STOCHASTIC MODELS IN NATURAL RESOURCES

Andrés Weintraub P.
Departamento de Ingeniería Industrial
Universidad de Chile

ICSP 2013
SP XIII
Bergamo, July 2013
Note:

Not a rigorous state of the art survey.

Related to papers in the session.

Will concentrate on Forestry problems.
Work with

- Antonio Alonso
- Laureano Escudero
- Monique Guinard
- Jean Paul Watson
- Roger Wets
- David Woodruff
Students:

- Jorge Alarcón
- Fernando Badilla
- Fernanda Bravo
- Jaime Gacitua
- Andrés Croume
- Martín Quinteros
- Ignacio Ríos
- Carlos Villa
- Luis Fernando Solari
- Francisco Castro
- Jorge Alarcon
Agenda

1) Areas to be covered
   a) Forestry
   b) Mining
   c) Aquaculture
      (Agriculture, Water Resources)
2) Deterministic Models
3) Determine uncertainty factors
4) Stochastic Models
5) Solution Approaches
6) Insights, conclusions
2) Basic Notions

a) Need for Efficiency (profits)
b) Scarcity of Resources
c) Concern for Environment
d) Problems with data
   ◦ Current
   ◦ Future

e) Uncertainties
   ◦ Markets
   ◦ Yields
   ◦ Time to execute task (transportation)
   ◦ Costs
   ◦ Natural Phenomena
     • Climate Change
     • Fires
     • Pests
3) The Deterministic approach

a. Problems that are modeled
b. Solution Algorithms
c. Applications and case studies
4) Stochastic Models

a) Representation of uncertainty in problems.
   Will represent most cases through scenarios.
b) Models and solution algorithms.
c) Analysis of solutions and insights.
d) Applications and case studies
e) Open problems
Forestry

Deterministic Models:
1) Strategic, Long term planning

Main decisions:
When to harvest each area $i$

$X_{ij}$: number of hectares managed with plan $j$.

Plan $j$: Harvest years 17,39

$\alpha_{ij}^t$: m$^3$/Ha if area $j$ is harvested in year $t$, under plan $j$. 
Main elements of model

\[
\begin{align*}
\text{Max} & \sum_i \sum_j e_{ij} x_{ij} \quad \text{(net returns)} \\
\sum_j x_{ij} & \leq A_i \quad \text{(area of unit i A_i)} \\
\sum_i \sum_j a_{ij}^t \cdot x_{ij} & = Vol^t \quad (m^3 \text{ harvested in t}) \\
Vol^t & \geq \text{lower Bound} \\
\text{Easy LP to solve used extensively. Usually run in a rolling horizon scheme}
\end{align*}
\]

One uncertainty considered
Yields (timber growth)

\[
a_{ij}^t \approx N(a_{ij}^t, \sigma_{ij}^t)
\]

worried about Lower Bound of timber demand.
New constraints

\[ P(\sum_{i} \sum_{j} a_{ij}^t \cdot x_{ij} \geq \text{Lower Bound}^t) \geq .95 \]

Chance constraints
Can be transformed into a deterministic nonlinear program (Weintraub & Vera 1991, Hof et al, 1993).
What can go wrong if use just average growth in a deterministic model?
May not have feasible solutions in later periods if growth is small and yields similar to demand.
Insighth

Is it worth to define a stochastic model? Only if supply is tight, close to demand. If not, forest has flexibility to recover. (Weintraub & Abramovich).
Fires

At strategic harvesting level.
Harvesting can act as fire blocking.

| A | C | B |

Fire start at A, expands directly to B (valuable timber).
If unit C is harvested, the fire goes in direction of B, but slower to reach B (alternate Path)

Heuristic Approach (Palma et. Al. Acuña et. Al)
- A random simulation model to evaluate fire consequences of harvest policies.
- A model to assign protection values to units.
- A harvesting model that includes protection values.
Increases productivity in later periods.
Tactical Models

Specific problem
Model 5 year horizon

Decisions:
  Units to harvest
  Roads to build for access
  Flows of timber
  (Andalaft et 2003)
17 independent forests, linked by demand constraints. Cannot be solved with CPLEX 3.0. Use Lagrangean relaxation and Strengthening of the LP formulation, used by firm 1990’s. CPLEX 9.0 has no problems.
Introduce Uncertainty in Markets
16 Scenarios (coordinated branching)
$T = \{1, 2, 3, 4\}; \ T^\circ = \{1, 2, 3\}$

$\Omega = \Omega^1 = \{10, 11, \ldots, 17\}$

$\Omega^2 = \{10, 11, 12\}; \ \Omega^4 = \{15, 16, 17\}$

$G^2 = \{2, 3, 4\}; \ G^3 = \{5, 6, 7, 8, 9\}$

Figure 1: Scenario Tree
Scenario : Expression of uncertain parameters for each period through Horizon.
The non-anticipativity constraints.
If scenario i and j are identical up to period t, then decisions under scenarios i and j must be equal up to period t. They share same information.
So, at node 2, decisions under scenarios 10, 11, 12, must be identical
First Challenge

How to generate scenarios, with probabilities.
One possibility: By logic, experience.
More rigorous: Simulate parameters as a random walk with convergence to a mean value, use historical statistics.

Other issues
How many scenarios needed to capture well uncertain results?.
More scenarios in area of concern e.g. low prices.
Can reduce scenarios by bundling.
Can use Cluster Analysis to determine similar bundles.
Solve each bundle exactly lead to subraph.
Solve set of subgraphs via PH.
Talk by Ignacio Rios and Roger Wets on copper price scenarios.
Talk by David Woodruff on Bundling in Mining.
Solution Approaches:

The non-anticipativity: Constraints make the problem hard to solve if large number of scenarios.

Direct Solution (Extensive Form)
Descomposition Approaches:

a) Coordinated Branching (Alonso Escudero).

Talk on Mining
Generate a Branching tree for each scenario
Branch on the same variable in each tree

![Tree Diagram]

**Variables**
- $x_1$
- $x_2, x_2$
- $x_3$

**BF tree $\mathcal{R}^1$**
- $x_1 = 0$
- $x_1 = 1$
- $x_2 = 0$
- $x_2 = 1$
- $x_3 = 0$
- $x_3 = 1$

**BF tree $\mathcal{R}^2$**
- $x_1 = 0$
- $x_1 = 1$
- $x_2 = 0$
- $x_2 = 1$
- $x_3 = 0$
- $x_3 = 1$

**BF tree $\mathcal{R}^3$**
- $x_1 = 0$
- $x_1 = 1$
- $x_2 = 0$
- $x_2 = 1$
- $x_3 = 0$
- $x_3 = 1$

**Variable Branching Order:** $x_1, x_2, x_2, x_3$

Figure 3: Branch and Fix Coordination
Table 2: Wood prices and corresponding lower (LB) and upper (UB) demand bounds, by period

<table>
<thead>
<tr>
<th>Sce. num.</th>
<th>P1 US$</th>
<th>LB m3</th>
<th>UB m3</th>
<th>P2 US$</th>
<th>LB m3</th>
<th>UB m3</th>
<th>P3 US$</th>
<th>LB m3</th>
<th>UB m3</th>
<th>P4 US$</th>
<th>LB m3</th>
<th>UB m3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>60</td>
<td>27000</td>
<td>50000</td>
<td>65</td>
<td>28000</td>
<td>52000</td>
<td>68</td>
<td>25000</td>
<td>50000</td>
</tr>
<tr>
<td>2</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>60</td>
<td>27000</td>
<td>50000</td>
<td>65</td>
<td>28000</td>
<td>52000</td>
<td>57</td>
<td>20000</td>
<td>51000</td>
</tr>
<tr>
<td>3</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>60</td>
<td>27000</td>
<td>50000</td>
<td>55</td>
<td>26000</td>
<td>50000</td>
<td>62</td>
<td>25000</td>
<td>48000</td>
</tr>
<tr>
<td>4</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>60</td>
<td>27000</td>
<td>50000</td>
<td>55</td>
<td>26000</td>
<td>50000</td>
<td>50</td>
<td>26000</td>
<td>49000</td>
</tr>
<tr>
<td>5</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>60</td>
<td>27000</td>
<td>50000</td>
<td>45</td>
<td>24000</td>
<td>44000</td>
<td>58</td>
<td>26000</td>
<td>48000</td>
</tr>
<tr>
<td>6</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>60</td>
<td>27000</td>
<td>50000</td>
<td>45</td>
<td>24000</td>
<td>44000</td>
<td>48</td>
<td>22000</td>
<td>40000</td>
</tr>
<tr>
<td>7</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>45</td>
<td>15000</td>
<td>33000</td>
<td>52</td>
<td>27000</td>
<td>45000</td>
<td>55</td>
<td>21000</td>
<td>50000</td>
</tr>
<tr>
<td>8</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>45</td>
<td>15000</td>
<td>33000</td>
<td>52</td>
<td>27000</td>
<td>45000</td>
<td>44</td>
<td>19000</td>
<td>42000</td>
</tr>
<tr>
<td>9</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>45</td>
<td>15000</td>
<td>33000</td>
<td>45</td>
<td>25000</td>
<td>41000</td>
<td>50</td>
<td>20000</td>
<td>36000</td>
</tr>
<tr>
<td>10</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>45</td>
<td>15000</td>
<td>33000</td>
<td>45</td>
<td>25000</td>
<td>41000</td>
<td>42</td>
<td>15000</td>
<td>30000</td>
</tr>
<tr>
<td>11</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>45</td>
<td>15000</td>
<td>33000</td>
<td>30</td>
<td>20000</td>
<td>32000</td>
<td>48</td>
<td>15000</td>
<td>28000</td>
</tr>
<tr>
<td>12</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>45</td>
<td>15000</td>
<td>33000</td>
<td>30</td>
<td>20000</td>
<td>32000</td>
<td>42</td>
<td>13000</td>
<td>24000</td>
</tr>
<tr>
<td>13</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>30</td>
<td>10000</td>
<td>18000</td>
<td>45</td>
<td>22000</td>
<td>40000</td>
<td>40</td>
<td>12000</td>
<td>18000</td>
</tr>
<tr>
<td>14</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>30</td>
<td>10000</td>
<td>18000</td>
<td>45</td>
<td>22000</td>
<td>40000</td>
<td>30</td>
<td>12000</td>
<td>18000</td>
</tr>
<tr>
<td>15</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>30</td>
<td>10000</td>
<td>18000</td>
<td>30</td>
<td>15000</td>
<td>25000</td>
<td>40</td>
<td>13000</td>
<td>18000</td>
</tr>
<tr>
<td>16</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>30</td>
<td>10000</td>
<td>18000</td>
<td>30</td>
<td>15000</td>
<td>25000</td>
<td>30</td>
<td>12000</td>
<td>17000</td>
</tr>
<tr>
<td>17</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>30</td>
<td>10000</td>
<td>18000</td>
<td>22</td>
<td>12000</td>
<td>22000</td>
<td>35</td>
<td>11000</td>
<td>16000</td>
</tr>
<tr>
<td>18</td>
<td>45</td>
<td>30000</td>
<td>40000</td>
<td>30</td>
<td>10000</td>
<td>18000</td>
<td>22</td>
<td>12000</td>
<td>22000</td>
<td>20</td>
<td>10000</td>
<td>15000</td>
</tr>
</tbody>
</table>
Consider 2 decision makers (dm)

Figure 5: Probabilities for Equiprobable Configuration
The deterministics DM considers the average value of uncertain parameters and maximizes net return.

The stochastic DM, maximizes expected return, such that the solution is feasible under all scenarios (could be 98%).

Consider three sets of scenarios:
- higher probability for scenarios with higher prices, for lower prices, equiprobable
Table 4: Equiprobable Scenarios

<table>
<thead>
<tr>
<th></th>
<th>$Z_{AVSC}$</th>
<th>$Z_{BFC}$</th>
<th>ABS GAP</th>
<th>REL GAP %</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC 1</td>
<td>7860376.2</td>
<td>8141684.5</td>
<td>281308.3</td>
<td>3.6</td>
</tr>
<tr>
<td>SC 2</td>
<td>74986706.4</td>
<td>7832291.9</td>
<td>335585.6</td>
<td>4.5</td>
</tr>
<tr>
<td>SC 3</td>
<td>7272765.9</td>
<td>7681257.3</td>
<td>408491.9</td>
<td>5.6</td>
</tr>
<tr>
<td>SC 4</td>
<td>6876035.1</td>
<td>7248863.7</td>
<td>372828.7</td>
<td>5.4</td>
</tr>
<tr>
<td>SC 5</td>
<td>6751277.3</td>
<td>7288986.5</td>
<td>537709.3</td>
<td>8.0</td>
</tr>
<tr>
<td>SC 6</td>
<td>6420668.3</td>
<td>6913420.5</td>
<td>492752.3</td>
<td>7.7</td>
</tr>
<tr>
<td>SC 7</td>
<td>6440966.1</td>
<td>6739744.0</td>
<td>298777.9</td>
<td>4.6</td>
</tr>
<tr>
<td>SC 8</td>
<td>6077296.2</td>
<td>6359584.0</td>
<td>282287.8</td>
<td>4.6</td>
</tr>
<tr>
<td>SC 9</td>
<td>6003190.1</td>
<td>6111671.1</td>
<td>108481.1</td>
<td>1.8</td>
</tr>
<tr>
<td>SC 10</td>
<td>Infeasible</td>
<td>5604078.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SC 11</td>
<td>Infeasible</td>
<td>4945591.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SC 12</td>
<td>Infeasible</td>
<td>4541990.9</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SC 13</td>
<td>Infeasible</td>
<td>4324647.4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SC 14</td>
<td>Infeasible</td>
<td>4149814.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SC 15</td>
<td>Infeasible</td>
<td>3335188.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SC 16</td>
<td>Infeasible</td>
<td>3067968.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SC 17</td>
<td>Infeasible</td>
<td>2866035.9</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SC 18</td>
<td>Infeasible</td>
<td>2593300.9</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$tt$</td>
<td>-</td>
<td>11942</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$Z_{IP}$</td>
<td>-</td>
<td>5541451.0</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Progressive Hedging (PH) (Rockafeller and Wets, 1991).

In this case, problem is decomposed by scenarios by relaxing (Penalized) the non anticipativity constraints (Woodruff, Wets, Watson).

If all variable continuous, algorithm converges. With integer variable no convergence proof.
Use Enhancements in implementatives of heuristics

- Linearize quadratic term in penalized objective function.

- Parameters need careful analysis.

- Mid gap tolerance, no need to solve each scenario problem to optimality. Small gap for termination leads to much faster solutions, no significant loss of optimality.
- Hot Starts solution of scenario problem in at itineration is used as starting solution for scenario problem at itination (n+1)

- Strengthening of formulation to make run time faster.

- Variable Fixing: If a variable satisfies non-anticipativity in one or several iterations e.g. in all scenarios, the variable has value 0 or 1, fix it.

- Suboptimal heuristic, no guarantee that is the value in optimal solution
Early Termination after a small (6-15) number of itinerations of PH, move to commercial MIP. Multiple fixed variables make problem easier.

Test case the same forest with uncertainty in prices and demands. Add uncertainty in growth. Run up to 324 scenarios.

Test to determine best enhancements.

- Hotstarts provide 11% faster solution.
- Gap 2% best compromise between computational effort and solution quality.
- Linearize quadratic penalization term in to just two segments.

- Number of PH itinerations not more than 10
<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Cplex EF value (1hr)</th>
<th>Cplex EF gap (1hr)</th>
<th>PH+ value</th>
<th>PH+ gap</th>
<th>PH+ Run time</th>
<th>PH+ vs. Cplex % value</th>
<th>Total Fixed Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>$4,928,180</td>
<td>0.31%</td>
<td>$4,920,078</td>
<td>0.47%</td>
<td>4m22s</td>
<td>-0.16%</td>
<td>2</td>
</tr>
<tr>
<td>64</td>
<td>$5,357,780</td>
<td>1.29%</td>
<td>$5,386,971</td>
<td>0.74%</td>
<td>13m1s</td>
<td>0.54%</td>
<td>1719</td>
</tr>
<tr>
<td>144</td>
<td>$5,266,830</td>
<td>2.12%</td>
<td>$5,287,935</td>
<td>1.68%</td>
<td>21m42s</td>
<td>0.40%</td>
<td>5051</td>
</tr>
<tr>
<td>162</td>
<td>$5,187,040</td>
<td>3.10%</td>
<td>$5,242,032</td>
<td>1.98%</td>
<td>22m54s</td>
<td>1.05%</td>
<td>823</td>
</tr>
<tr>
<td>216</td>
<td>$5,332,550</td>
<td>3.92%</td>
<td>$5,437,714</td>
<td>1.87%</td>
<td>37m11s</td>
<td>1.93%</td>
<td>1828</td>
</tr>
<tr>
<td>324</td>
<td>$5,545,260</td>
<td>2.91%</td>
<td>$5,536,196</td>
<td>2.99%</td>
<td>71m39s</td>
<td>-0.16%</td>
<td>740</td>
</tr>
</tbody>
</table>

Table 3
Results from 18 to 324 scenarios

Note:
CPLEX 10.2 runs for one hour
PH gives slightly better solutions except for 18 and 324 scenarios (very slightly worse)

PH+ much faster than CPLEX, but as number of scenarios gets larger, PH gets slower.

Needs to run scenarios in parallel
Challenge:

How to program effectively to run scenarios in parallel. Some results in one of the presentations (mining problem)
Lagrangean Relaxation

Can relax non-anticipativity constraints penalized.

Solve each escenario separately.

Get bound on optimal solution

Use heuristics to force non-anticipativity
At operational level

PULP

SAWTIMBER

EXPORT
Typical decisions.
Short term Harvesting.
Given demands for specific logs (length, width), obtained from trees.

Know units that can be harvested, and inventory in each.
Decisions:
- How many hectares to harvest in each unit.
- How to cut trees

Gabarro et.al: LP model with Colum Generation.
Model Assigns well trees to demands. Timber “wasted” goes from 15% to 1%.

Uncertainty (generally low, short horizon).
  ◦ Demands may change
  ◦ Sale prices
  ◦ Inventory may not be accurate
  ◦ Production problems
  ◦ Fires.

No explicit consideration of uncertainty in harvesting models. (Difficult models, not really needed).

Just rolling horizon
  ◦ Forest is flexible
  ◦ Have safety stocks
Daily Scheduling :_ Monthly planning
Daily Scheduling used successfully (Chile, Brazil, South Africa).
Uncertainty
  ‣ Travel times
  ‣ Machine, truck breakdowns
  ‣ Road interruptions due to weather
  ‣ Production uncertainty
No explicit models
  ‣ Fat solutions
  ‣ Add “extra” time for travel
  ‣ Have extra loaders, trucks
Environmental issues

Expressed as spatial constraints
Maximum opening size (adjacency)

Large areas of mature timber

Difficult to solve combinatorial problems. In many cases use heuristics (Tabu search, Caro et.al)
Large areas of mature timber
Uncertainty here more related to fuzzy constraints. Are these constraints really what planners want?

Operational constraints

- Not harvest near rivers (to reduce sediment)
- Not use tractors on fragile soils (erosion)
- Leave grown trees alongside roads (scenic beauty)

These rules related to inputs, impacts not well known.
How to reserve forest areas to maximize wildlife protection (Hof & Haight, 2009). Uncertainty in dispersal to neighboring areas, survival rates.

Approaches: MIP Simulation (for complex formulations of probabilities).
Climate change

Work in progress (Jordi Garcia)
Relate climate scenarios to timber yields.

Develop scenarios from climate change to timber availability.
Mining

Open pit and underground mines. Deterministic MIP models are used for long range mine extraction (Newman et. Al. Epstein et. al)

Open pit
Underground
MINERIA
Underground Mine

PRODUCTION

Block caving

LHD

Gravity (Tunnel)

Drain stage

ROCK SIZE REDUCTION

MAIN TRANSPORTATION (Train)

PLANTS (MILLS)

COMMERCIAL PRODUCTS

Gravity (Tunnel)
Uncertainties in mining

All papers are on mining.

- Prices

Future Uncertainties: Prices could be 2.5 to 6 $/pd.
  - % copper (grade)
  - Costs (energy) (Price related to costs).

Growing interest in incorporating uncertainty explicitly.

Codelco saved copper due to low prices.
Uncertainty in ore grade
Gholannejad and Osanloo 2007.
Block grade uncertainty.
Each block has probability distribution obtained through geostatistical simulation.
Transform probabilities of satisfying constraints into non-linear deterministic integer problem.
No computational results.
Aqualculture (Fishing)
Typical Uncertainties.

- Biomas (existing fish)
- Stock growth (mortality)
- Catch results
- Prices

Look for optimal harvest.
Use Dynamic Stochastic Programming
Can be seen as a Markow process
Determine optimal Harvest policies

Need to analyze if introduction of uncertainty modifies significantly decisions.
Salmon Farming

- Project with multiexport, large salmon production firms in Chile (Bravo et. al).
- Develop Supply Chain (Bravo et. Al)
- Deadly virus.
Salmon Farming

Fuente: www.multiexport.com
Uncertainties

- Markets: Prices – Demands (Production chain: 3 years).
- Chile Produced 40% of total world production.
- Costs
- Environmental constraints (use of lakes, seashore forbidden, contamination).
Chose as exercise virus

Scenarios: Fish surviving each Stage

Decision – How many fish need at start of chain.

- When to move fish to next stage.
Solution Approaches

1. Static Deterministic
2. Wait and see (Rolling Horizon)
3. Stochastic (Scenario Tree)
4. Perfect information
   Survival rates 95% to 99% in each of 4 stages.

Deterministic Model
60,000 Variables (13 0–1 variables)
Solve in seconds)

Stochastic Model (Extensive form) 950,000 variables
200,000 0–1 variables
Solves in 1,7 hours, GAP : 2.63%.
Main Results

Stochastic vs Deterministic
For Higher levels of survival. Deterministic gains up to 9%
For Medium levels stochastic gains up to 17%
For Low Levels Deterministic not feasible
Wait and see also leads to infeasibilities (less)

Perfect information leads to 7% better solution.
Conclusions

Why more intensive development of stochastic models.

1. Development of new algorithms
2. Power of hardware, software allows to solve larger problems
3. Some cases (mining) industry needs

Difficulties in implementation rigorously.
Generation of scenarios solution algorithms for large problems.

Not really in use so far in forestry, mining, fishing, (unlike other areas, like energy).