Quantifying the value of information in reservoir operations

K.A. Jusuf [a] and L. Raso [b]

[a]*Department of Environment Land Water and Planning (DELWP), Victoria, Australia*
[b]*Faculty of Technology, Policy and Management, Delft University of Technology, Delft, Netherlands*
Email: *Khalif.Jusuf@delwp.vic.gov.au*

Abstract: Operating a reservoir requires, at minimum, information about its current state, such as its storage level. In this case, the operation can be considered as purely reactive. Forecasts provide an additional benefit by introducing anticipation in the operation. Forecasts however, come at a cost, therefore it is helpful to know whether the added benefit of forecasts justifies their cost.

This work presents an innovative method for quantifying the value of information in reservoir operations. The method is based on tree-based model predictive control (TB-MPC). TB-MPC is an anticipatory method for controlling systems in real time. TB-MPC accounts for forecast uncertainty by representing forecasted scenarios in a tree structure, made of multiple out-branching. In this work, instead of using tree structures to control a system in real time, this paper proposes the use of tree structures as an analysis tool to estimate the overall operational performance under varying levels of information. This paper proposes the use of tree structures to emulate the decision-making process of a reservoir operator in how they respond to and/or anticipate all possible future conditions (scenarios) that the reservoir may be subjected to considering the information they have available.

Branching points in a tree structure occurs when a scenario can be distinguished by the operator from other scenarios. It represents the time when the reservoir operator is certain that one scenario is occurring instead of another based on the information that they have. Earlier branching points will result in improved operations, because they enable the operator to respond earlier to the occurring scenario and neglect the non-occurring ones. Better information (e.g. availability of forecasts, better measurements etc.) will move the branching point earlier in time.

In this paper, the application of this method is illustrated using simplified test cases. As indicated in the test cases, applying this method will require:

1. A method to generate ensembles of future inflows, such as an ARMA model, with parameters fitted to historical inflow observations. The ensemble size generated can be large and likely to be computationally demanding to process, therefore lumping similar ensemble members together is recommended to reduce computation times
2. A quantification of inflow observation and forecast uncertainty, such as a forecast error covariance matrix. This will influence the branching points of the tree structure, and
3. An objective function to optimise reservoir operations. This is to emulate the decision-making process of the reservoir operator.

The results of the test cases demonstrate that:

1. Better information results in tree structures with earlier branching points, and
2. Tree structures with earlier branching points result in improved reservoir operations

Using the method, the expected value of an information source (e.g. inflow forecasts) can be estimated by creating enough tree structures to sufficiently represent all conditions observed by the operator at present and all ensuing future scenarios the reservoir is subjected to. The overall value of information is then calculated by averaging the 'score' of the operations under all conditions according to their respective probabilities. If the objective function is expressed in monetary terms, this approach can potentially be used compare the value of having information against the cost of procuring and maintaining that information source.

Keywords: *Optimisation, operational water management, forecasting, value of information*

1. INTRODUCTION

Regulated water systems are often operated to meet multiple objectives that sometimes may conflict. An example of this conflict is to keep storage levels in a reservoir low enough to maintain spare capacity for flood protection, but also to keep it high enough to maintain reliability of water / hydropower supply.

Optimising the control decisions can provide a valuable contribution to effectively meet those objectives. Any improvement in the availability or quality of information will naturally improve the quality of decision making in operating the water systems.

This paper discusses a method to quantify the value of having inflow forecasts in reservoir operations. The method involves the application of a multistage stochastic optimisation framework to reflect the uncertainty faced by the operator, and how that uncertainty changes as more information is introduced. Simplified case studies are presented to illustrate the method.

2. BACKGROUND

In general, an operator of a water system needs to consider:

1. The objective of the operation. For example, maintaining water level at a set point, or maintaining reliability of supply while minimising flood risk
2. The state of their system at present. For example, the level of water in the reservoir.

The operation mentioned above can be considered as purely reactive (van Overloop, 2006). The operator looks at what is happening to their reservoir at present, and makes changes to the operation accordingly, to achieve the objective of the operation. Note that observations have uncertainties associated with them. For example, the operator can make a sub-optimal decision due to measurement error.

Forecasting allows water managers to approximate how their water systems will behave in the future. This allows them to adopt anticipatory operational decisions which may lead to better outcomes (van Andel, 2009). Streamflow forecasts are usually generated from rainfall runoff models, which utilise forecasted climate conditions as input to predict streamflow. This can be short term such as the forecasting involved in flood warning services, or long term such as seasonal streamflow outlooks. The focus of this work is short-term reservoir operations.

3. METHOD

3.1. Adopting real-time control methods

In the framework proposed in this work, quantification of operational performance involves adopting real-time control methods to emulate the decision-making process of the operator. The framework is based on Tree-Based Model Predictive Control (TB-MPC) (Raso et al., 2014), which is an evolution of Model Predictive Control (MPC). MPC optimises a control strategy (e.g. reservoir outflows) within specified constraints (e.g. spillway or pumping capacity) in response to a receding horizon of predicted disturbance (van Overloop, 2006).

In MPC, the predicted disturbance is deterministic. TB-MPC introduces an added complexity in that the predicted disturbance is an ensemble, structured as a tree. Ensemble forecast members tend to have small differences between each other at the early timesteps of the forecast horizon. Afterwards forecast uncertainty increases and the members diverge. A tree is generated from ensemble data by aggregating the scenarios with sufficiently small differences until a point in the horizon where they are deemed to be distinguishable from each other. In practice, this is to account for forecast uncertainty in controls while minimising computation time required for the optimisation of control action.

3.2. Tree-based multistage stochastic optimisation

In this work, instead of representing the uncertainty of a forecast ensemble used in real time control, the tree structure is used to represent the operator's perception of all possible future conditions their reservoir may be subjected to given a set of initial conditions they observe, and the information they have at hand over the control horizon. Therefore instead of being used as a control method, the tree structure is used as an assessment tool to estimate the utility of information sources.

For a simplified reservoir where the operator controls the outflows, the optimisation problem for a level of information I, is shown in Equation 1. In the equation, J^* is the overall cost function to be minimised, made of step-costs functions g, h is the optimisation horizon, u is the operational decision, v is the storage level, and q

is the reservoir inflow. n is the number of branches at the end of the horizon and p is their respective probabilities.

Equation 1 also specifies: (1) the tree structure that governs operational decisions, which is represented as a parent-branch relationship. Such a relationship defines for a given scenario (i), its parent scenario (P), and the point in time (B) where that member diverges from its parent. Note that the very first branch has itself as its parent and a branching point occurring in the first timestep, (2) a water balance of the reservoir, represented by model m, (3) constraints to the optimisation, such as the maximum and minimum volume and outflow capacity, (4) assumption of equal initial reservoir level for all scenarios, and (5) that a future trajectory of reservoir inflows is given for each observed initial flow.

Equation 1. Tree-based multistage stochastic optimisation of reservoir outflows

$$J^*(I, q_0, v_0) = \min_{u_{t,z}} \sum_{z=1}^{n} p(z) \cdot \left[\sum_{t=1}^{h-1} g_t(v_{t,z}, u_{t,z}, q_{t,z}) \right]$$

$where$:

(1) $u_{i,t} = u_{j,t}$ $when \begin{cases} P_I(i) = j \\ t < B_I(i) \end{cases}$ $i, j = 1, \dots, n$

(2) $v_{t,z} = m(v_{t-1,z}, u_{t,z}, q_{t,z})$

(3) $c(v_{t,z}, u_{t,z}, q_{t,z}) \leq 0$

(4) $v_{0,z} = v_0 \forall z \in \{1, \dots, n\}$

(5) $q_{t,z} | q_0$

3.3. Tree structure and optimisation

Tree structure is influenced by information. An adaptation of Bayes' Rule is used to conceptualise a way to represent the changing tree structures due to changing information. Bayes' rule governs how the probability of a given scenario change as more information is introduced. The method involves estimating the probability evolution of a scenario. Branching points occur in the tree structure when the probability of an occuring scenario increases to the level where the operator can distinguish the occuring scenario from the other non-occuring scenario. Details of how Bayes' rule is applied to develop a tree structure is explained in Appendix A.

The tree structure is then used in optimising reservoir operations. Optimisation is based on minimising or maximising an objective function which represents the goal of the reservoir operation. The equations and examples presented in this paper assumes an objective function to be minimised.

3.4. Expected value of information

The expected value of an information source (e.g. forecasts, more accurate gauges) can be estimated by running the tree generation and optimisation process under all foreseeable conditions.. In the context of reservoir outflow control, this is expressed in Equation 2. The foreseeable conditions are considered as all reservoir inflows (q_0), and water levels. (v_0) observed by the operator, whereas $p(q_0)$ and $p(v_0)$ their probability distributions. The probability distributions can be represented as histograms to simplify the computation.

Equation 2. Expected level of performance for a given level of information

$$V^*(I) = \int_{q_0} \int_{v_0} J^*(I, q_0, v_0) \cdot p(q_0) \cdot p(v_0) \cdot dq_0 dv_0$$

4. HOW INFORMATION AFFECT TREE STRUCTURES

4.1. Illustrative example

To illustrate how information affect tree structures that govern operational decisions, consider a situation in a simple reservoir where there are only two possible future scenarios, inflow (x_1) and no inflow (x_2). Whether or not there will be inflow, is determined by rain, so probability of x_1 is equal to the probability of rainfall, $P(x_1)$ = P_{rain}. Conversely, $P(x_2)$ = 1-P_{rain}. The operator of a storage has to release water from a storage to make room for this flow. Their aim is to keep the storage as full as possible without overtopping. If the operator only has

access to flow observations (as shown in Figure 1), they can only be sure that the flow event either will or will not occur at timestep 2.5, resulting in the control action illustrated in Figure 2.
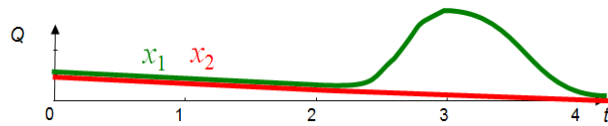

Figure 1. Example inflow observations received by an operator
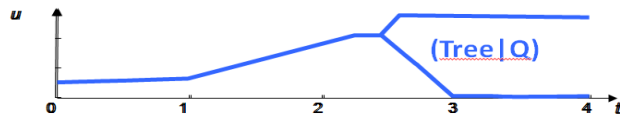

Figure 2. Example operational decision given the availability of observed inflow

However, if the operator also has access to rainfall observations (as illustrated in Figure 3), they will know whether the rainfall event that causes the flow event either will or will not occur at timestep 1 This results in the operational decision illustrated in Figure 4. It shows that that having rainfall observations in addition to flow observations is advantageous. The operator can anticipate the flow or no flow event sooner, as illustrated by the branching point occurring earlier. This allows more time for the operator to respond to the event that will eventually occur, reducing the risk of storage overtopping or excessive drawdown of storage.
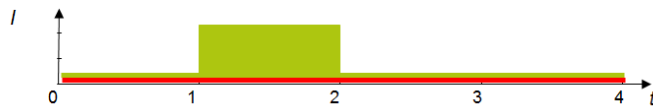

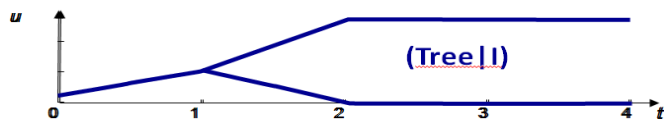**Figure 3.** Example rainfall observations received by an operator

Figure 4. Example control decision given the availability of rainfall observations

4.2. Test case using the proposed tree structure generation procedure

The example shown in Section 4.1 provides a simplified illustration of the concept behind the tree structure generation procedure presented in Appendix A. A test case was developed to confirm whether that procedure will produce tree structures with earlier branching points when better information is introduced. The test case assumes a 3 member inflow scenario ensemble which was arbitrarily created. These three scenarios are identical up until the 20th and 40th timestep, where they will diverge abruptly and significantly as shown in Figure 5.

There are two levels of information considered in the test case, one where only inflow measurements are available, and another where a forecast is also available. The variance of measurement error is assumed to be scalar with a value of 250,000 $(m^3/s)^2$. The covariance matrix of forecast error is assumed to be an identity matrix with 250,000 $(m^3/s)^2$ on the diagonal up to timestep 10, and 999,999,999 $(m^3/s)^2$ on the remaining diagonals. The covariance matrix essentially indicates that the forecast is able to predict up to 10 timesteps into the future with the same accuracy as the measurement. Beyond 10 timesteps, forecast accuracy is significantly diminished.

The tree structures generated from the procedure is shown in Table 1. The results show that if only measurements are available, the branching point occurs immediately after scenarios diverge. This is because the tree will only branch out only when divergence between possible scenarios are observable. If a forecast is also available, branching points occur 10 timesteps earlier, which is expected considering the assumed forecast error. The results of this test case demonstrates that better information will result in tree structures with earlier branching points.
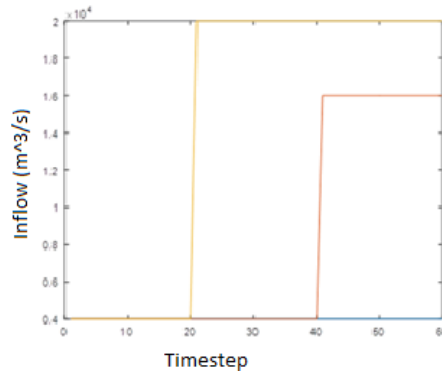
Figure 5. Inflow scenario ensemble for tree structure generation test case

Table 1. Tree structure (parent-branch relationship) generated from the tree structure generation test case

Scenario	1	2	3
Parent	1	1	1
Branch (Measurement only)	1	21	41
Branch (Forecast)	1	11	31

5. HOW TREE STRUCTURES AFFECT OPERATIONAL PERFORMANCE

To confirm that earlier branching point will result in improved operations, a simple test case was developed. The test case is based on a broader test case of TB-MPC in Salto Grande Reservoir, a multi-purpose reservoir in South America (Raso et al., 2014). This test case also demonstrates a simplified illustration of how the method may work in practice.

This test case uses a synthetically produced inflow scenario ensemble from an Autoregressive Moving Average (ARMA) model fitted to observed inflows to the reservoir. To obtain an adequate approximation of possible inflows, the ARMA model initially produces 2000 equiprobable inflow scenarios.

Only one inflow scenario ensemble is produced in the test case, with an arbitrarily selected initial observed inflow of 1800 m^3/s. As shown in Equation 2, to calculate the overall value of information, inflow scenario ensembles will need be created for all probable observed inflows. If the probability distribution of observed inflows is represented as a histogram, an inflow scenario ensemble is required for each histogram bin.

The inflow ensemble was consolidated using the method from Growe-Kuska et al. (2003) which lumps similar scenarios together, thereby combining their probabilities. The scenarios were reduced to two non-equiprobable scenarios. This was done to provide a clear illustration on the effect of branching points on operational performance and to minimise computation time. In practice, it is expected that the number of scenarios to be higher, to maintain an adequate representation of likely future inflows, while keeping operable computation times.

Figure 6 shows the two scenarios used in the test case. Inflows under both scenarios are generally very similar up to timestep 10, before diverging slightly between timesteps 10 and 25, and more significantly from timestep 25 onwards.
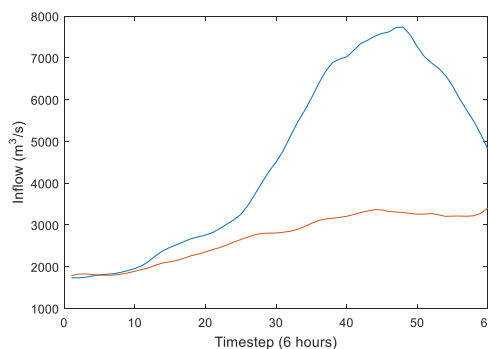

Figure 6. Two-member inflow scenario ensemble for operational performance test case

Four tree structures were arbitrarily created, each with a branching point of 1, 10, 20, and 30. The tree structures govern the optimisation of reservoir outflows to mainly keep water levels high enough to maximise hydropower outputs while keeping upstream and downstream levels low enough to minimise flood risks.

The results of the optimisation are shown in Table 2. Because there is almost no divergence between timestep 1 and 10, the score of the two trees are practically equal. The score increases slightly as the branching point increases to timestep 20, and further to timestep 25. The score increases more significantly with a branching point at timestep 30, where divergence between the two scenarios becomes much larger. This test case demonstrates that tree structures with earlier branching points result in more desirable objective function scores. The advantage of earlier branching points are most apparent when the scenarios significantly diverge.

Table 2. Optimisation results (note: lower values are more desirable)

Branching point	Objective function score (J^*)
1	4.18
10	4.18
20	4.27
30	5.83

6. CONCLUSIONS AND RECOMMENDATIONS

The method presented in this work uses tree structures to represent how information influences operational performance. The case studies which illustrates the method demonstrated that:

1. Better information results in tree structures with earlier branching points, and
2. Tree structures with earlier branching points result in improved reservoir operations.

This method also has the potential to assess the value of having information against the cost of procuring and maintaining that information source. This is done by expressing the objective function (J^*) in monetary terms and quantifying that for all foreseeable conditions to produce the expected value of information in monetary terms.

REFERENCES

Growe-Kuska, N., Heitsch, H., & Römisch, W. (2003). Scenario reduction and scenario tree construction for power management problems. Bologna: Power Tech Conference Proceedings, IEEE.

Raso, L., Schwanenberg, D., van de Giesen, N. C., & van Overloop, P. J. (2014). Short-term optimal operation of water systems using ensemble forecasts. *Advances in Water Resources 71*, 200–208.

Raso, L., van de Giesen, N., Stive, P., Schwanenberg, D., & van Overloop, P. J. (2013). Tree structure generation from ensemble forecasts for real time control. *Hydrological Processes*(27), 75-82.

van Andel, S.-J. (2009). *Anticipatory Water Management: Using ensemble weather forecasts for critical events.* Leiden: CRC Press/Balkema.

van Overloop, P.-J. (2006). *Model Predictive Control for Open Water Systems.* Amsterdam: IOS Press.

APPENDIX A: TREE STRUCTURE GENERATION

The probability of a scenario occurring given a level of information can be calculated using the Bayes' rule shown in Equation 3. $p_t(x)$ represents the prior probability of a scenario, and $p_{t+1}(x|I_t)$ represents the posterior probability distribution given an information level I_t.

Equation 3. Bayes' rule for continuous, 1-dimensional observations (Raso et al., 2013)

$$p_{t+1}(x|I_t) = \frac{f_{I_t}(I_t|x)p_t(x)}{\sum_{j=1}^{N} f_{I_t}(I_t|x_j) p_t(x_j)}$$

Equation 3 cannot be applied directly in tree generation, because a tree is generated at the initial timestep when the information is not yet available. To address this, Raso, et al. (2013) proposes using the Bayes' rule on average, at a given point in time, assuming a certain scenario is the occuring one. This is shown in

Equation 4. x_k can be considered as the scenario that is assumed to be the occuring one. Observations and forecasts for that occuring scenario will be taken from the distribution $f_{I_t}(I_t|x_j)$. The variance of the distribution is governed by the error of the information source (e.g. forecast error) over the horizon. The horizon is the timeframe after the present that the operator is considering when optimising their operational decisions. Integrating such functions as prescribed in Equation 4 is clearly complicated. To simplify, integration can computed using Monte Carlo integration by generating random samples of observation vectors.

Equation 4. 'On average' application of the Bayes' rule for a given information level and occurring scenario (Raso et al., 2013)

$$E\big[p_{t+1}\big(x|(I_t|x_k)\big)\big] = \int_{-\infty}^{\infty} p_{t+1}(x|I_t) \cdot f_{I_t}(I_t|x_k)dI_t$$

Equation 4 is to be applied for every timestep along the forecast horizon so that the probability dynamics of the occuring scenario x_k can be quantified. This is shown in Equation 5 where it defines how the probability of x_k will approach one as time moves forward towards the horizon, while the probability of other scenarios will approach zero. Note that ϕ can be considered as a dependence function in which

Equation 4 is rearranged so that its inputs are clearly presented.

Equation 5. Dynamic probability of scenarios (Raso et al., 2013)

$$(a) \; p_t(x) = p(x) \; for \; t = 1$$

$$(b) \; p_{t+1}(x) = \; \phi\big(p_t(x), f_{I_t}(I_t|x), x_k\big) \; for \; t = 1, \dots, H-1$$

This dynamic probability determines the point in time where one scenario is distinguishable from another. This can also be referred to as the bifurcation point of a tree. This is done by comparing two scenarios at a time, x_k which is the occuring one, and a different member x_j to calculate the probability of x_k in this reduced sample space. This is shown in Equation 6.

Equation 6. Occurring scenario probability in a reduced sample space (Raso et al., 2013)

$$p^r_t(x_k) = \frac{p_t(x_k)}{p_t(x_k) + p_t(x_j)}$$

From that, a distinguishability matrix can be produced. A distinguishability matrix defines, between every pair of scenarios, the time when the probability of the occuring member in the reduced sample space $p^r_t(x_k)$, exceeds a predefined confidence level p^*. This is shown in Equation 7. A probability confidence level of 0.95 is adopted in this work.

Equation 7. Distinguishability matrix (Raso, et al., 2013)

$$D_{k,j}(p^*) = t|min_t[p^r_t(x_k) > p^*] \; for \; \forall \; x_k \neq x_j$$

Once the distinguishability matrix has been created, a tree structure can be defined from it using the recipe prescribed in Raso, et al. (2013) below.

1. Find the maximum value of D (maxD) and the member corresponding to its row and column, called 'row' and 'column' member
2. The maxD value and the column member are assigned to the row member as branching point and parent, respectively
3. Matrix D is reduced, removing row and column of 'row member' (its row and its column)
4. The procedure from (1) to (3) is repeated until a single member remains
5. The last member is the root, its parent is itself and its branching point is 1