Hydroelectric Scheduling: Inflow Forecasting and Parallel Decomposition

Anderson Rodrigo de Queiroz — *ORIE*

Luana Medeiros Marangon Lima — *ORIE*

David P. Morton — *ORIE*

Elmira Popova — *ORIE*

informs 2010 Annual Meeting
Energizing the Future
NOVEMBER 7-10, 2010
AUSTIN, TEXAS

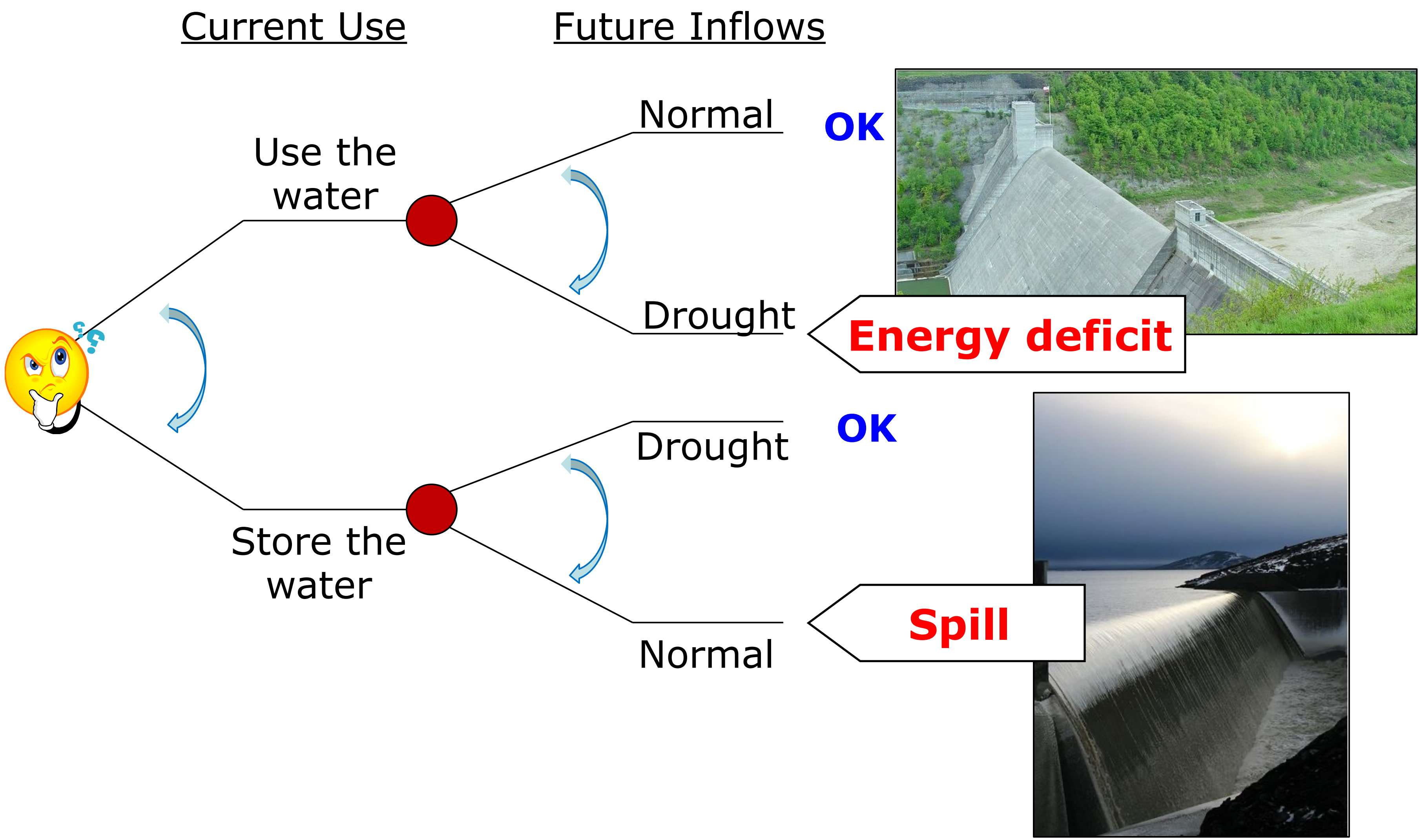
Outline

- Hydro Scheduling Problem Description
- Stochastic Models that Govern Inflows
- Problem Formulation
- Parallelization of the Decomposition Algorithm
- Future Research Goals

Hydrothermal Scheduling Problem

- Hydroelectricity is inexpensive to produce, yet depends on the supply of water (stochastic)
- Present decisions affect future conditions of the system and also future decisions (dynamic)
- Multiple interconnected reservoirs need to be scheduled for multi-period optimization (large-scale)

Decision Tree



Hydrothermal Scheduling: Brazilian Interconnected Power System

- 85% of electricity capacity comes from hydro
- Current model
  - optimization over 60 stages to determine the generator dispatch
  - agregated reservoir scheme
  - Forecast energy instead of inflow
  - Stochastic Dual Dynamic Programming
- Because of the big hydro dependency the inflows at the reservoirs are very important input parameters to the model

Probabilistic Model for the Inflows

Probabilistic Model for the Inflows

- We want to forecast future inflows to the system of reservoirs and rivers
- Model requirements to suit the optimization problem
  - Additive
  - Markov property
- We have the historical data for the monthly inflows per reservoir from 1931 to 2009

1	1931	178	371	326	479	332	226	125	89	112	192	153	215
1	1932	449	344	214	72	68	98	81	71	73	92	102	240
1	1933	287	161	147	108	84	72	69	64	63	67	72	141
1	1934	196	96	112	79	60	49	44	38	40	48	57	216
1	1935	242	381	184	169	121	98	73	73	69	86	86	92
1	1936	84	108	229	128	90	67	57	58	65	60	87	200
1	1937	349	231	170	121	128	95	74	62	57	136	158	378
1	1938	265	298	227	154	133	105	88	85	97	125	135	256

.
.
.

1	2005	254	208	176	114	119	93	84	65	64	65	106	176
1	2006	129	153	150	86	68	60	53	45	54	69	112	170
1	2007	363	274	132	100	81	70	59	54	42	50	83	88
1	2008	113	221	195	141	88	75	60	51	53	68	119	254
1	2009	255	268	225	177	114	93	90	75	88	129	00	00

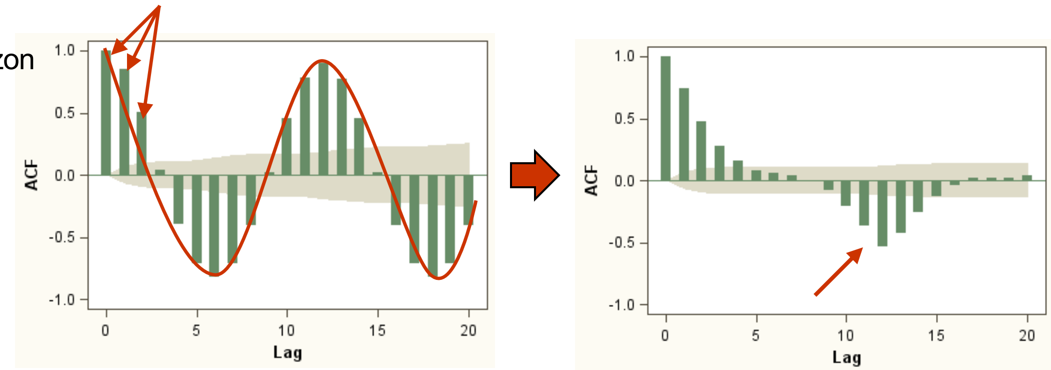
- 140 reservoirs → 140 variables

Generation Power Plants Geographically

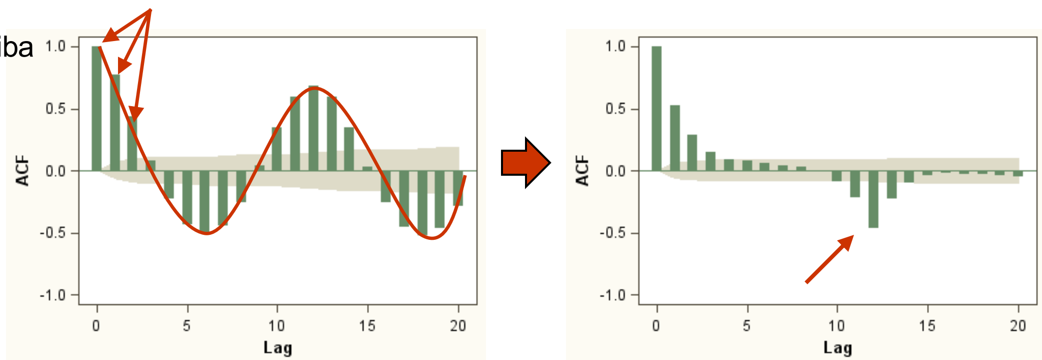
- The colored regions delineate the river basin boundaries
- The reservoirs operate in a cascade scheme
- To reduce problem size we aggregate reservoirs, and incremental inflows by basin
- Incremental inflow is the difference between a reservoir's natural inflow and that if the reservoir immediately upstream

Trend and Correlation Analysis

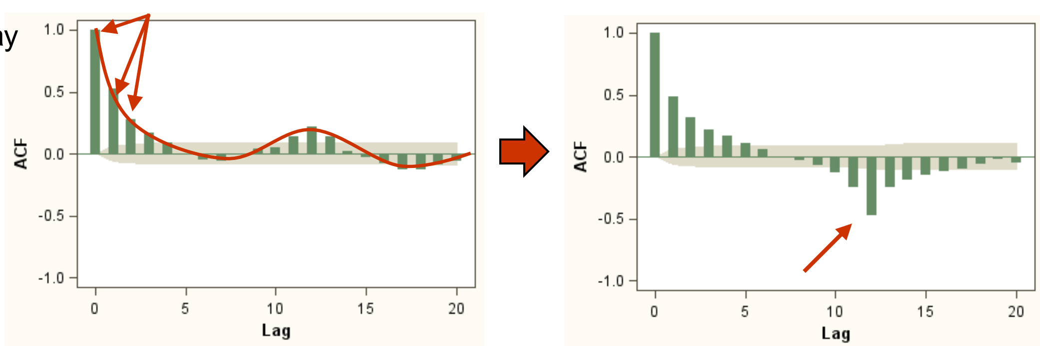
Amazon



Paranaiba



Uruguay



- Seasonality more or less pronounced depending on the basin
- Basins with greater seasonality exhibit non-stationary behavior
- Difference series at lag 12 does not completely eliminate seasonality

Univariate Model

- We fitted an ARMA model for each basin using the twice differenced series

- Conclusion
  - The log of the inflows give a better fit, but it leads to a non-additive model
  - The tails of the residual normality check do not look good even for the log model
  - There is correlation between the basins, so a multivariate analysis would be better to forecast the inflows.

Principal Component Analysis

First Component (9 basins)

Second Component (5 basins)

Third Component (1 basin)

The results show that geography explains 80% of the variability!

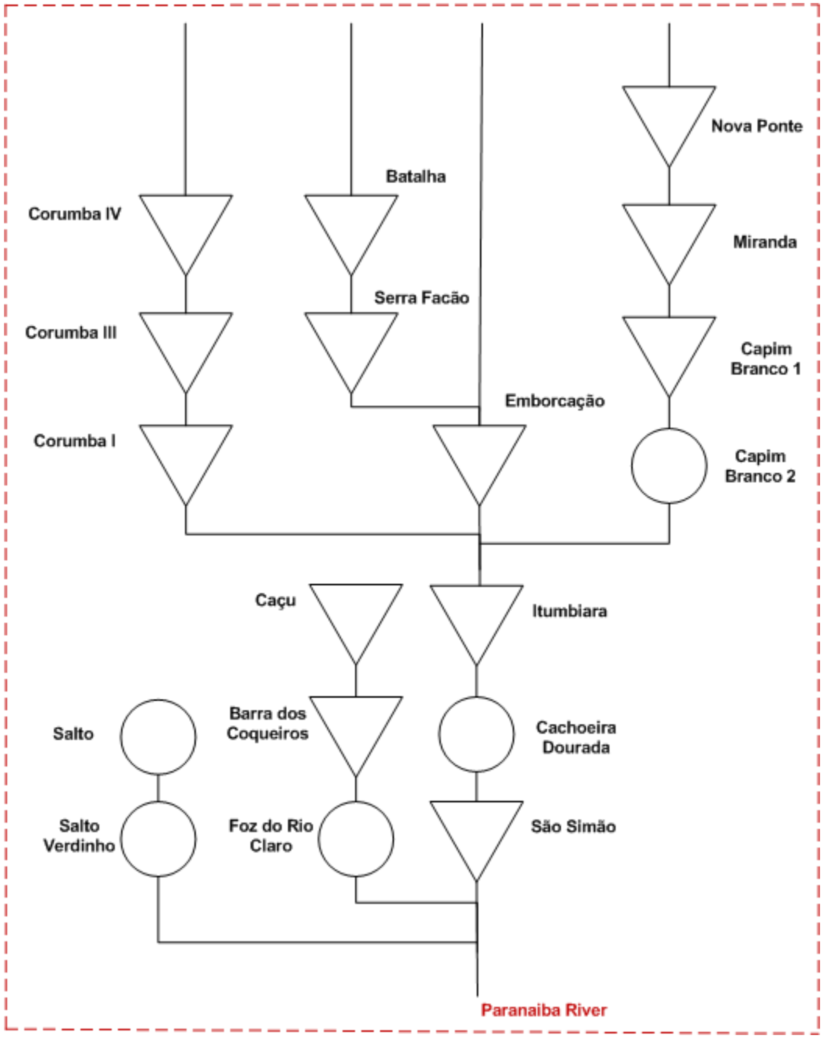
Future Goals

- Create a multivariate model for each component using Dynamic Linear Models
- Incorporate climate variables such as:
  - Precipitation
  - El Nino
  - Ocean temperature
- Generate scenario tree with the forecasted inflows as input for the optimization model.

Optimization Model

Hydrothermal Scheduling

- Objective is to minimize total expected cost to operate the system:
  - *Fuel costs for generating thermal power*
  - *Penalties for failure to meet demand*
- Decision variables: for each hydro plant, decision vector includes:
  - *turbined outflow volumes* GH_t
  - *spilled volumes* S_t
  - *reservoir storage level Vt*
- Uncertainty: future inflows b_t , b_{t+1} , … b_T



Stage t Problem Formulation

$$h_t(V_{t-1}, b_t^w) = \min \overbrace{\sum_{l \in L} c_{l,t} GT_{l,t}^w + \sum_{k \in K} u_{k,t} GD_{k,t}^w}^{\text{Present Cost}} + \overbrace{\sum_{w' \in W_t} p_t^{w'} h_{t+1}(V_t^w, b_{t+1}^{w'})}^{\text{Future cost function}}$$

(Water Balance) $\quad$ s.t. $\quad V_{i,t}^w + GH_{i,t}^w + S_{i,t}^w - \sum_{j \in M_i} (GH_{j,t}^w + S_{j,t}^w) = V_{i,t-1} + b_{i,t}^w \quad \forall i \in I$

(Meet Demand) $\quad \sum_{i \in I} \rho_i \, GH_{i,t}^w + \sum_{l \in L} GT_{l,t}^w + \sum_{k \in K} GD_{k,t}^w = D_t$

(Simple Bounds)

$$0 \le V_{i,t}^w \le \overline{V_{i,t}^w} \qquad \forall i \in I$$

$$0 \le GH_{i,t}^w \le \overline{GH_{i,t}^w} \qquad \forall i \in I$$

$$0 \le S_{i,t}^w \qquad \forall i \in I$$

$$0 \le GT_{l,t}^w \le \overline{GT_{l,t}^w} \qquad \forall l \in L$$

$$0 \le GD_{k,t}^w \qquad \forall k \in K$$

Stage t Benders Decomposition Master Program

- In a Benders' (or SDDP) decomposition algorithm, the stage t problem for each scenario is capture by the following master program

$$\begin{aligned}\min_{x_t,\theta_t} \quad & c_t x_t + \theta_t \\ \text{s.t.} \quad & A_t x_t = B_t x_{t-1} + b_t : \pi_t \\ & -G_t x_t + e\,\theta_t \geq g_t \quad : \alpha_t \\ & x_t \geq 0\end{aligned}$$

Where,

x_t: all stage t decision variables including: hydro generation, hydro storage, spillage, thermal generation...

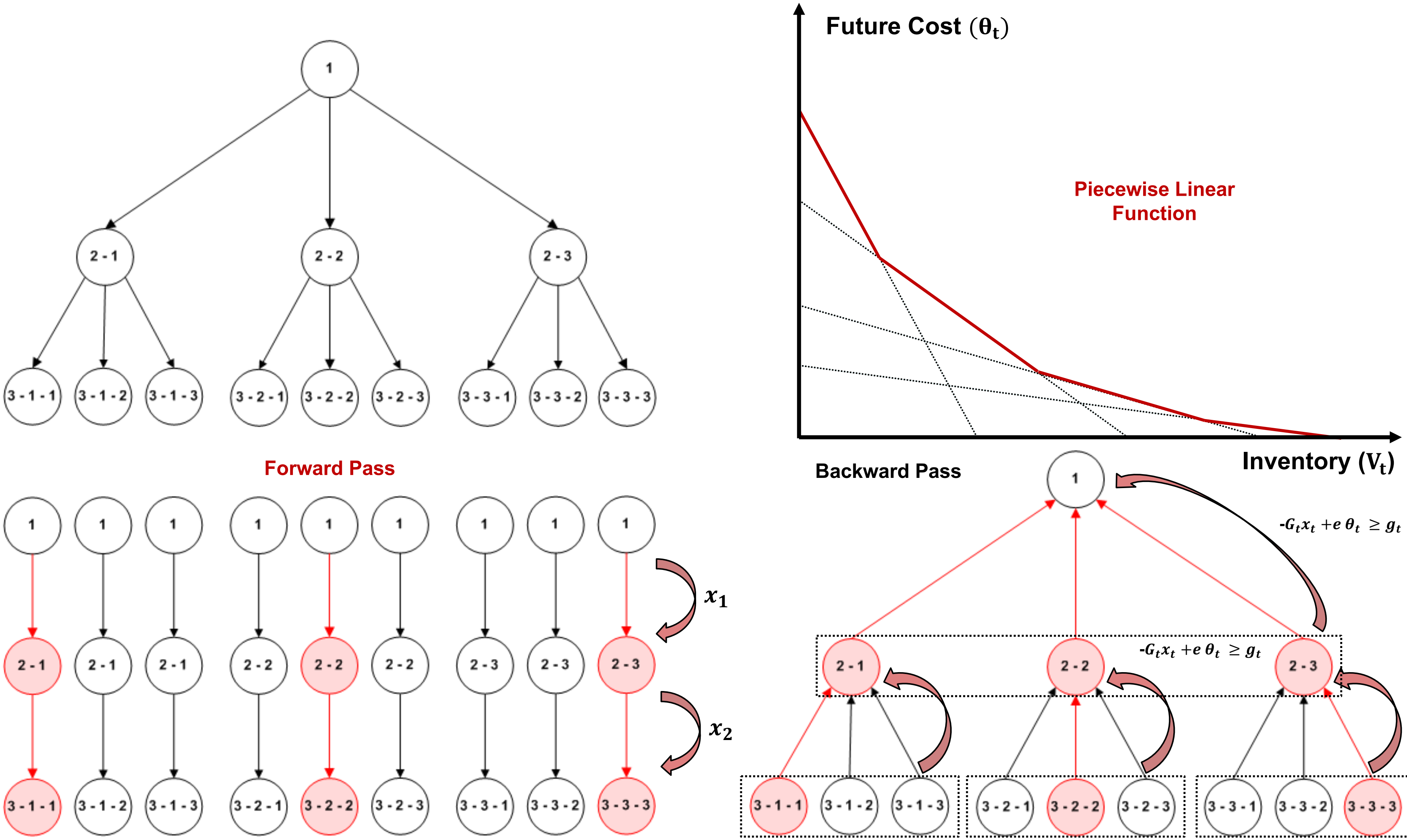
A_t: constraint matrix including water balance, meet demand, ...

b_t: stochastic inflow and deterministic demand

$B_t x_{t-1}$: is storage from last stage

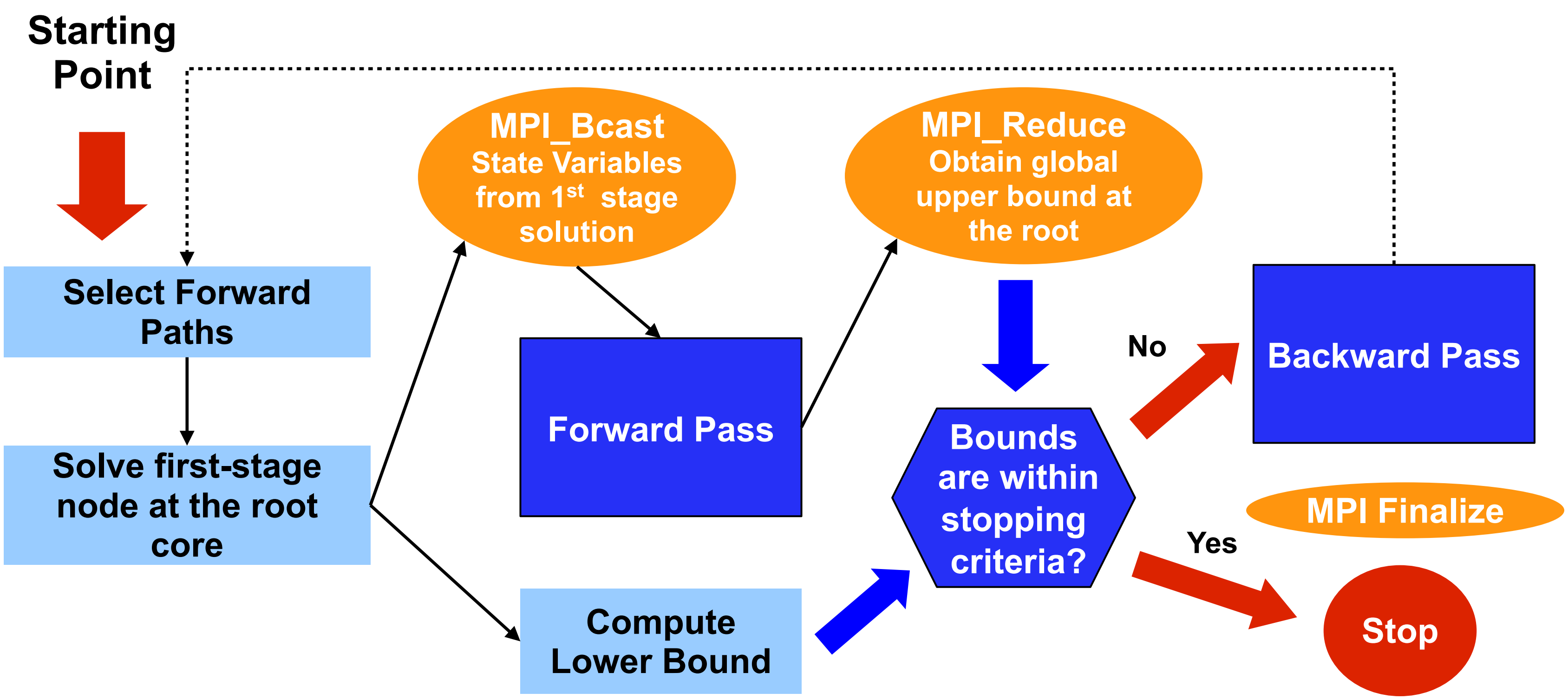
θ_t: future cost function

Sampling-Based Decomposition Algorithm

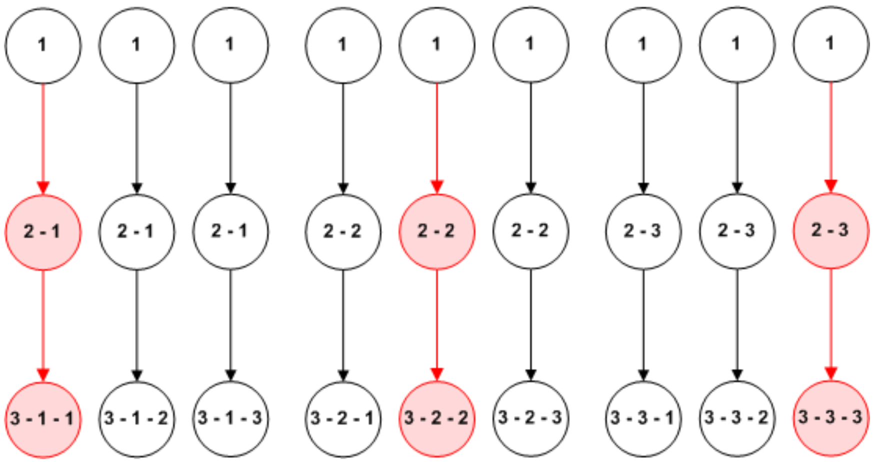


Algorithm Parallelization

- This type of algorithm can be parallelized, in both steps.
- MPI to communicate with the different cores.
- Synchronize using blocking collective communication calls.



Algorithm Parallelization *(cont.)*

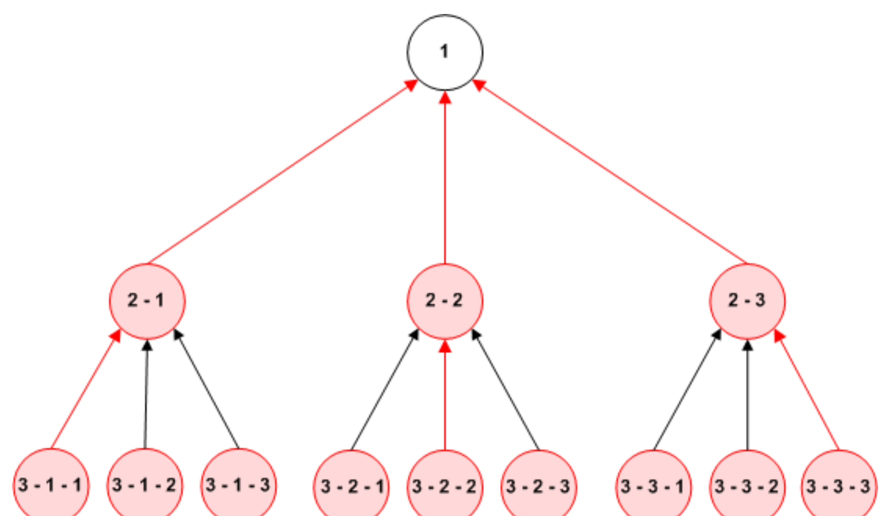


For each stage t = 2,...,T **Forward Pass**

- Check # of scenarios to solve
- **MPI_Scatterv** Send scenarios to other cores
- At each core: Solve each scen; Store state vars; Compute average upper bound
- **MPI_Gatherv** Get the state vars at the root
- **MPI_Bcast** State vars from root to other

For each stage t = T-1,…, 1 **Backward Step**

- Check # of scen to solve based on # of cores at the stage
- **MPI_Scatterv** Send scenarios to other cores
- At each core: Solve each scen; Store obj func value; Store Dual Prices
- **MPI_Gatherv** Get obj func values & dual prices at root
- Form cuts to use in the next stage
- **MPI_Bcast** cuts from root to others



Where are we right now?

- We have a univariate ARIMA model for the incremental inflows at each basin
- We have the serial and parallel version of the decomposition algorithm tested for a simple case with 3 stages and 2 possible scenarios at each stage
- The model considers inter-stage independency for inflows
- We are moving toward using real data and testing the algorithm on real-size instances
- Incorporate inflow forecasting into the optimization model
- Modify the algorithm to handle inter-stage dependency

Future Research Questions

- How the inflow forecasting model behaves with the addition of climate data for the river basins?
- Does the aggregated reservoir scheme behave well compared to a model with less aggregation or individual hydro plants?
- What is an adequate number of stages?
- Is monthly discretization appropriate?
- How does the algorithm scale with many processors available?
- Can we characterize the algorithm’s solution quality?

References

- Birge, J.R., Donohue, C.J., Holmes, D.F., Svintsitski, O.G., “ A parallel implementation of the nested decomposition algorithm for multistage stochastic linear programs”, Mathematical Programming 75 (1996) 327-352.
- Flach, B.C., Barroso, L.A., Pereira, M.V.F., “Long-term optimal allocation of hydrogeneration for a price-maker company in a competitive market: latestdevelopments and a stochastic dual dynamic programming approach” , IET Gener. Transm. Distrib., 2010, Vol. 4, Iss. 2, pp. 299–314
- Infanger, G., Morton, D.P., “Cut sharing for multistage stochastic linear programs with interstage dependency”, Mathematical Programming 75 (1996) 241-256.
- Pacheco, P.S. “Parallel Programming with MPI”, Morgan Kaufmann Publishers, 1997.
- Pereira, M.V.F., Pinto, L.M.V.G. “Multi-Stage Stochastic Optimization Applied to Energy Planning”, Mathematical Programming, No. 52, 1991.
- Pereira, M.V.F., Pinto, L.M.V.G. “Stochastic Optimization of a Multi-reservoir Hydroelectric System: A Decomposition Approach”, Water Resources Research, Vol.21, No. 6, Pages 779-792, June 1985.
- Philpott, A.B. and Guan, Z., “On the convergence of stochastic dual dynamic programming and related methods, Operations Research Letters, 36, 2008.
- Silva, E.L., Finardi, E.C., “Parallel Processing Applied to the Planning of Hydrothermal Systems” IEEE Transactions on Parallel and Distributed Systems, Vol. 14, No. 8, August 2003.
- Shapiro, A., "Analysis of Stochastic Dual Dynamic Programming Method", European Journal of Operational Research, to appear.
- West, M., Harrison, J., “Bayesian Forecasting and Dynamic Models”, Springer Series in Statistics, Second Edition 1997.
- Tsay, R.S., “Analysis of Financial Time Series”, Wiley-Interscience, Second Edition 2005.

Thank you!