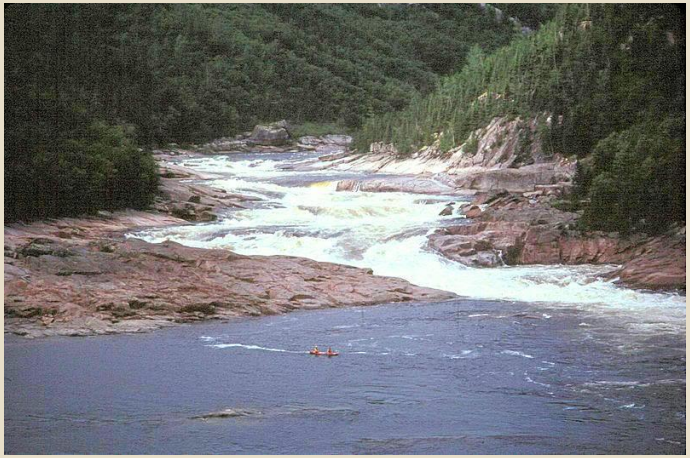
Correction of the under-dispersion of hydrological ensemble predictions and data assimilation

CENTRE DE RECHERCHES MATHÉMATIQUES

RioTinto



Richard Arsenault & Marco Latraverse

Centre de Recherches Mathématiques, Montréal, August 7th 2017

Presentation summary

- Problem context
- Hydrological modelling
- Ensemble forecasting
- Data Assimilation
- Expected outcome

Problem context

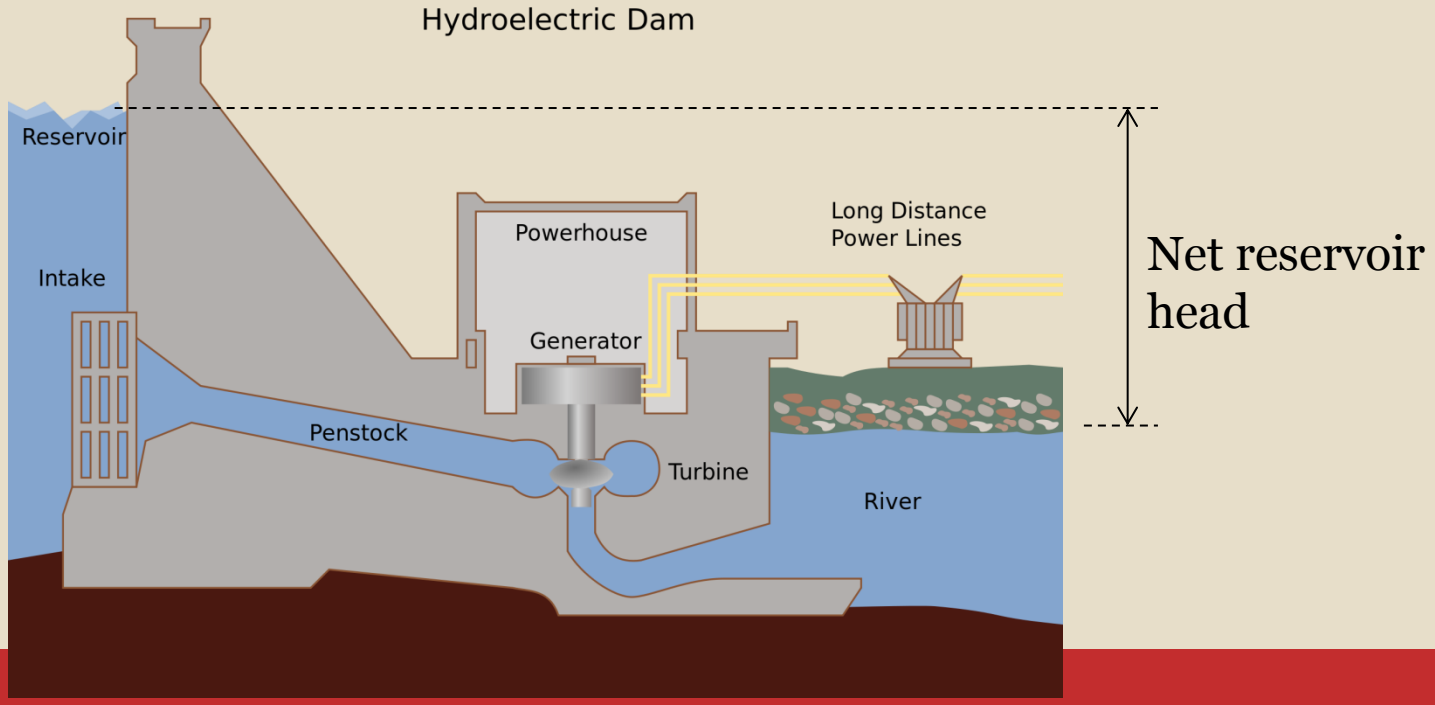
Problem context

- Rio Tinto Aluminium produces 90% of its own electricity
- 6 generating stations in our Quebec catchment (Saguenay-Lac-St-Jean region)
- Installed capacity of over 3000 MW

Problem context

Problem context

- Hydropower generation is a function of flow rate through the turbines and reservoir water level.
- Higher level = more efficient!
- Too high = risk of spilling if inflows are larger than expected



Problem context

- Our problem stems from the fact that weather is not perfectly known in advance. ☹
- This means reservoir inflows are unknown, but we can estimate them with a hydrological model and weather forecasts.
- Optimal management of hydropower reservoirs depends on good probabilistic forecasting to maximize expectancy of profits.

Hydrological modelling

- Deterministic black box model
- Empirical parameterized functions to represent hydrological processes
- Uses stochastic variables as inputs (precipitation, etc)
- Generates river flows and hydropower reservoir inflows.

Hydrological modelling

- Processes are time-dependent and processes are activated/deactivated through time

$$Q_t = M(w_t, IC_t) + \varepsilon_t$$

- Q_t: Observed inflows
- $M(\bullet)$: Hydrological model / inflow estimator
- w_t: Weather/climate inputs
- IC_t: Initial conditions (state variables)
- ε_t: Error term (includes error from hydrological model, weather observations and initial conditions)

Hydrological modelling

- The model states are persistent →Model has memory!

$$IC_{t+1}=IC_t + g(w_t)$$
$$IC_t=\{Sw_t,Uw_t,Sn_t...\}$$

- Sw_t : Soil Water balance
- Uw_t : Underground Water balance
- Sn_t : Snowpack water balance

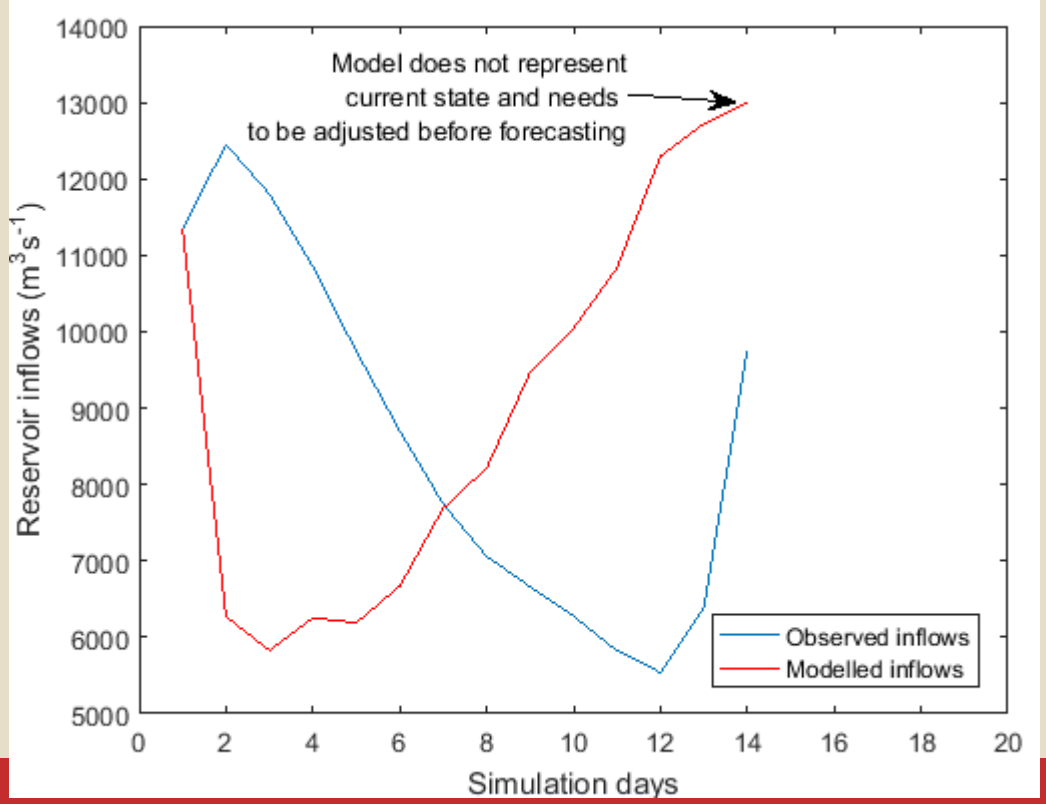
- Each is updated according to the active hydrological processes at each time step

Ensemble forecasting

- Single weather forecast for 14 days is guaranteed to be inaccurate.
- Solution → Use probabilistic forecasts!
- $\hat{Q}_t(1) = M(w_1, IC)$, $\hat{Q}_t(2) = M(w_2, IC)$, … , $\hat{Q}_t(n) = M(w_n, IC)$
- With n typically between 50 and 60.
- We want these probabilistic forecasts to have a mean bias of zero and adequate spread (variance)

Data Assimilation

- Oftentimes, the Initial Conditions (IC) do not represent the current state of the basin which causes an immediate bias in the simulation:



Data Assimilation

- We can modify the IC with a unique correction such as:

$$IC_t = IC_{t-1} + \Delta IC$$

- Or setting multiple initial states which will be used in the ensemble forecast to increase the number of members in the ensemble:

Ex. for ***n*** Weather time-series and ***m*** Initial Conditions:

$\hat{Q}_t(1,1) = M(w_1, IC_1)$, $\hat{Q}_t(2,1) = M(w_2, IC_1)$, … , $\hat{Q}_t(n,1) = M(w_n, IC_1)$

$\hat{Q}_t(1,2) = M(w_1, IC_2)$, $\hat{Q}_t(2,2) = M(w_2, IC_2)$, … , $\hat{Q}_t(n,2) = M(w_n, IC_2)$

$\hat{Q}_t(1,m) = M(w_1, IC_m)$, $\hat{Q}_t(2,m) = M(w_2, IC_m)$, … , $\hat{Q}_t(n,m) = M(w_n, IC_m)$

Expected outcome

- We need to find the best possible predictor for the 14 day reservoir inflow volumes during the summer-fall seasons, given the model's initial conditions.

- A few pointers:

- Which model initial condition state variables should be modified?
- By how much?
- Set constant error term, use a heuristic-based approach, select from prior estimated distributions, etc.?
- Run Monte Carlo simulations to evaluate improvements?

Thank you!

Questions?