the standard algorithm of SDDP is followed, and $M$ is chosen to be constant. Convergence of the algorithm depends on the average objective function generated in these $M$ inflow scenarios. As such, the value of $M$ should be large enough to ensure that the generated inflow scenarios are a good representation of the true probability distribution. The value of $M$ is chosen as 100 in this study.

Table 1. Specifications of the Dams and Their Hydropower Plants in the Karoon River-Reservoir System (Data Courtesy of Iran Water and Power Resources Development Co.)

<table>
<thead>
<tr>
<th>Dam</th>
<th>Normal water level (masl)</th>
<th>Minimum operating level (masl)</th>
<th>Installed capacity (MW)</th>
<th>Average tailwater level (masl)</th>
<th>Downstream instream flow demand ($10^6$ m$^3$ per week)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karoon4</td>
<td>1,025</td>
<td>996</td>
<td>1,000</td>
<td>845</td>
<td>20</td>
</tr>
<tr>
<td>Karoon3</td>
<td>840</td>
<td>800</td>
<td>2,000</td>
<td>645</td>
<td>20</td>
</tr>
<tr>
<td>Karoon1</td>
<td>532</td>
<td>500</td>
<td>2,000</td>
<td>369</td>
<td>20</td>
</tr>
<tr>
<td>Godarlandar</td>
<td>369</td>
<td>364</td>
<td>2,000</td>
<td>220</td>
<td>20</td>
</tr>
<tr>
<td>Gotvand</td>
<td>224</td>
<td>209</td>
<td>2,000</td>
<td>84</td>
<td>100</td>
</tr>
</tbody>
</table>

The convergence plots of the two models are shown in Fig. 4. A computer with Core i3-2.53GHz CPU and 4GB RAM has been used for running the SDDP models. Based on the classic convergence criterion, Model 1 has converged after 4 iterations and 14 min of runtime, whereas Model 2 has converged after 8 iterations and 26 min of runtime. For the sake of comparability of the results of the two models and in order to prevent premature stopping of the models, both models have been run to reach the iteration 20. As mentioned previously, this is common in SDDP applications to let the model continue up to a maximum number of iterations.

In order to investigate whether or not an error is introduced into the Benders cuts, because of the linear approximation of the Box-Cox transformation, further analysis has been carried out using Fig. 5. This figure presents the box plots of the percent of errors between the approximate expected value of future objective function (given by the Benders cuts) and the realized future objective function, in both Model 1 and Model 2. These plots represent how accurate the Benders cuts have been in predicting the realized future benefit functions in each stage for different scenarios in the two models. The box plots have been generated using the values of the errors considering 100 different synthetically generated inflow scenarios for each model. The box associated with any stage represents the accuracy of the Benders cuts in predicting the realized future benefit function for that particular stage in different scenarios.

Fig. 2. Normal probability plots in the Kherasan River of (a) the original and transformed inflow time series; (b) the model residuals of the 21st season in the models with and without transformation

Fig. 3. Multiple correlation coefficient of the linear regression shown in Eq. (31) as a function of the Box-Cox transformation parameter

The convergence plots of the two models are shown in Fig. 4. A computer with Core i3-2.53GHz CPU and 4GB RAM has been used for running the SDDP models. Based on the classic convergence criterion, Model 1 has converged after 4 iterations and 14 min of runtime, whereas Model 2 has converged after 8 iterations and 26 min of runtime. For the sake of comparability of the results of the two models and in order to prevent premature stopping of the models, both models have been run to reach the iteration 20. As mentioned previously, this is common in SDDP applications to let the model continue up to a maximum number of iterations.

In order to investigate whether or not an error is introduced into the Benders cuts, because of the linear approximation of the Box-Cox transformation, further analysis has been carried out using Fig. 5. This figure presents the box plots of the percent of errors between the approximate expected value of future objective function (given by the Benders cuts) and the realized future objective function, in both Model 1 and Model 2. These plots represent how accurate the Benders cuts have been in predicting the realized future benefit functions in each stage for different scenarios in the two models. The box plots have been generated using the values of the errors considering 100 different synthetically generated inflow scenarios for each model. The box associated with any stage represents the accuracy of the Benders cuts in predicting the realized future benefit function for that particular stage in different scenarios.

Note that the Benders cuts are an approximation (an upper bound, to be precise) of the expected future benefit function. On the other hand, the realized future benefit function could be either higher or lower than the expected future benefit function. Hence, one would expect that the approximation from Benders cuts might be either an underestimate (negative error) or an overestimate.
This implies that there have been no significant errors introduced by using the linear approximation, and the resulting Benders cuts have been as accurate as the Benders cuts in the model without the Box-Cox transformation.

As mentioned previously, the inflows from 2005 to 2014 have been utilized in this study for simulation. This period consists of nine different two-year scenarios. Comparison of the simulation results for the two models are shown in Table 2.

As is suggested by the results in Table 2, Model 1 has outperformed Model 2 and has resulted in achieving a 6.5% higher value for the objective function. Model 1’s application has also resulted in 0.5% more power generation, and in turn, Model 2 has resulted in 6.7% more agricultural allocation. However, the main differences between the performances of the two models have been in supplying instream flows and deviation from the end of horizon storage target. Model 1 has resulted in 92% less instream flow demand deficit and 70% less deviation from the end-of-horizon storage target.

In order to investigate the reasons behind the better performance of the SDDP model with Box-Cox transformation (i.e., Model 1) in comparison with Model2, influence of the Box-Cox transformation on the whole time series modeling procedure has been assessed. For this purpose, single-step forecasting has been carried out for the inflow series of the two reservoirs in the simulation period (2005–2014). Table 3 compares the performance of the calibrated inflow models, with and without Box-Cox Transformation. As stated previously, the inflow model with Box-Cox Transformation has been used in Model 1, and the inflow model without Box-Cox Transformation has been used in Model 2. Hence, Table 3 is essentially comparing the performance of the inflow model in Model 1 with that of Model 2. Presented in Table 3 are the correlation coefficient, the root mean square error (RMSE), and the mean absolute percentage error (MAPE), which are all estimated to compare the inflows forecasted using the inflow models and the observed inflows to the system in the simulation period. Correlation coefficient, RMSE, and MAPE are the most commonly used statistical

| Table 2. Optimization Runtime and the Average Two-Year Simulation Results from Nine Different Two-Year Simulations |
|----------------------------------|--------------|--------------|--------------|------------|----------------|----------------|--------------|----------------|----------------|
| Model                           | Runtime for 20 iterations (min) | Objective function (10^6 $) | Power generation (MWh) | Water allocated to agricultural (10^6 m^3) | Deficit in supplying instream flow demands | Deviation from end of horizon storage target | Deficit in supplying power demand (GWh) |
|----------------------------------|--------------|--------------|--------------|------------|----------------|----------------|--------------|----------------|
| Model 1                          | 91           | 558.42       | 26,533       | 1,854.47   | 7.66           | 5              | 35.20        | 1              | 0              |
| Model 2                          | 72           | 524.71       | 26,397       | 1,980.27   | 193.69         | 22             | 117.06       | 4              | 0              |

Fig. 5. Box plot of the error in approximating the realized future benefit function of each stage, based on 100 synthetically generated inflow scenarios, in (a) Model 1; (b) Model 2
Table 3. Comparison of the Performance of the Two PAR(4) Models in Forecasting the Inflows to the System in the Simulation Period

<table>
<thead>
<tr>
<th>Inflow model</th>
<th>Karoon River</th>
<th></th>
<th></th>
<th></th>
<th>Kheran River</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correlation</td>
<td>RMSE (%)</td>
<td>MAPE (%)</td>
<td>CRPS</td>
<td>Correlation</td>
<td>RMSE (%)</td>
<td>MAPE (%)</td>
<td>CRPS</td>
</tr>
<tr>
<td>With Box-Cox transformation</td>
<td>0.83</td>
<td>33.55</td>
<td>20.25</td>
<td>12.13</td>
<td>0.75</td>
<td>40.93</td>
<td>28.87</td>
<td>13.18</td>
</tr>
<tr>
<td>Without Box-Cox transformation</td>
<td>0.79</td>
<td>38.27</td>
<td>25.99</td>
<td>23.31</td>
<td>0.74</td>
<td>42.59</td>
<td>41.71</td>
<td>24.70</td>
</tr>
</tbody>
</table>

performance measures for the inflow forecasting models (Wang et al. 2009). Let \( q_{ot} \) and \( q_{ft} \) be the observed and forecasted inflows, respectively, and let \( \overline{q}_{ot} \) and \( \overline{q}_{ft} \) be the average of the observed and forecasted inflows, respectively. Correlation coefficient, RMSE, and MAPE have been estimated using the following equations:

\[
\text{Correlation} = \frac{\sum_{t=1}^{T} (q_{ot} - \overline{q}_{ot})(q_{ft} - \overline{q}_{ft})}{\sqrt{\sum_{t=1}^{T} (q_{ot} - \overline{q}_{ot})^2 \sum_{t=1}^{T} (q_{ft} - \overline{q}_{ft})^2}}
\]

(33)

\[
\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{T} (q_{ft} - q_{ot})^2}{T}}
\]

(34)

\[
\text{MAPE} = \frac{1}{T} \sum_{t=1}^{T} \frac{|q_{ft} - q_{ot}|}{q_{ot}}
\]

(35)

All three of these performance measures are based on comparing point forecasts with the observations. In order to compare the entire distribution of forecasts with the observations, another performance measure, namely continuous ranked probability score (CRPS), is used. CRPS is defined as follows:

\[
\text{CRPS} = \frac{1}{T} \sum_{t=1}^{T} \int_{-\infty}^{+\infty} [P_{\text{cum}}(x) - I_{q_{ot}}(x)]^2 dx
\]

(36)

where

\[
I_{q_{ot}}(x) = \begin{cases} 0 & x < q_{ot} \\ 1 & x \geq q_{ot} \end{cases}
\]

(37)

and \( P_{\text{cum}}(x) = \text{cumulative probability distribution of the forecasts.} \)

As shown in Table 3, all of the four performance measures indicate that the inflow model with Box-Cox Transformation (Model 1) has been more accurate in predicting the inflows to the system. Note that the more accurate an inflow model is, the better it represents the interstage dependency of the inflows. As a result, the accuracy of the inflow model affects the approximation of the expected future objective function, and that, in turn, affects the quality of the operation policies. Therefore, the better performance of Model 1 compared with Model 2 could be associated to the more accurate representation of interstage dependency of the inflows in the inflow model in Model 1, which is a result of using the transformed inflow time series.

In order to further assess these results, Fig. 6 shows the cumulative empirical probability distribution of the total inflows to the system for both of the inflow models with and without Box-Cox transformation compared to the historical cumulative probability distribution. The empirical probability distributions for the two models are estimated based on 1,000 two-year scenarios generated using each model. Fig. 6 shows that incorporating Box-Cox Transformation in the inflow model has resulted in better compatibility with historical pattern, and that, in turn, has likely resulted in the better accuracy of predicting the observed inflows to the system in 2005 to 2014, which is shown in Table 3. In other words, the inflow model without Box-Cox transformation (Model 2) predicts more inflows to the system compared to the inflows that are actually observed and consequently have resulted in poor operation policies.

Discussion and Concluding Remarks

A method was proposed in order to use the Box-Cox transformation in the inflow modeling of the SDDP algorithm. The results have shown that the proposed method for incorporating the Box-Cox transformation is effective and accurate. Based on the results obtained in this study, in order to model interstage dependency of the random variable(s) (inflow in this case), it is not necessary to have an analytic linear estimation of the derivatives for the interstage dependency model.

The proposed linear approximation for the expected value of next stage inflow is proven to be accurate. The proposed method seems to have the potential for incorporating other scenario generation techniques into the inflow model of SDDP. The key factor for such incorporation is the right selection of the inflow state variables. In the case considered in this paper, the transformed and the actual values of the previous stage inflow were considered as the inflow state variables.

The difference in the performance of the SDDP models with different inflow models could be associated to the difference in the accuracy of the incorporated inflow models in representing the interstage dependency of the inflows. It is shown that incorporating Box-Cox transformation improves the accuracy of the inflow model and hence improves the accuracy of the Benders cuts and consequently results in a better operation policy. Further research
can be dedicated to incorporating other nonlinear inflow models into SDDP.

Acknowledgments

The authors gratefully acknowledge the financial support provided by the Iran Water and Power Resources Development Company (grant #923566). The authors thank the reviewers, the Associate Editor, and the Editor for their constructive comments.

References


