

A TWO-STAGE STOCHASTIC PROGRAMMING MODEL FOR THE MID-TERM OIL REFINERY PLANNING UNDER UNCERTAIN DEMAND

Leonardo Macchiarulo

Federal University of Rio de Janeiro
Av. Athos da Silveira Ramos, 149, Rio de Janeiro - RJ, 21941-909
leomacchiarulo@ufrj.br

Juan Pablo Cajahuanca Luna

Federal University of Rio de Janeiro
Av. Athos da Silveira Ramos, 149, Rio de Janeiro - RJ, 21941-909
jpluna@pep.ufrj.br

Virgílio José Martins Ferreira Filho

Federal University of Rio de Janeiro
Av. Athos da Silveira Ramos, 149, Rio de Janeiro - RJ, 21941-909
virgilio@coppe.ufrj.br

ABSTRACT

This work proposes a two-stage stochastic mixed-integer programming model to deal with the mid-term refinery production planning problem. The model considers that decisions regarding the amount of refinery input to be processed in two production periods must be taken in a first stage, where complete and accurate knowledge of the market demands is not available. In a second stage, decisions regarding flow balancing among units, stock levels and production are taken. The challenge lies in defining the amounts of the different inputs that maximize the expected profit of the refinery. The Expected Value of Perfect Information and the Value of Stochastic Solution are computed with the goal of understanding the stochastic model capabilities and the cost of not considering the uncertain nature of the demand.

KEYWORDS. Refinery Planning. Two-Stage Stochastic Programming. Mixed Integer Programming.

Problem Definition, Mathematical Model, The Value of the Stochastic Program, Case Study, Conclusions, Introduction, Literature Review

1. Introduction

The production plan of an oil refinery must consider that the optimal decision of the processes to be used, and the number of ways to combine them, depends on the crude oil properties and the commercial products. Consequently, there is a trade-off between the use of higher quality and more costly types of crude oil that need less processing, and lower quality and less types of costly oil that need more processing.

Since the oil industry deals with commodities, it is highly influenced by many uncertainties [Hsieh and Chiang, 2001]. For instance, periodical variations of all instances, geopolitics issues, natural disasters, intelligent use of natural resources and local or worldwide economic crisis can impact the demand and/or the price of the oil derivatives. Also, unit problems can affect its production capacity.

The main goal of the work is to present a methodology based on stochastic programming that supports the decision makers in the mid-term oil refinery operations planning under uncertain demand. Within this approach, a two-stage recourse model [Birge and Louveaux, 2011] is formulated considering that decisions regarding the amounts of crude oil types and other refinery inputs are made before the realization of the random demand. In a second stage, when the demand is known, production decisions are taken. The objective of the model is to find the first stage decisions that maximize the expected profit across all scenarios (demand realizations).

As a means to assess the benefits of using a stochastic model instead of a deterministic one, the Expected Value of Perfect Information and the Value of the Stochastic Solution are computed. The former measures the maximum amount the decision maker is willing to pay in order to get accurate information on the future. The latter can be interpreted as the impact of not considering the stochastic nature of the problem.

2. Literature Review

Many studies in the area of refinery operations planning have been elaborated throughout the years. Below, some of them are briefly reviewed.

In the paper of [Liu and Sahinidis, 1996], the authors developed a two-stage stochastic linear programming model, considering products demand and price as the uncertainties. Investment decisions regarding refinery capacity expansion were modelled as first stage variables; supply, production and storage decisions as second stage variables.

[Moro et al., 1998] presented a generalized formulation for the tactical refinery planning problem. The authors considered the nonlinear relations of the property specification equations and tested the model using data from a real refinery.

The paper of [Pinto et al., 2000] proposed a nonlinear programming model for the refinery production planning problem and a mixed integer programming model for the crude oil scheduling problem.

In [Neiro and Pinto, 2005], it was proposed a mixed linear programming model to tackle the multiperiod problem. This model served as a basis for two stochastic programming models that considered fluctuations in the crude oil prices and in the products prices and demand.

The dissertation of [Macchiarulo, 2011] proposed a multistage stochastic programming model to deal with the multiperiod refinery operations planning. In the presented formulation, decisions regarding crude oil selection (type and quantity) must be taken for future consumption in the first stage; production and inventory decisions are taken in the subsequent stages. Due to the resulting large-scale program, the L-Shaped Method [Van Slyke and Wets, 1969] was applied as the main solution method.

In [Coelho, 2013], it was proposed a model for the refinery mid-term planning that considered the nonlinear relations of the property specification equations. The solution method was composed by a combination of heuristic, metaheuristics and exact methods.

[Yang and Barton, 2015] presented a two-stage mixed integer nonlinear model to deal with the integrated crude oil procurement and refinery production planning problem with uncertain crude oil quality. In the first stage, the optimal crude selections and purchase amounts are determined. Optimal operations decisions are taken in the second stage, after uncertainties are realized. Both sets of decisions are taken with the goal of maximizing the expected cost across all scenarios.

[Noh et al., 2018] developed a two-stage stochastic linear programming for the midterm production planning of oil refinery considering uncertain oil price, petroleum product prices, and product demand. In their model, the quantity of crude oil purchased, quantity of finished products sold and production capacity were the first stage decision variables; corrective actions associated with the product demand shortfall and/or surplus were the second stage decisions.

In the paper of [Boucheikhchoukh et al., 2022], it was proposed a mixed integer linear multiperiod refinery planning model considering the use of inventory tanks to mitigate the impact of process unit shutdowns.

[Zhang et al., 2024] presented a multistage mixed integer linear programming model to deal with the retrofit problem of a conventional fossil-based refinery into a hybrid biomass-based refinery for a planning horizon of 10 years.

3. Problem Definition

The focus of this work is the planning of the production operations under uncertain market demand. Moreover, the objective is to find the quantity of each refinery input to be processed, the stock levels, the production, and the flow balancing for each unit, campaign¹ and time period that maximizes the refinery expected operations profit. The refinery inputs can be types of crude oil, natural gas and processed products.

The figure 1 presents these concepts in a generalized input-output flow model of a unit. In the figure, the load Q in the campaign C of the unit U is composed by the input streams $i \in \{I_1^n, \dots, I_n^n\}$ produced in the campaigns $c \in \{C_1^n, \dots, C_n^n\}$ of the units $u \in \{U^1, \dots, U^n\}$. Considering the unit U as a processing unit and that there is no material loss, the load Q in the campaign C of the unit U must be equal to the sum of the quantities of its output streams $o \in \{O_1^n, \dots, O_n^n\}$. In case the unit U is a storage unit, the load Q in the campaign C of the unit U must be equal to the sum of the quantities of its output streams and the stored quantity. Finally, the output streams $o \in \{O_1^n, \dots, O_n^n\}$ will feed other units and campaigns accordingly or be delivered to the market.

In this work, the refinery operations planning problem is formulated as a two-stage stochastic programming with recourse [Birge and Louveaux, 2011]. Within this framework, one must decide in the first stage the inputs and their quantities to be consumed, before the realization of the market demand. In the second stage, when products demand is accurately known, decisions regarding the production must be taken for a set of periods. A generalized formulation for the two-stage stochastic programming problem with recourse² can be written as follows:

$$\begin{aligned} & \min_x \{c^T x + E_P Q(x, \xi(\omega))\} \\ & \text{Subject to:} \\ & Ax = b \\ & x \in X^{n_1} \end{aligned} \tag{1}$$

In the problem above, c is a vector of size n_1 , b a vector of size m_1 and A a matrix $m_1 \times n_1$. The vector x , with domain X^{n_1} , correspond to the first stage decisions and are independent from

¹A campaign is a process that defines not only the input streams, but also the operations variables of the unit, thus modifying the quantities generated of the output streams and their properties.

²The stochastic programming was firstly introduced in [Dantzig, 1955] and [Beale, 1955].

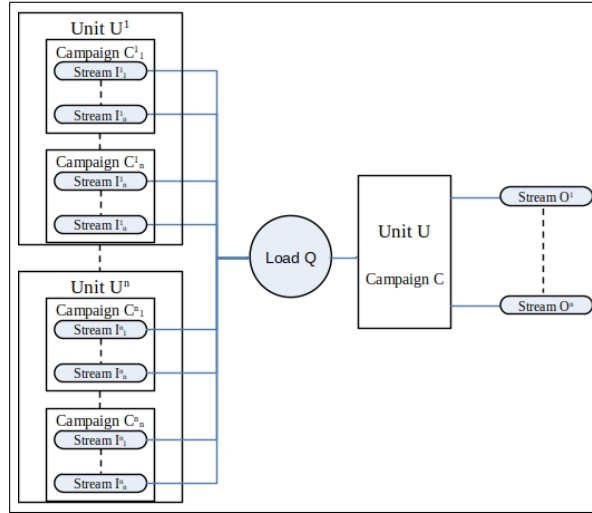


Figure 1: Unit input-output flow model

random vector $\omega \in \Omega$. The objective function minimizes the first stage decisions cost $c^T x$ and the expected value of the second stage decisions cost $E_P Q(x, \xi(\omega))$ with respect to the probability distribution P . By considering $\xi(\omega) = (q(\omega), h(\omega), T(\omega), W(\omega))$, a vector of size at most $N = \{n_2 + m_2 + m_2 \times n_1 + m_2 \times n_2\}$, the second stage problem can be defined as:

$$Q(x, \xi(\omega)) = \{\min_y q(\omega)^T y : W(\omega)y = h(\omega) - T(\omega)x, y \in Y^{n_2}\} \quad (2)$$

In the second stage problem, $q(\omega)$ is a vector of size n_2 , $W(\omega)$ is the recourse matrix of size $m_2 \times n_2$, the right-hand side $h(\omega)$ is a vector of size m_2 and $T(\omega)$ is the technology matrix $m_2 \times n_1$. The second stage decisions are represented by the vector y with domain Y^{n_2} and are taken only after the realization ω

4. Mathematical Model

In this section, the two-stage stochastic programming model is presented in its deterministic equivalent form, where each realization of the random demand represents a scenario. Despite resulting in an increase in the problem size, this approach avoids dealing with integrals and permits the use of well known mathematical optimization techniques. In summary, the model considers the following assumptions:

- There is no material loss inside or between units.
- Blending rules are given by linear relations and stream properties are constants.
- “Feed units” are responsible for receiving crude oil and other inputs and supplying them to storage units.
- “Delivery units” are responsible for receiving finished products from storage units and supplying them to the market.
- Market demand is the only source of uncertainty and its probability distribution is known.

Regarding the second assumption, it is considered that all the streams properties are constants and a linear blending rule is known despite the nonlinear nature of property specification. This approach not only avoids the use of nonlinear equations, but it is also commonly applied in the petroleum refining business through blending indices table [Al-Qahtan and Elkamel, 2010].

4.1. Nomenclature

Sets

U : Set of storage, mixer and processing units (indices: $u, u1, u2$).

U^P : Set of processing units. Subset of U .

U^S : Set of storage units. Subset of U .

U^M : Set of mixer units. Subset of U .

U^F : Set of feed units (index: u).

U^D : Set of delivery units (index: u).

C_u : Set of campaigns per unit u (indices: $c, c1, c2$).

$S_{u,c}$: Set of output streams per unit u and campaign c (indices: $s, s1$).

S^I : Set of refinery input streams (index: s).

S^P : Set of finished product streams (index: s).

$F_{u1,c1,s1,u,c}$: Set of streams $s1$ that flow from unit $u1$ and campaign $c1$ to unit u and campaign c .

L : Set of utilities (index: l).

P : Set of properties (index: p).

T : Set of periods (index: t).

SC : Set of scenarios (index: sc).

Parameters

$Y_{u,s,t}$: Cost for selecting the refinery input s in the feed unit u and period t .

$R_{u,s,t}$: Cost for consuming the refinery input s in the feed unit u and period t .

$R_{u,s,t}^{MIN}$: Minimum consumption quantity of the refinery input s in the feed unit u and period t .

$R_{u,s,t}^{MAX}$: Maximum consumption quantity of the refinery input s in the feed unit u and period t .

$K_{u,c,t}$: Cost for storing in the unit u , campaign c and period t .

$K_{u,c}^{INI}$: Initial stock quantity of the unit u in the campaign c .

$K_{u,c,t}^{MIN}$: Minimum stock quantity of the unit u in the campaign c and period t .

$K_{u,c,t}^{MAX}$: Maximum stock quantity of the unit u in the campaign c and period t .

$\kappa_{u,c,s1}$: Fraction of the stock in the unit u and campaign c that corresponds to the input stream $s1$.

$U_{u,t}^{MIN}$: Minimum processing capacity of the unit u in the period t .

$U_{u,t}^{MAX}$: Maximum processing capacity of the unit u in the period t .

$C_{u,c,t}^{MIN}$: Minimum processing capacity of the campaign c in the processing unit u and period t .

$C_{u,c,t}^{MAX}$: Maximum processing capacity of the campaign c in the processing unit u and period t .

$\gamma_{u,c,s}$: Yield of the output stream s in the processing unit u and campaign c .

$D_{u,s,t}^{sc}$: Demand of the finished product s in the delivery unit u , period t and scenario sc . It corresponds to a realization of the random vector $\mathbf{D}_{u,s,t}(\omega)$, whose probability distribution P is assumed to be known.

$H_{u,s,t}^{COST}$: Cost for the product s shortage, in the delivery unit u and period t .

$L_{u,c,l,t}^{COST}$: Cost for using the utility l in the unit u , campaign c and period t .

$P_{u,s,t}^{PR}$: Price of the product s in the delivery unit u and period t .

$P_{u,c,p}^{INI}$: Value of the property p of the initial stock $IN_{u,c}$ in the unit u , campaign c and period t .

$P_{u,c,s1,p,t}^{I}$: Value of the property p of the input stream $s1$ in the unit u , campaign c and period t .

$P_{u,c,p,t}^{MIN}$: Minimum specification value of the property p in the unit u , campaign c and period t .

$P_{u,c,p,t}^{MAX}$: Maximum specification value of the property p in the unit u , campaign c and period t .

$PROB^{sc}$: Probability of the scenario sc .

Main Variables

$y_{u,s,t}$: Decision of selecting or not, in the first stage, the refinery input s type in the feed unit u to be consumed in the period t .

$r_{u,s,t}$: Decision of selecting, in the first stage, the refinery input s quantity in the feed unit u to be consumed in the period t .

$f_{u1,c1,s1,u,c,t}^{sc}$: Quantity of the stream s that flows from the unit $u1$ and campaign $c1$ to the unit u and campaign c , in the period t and scenario sc .

$k_{u,c,t}^{sc}$: Stock quantity of the unit u in the campaign c , period t and scenario sc .

$h_{u,s,t}^{sc}$: Shortage quantity of the product s in the delivery unit u , period t and scenario sc .

4.2. Objective Function

$$\max \left\{ -rmc + \sum_{sc \in SC} [PROB^{sc} \times (rev^{sc} - utc^{sc} - stc^{sc} - shc^{sc})] \right\} \quad (3a)$$

$$rmc = \sum_{u \in U^F} \sum_{s \in S^I} \sum_{t \in T} (YCOST_{u,s,t} \times y_{u,s,t} + RCOST_{u,s,t} \times r_{u,s,t}) \quad (3b)$$

$$rev^{sc} = \sum_{u1 \in U} \sum_{c1 \in C_{u1}} \sum_{s \in S^P} \sum_{u \in U^D} \sum_{c \in C_u} \sum_{t \in T} PR_{u,s,t} \times f_{u1,c1,s,u,c,t}^{sc} \quad (3c)$$

$$utc^{sc} = \sum_{u1 \in U} \sum_{c1 \in C_{u1}} \sum_{s \in S_{u1,c1}} \sum_{u \in U^P} \sum_{c \in C_u} \sum_{t \in T} \sum_{l \in L} LCOST_{u,c,l,t} \times f_{u1,c1,s,u,c,t}^{sc} \quad (3d)$$

$$stc^{sc} = \sum_{u \in U^S} \sum_{c \in C_u} \sum_{t \in T} KCOST_{u,c,t} \times k_{u,c,t}^{sc} \quad (3e)$$

$$shc^{sc} = \sum_{u \in U^D} \sum_{s \in S^P} \sum_{t \in T} HCOST_{u,s,t} \times h_{u,s,t}^{sc} \quad (3f)$$

The objective function 3a maximizes the expected value of the operating profit over all the scenarios $sc \in SC$. For a given scenario rev^{sc} , the profit is given by the difference between the total revenue rev^{sc} and the sum of the operating costs: refinery input total cost rmc , utilities total cost utc^{sc} , inventory total cost stc and shortage total cost shc^{sc} .

4.3. Constraints

$$RMIN_{u,s,t} \times y_{u,s,t} \leq r_{u,s,t} \leq RMAX_{u,s,t} \times y_{u,s,t}, \quad (4)$$

for $u \in U^F, s \in S^I, t \in T$

$$r_{u,s,t} - \sum_{u2 \in U} \sum_{c2 \in C_{u2}} f_{u,c,s,u2,c2,t}^{sc} = 0, \quad (5)$$

for $u \in U^F, s \in S^I, t \in T, sc \in SC$

$$KMIN_{u,c,t} \leq k_{u,c,t}^{sc} \leq KMAX_{u,c,t}, \quad (6)$$

for $u \in U^S, c \in C_u, t \in T, sc \in SC$

$$f_{u1,c1,s1,u,c,t}^{sc} - \kappa_{u,c,s1} \times \sum_{u1 \in (U \cup U^F)} \sum_{c1 \in C_{u1}} \sum_{s1 \in S_{u1,c1}} f_{u1,c1,s1,u,c,t}^{sc} = 0, \quad (7)$$

for $u \in U^S, c \in C_u, t \in T, sc \in SC$

$$k_{u,c,t-1}^{sc} + \sum_{u1 \in (U \cup U^F)} \sum_{c1 \in C_{u1}} \sum_{s1 \in S_{u1,c1}} f_{u1,c1,s1,u,c,t}^{sc} - k_{u,c,t}^{sc} - \dots \\ \dots - \sum_{u2 \in (U \cup U^D)} \sum_{c2 \in C_{u2}} \sum_{s \in S_{u,c}} f_{u,c,s,u2,c2,t}^{sc} = 0, \quad (8)$$

for $u \in U^S, c \in C_u, t \in T, sc \in SC$

$$UMIN_{u,t} \leq \sum_{u1 \in U} \sum_{c1 \in C_{u1}} \sum_{s1 \in S_{u1,c1}} \sum_{c \in C_u} f_{u1,c1,s1,u,c,t}^{sc} \leq UMAX_{u,t}, \quad (9)$$

for $u \in U^P, t \in T, sc \in SC$

$$CMIN_{u,c,t} \leq \sum_{u1 \in U} \sum_{c1 \in C_{u1}} \sum_{s1 \in S_{u1,c1}} f_{u1,c1,s1,u,c,t}^{sc} \leq CMAX_{u,c,t}, \quad (10)$$

$$\text{for } u \in U^P, t \in T, sc \in SC \quad (11)$$

$$\sum_{u1 \in U} \sum_{c1 \in C_{u1}} \sum_{s1 \in S_{u1,c1}} \gamma_{u,c,s} \times f_{u1,c1,s1,u,c,t}^{sc} - \sum_{u2 \in U} \sum_{c2 \in C_{u2}} f_{u,c,s,u2,c2,t}^{sc} = 0, \quad (12)$$

for $u \in U^P, c \in C_u, s \in S_{u,c}, t \in T, sc \in SC$

$$\sum_{u1 \in U} \sum_{c1 \in C_{u1}} \sum_{s1 \in S_{u1,c1}} f_{u1,c1,s1,u,c,t}^{sc} - \sum_{u2 \in U} \sum_{c2 \in C_{u2}} \sum_{s \in S_{u,c}} f_{u,c,s,u2,c2,t}^{sc} = 0, \quad (13)$$

for $u \in U^M, c \in C_u, t \in T, sc \in SC$

$$PMIN_{u,c,p,t} \leq \frac{PINI_{u,c,p} \times k_{u,c,t-1}^{sc} + A_{u,c,p,t}^{sc}}{k_{u,c,t}^{sc} + B_{u,c,t}^{sc}} \leq PMAX_{u,c,p,t}, \quad (14)$$

$$A_{u,c,p,t}^{sc} = \sum_{u1 \in (U \cup U^F)} \sum_{c1 \in C_{u1}} \sum_{s1 \in S_{u1,c1}} PI_{u,c,s1,p,t} \times f_{u1,c1,s1,u,c,t}^{sc}$$

$$B_{u,c,t}^{sc} = \sum_{u2 \in (U \cup U^D)} \sum_{c2 \in C_{u2}} \sum_{s \in S_{u,c}} f_{u,c,s,u2,c2,t}^{sc}$$

for $u \in U^S, c \in C_u, p \in P, t \in T, sc \in SC$

$$\sum_{u1 \in U} \sum_{c1 \in C_{u1}} f_{u1,c1,s,u,c,t}^{sc} + h_{u,s,t}^{sc} - D_{u,s,t}^{sc} \geq 0, \quad (15)$$

for $u \in U^D, c \in C_u, s \in S^P, t \in T, sc \in SC$

$$y_{u,s,t} \in \{y_{u,s,t} \in \{0,1\} \mid u \in U^F, s \in S^I, t \in T\} \quad (16)$$

$$r_{u,s,t} \in \{i_{u,s,t} \geq 0 \mid u \in U^F, s \in S^I, t \in T\} \quad (17)$$

$$f_{u1,c1,s1,u,c,t}^{sc} \in \{f_{u1,c1,s1,u,c,t}^{sc} \geq 0 \mid (u1, c1, s1, u, c) \in F_{u1,c1,s1,u,c}, \quad (18)$$

$$t \in T, sc \in SC\} \quad (19)$$

$$k_{u,c,t}^{sc} \in \{k_{u,c,t}^{sc} \geq 0 \mid u \in U^S, c \in C_u, t \in T, sc \in SC\} \quad (20)$$

$$h_{u,s,t}^{sc} \in \{h_{u,s,t}^{sc} \geq 0 \mid u \in U^D, s \in S^P, t \in T, sc \in SC\} \quad (21)$$

In (4), the quantity of the refinery input s selected in the feed unit u and period t must meet the minimum and maximum consumption levels.

Constraint (5) considers that the refinery input s in the feed unit u , period t and scenario sc can be supplied to different units $u2$ and campaigns $c2$.

Constraint (6) defines that the stock of the storage unit u in the campaign c and period t must meet the minimum and maximum stock levels.

In (7), the stock of the storage unit u in the campaign c , period t and scenario sc is formed according to a specified recipe.

Constraint (8) considers that the stock of the storage unit u in the campaign c , period t and scenario sc is equal to the sum of the input streams with the stock from the previous period minus the sum of the output streams. In case it is the first period ($t = 1$), $k_{u,c,t-1}^{sc}$ is equal to the initial stock parameter $KINI_{u,c}$.

In (9), the total load of the processing unit u in the period t and scenario sc must meet the minimum and maximum processing levels.

Constraint (10) defines that the load of the campaign c in the processing unit u , period t and scenario sc must meet the minimum and maximum campaign levels.

In (12), the output stream s in the processing unit u , campaign c , period t and scenario sc is produced according to its yield regarding the input stream $s1$.

Constraint (13) considers that the sum of the input streams $s1$ of the mixer unit u in the campaign c , period t and scenario sc must be equal to the sum of all output streams generated.

In (14), the property p of the product in the storage unit u , campaign c , period t and scenario sc must meet the minimum and maximum specification levels.

Constraint (15) defines that the demand of the product s must be equal to the sum of the quantity in the delivery unit u , period t and scenario sc , with the shortage variable $h_{u,s,t}^{sc}$.

(16) to (21) are the domain of the main variables.

5. The Value of the Stochastic Program

Due to the computational weight that stochastic programming models naturally add, it is common to quantify the benefits of using them instead of deterministic ones. The two most used values in the literature are the Expected Value of Perfect Information (EVPI) and the Value of the Stochastic Solution (VSS) [Birge and Louveaux, 2011]. The former measures the maximum amount the decision maker is willing to pay in order to get complete and accurate information on the future. The latter quantifies the cost of using a deterministic solution instead of a stochastic one.

The EVPI is calculated by taking the difference between the recourse problem solution (RP) and the so-called wait-and-see solution (WS). The RP is obtained by solving the deterministic equivalent problem (as described in 4), while the WS is calculated by taking the expected value of the optimum solutions obtained from solving each scenario as an independent deterministic problem.

In order to calculate the VSS, it is firstly necessary to compute the expected value problem solution (EV), which is a program built upon the recourse problem, changing only the random variables by their respective mean values. Then, the optimal first stage decisions obtained in the expected value problem are then used to quantify the so-called expectation of the expected value problem (EEV). This is performed by fixing the first stage variables with the optimal values obtained in the EV problem and allowing the optimization problem to choose the optimal second stage variables. With the EEV in hands, the VSS is calculate by taking the difference between the EEV and the RP.

In summary, for maximization problems, $EVPI = WS - RP$ and $VSS = RP - EEV$. For stochastic programs with only uncertain right hand , [Madansky, 1960] established the following inequalities $EV \geq WS \geq RP \geq EEV^3$.

6. Case Study

The stochastic model proposed in this work allows decisions makers to not only plan but also perform analysis more robust when compared to other methods (e.g. scenario analysis). As an example, a brief analysis of results based on a real Brazilian refinery is presented in this section.

6.1. Refinery Configuration

Figure 2 shows the refinery production area layout considered in the experiments. The refinery scale and capacity mimics a real Brazilian refinery.

In this configuration, the refinery is fed with crude oil types FAZA and FAZB and liquefied natural gas (LNG). Also, products like aviation kerosene (QAV), diesel (S500 and S2000), naphthenic oils (NH-10, NH-20, NH-140) and asphalt (CAP) can be purchased in the market.

The refinery production is focused on attending the local demand of asphalts (CAP, CM-30 and CR-250), lubricants oils (NH10, NH20 and NH140), thermal insulation oil for power transformers (Isovolt-HP), fabric softener oil (OAF), marine fuel oils (MF-180 and MF-380), industry fuel oils (OCA1 and OCB1), kerosene-type jet fuel (QAV), diesel (Diesel S500 and Diesel S2000), gasoline and liquefied petroleum gas (LPG). For the experiments, it is assumed that the demand is normally distributed with mean $\mu = 0,0$ and standard deviation $\sigma = 0.25$. A realization in a scenario $sc \in SC$ is given as follows:

$$D_{u,s,t}^{sc} = DEM_s \times (1 + \mathcal{N}(\mu, \sigma^2)) , \quad (22)$$

for $u \in U^D$, $c \in C_u$, $s \in S^P$, $t \in T$, $sc \in SC$

The parameter DEM_s corresponds to a base demand.

³ $EV \geq WS$ is valid only in cases the technology and recourse matrices and the objective function coefficients are fixed.

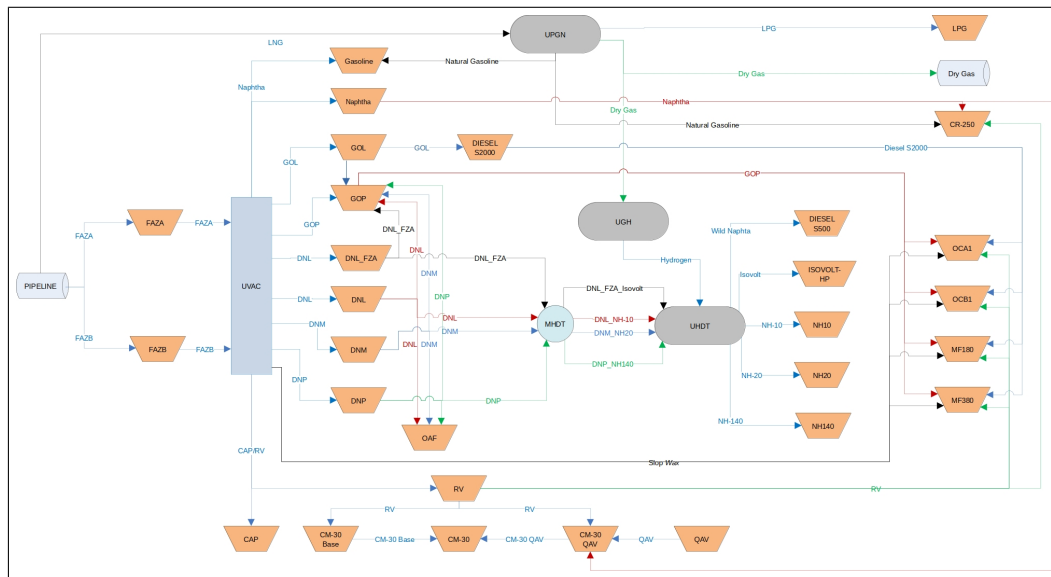


Figure 2: Refinery Configuration

6.2. Hardware and Software Information

The experiments were run on a computer with the following specification:

- Operating system: Linux Mint 21.3 x86_64.
- Kernel: 5.15.0-131-generic.
- CPU: Intel i5-8300H (8) @ 2.300GHz (set up to run at 3.000GHz).
- Memory: 15.830MiB.

The optimization model and the methods presented in this paper were implemented in Python 3.10.12, using the libraries PuLP 2.9.0 and pytups 0.88.1, and the solver CBC 2.10.7.

6.3. Computational Results

The EVPI displayed in table 1 indicates that, if it were possible to have complete and accurate knowledge about the future, the profit could have been on average 3.868,58 *m.u.* higher than the one found by the two-stage stochastic recourse program. But, since acquiring perfect information is not possible, the value of the stochastic solution can be considered the best result. It can be also concluded from the table 1 that, the gain of considering uncertain demand is on average 4.647.719,61 *m.u.* for two production periods.

Table 2 shows the economic results for two production periods. Most of the cost was related to refinery material inputs, followed by the expected shortage cost for approximately nine times less. This confirms that effective first stage decisions are a key element to reduce the losses that could occur from not meeting the demand. Another point of interest is how low the expected storage cost was when compared to the other two costs. This suggests that the “low-demand” scenarios had a much less impact to the refinery operations in this experiment.

In table 3, it is presented the optimal first stage decisions. In order to maximize the expected value of the profit, it was necessary to not only consume crude oil, but also order in the market amounts for diesel, naphthenic oils, asphalt and aviation kerosene. While the consume of the crude oil types permits to attend the demand for almost every scenario, the anticipated purchase

Table 1: EVPI and VSS

Item	Value (<i>m.u.</i>)
EV	99.665.104,88
WS	96.124.710,38
RP (Expected Profit)	96.120.841,80
EEV	91.473.122,19
EVPI = WS - RP	3.868,58
VSS = RP - EEV	4.647.719,61

Table 2: Operations economic results

Item	Value (<i>m.u.</i>)
Expected Profit	96.120.841,80
Expected Income	135.495.051,03
Expected Shortage Cost	3.845.839,13
Expected Storage Cost	418.417,24
Refinery Input Cost	35.109.952,87

Table 3: Optimal first stage decisions

Production Period	Refinery Input	Quantity (m^3)
1	FAZA	15.000,00
	FAZB	5.000,00
	LGN	4.500.000,00
	CAP	2.000,00
	DIESEL S2000	2.000,00
	DIESEL S500	1.000,00
	NH10	763,15
	NH140	1.878,40
	NH20	1.203,56
	QAV	211,26
2	FAZA	15.000,00
	FAZB	5.000,00
	LGN	4.500.000,00
	CAP	2.000,00
	DIESEL S2000	2.000,00
	DIESEL S500	1000,00
	NH10	795,64
	NH140	2000,00
	NH20	1.371,49
	QAV	214,84

of finished products minimizes the shortage cost in scenarios of high demand. In fact, along with inventory, it acts as a smoothing tool for demand deviations. Figures 3a and 3b illustrate respectively the revenue and total cost for each scenario derived from the optimal first stage decisions.

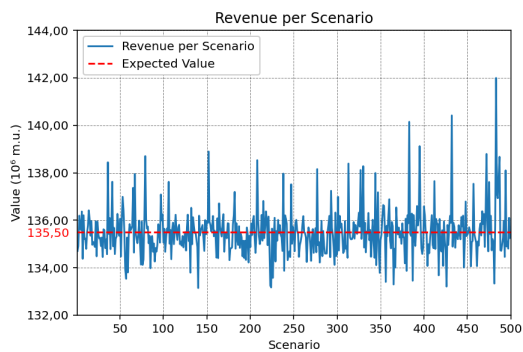
7. Conclusions

In this work, it was proposed a two-stage stochastic mixed integer programming with recourse model to tackle the tactical refinery planning problem. The goal was to define the amount of each refinery input to be consumed in two subsequent production periods that maximizes the expected value of the operating profit.

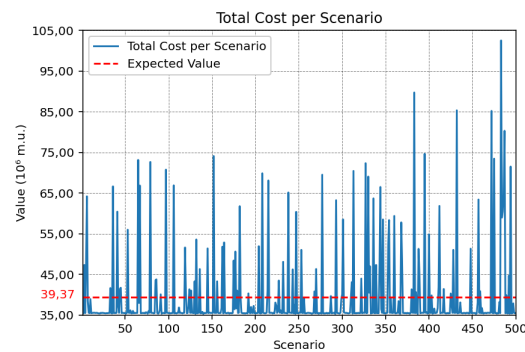
In order to evaluate the use of the stochastic model instead of a deterministic one, which is less costly in terms of computational demand, the EVPI and the VSS were computed. The analysis of the EVPI indicated that the stochastic model provided objective function solutions very close to the so-called wait-and-see solutions, making the former a good planning tool. The VSS experiment showed that the impact of not considering the random nature of the demand was relevant and could lead to higher costs, specially when considering the planning of multiple scenarios.

The experiments also showed that defining optimal refinery inputs is essential to minimize the effects of the market demand randomness, thus leading to better economic performance. This makes the stochastic model a valuable planning tool to support decision makers in the mid-term refinery production planning problem.

As suggestions for future work, the model could be adapted to a multi-stage one and other sources could be explored. Also, methods like the Sample Average Approximation [Shapiro and



(a) Revenue per scenario.



(b) Total cost per scenario.

Homem-de Mello, 1996] and the L-shaped algorithm [Van Slyke and Wets, 1969] could be used to increase solution quality and reduce computational effort.

References

- Al-Qahtan, K. Y. and Elkamel, A. (2010). *Planning and Integration of Refinery and Petrochemical Operations*. Wiley-VCH, Boschstr. 12, 69469 Weinheim, Germany.
- Beale, E. M. M. (1955). On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society, Series B*, 17(2):173–184.
- Birge, J. R. and Louveaux, F. (2011). *Introduction to Stochastic Programming*. Springer, New York, 2 edition.
- Boucheikhchoukh, A., Berger, V., Swartz, C. L., Deza, A., Nguyen, A., and Jaffer, S. (2022). Multiperiod refinery optimization for mitigating the impact of process unit shutdowns. *Computers & Chemical Engineering*, 164:107873.
- Coelho, O. d. S. J. (2013). *Programação Não-Linear Inteira Mista Aplicada ao Planejamento de Produção em Refinarias de Petróleo*. Tese de D.Sc., Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil.
- Dantzig, G. B. (1955). On minimizing a convex function subject to linear inequalities. *Management Science*, 1(3-4):197–206.
- Hsieh, S. and Chiang, C. C. (2001). Manufacturing-to-sale planning for fuel oil production. *The International Journal of Advanced Manufacturing Technology*, 19(4):303–311.
- Liu, M. L. and Sahinids, N. V. (1996). Optimization in processing planning under uncertainty. *Industrial and Engineering Chemistry Research*, 35(11):4154–4165.
- Macchiarulo, L. (2011). *Planejamento da produção de uma refinaria sob condições de incerteza*. Tese de M.Sc., Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil.
- Madansky, A. (1960). Inequalities for stochastic linear programming problems. *Management Science*, 6(2):197–204.
- Moro, L. F., Zanin, A. C., and Pinto, J. M. (1998). A planning model for refinery diesel production. *Computers and Chemical Engineering*, 22(1):1039–1042.

- Neiro, S. M. S. and Pinto, J. M. (2005). Multiperiod optimization for production planning of petroleum refineries. *Chemical Engineering Communications*, 192(1):62–88.
- Noh, N. M., Bahar, A., and Zainuddin, Z. M. (2018). Scenario based two-stage stochastic programming approach for the midterm production planning of oil refinery. *MATEMATIKA*, 34(3): 45–55.
- Pinto, J. M., Joly, M., and Moro, L. F. L. (2000). Planning and scheduling models for refinery operation. *Computers and Chemical Engineering*, 24(9-10):2259–2276.
- Shapiro, A. and Homem-de Mello, T. (1996). A simulation-based approach to two-stage stochastic programming with recourse. *Mathematical Programming*, 81(3):301–325.
- Van Slyke, R. M. and Wets, R. (1969). L-shaped linear programs with applications to optimal control and stochastic programming. *Siam Journal on Applied Mathematics*, 17:638–663. URL <https://api.semanticscholar.org/CorpusID:26352538>.
- Yang, Y. and Barton, P. I. (2015). Integrated crude selection and refinery optimization under uncertainty. *AIChE*, 62(4):1038–1053.
- Zhang, L., Torres, A. I., Chen, B., Yuan, Z., and Grossmann, I. E. (2024). Optimal retrofitting of conventional oil refinery into sustainable bio-refinery under uncertainty. *AIChE*, 70(4).