

A daily hydrological system management model that takes meteorological forecast errors into account

S. Krau¹, M. Latraverse², D. Tremblay² & A. Turgeon¹

¹École Polytechnique de Montréal, Canada

²Hydro-Québec, Canada

Abstract

Meteorological forecasts for the next few days are usually not detailed enough to determine the exact probability distribution of the forecast precipitation, and, hence, the probability distribution of the forecast inflow into the hydrological system. However, the hydrological system manager is responsible for assessing the risk of flooding and violating constraints when operating the reservoirs. This paper deals with the problems involved in modeling forecast errors and of determining an operating policy that takes such errors into account. This is accomplished by first building an inflow scenario tree that takes forecast errors into account and then by solving the reservoir management problem with that tree. This paper describes how the tree was built.

Keywords: meteorological, forecast, hydrological, management, risk, flooding, daily, stochastic, optimisation.

1 Introduction

The management of a hydrological system often implies the pursuit of various divergent objectives. There is a need to optimize power plant generation but also to regulate the flow of rivers, ensure a sufficient supply of drinking water, comply with vacationing-related constraints and ensure the safety of waterside communities as well as that of the structures. Management models most often take these different objectives into account through flexible constraints, i.e. constraints for which violation is tolerated, while minimizing the overall costs incurred by such potential violations. Daily management can then be reduced to a stochastic optimization problem related to a single objective. It consists in



determining for the current day the water released from the reservoirs which maximize the hope of future profits, with the latter defined as the difference between the value of the hydroelectric generation and the total cost of violating the constraints.

The management of a hydrological system becomes especially critical during periods of high runoff. Several constraints run the risk of being violated and operating margins become very tight. In these conditions, run-of-river systems are the most vulnerable due to their low storage capacity and the inhabited areas they cross. An adequate representation of the uncertainty of water inflows within a management model thus becomes critical.

The uncertainty of water inflows stems from errors in the calibration of the hydrological model parameters, uncertainty regarding the water conditions in catchment areas, and the inaccuracy of weather forecasts. Weather forecasts are generally based on a deterministic meteorological model that only proposes one precipitation and temperature scenario over a ten-day time frame. Obviously, such forecasts are not perfect and it would be useful to know its uncertainty in order to soundly manage the hydrological system.

We are proposing to formulate the daily management problem using a multistage stochastic mathematical program with recourse within which the uncertainty of the deterministic meteorologist forecast is represented. We will present the perturbation methods used to transform a deterministic meteorological forecast into a probabilistic meteorological forecast and build a water inflow scenario tree.

2 General description of the management problem

The hydrological system management problem has received substantial treatment in scientific literature. Starting with a general mathematical formulation, Labadie [8] shows that a water inflow management model can be written as a non-linear, non-convex and possibly large-scale stochastic optimization problem. The author presents the state of the art for the various approaches used in practice to solve the problem. We have chosen the linear programming approach, one that is commonly used (Loucks [11]) since it can serve to easily model the constraints that temporally link decisions, such as the constraints pertaining to the flow times between reservoirs or those associated with flow variation boundaries. This approach benefits in practice from the sturdiness and power of commercial linear programming software.

A linear management model is obtained by approximating separable non-linear functions from non-separable ones (generation functions) in the vicinity of a point given by non-linear functions in a piecemeal manner. The problem of finding the water release from the current day t which maximizes the hope of future profits is then formulated as a multistage linear stochastic mathematical program with recourse (Birge *et al.* [2]). This program can be reformulated into a deterministic linear mathematical program known as a *deterministic equivalent* (Birge *et al.* [2]). All that is needed is to know the discrete distributions of the daily inflows, and then build the resulting decision tree.



Note that the size of the constraint matrix increases linearly with the number of nodes in the inflow tree. Should the tree become too large in size, modeling compromises and/or the use of decomposition methods (Birge [3]) would then have to be used to solve the program.

3 Characterization of the meteorological forecast error

Uncertainty in meteorological forecasting is predominant compared to the other sources of uncertainty related to a hydrological forecasting model (Krzysztofowicz [7]). Coulibaly *et al.* [4] show that, even for fairly large errors, meteorological forecasting significantly improves the precision of inflow forecasts over a 7-day time frame. Numerous methods have been developed to characterize the uncertainty related to inflow forecasting (Tamea *et al.* [13]; Collischonn *et al.* [4]; Lefevre [10]). Maskey *et al.* [12] proposed an approach used to represent the uncertainty of meteorological forecasting within a deterministic hydrological model. To our knowledge, there is no management model that explicitly integrates uncertainty in meteorological forecasting.

The deterministic meteorological forecasting that we are considering is provided to Hydro-Québec by Environment Canada and takes the form of a set of grids at six-hour intervals over a nine-day time frame, for a total of 36 grids. The first 48 hours are generated by the regional climate model (15-km increments) and are available in six-hour blocks, while days 3 to 9 are generated by the global model (110-km increments) and are available in 12-hour blocks. Hydro-Québec forecasters increase the value over the first 24 hours and merge the regional and global models, while producing 6-hour blocks for the entire period, followed by an interpolation at every 10 km. The 36 proposed precipitation grids are then merged into 9 daily forecasting grids. The daily forecasting grids are set, i.e. they cover the same geographic region. At each of the grid points (spaced 10 km apart) is found a total amount of forecast precipitation along with the daily minimum and maximum forecast temperatures.

Given the current day t and a deterministic meteorological forecast issued over T days, there are T grids of daily meteorological forecasts issued. Let $\hat{\Pi}_{(t+T)|t}$, $\hat{Z}_{(t+k)|t}$, $\hat{T}_{i,j,(t+k)|t}^{\min}$, $\hat{T}_{i,j,(t+k)|t}^{\max}$ be respectively the daily forecast grid for day $t+T$, issued on day t , the amount of forecast precipitation at point (i,j) for day $(t+k)$ issued on day t , the minimum and maximum forecast temperature at point (i,j) for day $(t+k)$ issued on day t . Let (x_i, x_j) , I, J , be respectively latitude and longitude coordinates of grid points $(i, j) \in I \times J$, the number of indexes of longitudinal and latitudinal grid points.

We are proposing a characterization of forecasting errors for the amount of forecast precipitation and based on the history of variations between the forecasting grids and the observation grids. The observation grids are built based on an extrapolation of the values recorded at the observation stations located in the hydrological system catchments. Given eqn. (2) that links the amount of forecast precipitation to the precipitation forecast issued on day t for the day $(t+k)$, to grid point (i,j) .



$$z_{i,j,(t+k)} = \varepsilon_{i,j,(t+k)|t}(\hat{z}_{i,j,(t+k)}); \forall t; \forall (i, j) \in (I \times J); k = 0..T-1 \quad (1)$$

where:

$z_{i,j,(t+k)}$ is the random variable representing the quantity of precipitation observed at point (i, j) , for day $(t+k)$ and $\varepsilon_{i,j,(t+k)|t}(x)$ is the random function representing the uncertainty of a deterministic forecast issued on day t , at point (i, j) , for day $(t+k)$. Functions $\varepsilon_{i,j,(t+k)|t}(x)$ are spatially and temporally correlated and the full determination of their parameters is complex, even impossible, to realize given the field's spatial and temporal range.

The forecasting uncertainty is made up of three types of uncertainties: (1) uncertainty regarding the total amount of forecast precipitation over the forecasting time frame, (2) uncertainty regarding the temporal distribution of the precipitation, (3) uncertainty regarding the spatial distribution of the precipitation. We are representing the first two uncertainties by a perturbation (random data transformation) that we will call *temporal perturbation of the forecast precipitation*. We will call the uncertainty related to the spatial distribution of the forecast precipitation over the forecasting time frame *spatial perturbation of the forecast precipitation*. We assume that these two perturbations are independent. They are variations of the perturbations proposed by Latraverse [9] regarding the same forecasting grids over a two-day time frame. The perturbations that we are proposing can be used to transform deterministic meteorological forecasting into probabilistic forecasting in the form of a scenario tree.

4 Temporal perturbation of the forecast precipitation

Temporal perturbation consists in reasoning not on the amount of forecast precipitation at the grid points but on the total amount of precipitation on the grids. Latraverse [9] shows forecasting errors on these quantities are independent from one day to the next. The random variable q_{t+k} , associated with the total amount of precipitation observed on the grid for day $(t+k)$, verifies eqn. (3):

$$q_{t+k} = f(\hat{q}_{(t+k)|t}); k=0..T-1; \forall t. \quad (2)$$

with $\hat{q}_{(t+k)|t} = \sum_{i \in I} \sum_{j \in J} \hat{z}_{i,j,(t+k)|t}$, $q_{(t+k)|t} = \sum_{i \in I} \sum_{j \in J} z_{i,j,(t+k)|t}$ for $k=0..T-1$, $\forall t$,

and $f(\hat{q}_{(t+k)|t})$ a random variable. We propose to discretize the function $f(x)$ using discrete random variable $\Psi_1^V(\hat{q}_{(t+k)|t})$. Given a history of forecasting grids and observation grids and the following sets (see eqn. (4)):

$$\hat{A}_k = \{\hat{q}_{(t+k)|t}; t = \underline{t}_h.. \overline{t}_h - k\}, A_k = \{q_{(t+k)|t}; t = \underline{t}_h.. \overline{t}_h - k\}; k=0..T-1 \quad (3)$$



with $\overline{t_h}$, $\underline{t_h}$ and K , be respectively the history starting date, the history ending date and the number of days in typical forecasting time frame. Set \hat{A}_k is made up of the amount of forecast precipitation based on a forecast history and associated with k th day. Set A_k is made up of observed amounts of precipitation corresponding to elements of \hat{A}_k . Given \hat{E}_k^v respectively E_k^v , the subset of set \hat{A}_k made up of v elements that are closest to $\hat{q}_{(t+k)|t}$, respectively the subset of set A_k made up of the observed elements associated with elements of \hat{E}_k^v . Given an arbitrary division of E_k^v (see eqn. (4)) into N_k classes ($k = 0..T-1$):

$$E_k^v = C_k^1 \cup \dots \cup C_k^{N_k}; C_k^{i_k} \cap C_k^{j_k} = \emptyset; \forall i, j; i_k = 1..N_k; j_k = 1..N_k. \quad (4)$$

Let $c_k^{i_k}$ be the centroid of the class $C_k^{i_k}$ for $i_k = 1..N_k$. Random variable $\Psi_1^v(\hat{q}_{t+k})$ for a given t is defined by N_k amounts of precipitation and eqn. (5):

$$\text{Prob}[\Psi_1^v(\hat{q}_{t+k}) = c_k^{i_k}] = P_1(k, i_k) = \frac{|C_k^{i_k}|}{|E_k^v|}; i_k = 1..N_k; k = 0..T-1 \quad (5)$$

Given a deterministic meteorological forecast of the total amount of daily precipitations $\{\hat{q}_{t|t}, \hat{q}_{(t+1)|t}, \dots, \hat{q}_{(t+T-1)|t}\}$ issued at the start of current day t . Based on the distribution of discrete random variables $(\Psi_1^v(\hat{q}_{t+k}); k = 0..T-1)$, an event tree could be built that represents the scenarios of the amounts of probable precipitations with as many stages as days in the forecasting time frame, the tree becomes enormous in size, even if the forecasting time frame is limited to only about a dozen days. We resort to the use of a scenario tree reduction algorithm (Heitsch and Romisch [6]) as to only keep a limited number of scenarios forming a reduced tree that is statistically close to the original tree. The following notation is used for the remainder of the article: N_k^i , i_k^- and $d_k^{i_k}$ respectively the number of nodes in the reduced tree for stage k , the predecessor of i_k th node of stage k of the tree ($i_k = 1..N_k^i$, for $k = 0..T-1$) and the amount of precipitations associated with node i_k in the reduced tree ($k = 0..T-1$).

5 Spatial perturbation of the forecast precipitation

The uncertainty regarding the spatial distribution of precipitations is important for the characterization of forecasting errors in relation to the hydrological



system management problem. In fact, the risk of flooding or spillage in a hydrological system for a given amount of precipitation differs considerably depending on the distribution of precipitations in the catchment areas that make up the hydrological system under study.

We are assuming, for modeling purposes, that the spatial distribution of the points in the grid for all the deterministic forecasting grids remains invariant; only the geographic location of the forecasting grids changes. The spatial perturbation takes the form of a spatial translation of the deterministic meteorological forecasting grids. The parameters for these translations are calculated based on the forecasting errors on the location of the maximum precipitation point associated with the first forecasting grid. In reality, the forecasting grids cover a larger area than the catchments in the hydrological system, which allows, post-translation, to have full coverage of the hydrological system catchment areas through the translated grids.

The choice of this modeling resides in approximating the following meteorological phenomenon. Low-pressure systems generally follow the same direction. We are attempting in this modeling to capture the fact that the forecast low-pressure system can travel more to the north or more to the south, while keeping the original spatial distribution of the associated precipitation.

Let $(\bar{i}, \bar{j}), (\hat{x}_{\bar{j}}, \hat{y}_{\bar{j}}), (x_{\bar{j}}, y_{\bar{j}})$ be respectively the indices for the maximum precipitation grid point $(\hat{z}_{\bar{i}, \bar{j}, t|t} = \text{Max}\{\hat{z}_{i, j, t|t}; i \in I, j \in J\})$ on the first forecasting grid, the forecast longitude and latitude coordinates of grid point (\bar{i}, \bar{j}) and the random variable corresponding the coordinates observed at the grid point (\bar{i}, \bar{j}) .

The forecasting error equation on the location of the point of maximum precipitation is written $(x_{\bar{j}}, y_{\bar{j}}) = g(\hat{x}_{\bar{j}}, \hat{y}_{\bar{j}})$, with $g(\hat{x}_{\bar{j}}, \hat{y}_{\bar{j}})$ a random variable.

We are proposing to discretize function $g(\hat{x}_{\bar{j}}, \hat{y}_{\bar{j}})$ with discrete random variable $\Psi_2(\hat{x}_{\bar{j}}, \hat{y}_{\bar{j}})$. The assumption is made that the forecasting errors longitudinally and latitudinally from the maximum point of precipitation are independent and follow normal laws.

Given R ($R \gg 1$) points $(e_{r,1}, e_{r,2})$, generated from a forecast location from the maximum precipitation point and normal laws of errors on the latitude and longitude. Given A_l , $l = 1..L$, the different catchments of the hydrological system and $(u_l = (x_l, y_l))$, $l = 1..L$, the point that minimizes the aggregate of the distances between it and all the generated points part of catchment $(A_l, l = 1..L)$. Given $u^* = (x^*, y^*)$ the point minimizing the aggregate of the distances between it and all the generated points that are not part of any of the catchments in the hydrological system. Eqns (7) and (8) then define the random variable $\Psi_2(\hat{x}_{\bar{j}}, \hat{y}_{\bar{j}})$.



$$Prob\left[\Psi_2\left(\hat{x}_i, \hat{y}_j\right) = \left(x_l, y_l\right)\right] = P_2(l) = \frac{\sum_{r \mid (e_{r,1}, e_{r,2}) \in A_l} 1}{R}; l = 1..L \quad (6)$$

$$Prob\left[\Psi_2\left(\hat{x}_i, \hat{y}_j\right) = \left(x^*, y^*\right)\right] = P_2(L+1) = 1 - \sum_{l=1}^L P_2(l) \quad (7)$$

The spatial perturbation corresponds to $L+1$ movements for all the forecasting grids. Each movement is defined by the translation which correlates the maximum precipitation point associated with the first forecasting grid, with one of the $L+1$ locations. The forecast maximum and minimum temperatures at the grid points are updated based on their altitudes in relation to those they had prior to the translations. For instance, a grid point associated with a given geographic altitude will see its forecast temperature decrease if, after translation, said grid point has a greater altitude.

6 Example

The *temporal perturbation of the forecast precipitation* allows a tree of probable precipitations to be obtained. The *spatial perturbation of the forecast precipitation* generates series of grids of translated forecasts compared to the grids in the initial deterministic forecast. The composition of the two perturbations allows a scenario tree to be obtained of probabilistic-forecast grids.

The following figure (fig. 1) depicts results obtained by the described perturbation method for a given day of the past. The hydrological system considered in this example is the Gatineau system, a runoff-runoff system in Quebec. We can see on each grid (in black line) the Gatineau hydrological system frontier. At the beginning of the 14th October of 2003, a daily meteorological forecast grid has been issued (top left corner in fig.1). The other grids have been generated by the perturbation method. The intensity of the precipitation is shown by a color gradation.

7 Conclusion

The integration of meteorological forecasting uncertainty in a daily hydrological system management model can substantially improve the safety of hydrological system operation, especially when the system's water conditions are near their limits. The precise characterization of the uncertainty of the forecasting grids is complex and even impossible to achieve given the vast spatial and temporal field that needs to be considered. Aggregations (perturbations) are then needed to decrease the problem's complexity while attempting to capture the variability of future probable meteorological conditions. The management model we are proposing is a piecemeal linear model within which the representation of meteorological forecasting uncertainty takes the form of a tree of inflow scenarios.



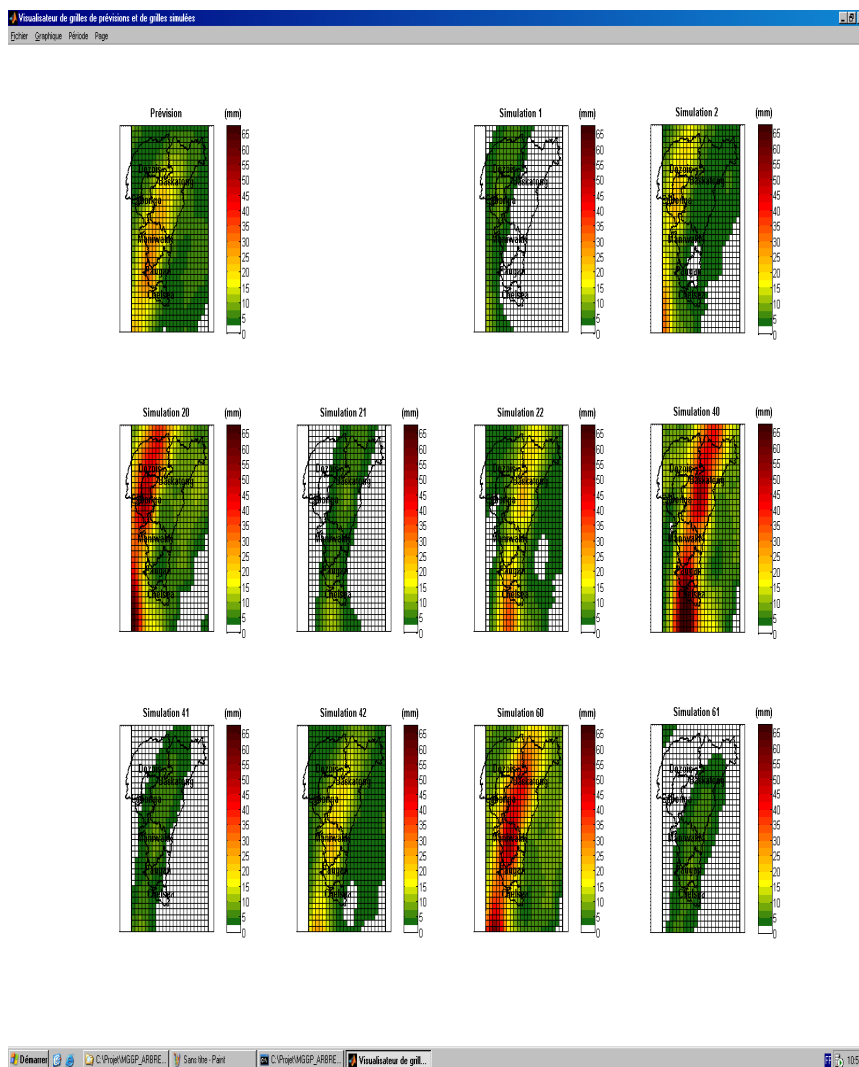


Figure 1: Example of grid perturbation method.

In the operational context, this management model is reset at the start of each day based on deterministic forecasting and must be solved fairly quickly so that decisions regarding reservoir water release are made on time. The transformation of the forecasting grid tree into an inflow scenario tree through the hydrological model is the most costly stage from the standpoint of computation time, which limits the number of meteorological scenarios that can be built for the tree. However, the parameters of the perturbation methods allow the meteorological scenarios that are to be represented in the forecast tree to be “chosen.” This allows the risk of flooding and/or spillage associated with extreme meteorological scenarios to be quantified over the forecasting time frame.

In our work, we confined ourselves to characterizing the uncertainty of precipitation forecasts, disregarding inflows resulting from the snowmelt. This therefore limits us to testing the approach in the summer or fall, when temperatures have very little influence on the transformation of precipitation into inflow. The proposed management model that is combined with the Hydrotel hydrological model is in the process of being tested on the Gatineau River hydrological system based on a history of several years of forecasting errors. If the results prove promising, we will integrate the notion of temperature in order to make the model applicable to the spring freshet.

References

- [1] Birge, J. R. and Louveaux, F., 1Springer Verlag, New York.1985.
- [2] Birge, J. R., Dempster, M., Gassmann, H., Gunn, E., King, A., and Wallace, S., A standard input format for multiperiod stochastic linear programs. Committee on algorithms newsletter, Vol. 17, pp. 1-19.1987.
- [3] Birge, J. R., Decomposition and partitioning methods for multi-stage stochastic linear programs. Operations Research, Vol. 33, pp. 989-1007.1985.
- [4] Collischonn, W., Hass, R., Andreolli I., Tucci, C. E. M. Forecasting River Uruguay flow using rainfall forecasts from a regional weather-prediction model. Journal of Hydrology.2005.
- [5] Coulibaly, P., Impact of meteorological predictions on real-time spring flow forecasting. Hydrological Processes, Vol.17, pp. 3791-3801.2003.
- [6] Heitsch, H., Romisch, W., Scenario reduction algorithms in stochastic programming. Computational Optimization and Applications, Vol. 24, pp. 187-206.2003.
- [7] Krzysztofowicz, R., Bayesian theory of probabilistic forecasting via deterministic hydrologic model. Water Resource Research, Vol. 35, pp. 2739-2750.1999.
- [8] Labadie, J. W., Optimal Operation of Multireservoir Systems: State-of-the-Art Review. Journal of Water Resources Planning and Management, Vol. 130, pp. 93-111.2004.
- [9] Latraverse, M., Prévision des apports à court et moyen terme et amélioration de la gestion des systèmes au fil de l'eau. Hydro-Quebec Production, Internal Report.2001.
- [10] Lefevre, M., Modélisation des erreurs pour le système de prévision PREVIS. Canadian Journal of Civil Engineering, Vol. 31, pp. 892-897.2004.
- [11] Loucks, D. P., Multiple-Reservoir Operation in North America. Water Science and Technology, Vol. 1, pp. 711-728.1981.
- [12] Maskey, S., Guinot V., Price, R. K., Treatment of precipitation uncertainty in rainfall-runoff modelling: a fuzzy set approach. Advances in Water Resources. 2004.
- [13] Tamea, S., Laio, F., Ridolfi, L., Probabilistic nonlinear prediction of river flows. Water Resources Research. Vol. 41, n.9, p. W09421.2005.



- [14] Turgeon, A., Solving a stochastic reservoir management problem with multilag autocorrelated inflows. *Water Resources Research*. Vol. 41, n. 12, p. W12414.2005.

