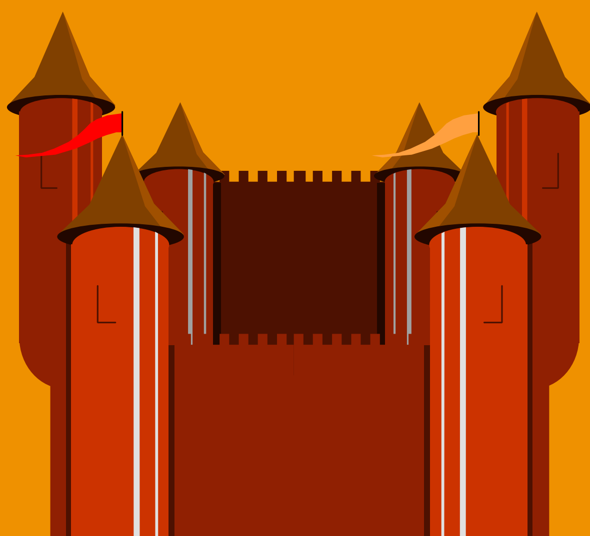
Approximate Dynamic Programming for High-Dimensional Problems in Energy Modeling

Workshop on Computatoinal Sustainability
Cornell University
June, 2009



Warren Powell
CASTLE Laboratory
Princeton University
http://www.castlelab.princeton.edu

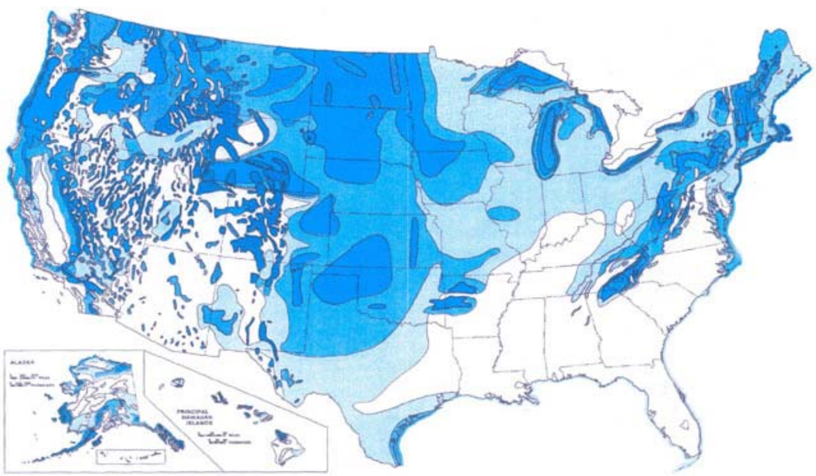
Goals for an energy policy model

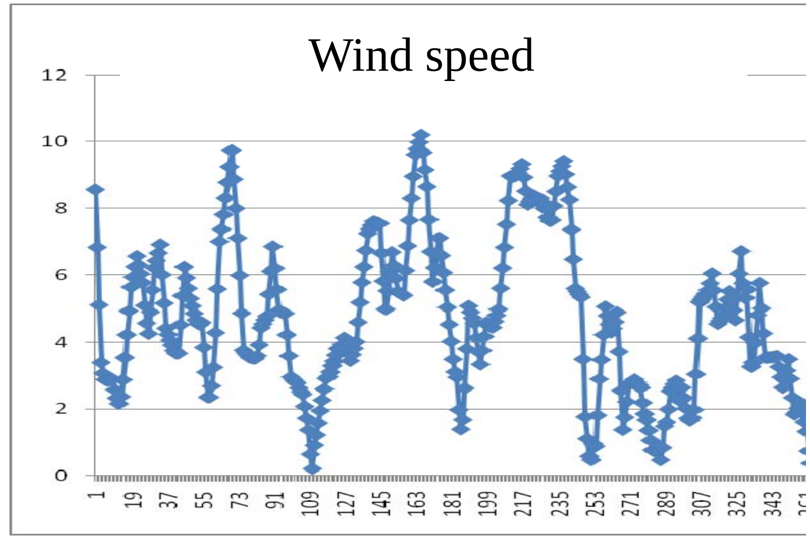
- Potential questions
  - » Policy questions
    - How do we design policies to achieve energy goals (e.g. 20% renewables by 2015) with a given probability?
    - How does the imposition of a carbon tax change the likelihood of meeting this goal?
    - What might happen if ethanol subsidies are reduced or eliminated?
  - » Energy economics
    - What is the best mix of energy generation technologies?
    - How is the economic value of wind affected by the presence of storage?
    - What is the best mix of storage technologies?

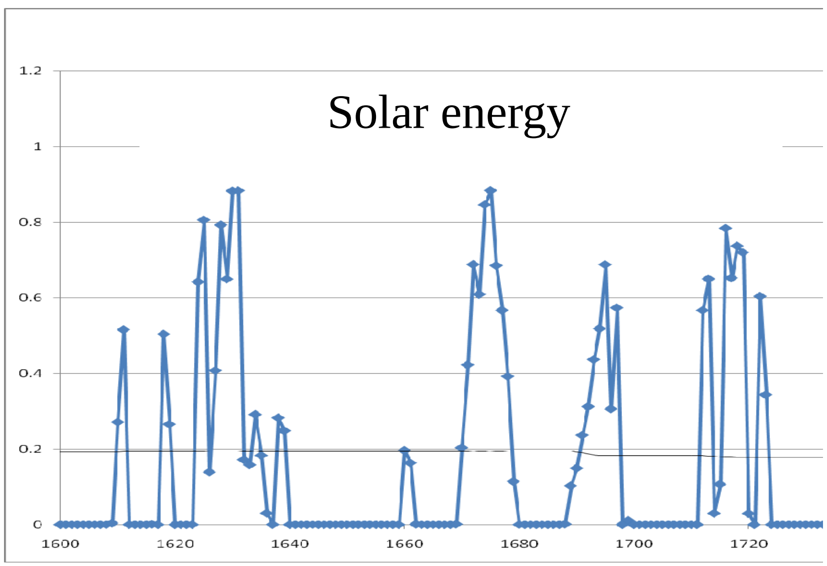
Goals for an energy policy model

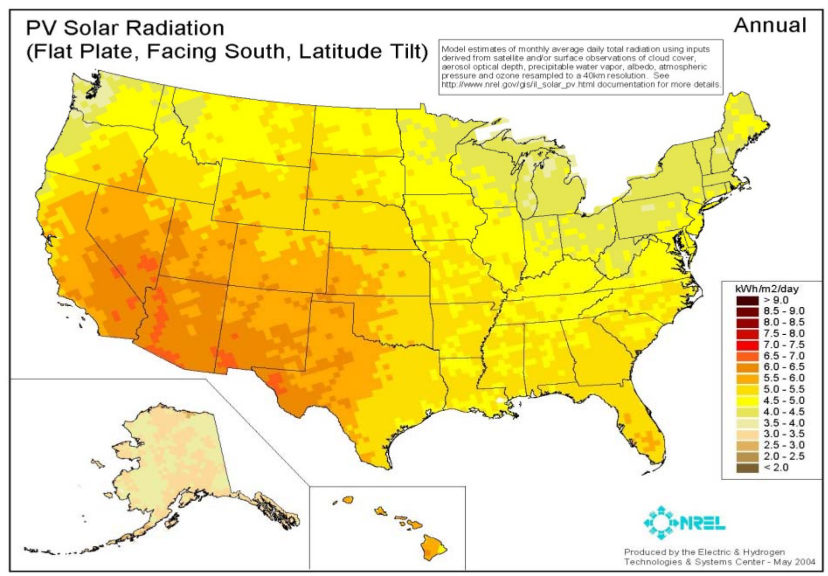
- Some challenges:
  - » The marginal value of wind and solar farms depends on the ability to work with intermittent supply.
  - » The impact of intermittent supply will be mitigated by the use of storage.
  - » Different storage technologies (batteries, flywheels, compressed air, pumped hydro) are each designed to serve different types of variations in supply and demand.
  - » The need for storage (and the value of wind and solar) depends on the entire portfolio of energy producing technologies.

Intermittent energy sources

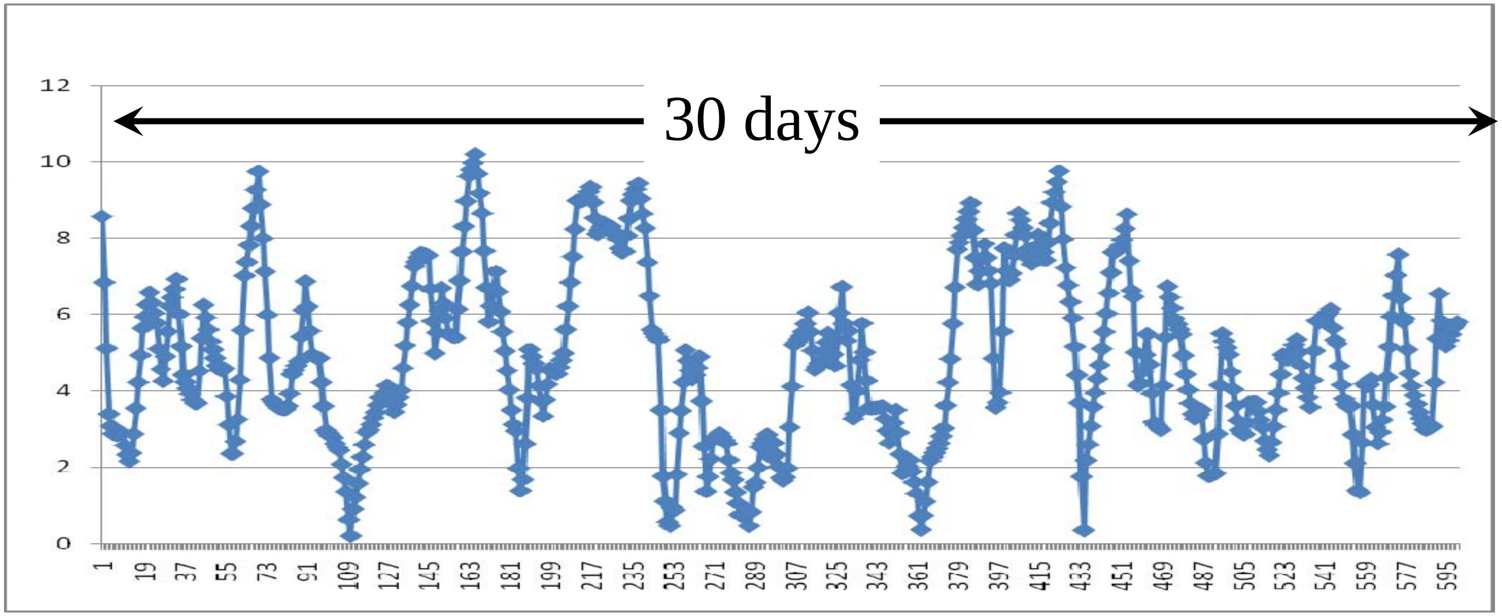




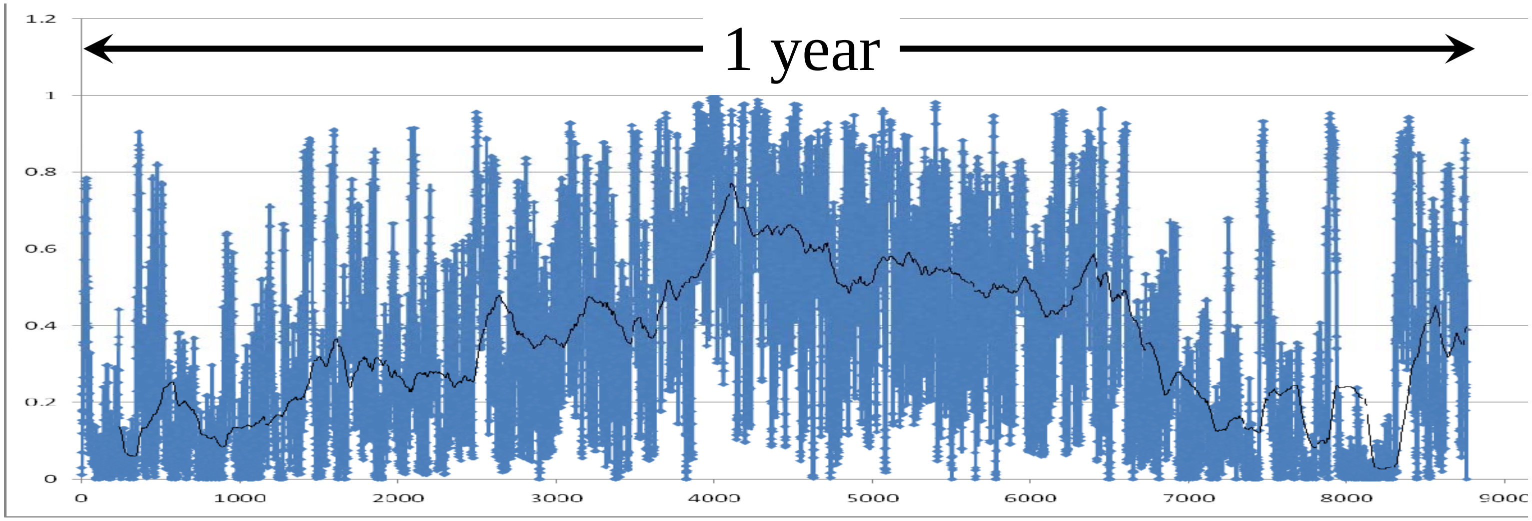




Wind

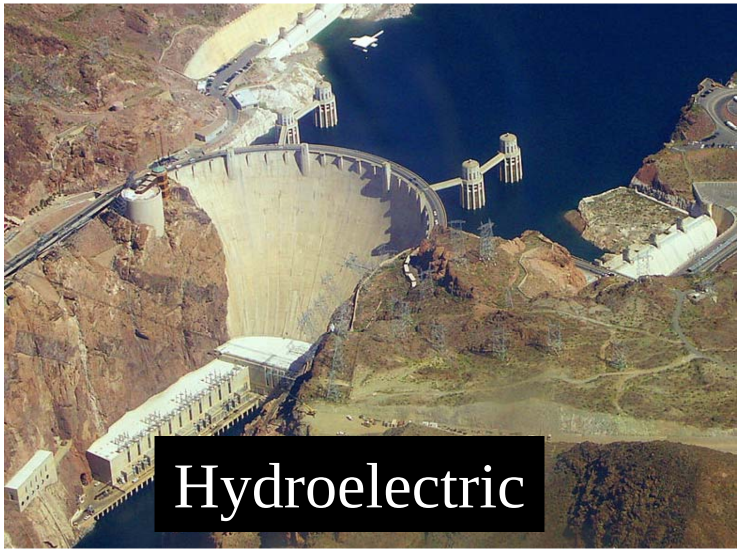


30 days
12
10
8
6
4
2
0
1
19
37
55
73
91
109
127
145
163
181
199
217
235
253
271
289
307
325
343
361
379
397
415
433
451
469
487
505
523
541
559
577
595

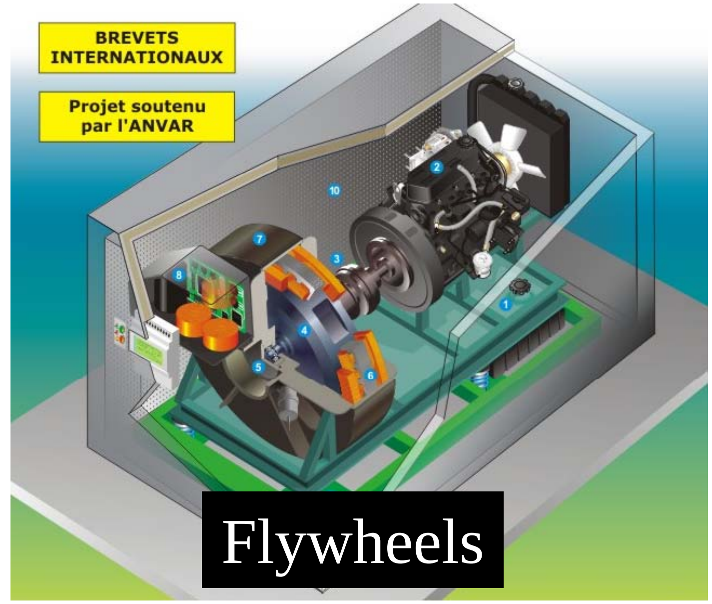


1 year
1.2
1
0.8
0.6
0.4
0.2
0
0
1000
2000
3000
4000
5000
6000
7000
8000

Storage

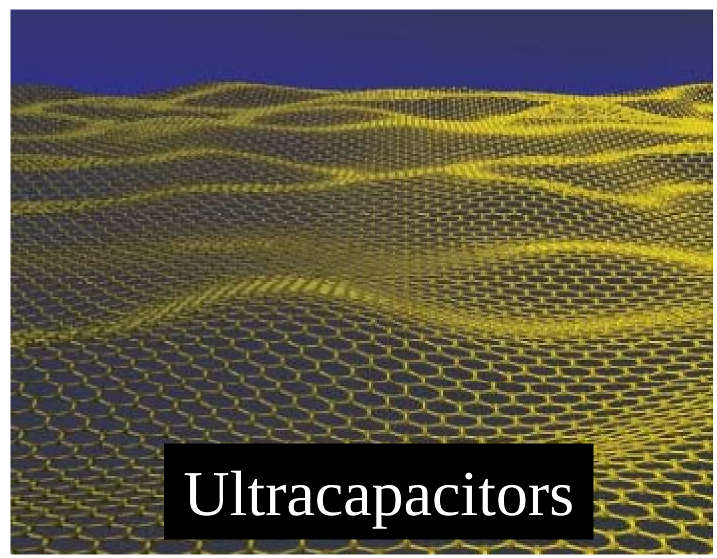


Hydroelectric



Flywheels

Batteries



Ultracapacitors

Long term uncertainties….

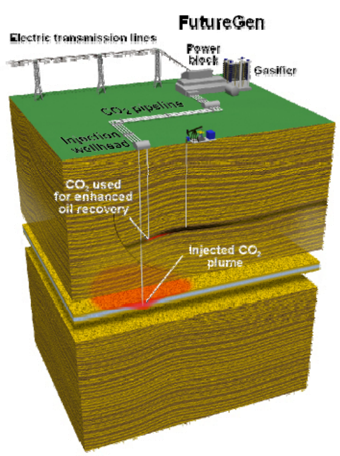
Tax policy

Price of oil

Batteries

2010 2015 2020 2025 2030

Solar panels



Carbon capture and sequestration

Climate change

Goals for an energy policy model

- Model capabilities we are looking for
  - » Multi-scale
    - Multiple time scales (hourly, daily, seasonal, annual, decade)
    - Multiple spatial scales
    - Multiple technologies (different coal-burning technologies, new wind turbines, …)
    - Multiple markets
      - – Transportation (commercial, commuter, home activities)
      - – Electricity use (heavy industrial, light industrial, business, residential)
      - – ….
  - » Stochastic (handles uncertainty)
    - Hourly fluctuations in wind, solar and demands
    - Daily variations in prices and rainfall
    - Seasonal changes in weather
    - Yearly changes in supplies, technologies and policies

Outline

- An optimization model for energy
- Introduction to approximate dynamic programming
- A blood management illustration
- "SMART" - Energy policy model

Energy resource modeling

System state: $S_t = (R_t, D_t, \rho_t)$

- R_t → Resource state
- D_t → Market demands
- ρ_t → "System parameters":
  - State of technology (costs, performance)
  - Climate, weather (temperature, rainfall, wind)
  - Government policies (tax rebates on solar panels)
  - Market prices (oil, coal)

Energy resource modeling

- The decision variables:




$$x_t^{cap} = \begin{pmatrix} \text{New capacity} \\ \text{Retired capacity} \\ \text{Storage capacity} \\ \textit{for each}\text{:} \\ \text{Type} \\ \text{Location} \\ \text{Technology} \end{pmatrix} \qquad x_t^{disp} = \begin{pmatrix} \text{Flow from:} \\ \quad \text{Resource to conversion} \\ \quad \text{Conversion to storage} \\ \quad \text{Storage to grid} \\ \quad \text{Conversion to grid} \\ \quad \text{Grid to intermediate uses} \\ \quad \text{Grid to final demand} \end{pmatrix}$$

Energy resource modeling

- Exogenous information:

$$W_t = \text{New information} = \left(\hat{R}_t, \hat{D}_t, \hat{\rho}_t\right)$$

$\hat{R}_t$ = Exogenous changes in capacity, reserves

$\hat{D}_t$ = New demands for energy by type

$\hat{\rho}_t$ = Exogenous changes in parameters.

Information can be:

Fine-grained:

Wind, solar, demand, prices, …

Coarse-grained:

Changes in technologies, policies, climate

Energy resource modeling

- The transition function

$$S_{t+1} = S^M(S_t, x_t, W_{t+1})$$

Also known as the "system model," "plant model" or just "model."

All the physics of the problem.

Introduction to ADP

- We can optimize decisions using Bellman's equation

$$V_t(S_t) = \max_{x \in \mathcal{X}} \left(C_t(S_t, x_t) + E\left\{ V_{t+1}(S_{t+1}) \mid S_t \right\} \right)$$

Introduction to ADP

- The computational challenge:

$$V_t(S_t) = \max_{x \in \mathcal{X}} \left(C_t(S_t, x_t) + E\left\{ V_{t+1}(S_{t+1}) \mid S_t \right\} \right)$$

How do we find $V_{t+1}(S_{t+1})$?

How do we compute the expectation?

How do we find the optimal solution?

Outline

- An optimization model for energy
- Introduction to approximate dynamic programming
- A blood management illustration
- “SMART” - Energy policy model

Introduction to ADP

- Classical ADP
  - » Most applications of ADP focus on the challenge of handling multidimensional state variables
  - » Start with

$$V_t(S_t) = \max_{x \in \mathcal{X}} \left(C_t(S_t, x_t) + E\left\{ V_{t+1}(S_{t+1}) \mid S_t \right\} \right)$$

  - » Now replace the value function with some sort of approximation

$$V_{t+1}(S_{t+1}) \approx \overline{V}_{t+1}(S_{t+1}) = \sum_{f \in \mathcal{F}} \theta_f \phi_f(S_{t+1})$$

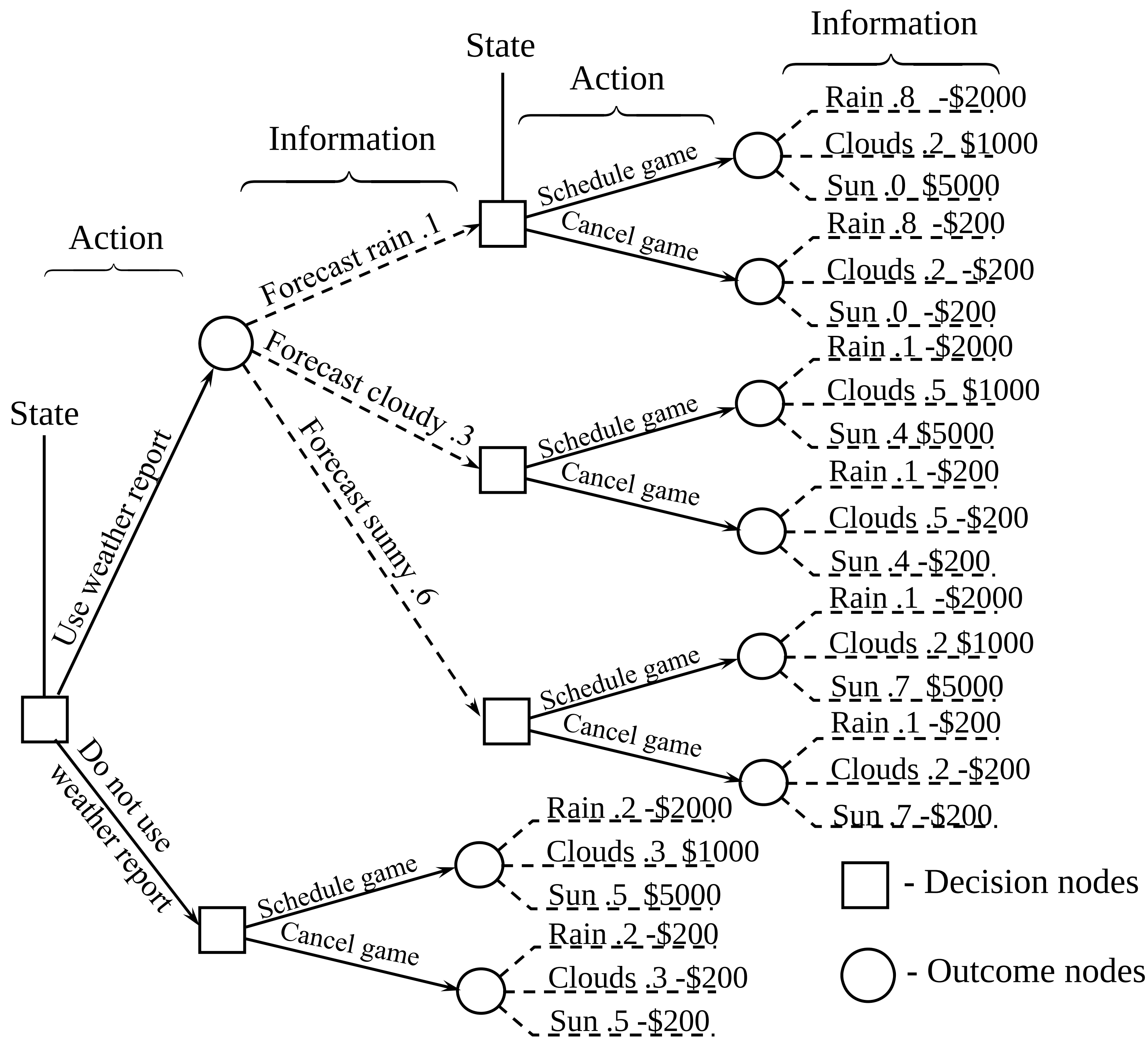
  - » May draw from the entire field of statistics/machine learning.

Introduction to ADP

- But this does not solve our problem
  - » Assume we have an approximate value function.
  - » We still have to solve a problem that looks like

$$V_t(S_t) = \max_{x \in \mathcal{X}} \left(C_t(S_t, x_t) + E \sum_{f \in \mathcal{F}} \theta_f \phi_f (S_{t+1}) \right)$$

  - » This means we still have to deal with a maximization problem (might be a linear, nonlinear or integer program) with an expectation. 


State
Action
Information
State
Action
Information
Use weather report
Do not use weather report
Forecast rain .1
Forecast cloudy .3
Forecast sunny .6
Schedule game
Cancel game
Rain .8 -$2000
Clouds .2 $1000
Sun .0 $5000
Rain .8 -$200
Clouds .2 -$200
Sun .0 -$200
Schedule game
Cancel game
Rain .1 -$2000
Clouds .5 $1000
Sun .4 $5000
Rain .1 -$200
Clouds .5 -$200
Sun .4 -$200
Schedule game
Cancel game
Rain .1 -$2000
Clouds .2 $1000
Sun .7 $5000
Rain .1 -$200
Clouds .2 -$200
Sun .7 -$200
Schedule game
Cancel game
Rain .2 -$2000
Clouds .3 $1000
Sun .5 $5000
Rain .2 -$200
Clouds .3 -$200
Sun .5 -$200
- Decision nodes
- Outcome nodes

The post-decision state

- New concept:
  - » The "pre-decision" state variable:
    - $S_t =$ The information required to make a decision x_t
    - Same as a "decision node" in a decision tree.
  - » The "post-decision" state variable:
    - $S_t^x =$ The state of what we know immediately after we make a decision.
    - Same as an "outcome node" in a decision tree.

The post-decision state

- An inventory problem:
  - » Our basic inventory equation:

$$R_{t+1} = \max\{0; R_t + x_t - \hat{D}_{t+1}\}$$

where

$$R_t = \text{Inventory at time } t.$$

$$x_t = \text{Amount we ordered.}$$

$$\hat{D}_{t+1} = \text{Demand in next time period.}$$

  - » Using pre- and post-decision states:

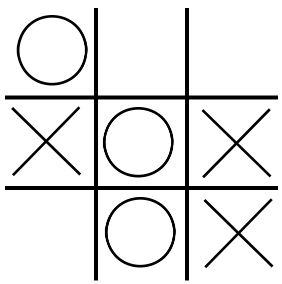
$$R_t^x = R_t + x_t \qquad \text{Post-decision state}$$

$$R_{t+1} = \max\{0; R_t^x - \hat{D}_{t+1}\} \qquad \text{Pre-decision state}$$

The post-decision state

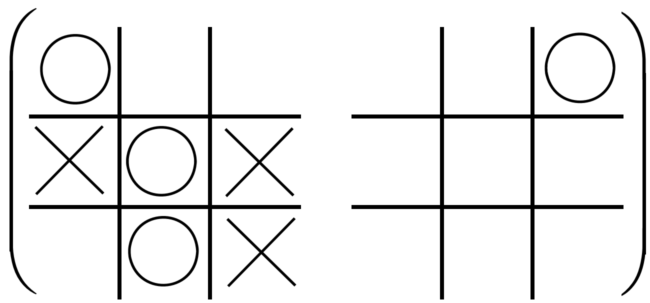
- Pre-decision, state-action, and post-decision

Pre-decision state



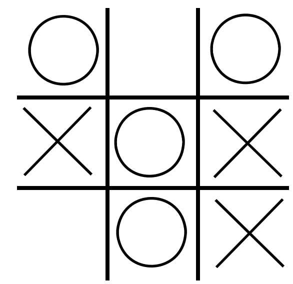
3^9 states

State Action



$3^9 \times 9$ state-action pairs

Post-decision state

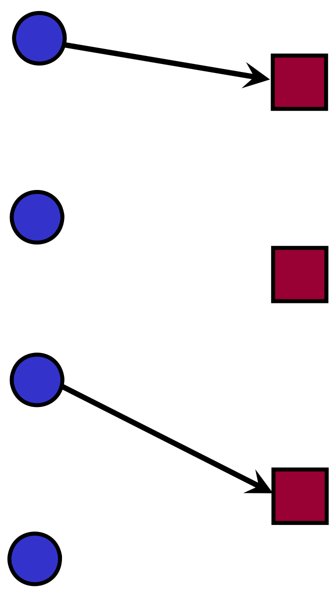


3^9 states

The post-decision state

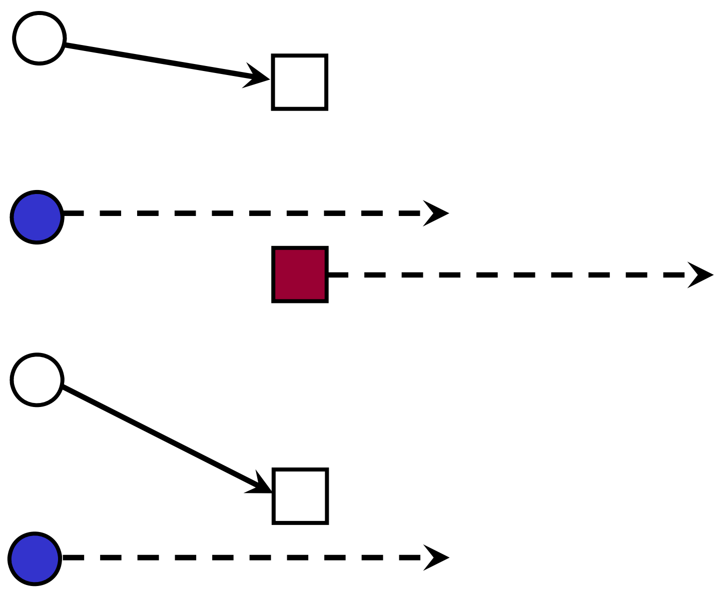
- Pre-decision: resources and demands

$$S_t = (R_t, D_t)$$

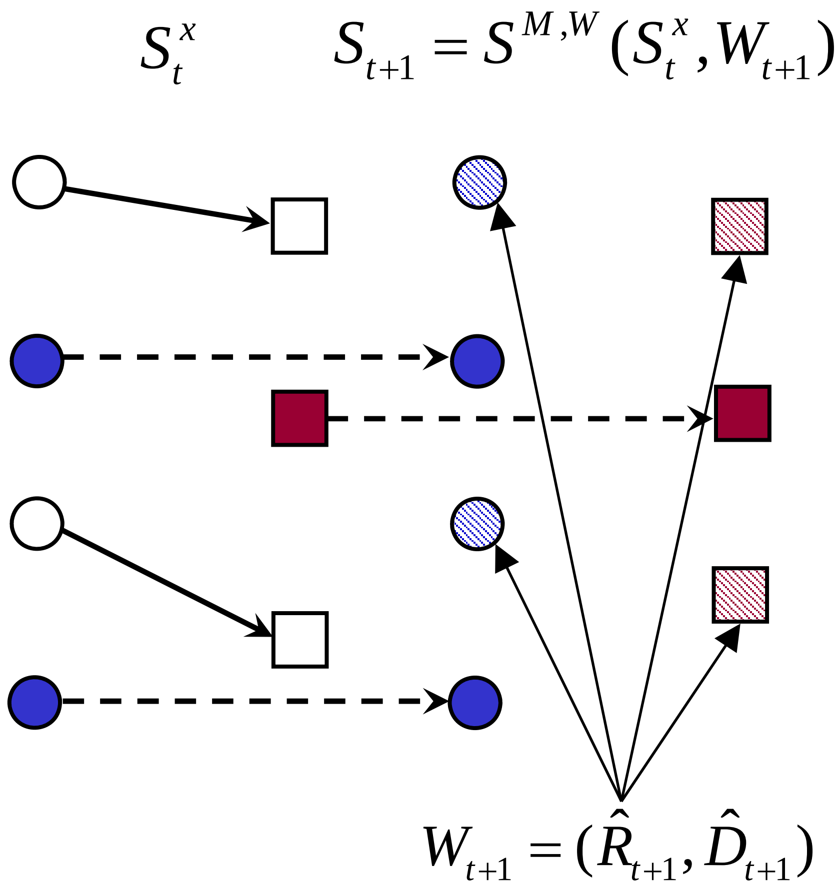


The post-decision state

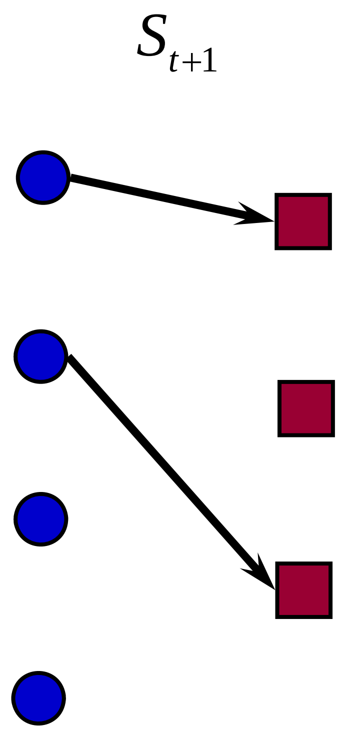
$$S_t^x = S^{M,x}(S_t, x_t)$$



The post-decision state



The post-decision state



The post-decision state

- Classical form of Bellman's equation:

$$V_t(S_t) = \max_{x \in \mathcal{X}} \left(C_t(S_t, x_t) + E\left\{ V_{t+1}(S_{t+1}) \mid S_t \right\} \right)$$

- Bellman's equations around pre- and post-decision states:
  - » Optimization problem (making the decision):

$$V_t(S_t) = \max_x \left(C_t(S_t, x_t) + V_t^x\left(S_t^{M,x}(S_t, x_t) \right) \right)$$

    - Note: this problem is deterministic!
  - » Expectation problem (incorporating uncertainty):

$$V_t^x(S_t^x) = E\left\{ V_{t+1}(S^{M,W}(S_t^x, W_{t+1})) \mid S_t^x \right\}$$

Introduction to ADP

- We first use the value function around the post-decision state variable, removing the expectation:

$$V_t(S_t) = \max_{x \in \mathcal{X}} \left(C_t(S_t, x_t) + V_t^x \left(S_t^x(S_t, x_t) \right) \right)$$

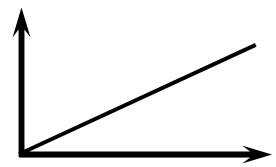
- We then replace the value function with an approximation that we estimate using machine learning techniques:

$$V_t(S_t) = \max_{x \in \mathcal{X}} \left(C_t(S_t, x_t) + \overline{V}_t \left(S_t^x(S_t, x_t) \right) \right)$$

The post-decision state

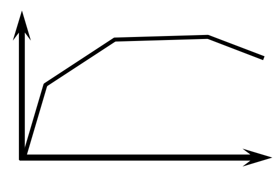
- Value function approximations:
  - » Linear (in the resource state):

$$\bar{V}_t(R_t^x) = \sum_{a \in A} \bar{v}_{ta} \cdot R_{ta}^x$$



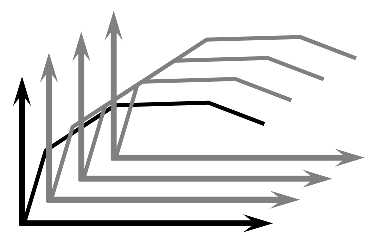
  - » Piecewise linear, separable:

$$\bar{V}_t(R_t^x) = \sum_{a \in A} \bar{V}_{ta}(R_{ta}^x)$$



  - » Indexed PWL separable:

$$\bar{V}_t(R_t^x) = \sum_{a \in A} \bar{V}_{ta}\left(R_{ta}^x \mid (features_t)\right)$$

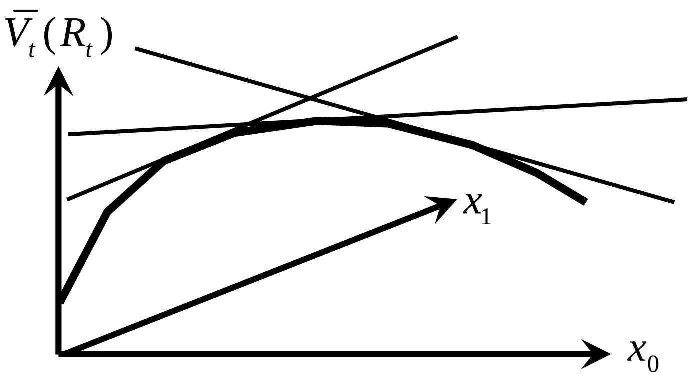


The post-decision state

- Value function approximations:
  - » Ridge regression (Klabjan and Adelman)

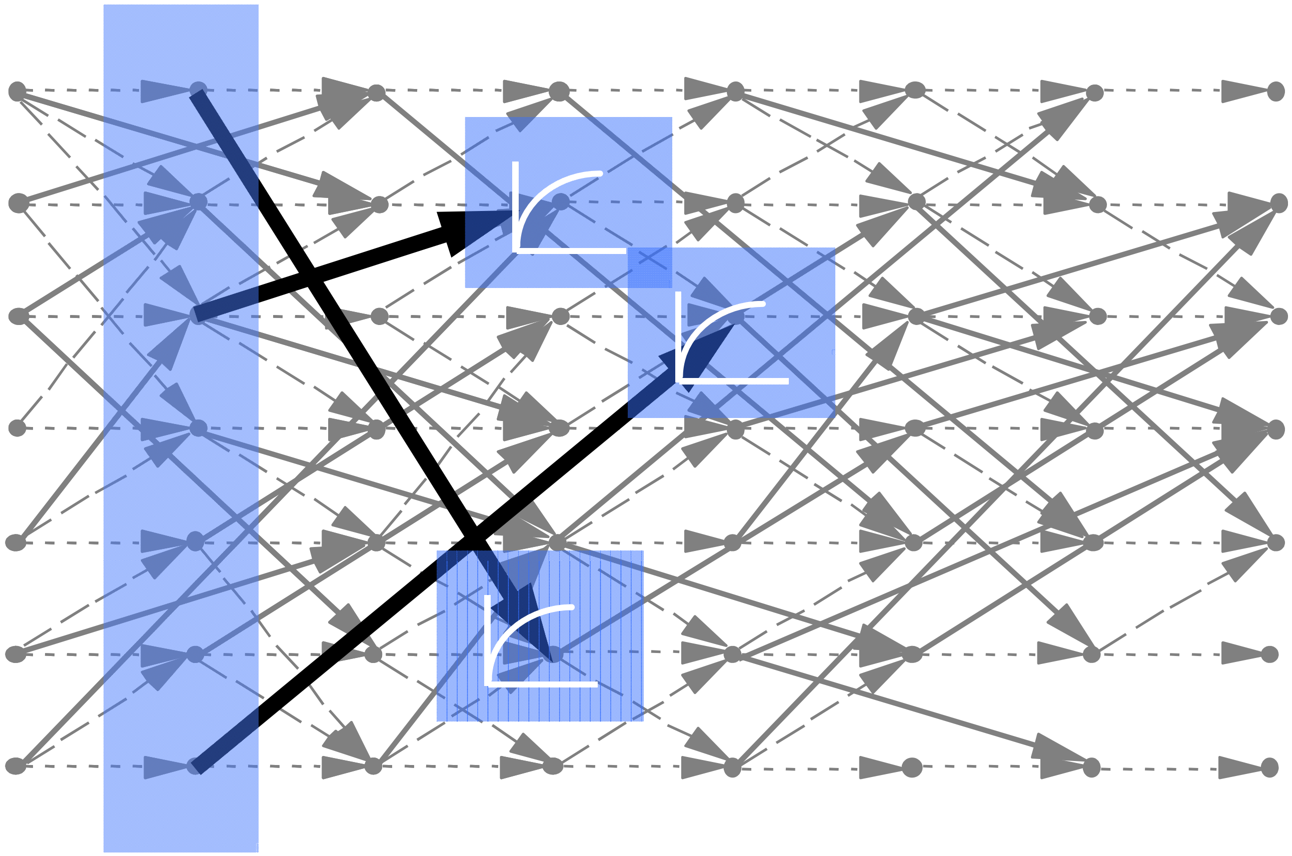
$$\bar{V}_t(R_t^x) = \sum_{f \in \mathcal{F}} \bar{V}_{tf}\left(\bar{R}_{tf}\right) \qquad \bar{R}_{tf} = \sum_{a \in \mathcal{A}_f} \theta_{fa} R_{ta}$$

  - » Benders cuts



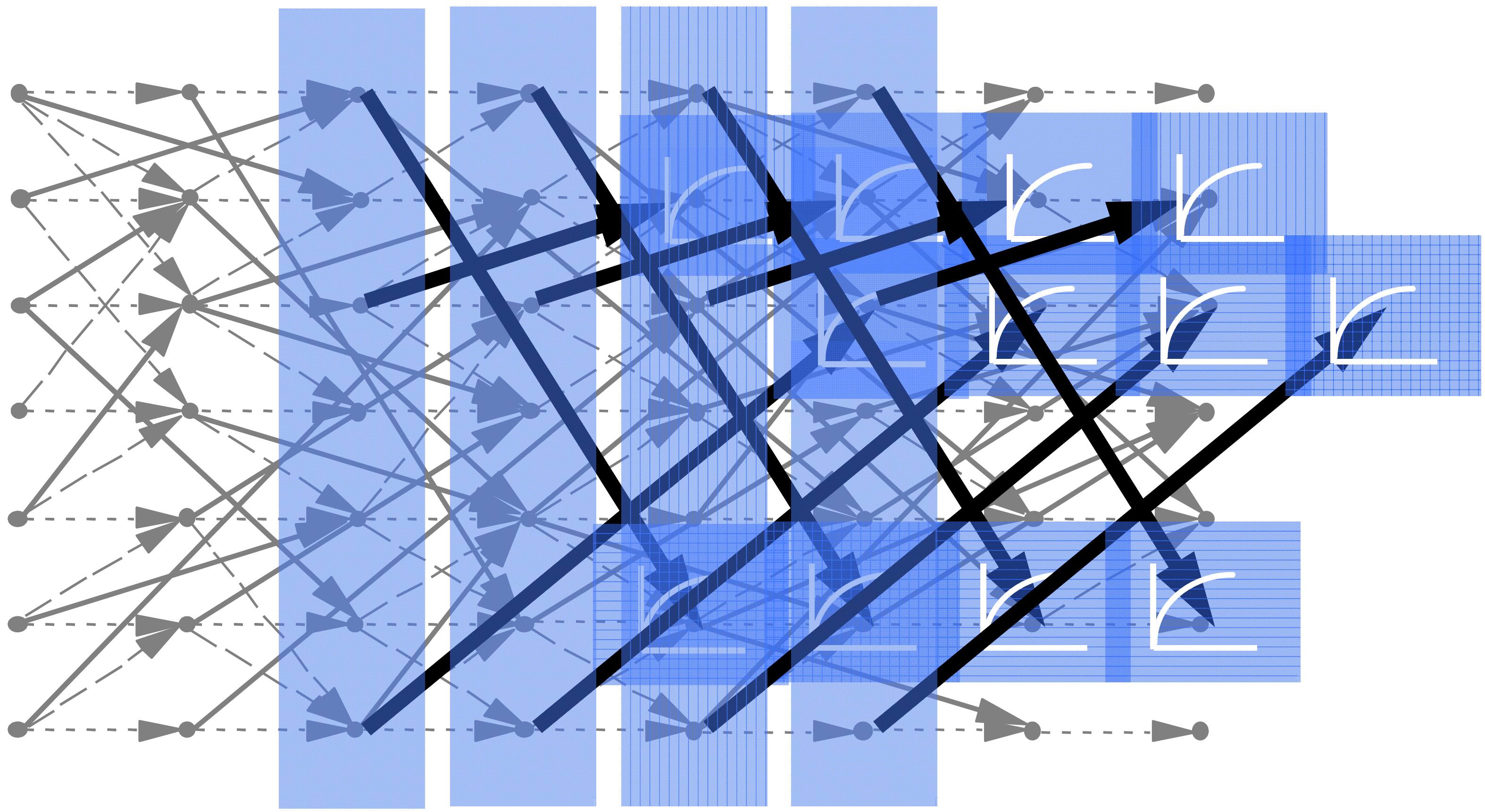
Making decisions

- Following an ADP policy



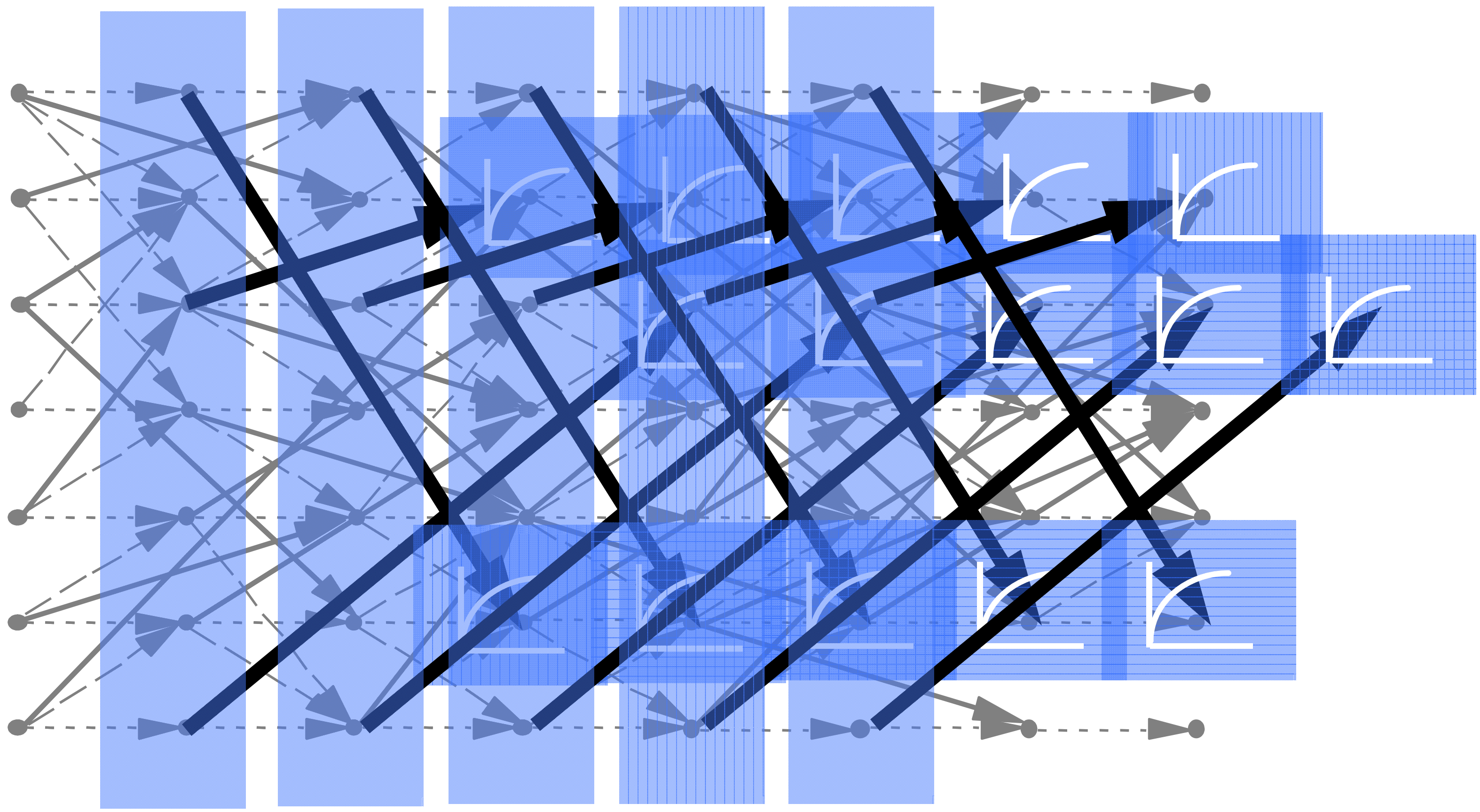
Making decisions

- Following an ADP policy



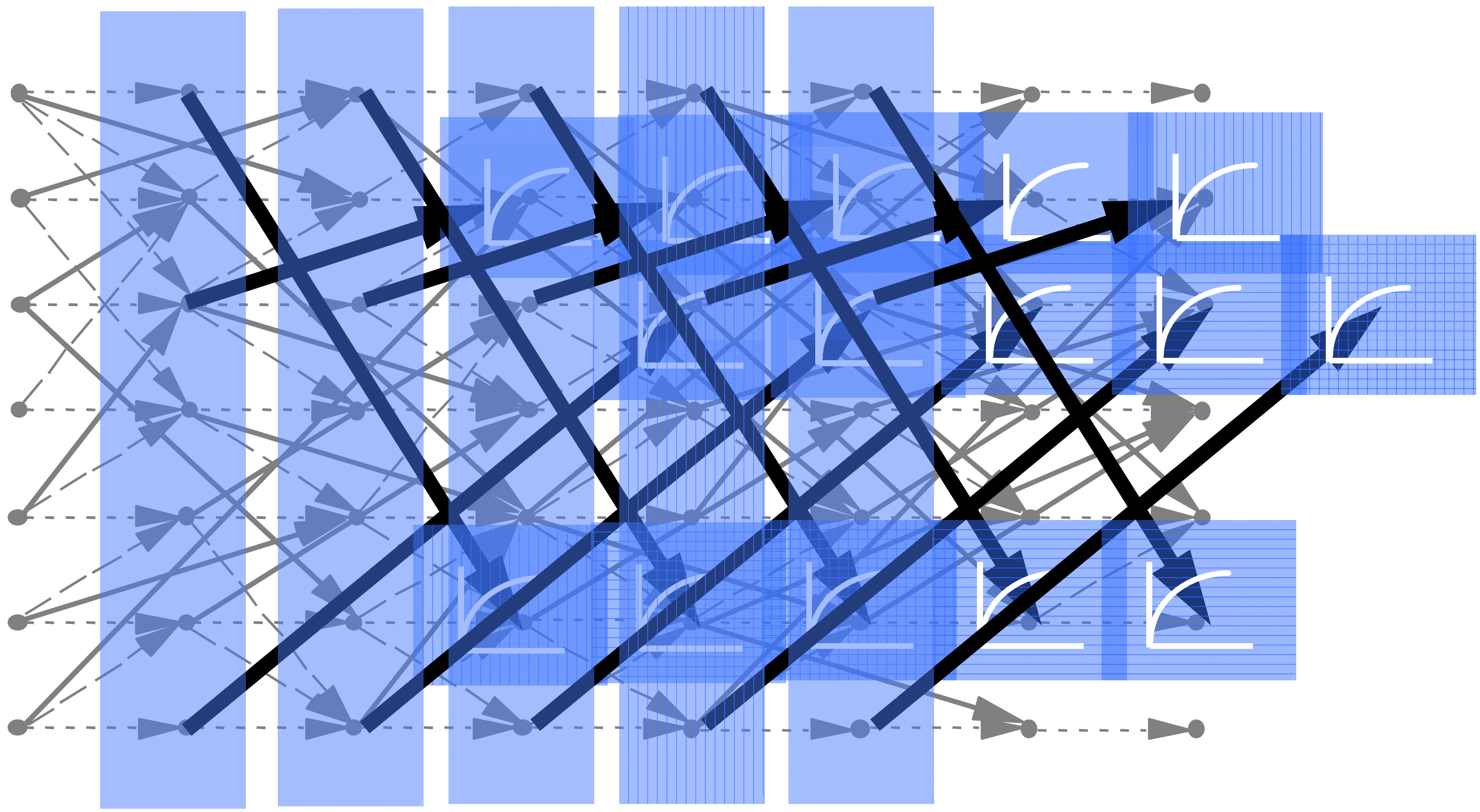
Making decisions

- Following an ADP policy



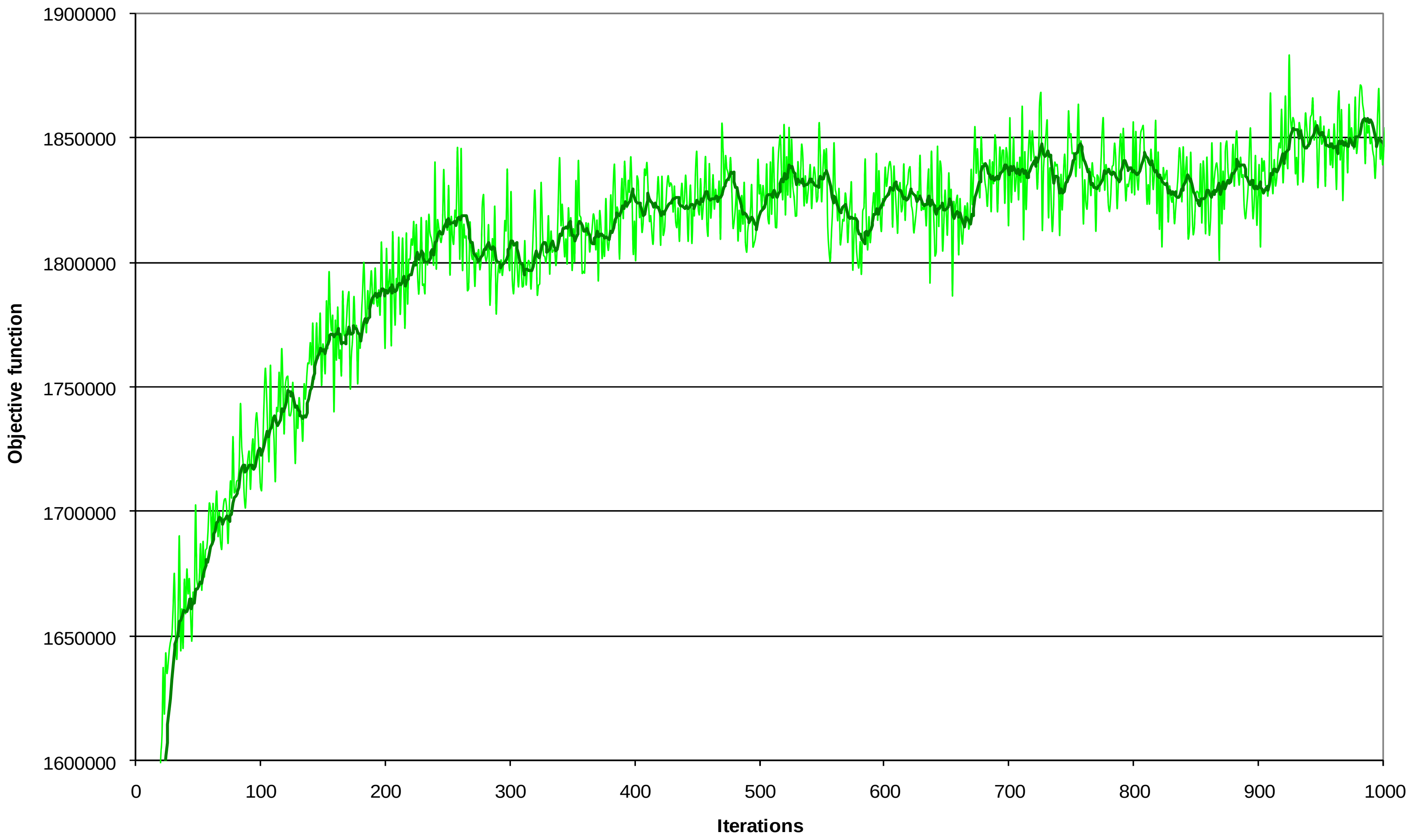
Making decisions

- Following an ADP policy



Approximate dynamic programming

- With luck, the objective function will improve steadily



The post-decision state

- Comparison to other methods:
  - » Classical MDP (value iteration)

$$V^n(S) = \max_x \left(C(S,x) + \gamma E V^{n-1}(S^M(S,x,W)) \right)$$

Approximate dynamic programming

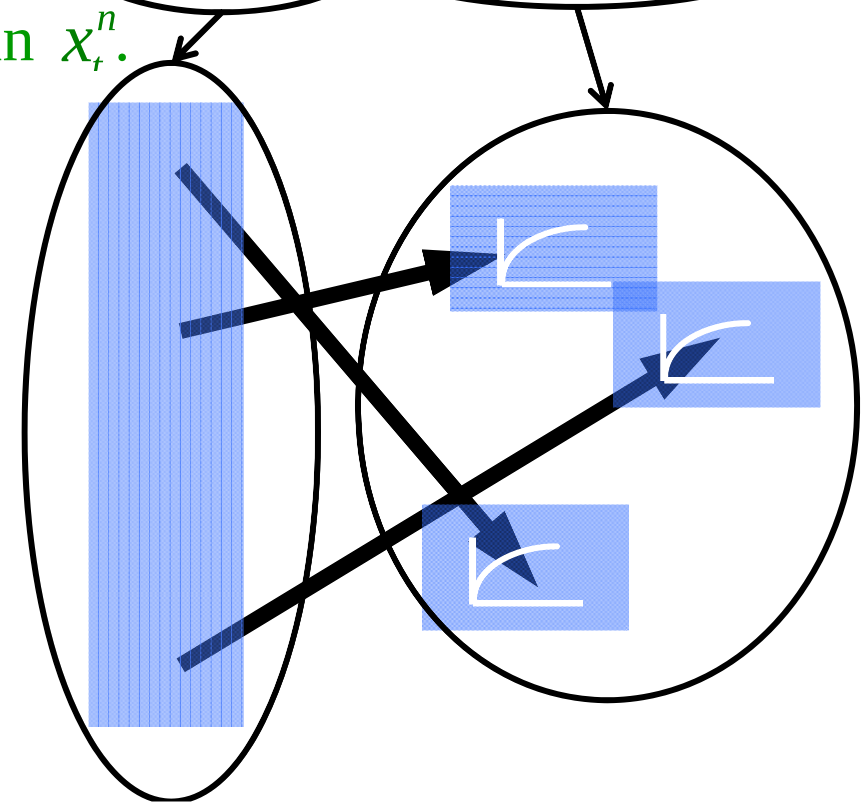
Step 1: Start with a pre-decision state S_t^n

Step 2: Solve the deterministic optimization using an approximate value function:

$$\hat{v}_t^n = \max_x \left(C_t(S_t^n, x_t) + \bar{V}_t^{n-1}(S^{M,x}(S_t^n, x_t)) \right)$$

to obtain x_t^n.

Deterministic optimization



Approximate dynamic programming

Step 1: Start with a pre-decision state S_t^n

Step 2: Solve the deterministic optimization using an approximate value function:

$$\hat{v}_t^n = \max_x \left(C_t(S_t^n, x_t) + \bar{V}_t^{n-1}(S^{M,x}(S_t^n, x_t)) \right)$$

to obtain x_t^n.

Deterministic optimization

Step 3: Update the value function approximation

$$\bar{V}_{t-1}^n(S_{t-1}^{x,n}) = (1-\alpha_{n-1})\bar{V}_{t-1}^{n-1}(S_{t-1}^{x,n}) + \alpha_{n-1}\hat{v}_t^n$$

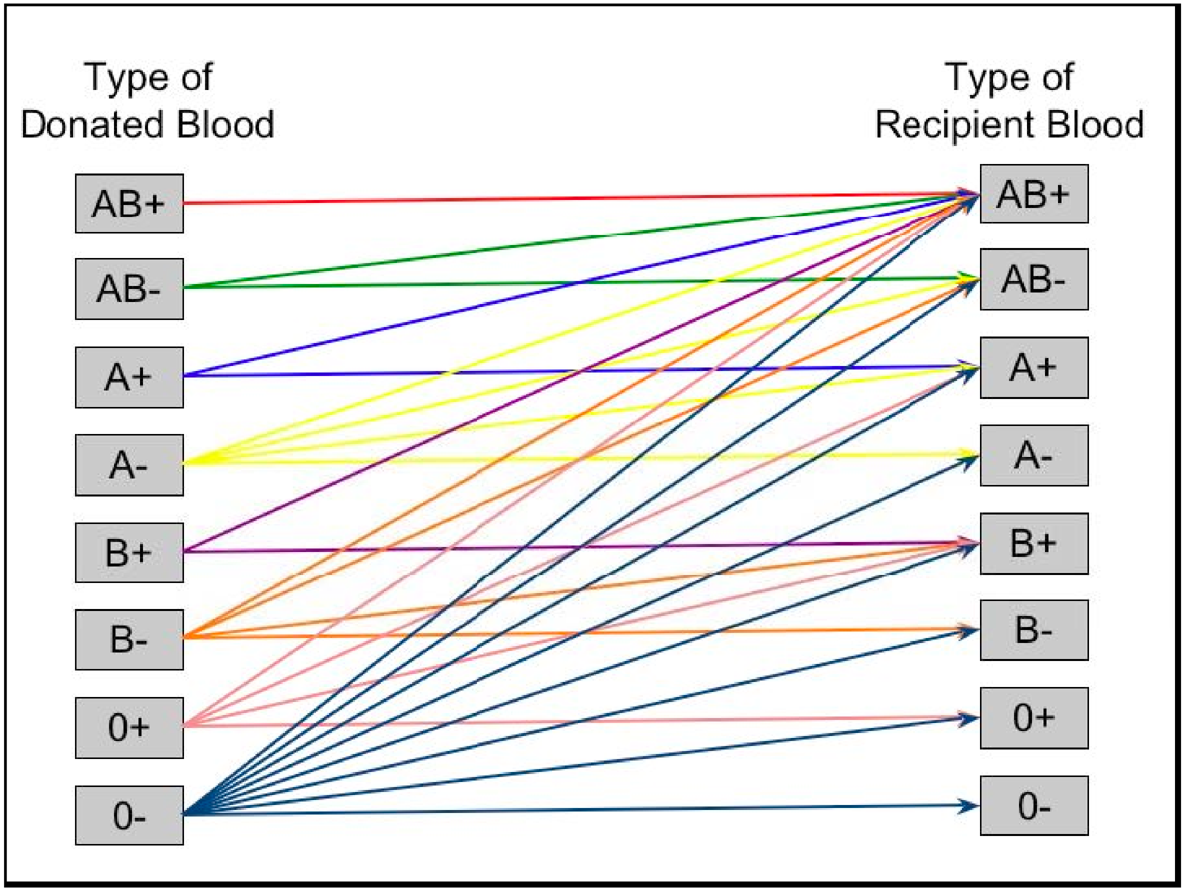
Recursive statistics

Outline

- An optimization model for energy
- Introduction to approximate dynamic programming
- A blood management illustration
- "SMART" - Energy policy model

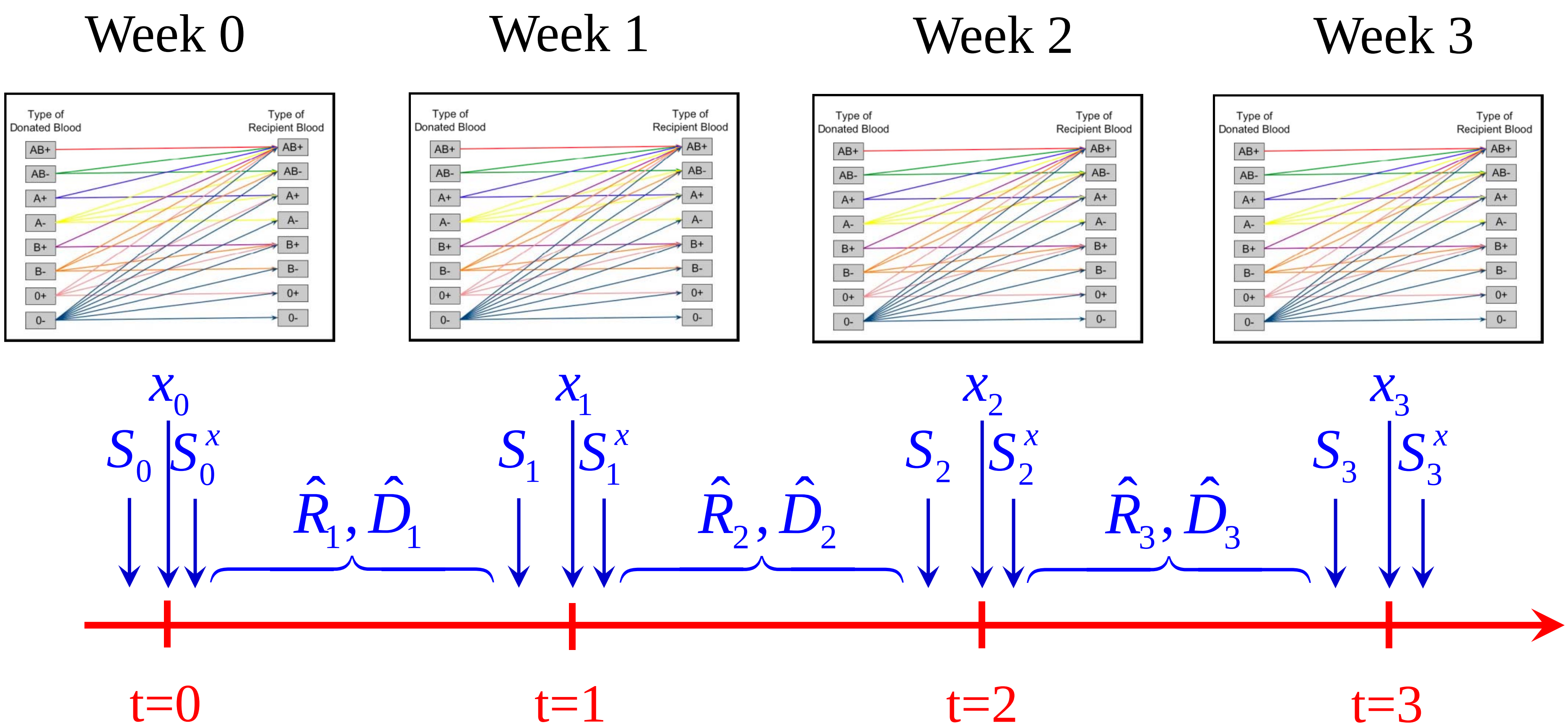
Blood management

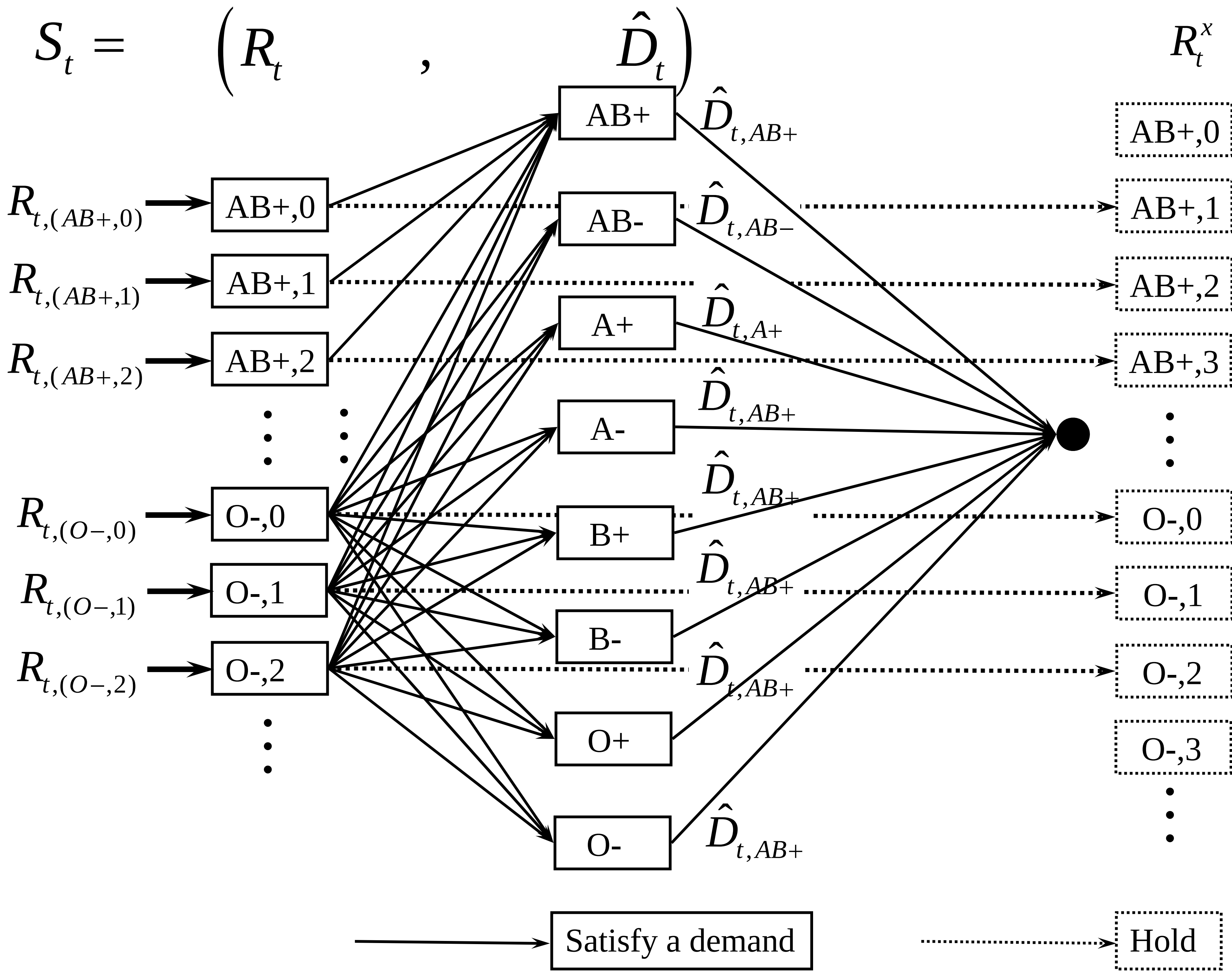
- Managing blood inventories



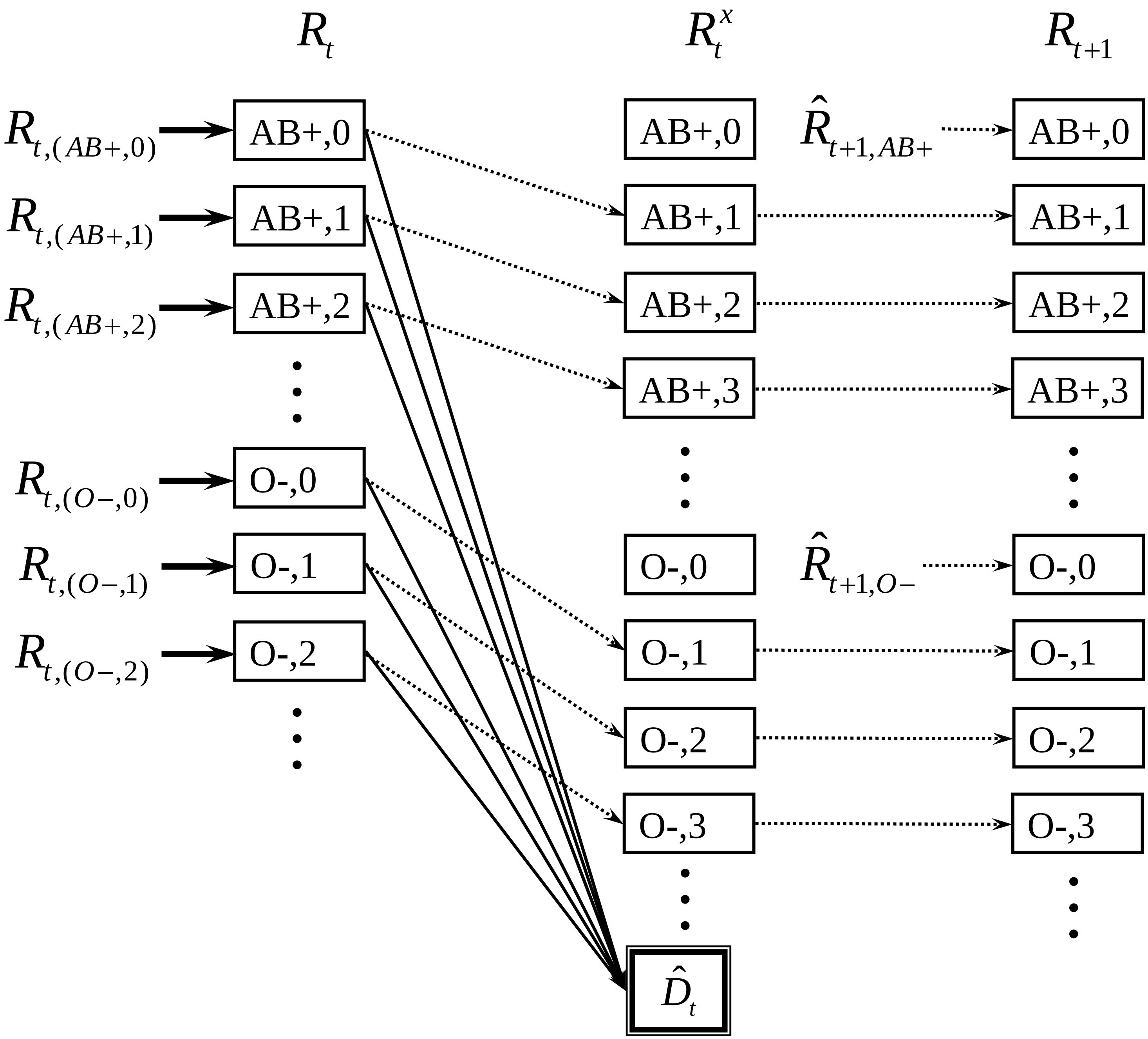
Blood management

- Managing blood inventories over time

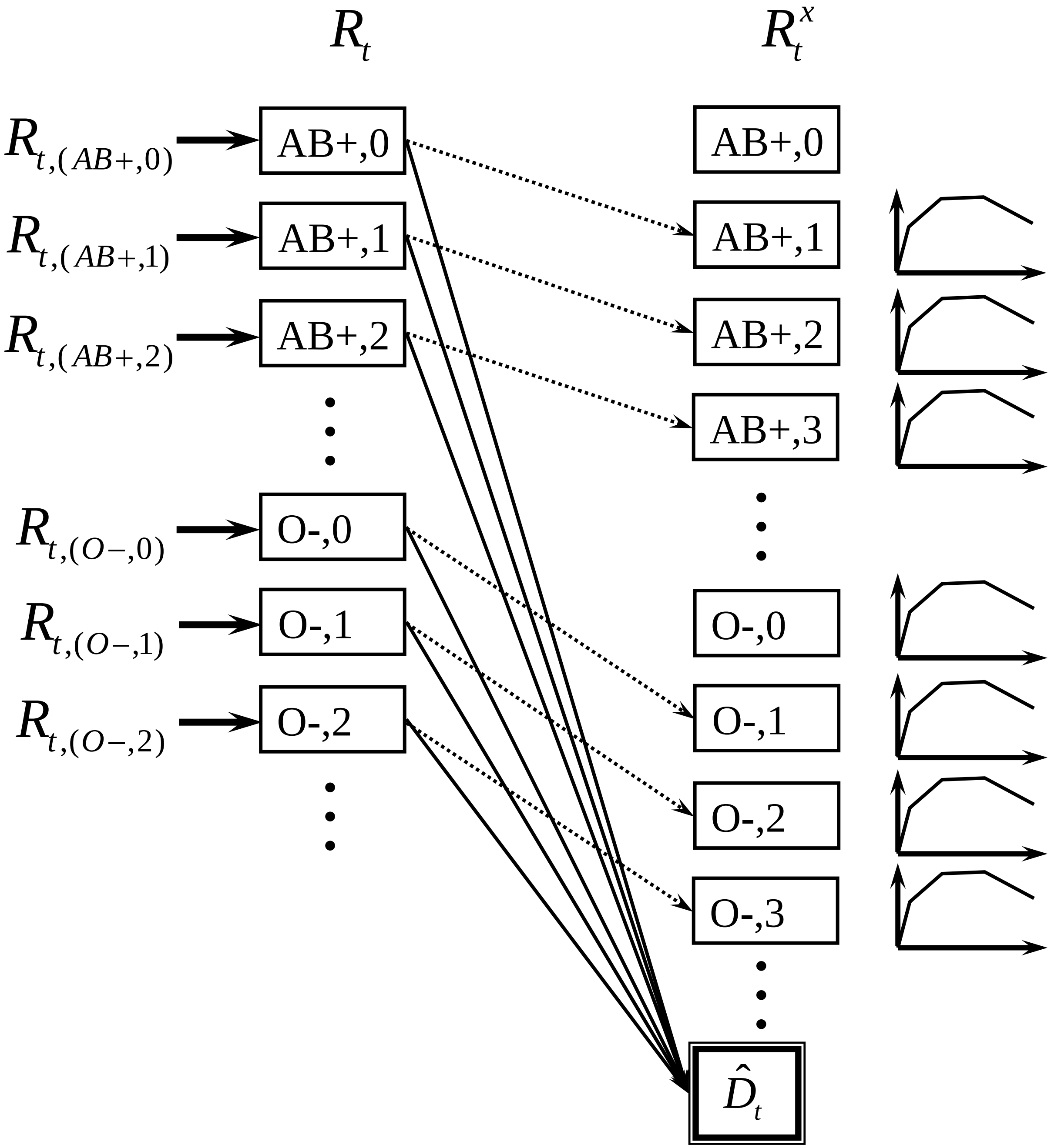




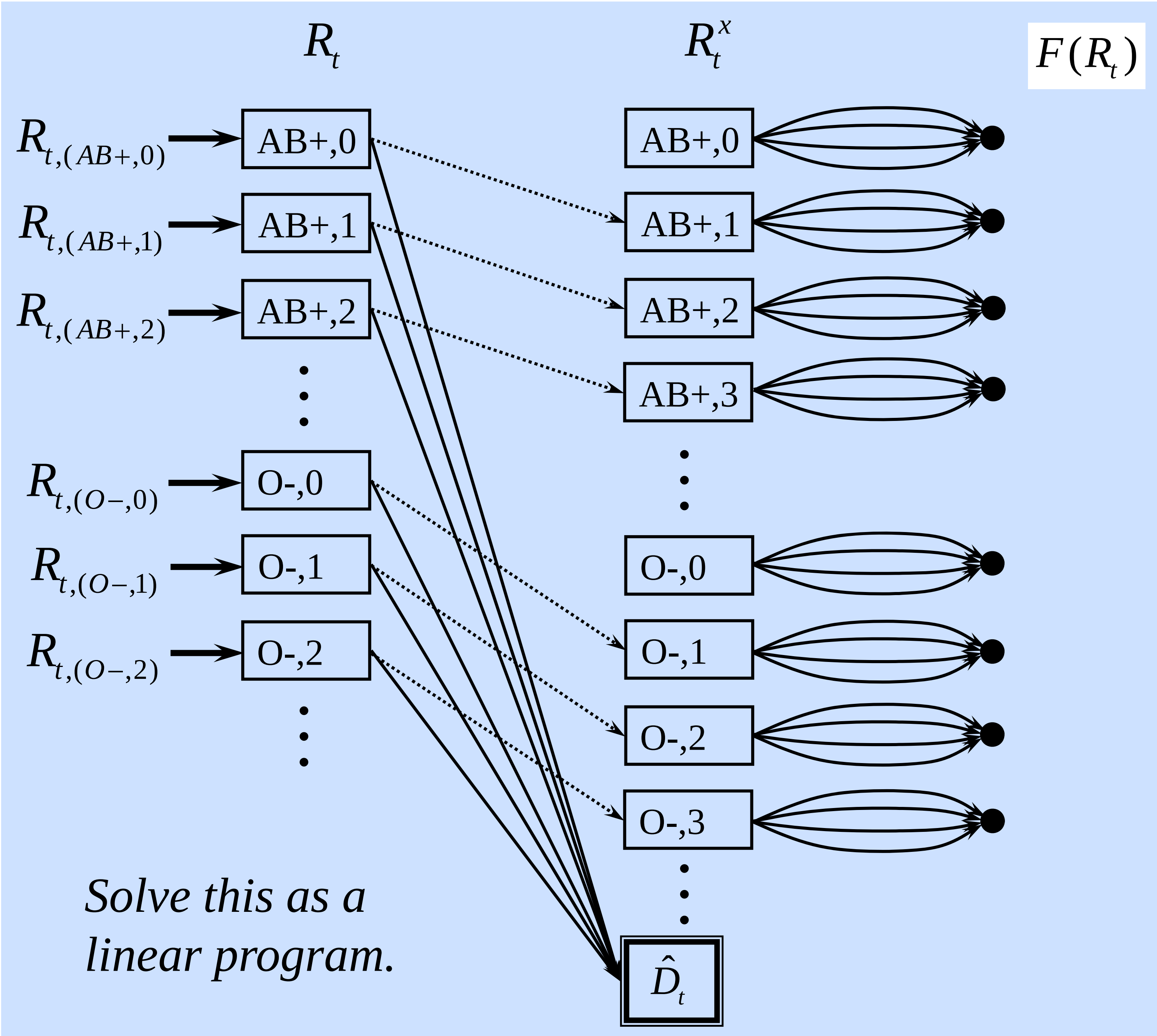
$S_t = (R_t , \hat{D}_t)$
R_t^x
$R_{t,(AB+,0)}$
$R_{t,(AB+,1)}$
$R_{t,(AB+,2)}$
$R_{t,(O-,0)}$
$R_{t,(O-,1)}$
$R_{t,(O-,2)}$
AB+,0
AB+,1
AB+,2
O-,0
O-,1
O-,2
AB+
AB-
A+
A-
B+
B-
O+
O-
$\hat{D}_{t,AB+}$
$\hat{D}_{t,AB-}$
$\hat{D}_{t,A+}$
$\hat{D}_{t,AB+}$
$\hat{D}_{t,AB+}$
$\hat{D}_{t,AB+}$
$\hat{D}_{t,AB+}$
$\hat{D}_{t,AB+}$
AB+,0
AB+,1
AB+,2
AB+,3
O-,0
O-,1
O-,2
O-,3
Satisfy a demand
Hold



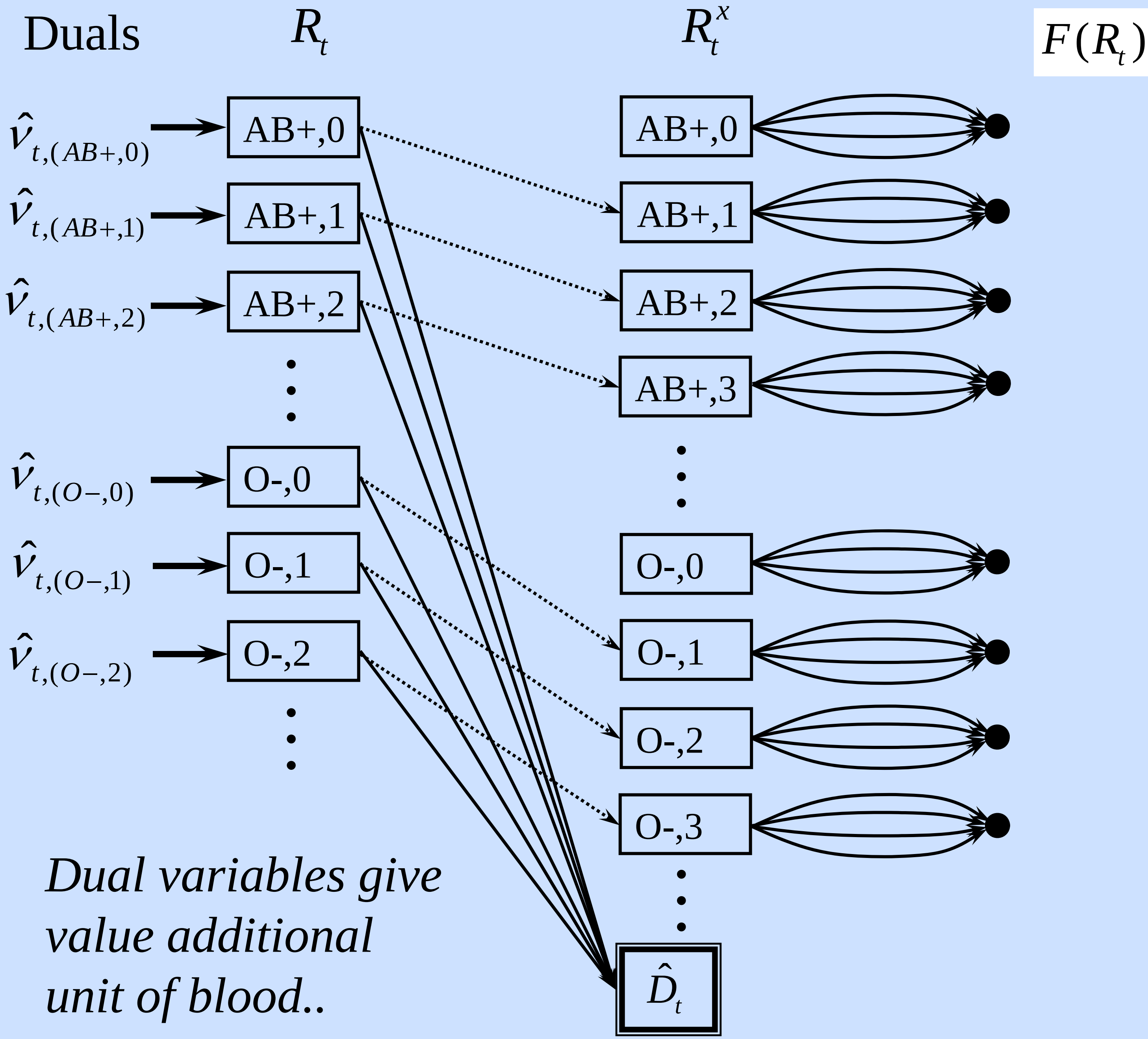
R_t
R_t^x
R_{t+1}
$R_{t,(AB+,0)}$
$R_{t,(AB+,1)}$
$R_{t,(AB+,2)}$
$R_{t,(O-,0)}$
$R_{t,(O-,1)}$
$R_{t,(O-,2)}$
AB+,0
AB+,1
AB+,2
AB+,3
O-,0
O-,1
O-,2
O-,3
$\hat{R}_{t+1,AB+}$
$\hat{R}_{t+1,O-}$
$\hat{D}_t$



R_t
R_t^x
$R_{t,(AB+,0)}$
$R_{t,(AB+,1)}$
$R_{t,(AB+,2)}$
$R_{t,(O-,0)}$
$R_{t,(O-,1)}$
$R_{t,(O-,2)}$
AB+,0
AB+,1
AB+,2
O-,0
O-,1
O-,2
AB+,0
AB+,1
AB+,2
AB+,3
O-,0
O-,1
O-,2
O-,3
$\hat{D}_t$



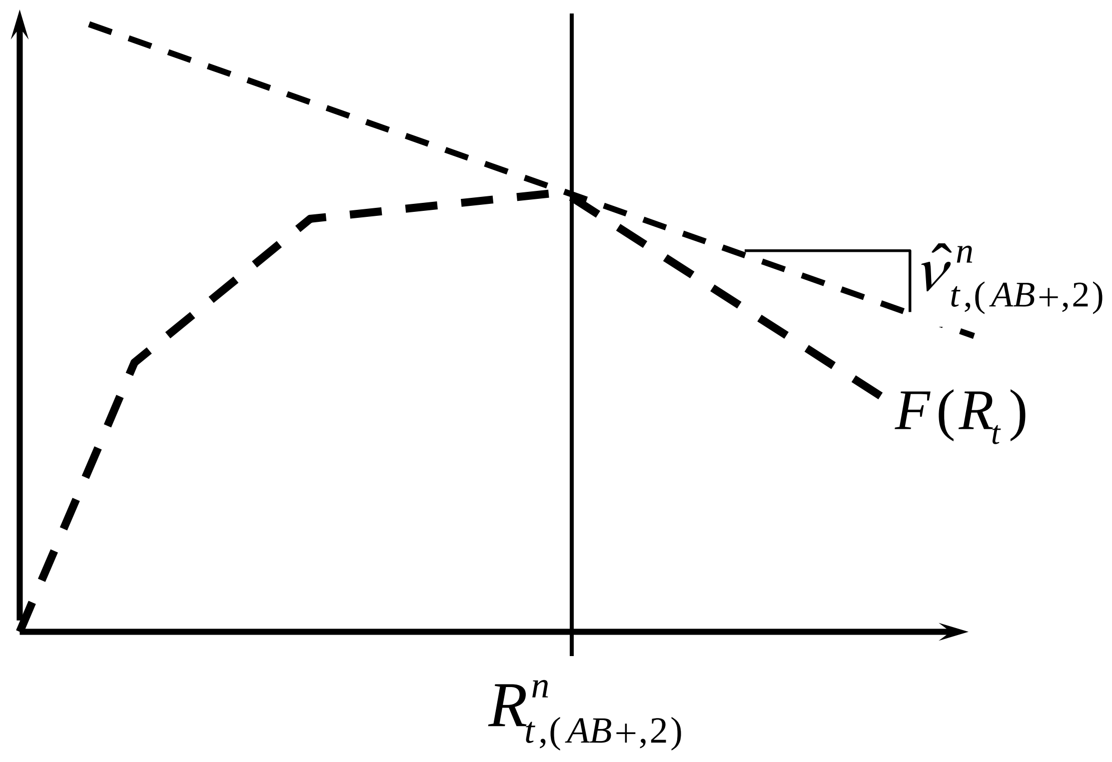
R_t
R_t^x
$F(R_t)$
$R_{t,(AB+,0)}$
$R_{t,(AB+,1)}$
$R_{t,(AB+,2)}$
$R_{t,(O-,0)}$
$R_{t,(O-,1)}$
$R_{t,(O-,2)}$
AB+,0
AB+,1
AB+,2
O-,0
O-,1
O-,2
AB+,0
AB+,1
AB+,2
AB+,3
O-,0
O-,1
O-,2
O-,3
$\hat{D}_t$
Solve this as a linear program.



Dual variables give value additional unit of blood..

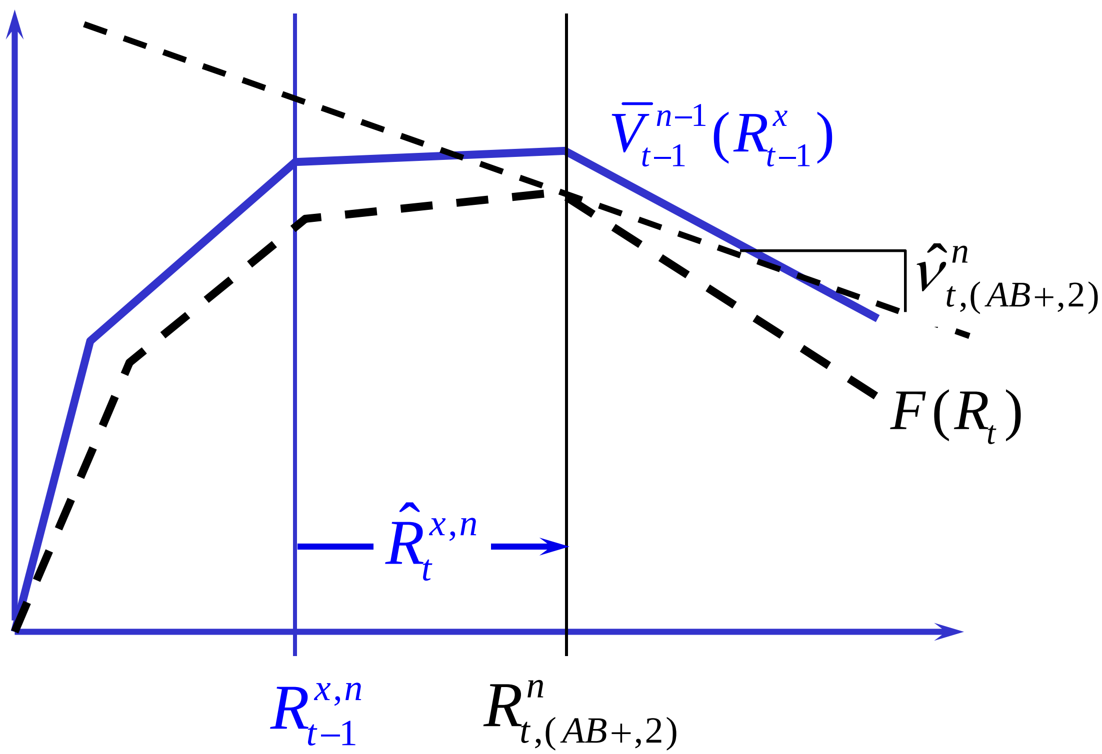
Updating the value function approximation

- Estimate the gradient at R_t^n



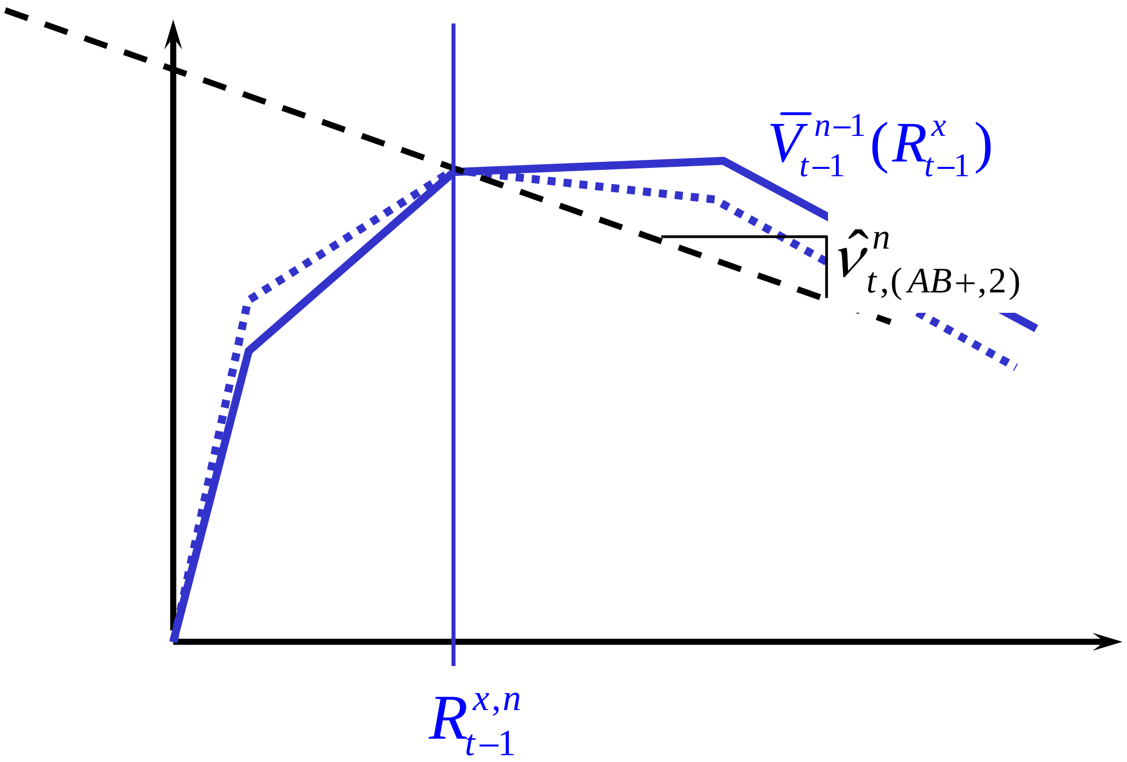
Updating the value function approximation

- Update the value function at $R_{t-1}^{x,n}$



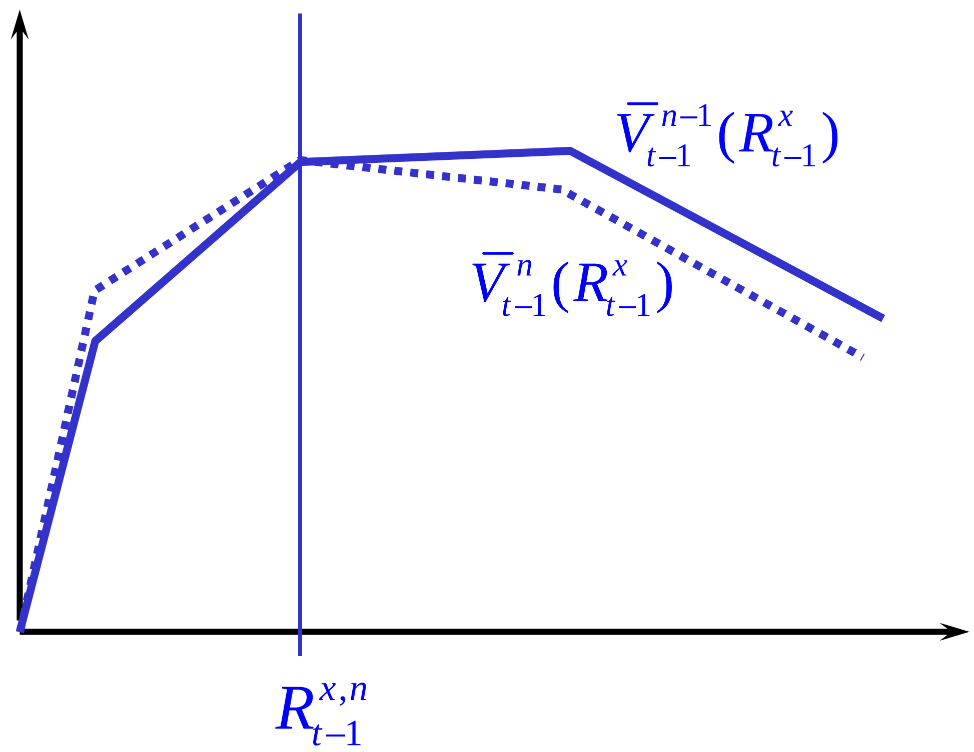
Updating the value function approximation

- Update the value function at $R_{t-1}^{x,n}$



Updating the value function approximation

- Update the value function at $R_{t-1}^{x,n}$



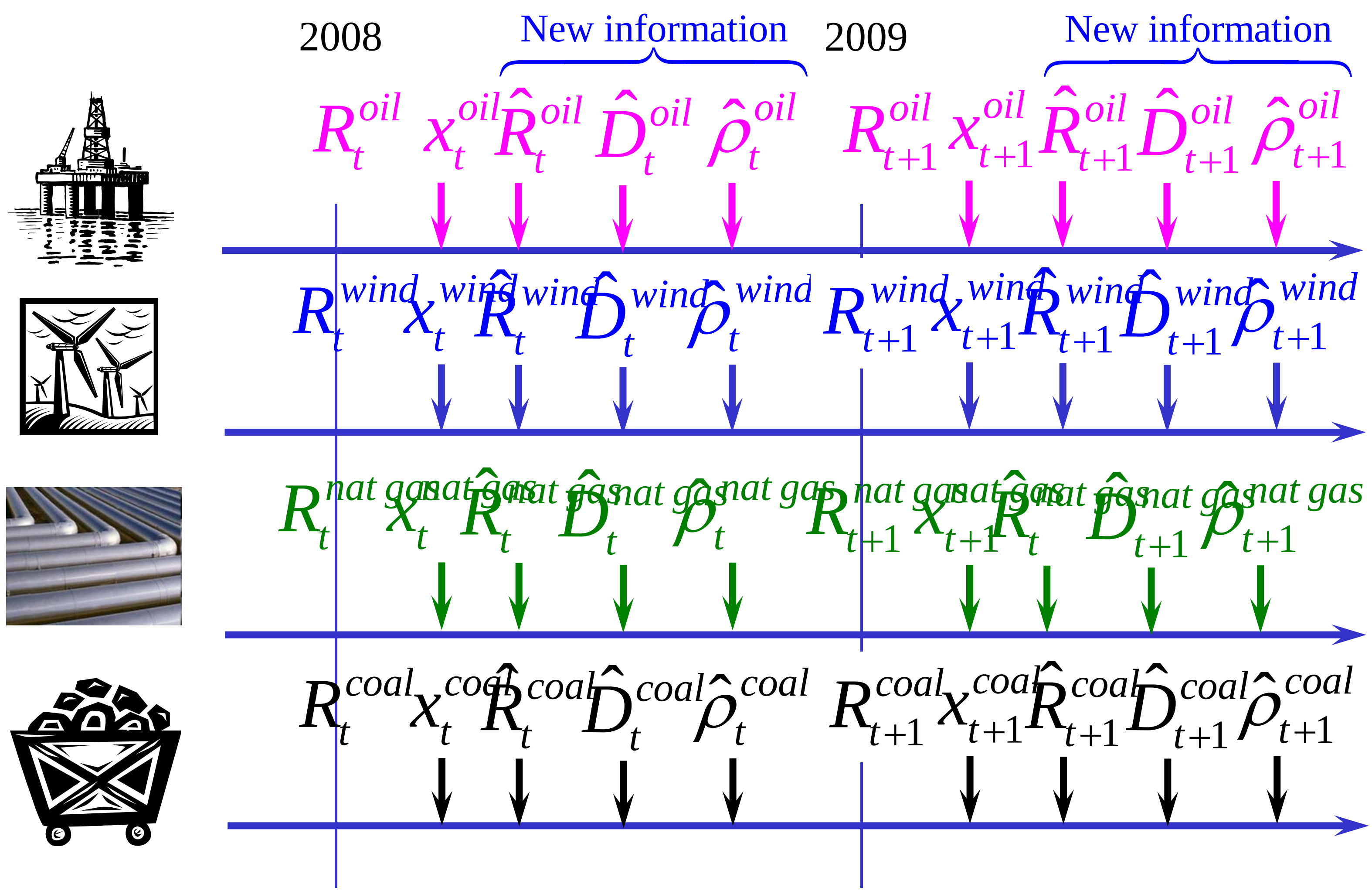
Outline

- An optimization model for energy
- Introduction to approximate dynamic programming
- A blood management illustration
- “SMART” - Energy policy model

SMART-Stochastic, multiscale model

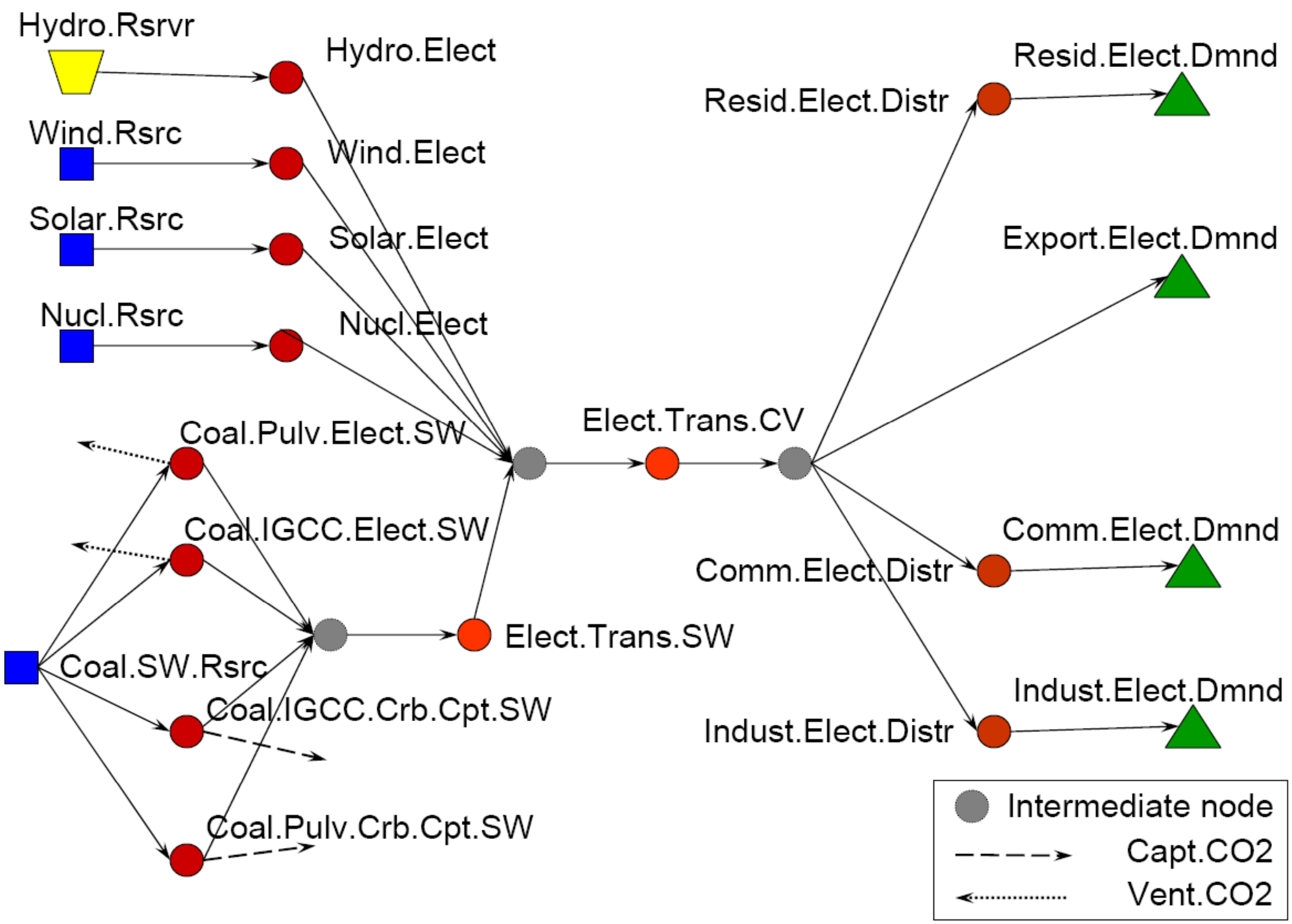
- SMART: A Stochastic, Multiscale Allocation model for energy Resources, Technology and policy
  - » Stochastic – able to handle different types of uncertainty:
    - Fine-grained – Daily fluctuations in wind, solar, demand, prices, …
    - Coarse-grained – Major climate variations, new government policies, technology breakthroughs
  - » Multiscale – able to handle different levels of detail:
    - Time scales – Hourly to yearly
    - Spatial scales – Aggregate to fine-grained disaggregate
    - Activities – Different types of demand patterns
  - » Decisions
    - Hourly dispatch decisions
    - Yearly investment decisions
    - Takes as input parameters characterizing government policies, performance of technologies, assumptions about climate

The annual investment problem



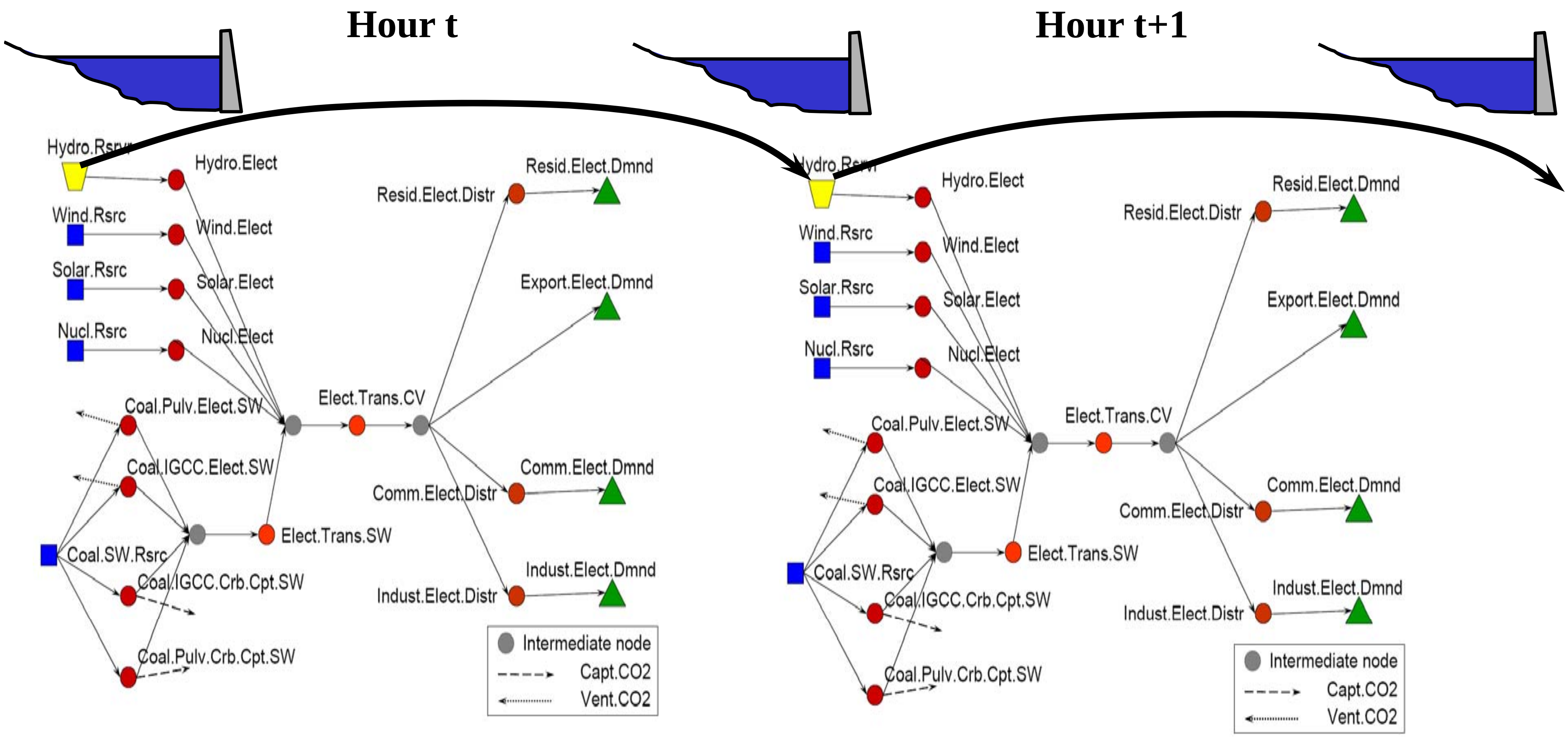
The hourly dispatch problem

Hourly electricity “dispatch” problem



The hourly dispatch problem

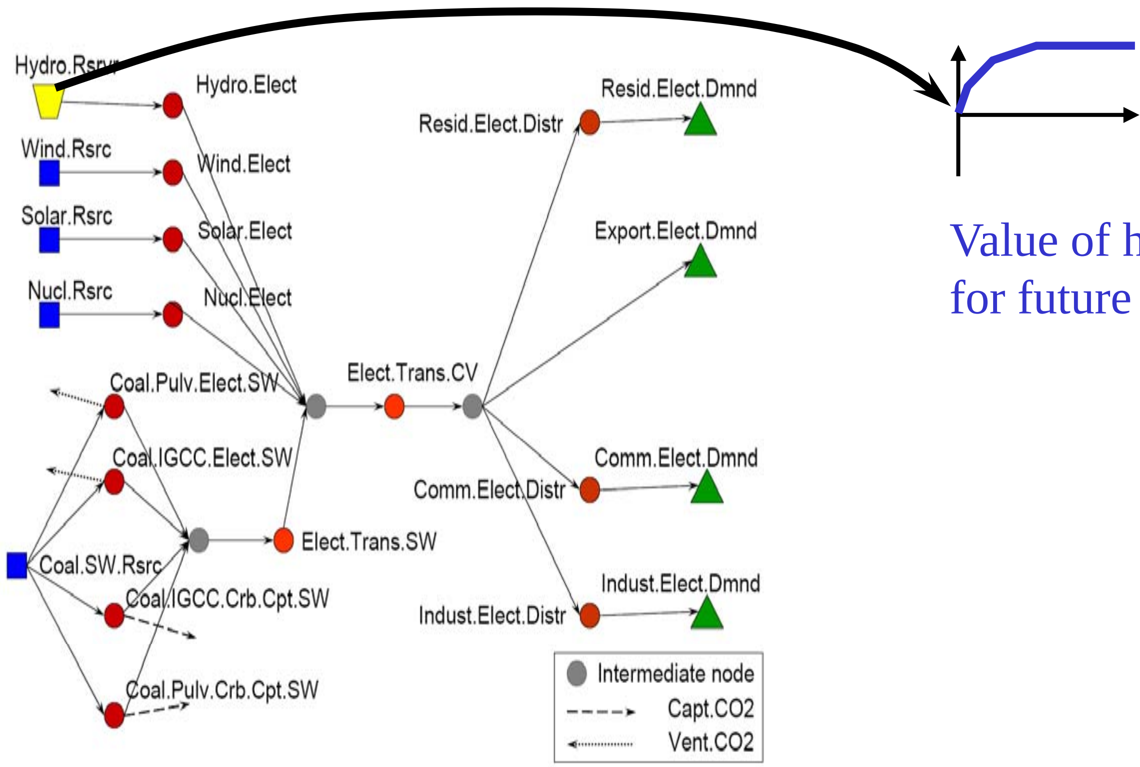
- Hourly model
  - » Decisions at time *t* impact *t+1* through the amount of water held in the reservoir.



The hourly dispatch problem

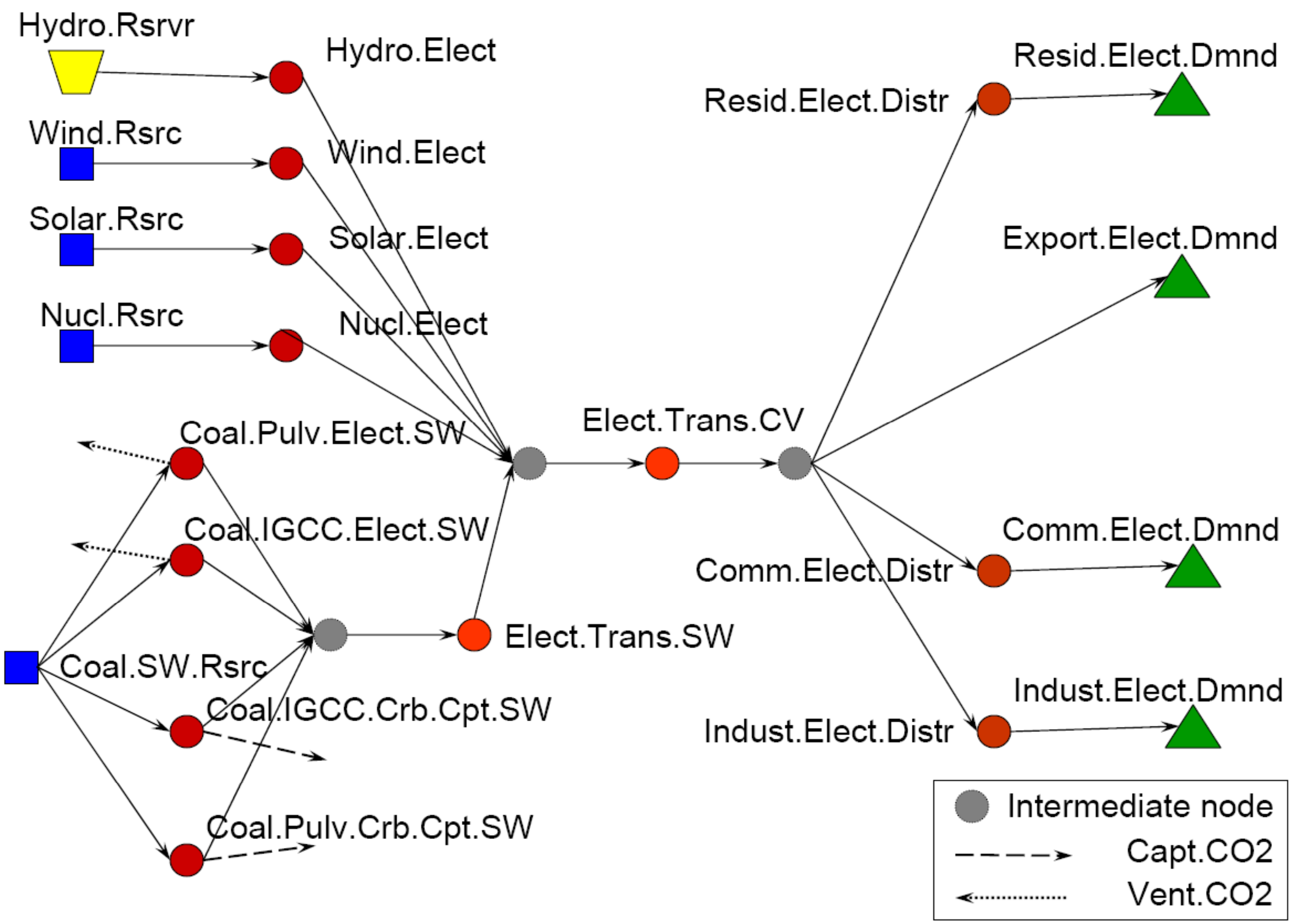
- Hourly model
  - » Decisions at time *t* impact *t+1* through the amount of water held in the reservoir.

Hour t

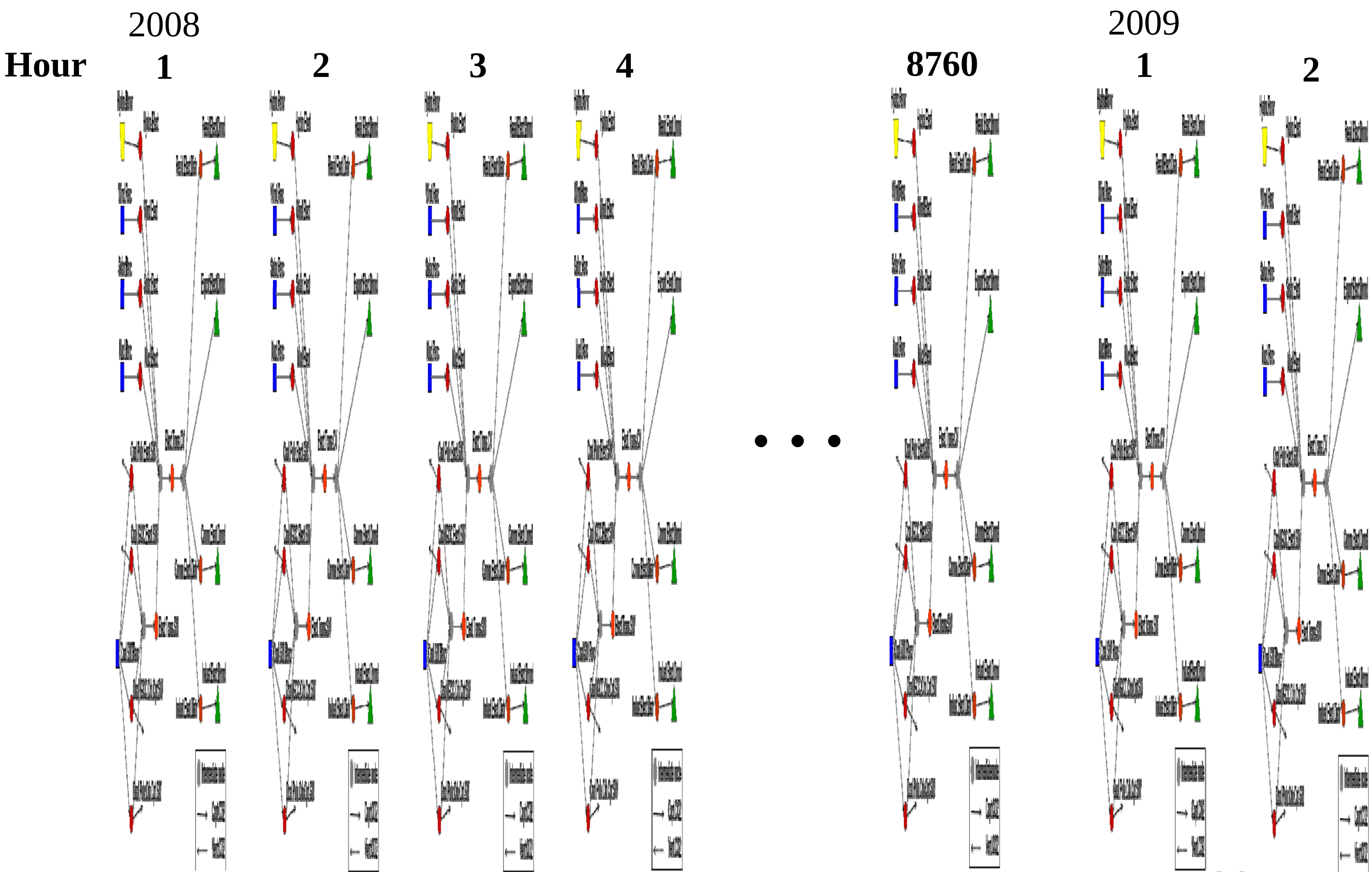


Value of holding water in the reservoir for future time periods.

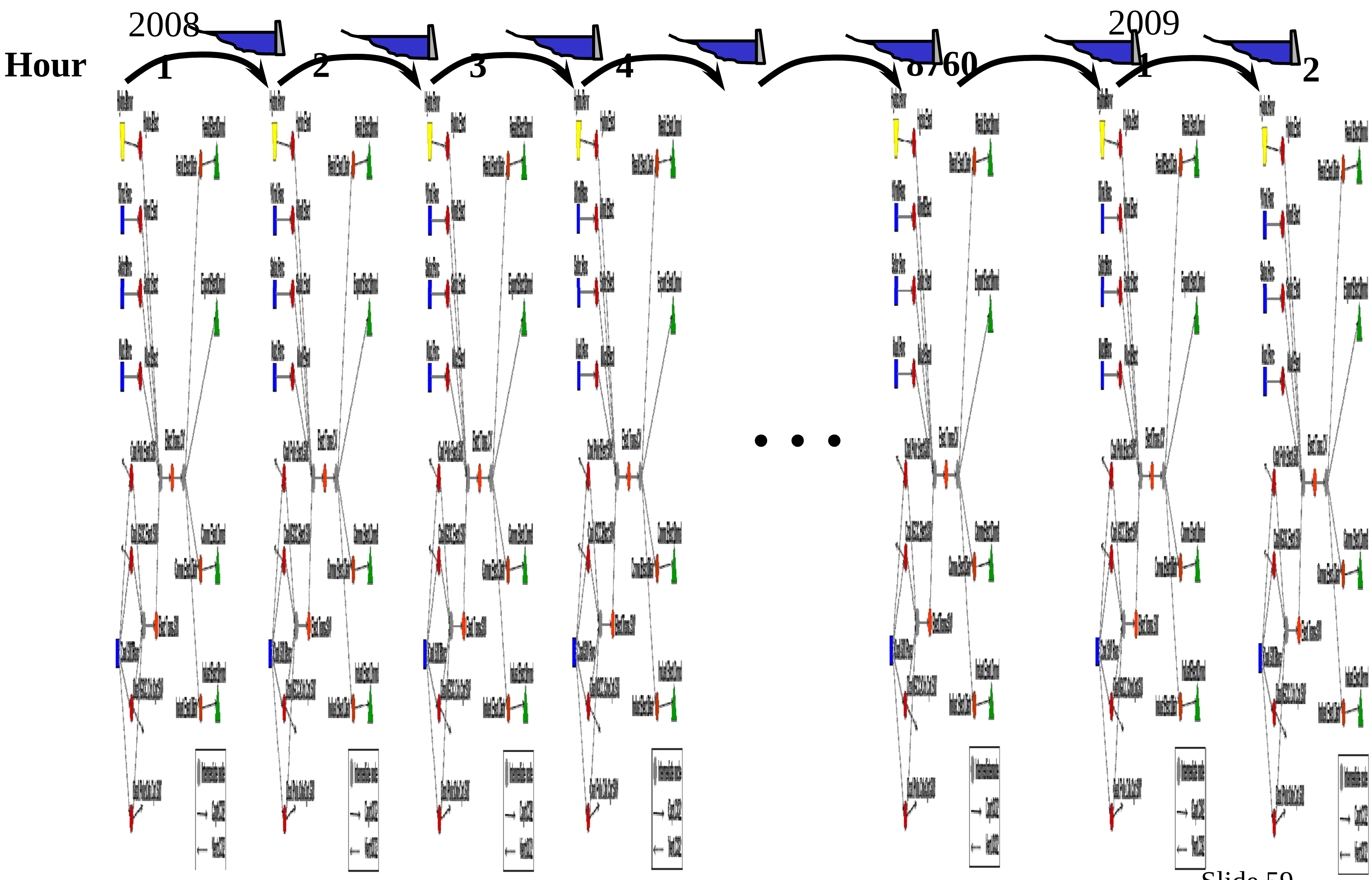
The hourly dispatch problem



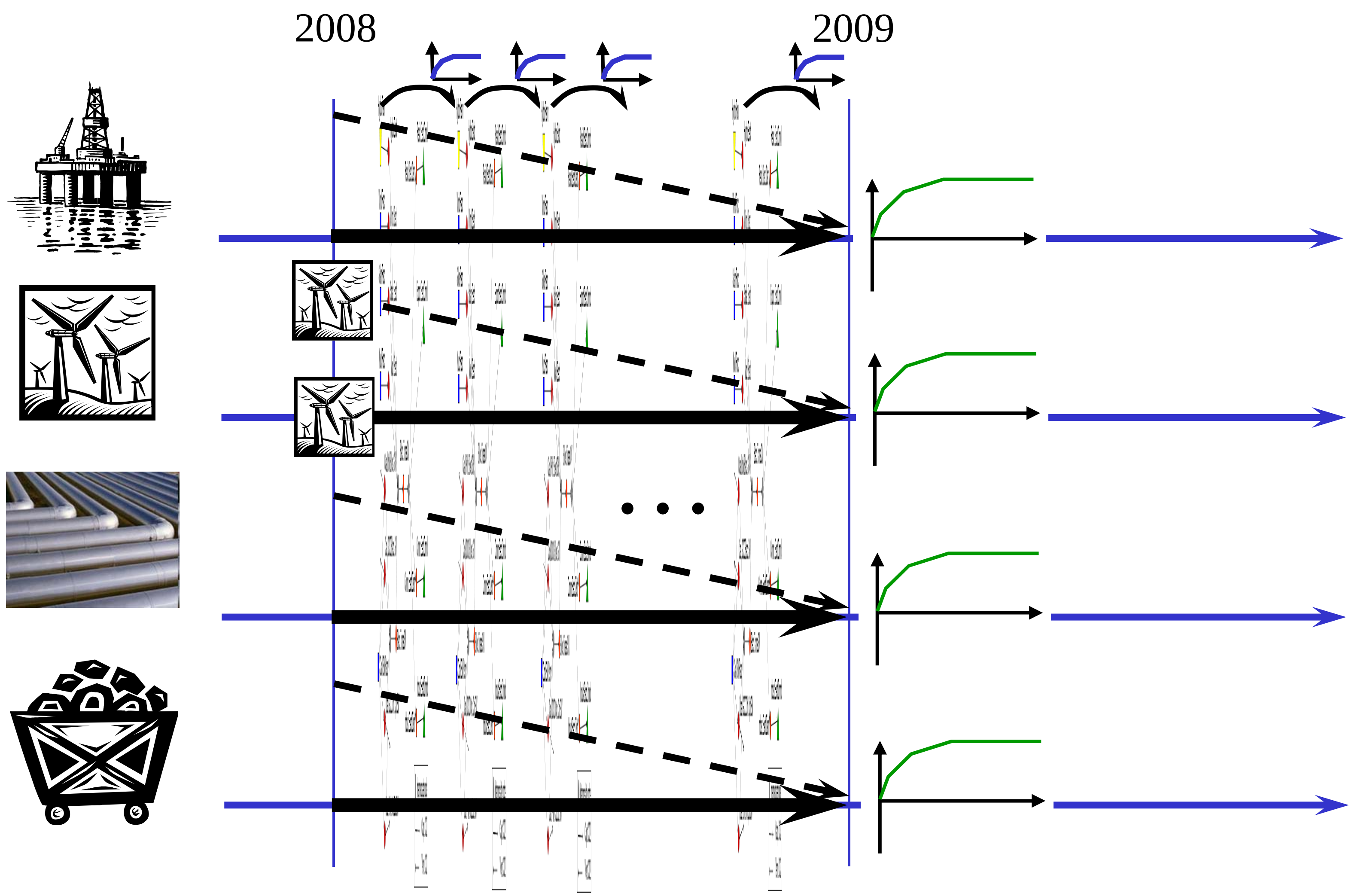
The hourly dispatch problem



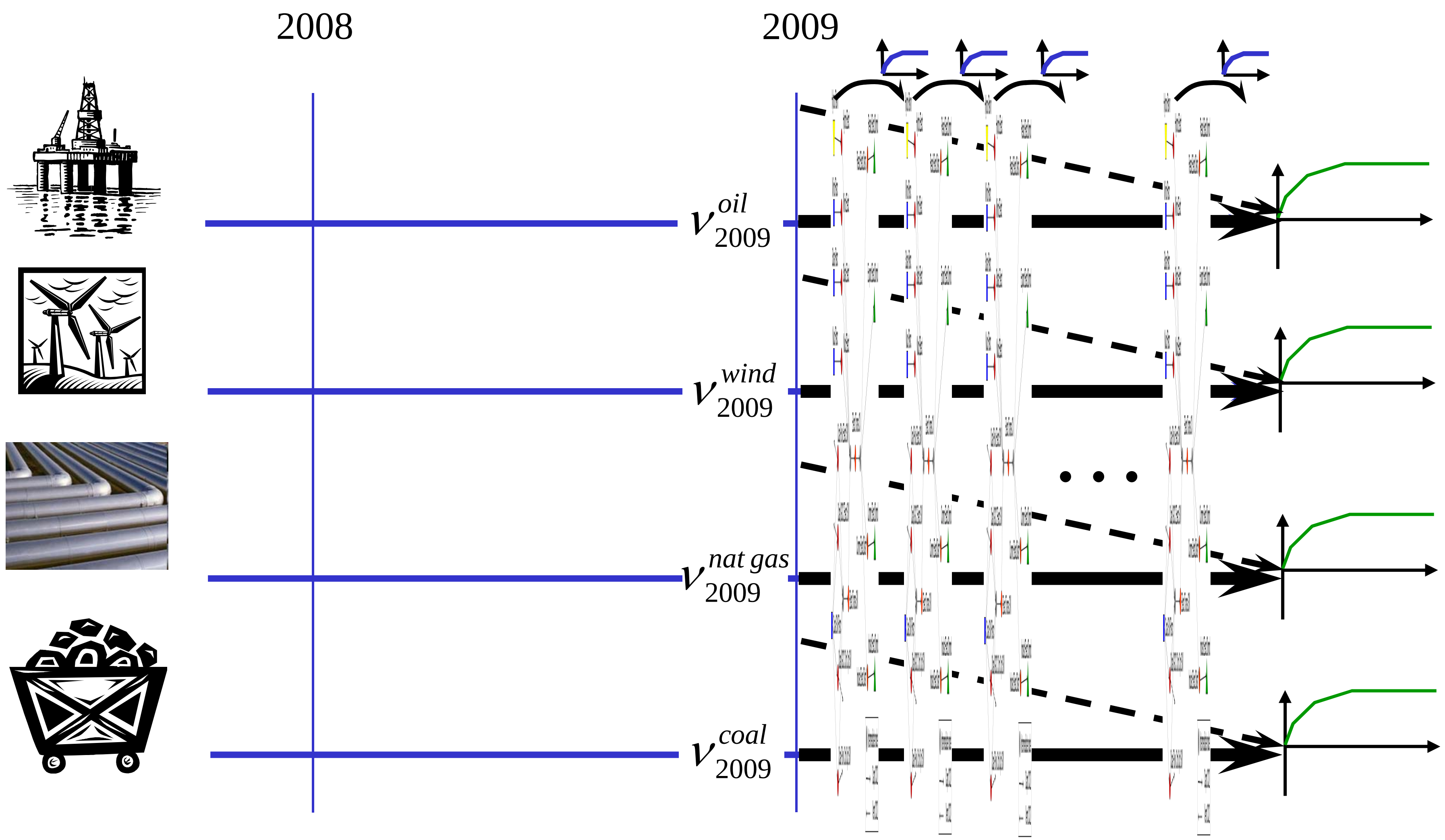
The hourly dispatch problem



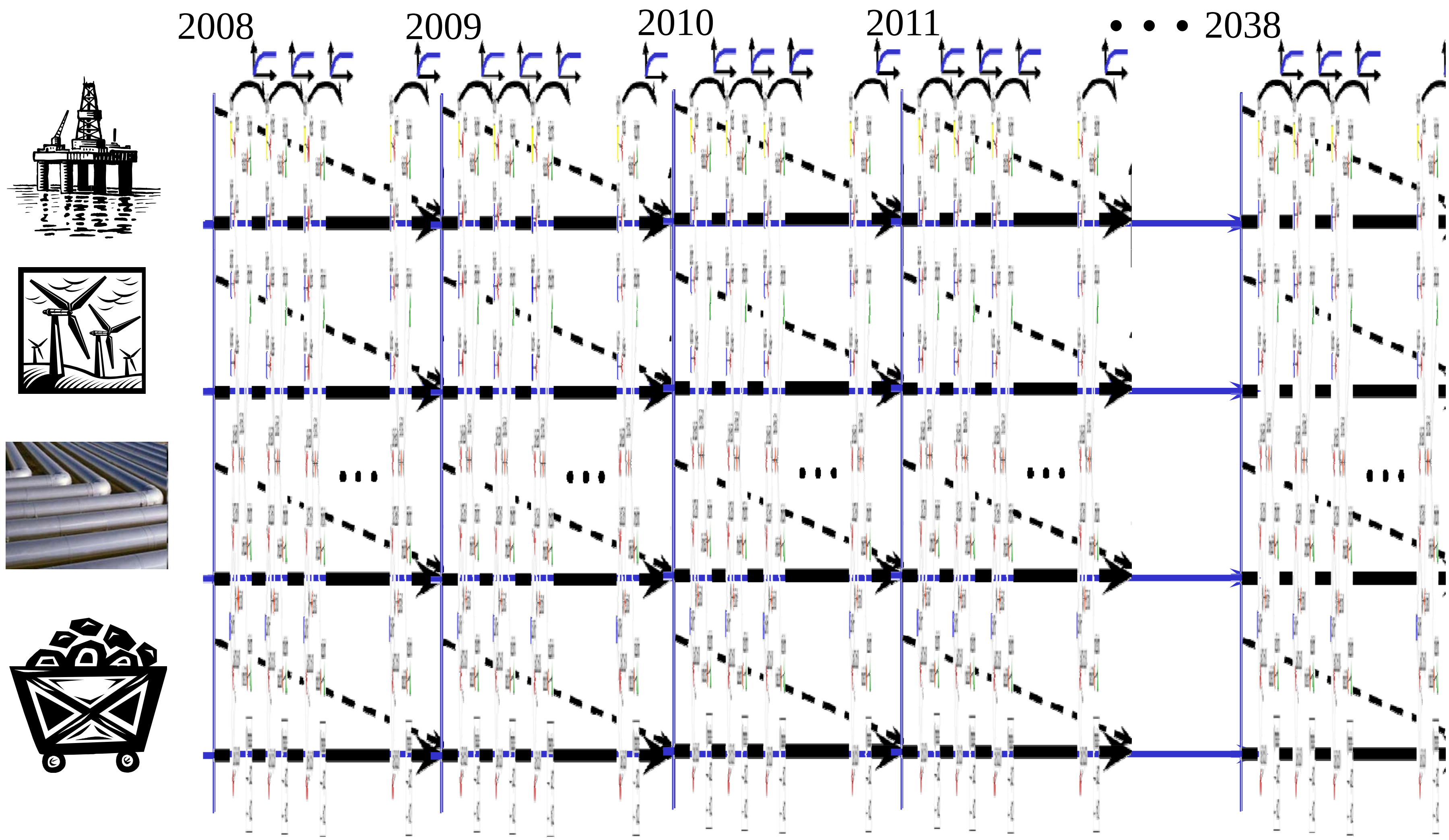
SMART-Stochastic, multiscale model



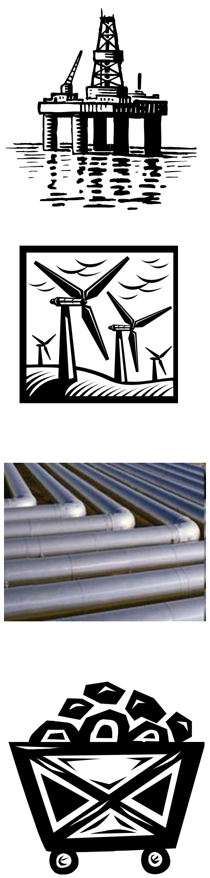
SMART-Stochastic, multiscale model

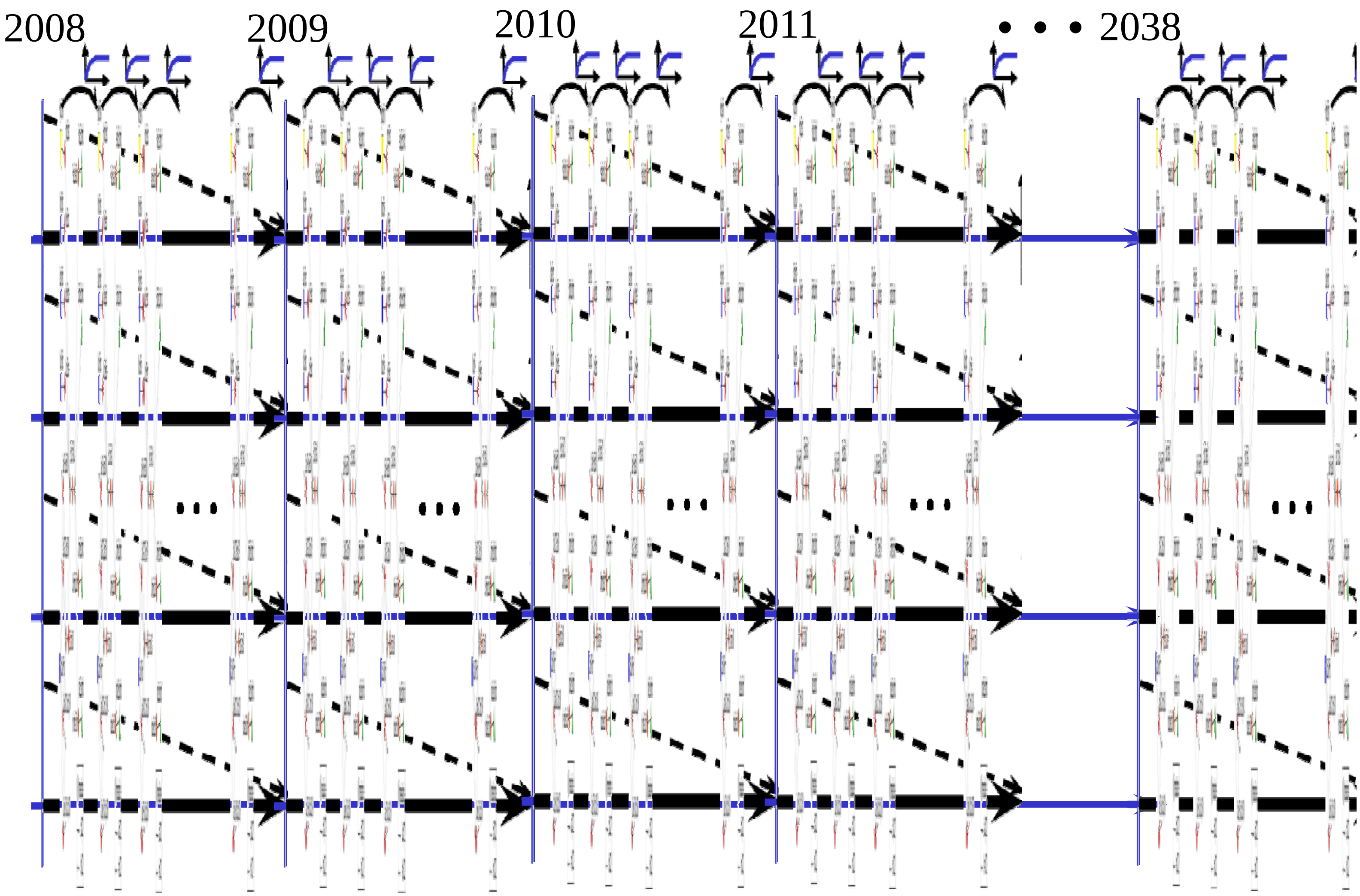


SMART-Stochastic, multiscale model

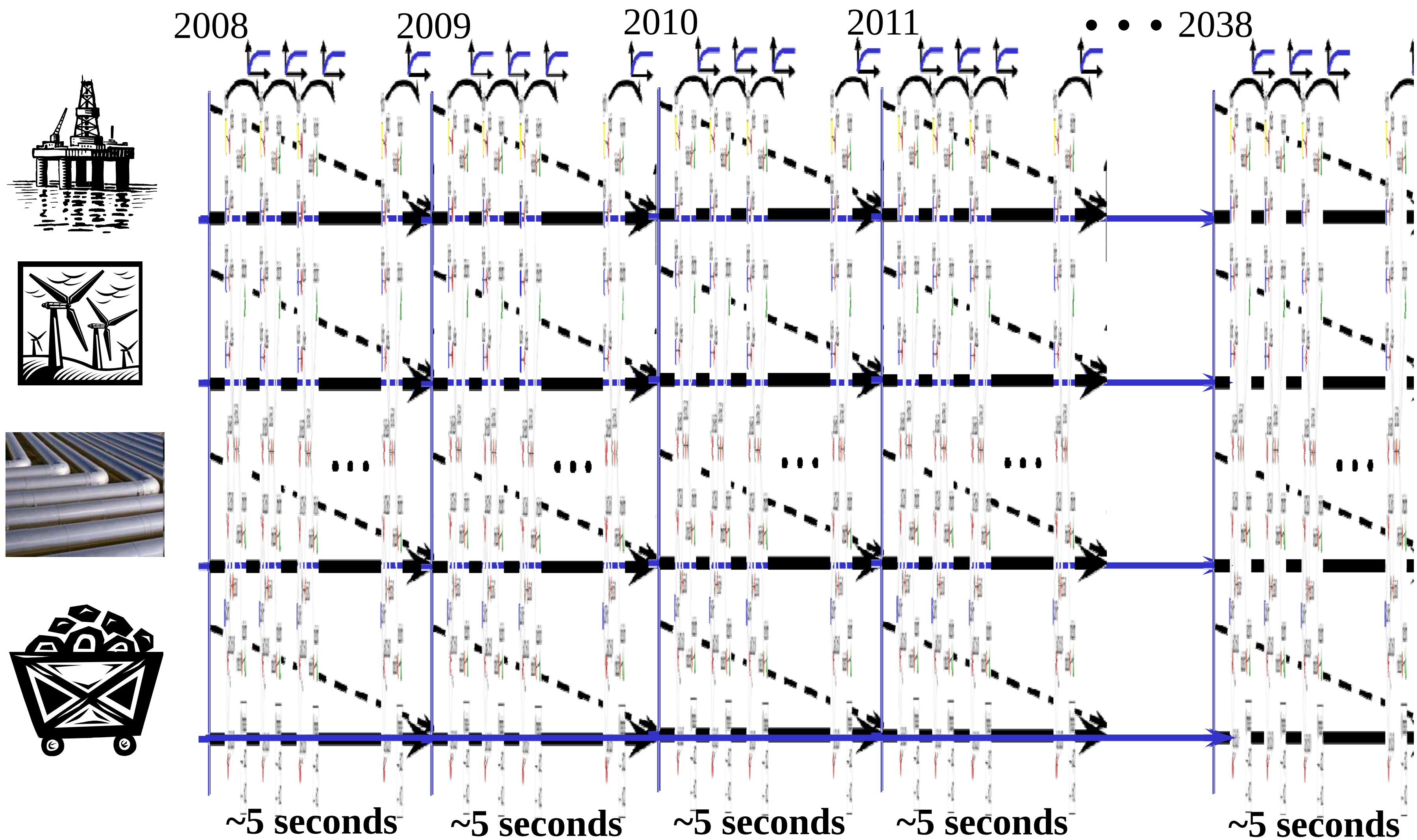


SMART-Stochastic, multiscale model



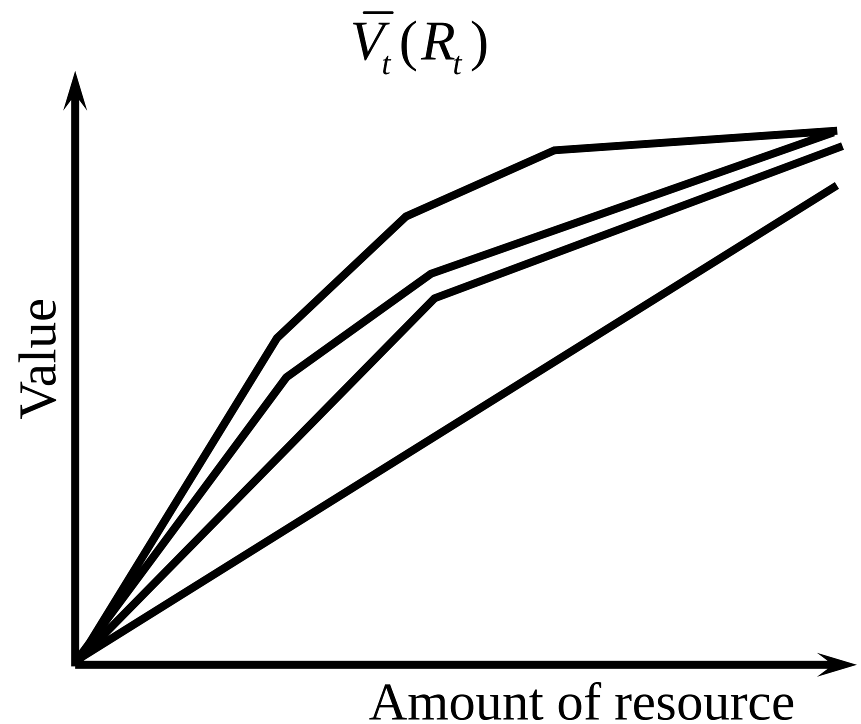


SMART-Stochastic, multiscale model



SMART-Stochastic, multiscale model

- Use statistical methods to learn the value of resources in the future.
- Resources may be:
  - » Stored energy
    - Hydro
    - Flywheel energy
    - …
  - » Storage capacity
    - Batteries
    - Flywheels
    - Compressed air
  - » Energy transmission capacity
    - Transmission lines
    - Gas lines
    - Shipping capacity
  - » Energy production sources
    - Wind mills
    - Solar panels
    - Nuclear power plants

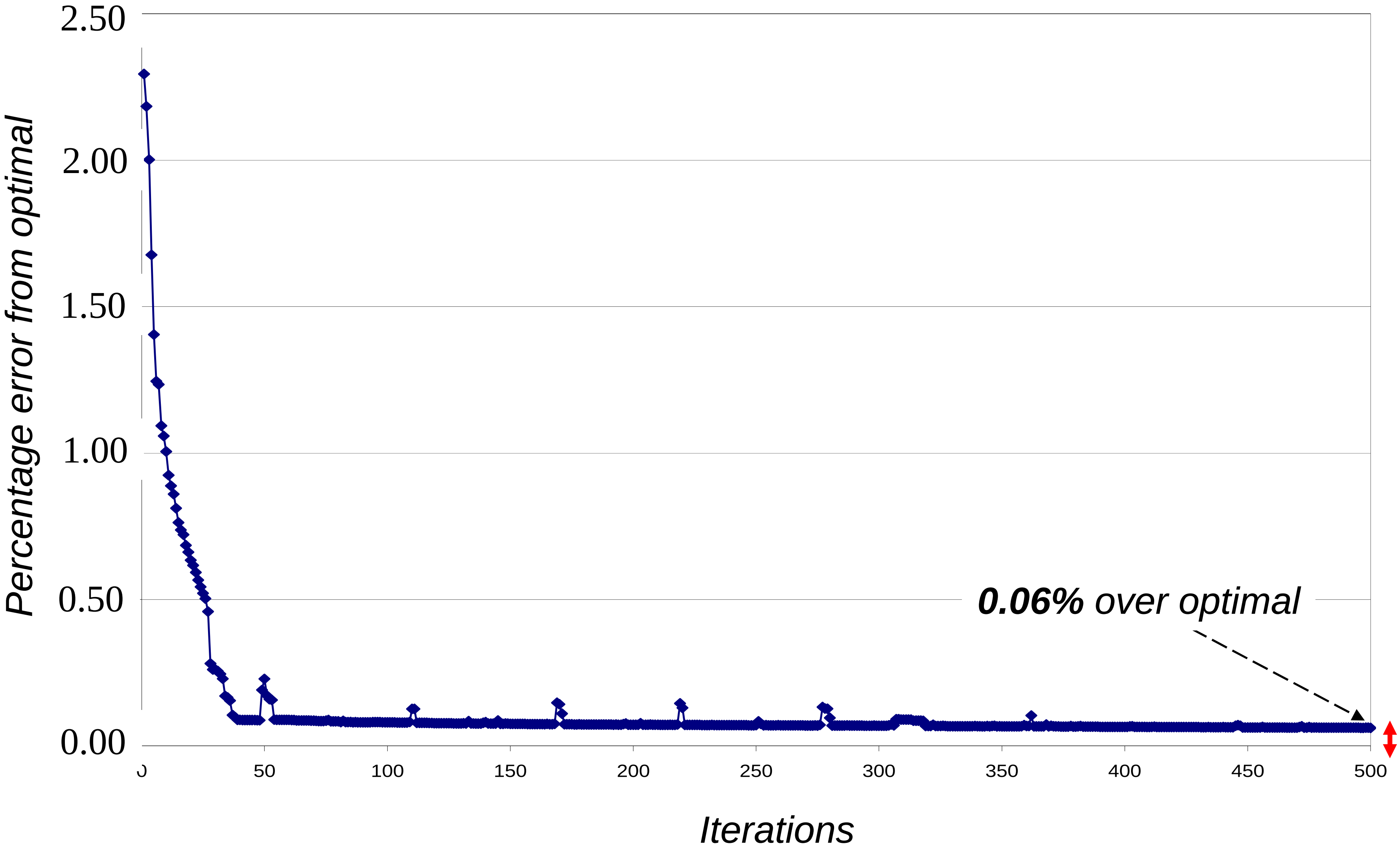


SMART-Stochastic, multiscale model

- Benchmarking
  - » Compare ADP to optimal LP for a deterministic problem
    - Annual model
      - – 8,760 hours over a single year
      - – Focus on ability to match hydro storage decisions
    - 20 year model
      - – 24 hour time increments over 20 years
      - – Focus on investment decisions
  - » Comparisons on stochastic model
    - Stochastic rainfall analysis
      - – How does ADP solution compare to LP?
    - Carbon tax policy analysis
      - – Demonstrate nonanticipativity

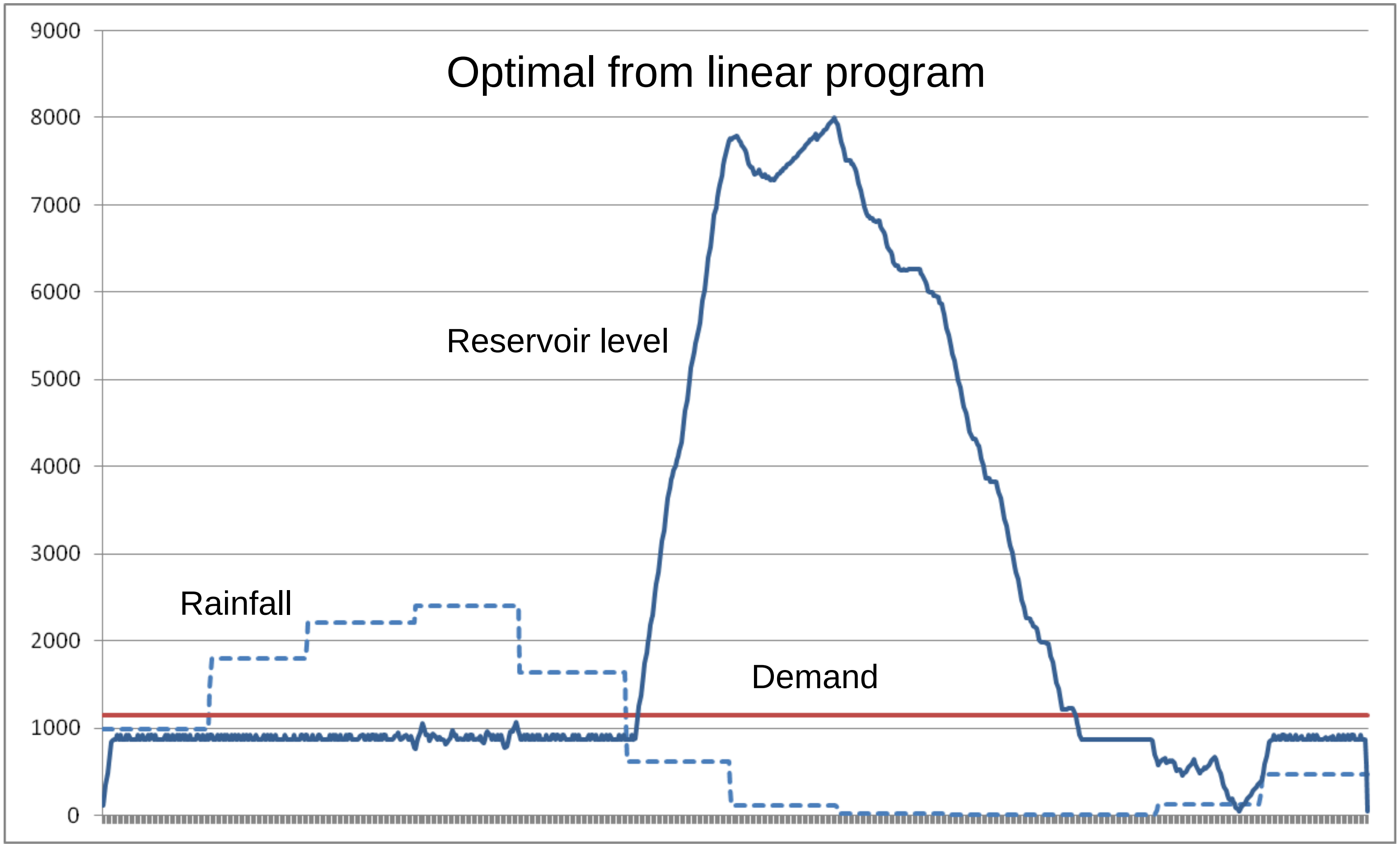
Benchmarking on hourly dispatch

- ADP objective function relative to optimal LP



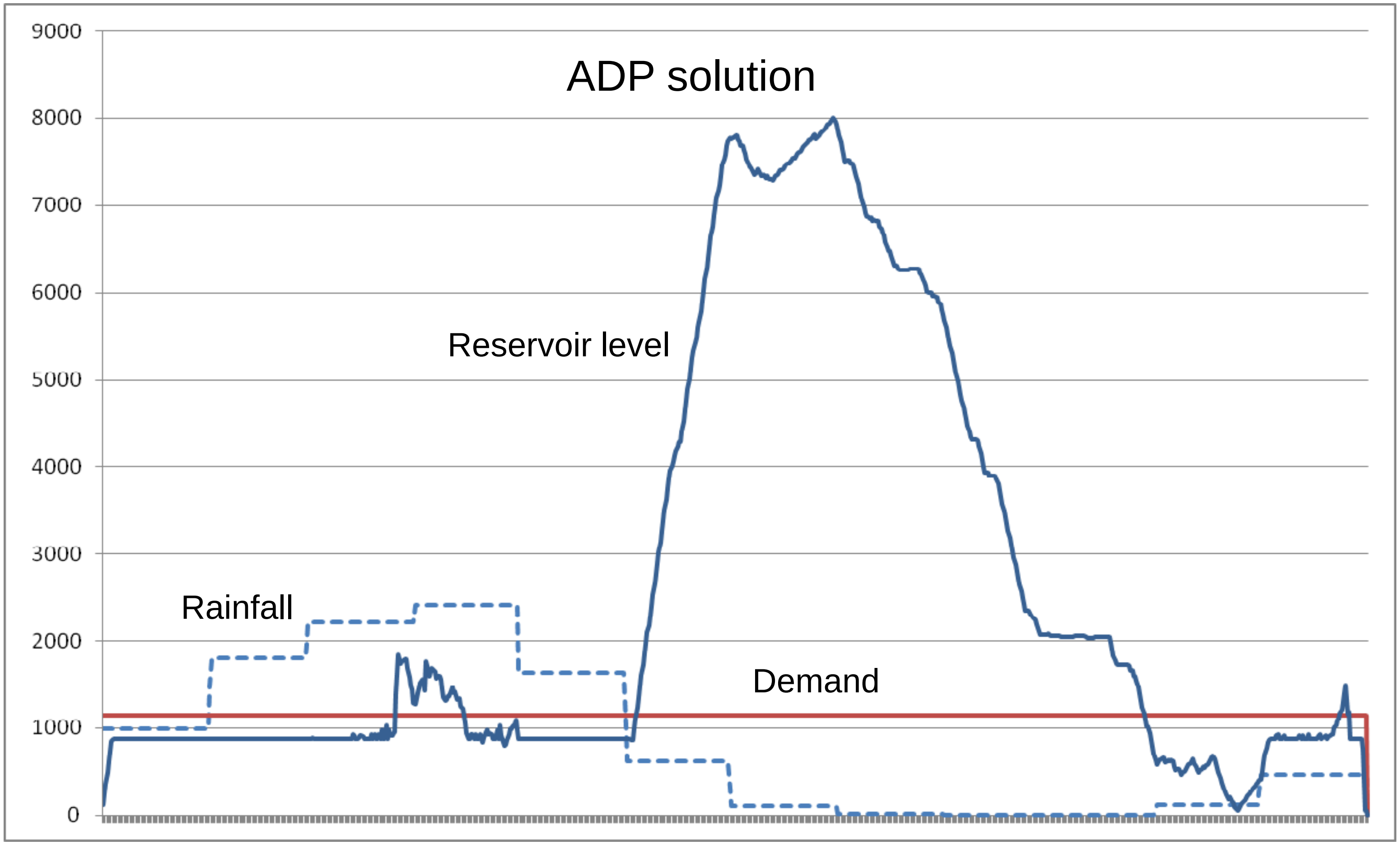
Benchmarking on hourly dispatch

- Optimal from linear program



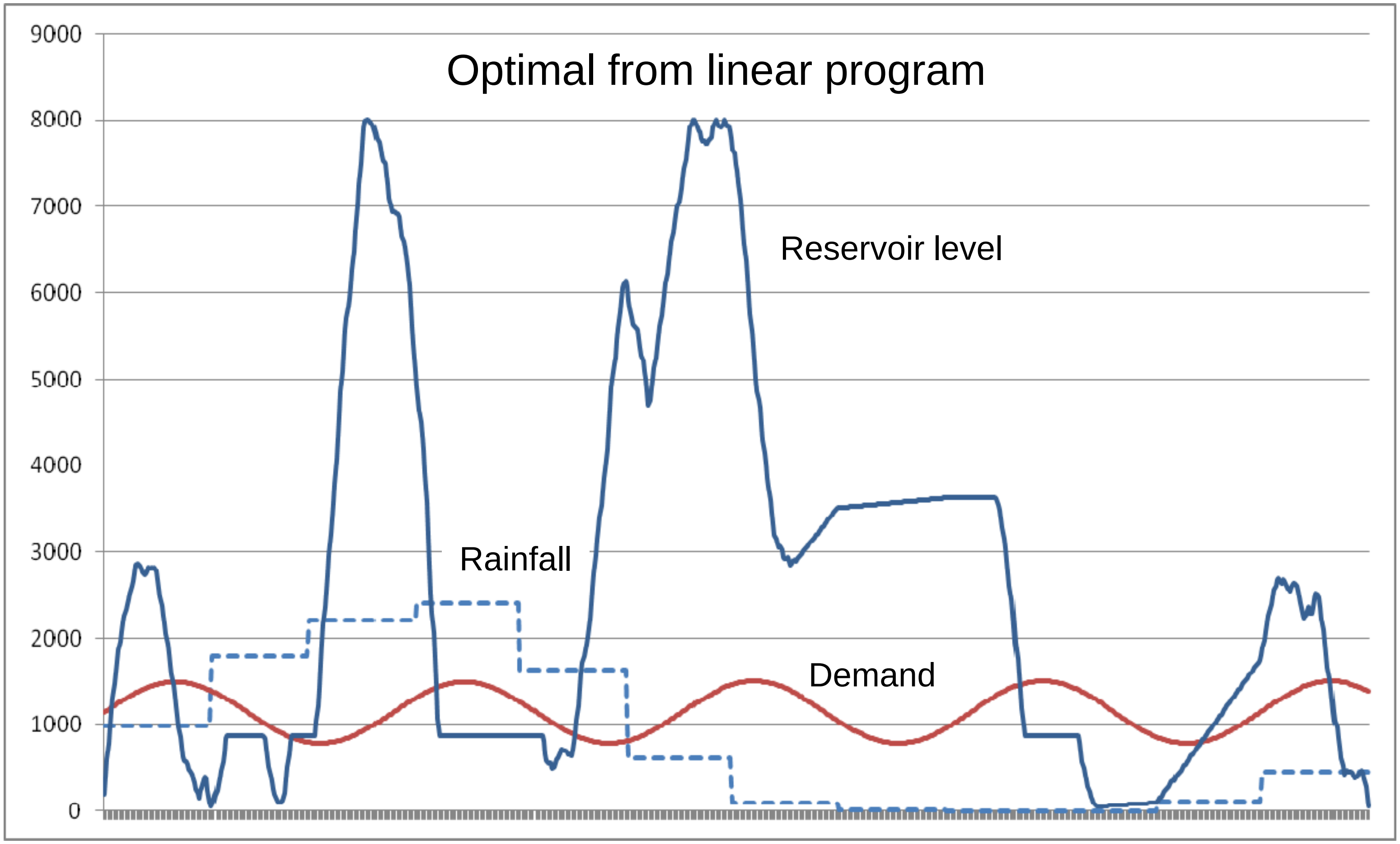
Benchmarking on hourly dispatch

- Approximate dynamic programming



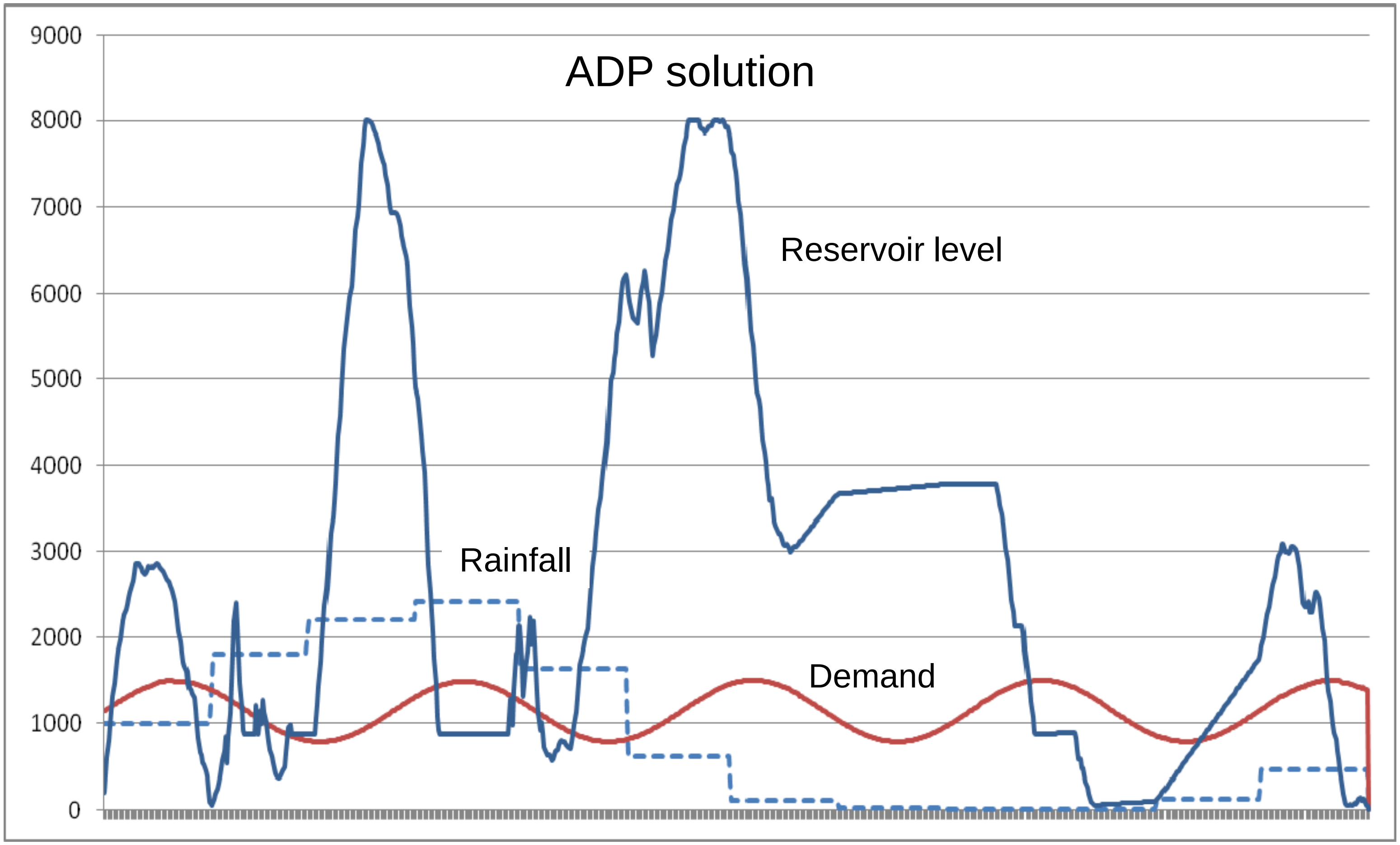
Benchmarking on hourly dispatch

- Optimal from linear program

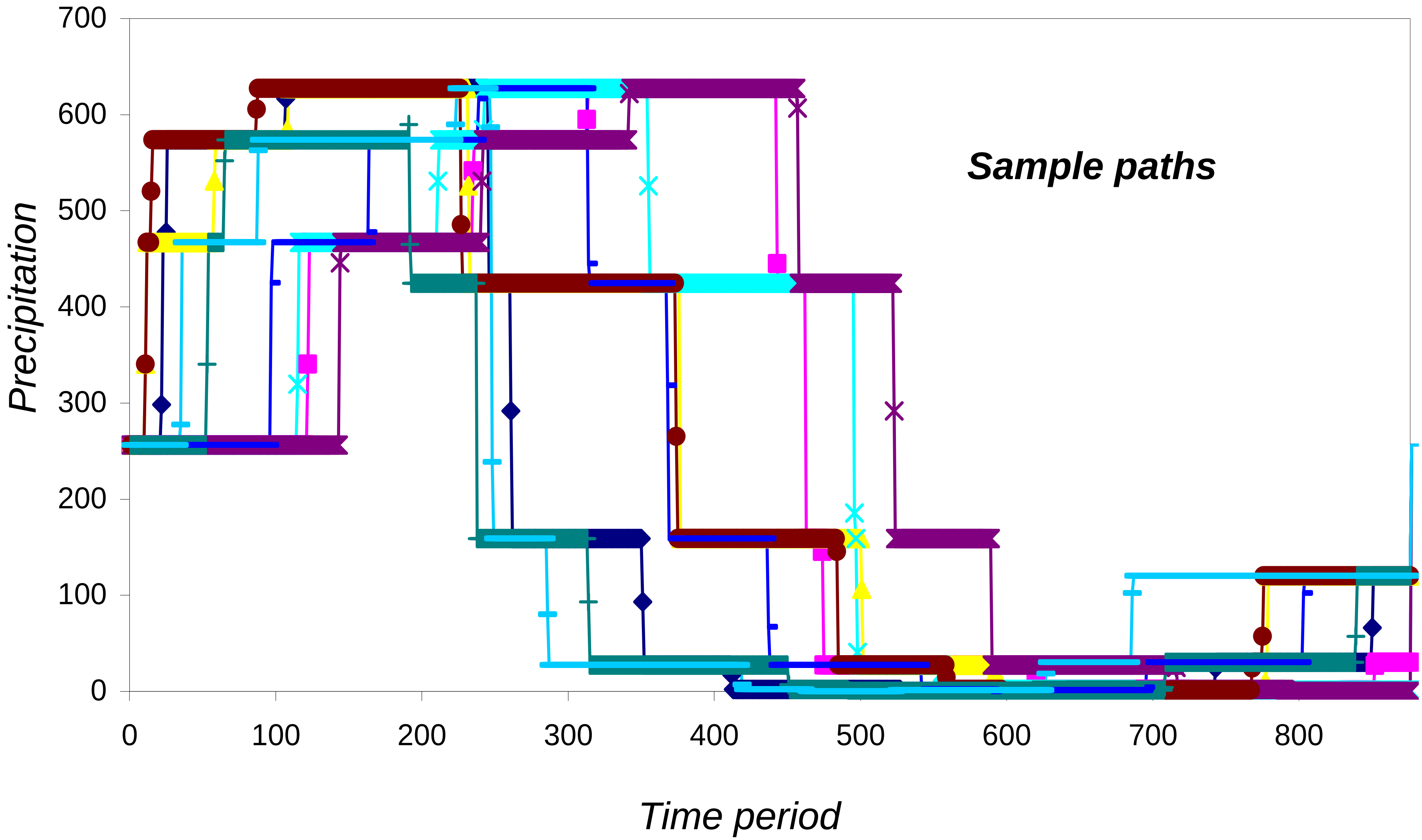


Benchmarking on hourly dispatch

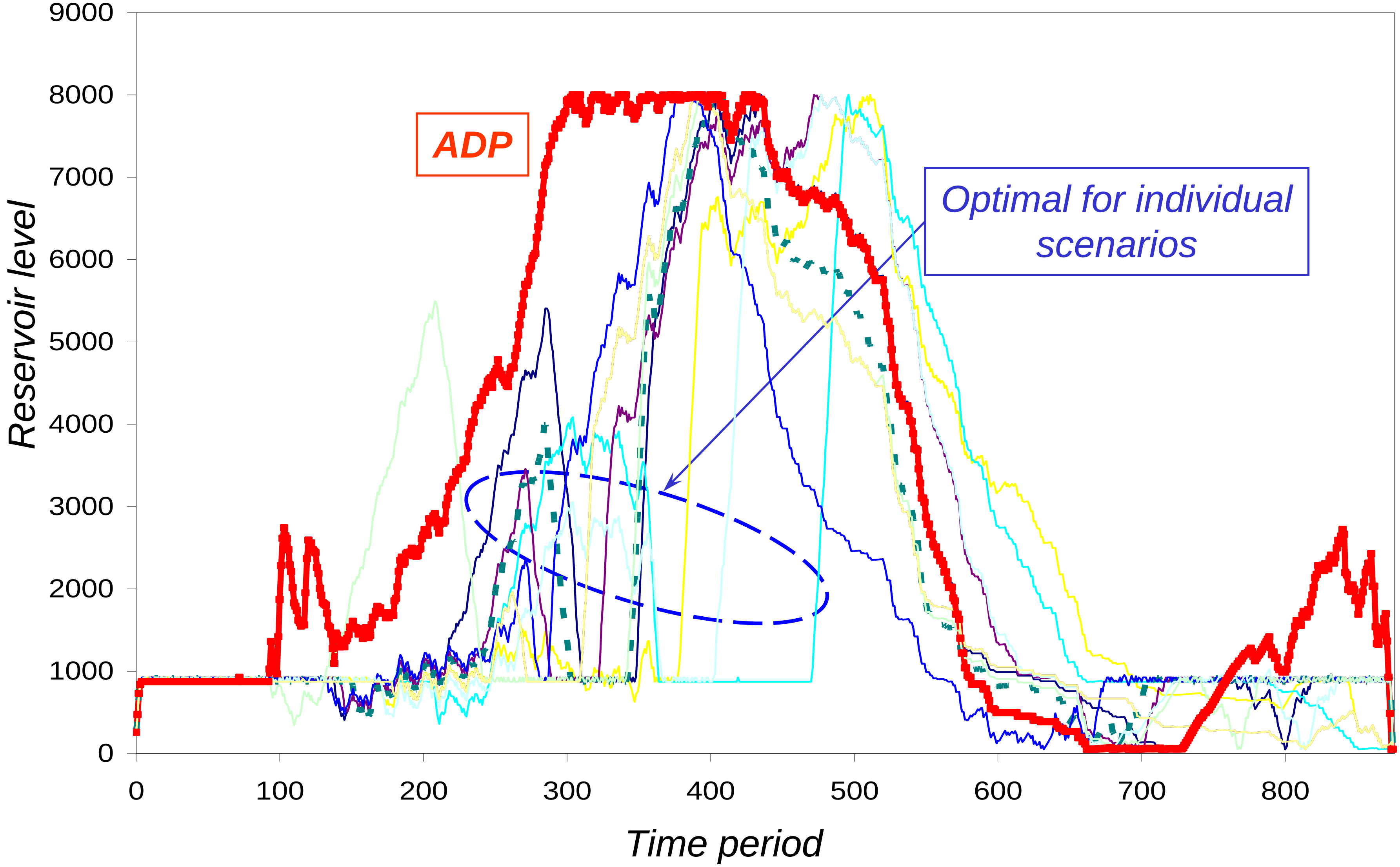
- Approximate dynamic programming



Stochastic rainfall



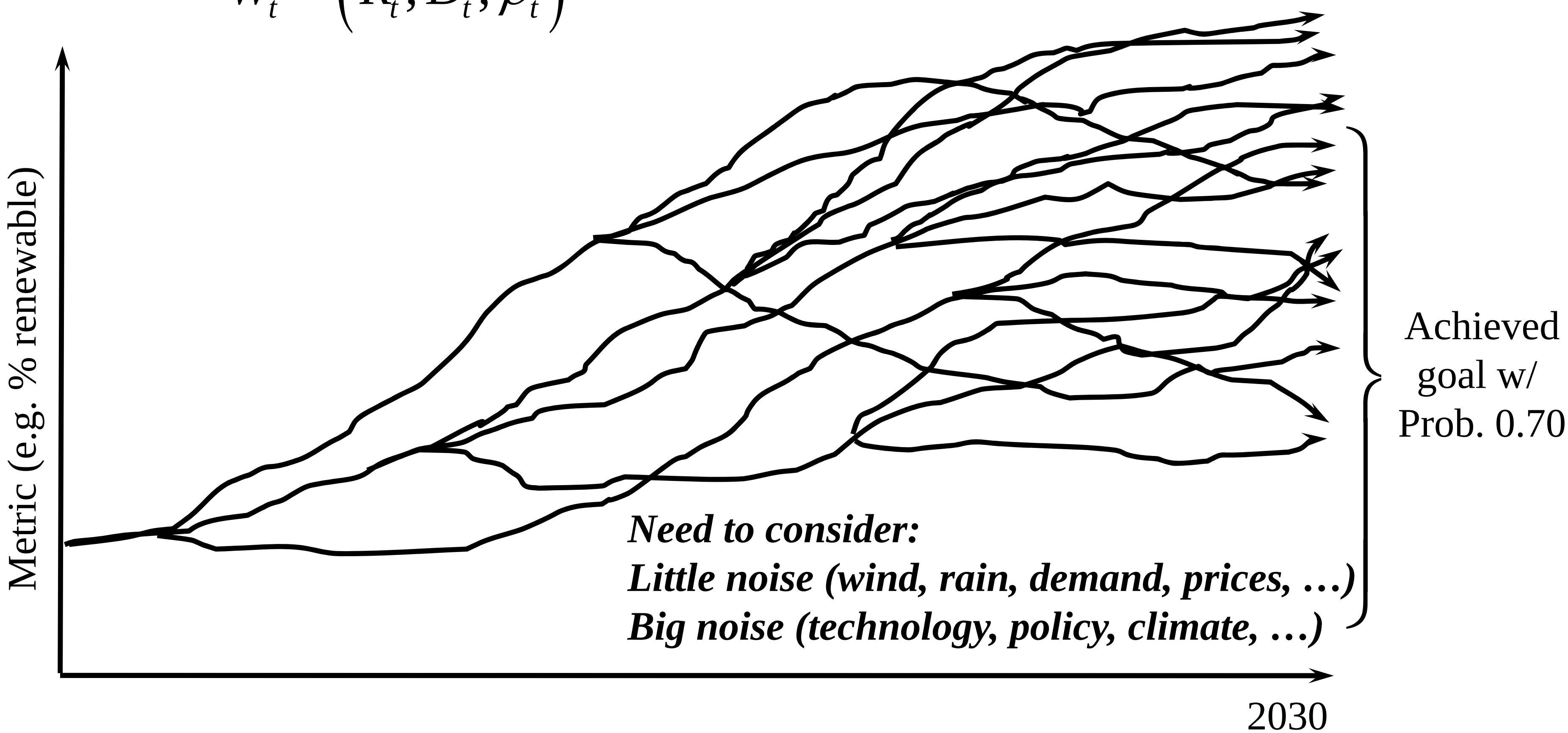
Stochastic rainfall



Energy policy modeling

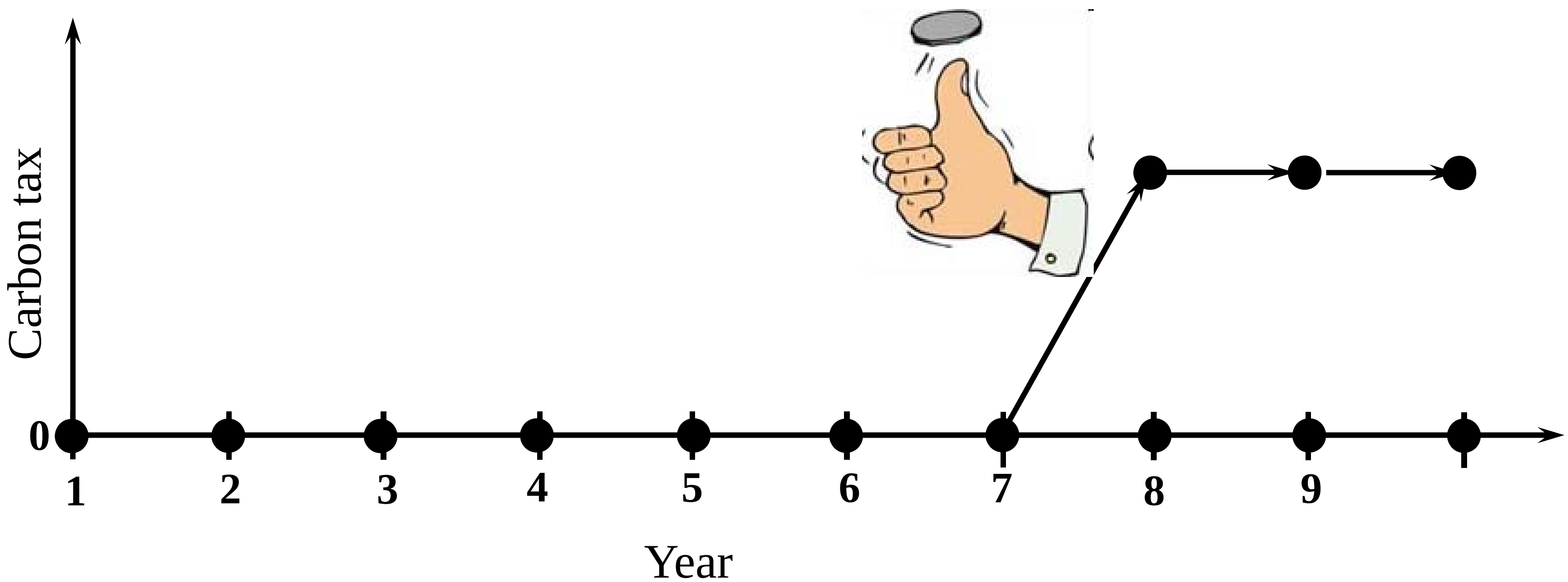
- Following sample paths
  - » Demands, prices, weather, technology, policies, …

$$W_t = \left(\hat{R}_t, \hat{D}_t, \hat{\rho}_t\right)$$

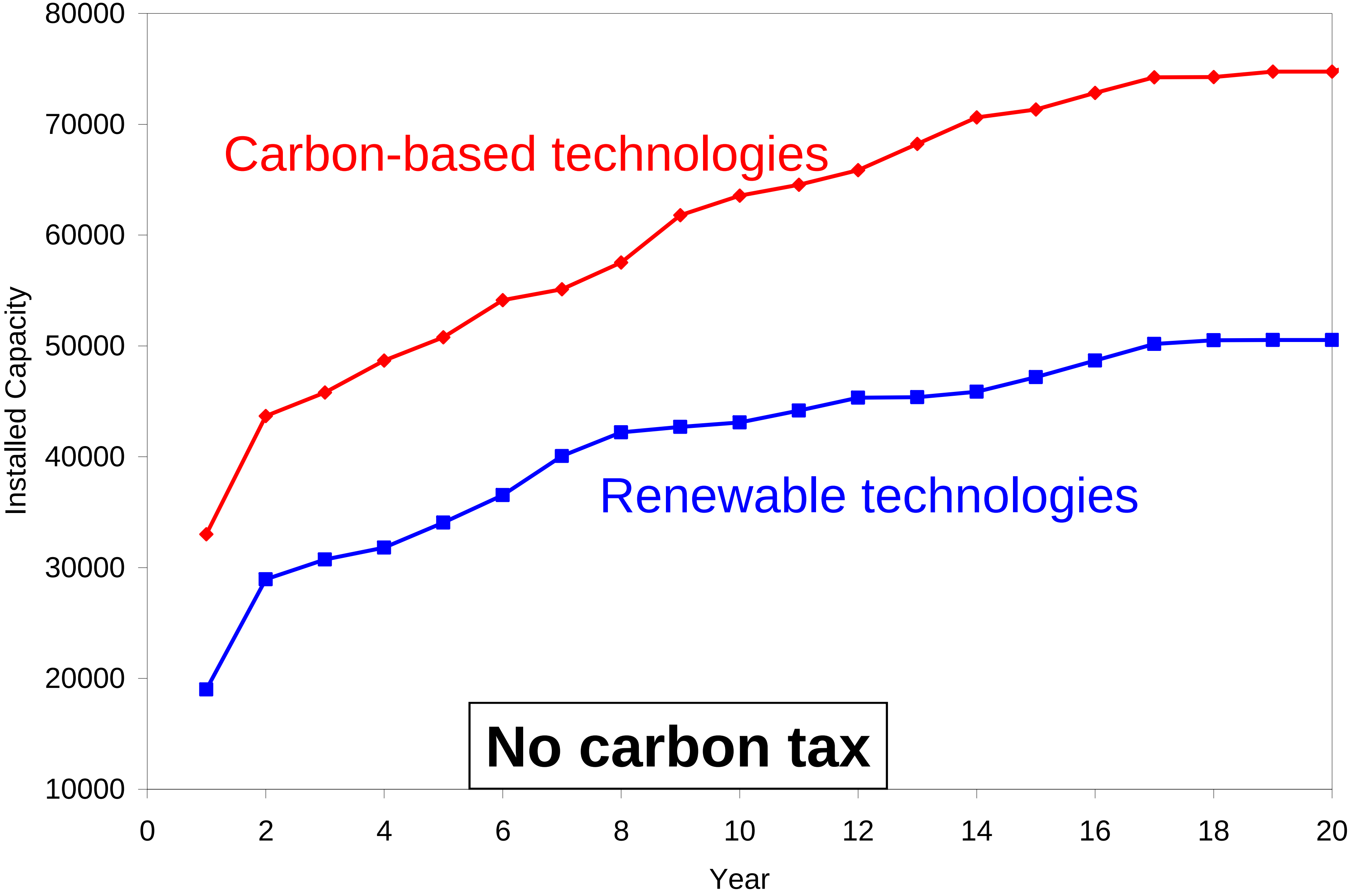


Energy policy modeling

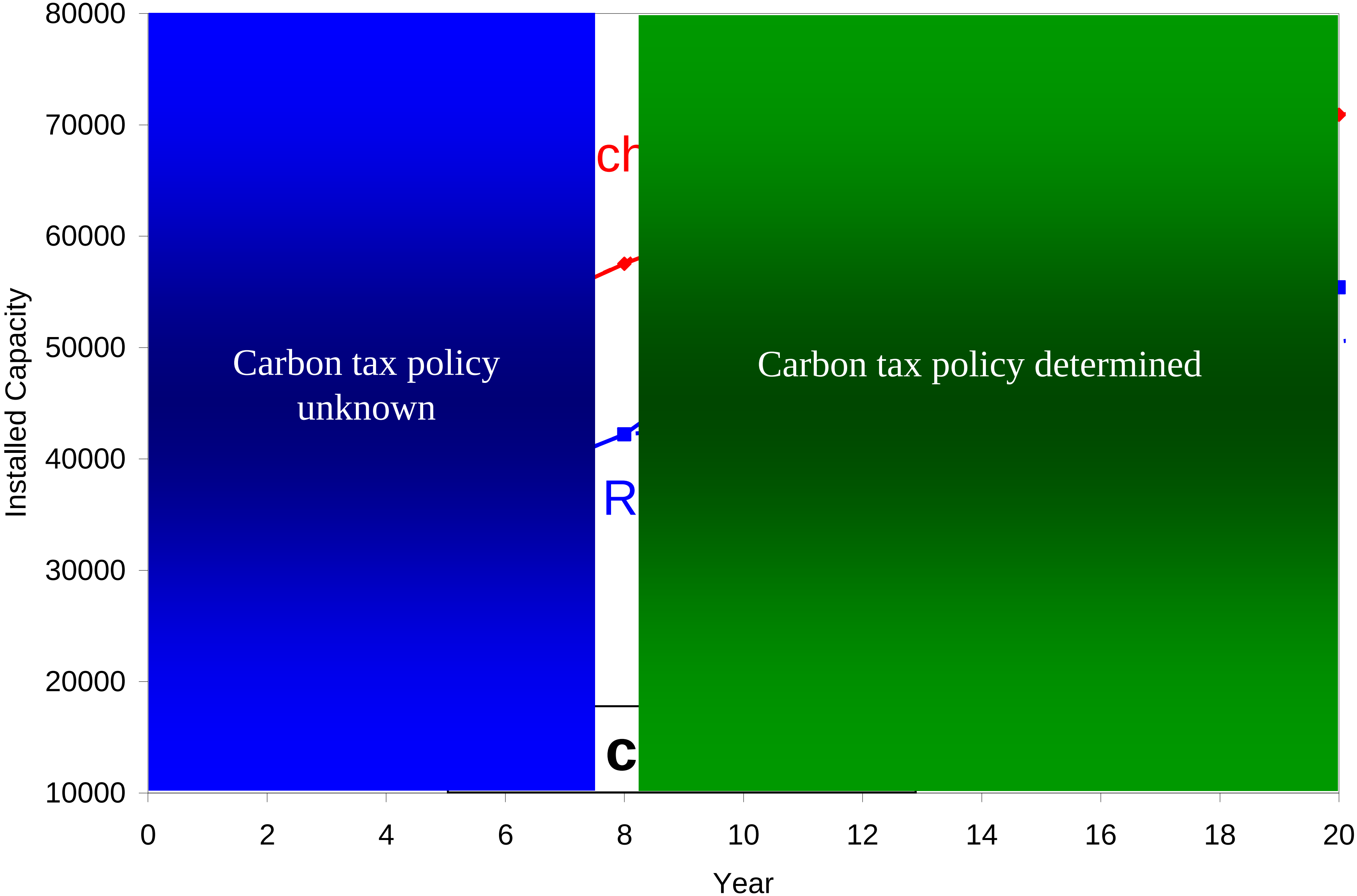
- Policy study:
  - » What is the effect of a potential (but uncertain) carbon tax in year 8?



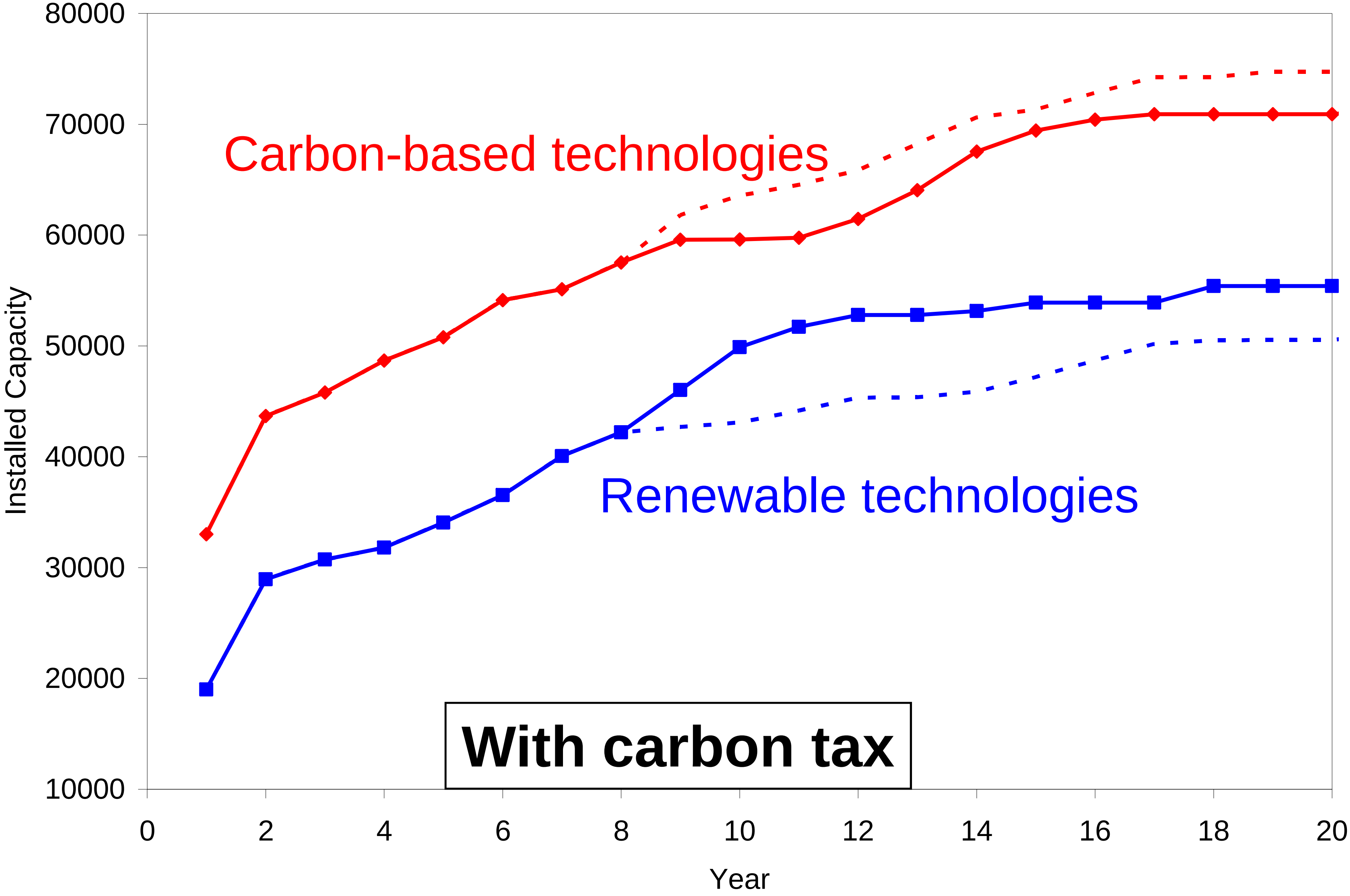
Energy policy modeling



Energy policy modeling



Energy policy modeling



Conclusions

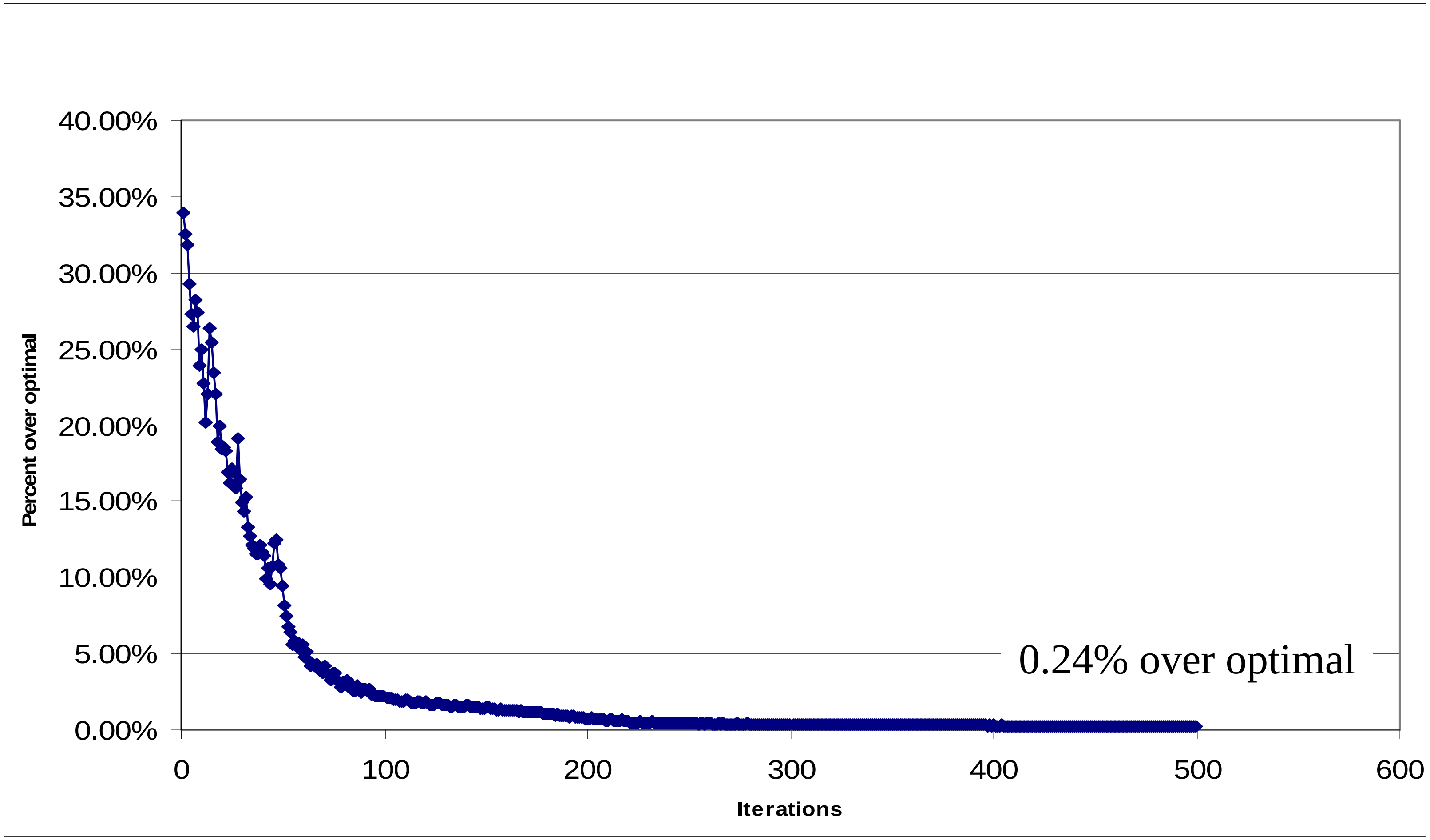
- Capabilities
  - » SMART can handle problems with over 300,000 time periods so that it can model hourly variations in a long-term energy investment model.
  - » It can simulate virtually any form of uncertainty, either provided through an exogenous scenario file or sampled from a probability distribution.
  - » Accurate modeling of climate, technology and markets requires access to exogenously provided scenarios.
  - » It properly models storage processes over time.
  - » Current tests are on an aggregate model, but the modeling framework (and library) is set up for spatially disaggregate problems.

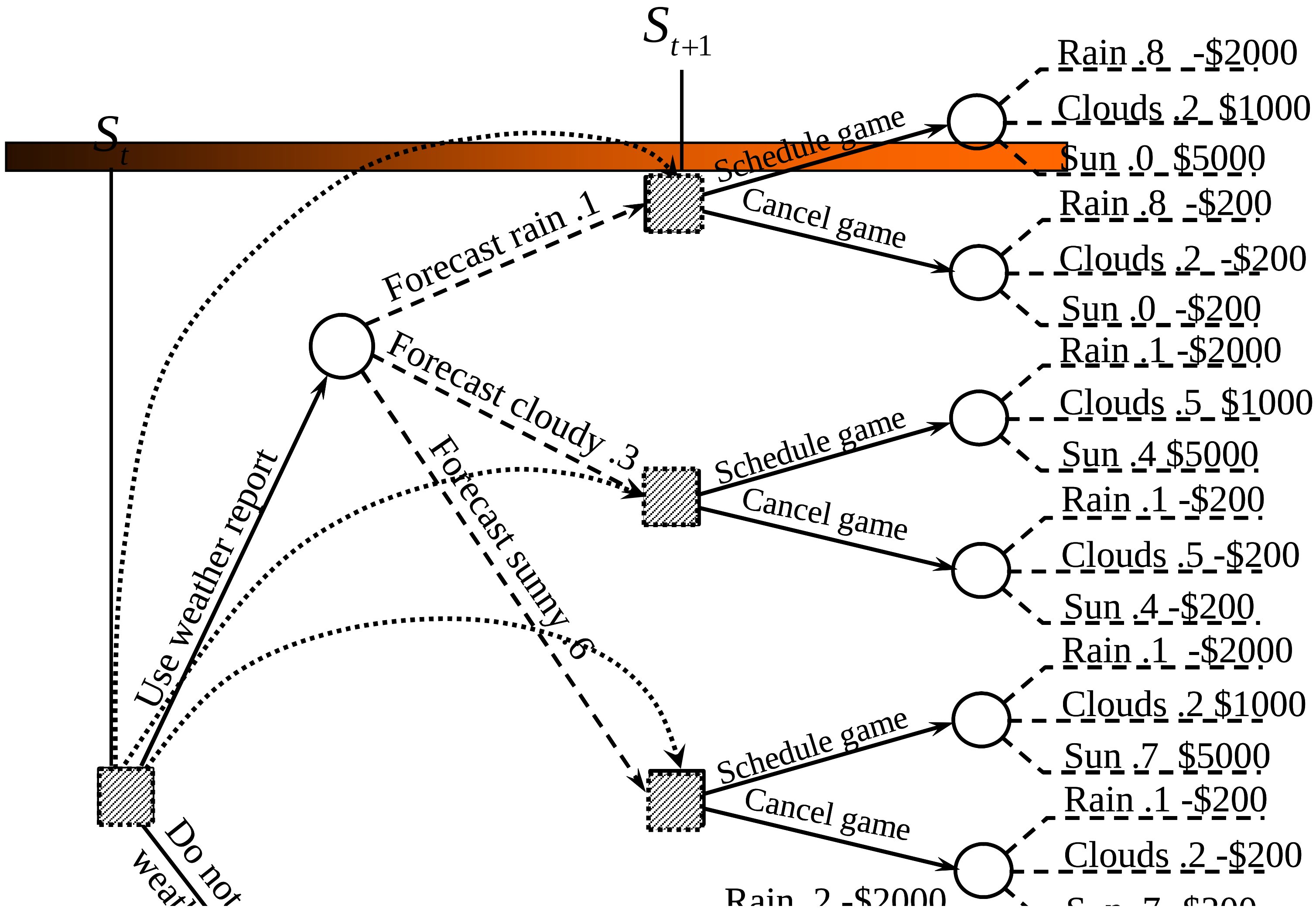
Conclusions

- Limitations
  - » Value function approximations capture the resource state vector, but are limited to very simple exogenous state variations.
  - » More research is needed to test the ability of the model to use multiple storage technologies.
  - » Extension to spatially disaggregate model will require significant engineering and data.
  - » Run times will start to become an issue for a spatially disaggregate model.

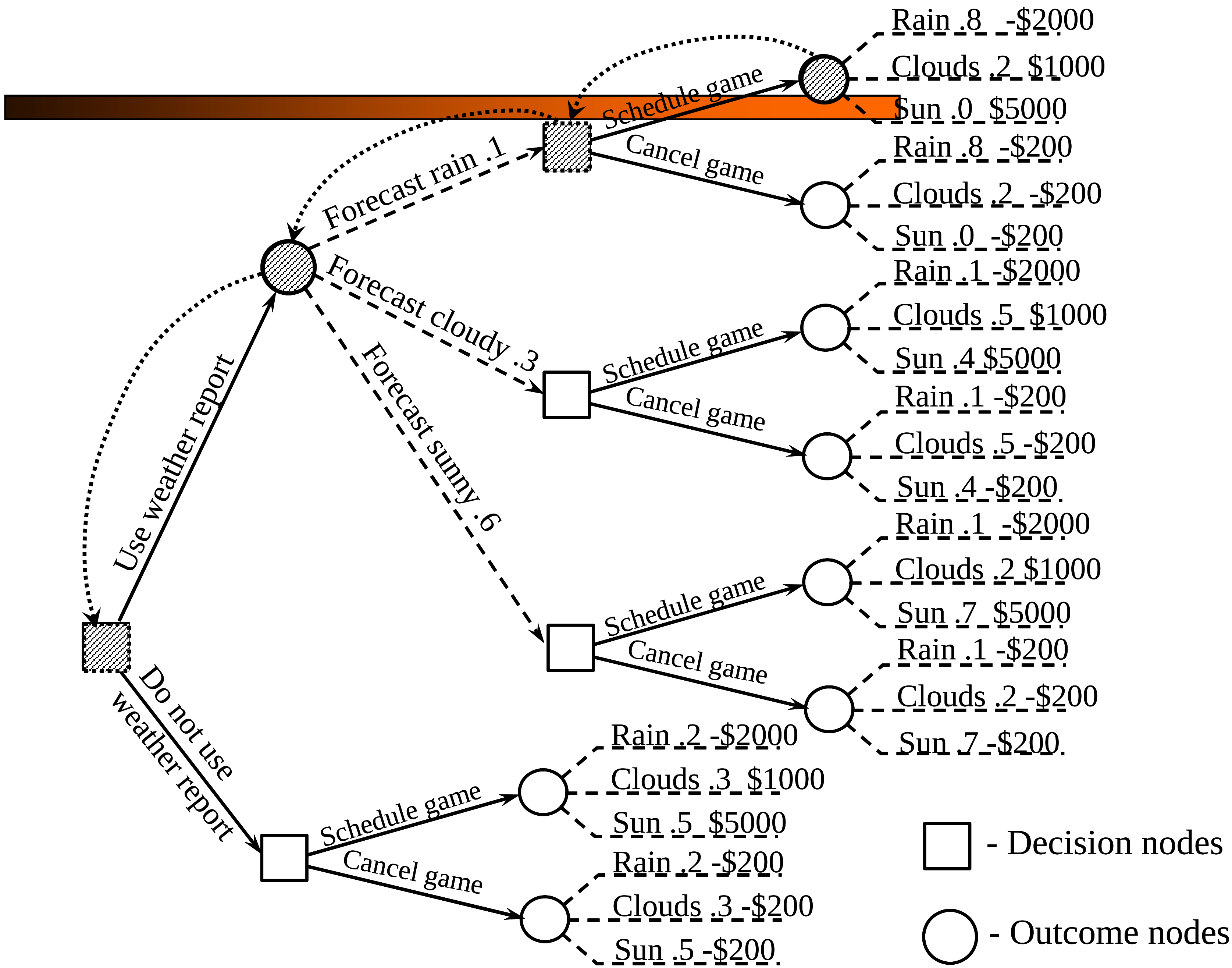
Multidecade energy model

- Optimal vs. ADP – daily model over 20 years





$$V_t(S_t) = \max_{x \in X}\left(C_t(S_t, x_t) + E\left\{V_{t+1}(S_{t+1}) \mid S_t\right\}\right)$$


Rain .8 -$2000
Clouds .2 $1000
Sun .0 $5000
Rain .8 -$200
Clouds .2 -$200
Sun .0 -$200
Rain .1 -$2000
Clouds .5 $1000
Sun .4 $5000
Rain .1 -$200
Clouds .5 -$200
Sun .4 -$200
Rain .1 -$2000
Clouds .2 $1000
Sun .7 $5000
Rain .1 -$200
Clouds .2 -$200
Sun .7 -$200
Schedule game
Cancel game
Forecast rain .1
Forecast cloudy .3
Forecast sunny .6
Use weather report
Do not use weather report
Rain .2 -$2000
Clouds .3 $1000
Sun .5 $5000
Rain .2 -$200
Clouds .3 -$200
Sun .5 -$200
- Decision nodes
- Outcome nodes

Demand modeling

- Commercial electric demand