An aggregation-decomposition bayesian stochastic optimization model for cascade hydropower reservoirs using medium-range precipitation forecasts

Yong Peng1, Wei Xu2 and Xiaoli Zhang3
1School of Hydraulic Engineering, Dalian University of Technology, Dalian, China, pengyong@dlut.edu.cn
2College of River and Ocean Engineering, Chongqing Jiaotong University, Chongqing, China xuwei19850711@163.com
3School of Water Conservancy, North China University of Water Resources and Electric, Zhengzhou, China, zxl436@163.com

Abstract. The forecast information is essential to improve the utilization efficiency of hydropower resources. To address the uncertainties of forecasting inflow, the Aggregation-Decomposition Bayesian Stochastic Dynamic Programming (AD-BSDP) model is presented in the present paper by using the 10-days precipitation value of the Quantitative Precipitation Forecasts from Global Forecast System (GFS-QPFs). The application in China’s Hun River cascade hydropower reservoirs shows that the GFS-QPFs are beneficial for hydropower generation and the performance of AD-BSDP is more efficiency and reliability than the others models.

1. Introduction
With the development of the social and economic, now the demand for electricity is increasing. The hydroelectric power, clean energy, has become more and more popular. In the hydropower generation operation, the Traditional Rule Curve uses water level as decision variable, the amount of information used in the decision diagram is limited. While considering rainfall and runoff forecasting information, in reservoir operation, is therefore an effective measure to improve the utilization of hydropower resources.

However, Quantitative Precipitation Forecasts (QPF) and runoff forecasts are not the accurate information. The accuracy of rainfall-runoff forecast model is affected by numbers of factors. Thus in the hydropower optimal operation, the uncertainty of forecasting inflow needs to be represented in the optimization model. Karamouz [1, 2] suggested the use of Bayesian Decision Theory incorporating new information in the interpretation of the flow probabilities. Karamouz and Vasilias [3] proposed a Bayesian Stochastic Dynamic Programming (BSDP) model, which embeds Bayesian theory in the SDP formulation. Kim and Palmer [4] constructed a BSDP model to investigate the value of seasonal flow forecasts in hydropower generation. Mujumdar and Nirmala [5] constructed a BSDP in cascade hydropower reservoirs by aggregated individual inflows to reduce the complexity of the uncertainties addressing.

In SDP model, especially in multi-reservoir cases, the complexity of the model and high computational costs due to curse of dimensionality are the important limiting factors for the application. The approach to overcome this problem is aggregate the storage and inflow of each reservoir to yield an equivalent reservoir, according to the hydrological or geographical features. While the way of
aggregation inflow and storage is an important factor to affect the efficiency of the equivalent reservoir. Tejada-Guibert et al [6] and Mujumdar and Nirmala [5] taken aggregate inflow as a state variable to reduce the complexity of the model and solve the computationally tractable, considering the storage of individual reservoirs separately. This approach spatially accounts the storage of each reservoir in the system, and the operation policies are derived for each individual reservoir. Comparing the aggregate inflow, the aggregate method of storage is more difficult for the bias of operation policies in each reservoir. Therefore the method, convert the storage and inflow of each reservoir into potential electrical energy, is presented to help multi-reservoir to derive operation policies [7-8]. While this approach lack of consideration the individual characteristics of the reservoirs and the amount of water to release from the aggregate reservoir. In order to solve this problem, another method is presented by using historical date or simulation date to build relationship among the individual reservoirs. The main purpose of this paper is to investigate the medium-range QPFs application in cascade hydropower reservoirs operation. There are two-fold to attain the goal: (1) the aggregation-decomposition method is implemented to overcome the complexity of the SDP model and high computational costs by aggregation cascade reservoirs as an equivalent reservoir. (2) the Bayesian Decision Theory is used to address the uncertainties of flow and forecast inflow. Based on the equivalent reservoir, the Aggregation-Decomposition Bayesian Stochastic Dynamic Programming (AD-BSDP) model is constructed. In order to verify the efficiency and reliability of the equivalent reservoir and Bayesian Decision Theory, the Aggregating Flow Stochastic Dynamic Programming (AF-SDP) and Aggregation-Decomposition Stochastic Dynamic Programming (AD-SDP) are constructed. This study takes China’s Hun River cascaded hydropower reservoirs as a case example. The real time QPFs (10 days lead time) of the Global Forecast System (QPFs-GFS) are utilized to forecast the 10 days inflow. Based on the performance of Traditional Rule Curve, the performances of the AF-SDP, AD-SDP and AD-BSDP by using medium-range QPFs information are evaluated and compared.

2. Case Study

2.1. Hun River Cascade Hydropower Reservoirs System

In this study, Hun River cascade hydropower reservoirs are taken as an application example. It is located in the lower reaches of the Hun river basin in the northeast of China. Fig. 1 shows the location of the cascade hydropower reservoirs and gauging stations. The basin covers an area of approximately 15000 km². The main features of this reservoirs system are shown in Fig.1 and Table 1.

![Figure 1. Location and main features of Hun River cascade reservoirs system](image)
2.2. Rainfall-Runoff Models

According to the hydrology and meteorology of the Hun River basin, the inflow is roughly divided into dry season and wet season. During dry season (from October to next April), the inflow, low and stable, is mainly supplied by underground runoff. During wet season (from May to September), the inflow, high and instable, is mainly affected by the rainfall. Therefore, the multiple linear regression model and Xinanjiang model [7] are applied to establish the inflow forecasting model during dry season and wet season respectively.

Table 1. Basic parameters of Hun River cascade hydropower reservoirs

| Characteristic                | Huanren | Huilong | Taipingshao |
|------------------------------|---------|---------|-------------|
| Basin area (km²)             | 10400   | 2100    | 528         |
| Total storage (Mm³)          | 3460    | 123     | 182         |
| Usable storage (Mm³)         | 2199    | 90      | 164         |
| Dead storage (Mm³)           | 1380    | 72      | 145         |
| Normal pool level (m)        | 300     | 221     | 191.5       |
| Dead pool level (m)          | 290     | 219     | 190         |
| Installed capacity (MW)      | 222     | 72      | 161         |
| Firm output (MW)             | 33      | 18      | 25          |
| Turbine capacity (m³/s)      | 416     | 330     | 534         |

2.3. Datas

The Global Forecast System (GFS) developed by the U.S. National Centers for Environmental Prediction, which is a global Numerical Weather Prediction (NWP) computer model. The data of GFS are available on NOAA FTP SERVERS for free, and such information forms the basis for non-state weather companies, e.g., Wunderground.com, Weatheronline.co.uk, and t7online.com. The 10 days QPFs-GFS data from 2001 to 2010 are obtained from t7online.com.

The observed precipitation and inflow originates from the Hun River cascade hydropower development authority. The observed precipitation data from 1968 to 2010 are obtained in the upstream of Huanren. The observed inflow data of each reservoir are also obtained from 1968 to 2010. The inflow of intermediate catchment from Huanren to Huilong (IC-HH) and Huilong to Taipingshao (IC-HT) are obtained by back stepping.

3. Reservoir Operation Optimization Model Object Function

3.1. Object Function

The performance measure $B(k)$ of cascade hydropower reservoirs should be a penalty function for a given combination of available state variables. If the calculated power generation $b(k)$ (MW) is less than the system firm output of $e$ (i.e., 76 MW), the value, $B(k)$ (MWH), will be penalized. The objective function of the Stochastic Dynamic Programming can be written as Eq. 2, which is the maximum expected generation of hydropower system return from the start time step to the end of the planning horizon [8].

$$B(k, q, l) = [b(k, q, l) - \alpha \cdot \max\{e - b(k, q, l), 0\}]^\beta \cdot \Delta t$$

$$\max \sum_{t=1}^{T} E[B(k, q, l)]$$

Where, $k$, $q$, and $l$, are the storage combination of individual reservoirs at the beginning and the end of the time step $t$, respectively; $q$, is the inflow combination; $\alpha$ and $\beta$ are the penalty factors; $\Delta t$ is the time in hours for decision interval $t$. 

3

ISAI2017
IOP Conf. Series: Journal of Physics: Conf. Series 887 (2017) 012005
doi:10.1088/1742-6596/887/1/012005
3.2. AF-SDP Model

In multi-reservoir system, application of SDP models bears high complexity and computational costs. Mujumdar and Nirmala\(^5\) takes the aggregate inflow and storage of individual reservoirs as state variables to reduce the complexity of the application. It takes the aggregate inflow as state variable instead of each corresponding to the inflow to an individual reservoir. Based on this method, the Aggregate Flow Stochastic Dynamic Programming (AF-SDP) model is constructed in this study.

In the AF-SDP model, the inflows are aggregated by Eq.3 and the aggregate inflow is discretized into a number of class intervals. Each class interval is represented by a representative value and the vector of the aggregate inflow class interval \(\hat{Q}_t\) is shown by Eq.4. The storages of individual reservoirs are discretized into a number of class intervals by Eq.5. The matrix, combination of the storage class intervals among the reservoirs, is shown in Eq.6.

\[
Q_t = \sum_{i=1}^{q_i} q_i^t \\
\hat{Q}_t = [Q_1^t, Q_2^t, \ldots, Q_n^t]
\]

\[
\begin{align*}
B_{z}^{t} &= A_{z}^{t} = \{v_1^z, v_2^z, \ldots, v_{\mu}^z\} \\
&\vdots \\
B_{z}^{t} &= A_{z}^{t} = \{v_1^z, v_2^z, \ldots, v_{\mu}^z\}
\end{align*}
\]

\[
\begin{bmatrix}
I_{t-1}^z \\
I_{t}^z \\
\vdots \\
I_{t}^z
\end{bmatrix}
= 
\begin{bmatrix}
 I_{t-1}^z \\
 I_{t}^z \\
 \vdots \\
 I_{t}^z
\end{bmatrix} = \begin{bmatrix}
 v_1^z & v_2^z & \cdots & v_{\mu}^z \\
 v_1^z & v_2^z & \cdots & v_{\mu}^z \\
 \vdots & \vdots & \vdots & \vdots \\
 v_1^z & v_2^z & \cdots & v_{\mu}^z
\end{bmatrix}
\]

Where superscripts 1, 2 and \(z\) refer to reservoir 1, reservoir 2 and reservoir \(z\). \(q_i^t\) is the inflow of reservoir \(z\) at time step \(t\). \(Q_i^t\) is the aggregate inflow of reservoirs system at time step \(t\). \(\hat{Q}_t\) is the vector of representative value of aggregate inflow discretization class intervals. \(\mu\) is the number of class intervals. \(B_{z}^{t}\) and \(A_{z}^{t}\) are the vector of representative value of the storage class intervals for reservoir \(z\) at the end of time step \(t-1\) and the beginning of time step \(t\) separately. \(\beta^z\) is the number of storage class intervals for reservoir \(z\). \(I_{t}^z\) and \(I_{t}^z\) represent the \(n\)th scenario of storage combination of the individual reservoir at the end time step \(t-1\) and the beginning time step \(t\) separately.

AF-SDP model recursive equation is defined in Eq.8.

\[
\begin{align*}
m &= \mu \\
n &= \prod_{z=1}^{Z} \beta^z
\end{align*}
\]

\[
f_{\text{opt}}\left(k_i^z, F_i^z\right) = \max_{l_t} \left\{ B_{z}^{t} \left(k_i^z, F_i^z, I_l^z\right) \right\} \\
+ \sum_{j=1}^{\mu} P_j^z \left[Q_i^t / F_i^t\right] \cdot f_{\text{opt}}\left(I_l^z, Q_i^t\right)
\]

Where, the \(m\) and \(n\) represent the computational costs of this model. The \(m\) is determined by the number of aggregate inflow discretization class intervals, and the \(n\) is determined by number of reservoir in system and the number of individual reservoir storage discretization class intervals. The
$P_{t+1}^f(Q_{t+1}^i/F_t^i)$ suggests that the probability of the inflow ($Q_{t+1}^i$) at time step $t+1$ belongs to class interval $f$, when the forecasting inflow ($F_t^i$) at time step $t$ belongs to class interval $j$.

3.3. AD-SDP model

Based on the AF-SDP model, the Aggregation-Disaggregation Stochastic Dynamic Programming (denoted by AD-SDP) aggregates the storage and inflow of the cascade reservoirs to yield an equivalent reservoir. The inflow of the equivalent reservoir is presented by aggregate inflow, which is mentioned above. The minimum and maximum reservoir storages of the equivalent reservoir are represented by Eq.9. The vectors of equivalent reservoir storage class interval are shown in Eq.10.

$$V_{\min} = \sum_{\alpha=1}^{\nu_{\min}} v_{\min}^\alpha$$
$$V_{\max} = \sum_{\alpha=1}^{\nu_{\max}} v_{\max}^\alpha$$

$$[L_{t-1}^a]
\begin{bmatrix}
L_{t-2}^a \\
\vdots \\
L_{t-1}^a
\end{bmatrix}
= [K_1^a]
\begin{bmatrix}K_1^a \\
\vdots \\
K_1^a
\end{bmatrix}
= [V_1^a]
\begin{bmatrix}V_1^a \\
\vdots \\
V_1^a
\end{bmatrix}$$

$$V_{\min} \leq V_1^a, V_2^a, \ldots, V_\phi^a \leq V_{\max}$$

Where the $v_{\min}^\alpha$ (Dead storage) and $v_{\max}^\alpha$ (Usable storage) represent the minimum storage and maximum storage of reservoir $z$. The superscripts $\phi$ is the number of the storage class intervals. $L_{t-1}^a$ and $K_1^a$ represent the representative value of the storage in class interval $\phi$ for equivalent reservoir at the end of time step $t-1$ and the beginning of time step $t$ separately.

The AD-SDP model is constructed by taking the aggregate storage and aggregate inflow as state variables. The $K_1^a$ and $L_{t-1}^a$ are the aggregate storages at the beginning and end of time step $t$. In order to measure the system performance in model iteration, the aggregate variables are disaggregated by the method of conditional expectation, which is introduced as below. The recursive equation of AD-SDP is defined in Eq.12.

$$f_{op}^t(K_t^a, F_t^a) = \max_{L_t^a} \left\{ B_t \left( K_t^a, F_t^a, L_t^a \right) \right. \right.$$
$$+ \sum_{j} P_{t+1}^j \left( Q_{t+1}^i / F_t^i \right) \cdot f_{op}^j \left( L_t^a, Q_{t+1}^i \right) \right. \right.$$

3.4. AD-BSDP model

In AF-SDP and AD-SDP model, these assume that the aggregate inflow constitutes a simple Markov Process and the uncertainty of forecasting inflow is not considered. In AD-BSDP model, the Bayesian Decision Theory (BDT) was used to address the uncertainty of forecasting inflow and the random in nature through both the posterior flow transition probability ($P(Q_{t+1}^i | F_t^i, Q_{t+1}^i)$) and the predictive probability of forecasts ($P(Q_{t+1}^i | Q_t^i)$). The recursive equation of AD-BSDP model is defined in Eq.13.

$$f_{op}^t(K_t^a, Q_{t+1}^i, F_t^a) = \max_{L_t^a} \left\{ \sum_{\tau} P_t^{\phi} \left[ Q_t^{\phi} / F_t^i \right] \cdot f_{op}^\tau \left( L_t^a, Q_t^{\phi} \right) \right. \right.$$
$$+ \sum_{j} P_{t+1}^j \left( Q_{t+1}^i / Q_t^{\phi} \right) \cdot f_{op}^j \left( L_t^a, Q_t^{\phi}, Q_{t+1}^i \right) \right. \right.$$

$$\right) \right.$$
4. Results

4.1. Quantitative Inflow Forecasting
The inflow forecasting models is applied to estimate the 10 days average inflow of Huanren. The observed rainfall and runoff data from 1968 to 2000 are used to calibrate the parameters of hydrological models. And the observed data from 2001 to 2010 are used to verify. Based on the work of calibration and verification, the forecasting data (10 days QPFs of GFS collected from 2001 to 2010) are applied to forecast the 10 days inflow of Huanren reservoir. The inflow of IC-HH and IC-HT are evaluated by the forecasting inflow of Huanren with the ratio of 10% and 5% respectively.

The Nash-Sutcliffe efficiency (NSE) is used to measure the performance of the hydrologic models during calibration, verification and forecasting. Results show that during dry season (from October to next April), the NSE value of calibration, verification and forecasting are only about 0.7. The main reason is hard to forecast the inflow during snowmelt for lack of meteorological data. Such as the data of temperature, the main factor affects snowmelt. During wet season (from May to September), the hydrological models perform well and the NSE value of calibration and verification exceed 0.9. The NSE value of forecasting reduces from 0.92 to 0.74 by convert the 10 days rainfall input data from observed to QPFs of GFS. Errors of the QPFs are the main source which influences accuracy of forecasting.

4.2. Simulation Results
The operating policies for Hun River cascade hydropower reservoirs are derived with the objects of maximize the total power supply and guarantee system stable running. Annual Total Hydropower Generation (ATHG), reliability and Annual Total Deficit Hydropower Generation (ATDHG), the important indicators for the managers in this study case, are chosen to study the performance of the system. Reliability of the system under a given policy is defined as the probability that the system output is satisfactory. The ATDHG is defined as the annual total deficit to the total firm power commitment during operation.

In order to comparing the performance of the operation policies (i.e., AF-SDP, AD-SDP, AD-BSDP), the existing Traditional Rule Curve (TRC), only taken the control water level as state variable, is selected as a model for comparison. The system ratio is the ratio by comparing the system indicator of operation policies with TRC. The operation policies are used to simulate the system operation by observed and forecasting inflow from 2001 to 2010 respectively. The ATHG of individual reservoirs and the system are listed in Table 2. The reliability and ATDHG of individual reservoirs and the system are listed in Table 3.

Results show that the performances of operation policies, deriving by AF-SDP, AD-SDP and AD-BSDP, are super than that of TRC. The state variable of TRC only is water level at current time step. The inflow information in future is not considered in strategy decision. So the decision-strategy is prone to spill, especially in wet season. The performance indicators of system ratio are demonstrated that the forecasting inflow information improved the utilization of hydropower resources in the study case.

The performance of AD-SDP is better than that of AF-SDP. The reasons are the AD-SDP model aggregates the reservoirs system as an equivalent reservoir and addresses the uncertainty of flow by Markov chain and forecasting inflow is treated as accurate information. In real-time operation, the obtained aggregate output needs to disaggregate among the reservoirs with the objectives of minimizing the spillages and maximizing the storage of the system. Thus, the disaggregate policies can control the release of Huanren to avoid or reduce spillages in Huilong and Taipingshao. Hence it produces higher ATHG and Reliability at Huilong and Taipingshao than AF-SDP.

The performance of the AD-BSDP is obviously better than that of AD-SDP. Because Quantitative Precipitation Forecasts (QPF) and runoff forecasts are not the accurate information. In the AD-SDP model, uncertainty of the flow was addressed by Markov chain and the forecasting inflow was treated as accurate information. While the AD-BSDP model used the BDT to address the uncertainties based on the equivalent reservoir.
Table 2. Simulation results of the Athg for observed and forecasting inflow from 2001 to 2010.

| Scenarios | Models | Inflow information | ATHG (MWH) | System ratio (%) |
|-----------|--------|--------------------|------------|------------------|
| 1         | TRC    | —                  | 372.58     | 986.39           |
| 2         | AF-SDP | observed           | 428.93     | 1043.57          |
| 3         | AD-SDP | observed           | 423.03     | 1078.42          |
| 4         | AD-SDP | observed           | 421.92     | 1099.76          |
| 5         | AF-SDP | forecasted         | 411.78     | 1013.65          |
| 6         | AD-SDP | forecasted         | 404.07     | 1050.09          |
| 7         | AD-SDP | forecasted         | 419.58     | 1086.24          |

Table 3. Simulation Results of the Reliability and Atdhg for Observed and Forecasting Inflow from 2001 To 2010

| Scenarios | Reliability (%) | System ratio (%) | ATDHG (KWH) | System ratio (%) |
|-----------|-----------------|------------------|-------------|------------------|
| 1         | 83.05           | 81.38            | 76.66       | 477              |
| 2         | 84.44           | 83.83            | 86.39       | 88.88            |
| 3         | 85.00           | 87.44            | 89.17       | 93.61            |
| 4         | 85.56           | 87.17            | 88.61       | 92.77            |
| 5         | 81.94           | 82.17            | 81.94       | 92.50            |
| 6         | 79.17           | 86.61            | 84.21       | 92.50            |
| 7         | 83.33           | 87.17            | 87.78       | 92.77            |

5. Summary and Conclusions

This study has focused on the deriving operation policies for cascade hydropower reservoirs by using medium-range precipitation forecasts. In view of the curse of dimensionality, the aggregation-decomposition method is used to overcome this problem by aggregation cascade reservoirs as an equivalent reservoir. Firstly the AF-SDP has been constructed by aggregating inflow, which is used to verify the efficiency and reliability of AD-SDP based on equivalent reservoir. In SDP model, the uncertainty of forecasting inflow is not considered. Then the AD-BSNP constructed by embedding the Bayesian Decision Theory in SDP to address the uncertainty of forecast inflow. Finally, the Hun River cascade hydropower reservoirs are taken as an example to evaluate the performance of the models. The analytical results indicate that the aggregation-disaggregation method is useful to overcome the curse of dimensionality in cascade hydropower reservoirs, and the approach, Bayesian Decision Theory, is critical for medium-range QPFs application in reservoir operation.

6. Acknowledgment

This work is supported by the National Natural Science Foundation of China (Grant No. 91547111, 51379027, 51609025) and Scientific Research Foundation for High-level Talents of North China University of Water Resources and Electric Power (Grant No. 201702013).

7. References

[1] Karamouz M, “Forecast uncertainty in reservoir operation,” In: Proceedings of the 15th Annual Water Resources Conference, Critical Water Issues and Computer Applications, 1988, ASCE, New York, pp: 265-268.

[2] Karamouz M, “Bayesian Decision Theory and Fuzzy Sets Theory in Systems Operation,” Proceedings of the 17th Annual Water Resources Conference, Optimizing the Resources for Water Management, 1990, ASCE, New York.
[3] Karamouz M and Vasiladiis HV, “Bayesian stochastic optimization of reservoir operation using uncertain forecasts,” Water Resources Research, vol. 28, May 1992, pp: 1221-1232.

[4] Kim YO and Palmer RN, “Value of Seasonal Flow Forecasts in Bayesian Stochastic Programming,” Journal of Water Resources Planning and Management, vol. 123, May 1997, pp: 327-335.

[5] Mujumdar, P.P. and Nirmala, B., 2007. “A Bayesian Stochastic Optimization Model for a Multi-Reservoir Hydropower System,” Water Resources Management, vol. 21, Dec. 2007, pp: 1465-1485, doi: 10.1007/s11269-006-9094-3.

[6] Tejada-Guibert J, Johnson S, Stedinger J. “The value of hydrologic information in stochastic dynamic programming models of a multireservoir system,” Water Resour Res, Vol. 31, Oct. 1995, pp: 2571-2579.

[7] Zhao, R. “The Xinanjiang model applied in China,” Journal of Hydrology, Vol, 135, Sep. 1992, pp: 371-381.

[8] Tang, G. et al., “Value of Medium-range Precipitation Forecasts in Inflow Prediction and Hydropower Optimization,” Water Resources Management, Vol. 24, Jan. 2010, pp: 2721-2742.