Introducing Uncertainty in Brazil's Oil Supply Chain

Juan Pablo Luna (UFRJ)
Claudia Sagastizábal (IMPA visiting researcher)
on behalf of OTIM-PBR team

Workshop AASS, April 1st 2016
OTIM-PBR: main features

General Goal: Introduce uncertainty in “PlanAb”, Petrobras’ planning tool for managing its supply chain (monthly decisions over half a year).

Specific goals:

• Representation of stochastic data, focusing on price and volume risk (“statistical” block).
• Definition of a risk-neutral model and direct solution of a small instance.
• Definition of a protocol for evaluating the results.
• The ultimate tool, for huge-scale problems, introducing a decomposition method.

Academic team: IMPA/UFRJ/UERJ/UFSC (10 persons), and a guest from NTNU Norway.
A Multidisciplinary Team

- IMPA: Mikhail Solodov, Jorge Zubelli
- UFRJ: Laura Bahiense, Carolina Effio, José Herskovits, Juan Pablo Luna
- UERJ: Welington de Oliveira
- UFSC: Marcelo Córdova, Erlon Finardi
- NTNU: Asgeir Tomassgard

- PETROBRAS:
  - Paulo Ribas (Supply chain department)
  - Flavia Schittine (OR department)
  - Sergio Bruno (Corporate risk department)
OTIM-PBR: main features

General Goal: Introduce uncertainty in “PlanAb”, Petrobras’ planning tool for managing its supply chain (monthly decisions over half a year).

Specific goals:

- Representation of stochastic data, focusing on price and volume risk (“statistical” block).
- Definition of a risk-neutral model and direct resolution of a small instance.
- Definition of a protocol for evaluating the results.
- The ultimate tool, for huge-scale problems, introducing a decomposition method.

Academic team: IMPA/UFRJ/UERJ/UFSC (10 persons), and a guest from NTNU Norway.
Computational model

Block 1: Statistical
Tool generating scenarios for:
  - International prices (oil and derivatives, correlated)
  - Oil volume arriving from the platforms each month.

Block 2: Optimization
Tool to determine the optimal policy for the whole supply chain of the company for the next month. It includes:
  - To which extent the company network (of production, transportation, commercialization) can be simplified?
  - How to handle uncertainty in the optimization problem (in the cost and in the right hand side)
  - Problem solution via decomposition method (huge scale problem
Computational model

Block 1: Statistical

Tool generating scenarios for:
- International prices (oil and derivatives, correlated)
- Oil volume arriving from the platforms each month.

Block 2: Optimization

Tool to determine the optimal policy for the whole supply chain of the company for the next month. It includes:
- To which extent the company network (of production, transportation, commercialization) can be simplified?
- How to handle uncertainty in the optimization problem (in the cost and in the right hand side)
- Problem solution via decomposition method (huge scale problem
Computational model

Block 1: Statistical

Tool generating scenarios for:

- International prices (oil and derivatives, correlated)
- Oil volume arriving from the platforms each month.

Technique: multivariate model + Kalman filter + SDP NLP solver, assessed by backtesting (Juan Pablo Luna).

Block 2: Optimization

Tool to determine the optimal policy for the whole supply chain of the company for the next month. It includes:

- To which extent the company network (of production, transportation, commercialization) can be simplified?
- How to handle uncertainty in the optimization problem (in the cost and in the right hand side)
- Problem solution via decomposition method (huge scale problem
The setting

Wide variety of commercial, industrial and logistics operations occur over the midstream segment.

Strong dependence among the operations and some gains can only be estimated by considering the whole supply chain.

Planning this segment is crucial to achieve the success of all operations.

Source: M. Maia, Petrobras
Brazil's supply chain

Source: M. Maia, Petrobras

- 23000 km pipelines
- >100 platforms
- > 40 terminals in Brazil
- >200 terminals abroad
- 12 refineries

Crude Oil:
- Imports (1)
- Exports (3)
- Production (5)

Oil Products:
- Imports (2)
- Exports (4)
- Market Selling (6)

~ 200 different crude oils
~ 50 different oil products

10 basins of oil production
Production flow PlanAb

Simplified Network

INTERNATIONAL MARKET (buy)

PLANT

INTERNATIONAL MARKET (sale)

DEMAND (Brazil)
Production flow
PlanAb **UNCERTAINTY**

**Simplified Network**

INTERNATIONAL MARKET (buy)

INTERNATIONAL MARKET (sale)

INTERNATIONAL MARKET (Brazil)

**Oil extraction**
Some comments

• Uncertainty increases the (already huge) size of the linear program (LP).

• Mounting the data for the deterministic model takes longer than solving the LP.

• Solving an aggregate model is not an option: keeping the level of detail of the deterministic model is important for the company.
Some comments

- Uncertainty increases the (already huge) size of the linear program (LP).

- Mounting the data for the deterministic model takes longer than solving the LP.

- Solving an aggregate model is not an option: keeping the level of detail of the deterministic model is important for the company.

**BUT**

- Changing CPLEX stopping tolerance from default $10^{-8}$ to $10^{-4}$ provides a good trade-off between accuracy and solving time: mean relative error on variables 0.007% solving times reduced in 18.56%
Statistical Block

OTIM-PBR
Sources of uncertainty
Calibration of price models

1. Oil, gasoline and diesel international prices
   - The model is a process defined by stochastic differential equations.
   - To determine the parameters defining the model (mean, trend, volatility, correlations) we maximize the proximity to historical data (likelihood) using a Kalman filter.
Sources of uncertainty

Calibration of price models

- There are several stochastic processes for modelling commodity prices.
- They should include important oil price features such as picks, seasonality, mean reversion, etc.
- The models are completely described by certain parameters that can be time dependent (mean, trend, volatility, correlations) and whose values need to be estimated (calibration).
Sources of uncertainty

Calibration of price models

- Agenda:
  1. Choose a model suitable for our purposes.
  2. Calibrate the model for fitting the available historical data.
  3. Generate future price scenarios.
Sources of uncertainty
Calibration of price models

- **Calibration:**
  - To determine the parameters defining the model we maximize the proximity to historical data (likelihood).
  - To compute the likelihood requires computing the joint probability density function that may not be available in closed form but can be estimated through Kalman Filters.
  - The likelihood function is nonconvex.
Schwartz - Smith Model

\[ \log(S_t) = a \chi_t + b \xi_t \]

\[ d\chi_t = -\text{diag}(\kappa) \chi_t dt + A^\chi dW_t^\chi \]

\[ d\xi_t = \mu dt + A^\xi dW_t^\xi \]

\[ dW_t^\chi dW_t^\xi = \hat{\rho}^\chi^\xi \]

\[ \delta = (a, b, \kappa, A^\chi, A^\xi, \lambda^\xi, \mu^\xi, \mu^*^\xi, \hat{\rho}^\chi^\xi) \]
Future Contract Price

\[ F^i_{t,T} = \mathbb{E}^* \left[S^i_T \mid \mathcal{F}_t \right] \]

\[
\log(F^i_{t,T}) = a \left[ e^{-k_i(T-t)} \chi^i_t \right] + b \xi^i_t + A^i(t,T)
\]

\[
A^i(t,T) = b \left[ (\mu_i - \lambda \xi_i)(T-t) \right] - a \left[ \frac{\lambda \chi_i}{k_i} (1 - e^{-k_i(T-t)}) \right]
\]

\[
+ \frac{1}{2} \left( (1 - e^{-2k_i(T-t)}) \left\| a e_i^\top A \chi \right\|^2 \right.
+ 2(1 - e^{-k_i(T-t)}) \frac{a b e_i^\top A \chi \hat{\rho} \chi \xi A \xi^\top e_i}{k_i}

\left. + \left\| b e_i^\top A \xi \right\|^2 (T-t) \right)
\]

Note the nonlinear relations between unknown variables
Discretization of Schwartz – Smith Model

\[
\begin{bmatrix}
\chi_{t+1} \\
\xi_{t+1}
\end{bmatrix} = \mathbb{E} \begin{bmatrix}
\begin{bmatrix}
\chi_{t+1} \\
\xi_{t+1}
\end{bmatrix} | \mathcal{F}_t
\end{bmatrix} + w_t
\]

\[
\mathbb{E} \begin{bmatrix}
\begin{bmatrix}
\chi_{t+1} \\
\xi_{t+1}
\end{bmatrix} | \mathcal{F}_t
\end{bmatrix} = \begin{bmatrix}
\text{diag}(e^{-k\Delta t}) \chi_t \\
\xi_t + \mu \Delta t
\end{bmatrix}
\]

\[
\text{Cov}(w_t) = \begin{bmatrix}
\tilde{\Sigma} \chi & \text{diag} \left( \left( \frac{1-e^{-k_i \Delta t}}{k_i} \right)_i \right) A \hat{\chi} \hat{\chi}^T A^T \\
A^T (\hat{\chi} \hat{\chi}^T) \Delta t \Sigma^\xi
\end{bmatrix}
\]

\[
\tilde{\Sigma}_{ij} = \frac{1 - e^{-(k_i + k_j)\Delta t}}{k_i + k_j} \Sigma_{ij}
\]

We need to keep positive definite this nonlinear matrix
Implementation Issues

1. Objective function (evaluated through a Kalman Filter) is highly nonlinear.
2. Objective function is defined only on certain set: feasible optimization methods must be used.
3. Nonlinear constraints must include the positive definiteness of correlation matrices.
4. Optimization methods are highly sensitive to gradient values: an accurate implementation is needed.
Numerical Results

• WTI, HO, RBOB future contracts (with 1 to 4 months of maturity) from Energy Information Administration-EIA. (http://www.eia.gov/dnav/pet/xls/PET_PR_FUT_S1-D.xls)

• Time period considered 1081 days (10/06/2011 to 22/09/2015)

• Three decks (1081, 720 and 359 days) for calibrating Schwartz – Smith.
Numerical Results

- We considered 1D and 3D Schwartz – Smith processes.

- The optimization problems were solved using FDIPA and FDIPA-SDP non linear programming algorithms.

- Numerical approximations of likelihood functions versus its exact evaluation.
Cenários Futuros (Caso 1081) RBOB 3D

Cenários Futuros (Caso 362-1081) RBOB 3D

Cenários Futuros (Caso 723-1081) RBOB 3D

RBOB 3D
Conclusions

• As expected, considering together all three assets produces better simulations of prices.

• Considering Schwartz – Smith models for more than one dimension leads to more challenging numerical problems that need sophisticated optimization solvers (and expertise from the optimization community).
Optimization Block

OTIM-PBR
Supply chain planning tool
Computational model

PlanAb

- Computational model for Petrobras supply chain
- Tool to determine the optimal planning for the whole supply chain of the company for the next month.

Currently PlanAb is a deterministic model that solves a large scale linear program:

\[
\begin{align*}
\min_{x,y} & \quad c^T x + q^T y \\
\text{s.t} & \quad Ax = b \\
& \quad Tx + Wy = h \\
& \quad x \leq x \leq \bar{x} \\
& \quad y \leq y \leq \bar{y} \\
& \quad y_i = \bar{y}_i \quad i \in I_V
\end{align*}
\]
Supply chain planning tool
Computational model

PlanAb

\[
\begin{aligned}
\min_{x,y} & \quad c^\top x + q^\top y \\
\text{s.t} & \quad Ax = b \\
& \quad Tx + Wy = h \\
& \quad x \leq x \leq \bar{x} \\
& \quad y \leq y \leq \bar{y} \\
& \quad y_i = \bar{y}_i \quad i \in I_V
\end{aligned}
\]

- Variable $x$ represents commercial transactions such as imports along the planning horizon (four months)
- Variable $y$ represents exports and logistic operations such as transportation and refinery oil/derivative production

Availability of oil volume from the plataforms
Most relevant output

Commercial transactions

Oil
- Imports
- Exports

Derivatives
- Imports
- Exports

Logistic operations

Refineries
- Type of consumed oil
- Type of produced derivatives

National market
- Which refinery sends products to which market
Objective function
PlanAb model

Current model:
\[
\min_{x,y} \ c^\top x + q^\top y
\]

Model (objective function)

Costs
• Oil imports
• Derivative imports
• Derivative storage
• Refining unit operations
• Transportation

Revenues
• Oil exports
• Derivative exports
• Derivative national sales

Prices are uncertain....

Proposed model:
\[
\min_{x,y(\xi)} \ c^\top x + E[q(\xi)^\top y(\xi)]
\]
1. Oil and derivatives balance
   - Includes all constraints involving the system balance of oil/derivatives and how they flow in the network
   - Balancing the derivatives in each terminal and refinery
     \[ Ax = b \]
     \[ Tx + Wy = h \]
   - Considering adjustments on the available oil for:
     - Processing in the refineries
     - Exportation, when the volume of oil arriving from the platforms is larger than \textit{foreseen}
     \[ y_i \leq y_i \leq \bar{y} \]
     \[ y_i = \bar{y}_i \quad i \in I_y \]
Sources of uncertainty

1. Oil, gasoline and diesel international prices
   (about 30 in total)
   Stochastic process: multivariate Schwartz-Smith
   (Juan Pablo’s talk)

   This uncertainty is represented by $\xi$

2. Availability of national oil
   The ratio between the foreseen and observed
   volumes follows a log-logistic probability
   distribution

   This uncertainty is represented by $\omega$

Two different sources of uncertainty!
PlanAb

- Deterministic

\[
\begin{align*}
\min_{x,y} & \quad c^\top x + q^\top y \\
\text{s.t.} & \quad Ax = b \\
& \quad Tx + Wy = h \\
& \quad x \leq x \leq \overline{x} \\
& \quad y \leq y \leq \overline{y} \\
& \quad y_i = \bar{y}_i \quad i \in I_V
\end{align*}
\]

- Stochastic

\[
\begin{align*}
\min_{x,y} & \quad c^\top x + \mathbb{E}[q(\xi)^\top y(\xi, \omega)] \\
\text{s.t.} & \quad Ax = b \\
& \quad Tx + Wy(\xi, \omega) = h \\
& \quad x \leq x \leq \overline{x} \\
& \quad y \leq y(\xi, \omega) \leq \overline{y} \\
& \quad y_i(\xi, \omega) = \bar{y}_i(\omega) \quad i \in I_V \\
& \quad \forall \xi \in \Xi \quad \omega \in \Omega
\end{align*}
\]

How large are these models?
Deterministic PlanAb

Implemented in AIMMS

- LP’s dimension:
  - 2.3 million of variables
  - 1.7 million of constraints

- Constraints’ matrix:
  - Sparse (0.0002% non-zero elements; approximately 7.3 million of elements)
  - Block-diagonal structure
Deterministic PlanAb

- PlanAb LP:
  - 2.3 million of variables, 1.7 million of constraints
  - 7.3 million of non-zero elements (0.0002% - sparse)
  - After presolve, size reduced approximately by 80%

<table>
<thead>
<tr>
<th>Model</th>
<th>Variables</th>
<th>Constraints</th>
<th>Non-zero elements</th>
<th>Presolve time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>2.3 million</td>
<td>1.7 million</td>
<td>7.3 million</td>
<td>0</td>
</tr>
<tr>
<td>Conservative</td>
<td>644,000</td>
<td>307,000</td>
<td>2.88 million</td>
<td>2.7</td>
</tr>
<tr>
<td>Automatic</td>
<td>385,000</td>
<td>94,000</td>
<td>2.19 million</td>
<td>11.1</td>
</tr>
<tr>
<td>Aggressive</td>
<td>385,000</td>
<td>92,000</td>
<td>2.14 million</td>
<td>13.2</td>
</tr>
</tbody>
</table>
Deterministic PlanAb

- Solution of PlanAb by CPLEX solver, using different methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Time AIMMS (sec)</th>
<th>Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primal simplex</td>
<td>600*</td>
<td>800 MB</td>
</tr>
<tr>
<td>Dual simplex</td>
<td>600*</td>
<td>800 MB</td>
</tr>
<tr>
<td>Network + Primal</td>
<td>600*</td>
<td>770 MB</td>
</tr>
<tr>
<td>Network + Dual</td>
<td>600*</td>
<td>820 MB</td>
</tr>
<tr>
<td>Barrier</td>
<td>180</td>
<td>1.25 GB</td>
</tr>
<tr>
<td>Barrier – Primal crossover</td>
<td>195</td>
<td>1.25 GB</td>
</tr>
<tr>
<td>Barrier – Dual crossover</td>
<td>195</td>
<td>1.25 GB</td>
</tr>
<tr>
<td>Sifting</td>
<td>600*</td>
<td>1.1 GB</td>
</tr>
<tr>
<td>Concurrent</td>
<td>352</td>
<td>1.8 GB</td>
</tr>
</tbody>
</table>

* Limit for time execution
Stochastic PlanAb

\[
\begin{align*}
\min_{x,y} & \quad c^\top x + \mathbb{E}[g(\xi)^\top y(\xi, \omega)] \\
\text{s.t.} & \quad Ax = b \\
& \quad Tx + W y(\xi, \omega) = h \\
& \quad x \leq \bar{x} \\
& \quad y \leq \bar{y} \\
& \quad y_i(\xi, \omega) = \bar{y}_i(\omega) \\
\forall \xi \in \Xi & \quad \omega \in \Omega
\end{align*}
\]

- Problem’s size depends on the number of scenarios of price and oil volume
  - Number of variables is $N$ times 2.3 million
  - Number of constraints is $N$ times 1.7 million

- Impossible to load the problem in a (powerful) computer for $N \geq 10!$ (Memory issues)
Two-stage decomposition
Stochastic PlanAb

- Consider finitely many scenarios of price and oil volume
- The problem is decomposed into two decision levels

\[
\begin{align*}
\min_x & \quad c^T x + \mathbb{E}[Q(x; \xi, \omega)] \\
\text{s.t.} & \quad Ax = b \\
& \quad \underline{x} \leq x \leq \bar{x}
\end{align*}
\]

with

\[
Q(x; \xi, \omega) \left\{ \begin{array}{l}
\min_y & \quad q(\xi)^T y \\
\text{s.t.} & \quad Wy = h - Tx \\
& \quad y \leq \bar{y} \\
& \quad y_i = \bar{y}_i(\omega) \quad i \in I_V
\end{array} \right.
\]
Two-stage decomposition

Stochastic PlanAb

- Consider finitely many scenarios of price and oil volume
- The problem is decomposed into two decision levels

\[
\begin{align*}
\min_x & \quad c^\top x + \sum_{i=1}^N p_i [Q(x; \xi_i, \omega_i)] \\
\text{s.t} & \quad Ax = b \\
& \quad x \leq x \leq \bar{x}
\end{align*}
\]

with

\[
Q(x; \xi, \omega) \left\{ \begin{array}{l}
\min_y & \quad q(\xi)^\top y \\
\text{s.t} & \quad Wy = h - Tx \\
& \quad y \leq \bar{y} \\
& \quad y_i = \bar{y}_i(\omega) \quad i \in I_V
\end{array} \right.
\]
Two-stage decomposition
Stochastic PlanAb

\[
\begin{align*}
\min \ c'x + \alpha \\
\text{st.} \quad Ax &= b \\
\text{cut}^i &< \alpha \\
i &= 1, 2, \ldots, k
\end{align*}
\]
Cutting-plane approximation

\[ E[Q(x; \xi, \omega)] \]

\[ x^1 \]
Cutting-plane approximation

\[ E[Q(x; \xi, \omega)] \]
Cutting-plane approximation

\[ E[Q(x;\xi,\omega)] \]

[Diagram showing points labeled \( X_1, X_3, X_2 \) with lines and curves intersecting at these points.]
Cutting-plane approximation
Cutting-plane approximation

\[ E[Q(x; \xi, \omega)] \]
• N linear programming problems must be solved for every first-stage decision.

• This is a difficult task for large values of N
  • We may solve the LPs in a approximate manner (inexact cuts)
  • Use more efficient cutting-plane methods, such as Bundle Methods:

  \[
  Q(x; \xi, \omega) \begin{cases} 
  \min_{y} & q(\xi)^{\top}y \\
  \text{s.t.} & Wy = h - Tx \\
  & y \leq \overline{y} \\
  & y_i = \overline{y}_i(\omega) \quad i \in I_V
  \end{cases}
  \]

  Convex proximal bundle methods in depth: a unified analysis for inexact oracles.
  W. de Oliveira, C Sagastizábal and C. Lemaréchal. 49
PlanAb with chance-constraints

- One manner to prevent the number of scenarios to be large is to handle the oil volume uncertainty by chance constraints.

\[
Q(x; \xi, \omega) = \begin{cases} 
\min_y q(\xi)^{\top}y \\
\text{s.t.} \\
Wy = h - Tx \\
y \leq \bar{y} \\
y_i = \bar{y}_i(\omega) \\
i \in I_V
\end{cases}
\]

The bounds \( \underline{y}_i^{LL} \), \( \bar{y}_i^{LL} \) are such that \( \mathbb{P}[\underline{y}_i^{LL} \leq y_i(\omega) \leq \bar{y}_i^{LL}] = 1 - p \quad (p \in (0, 1)) \)

Determining the bounds is not a difficult task, since the oil volume of one platform is independent from the other platforms.
The oil volume informed by the platform is only 38% reliable.
PlanAb with chance- constraints

\[ \bar{y}_i(\omega) \sim LL(\beta, \alpha, \gamma) \implies P[\bar{y}_i(\omega) \leq y] = \frac{(y - \gamma)^\alpha}{\beta^\alpha + (y - \gamma)^\alpha}, \text{ for } y \geq \gamma \]
Two-stage decomposition
Stochastic PlanAb + chance constraint
- Consider finitely many scenarios of \textbf{prices}
- The problem is decomposed into two decision levels

\[
\begin{aligned}
\min_x & \quad c^T x + \sum_{i=1}^{N} p_i [Q(x; \xi_i)] \\
\text{s.t.} & \quad Ax = b \\
& \quad x \leq x \leq \bar{x}
\end{aligned}
\]

with

\[
Q(x; \xi) = \left\{ \begin{array}{l}
\min_y & q(\xi)^T y \\
\text{s.t.} & Wy = h - Tx \\
y \leq \underline{y} \\
\underline{y}_i \leq y_i \leq \bar{y}_i \quad i \in I_V
\end{array} \right.
\]

The bounds $\underline{y}_i$, $\bar{y}_i$ are such that $P[\underline{y}_i \leq y_i(\omega) \leq \bar{y}_i] = 1 - p$ \quad (p \in (0, 1))
PlanAb + chance constraint
Conclusions

- Stochastic PlanAb
  - Price scenarios
  - Oil volume
    - can be modelled either by using scenarios or chance-constraints
    - follow independent log-logistic probability distributions

- The computational implementation of the stochastic PlanAb model, with price scenarios and chance-constraints for oil volumes is in progress
Good bye and thank you for coming.