A Hybrid of Stochastic Programming Approaches with Economic and Operational Risk Management for Petroleum Refinery Planning under Uncertainty

by

Cheng Seong Khor

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Applied Science
in
Chemical Engineering

Waterloo, Ontario, Canada, 2006

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

A Hybrid of Stochastic Programming Approaches with Economic and Operational Risk Management for Petroleum Refinery Planning under Uncertainty

The current situation of fluctuating high petroleum crude oil prices is a manifestation that markets and industries everywhere are impacted by the uncertainty and volatility of the petroleum industry. As the activity of petroleum refining is at the heart of the downstream sector of the petroleum industry, it is increasingly important for refineries to operate at an optimal level in the present turbulent, dynamic nature of the world economic environment. Refineries must assess the potential impact of significant primary changes that are posed by market demands for final products and their associated specifications; costs of purchasing the raw material crude oils and prices of the commercially saleable intermediates and products; and crude oil compositions and their relations to product yields; in addition to even be capable of exploring and tapping immediate market opportunities. Hence, this calls for a greater need in the strategic planning, tactical planning, and operations control of refineries in order to execute operating decisions that satisfy conflicting multiobjective goals of maximizing expected profit while simultaneously minimizing risk, on top of sustaining long-term viability and competitiveness. These decisions have to take into account uncertainties and constraints in factors such as the source and availability of crude oils as the raw material; the processing and blending options of the desired refined products that in turn depend on the uncertainties of the components' properties; and economic data such as prices of feedstock, chemicals, and commodities; production costs; distribution costs; and future market demand for finished products. Thus, acknowledging the shortcomings of deterministic models, this work proposed a hybrid of stochastic programming formulations for the optimal midterm production planning of a refinery that addresses three major sources of uncertainties, namely prices of crude oil and saleable products, product demands, and product yields. A systematic and explicit stochastic optimization technique was employed by utilizing slack variables to account for violations of constraints in order to increase model tractability. Four different approaches were

considered to ensure model and solution robustness: (1) the Markowitz's mean–variance (MV) model to handle randomness in the objective function coefficients of prices by minimizing the variance of the expected value or mean of the random coefficients, subject to a target profit constraint; (2): the two-stage stochastic programming with fixed recourse via scenario analysis approach to model randomness in the right-hand side and the left-hand side or technological coefficients by minimizing the expected recourse penalty costs due to constraints' violations; (3) incorporation of the Markowitz's mean–variance approach within the two-stage stochastic programming framework developed in (2) to minimize both the expectation and the variance of the recourse penalty costs; and (4) reformulation of the model developed in the third approach by utilizing the Mean–Absolute Deviation (MAD) as the measure of risk imposed by the recourse penalty costs. In the two-stage modelling approach that provided the framework for the proposed stochastic models, the deterministic first-stage planning variable(s) determined the amount of resources for the refinery production operations, that is, the crude oil supply. Subsequently, once the value of the planning variable had been decided and the random events had been realized, the corrective action or the recourse were implemented by selecting the random second-stage variables associated with operating decisions for improvements. Therefore, the overall objective in the bilevel approach to decision-making under uncertainty was to choose the planning variable of crude oil supply in such a way that the first-stage planning costs and the expected value of the random second-stage recourse costs were minimized. A representative numerical study was then illustrated, with the solutions compared and contrasted by several metrics derived from established relevant concepts, as follows. We found that the resulting outcome of the stochastic models' solutions consistently proposed higher expected profits than the deterministic model and the fuzzy linear fractional programming approach of Ravi and Reddy (1998) who worked on the same problem. Additionally, the stochastic models demonstrated increased robustness and reliability (or certainty) as measured by the coefficients of variation in comparison with the deterministic model.

Acknowledgements

First and foremost, I would like to thank my advisors/supervisors, Professor Dr. Peter L. Douglas and in particular, Professor Dr. Ali Elkamel, and also Professor Kumaraswamy "Ponnu" Ponnambalam (instructor for the course SYDE 760: Probabilistic Design, from which the original idea of this line of work was conceived), for earnestly providing mentorship to the best of their capability, even beyond the call of duty. They have been exceptionally patient and understanding; at times, simply being friends by lending their much-appreciated ears, in the course of my working on this thesis research. Working with them has been a truly enriching experience, not solely in the academic sense but also in contributing towards my personal development, possibly in more ways that they probably realized.

I am also grateful, beyond mere formal working associations, to fellow researchers and potentially life-long friends and collaborators of the petroleum refinery research group; primarily Khalid Al-Qahtani, Ibrahim Alhajri, and Mohammad Ba-Shammakh, in addition to Yusuf Ali Yusuf, Yousef Saif, and Aftab Khurram. I am also indebted to Masoud Mahootchi in appreciation of the many fruitful and valuable discussions especially at the initial stages of this work. All in all, thank you so very much for educating me not just about research, mathematical programming, and the myriad of refinery processes, but most importantly, about the meaning of a life well lived.

The encouragement and comradeship provided by the larger CO_2 (carbon dioxide) capture research group is equally instrumental, with due gratitude to Associate Professor Dr. Eric Croiset; Flora Wai Yin Chui, without whom I would not have survived my first term; officemates Mohammad Murad Chowdhuri and Guillermo "Memo" Ordorica; and Colin Alie for providing the laughter and the intellectually-simulating conversations, nonetheless.

I gratefully acknowledge financial support from Universiti Teknologi PETRONAS, my generous employer, for sponsoring my graduate studies at the University of Waterloo, Canada; more so, for granting me this opportunity and exposure so that I shall return with the knowledge and experience of educating and preparing successful future generations of Malaysians.

Last but definitely not the least, I extend my most heartfelt gratitude to my parents (beloved mother, Lee Swi Tin and father, Khor Geok Chuan) for sacrificing a great part of their life through unconditional love for this child to ensure that he grows up to become the person that he is today. I dearly hope that I am indeed that person who has made all those sacrifices and love worthwhile.

I am sorry that I am not able to name each and everyone who have touched my life in so many meaningful ways on these limited pages but I shall not forget your kind deeds, even your mere presence, in my mind and my heart. Indeed, I salute you and thank you all with the utmost sincerity and appreciation.

To my beloved mother and father and brother;
and all my respected teachers, past and present

Table of Contents

Appendices **249**

List of Tables

List of Illustrations

Nomenclature and Notations

Indices

i	for the set of materials or products
j	for the set of processes
t	for the set of time periods

Sets

I	set of materials or products
J	set of processes
T	set of time periods

Parameters

$d_{i,t}$	demand for product i in time period t
$d_{i,t}^{\mathrm{L}}, d_{i,t}^{\mathrm{U}}$	lower and upper bounds on the demand of product i during period t, respectively
$p_t^{\mathrm{L}}, p_t^{\mathrm{U}}$	lower and upper bounds on the availability of crude oil during period t, respectively
$I_{i,t}^{\mathrm{fmin}}, I_{i,t}^{\mathrm{fmax}}$	minimum and maximum required amount of inventory for material i at the end of each time period
$b_{i,j}$	stoichiometric coefficient for material i in process j
$\gamma_{i,t}$	unit sales price of product type i in time period t
λ_t	unit purchase price of crude oil in time period t
$\tilde{\gamma}_{i,t}$	value of the final inventory of material i in time period t

$\tilde{\lambda}_{i,t}$	value of the starting inventory of material i in time period t (may be taken as the material purchase price for a two-period model)
$\alpha_{j,t}$	variable-size cost coefficient for the investment cost of capacity expansion of process j in time period t
$\beta_{j,t}$	fixed-cost charge for the investment cost of capacity expansion of process j in time period t
r_t, o_t	cost per man-hour of regular and overtime labour in time period t

Variables

$x_{j,t}$	production capacity of process j ($j = 1, 2, \ldots, M$) during time period t
$x_{j,t-1}$	production capacity of process j ($j = 1, 2, \ldots, M$) during time period t–1
$y_{j,t}$	vector of binary variables denoting capacity expansion alternatives of process j in period t (1 if there is an expansion, 0 if otherwise)
$CE_{j,t}$	vector of capacity expansion of process j in time period t
$S_{i,t}$	amount of (commercial) product i ($i = 1, 2, \ldots, N$) sold in time period t
$L_{i,t}$	amount of lost demand for product i in time period t
P_t	amount of crude oil purchased in time period t
$I^{\mathrm{s}}_{i,t}, I^{\mathrm{f}}_{i,t}$	initial and final amount of inventory of material i in time period t
$H_{i,t}$	amount of product type i to be subcontracted or outsourced in time period t
R_t, O_t	regular and overtime working or production hours in time period t

Superscripts

$(\ )^{\mathrm{L}}$	lower bound
$(\ )^{\mathrm{U}}$	upper bound

Nomenclature and Notations for the Numerical Example (as depicted in Figure 8.1)

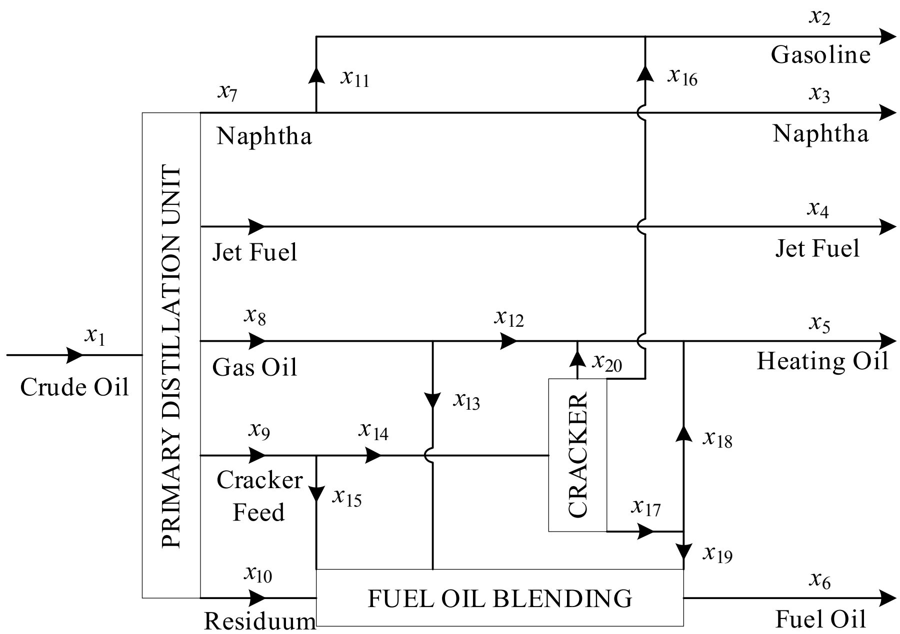


Figure 8.1. Simplified representation of a petroleum refinery for formulation of the deterministic linear program for midterm production planning

x_1	mass flow rate (in ton/day) of crude oil stream
x_2	mass flow rate (in ton/day) of gasoline in combined streams of x_{11} and x_{16}
x_3	mass flow rate (in ton/day) of naphtha stream after a splitter
x_4	mass flow rate (in ton/day) of jet fuel stream
x_5	mass flow rate (in ton/day) of heating oil stream
x_6	mass flow rate (in ton/day) of fuel oil stream
x_7	mass flow rate (in ton/day) of naphtha stream exiting the primary distillation unit (PDU)
x_8	mass flow rate (in ton/day) of gas oil stream
x_9	mass flow rate (in ton/day) of cracker feed stream
x_{10}	mass flow rate (in ton/day) of residuum stream
x_{11}	mass flow rate (in ton/day) of gasoline stream after splitting of naphtha stream exiting the PDU
x_{12}	mass flow rate (in ton/day) of gas oil stream after a splitter

x_{13} mass flow rate (in ton/day) of gas oil stream entering the fuel oil blending facility

x_{14} mass flow rate (in ton/day) of cracker feed stream after a splitter

x_{15} mass flow rate (in ton/day) of cracker feed stream entering the fuel oil blending facility

x_{16} mass flow rate (in ton/day) of gasoline stream exiting the cracker unit

x_{17} mass flow rate (in ton/day) of stream exiting the cracker unit into a splitter

x_{18} mass flow rate (in ton/day) of heating oil stream after splitting of cracker output

x_{19} mass flow rate (in ton/day) of cracker output stream

x_{19} mass flow rate (in ton/day) of heating oil stream exiting the cracker unit

CHAPTER 1

Introduction and Review of Current Modelling Practices and Related Literature

Chemical process design, planning, and operations problems are usually treated as deterministic problems with defined models and known constant parameters. In the real world, however, the chemical process industry is typically ridden with uncertainties in a multitude of factors spanning a wide range. These include market demands for products; prices of raw materials and saleable products; lead times and availabilities in the supply of raw materials as well as lead times or rates in the processing, production and distribution of final products; product yields; product qualities; capital, technology, competition, equipment, and facilities parameters such as reliability, availability, and failures (Subrahmanyam et al., 1994; Applequist et al., 2000; Jung et al., 2004; Sahinidis et al., 1989). Uncertainties might even arise in aspects as fundamental as thermodynamics, kinetics, and other modelling parameters, as noted by Ahmed (1998). These uncertainties could be present in the form of incomplete information, data variability, randomness, and others (Shapiro and Homem-de-Mello, 1998). Thus, uncertainties are inevitable and prevalent in mathematical models, parameters, and also in enforcing the planning model itself to specifications. Consequently, this renders models based on deterministic consideration to not always be optimal or even operable. In fact, Ben-Tal and Nemirovski (2000) stress that optimal solution of deterministic linear programming problems may become *severely* infeasible even if the nominal data is only slightly perturbed. This is supported by Sen and Higle (1999) who affirmed that under uncertainty, the deterministic formulation in which uncertain random variables are mathematically and statistically replaced by their expected values may not provide a solution that is feasible with respect to the random variables. Hence, the need to model uncertainty in process design, planning, scheduling, and operations activities has long been recognized as essential in the realm of chemical process systems engineering (PSE).

As a consequence of operating in such a rapidly changing dynamic and risky environment, in making planning decisions, a firm must not only be restricted to consideration of short-term economic criteria but ought to also identify and assess the impact of vital uncertainties aforementioned to its business in order to be able to develop coping strategies through implementation of contingency plans, to be effected as the uncertainties unfold. Since the selection of current decisions depends on decisions taken in previous time periods, it is essential to formulate planning decisions that not only maximize the expected profit, but also ensure future feasibility. This can be achieved by accounting for the minimization of economic risk involved in implementing a supposed optimal plan besides sustaining long-term viability and competitiveness (Cheng et al., 2003; Applequist, 2002; Applequist et al., 2000).

In fact, virtually all decision-making processes involve uncertain information, particularly when future events are considered. Apart from production planning and the related activity of process scheduling, other common engineering examples include applications in optimal control, real-time optimization, and capacity planning with the objective of expansion. Production planning applications are of particular interest due to their inherently uncertain nature, high economic incentives, and strategic importance. Furthermore, realistic production planning applications can be developed with well-established linear programming models, which can be extended to include uncertainties in parameters characterized by probability distribution functions, giving rise to the two-stage stochastic linear program, which forms the underpinning framework in the models proposed in this work.

In the chemical process systems engineering (PSE) literature, problems associated with the design, planning, and operations of process systems under uncertainty have been attracting considerable attention especially during the period of 1990s (Jung et al., 2004). Over time, from early works in the chemical engineering field addressing issues of uncertainties (for examples, see Grossmann and Sargent, 1978 and Malik and Hughes, 1979) to more recent works, numerous ideas have been proposed to formulate planning (and design) problems dealing with uncertain model parameters. In general, the solution approaches have proceeded along two main directions: (1) deterministic methods in which the emphasis is on ensuring the feasibility of the solutions over a given domain of the uncertain parameters, and (2) stochastic or probabilistic optimization techniques in which the objective is to optimize solutions that anticipate uncertainty of parameters that are described by probability distribution functions (Ierapetritou and Pistikopoulos, 1994b; Tarhan & Grossmann, 2005).

In the deterministic approach, the description of uncertainty is provided either by specific bounds on variables or by a finite number of fixed parameter values in terms of scenarios or time periods, transforming the process model to a deterministic approximation. These methods include:

(a) the "wait-and-see" approach, or sometimes referred to as scenario analysis or what-if analysis. It is characterized by discretization over the uncertain parameter space (for example, see Brauers and Weber, 1988);

(b) the use of multiperiod models, which is characterized by discretization over the time horizon (for examples, see Grossmann and Sargent, 1979; Grossmann et al., 1983; Sahinidis et al., 1989; Bok et al., 2000).

The model approximation can often be coupled with flexibility test or flexibility index problems as employed by Pistikopoulos and Grossmann (1988, 1989a, 1989b).

On the other hand, the more sophisticated stochastic optimization techniques take into account the detailed statistical properties of the parameter variations. These methods have evolved around two traditional forms of approaches, namely:

(a) the "here-and-now" approach of two-stage stochastic programming with recourse framework, originally proposed by Dantzig (1955) and Beale (1955) that is extendable to multiple stages. It is based on the postulation of general probability distribution functions describing process uncertainty with the objective of cost minimization or profit maximization due to violation of constraint(s) (examples of early work include Walkup and Wets, 1967; Wets, 1974; Grossmann and Sargent, 1978; Pai and Hughes, 1987, to mention only a few);

(b) the probabilistic modelling approach or also known as chance-constrained programming, originally introduced by Charnes and Cooper (1959), which includes in the constraints, the requirement that the probability of any constraint to be satisfied must be greater than the desired level (Gupta et al., 2000; Aseeri and Bagajewicz, 2004, again to mention only a few).

In addition, in a fairly more recent development, Ben-Tal and Nemirovski (2000) propose a robust optimization methodology for linear programming problems with uncertain data. In the realm of PSE, this approach has been adopted by Lin et al. (2004) to mixed-integer linear program (MILP) scheduling problems under bounded uncertainty in the coefficients of the objective function, the left-hand side parameters, and the right-hand side parameters of the inequalities considered via the introduction of a small number of auxiliary variables and constraints to determine the optimal schedule.

The two-stage stochastic programming approach has been proven to be most useful as a source of reliable design and planning information (Johns et al., 1978; Wellons and Reklaitis, 1989; Petkov and Maranas, 1998). As the name indicates, decisions are made in two stages in this modelling framework by loosely dividing time into "now" and "the future". The decision maker makes the first stage decision(s)\ prior to the realization of the uncertainty now and then makes the second stage recourse decision(s) contingent on the revealed information upon resolution of the uncertainty in the future. The first-stage decision variables are fixed while the second-stage operating variables are adjusted based on the realization of the uncertain parameters. Note that the stages do not necessarily correspond to periods in time. Each stage represents a decision epoch where decision makers have an opportunity to revise decisions based on the additional available information. For example, one can formulate a two-stage stochastic program for a multiperiod problem in which the second stage represents a group of periods in the remaining future (Cheng et al., 2005). Despite differences in individual details, most of the representative works in production planning of processes (see, for example, Ahmed & Sahinidis, 1998; Liu & Sahinidis, 1996; Petkov & Maranas, 1998; Ierapetritou & Pistikopoulos, 1994c, 1996c), including recent works in refinery planning (see, for example, Pongsakdi et al., in press; Neiro and Pinto, 2005; Aseeri and Bagajewicz, 2004), have followed the general structure of the two-stage stochastic programming framework, which provides an effective formulation for chemical process planning under uncertainty problems as will be demonstrated in this work.

It might be of interest to point out the differences between formulations of stochastic optimization problems that are derived from statistics and those that are motivated by decision-making under uncertainty. The analysis of "wait-and-see" solutions is mostly of interest in mathematical statistics in which information is collected and used *during* the decision process. Decision-making under uncertainty through stochastic programming is mostly concerned with problems that require a "here-and-now" decision, without making further observations of the quantities modelled as random variables. The solution must be found on the basis of the *a priori* information about these random quantities (Wets, 1989). Thus, the emphasis of stochastic programming lies in the methods of solution and the analytical solution properties whereas statistical decision theory stresses on procedures for constructing objectives and updating probabilities (Birge, 1997).

Additionally, two other notable approaches have also been proposed to deal with uncertainties in model parameters:

1. fuzzy programming as originally conceived in the seminal paper by Bellmann and Zadeh (1970) and popularized by Zimmermann (1991) with examples of application in PSE by Liu and Sahinidis (1997) and Ravi and Reddy (1998); and
2. the flexibility index analysis and optimization approach in design and operational planning problems. In the latter approach, flexibility is defined as the range of uncertain parameters that can be dealt with by a specific design or operational plan (Sahinidis, 2004). Flexibility thus refers to the ability of a system to readily adjust in order to meet the requirements of changing conditions. Some examples include the works of Pistikopoulos and Mazzuchi (1990); Straub and Grossmann (1993); and Ierapetritou and Pistikopoulos (1994a). This is a very much active major research area in PSE under the theme of integration of process design and control systems design and will not be addressed within the scope of this work.

1.1 APPROACHES TO MODELLING AND DECISION MAKING UNDER UNCERTAINTY IN OPERATIONS–PRODUCTION PLANNING AND SCHEDULING ACTIVITIES IN CHEMICAL PROCESS SYSTEMS ENGINEERING (PSE)

Operations and production planning activities in an industrial setting are crucial components of a supply chain. In fact, in his excellent review on single-site and multisite planning and scheduling, Shah (1998) considers medium-term or midterm planning as a special case of supply chain planning. In general, planning involves making optimal decisions about future events based on current information and available future projections. In the context of the chemical processing industry (CPI), typical decisions pertain to selection of new processes, expansion and/or shutdown policies of existing processes and facilities, and optimal operating patterns for production chains. These decisions have to be made in the face of the present inherently turbulent nature of business economic environments due to increasing competition, stringent production quality, fluctuating commodity prices and customer demands, and obsolescence in technology. In addition, companies ought to constantly recognize the potential benefits of new resources to be incorporated in conjunction with existing processes and facilities. The interaction of these situations provide incentives for companies in CPI to be concerned with the development of effective and efficient quantitative techniques and solutions for planning, as these are necessary tools in hedging against future contingencies for the eventual successful operation of even any modern-day enterprise, for that matter (Ierapetritou and Pistikopoulos, 1996; Sahinidis et al., 1989).

It is a well-recognized problem that production–manufacturing systems are subject to uncertainties presented by random events such as raw material variation, demand fluctuation, and equipment failures. The dynamic and random nature of product demands alone results in their forecasting being very difficult or sometimes even impossible. Despite the existence and availability of various planning models, managers often could not find one that is suitable for their needs. As a result, production is planned following an everyday practice without concern for achieving optimality. It is desirable to shift such experience-based decision making to an information-based data-driven decision-making model (Shapiro, 2004; 1999). This will require a systematic use of historical data and a theoretically sound mathematical model that is applicable to the real situation, with consideration for various possible operational and production uncertainties (Yin, K. K. et al., 2004). The present work is intended to contribute in these directions via the utilization of mathematical programming or optimization.

In the planning of chemical processes such as the vast array present in the operations of a petroleum refinery, we often have to deal with parameters that can vary during the operation and with parameters whose values are uncertain at the design stage. At this juncture, as stressed earlier, determining the right modeling tools is one of the most technologically challenging problems that operators and decision makers face today, as corroborated by Escudero et a. (1999). Probabilistic or stochastic methods and analyses have been demonstrated to be useful for screening the alternatives on the basis of the expected value of the economic criteria, typically the maximum expected profit or the minimum expected cost, and also the economic and financial risks involved. Several approaches have been reported in the literature addressing the problem of production planning under uncertainty. Extensive reviews addressing various issues in this area are available, for example, by Applequist et al. (1997) and by Cheng et al. (2005).

According to Gupta and Maranas (1999, 2003) and Vidal and Goetschalckx (1997), models of planning systems (with the term "planning" used here reflecting a general broad sense) can be broadly categorized into three distinct temporal classifications based on the addressed time frames or time horizons, namely strategic, tactical, and operational. A discussion of their features and characteristics from a practical perspective is provided by Shobrys and White (2000). The following aims to condense these views.

1. Long-range planning of capacity expansion and design models are termed as *strategic* or planning models (contrary to the aforementioned, the term "planning" is used here in a strict context to denote a long-term time horizon). They aim to identify the optimal timing, location, and extent of additional investments in processing networks over a relatively long time horizon ranging from the order of five to ten years. Thus, the decisions executed may affect access to raw materials, product slates, geographical markets, and obviously, production or distribution capacity. The strategic level requires approximate and aggregated data. For examples, see Sahinidis et al., (1989), Sahinidis and Grossmann (1991), and Norton and Grossmann (1994).

2. On the other extreme of the spectrum of planning models are short-term models classified as *scheduling* or *operational* planning models. These models are characterized by short time frames and therefore involve short-term decisions, typically less than one hour or one day, but could also stretch to a few days to one-to-two weeks to even two-to-three months. They address the exact full sequencing (timing) and volumes of the multifarious manufacturing tasks while accounting for the various resource and timing constraints, for instance, in the determination of the qualities of commodities to be produced by an oil refinery. Specifically, key decision variables involve the start time of an operation, and the duration and processing volume of the associated operating unit, under consideration for product demand, possible desire to keep major units operating continuously, and issues of containment. This operational level requires transactional data. For examples, see Shah et al. (1993), Xueya and Sargent (1996), and Karimi and McDonald (1997).

3. Medium-term or midterm or tactical planning models make up the third class of planning models. They are intermediate in nature and characteristically address planning horizons involving months, in a typical aggregation of two-to-six months, and up to one-to-two years. They execute the company-wide function of setting targets for operating performance, and coordinate activities across sales, materials management, manufacturing, and distribution. They consolidate features from both the strategic and operational models, including the amount and accuracy of data required. For instance, they account for the carryover of inventory over time and various key resource limitations, much like the short-term scheduling models; of which, an example within a petroleum refinery would be in deciding the type of crude oils to buy and the timing. On a contrasting note, similar to strategic planning models and unlike the operational models, they account for the presence of multiple production sites in the supply chain. In fact, refineries, with their typically large and complex manufacturing facilities, may also have a tactical planning process for each manufacturing site in order to coordinate activities across major units. The midterm planning models derive their value from this overlap and integration of modelling features. For examples, see McDonald and Karimi (1997) and Gupta and Maranas (2000).

A number of key decisions must be made during each of these time frames of days, months, and years in terms of the process operations. The crucial challenge is in providing the necessary theoretical, algorithmic, and computational support to aid optimal decision making accounting for future uncertainty primarily in product demands and other parameter variability.

As highlighted earlier, problems of design and planning of chemical processes and plants under uncertainty have been treated in the process systems engineering (PSE) literature using the well known decision problem model of two-stage stochastic programming with recourse. The two-stage programming strategy has been considered as an effective approach to the solution of process engineering problems such as production planning as it naturally differentiates between the following two sets (Acevedo and Pistikopoulos, 1998; Ruszczynski, 1997; Grossmann et al., 1983):

(i) the first-stage deterministic planning variables of resources representing the plan, that is, decisions that have to be made in advance and which remain fixed once selected, and

(ii) the second-stage stochastic operating or production variables, which are flexible and can be adjusted to represent operational decisions to achieve feasibility, depending on the observed event.

Under this framework, we pose the decision problem as one of maximizing (or minimizing, accordingly) an objective function consisting of two terms. The first corresponds to a contribution by the global or planning variables whose values are chosen independent of the uncertain parameters. The second term represents and quantifies the expected value of the contribution due to local or production variables, whose values will be adjusted in response to realization of specific values of the uncertain parameters. Generally, the objective function is a net present value of the associated investment, operating cost, and revenue streams. Thus, the objective in the two-stage modelling approach to decision under uncertainty, as reflected and defined in the objective function, is to choose the planning variables in such a way that the sum of the first-stage design costs and the expected value of the random second-stage recourse costs is minimized. Approaches differ primarily in how the expected value term is computed.

Moreover, the classification of the variables and constraints of a production planning problem (such as that addressed in this work) into two distinct categories, resulting in a two-stage hierarchical decision-making framework, can be effectively utilized for incorporating uncertainty in the dominant random parameter of product demands as dictated by market requirements, in addition to other parameters such as prices and yields, on a simultaneous basis, as will be demonstrated in this work. In this bilevel decision-making framework, the planning decisions are made “here-and-now” prior to the resolution of uncertainty, while the production decisions are postponed to a “wait-and-see” mode (Gupta and Maranas, 2000).

1.2 CLASSIFICATIONS OF UNCERTAINTY

According to Li (2004), uncertainty can be categorized based on different criteria. From the time horizon point-of-view, uncertainty can be present in short term, mid term, and long term. Short-term uncertainty typically involves day-to-day or week-to-week processing variations, for example in flow rates and temperatures; cancelled or rushed orders; and equipment failure; which requires the plant to respond within a short period of time (Subrahmanyam et al., 1994). Midterm uncertainty addresses time horizons spanning one to two years and incorporates features from both short-term and long-term uncertainties (Gupta and Maranas, 2003). Long-term uncertainty includes raw material or final product related issues of unit price fluctuations, seasonal demand variations, and production rate changes, occurring over longer time frames ranging from five to ten years (Sahinidis et al., 1989)

Li (2004) and Wendt et al. (2002) also classified uncertainties from the point-of-view of process operations, into two categories: external uncertainties and internal uncertainties. As indicated by its name, external uncertainties are exerted by outside factors but impacts on the process. Examples include feedstock condition such as feed composition and feed flowrate (for a petroleum refinery, this would be dictated by the type of crude oil intake for processing from the upstream exploration and production activities) and recycle flowrates as well as flows of utilities, the temperature and pressure of coupled operating units, and market conditions. Internal uncertainties arise from deficiency in the complete knowledge of the process. Some examples include yields of reactions, especially in processes with multiple reactions such as in a petroleum refinery; the kinetic parameters of reactions in units such as the fluidized-bed catalytic cracker (FCC); and the transfer rate of units such as the crude distillation unit (CDU). According to Goel and Grossmann (2004), Jonsbraten (1998) termed this class of uncertainty for planning problems as project exogenous uncertainty and project endogenous uncertainty to refer to external and internal uncertainties, respectively. As an aside, it is further noted that the scenario tree employed in modeling project exogenous uncertainty is independent of decisions made at preceding stages whereas the converse is true for its counterpart, that is, the scenario tree is dependent on prior decisions in modeling project endogenous uncertainty.

Uncertainties due to unknown input parameters are identified as (1) uncertain model parameters and (2) variable process parameters from the observability point-of-view (Rooney and Biegler, 2003). The exact values of uncertain model parameters are never known exactly for the design or planning problem although the expected values and confidence regions may be known. These include model parameters determined from (offline) experimental studies such as kinetic parameters of reactions as well as unmeasured and unobservable disturbances such as the influence of wind and sunshine. On the other hand, variable process parameters, although unknown at the design or planning stage, can be specified deterministically or measured accurately at later operating stages. Examples of these are (i) internal unmeasured disturbances such as feed flow rates, product demands, and process conditions and inputs (for example temperatures and pressures) and (2) external unmeasured uncertainty such as ambient conditions where an operation-of-interest takes place.

Table 1.1 summarizes the salient points on the three different categories to classify uncertainties and some associated examples.

1.3 MANAGEMENT OF PETROLEUM REFINERIES

1.3.1 Introduction to Petroleum Refinery and Refining Processes

Petroleum refining is a central key component and crucial link in the oil supply chain. It is where crude petroleum is transformed into products that can be used as transportation and industrial fuels, and for the manufacture of plastics, fibres, synthetic rubbers and many other useful commercial products. In general, a refinery is made up of several distinct parts as outlined in the following (Favennec and Pigeyre, 2001):

- the various processing units that separate crude oil into different fractions or cuts, upgrade and purify some of these cuts, and convert heavy fractions to light, more useful, fractions;
- utilities that refer to the systems and processes providing the refinery with fuel, flaring capability, electricity, steam, cooling water, effluent treatment, fire water, sweet water, compressed air, nitrogen, etc., all of which are necessary for the refinery's safe operation;

- the tankage area or tank farm where all crudes, finished products, and intermediates are stored prior to usage or disposal; and
- facilities for receipt of crude oil and for blending and despatch of finished products.

A simplified process flow diagram for a typical refinery is shown in Figure 1.1 while Table 1.2 provides a summary of the general processes that make up crude oil refining activities.

Table 1.1. Classification of Uncertainties (Li, 2004; Subrahmanyam et al., 1994; Rooney and Biegler, 2003; Dantus and High (1999))

Time-Horizon

Short Term

- Process variations, e.g., flow rates and temperatures
- Cancelled/Rushed orders
- Equipment failure

Mid Term (intermediate between short-term and long-term planning horizon)

LongTerm

- Unit price fluctuations
- Seasonal demand variations
- Production rate change
- Capital cost fluctuation

Process Operations

External (Exogenous)

- Sales uncertainty, e.g., unpredictable changes in prices and levels of demand of products
- Raw material purchase uncertainty, e.g., unpredictable changes in prices and levels of availability of raw materials (or feed stream), including raw material composition (i.e., feed composition)
- Economic factors, e.g., capital costs, manufacturing costs, direct costs, liability costs, and other less tangible costs
- Equipment purchase uncertainty, e.g., difficulties in predicting the cost and availability of equipment items
- Discrete uncertainty involving equipment reliability, e.g., uncertainty associated with the availability of an equipment item for normal operation, including other discrete random events
- Environmental impact, e.g., release factors and hazardous levels/values
- Regulatory uncertainty concerning laws, regulations, and standards, e.g., modification in emissions standards and new environmentally-motivated regulations
- Technology obsolescence
- Time uncertainty, e.g., delays in investment (perhaps due to projection that a project might hold promise of a better return/profit in the future in consideration of the current economic, political, and social situations)

Internal (Endogenous)

Manufacturing uncertainty, i.e., variations in processing parameters, e.g., yields and processing times

Observability (Rooney and Biegler, 2003)

Uncertain (process) model parameters

- From (offline) experimental studies, e.g., kinetic parameters (constants) of reactions, physical properties, and transfer coefficients
- Unmeasured and unobservable disturbances, e.g., influence of wind and sunshine

Variable process parameters

- Internal unmeasured disturbances, e.g., feed flow rates, stream quality, and process conditions and inputs (e.g., variations in temperatures and pressures)
- External unmeasured uncertainty, e.g., ambient conditions of operation

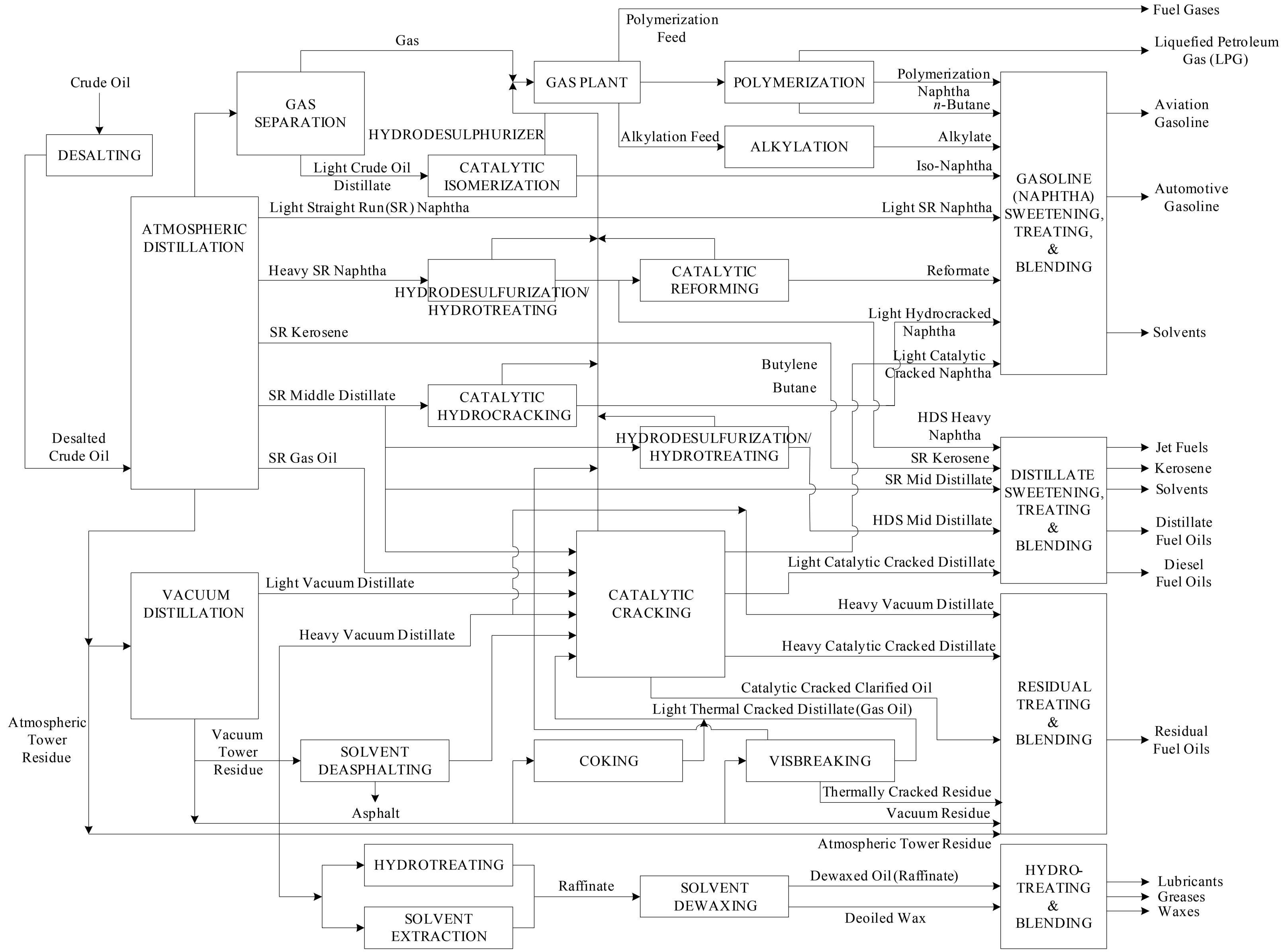


Figure 1.1. A simplified process flow diagram for a typical petroleum refinery (OSHA Technical Manual, http://www.osha.gov/dts/osta/otm/otm_iv/otm_iv_2.html, accessed on September 30, 2005; Gary and Handwerk, 1994)

Table 1.2. Summary of general processes of petroleum refining (OSHA Technical Manual, http://www.osha.gov/dts/osta/otm/otm_iv/otm_iv_2.html, accessed on September 30, 2005; Gary and Handwerk, 1994; Jechura, http://jechura.com/ChEN409/, accessed on October 17, 2005; Speight, 1998)

Process	Purpose	Function/ Action	Method	Feedstock	Catalyst	Process Variable	Product
Fractionation Processes							
Atmospheric crude distillation	Separate crudes into different boiling range fractions in preparation for further refining	Separation	Thermal	Crude oil	(none)	• Charge rate • Temperatures • Reflux rate • Process steam	• Fuel gas • Liquefied petroleum gas (LPG) • Light naphtha from butane (C_4) • Heavy naphtha • Jet fuel or kerosene • Light gas oil • Heavy gas oil • Residual/Residuum
Vacuum crude distillation	Separate crudes into fractions (same as atmospheric distillation) but without cracking	Separation	Thermal	Atmospheric tower residual	(none)	• Charge rate • Temperatures • Reflux rate • Process steam	• Gas oil • Lube feedstock • Residuum
Conversion Processes of Decomposition							
Catalytic cracking	Convert heavy oils into gasoline and lighter products (i.e., upgrade gasoline)	Alteration	Catalytic	• Gas oil • Cycle oil • Coke distillate • Residual fuel oil • Reduced crude	• Acid-treated natural aluminosilicates • Amorphous synthetic silica–alumina combinations • Crystalline synthetic silica–alumina, i.e., zeolites/molecular sieves	• Cracking temperature • Catalyst/Oil ratio • Space velocity • Catalyst type and activity (fresh) • Cracking zone's percentage of carbon on catalyst • Catalyst poisons • Recycle ratio	• For use as gasoline and petrochemical feedstock: −fuel gas, −propane, −butane, −naphtha, • Light gas oil • Heavy gas oil • Decanted oil

Hydrocracking	Convert heavy oils into gasoline and lighter HC products	Hydrogenation	Catalytic	• Gas oil • Cycle oil/Cracked oil • Coker distillate • Residual	Crystalline mixture of silica–alumina plus transition metals	• Reactor temperature and pressure • Space velocity • Hydrogen (H_2) consumption • Nitrogen (N_2) content • Hydrogen sulphide (H_2S) content	Lighter, higher quality products, e.g.,: • naphtha, • fuel gas, • light distillates from propane (C_3) and C_4
Hydrogen steam reforming	Produce hydrogen	Decomposition	Thermal/ Catalytic	• Desulfurized gas • Oxygen • Steam	Nickel-based on an α–aluminaor calcium aluminate support media	• Temperature • Pressure • Excess steam (determines equilibrium)	• H_2 • Carbon monoxide (CO) • Carbon dioxide (CO_2)
Steam cracking	Crack large hydrocarbon molecules	Decomposition (carbon rejection to redistributed hydrogen among various components)	Thermal	Atmospheric tower heavy fuel/distillate	(none)	• Temperature of superheated steam • Turbulence in reactor (prevents coke deposition)	• Cracked naphtha • Distillate • Coke • Residual
Coking	Convert heavy vacuum residuals into distillates (naphtha and gas oils) that may be catalytically upgraded (as feed to catalytic cracker)	Polymerization	Thermal	• Gas oil • Coke distillate	(none)	• Pressure • Temperature • Recycle ratio	• Gasoline • Petrochemical feedstock
Visbreaking	Reduce viscosity and pour points (or reduce amount of cutting stock required to dilute residual)	Decomposition	Thermal	Atmospheric tower residual	(none)	• Temperature • Pressure • Standard cutted stock amount for blending visbreaker tar • Fuel consumption • Feedstocks conversion and characteristics • Gas oil density and viscosity	• Distillate • Tar

Conversion Processes of Unification							
Alkylation	Unite light (i.e., low molecular weight) olefins (mainly butylenes) with isoparaffins to producte high-octane (i.e., higher molecular weight) gasoline	Combination	Catalytic	• Tower isobutane + isobutylene or + isopropylene • Cracker olefins	• Hydrofluoric acid • Sulphuric acid	• Reaction temperature • Type of olefin (propylene, butylenes, or pentene) • Isobutane concentration • Olefin space velocity (injection and mixing) • Catalyst (acid) type and strength	• Iso-octane (alkylate) gasoline blending stock • Normal propane and butane included in the feed • Tar
Grease Compounding	Combine/Blend metallic soaps (salts of long-chained fatty acids) and additives into lubricating oil medium	Combination	Thermal	• Lube oil • Fatty acid • Alkyl metal	(none)	• Metallic element (calcium, sodium, aluminium, lithium, etc.) in soaps and additives • Viscosity	Lubricating grease
Catalytic polymerization	Convert propylene and butylenes to a gasoline blending stock	Polymerization		• Propylene • Butylene	Phosphoric acid on crushed quartz or other porous solid carrier	• Temperature • Pressure • Feed composition • Catalyst activity • Recycle rate	• Gasoline blending stock • C_3 • C_4 • Fuel gas
Conversion Processes of Alteration or Rearrangement							
Catalytic Reforming	• Upgrade low-octane naphtha to produce a high octane (C_8) gasoline blending stock, • H_2 is produced as an important by-product	Alteration/ Dehydration	Catalytic	• Naphtha of gasoline boiling range, • Sometimes, other gasoline blending stocks, e.g., gasoline from fluid catalytic cracker (FCC), • Coker, • Hydrocracker naphtha	• Platinum on a silica or silica–alumina base, • Rhenium or another metal used with platinum	• Temperature, • Pressure, • Partial pressure of H_2, • Space velocity, • Catalyst activity • Catalyst poisons	• High C_8 gasoline blending stock or reformate • H_2 fuel gas • C_3 and C_4 • Aromatics

Isomerization	Product higher octane gasoline blending stock by converting straight chain to branch	Rearrangement	Catalytic	• Butane • Pentane (at 180°C) • Hexane	• Platinum on a silica or silica–alumina base • Molecular sieve	• Reactor temperature and pressure • Space velocity • Partial pressure of H_2 • Catalyst activity (i.e., catalyst type and age)	High percentage isobutene, isopentane, and isohexane for gasoline blending stock
Aromatics Production	• Separate high purity aromatics from solutions of other hydrocarbons • Production of low aromatic raffinate for jet fuel, high smoke point kerosene, and specialty products			Coke oven sources or petroleum (normally reformate)	(none)	• Temperatures • Densities • Efficiency of solvent • Reflux ratios • Composition of feed	• Benzene • Toluene • Xylene • Ethylbenzene • Nonane (C_9) + aromatics • Jet fuel • High smoke point kerosene • Specialty products
Treatment Processes							
Desalting	Remove contaminants (salt and suspended solids) from crude oil	Dehydration	Absorption	Crude oil	(none)	• pH, gravity, and viscosity of crude oil • Volume of wash water per volume of crudes • Temperature • Water/Oil ratio	Desalted crude oil
Hydrotreating	• Catalytically stabilize petroleum streams and/or • Remove objectionable objects by reacting with H_2	Hydrogenation	Catalytic	A wide range from naphtha to reduced crude of cracked hydrocarbons (HCs) to residuals	Cobalt and molybdenum oxides on alumina, nickel oxide, nickel thiomolybdate, tungsten, and nickel sulfides	• Temperature • Partial pressure of H_2 • Space velocity	• Stabilized poison-free feedstock, e.g., for cracker (FCC) feed • Distillate • Lube
Hydrodesulphuri-zation	Remove sulphur, contaminants	Treatment	Catalytic	High-sulphur residual/gas oil	Cobalt–molybdenum	• Temperature • Partial pressure of H_2 • Space velocity	Desulphurized olefins

Furfural extraction of lube oil stocks	Upgrade mid distillate and lubes to produce various grade of lubricating oils	Solvent extraction	Absorption	• Vacuum unit gas • Oils and deasphalted vacuum unit bottoms • Decanted oil from FCC • Cycle oils • Lube feedstocks	(none)	• Temperature gradient • Recycle rate • Solvent/Oil ratio • Charge rates	• High quality purified diesel and lube oil stock (raffinate) • Lube oil impurities (extract)
Solvent dewaxing of lube oils	Remove wax from lube oils, thus improving their pour points	Treatment	Cool/ Filter	Raffinate from furfural extraction of vacuum tower lube oils	(none)	• Temperature • Solvent efficiency • Solvent/Oil ratio • Solvent recycle rate	• Waxes • Wax-free (dewaxed) lube oils basestocl
Amine treating	Remove acidic contaminants	Treatment	Absorption	• Sour gas • Hydrocarbons with CO_2 and H_2S	(none)	• Flow rate • Composition • Temperature • Pressure • Degree of recovery	Acid-free gases and liquid HCs
Drying and sweetening	Remove water and sulphur compounds	Treatment	Absorption/ Thermal	• Liquid HCs • LPG • Alkyl feedstock	(none)	Solubility	Dry and sweet HCs
Phenol extraction	Improve viscosity index and colour	Solvent extraction	Absorption/ Thermal	Lube oil base stocks	(none)	Solubility	High quality lube oils
Solvent deasphalting	Remove asphalt	Treatment	Absorption	• Vacuum tower residual • Propane	(none)	• Temperature • Pressure • Concentration of mixtures • Solubility	• Heavy lube oil • Asphalt
Solvent extraction	Separate unsaturated oils	Solvent extraction	Absorption/ Precipitation	• Gas oil • Reformate • Distillate	(none)	• Solubility • Porosity	High-octane gasoline
Sweetening	• Remove H_2S, thiophene, and mercaptan to improve colour, odour, and oxidation stability • Convert mercaptan • Reduce CO_2 concentration	Treatment	Catalytic	Untreated distillate/gasoline	(none)	• Nature of fraction • Amount and type of impurities • End-product specifications	High quality distillate/gasoline

Petroleum refining is undoubtedly, one of the most complex chemical industries, comprising many different and complicated processes with various possible configurations and structures, as evidenced from Table 1.2. The critical objective of a refinery operation, as in any other business-oriented ventures, is to generate maximum profit by converting crude oils into valuable products such as gasoline, jet fuel, and diesel. Expectedly, there are many decisions to be considered to achieve optimal operation for a refinery. At the planning level, managers and executives need to decide the types of crude oil(s) to process, the types of products to produce, the operating route to use, the best operation mode for each process, the type of catalyst to select for each process, and others. At the process level, engineers and operators have to determine detailed operating conditions for each piece of equipment, namely temperatures, pressures, detailed process flow, and other values of processing parameters. All these decisions interact with one another; for example, temperature change in a reactor would result in different product yields and distribution as well as different utility consumption, hence different process performance would result. These are bound to implicate and affect the decisions made at the planning level to select raw material feeds for the processes involved and even possibly influence the overall operating scheme. Consequently, integration of refinery planning, scheduling, and operations optimization, or integrated (total) refinery optimization for short, is considered one of the most difficult and challenging applications of large-scale optimization but the expected outcome would be commensurable with the effort, time, and resources invested (Zhang and Zhu, 2000).

1.3.2 Production Planning and Scheduling

Production planning is the discipline related to the high level decision-making of macro-level problems for allocation of production capacity (or production levels) and production time (with less emphasis on the latter); raw materials, intermediate products, and final products inventories; labour and energy resources; as well as investment in new facilities. A coarse aggregation approach is typically employed, thus resulting in a loss of manufacturing detail such as the sequence or the order in which specific manufacturing steps are executed (Pekny and Reklaitis, 1998). If follows then that the primary objective of planning is to determine a feasible operating plan consisting of production goals that optimizes a suitable economic criterion, namely of maximizing total profit (or equivalently, of minimizing total costs), over a specific extended period of time into the future, typically in the order of a few months to a few years; given marketing forecasts for prices, market demands for products, and considerations of equipment availability and inventories (Reklaitis, 1982; Birewar and Grossmann, 1995; Grossmann et al., 2001; Bitran and Hax, 1977). In essence, its fundamental function is to develop a good set of operating goals for the future period. In the context of the hydrocarbon industry, planning requirements have become increasingly difficult and demanding arising from the need to produce more varied, higher-quality products while simultaneously meeting increasingly tighter environmental legislations and policies (Fisher and Zellhart, 1995).

On the other hand, production scheduling, in the context of the chemical processing industry, deals with lower level decision-making of micro-level problems embedded in the production planning problem that involves deciding on the methodology that determines the feasible sequence or order and timing in which various products are to be produced in each piece of equipment, so as to meet the production goals that are laid out by the planning model. Its major objective is to efficiently utilize the available equipment among the multiple types of products to be manufactured, to an extent necessary to satisfy the production goals by optimizing a suitable economic or systems performance criterion; typically over a short-term time horizon ranging from several shifts to several weeks. Scheduling functions specify the task(s) of each stage of production and this includes defining and projecting the inputs to and outputs from each production operation. It is particularly required whenever a processing system is used to produce multiple products by allocating the available production time between products. A key characteristic is the dynamic and extensive information required in scheduling activities to describe the manufacturing operations, the resource requirements, and the product demands. The sources of information are diverse and extend outside of the boundaries of the manufacturing organization itself since the information spans the technical, financial, and commercial domains. Furthermore, the data changes rapidly over time as customer orders, resource availability, and the manufacturing processes themselves undergo changes. Therefore, the resulting data complexity compels efficient management of information resources a necessary prerequisite for effective scheduling (Reklaitis, 1982; Birewar and Grossmann, 1995; Shobrys, 1995; Pekny and Reklaitis, 1998).

Hartmann (1998) and Grossmann et al. (2001) stresses the differences between a planning model and a scheduling model. In general, process manufacturing planning models consider economics of profit maximization of the operations by handling the issues of what to do and how to do it within longer time horizons. Process manufacturing scheduling models, on the other hand, consider feasibility of the operations for accomplishing a given number of orders or on completing required tasks within the shortest possible time, by addressing the question of when to do it. In particular, planning models ignore changeovers and treat products grouped into aggregated families. Conversely, scheduling models explicitly consider changeovers and consider products in greater detail, including the shipment of specific orders for specific products to specific customers.

Fisher and Zellhart (1995) also emphasizes that a planning model differs from a daily scheduling model or an operational process controller. For example, they point out that the product or process yields predicted or estimated in the planning model should not be expected to be used exactly in executing operating conditions. This is because planning models are almost always an average over time and not an accurate prediction of process conditions at any particular instant. As opposed to planning models, operations are not averaged over the scheduling period as time and operations move continuously from the beginning of the particular period to the end. The schedule is revised as needed so that it always starts from what is actually happening with revisions typically occur on each day or on each shift.

Scheduling can be viewed as a reality check on the planning process. The objective of scheduling is the implementation of the plan, subject to the variability that occurs in the real world. This variability could be present in the form of feedstock supplies and quality, the production process, customer requirements, or transportation. Schedulers assess how production upsets and other changes will force deviations from the plan, and they determine the actions to be taken in making corrections that would meet the plan objectives

Thus, scheduling appears to be the most active juncture of business and manufacturing systems. Schedulers are continually assessing how the capabilities of the production process compare to the needs of the business. On a daily basis, a scheduler has to react to process variability as well as business variability that can impact feedstock arrivals and product movements. Scheduling involves dynamic interactions with the business (marketing and customer service) as well as the manufacturing process and distribution. Additionally, human factors add unpredictability to these interactions. On the other hand, planning activities consider both business and manufacturing but the plans are updated less frequently and consider less detail. In addition to planning and scheduling functions, the third dimension in process plant management concerns operations control, in which applications focus on the manufacturing process with the deployment of distributed control systems (DCS) providing control capabilities for specific parts of the overall process (Shobrys, 1995).

Nevertheless, despite the differences, it is obvious that production planning, scheduling, and operations control are all closely-related activities. Decisions made at the production planning level have a great impact at the scheduling level, while the scheduling in itself determines the feasibility of executing the production plans with the resulting decisions dictating operations control. Ideally, all three activities should be analyzed and optimized simultaneously, thus calling for the need of the integration between planning, scheduling, and operational activities, with the expectation that this would greatly enhance the overall performance of not just the refinery or process plant concerned, but the parent governing organization as well. However, this is in general a difficult task given that for instance, even optimizing the scheduling problem in isolation for fixed production demands is a nontrivial problem, as highlighted by Birewar and Grossmann (1990) and emphasized in general by Bodington (19950. However, the recent survey by Grossmann et al. (2001) pointed out that the distinction between planning and scheduling functions is becoming increasingly blurred as evidenced by recent advances in the capability of the simultaneous optimization of planning and scheduling decisions, especially in the context of supply chain optimization problems. This has certainly promises greater hope for addressing the issue of integrated planning, scheduling, and operations, which provides the motivation for the following section.

1.3.3 The Need for Integration of Planning, Scheduling, and Operations Functions in Petroleum Refineries

For the process industry in general, Bassett et al. (1996) define the term *process operations* as tasks that must be addressed in managing a process plant so as to safely and efficiently manufacture a desired slate of products. As mentioned earlier, these tasks are principally composed of planning, scheduling, and operations, with the latter consisting of supervisory control, fault diagnosis, monitoring, regulatory control, and data acquisition and analysis. The tasks are conventionally viewed to be related in a hierarchical fashion with long-term strategic planning decisions imposing goals, targets, and constraints on midterm tactical decisions, which are in turn, implemented and supported via a number of operational execution functions. All these decision-making activities draw upon the enterprise information systems base, which forms the necessary foundation upon which other levels are grounded, as depicted in Figure 1.2, which is tailored for a petroleum refinery, but is in fact, sufficiently generic across all manufacturing entities. In addition, it is desirable to extend the scope of this hierarchy to include the highest level of strategic decision making, that is, the planning and design of production capacities required for future operation. While these levels can be viewed to constitute a hierarchy, the requirements of hierarchy dictate that these levels communicate bidirectionally, that is, in a two-way interactive dynamics between the different levels, with the lower levels communicating suitably aggregated performance limits and capacities to the upper levels. This is essentially the challenge of integrating the planning, scheduling, and operations functions of a process plant, primarily the flow of information between the various levels, in which petroleum refineries stand out as a prime example for the multifarious tasks involved that typically span several business and operation departments, handling large amount of data (Julka et al., 2002) in dealing with activities such as crude oil procurement, logistics of transportation, and scheduling of processes (for example, storage tanks and distillation units).

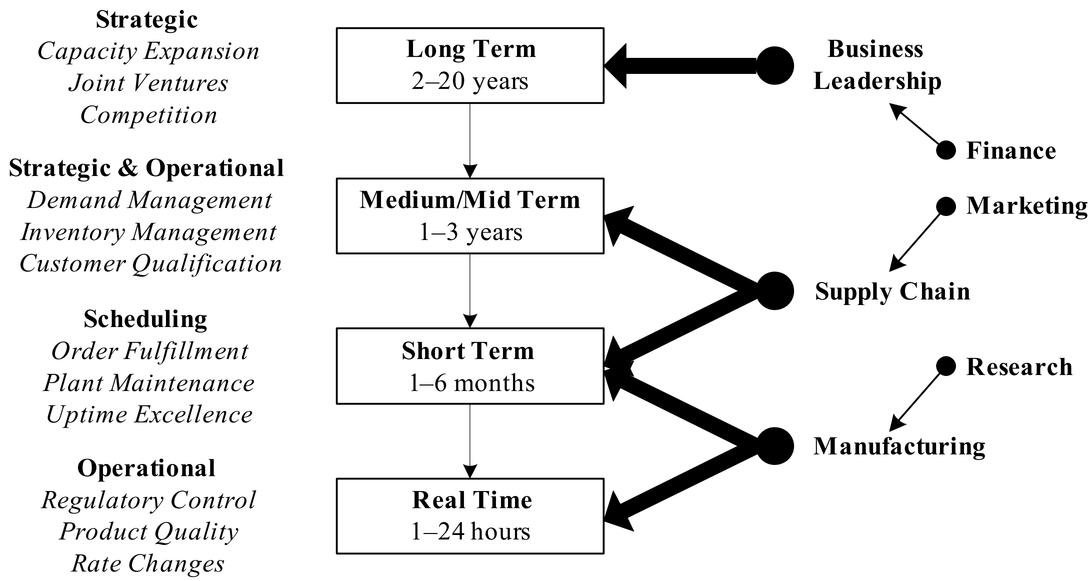


Figure 1.2. Typical functional hierarchies of corporate planning activities (McDonald, 1998)

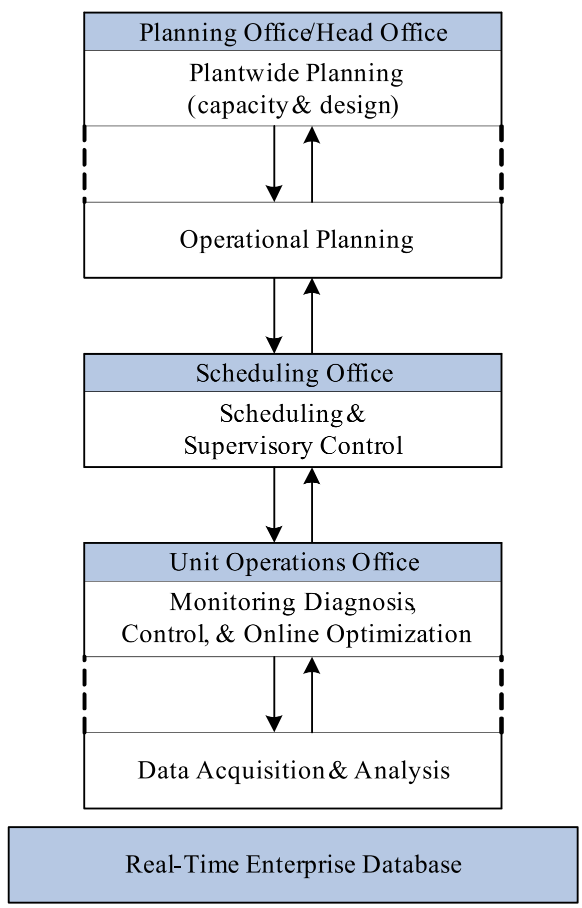


Figure 1.3. Structure of management activities in an enterprise, typically for a petroleum refinery (adapted from Li (2004) and Bassett et al. (1996))

Additionally, over the past decade or so, many companies have become aware that improvement in the performance of the supply chain from customer order to product delivery is essential to their continued success and sustainability, or even their mere survival as a business entity. In the petroleum industry, traditionally known for its sceptical view and slow response towards shifts in contemporary business practices (partly attributable to the high risk associated with its capital-intensive nature), there exists an increasingly widening organizational and operational gap in the supply chain between the activities of planning for the business (corporate planning) on one end and the scheduling and control of processes (operational planning) to meet commitments on the other end (Bodington, 1995). Julka et al. (2002) highlight one of the primary reasons as the incapability of currently available refinery decisions support systems (DSSs) in effectively performing the following functions of: (i) integrating all the decision-making activities within a refinery; (ii) interfacing with other co-existing DSSs; (iii) incorporating dynamic-state data from various sources within and outside of the refinery (for instance, from suppliers and vendors); and (iv) assisting functions of other departments concurrently.

In the wake of the onslaught of political and economic (and even social) globalization, coupled with the inherent complexities in the management of petroleum refineries as stressed earlier, there is immensely increased emphasis on integrated refinery optimization. This is typically declared in the overall objective of both corporate planning and operations planning activities to align production activities with business objectives in realizing a single ultimate goal of profit maximization. In simple terms, this basically trickles down to operating process units in such a way so as to generate maximum profit. On the one hand, rapid development of computing, information and communications technology (ICT (or just IT)), and its decreasing cost of deployment have tremendously aided and improved the manner in which refineries are operated. But on the other hand, they have triggered intense competition among refineries located both within the local geographic region and abroad, in executing the core activities of purchasing of crude oils and marketing of refined saleable commercial products. Moreover, refining activities are subjected to increasingly stringent environmental regulations such as allowable limits of sulphur content in gasoline and diesel. These regulations inevitably impose significant impacts on the profitability and the ensuing competitiveness of a refinery. These challenges are prompting refineries to continuously seek, demand, and implement various effective and efficient tools and technologies with the ultimate aim of total refinery optimization through integrated planning, scheduling, and operations functions (Li, 2004; Bodington, 1995). Of late, the current drive towards enterprise-wide optimization (Grossmann, 2005; McDonald, 1998, with the latter apparently appearing to be overlooked in related literature) offers an indication of renewed concerted effort towards this end, aided especially with the explosive improvement (which is still revolving) in scientific computing and information technology in recent years.

1.3.4 Planning, Scheduling, and Operations Practices in the Past

In the past, (strategic) planning in most non-integrated situations is performed by one entity close to the marketing and supply functions, but not part of them. Planning activities serve to consolidate feedstock purchases, commitments, and sales opportunities by attempting to set achievable targets for the plant. Scheduling is undertaken by another entity that stands between planning and operations. It attempts to produce a schedule that is feasible, if not optimal, to meet commitments. Process operations are handled by yet another entity, usually compartmentalized by processes, that operates the processes to the best capability, given the information available from planning and scheduling activities. The three entities have different objectives and possibly have different reward motivations and reward structures, which lead to different philosophies of what constitutes a job well done (Bodington, 1995)

On the whole, the petroleum industry has invested considerable effort in developing sophisticated mathematical programming models to help planners provide overall strategy and direction for refinery operations, crude oil evaluation, and other related tasks. Likewise, there has been substantial and extensive development and implementation of tools for scheduling. In addition, considerable efforts have been assumed in advanced process control for process plants to enable plants to run close to their optimal operating conditions. Unfortunately, a gap has existed between the three activities, with inadequate attention and effort invested to providing tools that aid the planner, the scheduler, and the operator in an integrated environment, as reported by Fisher and Zellhart (1995).

Typically, the refinery scheduler attempts to use the monthly linear programming (LP) model strategic plan to develop a detailed day-to-day schedule for refinery operations based on scheduled crude and feedstock arrivals, product liftings, and process plant availabilities and constraints. The schedule usually includes details of the operation of each process units, the transfer of intermediates to and from the tank farm, and product blending schedules. However, the scheduling is performed for each tank instead of for a pool. Moreover, most refinery schedulers have few extensive computing tools to accomplish this task. Many use spreadsheets that contain individual operating modes for the primary processes and for the main feedstocks, based on the same data employed in the LP model. The scheduler utilizes the spreadsheet to generate manufacturing plans on a daily or weekly basis. It has been reported that some even use just plain paper, pencil, and calculators as the their only aids in daily scheduling (Fisher and Zellhart, 1995). Compounding the problem is the fact that deficiencies in planning or operations often create problems that appear in the scheduling process. Operating deficiencies or inferior data on the status of the production process could potentially lead to customer service problems. These problems may also occur due to either a planning activity with an overly optimistic estimate of available capacity or a poor understanding of the production capabilities. Additionally, operations staff may not be effective or well-trained enough in executing activities as scheduled (Shobrys, 1995).

1.3.5 Mathematical Programming and Optimization Approach for Integration of Planning, Scheduling, and Operations Functions in Petroleum Refineries

The chemical process industry, as pointed out earlier, has been increasingly pursuing the use of computing technology to gather, organize, disseminate, and exploit enterprise information and to closely coordinate the decisions made at the various levels of the process operational hierarchy so as to optimize overall corporate objectives. In refinery management, computer software is commonly deployed nowadays to assist in terms of planning, scheduling, and control functions by executing effective decisions chiefly pertaining to crude oil selection, production planning, inventory control, and logistics of transport and despatch management. Continuous research and development in these aspects have gained important practical significance, as observed by Li (2004). In this respect, we support the notion advanced by Li (2004), Zhang and Zhu (2000), Bassett et al. (1996), and Bodington (1995), just to cite a few among many others, that the preferred approach for achieving integration of planning, scheduling, and operations functions is through the formulation and solution of suitably structured mathematical programming models as they have been proven to offer the most effective tools. Indeed, it is the governing theme of this work that mathematical optimization constructs offer the most effective framework for integration at the strategic, tactical, and operational levels of refineries. This shall provide the thrust for undertaking the current work in this thesis research with the ultimate objective of developing better management tools for decision-makers. In particular, we consider the mathematical programming approaches for modelling under uncertainty in the problem parameters of the midterm planning of a refinery.

As emphasized throughout the preceding discussion, the structure of the main management activities of an integrated refinery consists of three layers: (1) planning at the strategic level; (2) scheduling at the tactical level; and (3) unit operations at the operational level, as illustrated previously in Figure 1.3. First, the planning office, typically the head office, issues plans that are sent to the scheduling office as guidelines. The scheduling office then decides on the detailed daily or weekly schedules for each unit and subsequently sends these schedules to the unit operation office as operating guidelines (Li, 2004).

The head office produces plantwide high-level strategic plans and tactical plans for the refinery by considering plantwide factors in the form of market conditions, raw materials availability, and operating capacities. These plans deal with business decisions such as which units to run, which raw material(s) to process, and which products to produce. The high-level strategic plans, in general, relate to a period of several years. On the other hand, the tactical plans for local refinery management control are based on a refinery's strategic business plan and are executed on a monthly, quarterly, and annual budget. Further, according to Favennec (2001), the tactical plans include the use of management monitoring processes and the regular calculation of the commercial results that are often aligned with monthly accounting procedures.

On the whole, plantwide planning activities, with the underlying objective of seeking optimal operating strategy than can maximize total profit, are obviously crucial to the economics of a company (Li, 2004). In this work, several framework of midterm to long-term strategic planning models are developed for a typical medium-sized refinery with basic configuration.

The plantwide plans of the head office are then delivered to the scheduling office to act as a guideline. The scheduling office then determines the detailed timing of actions that are to be carried out in a plant within the specified ranges of the plantwide plans. Generally, the scheduling time horizons stretch from one week to ten days. The objective of scheduling is to seek feasible operating strategies that satisfy the planning requirements while simultaneously minimize the operating cost. Li (2004) highlighted that refinery scheduling pose one of the most challenging refinery management activities simply because the currently available technology and knowledge-base is still immature relative to the complexity that is demanded from it.

Refinery scheduling is further divided into three components as follows (Li, 2004):

- crude oil tankage area or tank farm scheduling that handles crude oil storing, transporting, and charging activities;
- refining area scheduling that establishes the various unit operations' operating conditions and flow rates of stream flows; and
- blending area scheduling that decides blending recipe for intermediate streams to produce products that meet quality specifications while maintaining appropriate product inventory levels.

Subsequently, the detailed scheduling results are sent to the unit operations office (which is typically housed in the plant's main control room, often dubbed as the heart of the plant) to enable the operators to run the units in such a way so as to realize the outlined scheduling objective. Various tools and performance criteria of the operational and logistical system for purposes of monitoring, diagnosis, control, and online optimization (or also known as real time optimization, RTO) of systems and processes are utilized to optimize the performance of the unit operations.

1.3.6 Current Persistent Issues in the Planning, Scheduling, and Operations Functions of Petroleum Refineries

According to Fisher and Zellhart (1995), planning and scheduling for a refinery typically encompass three areas: (1) crude oil management, (2) process unit optimization, and (3) product scheduling and blending.

Crude management entails crude segregation and crude unit operation. Process unit optimization deals with downstream (of the crude distillation unit (CDU)) process unit operations that handle crude unit intermediates. Product scheduling and blending handles the development of a product shipment schedule and an optimum blend recipe based on information from process unit optimization and current operating data.

A major problem in refinery planning is prevalent even at the very foundation: optimization of the CDU and its associated product yields. In addition to uncertainty surrounding the future price of crude oils, the actual composition of crude oils (or crudes, for short) is often only an educated guess. Crudes vary from shipment to shipment because of the mixture of sources actually shipped. It is expected that the quality of crudes does not significantly change over a short period of time, although this assumption could also render a plan to be inaccurate or worse, infeasible. If the actual crude composition does not closely agree with that modelled, then an error is committed that often propagates through the rest of a refinery planning model.

A second, equally common source of error in optimizing the submodel for the CDU is the assumption that the fractions from the distillation curve for the crude unit, or simply referred to as the crude cuts or the swing cuts of distillates, are produced as modelled. Frequently in practice, models are not even adjusted to show cut overlaps, all just because of wishing to take the easy way out in developing crude cut yields and distillates. One of the typical crude cutting procedures assigns distillation temperatures directly from the true boiling point crude analysis, in which no adjustment is made for the actual refinery degree of fractionation. This is a particular bad procedure for certain types of gasoline that have tight 90 percent point limits. The fractionation efficiency of gasoline and distillate components from all processes would have a significant effect in controlling aromatics and other types of hydrocarbons. Therefore, planners and decision makers ought to be more diligent by constantly reviewing the supposed optimized plans and comparing to actual situations in an effort to improve the prediction accuracy of their models.

The third component of the refinery planning and scheduling functions involve product scheduling and blending where this is usually handled by preparing both a short-range and a long-range plan, using the same model for the blending process. The long-range plan, typically covering 30 days, provides aggregate pools of products for a production schedule. The short-range plan, typically spanning seven days, fixes the blend schedule and creates recipes for the blender. Desired output from the long-range model includes (i) detailed product blend schedule; (ii) optimal blend recipes; (iii) predicted properties of blend recipes; (iv) product and component inventories; (v) component qualities, rundown rates, and costs; (vi) product prices; and (vii) equipment limits. For the short-range model, the desired output are: (i) a detailed product blend schedule; (ii) optimal blend recipes; (iii) predicted properties of blend recipes; and (iv) product and component inventories as a function of time (Fisher and Zellhart, 1995).

1.3.7 Petroleum Refinery Production and Operations Planning under Uncertainty

In the discussion in preceding sections, we emphasize our conviction in mathematical programming techniques under uncertainty, specifically stochastic programming methods, towards improvement in tools and methodologies for integrating the planning, scheduling, and operations functions of a refinery. This stems from the fact that it has long been recognized that traditional deterministic refinery planning models are not suitable for capturing the dynamic behaviour of the highly volatile oil and gas industry due to the presence of data uncertainties, in which exact information that will be needed in subsequent decision stages is not usually available to the decision maker when a decision must be made. Although majority of the works on optimization of refinery planning models are still based on deterministic programming, there has actually been quite a substantial body of work that addresses the issue of uncertainty in market conditions, mainly concerning product demands and prices (or costs) of crude oil and the saleable refined products. Table 1 attempts to provide a brief survey of recent works in the petroleum industry supply chain planning and optimization under uncertainty with focus on works purely addressing refinery production–operations planning.

Table 1.3. Recent works (in chronological order of descending recency) on petroleum industry supply chain planning and optimization under uncertainty with focus on refinery production–operations planning

Author (Year)	Application	Uncertainty Factor	Stochastic Modeling Approach
Pongsakdi et al. (in press)	Refinery production planning with emphasis on financial risk management	Product demands and prices	Two-stage SP with scenario analysis—a full deterministic model is run for parameters of each scenario and the results are used to fix the first-stage variables; then, the same model is rerun for the rest of the scenarios to obtain second-stage values
Neiro & Pinto (2005)	Multiperiod refinery production planning for selection of different crude oil types under uncertainty and crude oil handling constraints	Prices and demands of crude oil and products	Two-stage SP with scenario analysis for MINLP model
Aseeri & Bagajewicz (2004)	Measures and procedures for financial risk management in the planning of natural gas commercialization (in the Asia region)	Demand and prices	Two-stage SP with scenario analysis for MILP model (by varying transport process selection, expansion capacities, and production rates)
Aseeri et al. (2004)	Financial risk management of offshore oil (petroleum) infrastructure planning and scheduling to determine the sequence of oil platforms to build and the wells to drill as well as how to produce these wells over a period of time (with introduction of budgeting constraints that follow cash flow of the project, take care of the distribution of proceeds, and consider the possibility of taking loans against some built equity)	Oil prices and oil production (modelled via a productivity index)	Two-stage SP with sampling average algorithm (SAA) for MILP model

Table 1.3. Recent works (in chronological order of descending recency) on petroleum industry supply chain planning and optimization under uncertainty with focus on refinery production–operations planning (*continued*)

Author (Year)	Application	Uncertainty Factor	Stochastic Modeling Approach
Goel & Grossmann (2004)	Optimal investment and operational planning of offshore gas field developments under uncertainty in gas reserves	Size and initial deliverability of gas fields	Multistage SP as a sequence of two-stage SP using conditional non-anticipativity constraints incorporating decision-dependence of the scenario tree via hybrid mixed-integer/disjunctive programming
Lababidi et al. (2004)	Supply chain of a petrochemical company	Market demand, market prices, raw material costs, and production yields	Two-stage SP with scenario analysis for MINLP model
Li et al. (2005); Li (2004), Li et al. (2004)	Planning, scheduling, and economic analysis of refinery management with the integration of production and energy systems	Raw material costs, product demands, and other changing market conditions	Two-stage SP with penalty functions replaced by decision maker's service objectives of confidence level (probability of satisfying customer demands) and fill rate (proportion of demands met by plant) evaluated by loss functions
Jia & Ierapetritou (2003)	Mixed-integer linear programming model for gasoline blending and distribution scheduling		
Hsieh & Chiang (2001)	Manufacturing-to-sale planning system for refinery fuel oil production	Demand and cost	Fuzzy possibilities linear programming
Dempster et al. (2000)	Multiperiod supply, transformation, and distribution (STD) scheduling problem for strategic or tactical level planning of overall logistics operations in the petroleum industry	Product demands and spot supply costs/prices	Dynamic SP with scenario analysis
Escudero et al. (1999)	Multiperiod supply, transformation, and distribution (STD) scheduling problem	Product demand; product spot market supplying cost; product spot market selling price	Two-stage SP with scenario analysis based on partial recourse approach
Guldmann & Wong (1999)	Optimal selection of natural gas supply contracts by local gas distribution utilities	Weather variability	Simulation and response surface estimation via regression analysis of a large MILP and a much smaller NLP approximation of the MILP
Bok et al. (1998)	Investment planning in the South Korean petrochemical industry	Product demand	Two-stage SP for a multiperiod MINLP model

Table 1.3. Recent works (in chronological order of descending recency) on petroleum industry supply chain planning and optimization under uncertainty with focus on refinery production–operations planning (*continued*)

Author (Year)	Application	Uncertainty Factor	Stochastic Modeling Approach
Liu & Sahinidis (1997)	Process planning with example for a petrochemical complex	Material availabilities, product demands, material costs, product prices, process yields	Fuzzy programming for MILP or MINLP (depending on type of uncertainty)
Liu & Sahinidis (1996)	Process planning	Prices, demands, availabilities	Two-stage SP
Bopp et al. (1996)	Managing natural gas purchases	Demands, frequent price change	Scenario analysis
Clay and Grossmann (1994)	Stochastic planning with example on refinery planning (based on Edgar et al., 2001)	Market demands	Two-stage SP with fixed recourse

Nomenclature:
SP: stochastic programming
MILP: mixed-integer linear programming
NLP: nonlinear programming
MINLP: mixed-integer nonlinear programming

1.3.8 Factors of Uncertainty in Petroleum Refinery Production and Operations Planning

In the spirit of the recent work by Goel and Grossmann (2004), we classify possible factors of uncertainty in the planning of the production and operations of a petroleum refinery into two classes, namely the exogenous or external factors and the endogenous or internal factors, as shown in Table 1.4.

1.3.9 Production Capacity Planning of Petroleum Refineries

The planning and utilization of production capacity is one of the most important managerial responsibilities for managers in the manufacturing industry, including petroleum refineries. Such decisions have to be made in the face of uncertainty in several

Table 1.4. Possible factors of uncertainty in a petroleum refinery planning problem (Maiti et al., 2001; Liu and Sahinidis, 1997)

Exogenous (external) factors

- Availabilities of sources of crude oil (raw material) supply
- Economic data on feedstock, intermediates, finished products, utilities, and others
  - Prices of crude oil and chemicals
  - Production costs
  - Distribution costs
  - Market demands
- Production demands: final product volumes & specifications
- Location
- Budgets on capital investments for capacity expansion and new equipment purchases or replacements
- Investment costs of processes (for example, licence fees to be paid to process licensors providing use of a certain refining process technology such as UOP (Universal Oil Products)

Endogenous (internal) factors

- Properties of components
- Product/process yields
- Processing and blending options
- Machine availabilities

important parameters, with the most important of these uncertainties being market demand for the products being manufactured. Hence, manufacturing capacity planning has long attracted the attention of economists as well as researchers in the practice of its traditional domain, namely operations research and management science.

According to Escudero et al. (1993), there are two types of capacity planning problems. The more commonly discussed problem of deciding how much capacity to acquire and how to plan its utilization is a strategic problem that deserves careful analysis. On the other hand, in the tactical time horizon, the second-type of capacity problems are normally resolved through inventory buffers, additional workloads, or through alternate sourcing. Although new capacity cannot be acquired in this time horizon, it is often possible to develop alliances with other manufacturers or vendors to manage the production of uneven or unanticipated production volumes.

Sahinidis and Grossmann (1989) state that a considerable number of works has been reported, particularly in the operations research literature, concerning capacity expansion problems in several areas of application. A classic review on this subject can be found in Luss (1982). In the chemical engineering literature, a variety of methods has been applied to expansions of chemical plants, for example, (i) dynamic programming by Roberts (1964); (ii) branch-and-bound procedure combined with generalized reduced gradient of constrained nonlinear programming (NLP) algorithm by Himmelblau and Bickel (1980);

(iii) multiperiod mixed-integer linear programming (MILP) formulation by Grossmann and Santibanez (1980); (iv) goal programming by Shimizu and Takamatsu (1985) via a procedure of stepwise and subjective judgement process by the decision-maker by relaxation of flexible constraints and sensitivity analysis of linear programming, and (v) recursive MILP by Jimenez and Rudd (1987) to achieve an optimum integration sequence for a petrochemical industry. However, they are often ineffective for large-scale problems and are thus limited in the size of problems that can be handled. In addition to that, of particular interest is the problem of capacity expansion under uncertainty via the scenario analysis approach, of which Eppen et al. (1989) is a frequently-cited work that treats a real-world problem in the automobile industry by accounting for the expected downside risk.

However, capacity expansion will not be considered in this work since the primary objective is to focus on developing methodologies and tools for planning under uncertainty. It will therefore be left to future work.

1.3.10 Other Applications of Stochastic Programming Models in the Hydrocarbon Industry

Wallace and Fleten (2003) briefly discuss the applications of stochastic optimization models in the oil and gas (and petrochemicals) industry, in addition to the problem of refinery planning addressed in this work. They include the following:

- optimum oil field development to determine platform capacity for well drilling and production operations; number of wells including their placement and timing; and the production profile of wells, with stochasticity in random future oil prices (Jonsbraten (1998) described using scenarios;
- scheduling arrivals of tankers at a refinery for loading of gasoline for export;
- scheduling of gas fields production to decide on the location and timing of fields that should be developed and the ensuing pipelines that should be constructed;
- planning of gas storage facilities for contracted delivery, participation in potentially profitable spot markets, and others;
- portfolio management of natural gas contracts.

1.4 STOCHASTIC PROGRAMMING (SP) FOR OPTIMIZATION UNDER UNCERTAINTY

Stochastic programming (SP), as indicated by the section heading, is the subfield of mathematical programming that considers optimization in the presence of uncertainty (Dyer and Stougie, 2006). Within the context of modelling , it is the optimization branch explicitly concerned with models with random parameters (Birge and Louveaux, 1997). It is referred to as the study of practical procedures for decision making under the presence of uncertainties and risks (Uryasev and Pardalos, 2001). Further, according to Roger J.-B Wets (1996), arguably one of the most prominent theorists of the field, the motivation for modelling decision problems as stochastic programs is derived mostly from the search for a “robust” first-stage decision, that is, “a decision that will put the decision maker in a rather good position in whatever, or almost whatever, be the outcome of future events.” Even though 50 years have eclipsed since the pioneering seminal works of Dantzig (1955) and Beale (1955), George B. Dantzig still considers planning under uncertainty as the definitive open problem of utmost importance in the field of optimization (Horner, 1999). Thus, this augurs well for the consideration of uncertainty in the refinery production planning problem addressed in this work.

Stochastic programming deals with optimization problems that are characterized by two essential features: the uncertainty in the problem data and the sequence of decisions, in which some of the model parameters are considered random variables that take values from given or assumed discrete or continuous probability distributions. The decision must be made before the actual values of these random parameters are realized. The need for including uncertainty in complex decision models arose early in the history of mathematical programming. The first forms of model, involving an action followed by observation and reaction, that is, the two-stage stochastic programming with recourse, appeared independently in Dantzig (1955) and Beale (1955) (as also referred to in the preceding paragraph). (Dantzig uses the term “linear programs under uncertainty” while Beale refers to it as “linear programs with random coefficients”. The identical year of publication is a mere coincidence.) An alternative method, called chance- or probabilistic-constrained programming, was also developed quite early, principally by

Charnes and Cooper (1959). Both forms have their origin in statistical decision theory, but in contrast to decision theory, stochastic programming has emphasized methods of solution and analytical solution properties over procedures for constructing objectives and updating probabilities (Dupacova, 2002).

Stochastic programming with recourse is often used to model uncertainty, giving rise to large-scale mathematical programs that require the use of decomposition methods and approximation schemes for their solution. Surveys of developments and applications of stochastic programming can be found in Ermoliev and Wets (1988), Wets (1989), Birge and Wets (1991), Kall and Wallace (1994), and Ruszczynski and Shapiro (2003). There has been tremendous progress in stochastic optimization problems from both theoretical and practical perspectives, especially in stochastic linear programming, matching almost in parallel, its deterministic counterpart (Wets, 1996). This is illustrated by the successful use of stochastic programming approaches in a number of areas such as energy (particularly electricity generation) and production planning, telecommunications, forest and fishery harvest management, engineering, and transportation. Uryasev and Pardalos (2001) mention also that it was recently realized that the practical experience gained in stochastic programming can be expanded to a much larger spectrum of applications including financial modeling, asset-liability management, bond portfolio management, currency modelling, risk control, and probabilistic risk analysis.

Figure 1.3 depicts some of the more well-established optimization techniques under uncertainty with emphasis on chemical engineering applications, based on the recent review article by Sahinidis (2004). Interested readers are referred to Kall and Wallace (1994), Birge and Louveaux (1997), and Prekopa (1995) as standard basic references for the theory and application of multistage stochastic programs, in particular the two-stage program with fixed recourse that is widely adopted in model developments pertaining to this work.

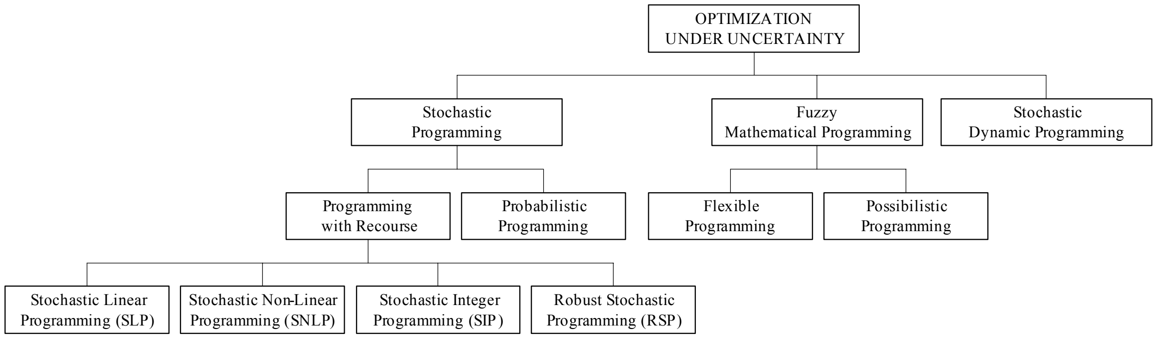


Figure 1.4. Established optimization techniques under uncertainty (with emphasis on chemical engineering applications as based on Sahinidis (2004))

1.4.1 Assessing the Need for Stochastic Programming Models: Advantages and Disadvantages

The starting point for many stochastic programming models is a deterministic linear programming (LP) model (or simply, a linear program). If some of the parameters in an LP are uncertain and the LP appears to be fairly sensitive to changes in the parameters, then it may be appropriate to consider an SP model (Sen and Higle, 1999). For example, consider a blending model for the production planning of a petroleum refinery that uses LP to recommend recipes to produce a crude oil blend with specific characteristics in the mixing tank preceding the crude distillation unit (CDU), by combining different types of crudes. In some instances, the content of these mixtures of crude oil may, or are even bound to, vary. If the optimal blend remains relatively unaffected within the range of variation, then we can justify the certainty assumption of LP. On the other hand, if the variations cause the optimal blend to vary substantially (which should be justifiably anticipated), then it may be worth pursuing the comparatively more complex and more computationally demanding stochastic programming model. In such a case, we can use LP sensitivity analysis for diagnostic purposes and stochastic programming to obtain an optimal blend. In a more general context, the distinction between deterministic models and stochastic programming models lies in the sense that in considering possible scenarios of a certain problem or a phenomenon, multiple scenarios with their associated data are optimized one at a time in deterministic models, as if they will occur with certainty. In contrast, a stochastic model considers the ensemble of all scenarios simultaneously, each with an associated probability of occurrence, as a probabilistic description of the future (Shapiro, 2004).

In many instances, we need stochastic programming models due to scant information pertinent to executing decisions. Such a situation is likely to arise, for example, with the introduction of new products or services. Consider the corporate planning or marketing arm of an integrated oil and gas company that wishes to introduce a new lubricant product from its refining activities. They may try to obtain information on the need for this product in multiple ways. They may inspect usage data of an existing similar product(s) in the market within their region and from a similar demographic region in a different part of the country. They could also obtain surrogate data from a computer simulation model. Finally, they could execute a market survey or perform a test within a small segment of the region. All of these approaches provide estimates of market demand for the new lubricant product, and these data estimates are likely to be different. With a stochastic programming model, the company can include all these alternative forecasts within one decision-making model to produce a more robust plan (Sen and Higle, 1999).

Moreover, stochastic programming has the additional benefit of allowing decision-makers to impose constraints reflecting their judgement of the risks associated with the firm's performance under various possible business and even non-business (for example, a socially-influenced event) scenarios. To illustrate this point, consider the set of constraints that state that losses by a firm in year n cannot exceed US\$$M$ million under any circumstance (or scenario). These constraints may alternatively be expressed as a single probabilistic constraint requiring that the probability associated with losses suffered by the firm surpassing US\$$M$ million in year n may not exceed a certain value p, say 0.05 (for convenience sake). Decision-makers may of course view such targets as being somewhat arbitrary, implying the need to apply methods of multiobjective optimization to systematically explore the tradeoff of maximum expected net revenues against risk targets. Therefore, by employing stochastic programming models, risk management is translated into systematic procedures for identifying efficient frontiers that describe the tradeoffs of expected return profit against explicit descriptions of risk exposure faced by the firm and/or the decision-maker (Shapiro, 2004).

From a computational point-of-view, stochastic programming provides a general framework to model path dependence of the stochastic process within an optimization model. Furthermore, it permits uncountably many states and actions, together with constraints, time-lags, and others. One of the important distinctions that should be highlighted is that unlike dynamic programming, stochastic programming separates the model formulation activity from the solution algorithm. One advantage of this separation is that it is not necessary for stochastic optimization models to all obey the same mathematical assumptions. This leads to a rich class of models for which a variety of algorithms can be developed. However, on the downside of the ledger, stochastic programming formulations can lead to large scale problems, thus methods based on approximation and decomposition becomes paramount for as measures of circumvention (Sen, 2001).

1.4.2 General Formulation of Optimization Model for Operating Chemical Processes under Uncertainty

Pintaric and Kravanja (2000) present a general formulation of the mathematical model for optimization of chemical processes under uncertainty, presented here in the following in a slightly revised form for a production planning problem:

$$\begin{aligned} \text{maximize} \quad & P(x,d,\theta) \\ \text{subject to} \quad & h(x,d,\theta)=0 \\ & g(x,d,\theta)\le 0 \\ & x\in X,\, d\in D,\, \theta\in TH \end{aligned} \tag{1.1}$$

where P is an objective function that usually represents the economic potential, for example, profit; x represents the vector of operating variables (for example, flowrates, compositions, temperatures, pressures); d is the vector of design or planning (size) variables (for example, area, volume, diameter, height, power); θ is the vector of uncertain parameters; and h and g are the vectors of process equality and inequality constraints, respectively. X, D, and TH are continuous feasible regions of operating variables, design variables, and uncertain parameters, respectively, defined by the lower and upper bounds. The most common approach in addressing the optimization problem in equation **(1.1)** is the two-stage stochastic programming formulation and this is addressed and elaborated later.

$$
\begin{array}{lll}
\text{design or planning stage:} & EP = \max\limits_{d} E_{\theta} P(d,\theta) & \\
\text{subject to:} & \text{operating stage:} & P(d,\theta) = \max\limits_{x} P(x,d,\theta) \\
 & \text{such that} & h(x,d,\theta) = 0 \\
 & & g(x,d,\theta) \le 0 \\
 & & x \in X, d \in D, \theta \in TH
\end{array}
\qquad \mathbf{(1.2)}
$$

where E_{θ} is the mathematical operator for the calculation of the expected value of the objective function EP over θ. The objective of the design stage is to select an optimal vector of design or planning variables, while the objective of the operating stage is to determine an optimal vector of operating variables at fixed design or planning decisions.

1.4.3 Overview of the Concept and Philosophy of Two-Stage Stochastic Programming with Recourse from the Perspective of Applications in Chemical Engineering

Under the standard two-stage stochastic programming paradigm, the decision variables of an optimization problem under uncertainty are partitioned into two sets according to whether they are decided (or implemented) before or after an outcome of the random variable(s) is observed. The *first-stage* variables are those that have to be decided before the actual realization of the uncertain parameters. They can be regarded as proactive and are often associated with planning issues such as capacity expansion decisions or in aggregate production planning. Subsequently, once the random events have presented themselves, further design, planning or operational policy improvements can be made by selecting, at a certain cost, the values of the *second-stage* or *recourse* variables.

Traditionally, the second-stage variables are interpreted as corrective measures against any infeasibility arising due to a particular realization of uncertainty, hence the term *recourse*. They allow the decision-maker to model a response to the observed outcome that constitutes his or her desired (and appropriate) recourse. The second-stage problem may also be operational-level decisions following a first-stage plan and realization of the uncertainty; hence they can be regarded as reactive. The corresponding second-stage cost is a random variable due to uncertainty. Thus, the overall objective is to choose the first-stage variables in a way that the sum of the first-stage costs and the expected value of the random second-stage recourse costs is minimized. This concept of recourse has been applied to linear, integer, and nonlinear programming problems (Sahinidis, 2004; Cheng et al., 2005).

For the two-stage stochastic recourse models, the expected value of the cost resulting from optimally adapting the plan according to the realizations of uncertain parameters is referred to as the *recourse function*. A problem is said to have *complete recourse* if the recourse cost for every possible uncertainty realization remains finite, independent of the nature of the first-stage decisions. In turn, if this statement is true only for the set of feasible first-stage decisions (that is, they satisfy the first-stage constraints, or in other words, the first-stage constraints are not violated), then the problem is said to have the slightly less restrictive property of *relatively complete recourse*. To ensure complete recourse in any problem, penalty functions (of costs) for deviations from constraint satisfaction of prescribed limits are used (Sen and Higle, 1999).

For example, in recourse planning, we model a response for each outcome of the random elements that might be observed. In actuality, this response will also depend upon the first-stage decisions. This type of planning in practice involves setting up policies that will help the organization adapt to the revealed outcomes. Specific to production and inventory systems, the first-stage decision might correspond to resource quantities, and demand might be modelled using random variables. When demand exceeds the amount available in stock, policy may dictate that customer demand be backlogged at the expense of some shortage costs. This policy constitutes a recourse response; to be precise, it is a simple recourse policy as according to the explanation in the preceding paragraph. The exact level of this response (that is, the amount backlogged) depends on the amounts produced and demanded. Under uncertainty, it is essential to adopt initial policies that will accommodate alternative outcomes. Consequently, this provides the conviction that modeling under uncertainty requires the incorporation of a recourse policy model (Cheng et al., 2005).

Typically, a two-stage stochastic planning problem is derived from a deterministic problem by identifying the decision variables that can be manipulated after the design, planning, and construction stage, and deferring the decision with respect to those variables until the operating stage. Since the manipulable variables or operating variables, that is, the recourse variables can be controlled in a way such that the best outcomes are obtained for the prevailing operating conditions established in the first-stage, a planning problem reduces to one in which the remaining decision variables are to be determined such that the expectation of an operationally-optimized economic objective is maximized (Pai and Hughes, 1987).

According to Sen and Higle (1999), the presence of uncertainty affects both feasibility and optimality of the optimization problem. In fact, formulating an appropriate objective function itself raises interesting modelling and algorithmic questions. Furthermore, in Section 1.1, we note that the many variants of the two-stage stochastic modelling approach lies primarily in the distinct approaches taken to evaluate the expected value term, which in principle (but not necessarily so) contains a multidimensional integral involving (possibly) the joint probability distribution of the uncertain parameters. However, this varies depending on the nature of the problem and information in the form of historical data that is available to the decision maker (Applequist et al., 2000). Gupta and Maranas (2003) rightly so pointed out that the main challenge in solving the two-stage stochastic program lies in evaluating the second-stage expectation term. Various techniques have been proposed in works addressing production planning under uncertainty by employing the two-stage decision problem model. Ierapetritou and Pistikopoulos (1994a, 1994b, 1996) proposed an algorithmic procedure using numerical Gaussian quadrature integration to approximate the multiple integrals of the expected profit with the corresponding quadrature points simultaneously obtained as a result of the optimization procedure. Liu and Sahinidis (1996) use a Monte-Carlo sampling approach to estimate the expectation of the objective function. Clay and Grossmann (1997) employ a sensitivity-based successive disaggregation algorithm without consideration of capacity expansion. Another approach in an earlier work is the technique employed by Friedman and Reklaitis (1975) that incorporates flexibility in the system by allowing for possible future additive corrections on the current decisions and optimize the system by applying an appropriate cost-for-correction in the objective function, which is essentially similar to the recourse model approach. It is deeply encouraging to note that recent applications of two-stage stochastic programming for solvinuncertainties and control risks explicitly, particularly the stochastic programming technique of two-stage stochastic linear programming with fixed recourse or in general also known as the scenario analysis technique. The underlying idea is to simultaneously consideg large-scale chemical production and process planning problems, as chronicled in Table 1.3, marvellously spans the entire full range of mathematical programming problems from linear programs (which no longer pose any serious computational complication with current availabilities of efficient algorithms and hardware computing prowess) and mixed-integer linear programs (MILP) to nonlinear programs (NLP) and mixed-integer nonlinear programs (MINLP).

1.4.4 The Classical Two-Stage Stochastic Linear Program with Fixed Recourse

Birge and Louveaux (1997) reiterated that stochastic programming is an attractive option for planning problems because it allows the decision maker to analyze r multiple scenarios of an uncertain future, each with an associated probability of occurrence. The model simultaneously determines an optimal contingency plan, due to the incorporation of recourse actions, for each scenario and an optimal plan that optimally hedges against these contingency plans. Optimization entails maximization of expected net profits or minimization of various expected costs, in which the term “expected” refers to multiplying net profits or costs associated with each scenario by its probability of occurrence (Lababidi et al., 2004).

On a more general note, a two-stage stochastic program with recourse is a special case of multistage stochastic program. As observed by Romisch and Schultz (2001) and Ahmed et al. (in press), much of the work in the area of multistage stochastic programs has focused on stochastic linear programs, which do not handle integer requirements or nonlinearities. This is mainly due to the highly desired property of convexity of the value function of stochastic linear programs; however, breaks down in the case of stochastic integer programs (Balasubramanian and Grossmann, 2004).

CHAPTER 2

Research Objectives

2.1 PROBLEM DESCRIPTION AND RESEARCH OBJECTIVES

The midterm production planning problem for petroleum refineries addressed in this work can be stated as follows. It is assumed that the physical resources of the plant are fixed and that the associated prices, costs, and market demand are externally imposed. Thus, the following are all assumed to be known or given (Reklaitis, 1982):

- types and number of production units and inventory facilities;
- the product slate;
- the costs of materials, labour, and energy;
- the product prices; as well as
- the product demands.

With these assumed given information, the principal governing objective of this research is to determine the optimal midterm (medium term) production planning of a petroleum refinery by computing the amount of materials being processed at each time in each unit and in each process stream, by explicitly and simultaneously accounting for the three major factors of uncertainty, namely: (i) market demand or product demand; (ii) prices of crude oil (the raw material) and the final saleable refining commercial products; and (iii) product yields of crude oil from chemical reactions in the primary distillation unit, or more commonly known nowadays as the crude distillation unit (CDU), of a typical petroleum refinery. A hybrid of various stochastic optimization modelling techniques within the framework of the classical two-stage stochastic programming with recourse model structure are applied to reformulate a deterministic refinery production planning problem into one that is both solution robust and model robust. This is accomplished by adopting the Markowitz's mean–variance approach in handling risk arising from variations in profit and penalty costs that are due to the recourse actions incurred as a result of violations of constraints subjected to the factors of uncertainty aforementioned.

A numerical study based on a representative example drawn from the literature is then presented and solved to optimality to demonstrate the effectiveness of implementing the stochastic models proposed. As indicated, we will achieve this through reformulating the deterministic linear programming (LP) model developed by Allen (1971) for refinery operations planning by introducing elements of uncertainty. Subsequently, the solutions are provided by implementing four approaches or techniques for decision-making under uncertainty, as elaborated in the following with a gradual increase of complexity:

- Approach 1: the Markowitz's mean–variance (*E*–*V* or MV) model to handle randomness in the objective function coefficients of prices by minimizing the variance of the expected value (or mean) of the random coefficients, subject to a target profit constraint;
- Approach 2: the two-stage stochastic programming with fixed recourse approach to model randomness in the right-hand side and left-hand side (or technological) coefficients by minimizing the expected recourse penalty costs due to violations of constraints;
- Approach 3: incorporation of the Markowitz's MV model within the two-stage stochastic programming framework developed in Approach 2 to minimize both the expectation and the variance of the recourse penalty costs; and
- Approach 4: reformulation of the model developed in Approach 3 by utilizing the Mean-Absolute Deviation (MAD) in place of variance as the measure of risk imposed by the recourse penalty costs.

Finally, the results obtained will be analyzed, interpreted, and compared and contrasted; the latter primarily with the work reported by Ravi and Reddy (1998), who made use of the fuzzy linear fractional goal programming approach in their proposed solution on the same deterministic model of Allen (1971).

Acknowledging the shortcomings of deterministic production planning models as addressed in the earlier section on review of the existing literature in the field, the novelty of the approaches in this work lies in striving to present an explicit method of stochastic programming under uncertainty by utilizing existing concepts of straightforward yet effective modelling techniques in formulating tractable stochastic models for application in the large-scale optimization problem of petroleum refinery production planning. In essence, the production plans are to be determined probabilistically within an environment of incomplete information. As events materialize with the uncertainties revealed, production rates of the refinery for the following planning period are to be determined in conditional fashion, taking account of accumulating experience and future possibilities, as denoted by the associated probabilities. The meaning and quantitative consequences to be assigned to each such event is to be determined beforehand from the scenario generation methodology that is applied, so that plans can be formulated and evaluated in advance for each possible contingency. Finally, these efforts are undertaken with the desirable intention of assessing potential future changes in the structure and requirements of the decision-making activity of production planning within a petroleum refinery (at least from a qualitative point-of-view, if not quantitatively through extending the optimization models in future work).

As a final remark, Shapiro (1993) praised stochastic programming with recourse models for offering rigorous formalism in evaluating the impact of uncertainty on production planning and scheduling plans. However, it is rightly cautioned that a great deal of artistry and problem specific analysis is required to effectively utilize the formalism since it can easily lead to producing complex models demanding heavy computational time. In line with this, justifications in using a specific modelling technique will be provided and comparisons made with other forms and techniques.

2.2 OVERVIEW OF THE THESIS

A major part of the remainder of the thesis shall focus on the implementation of the four approaches introduced in Section 2.1 for oil refinery planning under uncertainty and the resulting proposed general models. This is followed by the application of the stochastic models developed to a representative numerical example in order to test and demonstrate their suitability, effectiveness, and robustness for decision-making activities.

In Chapter 3, we present an exposition on the methodology of problem formulation for the planning of a large-scale process network under uncertainty, in view of application for the production planning of petroleum refineries. Rigorous and detailed mathematical treatment of the associated theories underpinning the concept of two-stage stochastic programming with recourse model is thoroughly discussed.

In Chapter 4, we demonstrate the general formulation of the deterministic production planning model for petroleum refineries. The framework and structure of the deterministic planning model is based mainly on the production planning model proposed by McDonald and Karimi (1997) and on the refinery planning model in the work by Ierapetritou and Pistikopoulos (1996).

Subsequently, in Chapter 5, we reformulate the deterministic model developed in the preceding chapter with the addition of stochastic dimensions to address uncertainties in commodity prices, market demand, and product yields. As elaborated in the previous section, four approaches are adopted, resulting in four stochastic models with gradual complexity and capability, especially pertaining to risk management.

We dedicate Chapter 6 for a discussion on the implementation of the models on the General Algebraic Modeling System (GAMS) platform. Here, we strive to succinctly justify and advocate the use of a high-level modelling language like GAMS (and its counterparts such as AMPL) for the ease of constructing and implementing a large-scale optimization model.

Chapter 7 then discusses the two metrics that we deem most suitable in evaluating the performance of the stochastic models and hence, the value of the venture of adopting stochastic programming itself. The first metric pertains to the concepts of solution robustness and model robustness as introduced in the seminal work by Mulvey et al. (1995) while the second metric employs the use of coefficient of variation, traditionally defined as the inverse ratio of data to the noise in the data. From this definition, it follows that a small value of C_v is desirable as we strive to minimize the presence of noise. Thus, C_v can be adopted as an indicator of the degree of uncertainty in a stochastic model.

Chapter 8 forms the heart of the thesis as it is essentially the test bed of the performance of the proposed stochastic models. Without loss of generality, we consider the deterministic midterm production planning model for a petroleum refinery by Allen (1971) and Ravi and Reddy (1988) as the base case or core model of our numerical study. First, the deterministic model is solved to optimality using GAMS with the results validated by LINDO. Then, based on the detailed steps outlined in Chapter 5, the base case model is reformulated according to the four approaches outlined to produce four convex nonlinear quadratic stochastic models. We then proceed to compute the optimal solution of the models and thereafter, a thorough analysis of the results obtained are established in an effort to investigate the robustness of the solutions and the models, by employing the aforementioned performance metrics.

Finally, we provide some concluding remarks with regards to the outlined research objectives that we have managed to achieve, followed by recommendations of promising future work to be undertaken, in Chapters 9 and 10, respectively. Miscellaneous supporting information is collected under Appendix and the thesis is brought to a close by a complete list of references detailing the multitude literature that has been cited in this work, as a respectful acknowledgement of the vast contributions of previous and present researchers to whom we owe our utmost gratitude.

CHAPTER 3

Methodology for Formulation of Mathematical Programming Models and Methods for Problems under Uncertainty

3.1 MOTIVATION FOR IMPLEMENTATION OF STOCHASTIC OPTIMIZATION MODELS AND METHODS

Westerberg (1996) advocates the view that optimization is a tool to aid decision-making through the selection of better values in solving a problem. In real-world applications, optimization problems in practice depend mostly on several model parameters, noise factors, uncontrollable parameters, and others, whose quantities are not given or even known, least of all fixed, at the planning stage. Typical examples from engineering, economics, and operations research include material parameters (for example, elasticity moduli, yield stresses, allowable stresses, moment capacities, specific gravity), external loadings, friction coefficients, moments of inertia, length of links, mass of links, location of the centre of gravity of links, manufacturing errors, tolerances, noise terms, product demand parameters, technological coefficients in input–output functions, and cost factors (and many others). Owing to the existence of several types of stochastic uncertainties, namely physical uncertainty, economic uncertainty, statistical uncertainty, and model uncertainty, these parameters must be modelled by random variables having a certain probability distribution. In most cases, at least certain moments of this distribution are known.

In order to cope with these uncertainties, a basic procedure in the engineering/economic practice is to replace first, the unknown parameters by some chosen nominal values, for example, statistically-computed expected or mean values, estimates, or merely guesses, of the parameters. Then, the resulting and mostly increasing deviation of the performance (output, behaviour) of the system or structure from the prescribed performance, that is, the tracking error, is compensated by (online) input corrections. However, the online correction of a system/structure is often time-consuming and causes mostly increasing expenses, typically in terms of correction or recourse costs.

Very large recourse costs may arise in case of damages or failures of a manufacturing or processing plant. This can be omitted to a large extent by taking into account upfront at the planning stage, the possible consequences of the tracking errors and the known prior and sample information about the random data of the problem. Hence, instead of relying on ordinary deterministic parameter optimization methods based on some nominal parameter values and then just applying some correction actions (conventionally via sensitivity analysis or deterministic scenario analysis), stochastic optimization methods should be applied. By incorporating the stochastic parameter variations into the optimization process, expensive and increasing online correction expenses can be omitted or at least reduced to a large extent (Marti, 2005).

3.2 AN OVERALL OUTLOOK ON FORMULATION OF STOCHASTIC OPTIMIZATION MODELS AND METHODS FOR THE REFINERY PRODUCTION PLANNING PROBLEM UNDER UNCERTAINTY

In principle, probabilistic or stochastic modelling is an iterative procedure that principally comprises the following three steps, as outlined by Diwekar (2002, 2003):

1. specify the identified uncertainties or randomness in key input parameters in terms of probability distributions;
2. perform sampling for the distribution of the specified random parameter in an iterative fashion;
3. propagate the effects of uncertainties through the model and apply suitable statistical techniques to analyze the results.

Planning under uncertainty requires the explicit representation of uncertain quantities within an underlying decision model. When the underlying model is a linear program, the representation of certain data elements as random variables results in a stochastic linear program (SLP) (Higle 1998). This provides the underlying framework for the deterministic refinery planning model problem considered in this work, subject to uncertainties in commodity prices, market demand, and product yields from crude oil. The fundamental idea behind SLP is the concept of recourse, that is, the ability to take corrective action(s) after a random event has taken place. Recourse models use corrective actions, usually in the form of penalty functions, to compensate for the violation of constraints arising during the realization of uncertainty. In other words, it is the opportunity to adapt a solution to the specific outcome observed. The two-stage model is one of the main paradigms of recourse models. The two-stage model divides the decision variables into two stages. The first-stage variables are those that have to be decided right now (or "here-and-now") before future actual realization of the uncertain parameters so as to accommodate any future uncertain parameter realizations or perhaps those that fall within some specified confidence limits. Hence, the two-stage formulation is also referred to as the "here-and-now" model. Then, the second-stage variables are those used as corrective measures, that is, recourse against any infeasibilities arising during the realization of the uncertainty. In the context of the refinery production planning problem under uncertainty, the raw material supply of crude oil(s) and the production rates are determined in the first stage or planning stage. Then, the effects of the uncertain parameters on system performance is established in the second stage or operating stage, in which decisions are made concerning the amount of production required to meet the actually realized product demand and product yields, or the amount of raw material required from other suppliers to meet production requirements, or the opposite situation of determining the inventory cost and space required to contain production surpluses.

It is apparent that the second stage is the most important part of the model since this is the stage at which the flexibility of the planning is checked, possibly by including consideration of variations of the operating variables to accommodate the uncertain parameter realizations. This part of the model is also the most computationally demanding, as remarks by Wellons and Reklaitis (1989).

3.3 FORMULATION OF STOCHASTIC OPTIMIZATION MODELS

Ponnambalam (2005) categorizes stochastic optimization problems into three classes of parameter randomness (in a loosely restricted order of increasing complexity) as follows: (1) randomness in the objective function coefficients; (2) randomness in the right-hand-side coefficients of constraints; and (3) randomness in the left-hand-side coefficients (also referred to as technological coefficients). He further discusses at least three conventional and widely-adopted methods of problem formulation and their associated challenges or difficulties when coefficients in a linear programming (LP) problem are random, as in the following.

1. In the Markowitz's mean–variance formulation approach, the LP becomes a quadratic programming (QP) problem that is somewhat harder to solve than an LP.
2. In the two-stage stochastic linear programs with recourse method, two major challenges are in (i) producing reasonable scenarios and their probabilities, and (ii) in the exponentially increasing size of the problem with the number of uncertain parameters. However, the results obtained are quite practical and mean–variance formulation can even be considered as a possible inclusion.
3. In the case of chance-constrained programming, the problem is quite easy to solve in the case of right-hand side (RHS) coefficients randomness but becomes nonlinear and increasingly difficult to solve in the case of left-hand side (LHS) randomness. Moreover, if the constraints have to be satisfied with joint probability, then the formulation becomes tedious and hard to solve. Furthermore, the decision maker has to arbitrarily choose a probability value to satisfy.

In addition to these three conventional approaches of stochastic optimization, the alternative method of robust optimization is also available in which an assumption on the perturbations of the uncertain coefficients are made.

The general approach for the formulation of stochastic optimization models can be found in any of the standard texts in the stochastic programming literature. The discussion here is based on Kall and Wallace (1994). First, the general form of the mathematical programming model is introduced as the following:

$$\begin{aligned} &\text{minimize} \quad g_0(x) \\ &\text{subject to} \quad g_i(x) \le 0,\ i = 1, 2, \cdots, m \\ &\qquad\qquad\quad x \in X \subset \mathbb{R}^n \end{aligned} \tag{3.1}$$

Here, it is understood that the set X as well as the functions g_i: $\mathbb{R}^n \to \mathbb{R}$, $i = 0, \ldots, m$, with $\mathbb{R}^n$ denoting the set of real n-vectors, are given by the modelling process.

If there are random parameters in **(3.1)**, they may lead to the problem

$$\begin{aligned} &\text{"minimize"} \quad g_0(x, \tilde{\xi}) \\ &\text{subject to} \quad g_i(x, \tilde{\xi}) \le 0,\ i = 1, 2, \cdots, m \\ &\qquad\qquad\quad x \in X \subset \mathbb{R}^n \end{aligned} \tag{3.2}$$

where $\tilde{\xi}$ is a random vector varying over a set $\Xi \subset \mathbb{R}^k$. More precisely, it is assumed throughout that a family F of events, which are subsets of Ξ, and the probability distribution P on F, are given. Hence, for every subset $A \subset \Xi$ that is an event, as denoted by $A \in F$, the probability $P(A)$ is known. Furthermore, it is assumed that the functions $g_i(x, \cdot)$: $\Xi \to \mathbb{R}$ $\forall x$, i are random variables themselves, and that the probability distribution P is independent of x.

However, problem **(3.2)** is not well defined since the meanings of "minimize" as well as of the constraints are not clear at all, if we think of taking a decision on x before knowing the realization of $\tilde{\xi}$. Therefore, a revision of the modelling process is necessary, leading to the so-called *deterministic equivalent* for **(3.2)** (Kall & Wallace, 1994).

3.3.1 The Deterministic Equivalent Model

If the following variable is defined:

$$g_i^+(x, \xi) = \begin{cases} 0 & \text{if } g_i(x, \xi) \le 0, \\ g_i(x, \xi) & \text{otherwise} \end{cases} \tag{3.3}$$

then, the ith constraint is violated if and only if $g_i^+(x,\xi) > 0$ for a given decision x and realization ξ of $\tilde{\xi}$. Hence, we could provide for each constraint, a *recourse* or *second-stage activity* $y_i(\xi)$ that, after observing the realization ξ, is chosen so as to compensate for the violation of a particular constraint, if there is one, by satisfying the condition $g_i(x,\xi) - y_i(\xi) \le 0$. This extra effort is assumed to cause an extra cost or penalty of q_i per unit, in which the additional costs termed as *recourse function* amount to:

$$Q(x,\xi) = \min_y \left\{ \sum_{i=1}^{m} q_i y_i(\xi) \middle| y_i(\xi) \ge g_i^+(x,\xi), \quad i = 1,\cdots,m \right\} \tag{3.4}$$

Thus, this yields a total cost comprising *first-stage* and *recourse* cost of

$$f_0(x,\xi) = g_0(x,\xi) + Q(x,\xi) \tag{3.5}$$

Instead of **(3.4)**, we might consider a more general linear recourse program with a recourse vector $y(\xi) \in Y \in \mathbb{R}^{\bar{n}}$ (Y is some given polyhedral set, that is, a convex set with linear inequalities, such as $\{y | y \ge 0\}$), an arbitrary fixed $m \times \bar{n}$ matrix W (the *recourse matrix*), and a corresponding unit cost vector $q \in \mathbb{R}^{\bar{n}}$. This results in the following recourse function for **(3.5)**:

$$Q(x,\xi) = \min_y \left\{ q^T y \middle| Wy \ge g^+(x,\xi), y \in Y \right\} \tag{3.6}$$

where $g^+(x,\xi) = \left(g_1^+(x,\xi),\cdots,g_m^+(x,\xi)\right)^T$.

If we consider a situation of a factory producing m products, $g_i(x,\xi)$ could be understood as the difference between the demand and the output of a product i. Then, $g_i^+(x,\xi) > 0$ means that there is a shortage in product i, relative to the demand. If the factory is assumed to be committed to cover the demands, problem **(3.4)** could, for instance, be interpreted as buying the shortage of the products at the market. Problem **(3.6)** instead could result from a *second-stage* or *emergency* production program, carried through with the factor input y and a technology represented by the matrix W. Choosing $W = I$, the $m \times m$ identity matrix, **(3.4)** turns out to be a special case of **(3.6)**.

We could also consider a nonlinear recourse program to define the recourse function for **(3.5)**; for instance, $Q(x,\xi)$ could be chosen as:

$$Q(x,\xi) = \min\left\{q(y) \middle| H_i(y) \geq g_i^+(x,\xi), \quad i = 1,\cdots,m; y \in Y \subset \mathbb{R}^{\bar{n}}\right\} \quad \textbf{(3.7)}$$

where $q : \mathbb{R}^{\bar{n}} \to \mathbb{R}$ and $H_i : \mathbb{R}^{\bar{n}} \to \mathbb{R}$ are supposed to be given.

In any case, if it is meaningful and acceptable to the decision maker to minimize the expected value of the total costs (that is, consisting of the first-stage and the recourse costs) then instead of problem **(3.2)**, we could consider its deterministic equivalent, *the two-stage stochastic program with recourse*:

$$\min_{x \in X} E_{\tilde{\xi}} f_0\left(x,\tilde{\xi}\right) = \min_{x \in X} E_{\tilde{\xi}}\left\{g_0\left(x,\tilde{\xi}\right) + Q\left(x,\tilde{\xi}\right)\right\} \quad \textbf{(3.8)}$$

The above two-stage problem is immediately extendable to the *multistage recourse program* as follows: instead of the two decisions x and y, to be taken at stages 1 and 2, we are now faced with $K + 1$ sequential decisions $x_0, x_1, \cdots, x_K \left(x_\tau \in \mathbb{R}^{\bar{n}_\tau}\right)$, to be taken at the subsequent stages $\tau = 0, 1, \cdots, K$. The term "stages" can, but need not necessarily or strictly, be interpreted as time periods (depending on the context and nature of the problem addressed). However, an extensive discussion on the multistage recourse problems will not be presented here as this formulation is not considered in the present work.

3.4 RECOURSE PROBLEMS AND MODELS

The term “recourse” refers to the opportunity to adapt a solution to the specific outcome observed. Recourse models result when some of the decisions must be fixed before information relevant to the uncertainties is available, while some of them can be delayed until afterward to allow for the mentioned opportunity for recourse action to be taken. For instance, in this problem of refinery planning under uncertainty, in addressing market demand uncertainty, the resource quantities, that is, the crude oil supply, must be determined fairly early, but the production quantities can be delayed until after the demand is known. In this sense, the production quantities may be thought of as being “flexible” or “adaptive” while the resource quantities of crude oil are not (Higle, 2005).

As elaborated earlier somewhat more mathematically, the general recourse model is to choose some initial decision that minimizes current costs plus the expected value of future recourse actions. With a finite number of second-stage realizations and all linear functions, the full deterministic equivalent linear program or extensive form can always be formed (Birge and Louveaux, 1997).

3.5 COMPONENTS OF A RECOURSE PROBLEM

Recourse problems are generally characterized by the following three properties: (1) scenario tree, (2) scenario problems, and (3) nonanticipativity constraints. A *scenario* is one specific, complete, realization of the stochastic elements that might appear during the course of the problem, for instance, a possible sequence of demand over the duration of the problem. The *scenario tree* is a structured distributional representation of the stochastic elements and the manner in which they may evolve over the period of time represented in the problem. A *scenario problem* is associated with a particular scenario and may be looked upon as a deterministic optimization problem.

Nonanticipativity constraints are specific conditions that may be necessary to include depending on the manner in which a problem is formulated so as to ensure that the decision sequence honours the information structure associated with the scenario tree.

They are the only constraints linking the separate scenarios (or linking decisions across scenarios in each period). Without them, a problem would decompose into a separate problem for each scenario, maintaining the structure of that problem (Birge & Louveaux, 1997). They impose the condition that scenarios that share the same history of information until a particular decision epoch, should also make the same decisions. In reality, the nonanticipativity constraints ensure that the solutions obtained are implementable, that is, actions that must be taken at a specific point in time depend only on information that is available at that time. The future is uncertain; therefore, today's decision cannot take advantage of knowledge of the future, thus, they are independent of each other. Because of this, the terms "nonanticipativity" and "implementability" are sometimes used interchangeably. These nonanticipativity constraints, which are derived from the scenario tree, are a distinguishing characteristic of stochastic programs; solution methods are typically designed to exploit their structure (Higle, 2005; http://www-fp.mcs.anl.gov/otc/Guide/OptWeb/continuous/constrained/stochastic/index.html, accessed November 2, 2005).

3.6 FORMULATION OF THE TWO-STAGE STOCHASTIC LINEAR PROGRAM (SLP) WITH RECOURSE PROBLEMS

The classical two-stage stochastic linear program (SLP) with fixed recourse as originally proposed in the seminal works of Dantzig (1955) and Beale (1955) has the following general form:

$$\begin{aligned} &\text{minimize } c^T x + E_\xi\left[Q\left(x, \xi(\omega)\right)\right] \\ &\text{subject to } Ax = b \\ &\qquad\qquad x \in X \geq 0 \end{aligned} \tag{3.9}$$

$$\begin{aligned} \text{where} \quad & Q\left(x, \xi(\omega)\right) = \text{minimize } q^T(\omega) y(\omega) \\ & \text{subject to } W(\omega) y(\omega) = h(\omega) - T(\omega) x \\ & \qquad\qquad y(\omega) \geq 0 \end{aligned} \tag{3.10}$$

where $x \in \mathbb{R}^n$ is the vector of first-stage decision variables of size $(n \times 1)$;

c, A, and b are the first-stage data of sizes $(n \times 1)$; $(m \times n)$; and $(m \times 1)$ respectively;

$\omega \in \Omega$ represent random events;

$q(\omega)$, $h(\omega)$, and $T(\omega)$ are the second-stage data of sizes $(k \times 1)$; $(l \times 1)$; and $(l \times k)$ respectively;

$W(\omega)$ is the random recourse coefficient matrix of size $(l \times k)$;

$y(\omega)$ is the vector of second-stage decision variables.

In this classical SLP model, the problem in **(3.9)** represents the first-stage model while **(3.10)** corresponds to the second-stage model. x is also referred to as the "here-and-now" decision. Note in particular that x does not respond to ω as it is effectively determined before any information regarding the random or uncertain data is obtained, that is, before the actual realization of the uncertain parameters. In other words, x is characterized as being scenario-independent and its optimal value is not conditional on the realization of the uncertain parameters. Thus, they are the design or planning variables (depending on the context of the problem addressed). Variables in this set cannot be adjusted once a specific realization of the uncertain data is observed. c is the column vector of cost coefficient at the first stage. A is the first-stage coefficient matrix with b as the corresponding right-hand side vectors.

In contrast, the second-stage variable y is determined only after observations regarding ω have been obtained. For a given realization of ω, the second-stage problem data of q, h, and T become known. Each component of q, h, and T is thus a possible random variable. q is the vector of recourse cost coefficient vectors at the second stage. As stated, W is the random second-stage recourse coefficient matrix with h as the corresponding right-hand side vectors. T is the matrix that ties the two stages together while ω denotes scenarios in the future. If $T_{i\cdot}(\omega)$ is the ith row of the matrix $T(\omega)$, then combining the stochastic components of the second-stage data results in a vector of a particular realization of the random data $\xi(\omega) = \left(q^T(\omega), h^T(\omega), T_{1\cdot}(\omega), \cdots, T_{l\cdot}(\omega)\right)$, which is also random with potentially up to $N = k + l + (l \times n)$ components. In other words, a single random event or state of the world ω influences several random variables that are all components of a *single* random vector ξ. This is one of the most profound feature of a two-stage stochastic program (Birge and Louveaux, 1997; Higle, 2005; Uryasev, 2005, www.ise.ufl.edu/esi6912/FALL2005/DOCS/notes8.ppt, accessed November 11, 2005; Lai et al., 2005; Mulvey et al., 1995).

In the literature, the problem in **(3.10)** is variously termed as the second-stage problem, the subproblem, or the recourse subproblem. It allows some operational recourse or corrective actions to take place to improve the objective and correct any infeasibility. Essentially, the goal of a two-stage model is to identify a first-stage solution that is well-hedged or simply well-positioned against all possible realizations of ω.

From an application point-of-view especially in the field of operations research and management science, stochastic programs seek to minimize the cost of the first-stage decision plus the expected cost of the second-stage recourse decision. (A contrasting stand is usually taken in the financial engineering field in which the objective is typically to maximize profit, hence leading to a stochastic maximization program.) The first linear program minimizes the first-stage direct costs $c^T x$ plus the expected recourse cost $E_\xi\left[Q\left(x,\xi(\omega)\right)\right]$ over all possible scenarios while meeting the first-stage constraints, $Ax \geq b$. The recourse cost Q depends on both x and ω. The second linear program describes how to choose y in which a different decision is taken for each random scenario ω. It minimizes the cost $q^T y$ subject to some recourse constraint $Tx + Wy = h$. As mentioned earlier, this random constraint can be thought of as requiring some action to correct the system after the random event occurs. It is the goal constraint in which violations of this constraint are allowed, but the associated penalty cost as given by $q^T y$ will influence the choice of x. Thus, $q^T y$ is the recourse penalty cost or the second-stage value function, or just simply referred to as the recourse function, while $E_\xi\left[Q\left(x,\xi(\omega)\right)\right]$ denotes the expected value of the recourse function.

In the case of refinery production planning under market demand uncertainty, x_1, for example, might correspond to the resource levels of crude oil to be acquired and h_1 corresponds to a specific data scenario, notably the actual demands for the various refining products. The decision y_1 adapts to the specific combination of x_1 and h_1 obtained. In the event that the initial decision x_1 is coupled with a "bad" outcome, the variable y_1 offers an opportunity to recover to the fullest extent possible. For example, the random constraint would require the purchase of enough mass or volume of products as the second-stage recourse measure to supplement the original amount that has been produced in order to meet the demand, or it would require the expense cost of storage of production surplus. Thus, recourse problems are always presented in the form of two or more decision stages (Higle, 2005).

Additionally, Sen and Higle (1999) pointed out that this formulation emphasizes the time-staged nature of the decision-making problem, that is, the selection of x is followed by the selection of y, which is undertaken in response to the scenario that unfolds. Thus, the first-stage decision x represents the immediate commitment made, while the second-stage decision y is delayed until additional information is obtained. (For this reason, when solving a recourse problem, one typically reports only the first-stage decision vector.)

Much of the difficulty associated with recourse models may be traced to difficulties with evaluating and approximating the recourse function. In essence, the difficulty in solving the recourse problem may be attributed to the evaluation of the expectation of the random linear program value function $Q\left(x, \xi(\omega)\right)$ that involves multidimensional integration. Notwithstanding the impracticality of the multidimensional integration of this particular function, the recourse function possesses one of the most sought-after properties in all of mathematical programming, namely convexity (Sen and Higle, 1999).

Higle (2005) made an interesting analytical remark that an optimal solution will tend to have the property that x leaves the second-stage decision in a position to exploit advantageous outcomes of ω without excessive vulnerability to disadvantageous outcomes. She further noted that in such a case, careful attention to the specific outcomes used to model the uncertainty is necessary. An overly coarse model may result in a failure to adequately represent outcomes that should influence the first-stage decision, leaving the second-stage in an un-modelled state of vulnerability. On the other hand, an excessively fine and detailed model may typically result in increased computational burden.

As presented without assuming any additional properties or structure on **(3.10)**, this formulation is referred to as the two-stage SLP with "general recourse" problem. Further, there is specific structure in the recourse subproblem that can be exploited for computational advantage. The following section describes the specific subproblem structure type of fixed recourse, which extensively forms the underpinning structure of most models proposed in this work.

3.7 FORMULATION OF THE TWO-STAGE STOCHASTIC LINEAR PROGRAM (SLP) WITH FIXED RECOURSE PROBLEMS

A special case of the recourse model, known as the fixed recourse model, arises when the constraint coefficients matrix $W(\omega)$ in the second-stage problem is not subject to uncertainty, that is, it is fixed and hence is denoted simply as the matrix W. This is by far, the most widely used form of the recourse models since most of the theory and computational methods have been developed for this class of linear two-stage problems (Ermoliev and Wets, 1988). For this, the recourse subproblem becomes:

$$\begin{aligned} Q\left(x,\xi(\omega)\right) = \text{minimize } & q^T(\omega)y(\omega) \\ \text{subject to } & Wy(\omega) = h(\omega) - T(\omega)x \\ & y(\omega) \geq 0 \end{aligned} \quad \textbf{(3.11)}$$

Reduction of the classical two-stage SLP with fixed recourse to the deterministic equivalent program (DEP) yields

$$\begin{aligned} \text{minimize } & c^T x + \mathcal{Q}(x) \\ \text{subject to } & Ax = b, \\ & x \geq 0 \end{aligned} \quad \textbf{(3.12)}$$

where

$$\mathcal{Q}(x) = E_\xi\left[Q\left(x,\xi(\omega)\right)\right] \quad \textbf{(3.13)}$$

with the recourse function given by

$$Q\left(x,\xi(\omega)\right) = \min_y \left\{q^T(\omega)y \,\middle|\, Wy = h(\omega) - T(\omega)x,\ y \geq 0\right\} \quad \textbf{(3.14)}$$

and

$$\xi(\omega) = \left(q^T(\omega), h^T(\omega), T_{1\cdot}(\omega), \cdots, T_{l\cdot}(\omega)\right) \tag{3.15}$$

Since the expected value of the minimum recourse cost $Q\left(x, \xi(\omega)\right)$ modifies the objective of the first-stage problem, the whole model has a certain internal dynamical structure: in computing an optimal first-stage decision x, we have to take into account not only the direct first stage profit c^Tx but also the expected value of the future recourse cost. If there is no feasible solution to **(3.11)**, we assume that $Q\left(x, \xi(\omega)\right) \to +\infty$, and this should also be considered in the first-stage decision (Ermoliev and Wets, 1988).

Further to that, Ermoliev and Wets (1988) also highlighted the widespread interest in stochastic programming problems with recourse due to their vast application to modelling decision problems that involve random data. If some constraints, for example, $Tx = h$ in a linear programming problem include random coefficients in T or h and a decision has to be taken before knowing the realizations $T(\omega)$ and $h(\omega)$ of T and h, it is generally impossible to require that the equality

$$T(\omega)x = h(\omega) \tag{3.16}$$

be satisfied for each realization of the stochastic constraint parameters. The problem with recourse is a way of overcoming these modelling difficulties; the recourse decision y may be interpreted as a correction in **(3.16)**, and the recourse cost $Q\left(x, \xi(\omega)\right)$ as a penalty for discrepancy in **(3.16)**.

3.8 FORMULATION OF THE DETERMINISTIC EQUIVALENT PROGRAM FOR THE TWO-STAGE STOCHASTIC LINEAR PROGRAM (SLP) WITH FIXED RECOURSE AND DISCRETE RANDOM VECTORS

In the following, the framework of the two-stage SLP with recourse model for random vectors with discrete distributions is considered and examined in more detail for the case of a discretely distributed random vector ξ (or equivalently stated as discretely distributed random elements), attaining values of:

$$\xi_1 = (q_1, h_1, T_1) \text{ with probability } p_1 > 0,$$
$$\xi_2 = (q_2, h_2, T_2) \text{ with probability } p_2 > 0,$$
$$\cdots$$
$$\xi_L = (q_L, h_L, T_L) \text{ with probability } p_L > 0,$$

where
$$\sum_{l=1}^{L} p_l = 1 \qquad \textbf{(3.17)}$$

In this case, the two-stage problem of **(3.9)** and **(3.10)** takes on the form

$$\begin{aligned} \text{minimize } & c^T x + \sum_{l=1}^{L} p_l Q(x, \xi_l) \\ \text{subject to } & Ax = b \\ & x \geq 0 \end{aligned} \qquad \textbf{(3.18)}$$

where $Q(x, \xi^l)$ is the minimum objective value in the recourse problem

$$\begin{aligned} \text{minimize } & (q_l)^T y \\ \text{subject to } & Wy = h_l - T_l x \\ & y \geq 0,\ l = 1, 2, 3, \cdots, L \end{aligned} \qquad \textbf{(3.19)}$$

If we denote the solutions to problem **(3.20)** at a given x as $\hat{y}_l(x)$ for $l = 1, 2, \cdots, L$, we can express the first-stage objective as

$$c^T x + \sum_{l=1}^{L} p_l (q_l)^T \hat{y}^l (x) \qquad \textbf{(3.20)}$$

Although the solutions $\hat{y}_l(x)$ depend on x in a rather involved way, instead of considering **(3.19)** and **(3.20)** as a bilevel problem, we can put together the first-stage problem **(3.19)** and all realizations of the second-stage problem **(3.20)** into a large linear programming model:

$$\begin{array}{lccccccccccc}
\text{minimize} & c^T x & + & p_1(q_1)^T y_1 & + & p_2(q_2)^T y_2 & + & \cdots & + & p_L(q_L)^T y_L & & \\
\text{subject to} & Ax & + & & & & & & & & = & b \\
& T_1 x & + & W y_1 & & & & & & & = & h_1 \\
& T_2 x & + & & & W y_2 & & & & & = & h_2 \\
& \vdots & \vdots & & & & & \ddots & & & \vdots & \vdots \\
& T_L x & + & & & & & & & W y_L & = & h_L \\
& x \geq 0 & & y_1 \geq 0 & & y_2 \geq 0 & & & & y_L \geq 0 & &
\end{array} \qquad \textbf{(3.21)}$$

Problems **(3.19)–(3.21)** are equivalent in the sense that they have the same set of solutions for the first-stage decision vector x in **(3.18)** and in which the optimal values of $y_1, y_2, \cdots, y_L$ in **(3.21)** are solutions to the realizations of the second-stage recourse problem **(3.20)** at the optimal x. Hence, **(3.21)** is referred to as the deterministic equivalent program of the two-stage SLP of **(3.19)–(3.20)**.

It is noted that the nonanticipativity constraint is met. There is only one first-stage decision x whereas there are L second-stage decisions, one for each scenario. The first-stage decision cannot anticipate one scenario over another and must be feasible for each scenario, that is, the conditions imposed by $Ax = b$ and $Wy = h_l - T_l x$ for $l = 1, 2, 3, \cdots, L$. Since all the decisions x and y_i are solved simultaneously, x is thus chosen to be optimal, in some sense, over all the anticipated scenarios.

Another feature of the deterministic equivalent problem worth noting is that since the T and W matrices are repeated for every scenario in the model, the size of the problem increases linearly with the number of scenarios. Since the structure of the matrices remains the same and because the constraint matrix has a special shape, solution algorithms can take advantage of these properties. Taking uncertainty into account leads to more robust solutions but also requires more computational effort to obtain the solution (http://www-fp.mcs.anl.gov/otc/Guide/OptWeb/continuous/constrained/stochastic/index.html, accessed November 2, 2005). For further elaboration on the properties of the deterministic equivalent program of stochastic programs with fixed recourse, the interested reader is referred to the excellent extensive survey in the now classical paper by Wets (1974).

3.9 NONANTICIPATIVE POLICIES

As emphasized earlier, one of the more important notions incorporated within a stochastic programming formulation is that of *implementability* or nonanticipativity. Nonanticipativity of the decision process is an inherent component of stochastic optimization problems; this concept essentially and fundamentally distinguishes stochastic from deterministic optimization problems (Wets, http://www.math.ucdavis.edu/~rjbw/ARTICLES/ref2Circ.pdf, accessed on April 9, 2006). It reflects the requirement that under uncertainty, the design or planning decisions x must be implemented before an outcome of the random variable is observed. That is, the planning decision is made while the random variable is still unknown, and therefore, it cannot be based on any particular outcome of the random variable. In the two-stage SLP, this implies that the first-stage or first-period decision x is independent of which second-stage or second-period scenario actually occurs. Thus, the wait-and-see approach, which is anticipative, is not an appropriate decision-making framework for planning. On the other hand, the adaptive here-and-now approach embodied in the two-stage SLP with general recourse provides planning decisions x that are not dependent on any outcome of the random variable and are hence nonanticipative (Sen and Higle, 1999).

The scenario tree portrayed in Figure 3.1 illustrates this concept of nonanticipativity or implementability. Since information is revealed sequentially, two or more scenarios may share a common sequence of outcomes for the first t periods, with $t < T$, where T denotes the number of periods. For example, scenarios 1 and 2, which correspond to the paths $a \rightarrow b \rightarrow d \rightarrow h$ and $a \rightarrow b \rightarrow d \rightarrow i$, share the same sequence of outcomes in the first two periods, that is $a \rightarrow b$, $b \rightarrow d$ and hence, these two scenarios are indistinguishable until the third period. To maintain implementability, the decisions associated with these two scenarios must be identical in the first two periods. In general, if two scenarios share the same sequence of nodes during the first t periods, they ought to share the same information base during these periods. Consequently, decisions associated with these scenarios must be identical through period t. This is essentially the requirement of the nonanticipativity condition as implicitly honoured in the formulation of the scenario tree in Figure 3.1.

In simple words, nonanticipativity indicates that today's decisions cannot "anticipate" specific occurrences of future random events. Therefore, careful consideration must be given to the timing of all random events, hence rendering the stochastic structure as a secondary characteristic that can only be defined after the temporal structure has been determined (Gassmann, 1998).

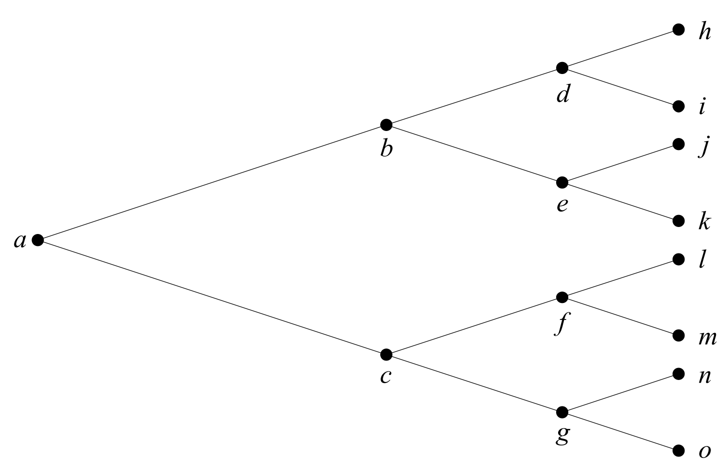


Figure 3.1. The scenario tree is a useful mechanism for depicting the manner in which events may unfold. It can also be utilized to guide the formulation of a multistage stochastic (linear) programming model (Sen and Higle, 1999).

3.10 REPRESENTATIONS OF THE STOCHASTIC PARAMETERS

A key component in formulating stochastic optimization models for decision making under uncertainty is the representation of the stochastic model parameters. According to Escudero et al. (1999), three approaches are conventionally and widely employed to represent and analyze uncertainty or randomness in this type of parameters or data: (1) by its average or mean value, that is, its expected value, (2) in terms of the continuous probability distribution that most aptly describe each item, or (3) based on a representative collection of unplanned events, termed as scenarios, which in precise probabilistic terms, corresponds to a discrete distribution given by a finite probability space (Henrion et al., 2001).

The characterization of each of these parameters as a unique mean value appears to be a vague exercise because it obviously could not be representative of the situation in all cases. The alternative of considering the continuous probability distributions of the parameters implicit in the model definition is a realistic and accurate approach (see for example, Gupta and Maranas, 2003; Petkov and Maranas, 1998; Ierapetritou and Pistikopouskos, 1996c; Wellons and Reklaitis, 1989). However, this approach usually leads to a very complicated model because it typically requires the following information: (a) statistical knowledge of a large number of historical data sets; (b) a given or known continuous probability distribution assumption of estimation; (c) knowledge of certain types of relationships or correlations among the variables; (d) methods of complicated algorithms to solve the formulated model, essentially to evaluate the numerical integration of the expectation terms; and (e) sound knowledge of statistical theory.

Therefore, the third alternative of scenario analysis is advocated as the most promising and practical alternative (within the refinery planning under uncertainty literature, see for example, Neiro & Pinto, 2005; Pongsakdi et al. (in press); Dempster et al., 2000; and Escudero et al., 1999; for a representative work in the operations research literature, see Eppen et al., 1989). The scenario-based technique attempts to represent a random parameter by forecasting all of its possible and likely future outcomes, typically in a scenario tree (or other methods) A representative scenario tree can be constructed, in general, by adopting the following approaches (among many other variants that have been proposed in the literature): (a) the decision-maker, who qualifies as an expert in addressing the problem at hand, defines all the scenario items; (b) the decision-maker defines the scenario set, and computer codes are written to select a representative subset; (c) the decision-maker defines a typical basic scenario and the variability of the parameters, and computer simulations, typically based on Monte Carlo methods, are employed to create the scenario tree. This third option is also useful if a reduction of the state space is desired (Bonfill et al., 2004).

The scenario analysis approach enables the user to define relationships among the realizations of the parameters and between consecutive time periods (or groups of periods), for instance, as joint distributions to account for correlations among parameters. The likelihood of realization of each scenario depends on its assigned relative weight, that is, in effect, its probability. A scenario is defined by a given realization of the parameters along the time horizon. In the model developed in this work, a scenario is given by values of the product demand, commodity prices, and product yields for three possible representative outcomes generalized as the 'realistic" event, the "optimistic" event, and the "pessimistic" event, as well as its probability of happening, together with the deterministic data (Escudero et al., 1999).

However, the infamous shortcoming of the scenario analysis or also known as the progressive hedging approach (Birge and Wets, 1991) is that the number of scenarios increases exponentially with the number of random parameters, resulting in an exponential increase in the problem size. As a result, the computational strategy becomes expensive because the computation time generally increases polynomially (quadratically or even cubically) with the size of the optimization problem (Biegler, 1993). In this aspect, continuous probability distributions for the uncertain parameters could be considered to circumvent this difficulty, in which, a substantial reduction in the size of the problem is usually accomplished at the expense of introducing nonlinearities into the problem through multivariate integration over the continuous probability space. In addition, the continuous distribution-based approach is used particularly in cases where a natural set of discrete scenarios cannot be identified and only a continuous range of potential futures can be predicted. By assigning a probability distribution to the continuous range of potential outcomes, the need to forecast exact scenarios is obviated.

Typically, the distribution-based approach is adopted by modelling the uncertainty as being normally distributed with a specified mean and standard deviation.

Nevertheless, Subrahmanyam et al. (1994) argued that since in many cases in the industry, no sufficiently detailed forecast is available anyway to adequately construct a continuous distribution, a discrete distribution for the uncertain parameters in the form of scenarios still emerges as the most realistic and practical approach. Furthermore, realizations of the random variables in a refinery planning problem generally correspond to a finite number of representative scenarios that need to be taken into account in the search for a "hedging" solution for the optimization under uncertainty problem. This is particularly so when no statistical information is readily available about the uncertain unknown parameters. From the solution perspective, this is an advantageous approach as it eliminates the cumbersome handling of the nonlinear terms introduced by continuous distributions, as stressed earlier. If a continuous distribution is indeed available, it can be approximated by a set of scenarios too, as depicted in Figure 3.2. The continuous distribution may be discretized into a number of parameter values with the associated probabilities given by the corresponding area under the probability distribution function. It is important to note that consideration in selecting the number of scenarios to represent uncertainty or randomness in the parameters is a trade-off between model accuracy and computational efficiency in which a larger number of scenarios would give higher solution accuracy but is computationally expensive at the same time.

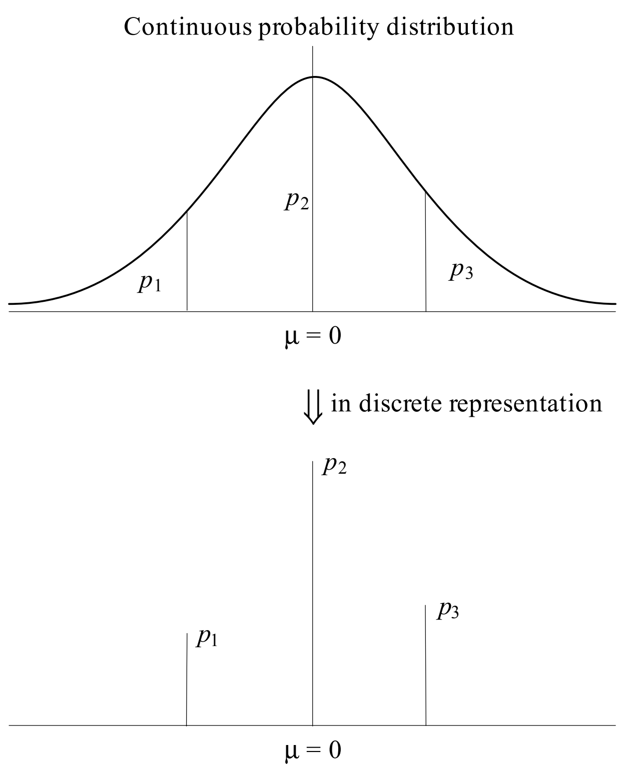


Figure 3.2. Discrete representation of a continuous probability distribution

3.10.1 Accurate Approach for Representation of the Stochastic Parameters via Continuous Probability Distributions

As mentioned in earlier sections, Gupta and Maranas (2003) identified the evaluation of the expectation of the inner recourse problem as the main challenge associated with solving two-stage stochastic problems. For a scenario-based description of uncertainty, this can be achieved by explicitly associating a second-stage variable with each scenario and solving the resulting large-scale extensive formulation (Birge & Louveaux, 1997) by efficient solution techniques such as Dantzig−Wolfe (1960) decomposition and Benders (1962) decomposition. For continuous probability distributions, this challenge has been primarily resolved through the actually similar methodology of explicit/implicit discretization of the probability space in approximating the multivariate probability integrals. The two most commonly used discretization techniques in the chemical process systems engineering (PSE) literature are Monte Carlo sampling techniques (Liu & Sahinidis, 1996; Diwekar & Kalagnanam 1997) (which is also used for discrete distributions) and the Gaussian quadrature formula for approximation of integral evaluation of the expectation terms (Acevedo and Pistikopoulos, 1998; Ierapetritou and Pistikopoulos, 1994c, 1996c; Ierapetritou et al., 1996a; Straub & Grossmann, 1990). The primary advantage of these methods lies in their relative insensitivity to the form of the underlying probability distribution of the uncertain parameter. The Monte Carlo approach lacks in terms of accuracy but avoids the high-dimensional numerical integration since the expectations can be expressed as finite sums, with each constraint duplicated for each scenario, in which a second-stage variable can be associated with each realization of the random parameters (Bonfill et al., 2004).

However, the major downside to them, as also in the scenario-based approach, is the exponential increase in the problem size with the increasing number of uncertain parameters and the scenarios considered due to the nested structure of the two-stage formulation (Shah, 1998). This directly translates to an intensive and excessively large increase in computational requirements, rendering a limit to the practical commercial applicability of these techniques. Petkov and Maranas (1998) propose a methodology to narrow this computational gap, whose work is extended by Gupta and Maranas (2003, 2000). The approach explicitly solves the inner recourse problem analytically for the second-stage variables in terms of the first-stage variables. This is followed by analytical integration over all realizations of the random variables for the evaluation of the expectation terms. By the explicit solution of the inner problem followed by analytical integration over all product demand realizations, the need for discretization of the probability space is obviated. The stochastic attributes of the problem are translated into a resulting equivalent deterministic program at the expense of introducing nonlinearities into the optimization problem. This obviated the need for discretization of the probability space and hence, reduces the associated computational burden.

3.11 SCENARIO CONSTRUCTION

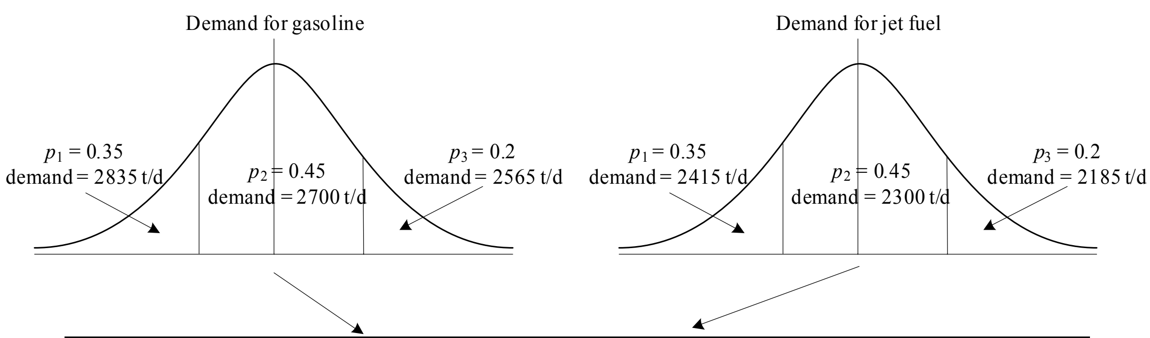
The issue of modelling the stochastic elements is perhaps the most crucial in stochastic optimization. To accomplish this, scenario analysis offers an effective and easily understood tool for addressing the stochastic elements in a multi-stage model. A scenario can be defined as a single deterministic realization of all uncertainties over the planning horizon. Ideally, the process constructs scenarios that represent the universe of possible outcomes (Glynn and Iglehart, 1989; Dantzig and Infanger, 1993). This objective differs from generation of a single scenario, for instance, as carried out in forecasting techniques (or trading strategies in financial practices) (Mulvey et al., 1997).

Each scenario corresponds to a particular outcome of the random elements in a random vector. It is largely a matter of notational convenience that we refer to these vectors and matrices as being random. In most cases, only a small number of the elements are actually random; the rest are constant (the latter are termed as degenerate random variables). In defining the set of scenarios, it is necessary to identify all possible outcomes of the random elements. This consists of identifying the values of those elements that are fixed and the set of all possible outcomes of those random or uncertain elements that vary. In undertaking the latter task, it is important to note the distinctions between models of dependent and independent random variables, which are elaborated as follows (Sen and Higle, 1999).

From a modelling perspective, dependence results when the random elements are subject to a common influence and are most easily described using joint distributions. For example, in a hydroelectric-power-planning model, all hydrological reserves are influenced by the weather. In wet years, reservoirs will tend to be full; in dry years, they will tend toward lower levels. In such a case, it would be convenient to model wet periods and dry periods (or even multiple degrees of wet and dry periods) and to specify the set of reservoir levels that correspond to each type of period. By specifying the probability with which each type of period occurs, one obtains a joint distribution on the reservoir levels (Sen and Higle, 1999).

Independent random variables result when there is no apparent link between the various elements. In this case, one can most easily describe the random elements individually using marginal distributions. For example, in the telecommunication network planning example, the number of calls initiated between any pair of nodes is generally not influenced by the calls between any other pair. Thus, one models the pairwise demand as independent random variables using distributions appropriate to the application. (For example, if it is reasonable to assume that arrival of calls follow a Poisson process, then a Poisson distribution is appropriate.) In this case, a scenario identifies a value for each realization. With independent random variables, the set of all possible outcomes is the Cartesian product of the elemental outcomes for each random variable. The probability associated with any given outcome is the product of the corresponding marginal probabilities. For example, if there are two random variables with five outcomes each and one random variable with four outcomes and the random variables are independent, there are $5 \times 5 \times 4 = 100$ possible scenarios being modeled. It is easy to see that with independent random variables, the number of possible scenarios grows exponentially in the number of random elements (Sen and Higle, 1999).

For the refinery planning problem under uncertainty, consider, for instance, the case of demand uncertainty for two products, gasoline and jet fuel as shown in Figure 3.3. The demand for each is described by three discrete points with point probability associated with each of them. A unique combination of two such points, one from each distribution of gasoline or jet fuel, constitutes a scenario. Assuming that demands for gasoline and jet fuel are completely independent, the associated joint probability of occurrence of both points are given by the product of the individual probabilities. Therefore, the number of scenarios that can be generated is equal to the number of combinations that are possible by considering every pair. Mathematically, this is given by the number of elements of each scenario, that is, the number of possible states in each scenario, raised to the power of the number of random parameters. In this example, a scenario incorporates the possibility of the states of “average” (realistic), “above average” (the optimistic), and “below average” (pessimistic) (number of states = 3) in demand uncertainty for products gasoline and jet fuel (random parameters = 2). Therefore, the total number of scenarios is given by 3 × 3 = 9 scenarios. Notice that the number of scenarios grows exponentially with the number of random parameters and this presents a potential problem. A useful technique is to generate scenarios by employing Monte-Carlo type sampling method for the independent random parameters. A finite number of scenarios thus generated (which will be small in number when compared to the total number of scenarios) is then included in the planning analysis. The selection of scenarios is weighted by their probabilities in which scenarios with higher probability are more likely to be realized. Additionally, the scenario approach to uncertainty allows the designer to readily use intuitive forecasts in the model where a realization of a scenario at any point of time may be easily implemented, without having to cumbersomely deal with continuous distributions (Subrahmanyam et al., 1994).



Scenario	Demand for gasoline (t/d)	Demand for jet fuel (t/d)	Probability	
1	2835	2415	0.35 × 0.35	= 0.1225
2	2835	2300	0.35× 0.45	= 0.1575
3	2835	2185	0.35 × 0.2	= 0.07
4	2700	2415	0.45 × 0.35	= 0.1575
5	2700	2300	0.45 × 0.45	= 0.2025
6	2700	2185	0.45 × 0.2	= 0.09
7	2565	2415	0.2 × 0.35	= 0.07
8	2565	2300	0.2 × 0.45	= 0.09
9	2565	2185	0.2 × 0.2	= 0.04
			Σ(Probability)	= 1.000

Figure 3.3. Scenario generation derived from discrete probability distributions (based on Subrahmanyam et al., 1994)

CHAPTER 4

General Formulation of the Deterministic Midterm Production Planning Model for Petroleum Refineries

4.1 ESTABLISHMENT OF NOMENCLATURE AND NOTATIONS FOR THE DETERMINISTIC APPROACH

4.1.1 Indices

i	for the set of materials or products
j	for the set of processes
t	for the set of time periods

4.1.2 Sets

I	set of materials or products
J	set of processes
T	set of time periods

4.1.3 Parameters

$d_{i,t}$	demand for product i in time period t
$d_{i,t}^{\mathrm{L}}, d_{i,t}^{\mathrm{U}}$	lower and upper bounds on the demand of product i during period t, respectively
$p_t^{\mathrm{L}}, p_t^{\mathrm{U}}$	lower and upper bounds on the availability of crude oil during period t, respectively
$I_{i,t}^{\mathrm{fmin}}, I_{i,t}^{\mathrm{fmax}}$	minimum and maximum required amount of inventory for material i at the end of each time period

$b_{i,j}$	stoichiometric coefficient for material i in process j
$\gamma_{i,t}$	unit sales price of product type i in time period t
λ_t	unit purchase price of crude oil in time period t
$\tilde{\gamma}_{i,t}$	value of the final inventory of material i in time period t
$\tilde{\lambda}_{i,t}$	value of the starting inventory of material i in time period t (may be taken as the material purchase price for a two-period model)
$\alpha_{j,t}$	variable-size cost coefficient for the investment cost of capacity expansion of process j in time period t
$\beta_{j,t}$	fixed-cost charge for the investment cost of capacity expansion of process j in time period t
r_t, o_t	cost per man-hour of regular and overtime labour in time period t

4.1.4 Variables

$x_{j,t}$	production capacity of process j ($j = 1, 2, \ldots, M$) during time period t
$x_{j,t-1}$	production capacity of process j ($j = 1, 2, \ldots, M$) during time period $t-1$
$y_{j,t}$	vector of binary variables denoting capacity expansion alternatives of process j in period t (1 if there is an expansion, 0 if otherwise)
$CE_{j,t}$	vector of capacity expansion of process j in time period t
$S_{i,t}$	amount of (commercial) product i ($i = 1, 2, \ldots, N$) sold in time period t
$L_{i,t}$	amount of lost demand for product i in time period t
P_t	amount of crude oil purchased in time period t
$I_{i,t}^{\mathrm{s}}, I_{i,t}^{\mathrm{f}}$	initial and final amount of inventory of material i in time period t
$H_{i,t}$	amount of product type i to be subcontracted or outsourced in time period t
R_t, O_t	regular and overtime working or production hours in time period t

4.1.5 Superscripts

$(\ )^L$ lower bound

$(\ )^U$ upper bound

4.2 LINEAR PROGRAMMING (LP) FORMULATION OF THE DETERMINISTIC MODEL

The basic framework for the deterministic linear production planning model for a petroleum refinery will be mainly derived from models formulated by McDonald and Karimi (1997) and Ierapetritou and Pistikopoulos (1994a), apart from the models specific to refinery planning as proposed by Pongsakdi et al. (in press), Dempster et al. (2000), and Escudero et al. (1999). In addition, we consider the remarks by Kallrath (2002) that a refinery planning model should comprise the following constraints:

- flow of crude oil and components for blending operations (in the form of linear material balances);
- proportional composition of flow streams (in the form of nonlinear equations);
- quality constraints and capacity limits of processing/production units and storage tanks (in the form of inequalities denoting suitable lower and upper bounds); and
- assignment of processing/production units and storage tanks (in the form of equations and inequalities involving binary variables).

Consider the production planning problem of a typical oil refinery operation with a network of M continuous processes and N materials as shown in Figure 4.1. Let $j \in J = (1, 2,\ldots, M)$ index the set of continuous processes whereas $i \in I = (1, 2,\ldots, N)$ index the set of materials. These products are produced during n time periods indexed by $t \in T = (1, 2,\ldots, n)$ to meet a prespecified level of demand during each period. Given also are the prices and availabilities of materials as well as investment and operating cost data over a time period. The problem then consists of determining (i) the production profiles; (ii) sales and purchases of chemical products; and (iii) capacity expansions for the existing processes over each time period that will maximize profit over the time period by also ensuring future feasibility. A typical aggregated linear planning model consists of the following sets of constraints.

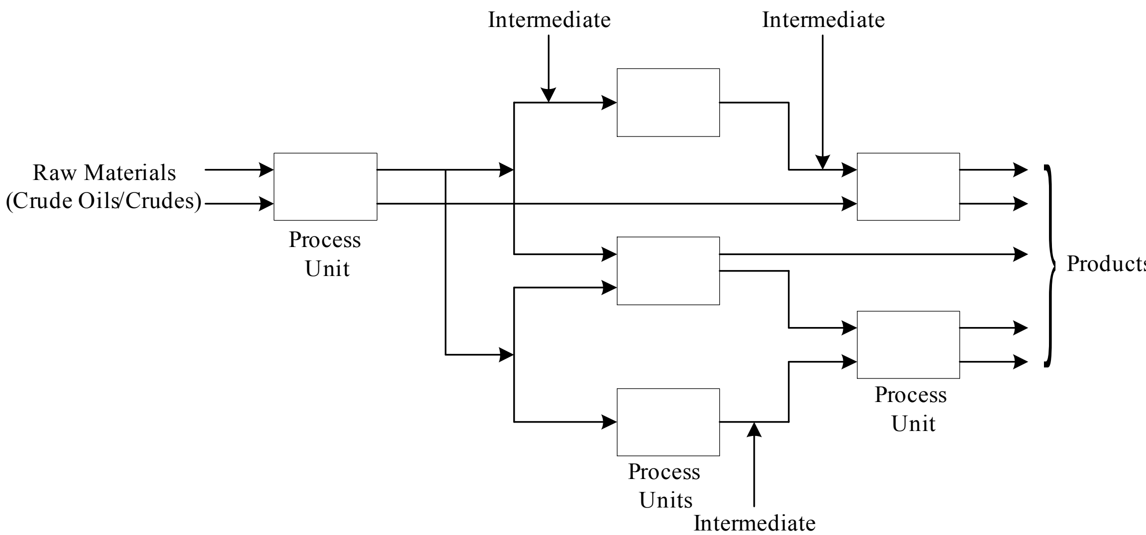


Figure 4.1. A network of processes and materials of a typical oil refinery operation (based on Ierapetritou and Pistikopoulos, 1994a)

4.2.1 Production Capacity Constraints

$$x_{j,t} = x_{j,t-1} + CE_{j,t} \qquad \forall j \in J \tag{4.1}$$

$$y_{j,t} CE^{\mathrm{L}}_{j,t} \leq CE_{j,t} \leq y_{j,t} CE^{\mathrm{U}}_{j,t} \qquad \forall j \in J, \forall t \in T \tag{4.2}$$

where
$$y_{j,t} = \begin{cases} 1 & \text{if there is an expansion} \\ 0 & \text{otherwise} \end{cases} \tag{4.3}$$

where $x_{j,t}$ denotes the production capacity of process j during time period t; $CE_{j,t}$ represents the (potential) capacity expansion of process j in t; $y_{j,t}$ are binary variables deciding on expansion of process j in period t; and $CE^{\mathrm{L}}_{j,t}$ and $CE^{\mathrm{U}}_{j,t}$ are the constant lower and upper bounds of the capacity expansion variables $CE_{j,t}$, respectively. If capacity expansion is not considered, then equation **(4.1)** becomes

$$x_{j,t} \leq x_{j,t-1} \qquad \forall j \in J \tag{4.4}$$

4.2.2 Demand Constraints

$$S_{i,t} + L_{i,t} = d_{i,t}, \quad \forall i \in I_{\mathrm{P}}, \forall t \in T \tag{4.5}$$

$$d_{i,t}^{\mathrm{L}} \leq S_{i,t} \leq d_{i,t}^{\mathrm{U}}, \quad \forall i \in I_{\mathrm{P}}, \forall t \in T \tag{4.6}$$

where $S_{i,t}$ denotes the amount of product i sold in time period t; $L_{i,t}$ is the amount of lost demand for product i in time period t; $d_{i,t}$ is the demand for product i in time period t; and $d_{i,t}^{\mathrm{L}}$ and $d_{i,t}^{\mathrm{U}}$ are the lower and upper bounds on the demand of product i during period t, respectively.

4.2.3 Availability Constraints

$$p_t^{\mathrm{L}} \leq P_t \leq p_t^{\mathrm{U}} \qquad i = 1, 2, ..., N;\ t = 1, 2, ..., T \tag{4.7}$$

where P_t denotes the amount of crude oil purchased in time period t while p_t^{L} and p_t^{U} are the lower and upper bounds of the availability of crude oil during period t, respectively.

Note that an instance for which the bounds defined by **(4.6)** and **(4.7)** could arise is in the case of long-term contracts in which fixed amounts of sales or purchases/procurements are committed over several time periods (Iyer and Grossmann, 1998).

4.2.4 Inventory Requirements

In addition to the amount of materials purchased and/or produced, a certain level of inventory must be maintained at both time periods to ensure material availability. If the starting (initial) and final amount of inventory of material i in time period t is represented with the variables $I^{\mathrm{s}}_{i,t}$ and $I^{\mathrm{f}}_{i,t}$, respectively, the following conditions hold:

$$I^{\mathrm{f}}_{i,t} = I^{\mathrm{s}}_{i,t+1} \quad i = 1,\ldots,N,\ t = 1,\ldots,T \qquad \textbf{(4.8)}$$

$$I^{\mathrm{fmin}}_{i,t} \le I^{\mathrm{f}}_{i,t} \le I^{\mathrm{fmax}}_{i,t} \qquad i = 1,2,\ldots,N \qquad \textbf{(4.9)}$$

where $I^{\mathrm{fmin}}_{i,t}$ and $I^{\mathrm{fmax}}_{i,t}$ are the minimum and maximum required amount of inventory for material i at the end of each time period. Equation **(4.8)** simply states that $I^{\mathrm{s}}_{i,t+1}$, the starting inventory of material i in time period t+1 is the same as $I^{\mathrm{f}}_{i,t}$, the inventory of material i at the end of the preceding period t (if t = 1, then $I^{\mathrm{f}}_{i,t} = I^{\mathrm{f}}_{i,1}$ denotes the initial inventory).

4.2.5 Material Balances (or Mass Balances)

$$P_t + I^{\mathrm{s}}_{i,t} + \sum_{j \in J} b_{i,j} x_{j,t} - S_{i,t} - I^{\mathrm{f}}_{i,t} = 0, \quad i \in I,\ t \in T \qquad \textbf{(4.10)}$$

where $b_{i,j}$ is the stoichiometric coefficient for material i in process j. These balances can be further classified into three categories, namely for (i) fixed production yields; (ii) for fixed blends or splits; and (iii) for unrestricted balances, as accounted for explicitly by the numerical example to be studied later.

4.2.6 Objective Function

A profit function over the time horizon is considered as the difference between the revenue due to product sales and the overall cost, which consists of cost of raw materials, operating cost, investment cost, and inventory cost:

$$\text{Profit} = \sum_{t \in T} \left[\begin{array}{l} \sum_{i \in I} \gamma_{i,t} S_{i,t} + \sum_{i \in I} \tilde{\gamma}_{i,t} I^{\mathrm{f}}_{i,t} - \sum_{i \in I} \lambda_{i,t} P_{i,t} - \sum_{i \in I} \tilde{\lambda}_{i,t} I^{\mathrm{s}}_{i,t} - \sum_{j \in J} C_{j,t} x_{j,t} - \sum_{i \in I} h_{i,t} H_{i,t} \\ -\sum_{j \in J} \left(\alpha_{j,t} CE_{j,t} + \beta_{j,t} y_{j,t} \right) - \left(r_t R_t + o_t O_t \right) \end{array} \right] \quad \textbf{(4.11)}$$

where for the purpose of clearer presentation, the definition of each term is presented as follows:

$\gamma_{i,t}$ = unit sales price of product type i in time period t;

$\lambda_{i,t}$ = unit purchase price of product type i in time period t;

$\tilde{\gamma}_{i,t}$ = value of the final inventory of material i in time period t;

$\tilde{\lambda}_{i,t}$ = value of the starting inventory of material i in time period t (may be taken as the material purchase price for a two-period model);

$C_{j,t}$ = operating cost of process j in time period t;

$h_{i,t}$ = unit cost of subcontracting or outsourcing the production of product type i in time period t;

$H_{i,t}$ = amount of product type i to be subcontracted or outsourced in time period t;

$\alpha_{j,t}$ = variable-size cost coefficient for the investment cost of capacity expansion of process j in time period t;

$\beta_{j,t}$ = fixed-cost charge for the investment cost of capacity expansion of process j in time period t;

r_t, o_t = cost per man-hour of regular and overtime labour in time period t, respectively;

R_t, O_t = regular and overtime working or production hours in time period t, respectively.

CHAPTER 5

General Formulation of the Stochastic Midterm Production Planning Model under Uncertainty for Petroleum Refineries

The refinery production planning problem under uncertainty differs from the deterministic problem in that some (or even all) of the planning parameters or coefficients are considered to be random variables. The production planning objective function must now not only represent the net profit to be derived from the sales of refined products (based on the amount of crude oil purchased); it must also reflect a measure of system performance. The ultimate goal of the planning problem is then to determine the maximum profit expected by implementing a production planning scheme that will operate in a feasible manner while accounting for the expected revenue loss mainly due to unmet demand and to a lesser degree, surplus of production (Wellons and Reklaitis, 1989).

5.1 STOCHASTIC PARAMETERS

The following are the uncertain or random parameters considered in this work:

- market demand or product demand (where similar meaning is implied in the interchangeability of both terms);
- prices of crude oil (the raw material) and the final saleable products, referred to collectively as the prices of commodities; and
- product yields of crude oil from chemical reactions in the primary distillation unit of a typical petroleum refinery.

In spite of the resulting exponential increase in problem size, the scenario analysis approach has been extensively applied and invoked in the open literature and has been proven to provide reliable and practical results (Gupta and Maranas, 2003; 2000). Hence, in this work, it is adopted for describing uncertainty in the stochastic parameters. Representative scenarios are constructed to model uncertainty in prices, product demand, and production yields. This is accomplished within one of the most widely-used structures for decision making under uncertainty, that is, the two-stage stochastic programming framework. To reemphasize, in this framework, the decision variables of the problem are partitioned into two sets. The first-stage planning variables correspond to decisions that need to be made prior to resolution of uncertainty (the "here-and-now" decisions). Subsequently, based on these decisions and the realization of the random events, the second-stage operating decisions are made subject to the restrictions of the second-stage recourse problem (the "wait-and-see" decisions). The presence of uncertainty is translated into the stochastic nature of the recourse penalty costs associated with the second-stage decisions. Therefore, the objective function consists of the sum of profits or costs determined by the first-stage decisions and the expected second-stage recourse costs (Gupta and Maranas, 2000).

5.2 APPROACHES UNDER UNCERTAINTY

The stochastic model is developed with the objective of yielding a solution that is less sensitive to the presence of uncertainties. The model attempts to minimize variation in profits and costs that arise due to operation under unplanned events or scenarios. Two methods to deal with uncertainty, discussed in light of the recent presentation by Nelissen (http://www.gams.com/presentations/present_uncertainty.pdf, accessed December 17, 2005), are employed and combined in the proposed stochastic models in this work:

- the static model of Markowitz's mean–variance (*E*–*V* or MV) approach to handle randomness in the objective coefficients of prices by minimizing the expected cost and the variance (for the given expected value or mean of the objective function); and
- the dynamic approach of two-stage stochastic programming with fixed recourse to handle randomness in the left-hand side (LHS) (or also known as technological coefficients) and the right-hand side (RHS) coefficients by trading off between maximizing profit and minimizing the impacts of the associated recourse penalty costs through accounting for their expected values and deviations; deviations are

representation of the risks taken and in this work, they are quantified by two measures, namely the variance and the mean-absolute deviation (MAD).

The following four approaches and approximation schemes are implemented in the formulation of the stochastic models:

1. Approach 1: the Markowitz's mean–variance model to handle randomness in the objective coefficients of prices by minimizing the variance of the expected value or mean of the random coefficients subject to a target value constraint;

2. Approach 2: the two-stage stochastic programming with fixed recourse approach to model randomness in the right-hand side and left-hand side (or technological) coefficients by minimizing the expected recourse penalty cost due to violations of constraints;

3. Approach 3: incorporation of the Markowitz's *E–V* approach within the two-stage stochastic programming framework developed in Approach 2 in order to minimize both the expectation and the variance of the recourse penalty costs; this results in a stochastic quadratic programming model with fixed recourse; and

4. Approach 4: reformulation of the model developed in Approach 3 by utilizing the Mean–Absolute Deviation (MAD) as the measure of risk imposed by the recourse penalty costs

As an overview, in the two-stage stochastic programming approach to a production planning problem under uncertainty, the decision variables are classified into two sets. The first-stage variables, which are often known as planning variables themselves, are those that have to be decided before the actual realization of the uncertain parameters. Planning variables depend only on the fixed and structural constraints that are independent of uncertainty. These first-stage planning variables are typically the amount of raw materials needed, the production rates required, and others. Subsequently, once the values of the planning variables have been decided and the random events have presented themselves, further operational policy improvements can be made by selecting, at a certain cost, the values of the second-stage recourse variables, also known as the control or operating variables for implementing corrective actions. Due to uncertainty, the second-stage cost is a random variable. These second-stage recourse variables typically determine the amount of products to be purchased from other producers (or to be outsourced) to meet the market demand actually realized or the amount of raw material required from other suppliers to achieve production requirements (Li et al., 2004).

5.3 GENERAL TECHNIQUES FOR MODELLING UNCERTAINTY

In general, uncertainties in commodity prices, future product demand, and process yields can be modelled in either of these two ways: first, by considering a specified range T defined as follows:

$$\theta \in T,\ T = \left\{ \theta \middle| \theta^{L} \le \theta \le \theta^{U} \right\} \qquad \textbf{(5.1)}$$

where
$\theta^{L} = \theta^{N} - \Delta\theta^{-}$,
$\theta^{U} = \theta^{N} + \Delta\theta^{+}$,
θ^{N} = vector of the nominal values of the uncertain parameters,
$\Delta\theta^{-}, \Delta\theta^{+}$ = expected positive and negative deviations, respectively;

or second, by providing probability distribution functions (Ierapetritou and Pistikopouslos, 1994a; Ponnambalam, 2005). This work adopts the former technique throughout.

5.4 ESTABLISHMENT OF NOMENCLATURE AND NOTATIONS FOR THE STOCHASTIC APPROACH

5.4.1 Indices

s for the set of scenarios

k for the set of products with yield uncertainty

5.4.2 Sets

I set of materials or products

K set of products with yield uncertainty

5.4.3 Stochastic Parameters

p_s probability of scenario s

$\gamma_{i,s,t}$ unit sales price of product type i in time period t per realization of scenario s

$\lambda_{t,s}$ unit purchase price of crude oil in time period t per realization of scenario s

$d_{i,s,t}$ demand for product i in time period t per realization of scenario s

5.4.3.1 Recourse Parameters

c_i^+ fixed unit penalty cost for shortfall in production (underproduction) of product type i

c_i^- fixed unit penalty cost for surplus in production (overproduction) of product type i

$q_{i,k}^+$ fixed unit penalty cost for shortage in yields from material i for product type k

$q_{i,k}^-$ fixed unit penalty cost for excess in yields from material i for product type k

5.4.4 Stochastic Recourse Variables (Second-Stage Decision Variables)

$z_{i,s}^+$ amount of underproduction of product type i per realization of scenario s

$z_{i,s}^-$ amount of overproduction of product type i per realization of scenario s

$y_{i,k,s}^+$ amount of shortage in yields from material i for product type k per realization of scenario s

$y_{i,k,s}^-$ amount of excess in yields from material i for product type k per realization of scenario s

5.5 APPROACH 1: RISK MODEL I BASED ON THE MARKOWITZ'S MEAN–VARIANCE APPROACH TO HANDLE RANDOMNESS IN THE OBJECTIVE FUNCTION COEFFICIENTS OF PRICES

5.5.1 Uncertainty in the Price of Crude Oil

The rapid rise in world petroleum crude oil prices in years 2004 to 2005, which triggered off a similar trend in petroleum-based fuel prices, has cast much uncertainty in the forecasting of future oil prices. This is a direct result of various events of global impact including (but certainly not confined to) increased demand from the emerging economic power of China; political conflicts and instability in major hydrocarbon resources and petroleum supplier countries in the Middle East; high gasoline demand from North America; an all-time thirty-year low of oil stocks in the Organization of Economic Co-operation and Development (OECD) countries; supply uncertainty from Iraq, Nigeria, Russia, and Venezuela; and the disparity between crude availability and refining capacity. Coupled with the peaking in the world oil production and consumption (Hirsch et al., 2006), it has indeed become highly pertinent to take into account the factor of crude oil price uncertainty in refinery production planning, arguably the heart of the downstream processing sector of the petroleum industry, as equally emphasized by Neiro and Pinto (2005). All these factors compounds the intricacies of the crude oil price determination process, this inevitably necessitates extending the price analysis beyond the markets for petroleum. Didziulis (1990) reports that crude oil prices are determined in two closely-related markets, namely the crude oil markets and the refined products markets. As the raw material used in refineries in joint-products processes, for a given level of supply, the value of petroleum or crude oil lies in the value of refined products derived from it. Further elaboration on this issue is obviously beyond the scope of this work and the interested reader is referred to the rich literature available on this subject.

As readily recognized, a competing issue facing decision-makers in oil refineries is cost, which is directly related to price. To guarantee supplies of crude oil and availability of transportation, decision-makers must effectively pay a premium. They can purchase crude on a "spot" basis in each period at spot prices that could be lower than market prices or firm prices. Spot prices, however, cannot be determined in advance. In addition, crude availability at the spot price is by no means certain or guaranteed, nor can the quantity of crude oil available be predicted. Similarly, transportation capacity can be purchased in an "interruptible" basis in each period at a potentially lower cost than for firm purchases, but again, availability and capacity are neither guaranteed nor predictable in advance (Bopp et al., 1996).

The Petroleum Division of the Energy Information Administration (EIA) of the Department of Energy (DOE), United States of America maintains an excellent website providing recent and current information on price data of crude oil and its refined products at http://www.eia.doe.gov/pub/oil_gas/petroleum/analysis_publications/oil_market_basics/Price_links.htm (accessed December 27, 2005) while historical price data are accessible at http://www.eia.doe.gov/neic/historic/hpetroleum.htm (accessed December 28, 2005). Specific information on crude oil price in chronological order dated since 1970 until the daily present time is available at http://www.eia.doe.gov/emeu/cabs/chron.html (accessed December 27, 2005). To model the uncertainty in crude oil price in this work, historical data on a daily basis for the Brent crude oil and the West Texas Intermediate (WTI) at Cushing crude oil for the years 2004 and 2005 is analyzed and is considered to be representative for midterm production planning activities (the justification for this reasoning follows). Figure 5.1 depicts the daily price for both types of crude oil for the considered period of January 5, 2004–December 30, 2005. The complete numerical data for this period and the associated computed analytical results are provided in detail in Appendix B.

As stated earlier, the trend of oil price in this two-year period is considered to be representative as it captured the four major events of spikes in oil price experienced since the triggering of the rising oil price phenomenon, namely:

1. on October 22, 2004: the cumulative effects of the war on Iraq launched by the government of the U.S.A. rapid increases in global demand for crude oil, constrained capacity of the Organization of the Petroleum Exporting Countries (OPEC), and low worldwide inventories resulted in an all-time high of $55.17 per barrel for crude oil contract price as reported by the New York Mercantile Exchange (NYMEX) for the WTI. The aftermath effects of Hurricane Ivan forced

the temporary termination of natural gas and crude oil production from the Gulf Coast;

2. on April 4, 2005: Chevron-Texaco, the major oil company with ownership of the fourth largest non-state-owned oil reserves in the world, agreed to buy Unocal, a medium-size U.S.-based oil company. It was the largest merger–acquisition exercise in the oil and gas industry since 2001;
3. on July 5, 2005: Tropical storm Cindy interrupted oil and natural gas production in the U.S.'s Gulf of Mexico region. The storm shut off oil and gas platforms, forced the closure of the Louisiana Offshore Oil Port (the largest U.S. oil-import terminal) and some refineries also ceased operations;
4. on August 28, 2005: Hurricane Katrina hit the U.S. Gulf of Mexico region near New Orleans, resulting in a severe impact on the local oil and natural gas production: shut down of key hydrocarbons infrastructure including the Louisiana Offshore Oil Platform, the Capline crude oil pipeline, and the Colonial and Plantation oil products pipelines; and disruption in operations of oil refineries. The U.S. government announced that it would loan out crude oil from the Strategic Petroleum Reserve to alleviate the situation and members of the International Energy Agency (IEA) pledged offers of emergency reserves to the U.S..

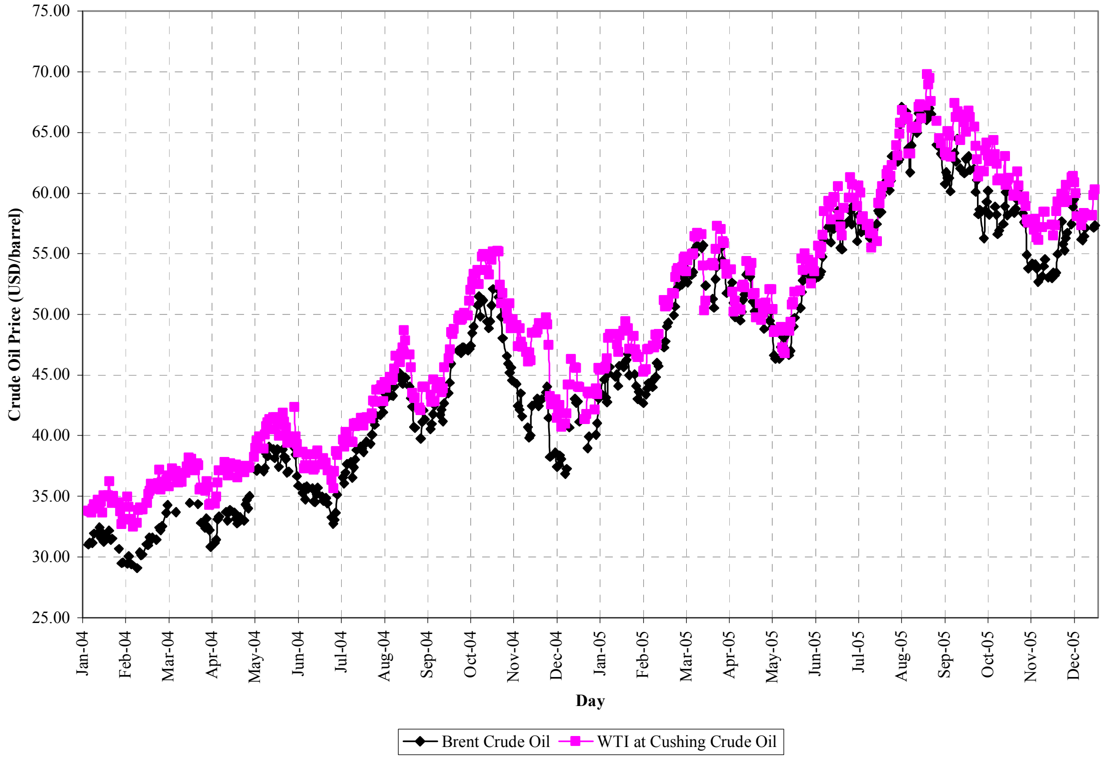


Figure 5.1. Daily crude oil price data for the period January 5, 2004–December 30, 2005 (Energy Information Administration (EIA), 2005)

From the data, the mean price of crude oil is determined. Three different statistics are then employed to investigate and portray the degree of variations in the data that reflect price uncertainty: (a) the standard error, that is, the percentage of difference from the mean or average value; (b) the sample standard deviation, that is, the square root of sample variance, which measures how widely values are dispersed from the mean (Ross, 2004); and (c) the maximum price. The statistical analysis computation is executed by utilizing the *Descriptive Statistics* tool for *Data Analysis* in the Microsoft® Excel spreadsheet software package (Microsoft Corporation, 2001). The results of the analysis are summarized in Table 5.1.

Table 5.1. Statistics of daily crude oil price data for the period of January 5, 2004–December 30, 2005 (Energy Information Administration (EIA), 2005)

Crude Oil Type	Mean Price (US$/barrel)	Standard Error	Standard Deviation	Maximum Price
Brent	47.04	0.4704	9.990	67.09
West Texas Intermediate (WTI) at Cushing	48.94	0.4325	9.671	69.82

5.5.2 Uncertainty in the Prices of the Major Saleable Refining Products of Gasoline, Naphtha, Jet Fuel, Heating Oil, and Fuel Oil

Table 5.2 provides statistics of daily price data for the period January 5, 2004–December 26, 2005 (Energy Information Administration (EIA), 2006) for the major saleable refining products considered in the numerical study to be presented in the following chapter. The products are gasoline, naphtha, jet fuel, heating oil, and fuel oil.

Table 5.2. Statistics of daily price data for the major saleable refining products of gasoline, naphtha, jet fuel, heating oil, and fuel oil for the period of January 5, 2004–December 30, 2005 (Energy Information Administration (EIA), 2005)

Product	Mean Price (cent/gallon)	Standard Error	Standard Deviation	Maximum Price	Remark
Gasoline	210.46 cent/gallon	3.139	32.01	311.7	Daily USA retail gasoline price data for all grades and all formulations
(*Source*: *Retail Gasoline Historical Prices*, http://www.eia.doe.gov/oil_gas/petroleum/data_publications/wrgp/mogas_history.html, accessed on January 23, 2006)					
Naphtha	$343/ton MOPJ*		(not available)		
(*Source*: *Naphtha Remained Firm at $340–350*, ChemLOCUS Chem Journal: Korea Chemical Market Information, http://www.chemlocus.com/news/daily_read.htm?menu=D1&Sequence=6538&cpage=14&sub=, accessed on January 24, 2006)					
Jet Fuel (Kerosene-type)	143.1909 cent/gallon	1.800 172	40.091 74	313.03	Daily USA Gulf Coast kerosene-type jet fuel spot price FOB (free-on-board)
(*Source*: *Historical Petroleum Price Data—Other Product Prices*, http://www.eia.doe.gov/neic/historic/hpetroleum2.htm#Other, accessed on January 23, 2006)					
Heating Oil	204.4566 cent/gallon	4.902 018	33.962 18	269.159	Weekly USA No. 2 heating oil residential price (excluding tax)
(*Source*: *Heating Oil and Propane Update*, http://tonto.eia.doe.gov/oog/info/hopu/hopu.asp, accessed on January 23, 2006)					
Fuel Oil	86.813 04 cent/gallon	3.978 387	19.079 67	126.7	Monthly USA residual fuel oil retail sales by all sellers (until November 30, 2005)
(*Source*: *Residual Fuel Oil Prices by Sales Type*, http://tonto.eia.doe.gov/dnav/pet/pet_pri_resid_dcu_nus_m.htm, accessed on January 24, 2006)					

**Note*: MOPJ stands for “Mean of Platts Japan”, one of the industry’s standards for reporting prices of petroleum products. For further information, the reader is referred to Platts Methodology and Specifications Guide for Asian Naphtha at http://www.platts.com/Content/Oil/Resources/Methodology%20&%20Specifications/method_asian_naptha_2004.pdf (accessed on March 29, 2006).

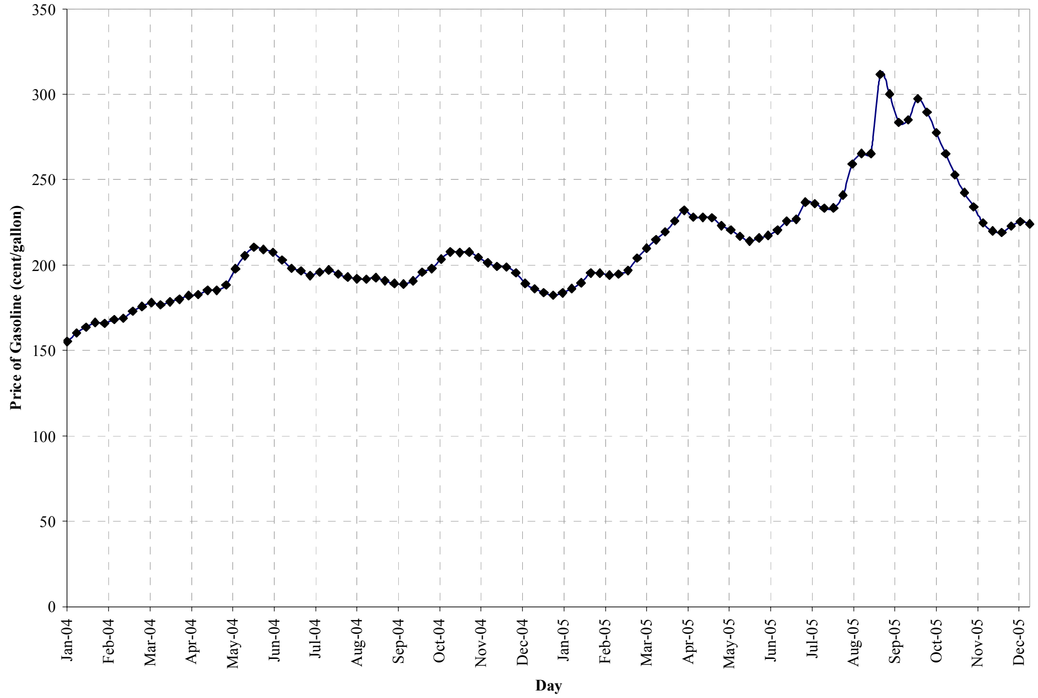


Figure 5.2. Weekly USA retail gasoline price (cents per gallon) for all grades and all formulations for the period of January 5, 2004–December 26, 2005 (Energy Information Administration (EIA), *Retail Gasoline Historical Prices*, http://www.eia.doe.gov/oil_gas/petroleum/data_publications/wrgp/mogas_history.html, accessed on January 23, 2006).

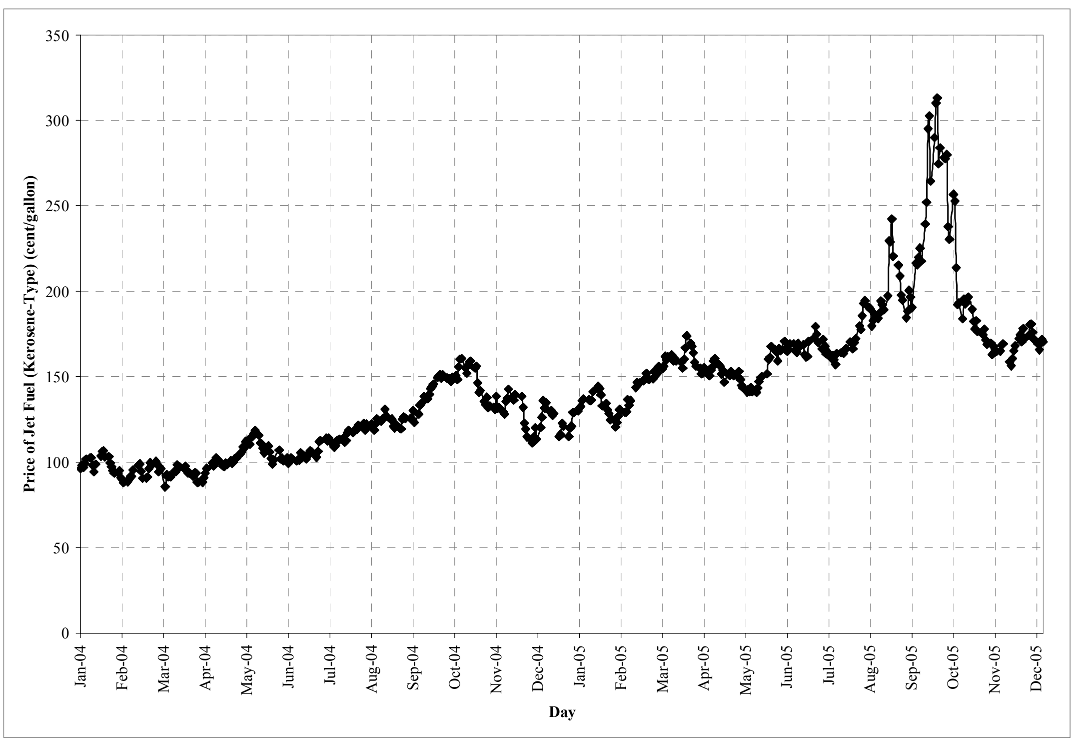


Figure 5.3. Daily USA Gulf Coast kerosene-type jet fuel spot price FOB (free-on-board) for the period of January 5, 2004–December 23, 2005 (Energy Information Administration (EIA), *Historical Petroleum Price Data—Other Product Prices*, http://www.eia.doe.gov/neic/historic/hpetroleum2.htm#Other, accessed on January 23, 2006).

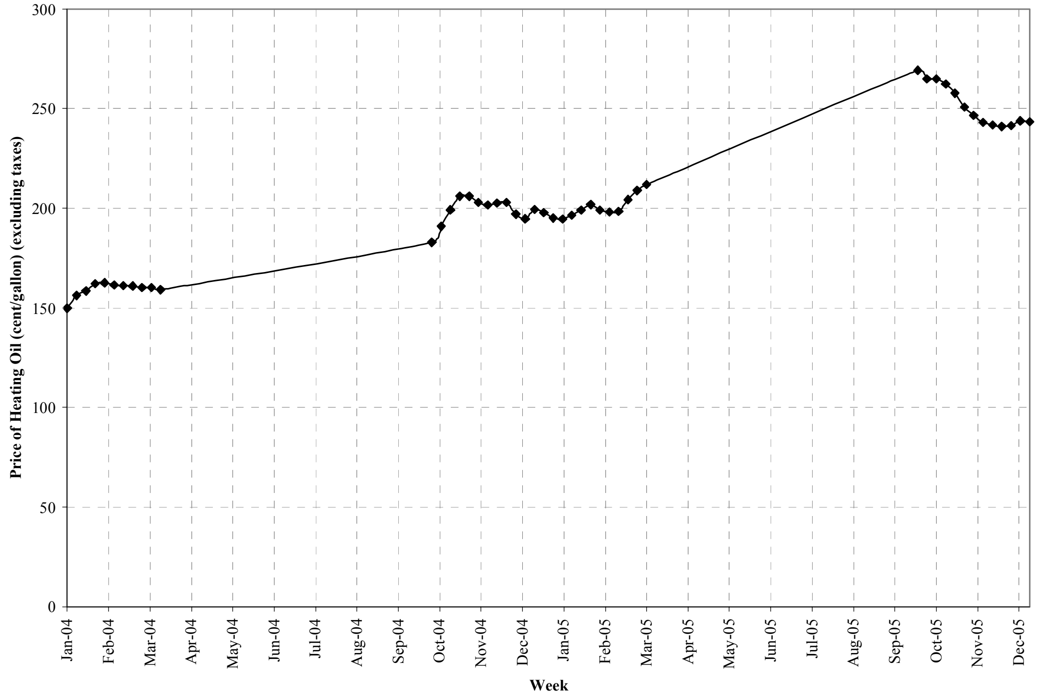


Figure 5.4. Weekly USA No. 2 heating oil residential price (cents per gallon excluding taxes) for the period of January 5, 2004–December 26, 2005 (Energy Information Administration (EIA), *Heating Oil and Propane Update* at http://tonto.eia.doe.gov/oog/info/hopu/hopu.asp, accessed on January 23, 2006). (*Additional note*: The No. 2 heating oil is a distillate fuel oil for use in atomizing type burners for domestic heating or for use in medium capacity commercial–industrial burner units.)

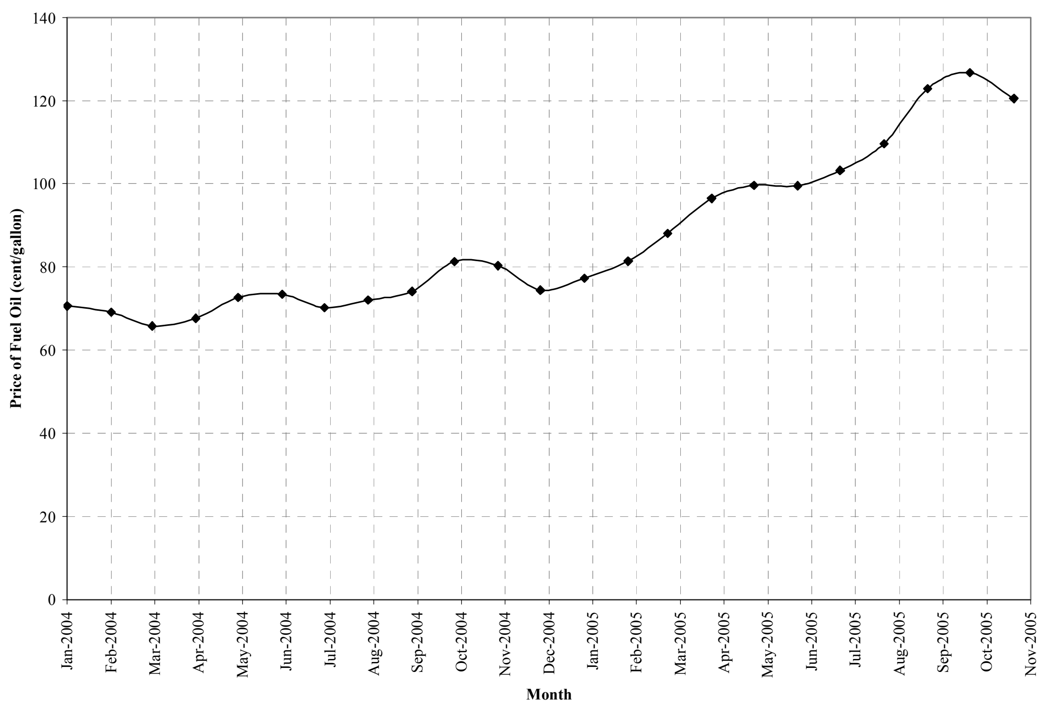


Figure 5.5. Monthly USA residual fuel oil retail sales by all sellers (cents per gallon) for the period of January 5, 2004–November 30, 2005 (Energy Information Administration (EIA), *Residual Fuel Oil Prices by Sales Type*, http://tonto.eia.doe.gov/dnav/pet/pet_pri_resid_dcu_nus_m.htm, accessed on January 24, 2006. Note that there is no data available for the following periods: (i) between March 16, 2004 and October 3, 2004 and (ii) between March 15, 2005 and October 2, 2005.

5.5.3 Reportage of Oil Prices

Agencies specializing in reporting prices of petroleum and petroleum products market such as *Platts Oilgram Journal* (McGraw-Hill), *Argus*, and the *London Oil Report* obtain price levels estimates from the trading houses. The resulting price estimates become the market reference prices that are used to set prices for other transactions. Typically, prices for more than 14 types of products are quoted daily. For certain products, there are three forms of price quotations, depending on whether shipment is by cargo or by barge and whether the price if on the basis of free-on-board (FOB) or cost, insurance, and freight (CIF).

These published prices represent the estimated value at a particular time of a cargo of a standard product of known characteristics; for instance, for gas oil, these properties would be mass in tonnage, relative density, sulphur content, cetane index, and others. Thus, the published quotations are far from being representative of the variety of products that are actually traded.

The selling price for a cargo is agreed in terms of a differential from an agreed quotation. This adjustment factor principally takes into account the tonnage, the method of transportation, and the quality, plus all other aspects relevant to any commercial transaction (Favennec, 2001)

5.5.4 Stochastic Modelling of Randomness in the Objective Function Coefficients of Prices for the General Deterministic Model

The classical approach to model the tradeoffs between expectation and variability in a stochastic optimization problem is to employ variance as the measure of variability or dispersion. This gives rise to adopting the well-known mean–variance (*E*–*V*) portfolio optimization model of Markowitz (1952, 1959), conveniently referred to as the MV or *E*–*V* approach. The goal of the Markowitz model consists of two criteria, namely to maximize the first criterion of expected profit while appending a limiting constraint on the magnitude of the second criterion of risk, which is measured by using variance

(Eppen et al., 1989). Malcolm and Zenios (1994) presented an application of this approach by adopting the robust optimization framework proposed by Mulvey (1995) to the problem of capacity expansion of power systems. The problem was formulated as a large-scale nonlinear program with the variance of the scenario-dependent costs included in the objective function. Another application using variance is presented by Bok et al. (1998), also within a robust optimization model, for investment in the long-range capacity expansion of chemical process networks under uncertain demand. A brief review of the Markowitz's model is presented in Appendix A.

In his model, Markowitz (1952) introduced the concept of portfolio management theory using a mathematical programming approach. An investor has a choice between various financial instruments whose rate of return is uncertain. Theoretically, the investor should maximize expected utility, but this utility function is not usually available. Instead, the Markowitz approach is to "draw" the so-called efficient frontier, in which for a given expected return, one solves a quadratic program that identifies the portfolio minimizing variance. (Alternatively, the efficient frontier can be interpreted as the solution that minimizes variance for an each expected return.) A plot of expectation (typically, expected profit) versus variance is produced. The onus is then on the decision maker to choose a point on this efficient frontier corresponding to the desired profit with the associated bearable amount of risk (Wets, 1996).

It follows that the application of portfolio theory to the selection of production planning programs will involve the determination of sets of programs that are efficient in the return–risk (or profit–risk) space. In this approach, portfolio variance is used as the measure of risk. An efficient portfolio is defined as the minimum variance portfolio that yields a specified level of expected return or profit under relevant constraints characterizing the decision space. The efficient frontier is obtained by solving the problem for a set of exogenously specified expected income or profit levels and joining the optimum solutions in the expected profit–variance (E–V) space. This framework, which leads to a convex quadratic programming problem, is a widely used portfolio management technique in the finance literature especially because of its desirable theoretical properties (Cabrini et al., 2004).

As argued in the previous section, since prices are likely to be uncertain and to show variability due to a multitude of possible reasons, it is desired to assume a distribution of possible prices rather than a fixed price (typically the expected value or mean) for each product. The goal of the stochastic model then is to determine production quantities (as given by the variables x_i, with i indicating the product type) that meet specified requirements, as dictated by market demand, by simultaneously minimizing the various expected production costs.

5.5.5 Sampling Methodology by Scenario Generation for the Recourse Model under Price Uncertainty

A collection of scenarios that best captures and describes the trend of prices of the different types of crude oil as raw material feeds and prices of the saleable refining products for a reasonable period of time are generated based on available historical data as presented in Section 5.6.1. The solution is bound to be more robust and representative with more scenarios considered; however, as cautioned earlier, the major pitfall with the recourse problem via scenario analysis is the explosive nature of exponential increase in problem size with the number of uncertain parameters. Weights representing an *a priori* probability measure can be assigned to all possible outcomes ω of the outcome space Ω. As scenarios represent every possible environment ω that becomes an element of the probability space, the associated probabilities of p_s with index s = 1, 2, …, NS denoting the sth scenario are assigned to each scenario respectively to reflect the corresponding likelihood of each scenario of state of the world being realized (Ermoliev and Wets, 1988), with $\sum_{s \in S} p_s = 1$. Table 5.3 summarizes attributes of the scenarios constructed for modelling price uncertainty whereas Table 5.4 presents the scenario construction to model price uncertainty for a refinery producing i = 1, 2, …, N commercial products. Note that the price of the raw material crude oil is expressed as a negative coefficient because it is a cost term.

Table 5.3. Attributes of the scenario construction for modelling price uncertainty for product *i*

	Price Uncertainty: Objective Function Coefficient of Prices ($/ton)			
	Scenario 1 ($s = 1$)	Scenario 2 ($s = 2$)	…	Scenario NS ($s = NS$)
Percentage of deviation from the expected value	$+\chi_1$%	$-\chi_2$%	…	$+\chi_{NS}$%
Price of product i in scenario s ($/ton) $c_{i,s}$	$c_{i,1}$	$c_{i,2}$	…	$c_{i,NS}$
Probability p_s	p_1	p_2	…	p_{NS}

Table 5.4. Representative scenarios of price uncertainty in the refinery planning under uncertainty problem

Material/ Product	Price Uncertainty: Objective Function Coefficient of Prices ($/ton)			
	Scenario 1	Scenario 2	…	Scenario NS
Product 1 (typically the raw material, crude oil)	$c_{1,1}$	$c_{1,2}$	…	$c_{1,NS}$
Product 2	$c_{2,1}$	$c_{2,2}$	…	$c_{2,NS}$
Product 3	$c_{3,1}$	$c_{3,2}$	…	$c_{3,NS}$
⋮	⋮	⋮	…	⋮
Product N	$c_{N,1}$	$c_{N,2}$	…	$c_{N,NS}$
Probability p_s	p_1	p_2	…	p_{NS}

5.5.6 Expectation of the Objective Function

As mentioned earlier, the classical mean–variance approach devised by Markowitz (1952, 1959) is adopted in an attempt to maximize profit by minimizing the variance for the given expected value (mean) of the objective function. To represent the different scenarios accounting for uncertainty in prices, the price-related random objective coefficients comprising (i) $\lambda_{i,t}$ for the raw materials costs of different types of crude oil that can be handled by the crude distillation unit of a refinery and (ii) $\gamma_{i,t}$ for the prices of saleable refined products are added with the index s to denote the scenarios, taking into account the probabilities of realization of each scenario. For ease of reference, both groups of price (or cost) parameters are redefined as the parameter $c_{i,s,t}$ or $c_{i's,t}$, in which the only minor difference between the two is in the use of the index i' and the corresponding set of I' to refer to "products" that are actually the raw materials crude oils as distinguished from the index i used to indicate saleable products.

For any constants *a* and *b*, using the identity for expectations gives the following:

$$E\left(aX \pm bY\right) = aE[X] \pm bE[Y] \qquad \textbf{(5.2)}$$

Since the objective function given by equation **(4.8)** is linear, it is straightforward to show that the expectation of the objective function with random price coefficients is given by the following:

$$\begin{aligned}
E\left[z_0\right] &= E\left[\text{Profit}\right] \\
&= E\left[\sum_{t\in T}\left(\begin{array}{l}\sum_{i\in I}\sum_{s\in S} p_s c_{i,s,t} S_{i,t} + \sum_{i\in I}\tilde{\gamma}_{i,t} I^{\mathrm{f}}_{i,t} - \sum_{i'\in I'}\sum_{s\in S} p_s c_{i',s,t} P_t - \sum_{i\in I}\tilde{\lambda}_{i,t} I^{\mathrm{s}}_{i,t} \\ -\sum_{j\in J} C_{j,t} x_{j,t} - \sum_{i\in I} h_{i,t} H_{i,t} - \sum_{j\in J}\left(\alpha_{j,t} CE_{j,t} + \beta_{j,t} y_{j,t}\right) - \left(r_t R_t + o_t O_t\right)\end{array}\right)\right] \\
&= E\left[\sum_{t\in T}\sum_{i\in I}\sum_{s\in S} p_s c_{i,s,t} S_{i,t}\right] + E\left[\sum_{t\in T}\sum_{i\in I}\tilde{\gamma}_{i,t} I^{\mathrm{f}}_{i,t}\right] - E\left[\sum_{t\in T}\sum_{i'\in I'}\sum_{s\in S} p_s c_{i',s,t} P_t\right] \\
&\quad - E\left[\sum_{t\in T}\sum_{i\in I}\tilde{\lambda}_{i,t} I^{\mathrm{s}}_{i,t}\right] - E\left[\sum_{t\in T}\sum_{i\in I} h_{i,t} H_{i,t}\right] - E\left[\sum_{t\in T}\sum_{j\in J}\left(\alpha_{j,t} CE_{j,t} + \beta_{j,t} y_{j,t}\right)\right] \\
&\quad - E\left[\sum_{t\in T}\left(r_t R_t + o_t O_t\right)\right] \\
&= \sum_{t\in T}\sum_{i\in I}\sum_{s\in S} p_s c_{i,s,t} S_{i,t} + \sum_{t\in T}\sum_{i\in I}\tilde{\gamma}_{i,t} I^{\mathrm{f}}_{i,t} - \sum_{t\in T}\sum_{i'\in I'}\sum_{s\in S} p_s c_{i',s,t} P_t - \sum_{t\in T}\sum_{i\in I}\tilde{\lambda}_{i,t} I^{\mathrm{s}}_{i,t} \\
&\quad -\sum_{t\in T}\sum_{i\in I} h_{i,t} H_{i,t} - \sum_{t\in T}\sum_{j\in J}\left(\alpha_{j,t} CE_{j,t} + \beta_{j,t} y_{j,t}\right) - \sum_{t\in T}\left(r_t R_t + o_t O_t\right)
\end{aligned}$$

$$E\left[z_0\right] = \sum_{t\in T}\left[\begin{array}{l}\sum_{i\in I}\sum_{s\in S} p_s c_{i,s,t} S_{i,t} + \sum_{i\in I}\tilde{\gamma}_{i,t} I^{\mathrm{f}}_{i,t} - \sum_{i'\in I'}\sum_{s\in S} p_s c_{i',s,t} P_t - \sum_{i\in I}\tilde{\lambda}_{i,t} I^{\mathrm{s}}_{i,t} - \sum_{i\in I} h_{i,t} H_{i,t} \\ -\sum_{j\in J}\left(\alpha_{j,t} CE_{j,t} + \beta_{j,t} y_{j,t}\right) + r_t R_t + o_t O_t\end{array}\right] \qquad \textbf{(5.3)}$$

It is pertinent to point out that consideration of the expected value of profit alone as the objective function, which is characteristic of stochastic linear programs, is obviously inappropriate for moderate and high-risk decisions under uncertainty since it is well acknowledged that most decision makers are risk averse for important decisions. As stressed by Mulvey et al. (1995), the expected value objective ignores both the risk attribute of the decision maker and the distribution of the objective values. Hence, variance of the objective function ought to be considered as a risk measure of the objective function, which is the second major component of the Markowitz's mean–variance approach adopted in this Risk Model I.

5.5.7 Variance of the Objective Function

For any constants *a* and *b*, using the identity for variances gives the following:

$$V\left(aX \pm bY\right) = a^2 V(x) + b^2 V(Y) \quad \textbf{(5.4)}$$

Noting that it is the coefficients of the objective function that are random and not the deterministic production mass variables x_i for each product type *i*, thus variance for the expected value of the objective function as shown in equation **(5.3)** is expressed as:

$$V\left(z_0\right) = V\left\{\sum_{t\in T}\left[\begin{array}{l}\sum_{i\in I}\gamma_{i,s,t}S_{i,t} + \sum_{i\in I}\tilde{\gamma}_{i,t}I_{i,t}^{\mathrm{f}} - \sum_{i\in I}\lambda_{i,s,t}P_{i,t} - \sum_{i\in I}\tilde{\lambda}_{i,t}I_{i,t}^{\mathrm{s}} - \sum_{j\in J}C_{j,t}x_{j,t} \\ -\sum_{i\in I}h_{i,t}H_{i,t} - \sum_{j\in J}\left(\alpha_{j,t}CE_{j,t} + \beta_{j,t}y_{j,t}\right) - \left(r_t R_t + o_t O_t\right)\end{array}\right]\right\} \quad \textbf{(5.5)}$$

Similar to the derivation of the expectation of the objective function, the random price coefficients are collectively redefined as $c_{i,s,t}$ to give the following relation for the variance of profit:

$$\begin{aligned} V\left(z_0\right) = {} & V\left[\sum_{t\in T}\sum_{i\in I}c_{i,s,t}S_{i,t}\right] + \cancelto{0}{V\left[\sum_{t\in T}\sum_{i\in I}\tilde{\gamma}_{i,t}I_{i,t}^{\mathrm{f}}\right]} + V\left[\sum_{t\in T}\sum_{i'\in I'}c_{i',s,t}P_{i',t}\right] \\ & + \cancelto{0}{V\left[\sum_{t\in T}\sum_{i\in I}\tilde{\lambda}_{i,t}I_{i,t}^{\mathrm{s}}\right]} + \cancelto{0}{V\left[\sum_{j\in J}C_{j,t}x_{j,t}\right]} + \cancelto{0}{V\left[\sum_{t\in T}\sum_{i\in I}h_{i,t}H_{i,t}\right]} \\ & + \cancelto{0}{V\left[\sum_{t\in T}\sum_{j\in J}\left(\alpha_{j,t}CE_{j,t} + \beta_{j,t}y_{j,t}\right)\right]} + \cancelto{0}{V\left[\sum_{t\in T}\left(r_t R_t + o_t O_t\right)\right]} \quad \textbf{(5.6)} \\ = {} & \sum_{t\in T}\sum_{i\in I}\left(S_{i,t}\right)^2 V\left(c_{i,s,t}\right) + \sum_{t\in T}\sum_{i'\in I'}\left(P_{i',t}\right)^2 V\left(c_{i',s,t}\right) \\ V\left(z_0\right) = {} & \sum_{t\in T}\sum_{i\in I}S_{i,t}^2 V\left(c_{i,s,t}\right) + \sum_{t\in T}\sum_{i'\in I'}P_{i',t}^2 V\left(c_{i',s,t}\right) \end{aligned}$$

Although the above derivation is mathematically and statistically sound, it does not explicitly evaluate variances of the random price coefficients as given by $V\left(c_{i,s,t}\right)$ and $V\left(c_{i',s,t}\right)$. We therefore consider an alternative formulation using the following definition for variance of *X* from Markowitz (1952) (but the definition should be easily available in any standard text on statistics):

$$V = p_1\left(x_1 - E\right)^2 + p_2\left(x_2 - E\right)^2 + p_3\left(x_3 - E\right)^2 + \cdots + p_n\left(x_n - E\right)^2 \qquad \textbf{(5.7)}$$

If we consider *NS* number of different scenarios in set *S*, then variation in profit is given by the probabilistically-weighted summation of the squared deviation of the objective function of a scenario from the expected value of the objective function, as depicted below:

$$\begin{aligned} V\left(z_0\right) &= \sum_{s \in S} p_s\left(z_{0,s} - E\left[z_0\right]\right)^2 \\ V\left(z_0\right) &= p_{s_1}\left(z_{s_1} - E\left[z_0\right]\right)^2 + p_{s_2}\left(z_{s_2} - E\left[z_0\right]\right)^2 + \cdots + p_{s_\mu}\left(z_{s_\mu} - E\left[z_0\right]\right)^2 + \cdots \\ &\quad + p_{s_{NS}}\left(z_{s_{NS}} - E\left[z_0\right]\right)^2 \end{aligned} \qquad \textbf{(5.6)}$$

where s_μ refers to the average scenario or the "most likely" scenario in which the coefficients take on the expected values, thus resulting in the expression $z_{s_\mu} - E\left[z_0\right]$ equals to zero and yielding:

$$\begin{aligned} V\left(z_0\right) &= \sum_{s \in S} p_s\left(z_{0,s} - E\left[z_0\right]\right)^2 \\ V\left(z_0\right) &= p_{s_1}\left(z_{s_1} - E\left[z_0\right]\right)^2 + p_{s_2}\left(z_{s_2} - E\left[z_0\right]\right)^2 + \cdots p_{s_\omega}\left(z_{s_\omega} - E\left[z_0\right]\right)^2 \end{aligned} \qquad \textbf{(5.7)}$$

5.5.8 Risk Model I

In the spirit of the Markowitz's mean–variance approach, the objective function for the stochastic model can now be formulated as:

$$\begin{aligned} &\text{maximize } z_1 = E[z_0] - \theta_1 V(z_0) \\ &\text{s.t.} \qquad \text{constraints } \mathbf{(4.1)} - \mathbf{(4.7)} \end{aligned} \qquad \mathbf{(5.8)}$$

that is, the model is subject to the same set of constraints as the deterministic model, with θ_1 as the risk parameter or risk factor associated with risk reduction for the expected profit (for convenience, θ_1 will henceforth simply be referred to as the profit risk factor).

For the reason of obtaining a term that is dimensionally consistent with the expected value term, the standard deviation of z_0 may be considered instead of the variance as a risk measure to reflect dispersions of the random objective function. In the case of variance, the difference in dimensionality is taken care of by the risk factor θ_1. This is concurred by Markowitz (1952) where it is stated that even though variance is the more well-known measure of dispersion about the expected, if instead of variance, an investor was concerned with standard error or with the coefficient of dispersion σ/E (also known as the coefficient of variation in more recent literature), the choice would still lie in the set of efficient portfolios. (As an aside, a very recent paper by Kristoffersen (2005) discusses a variety of risk measures in the two-stage stochastic linear programming approach.)

Therefore, the objective function considering standard deviation, which would require a different risk factor θ_1', is expressed as:

$$\text{maximize } z_1 = E[z_0] - \theta_1' \sqrt{V(z_0)} \qquad \mathbf{(5.9)}$$

Note that the solution convergence is expected to be different for models **(5.8)** and **(5.9)** due to the presence of the square root operation of variance in computing standard deviation, which would be expected to increase the computation time. In this regard, model **(5.8)** is preferred to model **(5.9)**.

However, the primary difficulty in executing both models **(5.8)** and **(5.9)** is in determining a suitable range of values for the profit risk factors θ_1 and θ_1', respectively that will cater to decision makers who are either risk-prone or risk-averse. An approach to overcome this is proposed, among others, by Terwiesch et al. (1994) and Ponnambalam (2005), in which the variance or the standard deviation of the objective function is minimized as follows:

$$\text{minimize } z_1 = V(z_0) \tag{5.10}$$

or

$$\text{minimize } z_1 = \sqrt{V(z_0)} \tag{5.11}$$

while adding the inequality constraint for the mean of the objective function (as given by equation **(5.1)**) that sets a certain target value for the desired profit to be achieved:

$$E[z_0] \geq \text{Target objective function value} \tag{5.12}$$

The process of maximizing or minimizing one objective while specifying constraints on another is a widespread practice in dealing with problems with two objectives (Eppen et al., 1989) such as in this model, in which the two objectives considered are profit and the risk associated with it.

In order to take advantage of the faster convergence rate in computing variance as compared to standard deviation, the objective function given by equation **(5.10)** is adopted in our model. Thus, in the model, the profit risk factor is now controlled by specifying the desired profit that will give the target objective function value, with the corresponding risk reflected by the variance expressed in the new objective function **(5.10)**. Mathematically, it is also noted that the target value constraint is needed;

otherwise, the optimization problem would simply compute the objective function value to be zero as that obviously corresponds to the minimum value of variance. Note that an equivalent representation of the objective functions **(5.10)** and **(5.11)** is given by the negative of the corresponding maximization problem, which thus gives Risk Model I:

$$\begin{aligned}
&\text{maximize } z_1 = -\text{Var}(z_0) \\
&\text{s.t.} \quad E\left[z_0\right] \geq \text{Target objective function value} \\
&\qquad\quad \text{constraints } \mathbf{(1)-(7)}
\end{aligned} \tag{5.13}$$

while the equivalent expression for the inferiorly preferred objective function due to an expected longer computation time is given by:

$$\text{maximize } z_1 = -\sqrt{\text{Var}(z_0)} \tag{5.14}$$

To determine the suitable range for the target objective function value (that is, the desired profit), a test value is assumed and the corresponding solution is computed. Then, the test value is increased or decreased, with the solution computed each time in order to investigate and establish the range of target objective function values that ensures solution feasibility. It is noted that the maximum target objective function value to maintain solution feasibility should be fairly well approximated by the optimal objective function value of the deterministic model.

As emphasized in the 1952 seminal paper by Markowitz, it is useful to keep in mind that the decision with maximum expected return or profit is not necessarily the one with minimum variance. There is a rate or trade-off at which an investor can gain expected profit by taking on variance or risk, or reduce risk by giving up expected profit. In essence, this trade-off will be demonstrated by the profit gained for different values of the profit risk factor θ_1 (or θ_1') specified.

As an additional note, the formulated model is alternatively known as a two-stage risk-based programming approach. The model is now complete and is solved to optimality based on a numerical example with the obtained results discussed, as to be found in the subsequent part of the paper where a representative example is presented. In the immediate following section on an alternative modelling approach, this Markowitz's mean–variance model as developed here, is implemented within a two-stage stochastic programming framework and is extended to become a recourse model.

5.6 APPROACH 2: THE EXPECTATION MODEL AS A COMBINATION OF THE MARKOWITZ'S MEAN–VARIANCE APPROACH AND THE TWO-STAGE STOCHASTIC PROGRAMMING WITH FIXED RECOURSE FRAMEWORK

In this approach, the existing deterministic model is reformulated in an attempt to minimize the expected value of the recourse penalty costs due to violations of constraints by incorporating the Markowitz's mean–variance model within a two-stage stochastic programming with fixed recourse framework. The model attempts to handle uncertainty in the random objective coefficients of prices, the random left-hand side (LHS) (technological) coefficients, and the random right-hand side (RHS) coefficients. One of the primary motivations for adopting the recourse problem model of minimizing only the expected penalty costs is to avoid the computationally more demanding nonlinear quadratic programming problem that arises with the simultaneous minimization of both the expected value and variance of the recourse penalty costs. (The latter, that is, the approach of minimizing both the expected value and the variance of the penalty costs, will be addressed in Approach 3 in the following section.) Additionally, as intended in the model development of Approach 1 in the preceding section, the aim for the inclusion of the Markowitz's mean–variance Model 1 is to account for randomness in the objective coefficients due to uncertainty arising in prices of crude oil and the commodities (which comprises) gasoline, naphtha, jet fuel, heating oil, fuel oil, and cracker feed.

5.6.1 Two-Stage Stochastic Programming with Fixed Recourse Framework to Model Randomness in the Right-Hand-Side (RHS) Coefficients of Product Demand Constraints

With the onslaught of an impending energy crisis due to rising crude oil prices, energy over-consumption, and depletion of hydrocarbon resources worldwide, the oil and gas industry has become increasingly competitive with market demand, essentially driven by customers consumption needs, imposing a significant complexity in production requirements. Jung et al. (2004) noted that among the many factors contributing to uncertainties in a typical example of the chemical process industry (of which the production planning of a petroleum refinery is a chief example), product demand uncertainty, as dictated by market demand, may well hold the dominant impact on profits and customer satisfaction. This is equally emphasized by a series of papers addressing chemical production planning under uncertainty by Gupta and Maranas (2000, 2003) that identified product demand as one of the key sources of uncertainties in the wider context of any production–distribution system. It is further noted that product demand fluctuations over medium-term (1–2 years) to long-term (5–10 years) planning horizons may be significant. Hence, deterministic planning and scheduling models may yield unrealistic results by not capturing the effect of demand variability on the trade-off between lost sales and inventory holding costs.

Demand uncertainty, leading to cause of failure in accounting for significant demand fluctuations by incorporating a stochastic description of product demand, can result in over- or under-production, with resultant excess inventories or unsatisfied customer demand, respectively. The latter could also result in incurred cost due to outsourcing or external purchasing for production make-up. Excess inventory incurs unnecessarily high inventory holding charges, while the consequences of the inability to meet customer needs eventually translates to both loss of profit and potentially, the long term loss of market share. These are highly undesirable scenarios particularly in current market settings where the profit margins are extremely tight. In terms of strategic corporate planning, the former scenario results as a failure in effectively managing the downside risk exposure of a company while the latter corresponds to a failure in recognizing an opportunity to capture additional market share (Gupta and Maranas, 2000; 2003).

Even though it has been increasingly common for firms to subcontract or outsource certain amount of production in order to meet increasing customer demand, in general, they are still viewed as an additional cost to firms, as mentioned earlier. The expected value of the ability to meet the product needs of customers is traditionally called the customer satisfaction index. Under competitive market conditions, customer satisfaction level, also known as service level, is recognized as an important index that must be monitored and maintained at a high level. Thus, deterministic planning models, which do not recognize the uncertainty in future demand forecasts, can be expected to result in inferior planning decisions as compared to stochastic models that explicitly considers uncertainty (Gupta et al., 2000).

5.6.1.1 Rationale for Adopting the Two-Stage Stochastic Programming Framework to Model Uncertainty in Product demand

From the arguments presented in the previous section, it is highly evident that in production systems, demand forecasts are often critical to the planning process. When demand is assumed to be known with certainty, an optimal deterministic production plan can easily be obtained, which in turn leads to an optimal capacity plan. But, as readily acknowledged, in reality, demand is rarely known with absolute certainty. Consequently, production planning decisions are usually postponed until better information is available. However, capacity plans cannot be postponed, and hence cannot rely on the production plan. Indeed, as demand varies from week to week, there may not be a unique production plan. Thus, the two-stage nature of the production planning process is apparent, as advocated by Higle and Sen (1996).

5.6.1.2 Sampling Methodology by Scenario Generation for the Recourse Model under Market Demand Uncertainty

Uncertainty in market demand introduces randomness in constraints for production demand, that is, the production requirements of intermediates and final saleable products as given by equation **(4.3)**. The sampling methodology employed for scenario generation for the recourse model under demand uncertainty is similar to the case of price uncertainty addressed in the previous section. Table 5.5 summarizes attributes of scenarios constructed for modelling demand uncertainty.

Considering that there are numerous minor variations within the many complex refining processes that cause random variations in the production of the major final saleable commercial products, it is not unreasonable to assume by virtue of the Central Limit Theorem that these cumulative minor random effects will be approximately normally distributed. With demand $d_{i,s}$ as the random variable where i denotes the product type and s indicates the corresponding scenario considered, the following relationship represents the random demand:

$$d_{is} = z\sigma_d + \mu_d \qquad \textbf{(5.15)}$$

where z is the variable for the standard normal distribution with mean 0 and variance 1 while μ_d and σ_d are the mean and variance for the distribution of demand. Similar to Approach 1, a collection of representative events or scenarios of market demand uncertainty for i = 1, 2, …, NS products with associated probabilities to indicate their comparative frequency of occurrence (Ermoliev and Wets, 1988) are depicted in Table 5.6.

Table 5.5. Attributes of the scenario construction for modelling market demand uncertainty for product *i*

	Demand Uncertainty: Right-Hand-Side Coefficient of Constraints for Product i Demand (ton/day) under Scenario s			
	Scenario 1 (s = 1)	Scenario 2 (s = 2)	…	Scenario NS
Percentage of deviation from the expected value	$+\delta_1$%	$-\delta_2$%	…	$+\delta_{NS}$%
Demand for product i (ton/day) $d_{i,s}$	$d_{i,1}$	$d_{i,2}$	…	$d_{i,NS}$
Probability p_s	p_1	p_2	…	p_{NS}

Table 5.6. Representative scenarios of market demand uncertainty in the refinery planning under uncertainty problem

Product type i	Demand Uncertainty: Right-Hand-Side Coefficient of Constraints for Product i Demand (ton/day) under Scenario s, $d_{i,s}$			
	Scenario 1	Scenario 2	$\cdots$	Scenario NS
Product 1	$d_{1,1}$	$d_{1,2}$	$\cdots$	$d_{1,NS}$
Product 2	$d_{2,1}$	$d_{2,2}$	$\cdots$	$d_{2,NS}$
Product 3	$d_{3,1}$	$d_{3,2}$	$\cdots$	$d_{3,NS}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
Product N	$d_{N,1}$	$d_{N,2}$	$\cdots$	$d_{N,NS}$
Probability p_s	p_1	p_2	$\cdots$	p_{NS}

5.6.1.3 Modelling Uncertainty in Product Demand by Slack Variables in the Stochastic Constraints and Penalty Functions in the Objective Function due to Production Shortfalls and Surpluses

As emphasized earlier, one of the main consequences of uncertainty within the context of decision-making is the possibility of infeasibility in the future. The two-stage recourse models provide the liberty of addressing this issue by postponing some decisions into the second stage; however, this comes at the expense of the use of corresponding penalties in the objective function, as reiterated by Sen and Higle (1999). Decisions that can be delayed until after information about the uncertain data is available almost definitely offer an opportunity to adjust or adapt to the new information received. Although it is acknowledged that it is typically beyond our control whether decisions can or cannot be delayed, there is generally value associated with delaying a decision, when it is possible to do so, until after additional information is obtained; this is advocated by Higle (2005).

In devising the appropriate penalty functions, we resort to the introduction of some compensating slack variables in the probabilistic constraints to eliminate the possibility of second-stage infeasibility. In addition to that, the recourse-based modelling philosophy requires the decision maker to impute a price as a penalty to remedial activities that are undertaken in response to the randomness. For some applications such as in production planning and inventory models, these costs are standard. (However, in some situations, it may be more appropriate to accept the possibility of infeasibility under some circumstances, provided the probability of this event is restricted below a given threshold) (Sen and Higle, 1999).

As pointed out by Clay and Grossmann (1997), compensating slack variables accounting for shortfall and/or surplus in production are introduced in stochastic constraints with the following results: (i) inequality constraints are replaced with equality constraints; (ii) numerical feasibility of the stochastic constraints can be ensured for all events; and (iii) penalties for feasibility violations can be added to the objective function. Since a probability can be assigned to each realization of the stochastic parameter vector (that is, to each scenario), the probability of feasible operation can be measured. Further according to Clay and Grossmann (1997), assigning penalties to the feasibility slack activities in the objective function is similar to the "discrepancy cost" approach suggested by Dempster (1980). Using Dempster's approach, one assigns a cost to the violation of any of the constraint conditions. In the production planning context, one example would be to add a slack variable for producing less than the minimum demand for a product, and then penalizing this slack based on the cost of purchasing this makeup product from the outside market; likewise, for the condition of surplus production with respect to the market demand, the slack variable is penalized based on the inventory cost for holding or storing the excess of production.

In other words, the stochastic nature of a production requirement constraint is handled accordingly by noting that there is an added cost associated with infeasibility of any stochastic constraint, as equally noted, among others, by Evers (1967) and Wets (1983). In addition, infeasibility requires appropriate action to be taken, hence, giving rise to the notion of recourse and the subsequent construction of the desired recourse problems or models. Thus, the principle that applies is that infeasibility due to violation(s) of the constraints will be acceptable, but this is penalized through the introduction of slack variables modelled as the expected shortfalls (or shortages) or surpluses in production. The penalty for infeasibility is included in the objective function as a result of uncertainty in market demand leading to randomness in production requirements. The penalty terms in the objective function is handled by maximizing the expected profit while minimizing the expected value or mean of the recourse penalty costs. (Later in Approach 3, the variance of the expected recourse penalty costs is minimized as well to reflect the level of risk undertaken by a decision maker).

The use of penalty functions in the objective function for stochastic models was pioneered by Evers (1967) as a technique of accounting for losses due to infeasibility. In our context of production planning of chemical plants (such as an oil refinery) under the exogenous uncertainty of product demand, a penalty term in the objective function is employed to quantify the effect of missed revenues and loss of customer confidence. It typically assumes the following form:

$$-\gamma\sum_{i=1}^{N} P_i \max\left[0,\left(\theta_i - Q_i\right)\right] \qquad \textbf{(5.16)}$$

where γ is the penalty coefficient whose value determines the relative weight attributed to production shortfalls as a fraction of the profit margins (Wellons and Reklaitis, 1989; Birewar and Grossmann, 1990; Ierapetritou and Pistikopoulos, 1996; Petkov and Maranas, 1998).

Based on the concepts presented, the penalty coefficients in the stochastic refinery planning model are supposed to be proportional to the respective shortfalls or surpluses in products. These penalties are interpreted and assumed accordingly, per unit of undeliverable or overproduced products, as follows:

c_i^+: the fixed penalty cost paid per unit of demand $d_{i,s}$ that cannot be delivered or satisfied by production and thus, is considered as cost of lost demand, or if it is to be obtained from other sources, then it is the cost of purchasing in the open market to meet the shortfall in unsatisfied production requirement demand;

c_i^- : the fixed penalty cost paid per unit of the products produced in excess of $d_{i,s}$ and is typically the cost of inventory to store the production surplus that exceeds demand.

It is noted that inventory cost should always be lower than the cost of purchasing a commodity in the open market as otherwise, it would be comparatively more economical for the refinery to outsource their production, thus defeating the purpose of setting up an inventory system.

As pointed out by Kall and Wallace (1994), the penalty costs incurred due to violations in the constraints are actually determined after the observation of the random data; hence, they are recourse costs that are imposed on the second-stage variables. Based on statistical concepts, in a case involving repeated execution of the production as it is in the operations of a refinery, it becomes appropriate to apply an expected value criterion. More precisely, the objective of the model is to maximize the sum of the original first-stage profit from sales terms while minimizing the expected recourse costs of the second-stage variables. To accomplish this, Risk Model I developed in Approach 1 based on the Markowitz's MV approach is reformulated to incorporate the penalty terms. Accordingly, the following non-negative second-stage recourse slack variables are introduced and defined as follows:

$z_{i,s}^{+}$: the amount of unsatisfied demand of product i due to shortfall in supply or underproduction (shortages in production) per realization of scenario s;

$z_{i,s}^{-}$: the amount of extra product i produced due to surplus in supply or overproduction (excesses in production) per realization of scenario s;

where $z_{i,s}^{+} \equiv \max(0, z)$ is the positive part of z while $z_{i,s}^{-} \equiv \max(0, -z)$ is the negative part of z. Thus, the expected recourse penalty for the second-stage costs due to uncertainty in product i demand for all considered scenarios generated is given by:

$$E_{s,\text{demand}} = \sum_{i \in I} \sum_{s \in S} p_s \left(c_i^{+} z_{i,s}^{+} + c_i^{-} z_{i,s}^{-} \right) \tag{5.17}$$

To ensure that the original information structure associated with the decision process sequence is honoured, for each of the products whose demand is uncertain, s new constraints to model the s number of scenarios generated for each product are added to the stochastic model in place of the original deterministic constraint. Herein lies a demonstration of the fact that the size of a recourse model formulated to handle uncertainty increases exponentially since the total number of scenarios grows exponentially with the number of random parameters. This step ensures compliance with the notion that while some decisions (as presented by the associated decision variables) can respond to a specific scenario; other decisions represented by other constraints cannot do so (Higle, 2005). In general, the new constraints are expressed as:

$$x_i + z_{i,s}^{+} - z_{i,s}^{-} = d_{i,s}, \quad i \in I, s \in S \tag{5.18}$$

or in graphical representation:

Deterministic model constraint	Stochastic model constraints	
	$x_i + z_{i1}^+ - z_{i1}^- = d_{i1}$	Scenario 1
$x_i \le d_i$	$x_i + z_{i2}^+ - z_{i2}^- = d_{i2}$	Scenario 2
	$\vdots$	$\vdots$
	$x_i + z_{i,s}^+ - z_{i,s}^- = d_{i,s}$	Scenario s

Figure 5.6. Graphical representation of the transformation of a deterministic model's constraints into a correspondingly formulated stochastic model's constraints that capture its possible scenarios

5.6.1.4 The surplus penalty c_i^+ and the shortfall penalty c_i^-

In general, the $z_{i,s}^+$ and $z_{i,s}^-$ recourse variables are used in stochastic linear programming with simple recourse framework to obtain equality for all stochastic constraints. These variables appear only once in the formulation and must be dealt with in a manner that recognizes their economic consequences (Wets, 1983). As highlighted earlier, the role of the recourse variables in the formulation is to obtain feasibility for the various possible realizations of the stochastic demand constraints. Any deviation from the equality of the stochastic constraints will raise a net cost. Therefore, the surplus stock $z_{i,s}^-$ that exists at the end of a particular period is actually a liability since the cost of producing or obtaining the surplus cannot be recovered. This excess production is dealt with by assuming that the firm can realize a salvage value for it at the end of each period or is simply sent to inventory. In the case of salvaging, the company will realize a cash inflow from the sale of the surplus and the penalty will be negative since the total cost is reduced (the cost of producing surplus has already been accounted for in the first-stage production amount variables). The salvage value must be less than the cost of production or subcontracting (or outsourcing) since otherwise, the refinery will always have an incentive to overproduce, secure in the knowledge that the variable cost of production can be later recaptured.

When the refinery underproduces, it must scramble to fill excess demand and incurs a cost approximately equal to the cost of producing the required stock through subcontracting/outsourcing or overtime labour or by purchasing the product from the open market. If the refinery is operating at a level that is close to capacity (or if it is not close to capacity but has to incur an additional setup cost), this will usually result in the firm incurring costs that exceed normal production costs. The size of the penalty will then depend on the cost of the least expensive and available alternative production route or method (Kira et al., 1997).

5.6.1.5 Limitations of Approach 2

It is well-acknowledged that the scenario approach adopted in formulating recourse models results in an exponential growth in the problem size as the number of scenarios increases exponentially with the number of uncertain parameters to be modelled. Additionally, employing penalty functions in modelling violations of constraints with random parameters is also large restricted in that many new non-negative slack variables $z_{i,s}^{+}$ and $z_{i,s}^{-}$, accounting for the constraints' violations, must be added.

5.6.1.6 A Note on a More General Penalty Function for Production Shortfalls

Birewar and Grossmann (1990) pointed out that the penalty function described simply as the product of the cost of production shortfall and the shortfall quantity ($\text{penalty}_{i,t}$ = cost $\Omega_{it} \times \text{shortfall}_{i,t}$, $i = 1, 2, \ldots, N_{\text{p}}$, $t = 1, 2, \ldots, T$) may not be able to adequately represent loss of consumer satisfaction due to the shortfalls in orders booked. For example, it is more realistic to assume that the degree of consumer dissatisfaction will increase as the percentage shortfall increases. In other words, the constant of proportionality would rise as the percentage shortfall increases. For example, if the shortfall SF_{it} for product i in interval t is less than $\sigma_{i,t}^{1}$, then the penalty is proportional to the shortfall and the constant of proportionality or the penalty constant is $\Omega^1_{i,t}$. For a shortfall between $\sigma^1_{i,t}$ and $\sigma^2_{i,t}$, the penalty constant is given by $\Omega^2_{i,t}$. For a shortfall greater than $\sigma^2_{i,t}$, the penalty constant is given by $\Omega^3_{i,t}$.

The total penalty for each product i in interval t is defined by the following three groups of constraints:

$$\mathrm{PN}_{i,t} \geq \mathrm{SF}_{i,t}\Omega^1_{i,t}, \quad i = 1,2,\ldots,N_\mathrm{p},\; t = 1,2,\ldots,T \qquad \textbf{(5.19)}$$

$$\mathrm{PN}_{i,t} \geq \sigma^1_{i,t}\Omega^1_{i,t} + \left(\mathrm{SF}_{i,t} - \sigma^1_{it}\right)\Omega^2_{i,t}, \quad i = 1,2,\ldots,N_\mathrm{p},\; t = 1,2,\ldots,T \qquad \textbf{(5.20)}$$

$$\mathrm{PN}_{i,t} \geq \sigma^1_{i,t}\Omega^1_{i,t} + \sigma^2_{i,t}\Omega^2_{i,t} + \left(\mathrm{SF}_{i,t} - \sigma^2_{i,t}\right)\Omega^3_{i,t}, \quad i = 1,2,\ldots,N_\mathrm{p},\; t = 1,2,\ldots,T \qquad \textbf{(5.21)}$$

provided

$$\Omega^1_{i,t} \leq \Omega^2_{i,t} \leq \Omega^3_{i,t}, \quad i = 1,2,\ldots,N_\mathrm{p},\; t = 1,2,\ldots,T \qquad \textbf{(5.22)}$$

Similarly, any such group linear constraints can be used to replace the penalty constraints in the model(s). This approach will be considered in future work.

5.6.2 Two-Stage Stochastic Programming with Fixed Recourse Framework to Model Randomness in the Left-Hand-Side (LHS) or Technology/Technological Coefficients of Product Yield Constraints

5.6.2.1 Uncertainty in Product Yields

The different types of petroleum crude oil and their associated values are defined, identified, and distinguished according to their yields structure or pattern besides the qualities of their refined useful products (OSHA Technical Manual, http://www.osha.gov/dts/osta/otm/otm_iv/otm_iv_2.html, accessed on September 30, 2005). Thus, different types of crude oil would lead to varying degrees of product yields. The yield pattern is dependent upon complex interaction of feed characteristics and reactor conditions that determine severity of operation (Gary and Handwerk, 1994; http://jechura.com/ChEN409/, accessed on October 17, 2005).

To determine the value of crude oil, comprehensive compilations of laboratory and pilot plant data that define the properties of the specific crude oil are undertaken. These data are termed as crude assays. A similar method is also used to establish the processing parameters of a particular crude oil. At a minimum, the assay should contain a true boiling point curve and a specific gravity curve for the crude oil. Most crude oil suppliers, however, extend the scope of the assay to include sulphur contents, viscosity, pour points, and many other properties (Jones, 1995; Gary and Handwerk, 1994; Speight, 1998).

In the literature, the term product yields (as used in Li (2004), for example) is also variably referred to as production yields (Pongsakdi et al., in press; Lababidi et al., 2004), processing yields (Fisher and Zellhart, 1995), or process yields (Bassett et al., 1997). Estimating the yields of the desired fractions that might be obtained from a single crude is a fairly simple task. However, the refiner is rarely processing a single crude but a mixture of a number of crudes. Assays are usually available for single crudes, but only for very few blends, and these are unlikely to be the ones of interest. Performing a complete assay of a crude is therefore an expensive and time consuming procedure. The blend being charged to the crude distillation unit, typically the first processing unit encountered in refining processes, could change significantly before an assay could be completed. The refiner, therefore, must have some other means of estimating the amounts of the various streams that he should gain (that is, the product yields) from his current blend of crude oils. Fortunately, computer programs are available that can take crude assay data and derive from them, a complex of pseudo-hydrocarbon components that will satisfactorily represent the actual crude. For a complete treatment of this subject, the reader is referred to the text by Maples (1993).

Liou et al. (1989) reported that uncertainties in yield of reactors, or more precisely, product yield of chemical reactions that take place in reactors, are mainly due to the different impact of two factors, namely (1) nonideal flow and/or (2) catalytic pellet transport limitations, on the performance of the laboratory or pilot plant reactors in the scale-up procedure to commercial reactors. This problem is even more prominent and demands increased attention when a large number of chemical reactions occur simultaneously, as in the multiple unit operations that make up a petroleum refinery.

The factor of nonideal flow refers to deviation from ideal plug flow modelling used in pilot plant studies during the scale-up procedure to industrial-standard reactors. The yield in a relatively shorter laboratory reactor is affected by axial dispersion more than that of a full-scale industrial reactor operating at the same residence time distribution. An *a priori* estimate of the maximum deviation in the yield of a laboratory reactor from that of a full-scale reactor is important in estimating the uncertainty involved in the scale-up practice. This information is vital for predicting the minimum length of a laboratory reactor for specified operating conditions so that the yield of a desired product will not deviate from that of a commercial reactor by more than some specified value.

The second factor of transport limitations in catalytic pellets affects the yield of a desired product when many isothermal reactions occur simultaneously. The yield in a laboratory reactor is typically higher than that of an industrial unit as the catalytic pellets used are usually smaller. Therefore, it is pertinent to model in order to predict the maximum uncertainty that may be introduced by this difference. It is reported that research has been carried out to determine the effect of diffusion on the local yield of a desired product for a system with an arbitrary number of first-order isothermal reactions in which the involved prediction requires knowledge of all rate constants.

As an example, the analysis and scale-up of a laboratory packed bed reactor is often complicated by the presence of intraparticle diffusion resistance, which results in the intrinsic kinetics becoming unclear. The difference in particle size and consequently, in intraparticle diffusion, introduces an uncertainty in the scale-up procedure. Therefore, it is crucial to be able to predict *a priori*, the maximum impact of both intraparticle diffusion and axial dispersion on reactors with simultaneous multiple reactions (Liou et al., 1989).

On the other hand, at the operational level, the phenomenon of catalyst deactivation may affect the yield of reactions, thereby introducing significant uncertainties in modelling product yields. By definition, any process, chemical or physical, which decreases the intrinsic activity of a catalyst, can be classified as deactivation. In many cases involving more complex catalysts or reactions, deactivation can be accompanied by large changes in selectivity as well and hence, lead to uncertainty in related products' yields and conversion (Petersen and Bell, 1987). In general, catalyst deactivation can be caused by coke formation; contamination of the active sites, agglomeration, and poisoning of the catalyst. As an illustration, coke formation or coking, widely experienced in the catalysis of hydrocarbon conversions, can deactivate both metallic and acid catalytic active sites for hydrocarbon reactions. Accumulation of such carbonaceous deposits affects selectivity in hydrocarbon conversions, thus resulting in uncertainty in product yields (Sermon et al., 1996).

In addition, from the modelling point-of-view, Yang et al. (1996) remarked that product yields for many reactions of organic species or compounds are known to be relatively uncertain or random due to lack of data or to the lumping procedures used to condense mechanisms of the reactions. For instance, feedstocks of fluidized catalytic crackers, the major refinery unit operation for gasoline production, consist of thousands of components, thus rendering the estimation of intrinsic kinetic constants to be very difficult. Therefore, the lumping of components according to the boiling point range is generally accepted, although as noted, this is bound to result in uncertainty of the product yields (Alvarez-Ramirez, 2004).

Nevertheless, Fisher and Zellhart (1995) cautioned that planners and users of a planning model must recognize that the developed planning model and the entailing plan is a forecast for an uncertain future. Therefore, an excessive amount of time should not be spent in trying to estimate product yields that are accurate to a very high degree. Although product yields should always be as accurate as possible, it is not to the extent that the curse of "paralysis of analysis" sets in due to concern that the yields are not perfect. In general, it is acknowledged that yields within one percent of the actual (correct) value are acceptable. Yields for each refinery process should be inspected for mass balances, hydrogen balances, and balances for other significant materials such as sulphur, where and when applicable.

There is relatively few works addressing uncertainty in product yields of chemical reactions in the process systems engineering (PSE) literature pertaining to production planning. The recent work of Pongsakdi et al. (in press) mentions accounting for this factor of uncertainty; however, neither an explicit representation in the stochastic model formulation nor a detailed discussion of yield uncertainty is presented. Uncertainty in yield from a process, termed as productivities, appears to be an active work in the Grossmann and coworkers research group at Carnegie Mellon University, with Goel (2004) classifying yield in a process network problem under uncertainty as an endogenous parameter, an under-treated issue in uncertainty modelling in which the implication is that the structure of the scenario tree generated is dependent of when decisions are made. This work has been extended by Tarhan and Grossmann (2005).

5.6.2.2 Product Yields for Petroleum Refining Processes

As product yields from a process feature as one of the primary sources of uncertainty in the midterm planning of a refinery, it is deemed worthwhile to attempt to identify the factors that influence the outcome of the yield pattern of a specific refining process, and which directly (or possibly indirectly) contribute to uncertainty in product yields for the particular process. This information is summarized in Table 5.7.

5.6.2.3 Product Yields from the Crude Distillation Unit (CDU)

The crude distillation unit (CDU) is the primary unit operation for the initial fractionation of crude oil. A representation of the true boiling point (TBP) distillation curve for the CDU is obtained by associating the temperature scale and the distilled percentages as shown in Figure 5.7. In this diagram, each rectangular area represents the yield of the different cuts from the crude. With this representation, it is possible to situate the types of petroleum cuts and the corresponding temperature limits or cut points as that obtained in refineries. An alternative representation of boiling temperature against cumulative volume percentage, which is perhaps more commonly encountered, is shown in Figure 5.8 for the Arabian Light crude oil.

Table 5.7: Factors influencing the yield pattern of processes in petroleum refining (Gary and Handwerk, 1994)

Refinery Process	Factors influencing Yield
Crude distillation	• Crude oil type • Feed characteristics • Reactor conditions/Operating severity, e.g., temperature and pressure
Coking	• Temperature (high) • Pressure (low) • Feed characteristics especially carbon residue
Catalytic cracking	Catalyst type
Catalytic hydrocracking	• Crude oil type • Previous processing operations • Catalyst type and activity • Operating conditions, e.g., temperature and pressure
Catalytic reforming	• Reactor pressure • Catalyst type and activity • Feed quality
Isomerization	• Feed properties • Operating severity
Alkylation	• Isobutane/Olefin ratio • Temperature

In current industrial practice of refinery planning and optimization using linear programming, the CDU is modelled based on the stream TBP cut point scheme using the technique of swing cut modelling. (Sahdev et al., 2004, www.cheresources.com/refinery_planning_optimization.shtml, accessed February 22, 2006). Swing cuts are increasingly used in linear programming assay tables to represent the refinery's flexibility to alter cut-points for optimization of side-cut yields and properties. (Tucker, Michael A, www.kbcat.com/pdfs/tech/tp_002.pdf, accessed February 22, 2006).

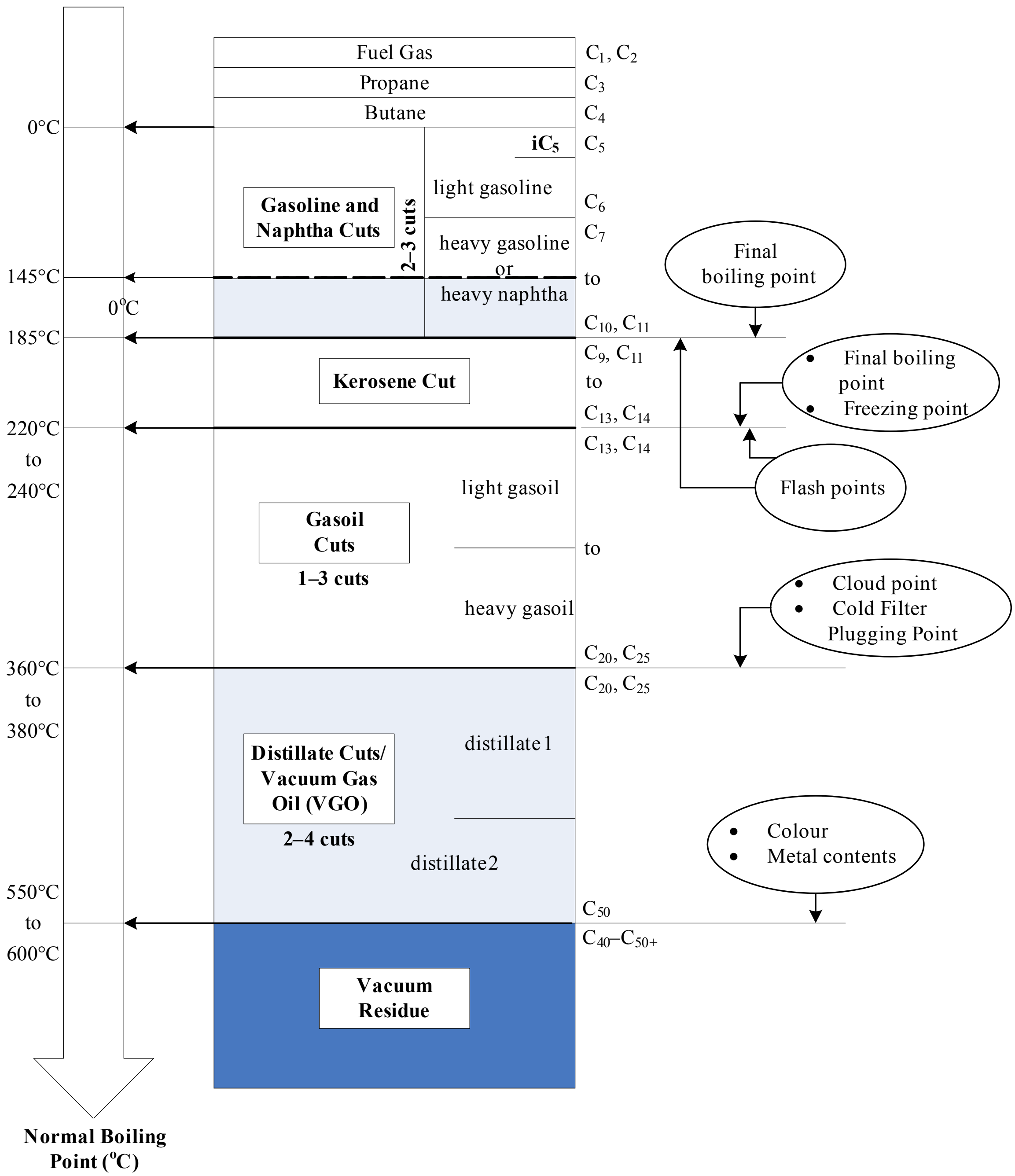


Figure 5.7. Representation of the true boiling point (TBP) distillation curve for the crude distillation unit (CDU) (taken from ENSPM Formation Industrie, 1993)

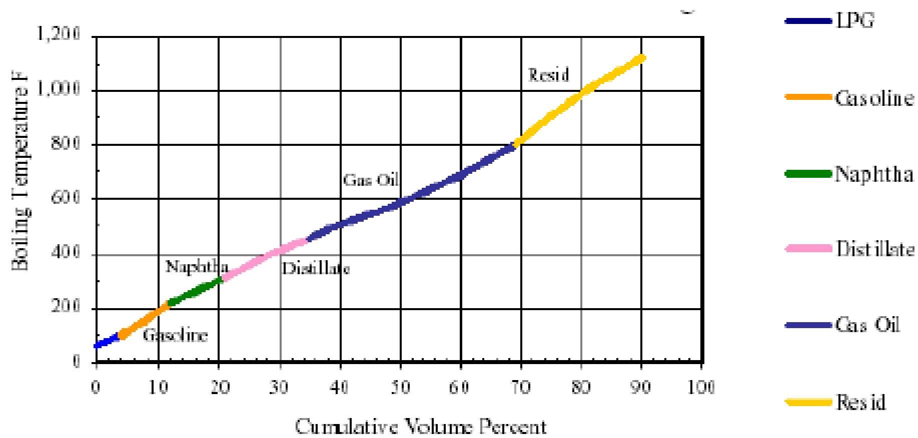


Figure 5.8. Fractions from the distillation curve for the Arabian Light crude oil (taken from Jechura, http://jechura.com/ChEN409/, accessed on October 17, 2005)

5.6.2.4 Sampling Methodology by Scenario Generation for the Recourse Model under Product Yields Uncertainty

Uncertainty in product yields introduces randomness in the mass balances as given by equation **(4.7)**. The sampling methodology employed for scenario generation for the recourse model under product yields uncertainty is similar to the case of demand uncertainty addressed in the previous section. Table 5.8 summarizes attributes of scenarios constructed for modelling product yields uncertainty whereas Table 5.9 presents the scenario construction to model yield uncertainty of products $k = 1, 2, 3, \ldots, N_c$ from material i. Note that in order to ensure that the material balances are satisfied, the summation of yields must equal to unity. Therefore, if there are N_c number of products with randomness in yield, then the yield for the N_cth product considered is computed as the difference of the summation of yields for the (N_c–1) products subtracted from 1, that is, $y_{i,N_c,s} = 1 - \sum_{k=1}^{N_c} y_{i,k,s}$. In most situations pertaining to chemical processes planning, doing so would not distort the physics of the problem as usually, there is provision to account for yield losses. In the case of petroleum refining, the N_cth product usually refers to a product at the "bottom of the barrel" of a certain processing unit, which possesses relatively insignificant commercial value compared to the yields of the other products produced by the same unit.

Table 5.8. Attributes of the scenario construction for modelling product yields uncertainty from material *i*

	Yield Uncertainty: Left-Hand Side Coefficient of Mass Balance for Product k Yield from Material i (unitless) under Scenario s			
	Scenario 1 (s = 1)	Scenario 2 (s = 2)	$\cdots$	Scenario NS
Percentage of deviation from the expected value	$+\psi_1\%$	$-\psi_2\%$	$\cdots$	$+\psi_{NS}\%$
Yield of product k from material i (ton/day) $y_{i,k,s}$	$y_{i,k,1}$	$y_{i,k,2}$	$\cdots$	$y_{i,k,NS}$
Probability p_s	p_1	p_2	$\cdots$	p_{NS}

Table 5.9. Representative scenarios of product yields uncertainty in the refinery planning under uncertainty problem

	Yield Uncertainty: Left-Hand Side Coefficient of Mass Balance for Product k Yield from Material i (unitless) under Scenario s, $y_{i,k,NS}$			
Product type j	Scenario 1	Scenario 2	$\cdots$	Scenario s
Product 1	$y_{i,1,1}$	$y_{i,1,2}$	$\cdots$	$y_{i,1,NS}$
Product 2	$y_{i,2,1}$	$y_{i,2,2}$	$\cdots$	$y_{i,2,NS}$
Product 3	$y_{i,3,1}$	$y_{i,3,2}$	$\cdots$	$y_{i,2,NS}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
Product N_c	$1-\sum_{k=1}^{N_c} y_{i,k,1}$	$1-\sum_{k=1}^{N_c} y_{i,k,2}$	$\cdots$	$1-\sum_{k=1}^{N_c} y_{i,k,NS}$
Probability p_s	p_1	p_2	$\cdots$	p_{NS}

5.6.2.5 Modelling Uncertainty in Product Yields by Slack Variables and Penalty Functions for Shortages and Excesses in Yields

To be consistent with the definitions of the variables accounting for production shortfalls ($z_{i,s}^{+}$) and surpluses ($z_{i,s}^{-}$) in addressing demand uncertainty, variables to denote the deviation from the expected value (mean) of the yield of product j from material i are defined as follows. A positive deviation refers to a shortage in product yield; conversely, a negative deviation denotes an excess in product yield. Therefore, the variables are properly defined as below:

$y_{i,k,s}^{+}$: the amount of shortage in yields from material i (from the expected value) for product type k per realization of scenario s,

$y^{-}_{i,k,s}$: the amount of excess in yields from material i (from the expected value) for product type k per realization of scenario s,

The amount of shortage of yields from crude oil can be equivalently interpreted as the amount of material to be added to compensate for the associated shortage in product k; in contrary, the amount of excess of yields indicates the amount of material to be reduced to account for the associated surplus in product k.

Based on the similar approach of adopting penalty functions for modelling demand uncertainty, it is assumed that a fixed penalty cost of $q^{+}_{i,k}$ is incurred per unit of $y^{+}_{i,k,s}$ amount of shortage of yield from material i, and a fixed penalty cost of $q^{-}_{i,k}$ for per unit of excess of yield from material i by the amount $y^{-}_{i,k,s}$. Thus, the expected recourse penalty for the second-stage costs due to uncertainty in yield of product k from material i for all considered scenarios generated is given by:

$$E_{s,\mathrm{yield}} = \sum_{i\in I}\sum_{s\in S} p_s \left(q^{+}_{i,j} y^{+}_{i,k,s} + q^{-}_{i,j} y^{-}_{i,k,s} \right) \qquad \textbf{(5.23)}$$

As in the case of product demand uncertainty, to ensure that the original information structure associated with the decision process sequence is honoured, N_s new constraints (in place of the original single deterministic fixed yield constraint) to account for the N_s number of scenarios dealing with product yield uncertainty are introduced for *each* product whose yield is uncertain (Higle, 2005). The general form of the new constraints is:

$$T_i x_1 + x_i + y^{+}_{i,k,s} - y^{-}_{i,k,s} = 0, \quad \forall i \in I, \forall k \in K, \forall s \in S \qquad \textbf{(5.24)}$$

To obtain realistic values for the yield deviation terms, upper bounds of five (5) percent of the crude oil mass flowrate are imposed as an estimate of the maximum value that these terms could sensibly assume.

The combination of Tables 5.4, 5.6, and 5.9 as given by Table 5.10 completes the scenario formulation in order to *simultaneously* model uncertainties in commodity prices, product demand, and product yields as represented by randomness in the coefficients of the objective function and the RHS and the LHS of the constraints, respectively. The major assumption that prices, demand, and yields in each scenario are highly correlated enables the combination of all the scenarios in a stochastic programming model with discrete random variables. Otherwise, the computation would involve the construction of joint probability distribution of random variables that are made up of scenarios depicting all possible combinations of the three parameters of prices, demand, and yields. This will be considered in future work.

The corresponding expected recourse penalty for the second-stage costs due to uncertainties in both demand and yields is given by:

$$
\begin{aligned}
E_s &= E_{s,\text{demand}} + E_{s,\text{yield}} \\
&= \sum_{i \in I} \sum_{s \in S} p_s \left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^- \right) + \sum_{i \in I} \sum_{s \in S} p_s \left(q_i^+ y_{i,k,s}^+ + q_i^- y_{i,k,s}^- \right) \\
&= \sum_{i \in I} \sum_{s \in S} p_s \left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^- \right) + p_s \left(q_i^+ y_{i,k,s}^+ + q_i^- y_{i,k,s}^- \right) \\
E_s &= \sum_{i \in I} \sum_{s \in S} p_s \left[\left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^- \right) + \left(q_i^+ y_{i,k,s}^+ + q_i^- y_{i,k,s}^- \right) \right] \\
&= \sum_{i \in I} \sum_{s \in S} p_s \xi_s \quad \text{where } \xi_{i,s} = \left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^- \right) + \left(q_i^+ y_{i,k,s}^+ + q_i^- y_{i,k,s}^- \right) \\
E_s &= \sum_{i \in I} \sum_{s \in S} E_{i,s}
\end{aligned}
\qquad \textbf{(5.25)}
$$

The expression for $E_{s'}$ as given by equation **(5.25)** clearly formalizes and illustrates one of the most profound concepts of the recourse model, that is, the sole random variable is now redefined to be the scenarios and no longer the separate random variables of coefficients denoting prices, demand, and yields as considered earlier. Mathematically, single random vectors of the recourse variables $\tilde{\xi} = \left(z_{i,s}^+, z_{i,s}^-, y_{i,k,s}^+, y_{i,k,s}^- \right)$ are used in place of the four single random variables $z_{i,s}^+$, $z_{i,s}^-$, $y_{i,k,s}^+$, and $y_{i,k,s}^-$, in which vectors $\tilde{\xi}$ are random variables themselves (denoted here by the wavy line above the symbol). $\tilde{\xi}$ is described by a finite discrete distribution of $\left\{ (\xi_s, p_s), s = 1,2,3,\dots,s \middle| p_s > 0 \; \forall s \right\}$ as depicted by the discrete probabilities in Table 5.10 (Kall and Wallace, 1994).

Table 5.10. Complete scenario formulation for the refinery production planning under uncertainty in commodity prices, market demand for products, and product yields problem

Material/Product	Scenario 1	Scenario 2	$\cdots$	Scenario NS
Price Uncertainty: Objective Function Coefficient of Prices ($/ton)				
Product 1	$c_{1,1}$	$c_{1,2}$	$\cdots$	$c_{1,NS}$
Product 2	$c_{2,1}$	$c_{2,2}$	$\cdots$	$c_{2,NS}$
Product 3	$c_{3,1}$	$c_{3,2}$	$\cdots$	$c_{3,NS}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
Product N	$c_{N,1}$	$c_{N,2}$	$\cdots$	$c_{N,NS}$
Demand Uncertainty: Right-Hand-Side Coefficient of Constraints for Product i Demand (ton/day) under Scenario s, $d_{i,s}$				
Product 1	$d_{1,1}$	$d_{1,2}$	$\cdots$	$d_{1,NS}$
Product 2	$d_{2,1}$	$d_{2,2}$	$\cdots$	$d_{2,NS}$
Product 3	$d_{3,1}$	$d_{3,2}$	$\cdots$	$d_{3,NS}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
Product N	$d_{N,1}$	$d_{N,2}$	$\cdots$	$d_{N,NS}$
Yield Uncertainty: Left-Hand Side Coefficient of Mass Balance for Product j Yield from Material i (unitless) under Scenario s, $y_{i,k,s}$				
Product 1	$y_{i,1,1}$	$y_{i,1,2}$	$\cdots$	$y_{i,1,NS}$
Product 2	$y_{i,2,1}$	$y_{i,2,2}$	$\cdots$	$y_{i,2,NS}$
Product 3	$y_{i,3,1}$	$y_{i,3,2}$	$\cdots$	$y_{i,2,NS}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
Product N_c	$1-\sum_{k=1}^{N_c} y_{i,k,1}$	$1-\sum_{k=1}^{N_c} y_{i,k,2}$	$\cdots$	$1-\sum_{k=1}^{N_c} y_{i,k,NS}$
Probability p_s	p_1	p_2	$\cdots$	p_{NS}

5.6.2.6 Expectation Model I

A new reformulated objective function z_2 is now proposed, consisting of the sum of the following components: (1) maximization of the expected net profit from product sales subtracting the raw material costs of purchasing crude oil and the operating costs; (2) minimization of the sum of variance in profit; and (3) minimization of the sum of expected recourse penalty costs due to shortfalls or surpluses in production and shortages or excesses of product yields from certain materials. The mathematical expression for this new objective function is presented as:

$$\text{maximize } z_2 = z_1 - E_{s'} = E[z_0] - \theta_1 V(z_0) - E_{s'} \quad \textbf{(5.26)}$$

where

$$E[z_0] = \sum_{t \in T} \left[\begin{array}{l} \sum_{i \in I} \sum_{s \in S} p_s \gamma_{i,s,t} S_{i,t} + \sum_{i \in I} \tilde{\gamma}_{i,t} I^{\mathrm{f}}_{i,t} - \sum_{i \in I} \sum_{s \in S} p_s \lambda_{s,t} P_{i,t} \\ - \sum_{i \in I} \tilde{\lambda}_{i,t} I^{\mathrm{s}}_{i,t} - \sum_{i \in I} h_{i,t} H_{i,t} - \sum_{j \in J} \left(\alpha_{j,t} CE_{j,t} + \beta_{j,t} y_{j,t} \right) \end{array} \right] - \sum_{t \in T} \left(r_t R_t + o_t O_t \right)$$

$$V(z_0) = \sum_{t \in T} \sum_{i \in I} S^2_{i,t} V\left(\gamma_{i,s,t}\right) - \sum_{t \in T} \sum_{i \in I} P^2_{i,t} V\left(\lambda_{s,t}\right)$$

subject to

deterministic constraints (first stage):

(4.1) and **(4.2)** for production capacity;

(4.7) for mass balances except for fixed production yields (which are random);

(4.4) for availability constraints;

(4.5) and **(4.6)** for inventory requirements;

stochastic constraints (second stage):

(5.18) for demand constraints

$$x_i + z^+_{i,s} - z^-_{i,s} = d_{i,s}, \quad i \in I,\ s \in S\ ;$$

(5.24) for mass balances for fixed production yields

$$T_i x_1 + x_i + y^+_{i,k,s} - y^-_{i,k,s} = 0, \quad i \in I, k \in K, s \in S\ .$$

Solution of the first-stage variables provides decisions on the flowrate of production streams. Historical data of actual commodity prices, market demand, and product yields are considered, and depending on which scenario occurs, appropriate production will be executed in order to satisfy the realized prices, demand, and yields. These are the second-stage recourse decisions that are clearly constrained by what has been produced in the first-stage (apart from being constrained or depended upon by the corresponding scenario).

Theoretically, the solution is, in general, likely to be more representative or more robust with more scenarios considered but at the expense of being computationally expensive (that is, increase in computation time). Furthermore, with considerable number of scenarios taken into account, typically in the hundreds (for example, Pongsakdi et al. (in press) considered 600 scenarios), more "noise" is present in the data. A more practical approach is perhaps to compute the expected values of the data obtained from the first round of scenario generation, and then to subject these expected values to a second round of scenario generation in order to obtain a more robust solution, although with a trade-off in the increment of solution time. Another approach that can be considered is the use of the concept of principal component analysis (PCA) in distilling the hundreds of scenarios into a smaller number of representative scenarios, thus reducing both the amount of computational time and the presence of noisy data.

5.6.2.7 Expectation Model II

As remarked in Approach 1, a potential complication with Expectation Model I lies in computing a suitable range of values for the profit risk factor θ_1. Therefore, the proposed alternative modelling strategy of minimizing the variance, or in keeping to the maximization problem, becomes maximizing the negative of the variance, while adding a target value constraint for the mean of the original profit objective function; this is employed as follows for Expectation Model II:

$$\begin{aligned} &\text{maximize} && z_2 = -V(z_0) - E_{s'} \\ &\text{subject to} && E\left[z_0\right] \geq \text{Target objective function value} \\ & && \text{deterministic and stochastic constraints } \mathbf{(4.1)} - \mathbf{(4.7)} \end{aligned} \qquad \mathbf{(5.27)}$$

5.7 APPROACH 3: RISK MODEL II WITH VARIANCE AS THE MEASURE OF RISK OF THE RECOURSE PENALTY COSTS

5.7.1 Two-Stage Stochastic Programming with Fixed Recourse to Model Uncertainty in Prices, Demand, and Product Yields by Simultaneous Minimization of the Expected Value and the Variance of the Recourse Penalty Costs

As highlighted earlier, Mulvey et al. (1995) stress the inappropriateness of models with the expected value objective since they ignore both the risk attitude of the decision-maker and the distribution of the objective values ξ_s (as given by equation **(5.25)** for

Expectation Models I and II developed in the previous section). Higle (2005) equally advocated this by noting that expected values are risk-neutral models; hence they do not always provide a satisfactory model for decision-making under risk. This is especially so in the settings of moderate and high-risk decision-making under uncertainty as most decision-makers are risk-averse in making (relatively) important decisions.

Therefore, in this third approach, the expected value model developed in Approach 2 is extended to incorporate the measure of economic risk associated with an investment alternative. This is accomplished by reformulating the recourse penalty terms, again in the spirit of the Markowitz's mean–variance approach, as similar to the profit terms. The resulting risk model obtained is one in which in maximizing expected profit by minimizing its deviation (as computed by its variance), the expected value of the recourse penalty costs is minimized as well as its deviation by computing a parameterized function of the variance of the penalty costs. In general, this is the classical approach long favoured among financial planners in the field of financial engineering to enable decision-makers to investigate the tradeoffs between expectations and variances of costs associated with their decisions (although by stating this, we do not intend in any way to conceal the fact that there has also been published works doubting the practicality of the MV approach, among others, by Michaud and Michaud (2006), Michaud (1998), and Jobson and Korkie (1981)). According to Luenberger (1998), this approach enables the tradeoffs between the means and the variances to be explicit while Mulvey et al. (1995) perceives the incorporation of the variance term as an indicator of model robustness, as will be analyzed through the numerical example on refinery midterm planning. Note that we did not consider adopting the robust optimization model in the form originally proposed by Mulvey et al. (1995) following the argument made by Sen and Higle (1999) that the robust models are, in general, structurally unrelated to solutions obtained from the recourse model based on Markowitz's mean–variance approach. Instead, they further stressed that the robust optimization models are instead dominated by solutions derived from the inferior least-cost model.

A brief review on the concept of variance for a random variable follows. If X is a random variable (that is, a variable whose value is decided by chance) that can take on a finite number of values x_1, x_2, x_3, …, x_n, with the associated probability of such occurrences given by p_1, p_2, p_3, …, p_n, respectively, then the expected value or mean of X is defined to be:

$$E = p_1 x_1 + p_2 x_2 + p_3 x_3 + \cdots + p_n x_n \qquad \textbf{(5.28)}$$

From the definition of variance of X introduced earlier in equation **(5.5)** (Markowitz, 1952), the parameterized function of the variance for the various expected recourse penalty for the second-stage costs V_s is thus derived as:

$$\begin{aligned} V_s &= \sum_{s \in S} p_s \left(\xi_s - E_{s'} \right)^2 \\ &= \sum_{s \in S} p_s \left(\xi_s - \sum_{s' \in S'} p_{s'} \xi_{s'} \right)^2 \qquad \textbf{(5.29)} \\ V_s &= \sum_{s \in S} p_s \left\{ \sum_{i \in I} \begin{bmatrix} \left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^- \right) \\ + \left(q_i^+ y_{i,k,s}^+ + q_i^- y_{i,k,s}^- \right) \end{bmatrix} - \sum_{i \in I} \sum_{s' \in S'} p_{s'} \begin{bmatrix} \left(c_i^+ z_{i,s'}^+ + c_i^- z_{i,s'}^- \right) \\ + \left(q_i^+ y_{i,k,s'}^+ + q_i^- y_{i,k,s'}^- \right) \end{bmatrix} \right\}^2 \end{aligned}$$

Note that the index s' and the corresponding set S' is merely used to denote scenarios for the evaluation of the inner expectation term in order to distinguish from the original index s used to represent the scenarios. The variance V_s is weighted by the risk tradeoff parameter θ_2 that is varied over the entire range of $(0, \infty)$ to generate a set of feasible decisions that have maximum return for a given level of risk. This feasible decisions set is equivalent to the "efficient frontier" portfolios introduced by Markowitz (1952; 1959) for financial investment applications. The parameter θ_2 can be seen as reflecting the decision maker's attitude towards variability, that is, in more explicit terms, the risk attitude of the decision maker. Hence, the following is the mathematical description of Risk Model II, as a result of the recourse reformulation of Expectation Model I developed in Approach 2, utilizing variance as the measure of risk by minimizing the variance in the expected recourse penalty costs:

$$\begin{array}{ll} \text{maximize} & z_3 = z_2 - \theta_2 V_s = E[z_0] - \theta_1 V(z_0) - E_{s'} - \theta_2 V_s \\ \text{subject to} & \text{deterministic and stochastic constraints } \mathbf{(4.1)-(4.7)} \end{array} \quad \mathbf{(5.30)}$$

As denoted, Risk Model II is subject to the same set of constraints as for Expectation Model I outlined in Approach 2.

It is desirable to demonstrate Risk Model II as possessing robustness both in terms of its solution (*solution robust*) as well as the model itself (*model robust*). According to Mulvey et al. (1995), a solution to a stochastic optimization model is defined as solution robust if it remains *close* to optimality for all scenarios of the input data; and model robust if it remains *almost* feasible for all data scenarios. In refinery planning, model feasibility is as pertinent as solution optimality. In mitigating demand uncertainty, model feasibility is measured by expected surpluses and shortfalls in production, in which each denotes situation of excess production and unmet demand, respectively. Values of expected unmet demand should be minimized in order to gain customer demand satisfaction, while excess production should be simultaneously minimized to contribute to better inventory management.

Risk Model II is characterized by solutions in the multiobjective space as defined by the expected recourse penalty costs and the variance of the recourse costs. In the model, a measure of solution robustness is obtained by varying the penalty parameter θ_2 and observing the corresponding changes in the expected value of recourse penalty costs and expected feasibility. The model does not serve to present an absolute optimal solution that corresponds to the best possible outcome (typically in terms of maximum profit and minimum cost) desired by any decision-maker; it is merely a tool to facilitate a decision-maker in determining the choice that constitutes the best decision. As pointed out by Applequist et al. (2000), while the concept of using variance of the objective function value as a measure of risk is sound, such an approach requires the specification of the values of the one or more penalty of trade-off parameters (such as given by θ_1 and θ_2 in this Risk Model II). It is therefore left entirely in the hands of the decision-maker to effectively choose from a family of solutions corresponding to different values of these trade-off parameters.

5.7.2 Limitations of Approach 3

This approach results in an even larger nonlinear programming (NLP) problem than Approach 2, involving quadratic terms introduced by variance of the expected recourse penalty costs, which adds to the computational burden in the solution. Therefore, in general, it is undesirable to consider higher moments (beyond variance) in stochastic modelling approach as even the consideration of variance already requires the solution of a nonlinear program.

It should also be pointed out that incorporating the variance of the recourse function as part of the objective function could potentially cause the problem to lose convexity. Thus, without employing the techniques of global optimization, one is liable to get trapped in local optima solutions. In some models, incorporating the variance of the recourse function into the objective function leads to poor design or planning for cases in which the variance is small, but the design is unnecessarily expensive (Sen, 2001). Furthermore, Sen and Higle (1999) remarked that this approach that has its roots in the Markowitz's model, which is based on assumptions such as normally distributed returns, that may not necessarily hold in some applications.

In addition, since variance is a symmetric risk measure, profits both below and above the target levels are penalized equally, when it is actually desirable to only penalize profits of the former, that is, profits that are below the target (Barbaro and Bagajewicz, 2004). In other words, constraining or minimizing the variance of key performance metrics to achieve robustness, which in this case are the profit and the recourse penalty costs, may result in models that overcompensate for uncertainty, as reported by Samsatli et al. (1998). They therefore propose a general approach to robustness that can be tailored for various types of constraints to be imposed on the system and on specific suitable performance metrics. Other potentially more representative risk measures should also be considered with Kristofferson (2005) providing a recent review of a wide choice of risk measures applicable within a two-stage stochastic optimization framework (as highlighted earlier in Section 5.5.7).

5.7.3 A Brief Review of Risk Modelling in Chemical Process Systems Engineering (PSE)

The subject of risk modelling, mainly incorporated through stochastic optimization models, has been receiving increasing attention from the chemical process systems engineering (PSE) community since the publication of one of the earliest papers (if not the earliest) addressing this issue by Applequist et al. (2000) of the PSE research group at Purdue University. Of late, the Miguel Bagajewicz research group at University of Oklahoma, for instance, has produced a steady stream of publications addressing research problems related to financial risk management in planning under uncertainty, mainly with applications in the oil and gas industry. This can be found in the works of Pongsakdi et al. (in press), Barbaro and Bagajewicz (2004a, 2004b), and Aseeri and Bagajewicz (2004). Other works from Bagajewicz and co-workers that also incorporate the concept of financial risk management include Guillen et al. (2005), Aseeri et al. (2005), Bonfill et al. (2004),), and Romero et al. (2003).

5.8 APPROACH 4: RISK MODEL III WITH MEAN-ABSOLUTE DEVIATION (MAD) AS THE MEASURE OF RISK IMPOSED BY THE RECOURSE PENALTY COSTS

In this proposed fourth approach, we attempt to formulate a two-stage stochastic programming with fixed recourse framework to model the same three factors of uncertainties (namely commodity prices, market demand, and product yields) by minimizing the mean-absolute deviation (MAD) of the various expected recourse penalty for the second-stage costs. In essence, this model replaces the variance term in the objective function of Risk Model II with the MAD term.

5.8.1 The Mean-Absolute Deviation (MAD)

The mean-absolute deviation (MAD) (often inaccurately called the mean deviation) is defined as:

$$\text{MAD} = \frac{1}{N}\sum_{i=1}^{N} f_i \left| x_i - \bar{x} \right| \tag{5.31}$$

where N is the sample size, x_i are the values of the samples, $\bar{x}$ is the mean, and f_i is the absolute frequency.

In their pioneering work, Konno and Yamazaki (1991) proposed a mean-absolute deviation (MAD) portfolio model to formulate a large scale portfolio optimization problem. This serves as an alternative measure of risk to the standard Markowitz's mean–variance portfolio selection model, which models risk by the variance of the rate of return of a portfolio, thus leading to a nonlinear convex quadratic programming (QP) problem. Although both measures are almost equivalent from a mathematical point-of-view, they are substantially different from a computational point-of-view in the following ways. First, the use of MAD, essentially as a proposed linearization method of the objective function to produce an equivalent linear programming (LP) problem, serves to overcome the computational difficulties of the QP portfolio model and therefore, enables large-scale problems to be solved faster and more efficiently. This is a situation that held true at least until the mid-1990s before greater advancements in computer technology increasingly narrowed the gap in speed between the computational solution of an LP and a convex QP. Second, since the model can be casted into a linear programming (LP) problem, it can be solved much faster than a corresponding MV model. Third, the LP formulation has computational advantages over the QP formulation when integer constraints and nonconvex functions are considered. Thus, in the area of investment portfolios within the financial engineering field, the LP formulation is more suitable in handling problems associated with real transaction environments. Moreoever, it is further reported that the minimization of MAD provides similar results as the Markowitz formulation if the return is multivariate normally distributed (Konno and Wijayanayake, 2002; Konno and Koshizuka, 2005). Appendix C presents a theoretical treatment of MAD based on Konno and Yamazaki (1991).

Three additional fundamental difficulties also exist with the MV model: (i) the assumption that returns are normally distributed about the mean, which is not required in the MAD model, (ii) the required storage and calculation of a usually dense variance covariance matrix; which again is not required in the MAD model formulation and consequently, its estimation is avoided; and (iii) the use of variance as a measure of risk equally penalizes both upside and downside variation; when it should only be the latter that is undesirable and thus penalized, as stressed earlier in Section 5.7.2 and demonstrated graphically in Figure 5.9 (Konno and Yamazaki, 1991; Simaan, 1997; Speranza, 1996; Murtagh, http://www.esc.auckland.ac.n/Organisations/ORSNZ/conf37/Papers/Murtagh.pdf, accessed on November 12, 2005).

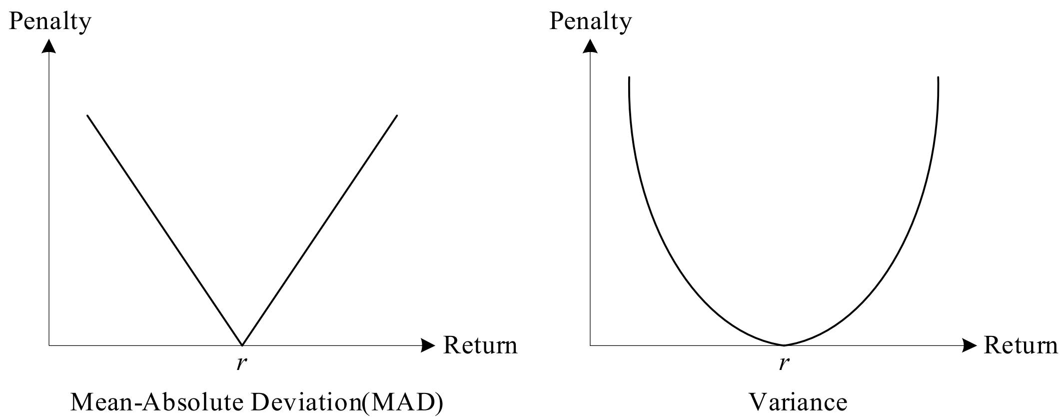


Figure 5.9. Penalty functions for mean-absolute-deviation (MAD) and variance minimization (based on Zenios and Kang (1993) and Samsatli et al. (1998)).

Ogryczak and Ruszczynski (1999) further demonstrated that MAD is an authentic measure of risk in view of its compatibility with von Neumann's principle of maximization of expected utility (MEU) under risk aversion; a result corroborated by Speranza (1996). This substantiates the solid economic foundation of the theoretical properties of MAD (Konno and Koshizuka, 2005).

5.8.2 Two-Stage Stochastic Programming with Fixed Recourse to Model Uncertainty in Prices, Demand, and Product Yields by Simultaneous Minimization of the Expected Value and the Mean-Absolute Deviation of the Recourse Penalty Costs

Konno and Yamazaki (1991) present the absolute deviation function, denoted as L_1 risk, as:

$$W(\mathbf{x}) = E\left[\left|\sum_{j=1}^{n} R_j x_j - E\left[\sum_{j=1}^{n} R_j x_j\right]\right|\right] \quad \textbf{(5.32)}$$

From equation **(5.25)** of Approach 2, the expected recourse penalty for the second-stage costs due to the combined uncertainties in market demand and product yields from crude oil is given by:

$$E_{s'} = \sum_{i\in I}\sum_{s'\in S'} p_{s'}\left[\left(c_i^+ z_{i,s'}^+ + c_i^- z_{i,s'}^-\right) + \left(q_i^+ y_{i,k,s'}^+ + q_i^- y_{i,k,s'}^-\right)\right] \quad \textbf{(5.25)}$$

Thus, the corresponding mean-absolute deviation (MAD) of the expected penalty costs due to violations of constraints for maximum product demand and product yields as a result of randomness in both demand and yields is formulated as:

$$\begin{aligned} W(p_s) &= \sum_{s\in S} p_s \left|\xi_s - E_{s'}\right| \\ &= \sum_{s\in S} p_s \left|\xi_s - \sum_{s'\in S} p_{s'}\xi_{s'}\right| \\ W(p_s) &= \sum_{s\in S} p_s \left|\sum_{i\in I}\begin{bmatrix}\left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^-\right) \\ +\left(q_i^+ y_{i,s}^+ + q_i^- y_{i,s}^-\right)\end{bmatrix} - \sum_{i\in I}\sum_{s'\in S'} p_{s'}\begin{bmatrix}\left(c_i^+ z_{i,s'}^+ + c_i^- z_{i,s'}^-\right) \\ +\left(q_i^+ y_{i,k,s'}^+ + q_i^- y_{i,k,s'}^-\right)\end{bmatrix}\right| \end{aligned} \quad \textbf{(5.33)}$$

Since this function is not linear, it is linearized by adopting the transformation procedure proposed by Konno and Yamazaki (1991) and revisited in Papahristodoulou and Dotzauer (2004). The variables $Y_{ij} \geq 0$ are defined, in which these Y_{ij} variables can be interpreted as linear mappings of the nonlinear expression given by $\left|\left(c_i^+ z_{is}^+ + c_i^- z_{is}^-\right) + \left(q_i^+ y_{is}^+ + q_i^- y_{is}^-\right) - \sum_{i\in I}\sum_{s'\in S'} p_{s'}\left[\left(c_i^+ z_{i,s'}^+ + c_i^- z_{i,s'}^-\right) + \left(q_i^+ y_{i,k,s'}^+ + q_i^- y_{i,k,s'}^-\right)\right]\right|$. Thus, equation **(5.33)** is rewritten simply as

$$Y_{ij} = W(p_s) = \sum_{s\in S} p_s \left| \sum_{i\in I} \begin{bmatrix} \left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^-\right) \\ + \left(q_i^+ y_{i,s}^+ + q_i^- y_{i,s}^-\right) \end{bmatrix} - \sum_{i\in I}\sum_{s'\in S'} p_{s'} \begin{bmatrix} \left(c_i^+ z_{i,s'}^+ + c_i^- z_{i,s'}^-\right) \\ + \left(q_i^+ y_{i,k,s'}^+ + q_i^- y_{i,k,s'}^-\right) \end{bmatrix} \right| \quad \textbf{(5.34)}$$

subjected to the following three constraints:

$$Y_{ij} \geq -\sum_{s\in S} p_s \left\{ \sum_{i\in I} \begin{bmatrix} \left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^-\right) \\ + \left(q_i^+ y_{i,s}^+ + q_i^- y_{i,s}^-\right) \end{bmatrix} - \sum_{i\in I}\sum_{s'\in S'} p_{s'} \begin{bmatrix} \left(c_i^+ z_{i,s'}^+ + c_i^- z_{i,s'}^-\right) \\ + \left(q_i^+ y_{i,k,s'}^+ + q_i^- y_{i,k,s'}^-\right) \end{bmatrix} \right\} \quad \textbf{(5.35)}$$

$$Y_{ij} \geq \sum_{s\in S} p_s \left\{ \sum_{i\in I} \begin{bmatrix} \left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^-\right) \\ + \left(q_i^+ y_{i,s}^+ + q_i^- y_{i,s}^-\right) \end{bmatrix} - \sum_{i\in I}\sum_{s'\in S'} p_{s'} \begin{bmatrix} \left(c_i^+ z_{i,s'}^+ + c_i^- z_{i,s'}^-\right) \\ + \left(q_i^+ y_{i,k,s'}^+ + q_i^- y_{i,k,s'}^-\right) \end{bmatrix} \right\} \quad \textbf{(5.36)}$$

and the non-negativity constraints for Y_{ij}:

$$Y_{ij} \geq 0 \quad \textbf{(5.37)}$$

Similar to the formulation of Risk Model II in Approach 3 that utilizes variance as the measure of risk for the recourse penalty costs, the adoption of MAD introduces the risk parameter θ_3, varied over the entire range of $(0, \infty)$ to consider its trade-offs with the expected profit term, the profit variability term, and the expected recourse term in the objective function. Therefore, the reformulated mathematical program for Risk Model III, which utilizes MAD as the measure of risk, is given by:

$$\text{maximize } z_3 = z_2 - \theta_3 W(p_s)$$
$$= E[z_0] - \theta_1 \mathrm{Var}(z_0) - E_{s'} - \theta_3 W(p_s)$$
$$\text{maximize } z_3 = E[z_0] - \theta_1 \mathrm{Var}(z_0) - \sum_{i\in I}\sum_{s\in S} p_s \left[\left(c_i^+ z_{is}^+ + c_i^- z_{is}^- \right) + \left(q_i^+ y_{is}^+ + q_i^- y_{is}^- \right) \right] \quad \textbf{(5.38)}$$
$$- \theta_3 \sum_{i\in I}\sum_{s\in S} p_s \left\| \begin{bmatrix} \left(c_i^+ z_{is}^+ + c_i^- z_{is}^- \right) \\ + \left(q_i^+ y_{is}^+ + q_i^- y_{is}^- \right) \end{bmatrix} - \sum_{i\in I}\sum_{s\in S} p_s \begin{bmatrix} \left(c_i^+ z_{is}^+ + c_i^- z_{is}^- \right) \\ + \left(q_i^+ y_{is}^+ + q_i^- y_{is}^- \right) \end{bmatrix} \right\|$$

subject to the same set of constraints as for Expectation Model I and Risk Model II with the addition of constraints **(5.35)**–**(5.37)**.

Note that the risk parameters θ_1 and θ_3, in their dual role as the scaling factors for the variance term and the MAD term respectively, would mathematically necessitate that the value of θ_1 is (much) smaller than θ_3 ($\theta_1 < \theta_3$) since θ_1 is required to scale down the squared operation involved in computing variance whereas the MAD term is in the same dimension as the expectation terms.

CHAPTER 6

Model Implementation on the General Algebraic Modeling System (GAMS)

In the past, users have to spend a substantial amount of time on computer coding to solve mathematical programming problems. However, major progress has been achieved in recent times in the development of mathematical optimization algorithms and computer codes, thus reducing significant amount of time required to form and implement solution procedures. This enables more time to be focussed on developing strong model formulations rather than developing coding and solver development.

The formulation and solution of major types of mathematical programming problems with increasingly larger scale can now be effectively performed with modelling systems such as GAMS (General Algebraic Modeling System) (Brooke et al., 1998) and AMPL (Fourer et al., 1992). While these require that the model be expressed explicitly in algebraic form, they have the advantage of being automatically interfaced with codes of solvers for solving the various types of problems that may be encountered. The modelling platform GAMS, for instance, has a library of solvers with the capability of providing global solutions for linear programs (LP), integer linear programs (ILP) and mixed integer linear programs (MILP), as well as determining local optima of nonlinear programs (NLP), integer nonlinear programs (INLP), and mixed integer nonlinear programs (MINLP) that have nonlinearities in continuous variables (Rardin, 1999). GAMS can also perform automatic differentiation and allow the use of indexed equations, which greatly facilitates and enhances the generation of large scale models. Furthermore, these modelling systems are now widely available on desktop personal computers (PCs) (Grossmann et al., http://egon.cheme.cmu.edu/papers.html, accessed on December 10, 2005).

In essence, GAMS allows the user to almost exclusively concentrate on modelling a problem by making the setup simple: defining variables, equations, and data, and then selecting an appropriate solver. In fact, GAMS possess the capability of providing a default solver that is determined to be (most) suitable for the structure of the problem at hand. The problem formulation can be altered with speed and ease, and the specification of different solvers to be tested requires only minimal effort of a single line coding. All solution algorithms can be deployed without requiring any change in existing models, thus resulting in significant reduction of time dedicated to developing and executing computational experiments.

As explained, GAMS is a modelling system for optimization that provides an interface with a variety of different algorithms of solvers. Models are supplied by the user to GAMS in an input file in the form of algebraic equations using a higher level language. GAMS then compiles the model and interfaces automatically with a solver, which is an optimization algorithm. The compiled model as well as the solution computed by the solver is subsequently reported back to the user through an output file. The simple diagram below, taken from Grossmann (1991, http://www.che.boun.edu.tr/che477/gms-mod.html, accessed on September 30, 2005), illustrates this process.

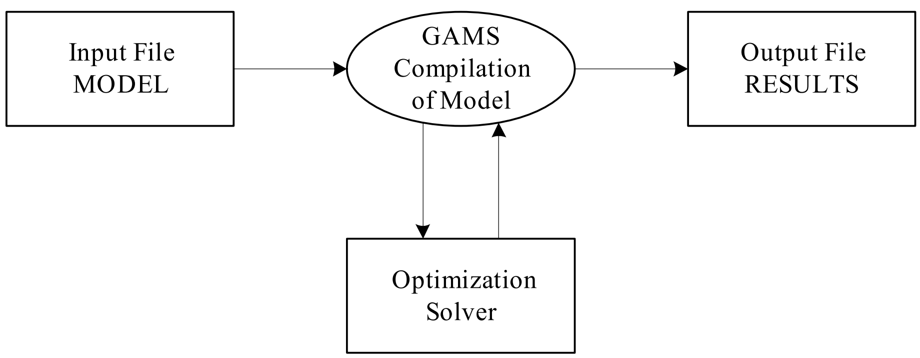


Figure 6.1. Framework of the GAMS modelling system

Note that in this work, the objective functions of the proposed stochastic models are made up of convex functions as given by the expectation operation, the variance operation, the recourse function, or the Mean-Absolute Deviation (MAD) expression, in various forms of nonnegative-weighted combinations as stipulated in the formulation of the respective models. We therefore conclude that all the models possess the highly sought-after mathematical programming property of convexity based on the theorem that states "any $f(\mathbf{x})$ formed as the nonnegative-weighted ($\alpha_i \geq 0$) sum $f(\mathbf{x}) = \sum_{i=1}^{k} \alpha_i g_i(\mathbf{x})$ of convex functions $g_i(\mathbf{x})$, $i = 1, \ldots, k$, is itself convex" (Rardin, 1998). Since the local optimal of a convex function is also the global optimal, a starting value to initialize the solution in GAMS would not be required

Based on the established ease and advantages of using GAMS, numerical studies of the deterministic model and the proposed stochastic models are coded and implemented using GAMS Integrated Development Environment (IDE) version 2.0.19.0 (Module GAMS Rev 130) for Windows platform. The models are then solved using CPLEX 9.0 (ILOG CPLEX Division, http://www.gams.com/dd/docs/solvers/cplex.pdf, accessed February 12, 2006) for the linear deterministic case and CONOPT 3 (Drud, 1996, http://www.gams.com/solvers/conopt.pdf, accessed on January 10, 2006) for the five nonlinear stochastic cases on a Pentium IV, 1.40 GHz, 512 MB of RAM machine (a notebook computer instead of a desktop, to be precise).

CPLEX is incidentally the default solver in GAMS for handling linear programs (LP), whose algorithm is based on the interior point methods that were first introduced by Karmarkar (1984). For nonlinear programming (NLP) problems, CONOPT 3 is also the GAMS default solver, in which it is based on a feasible path generalized reduced gradient method with restoration. The solutions generated for the deterministic equivalent formulation of the stochastic problems consist of: (i) the first-stage decision variables of production flowrates for all process streams and (ii) the second-stage recourse variables of production deviations due to randomness in demands and yields. Due to the nonlinearities of the stochastic models, starting values for the first-stage decision variables have been initialized to the optimal solutions obtained from the deterministic model in an effort to ensure solutions of global optimality. Although a global optimum could not be guaranteed due to the general nonconvexities of the problems, multiple local solutions have not been detected under tests of varied initial conditions. This also indicates that CONOPT 3 is a robust solver for the nonlinear nonconvex stochastic models.

In addition, it may be useful to note that nonlinear optimization algorithms often search in the space defined by superbasic variables, which are variables that are not in the basis but whose values are between the upper and lower bounds. If an infeasible solution is found by the solver used in GAMS, it is most likely due to the non-existent of a superbasic variable, in which the facility within GAMS would readily report.

CHAPTER 7

Analysis of Results from the Stochastic Models

In the context of production planning, robustness can be generally defined as a measure of resilience of the planning model to respond in the face of parameter uncertainty and unplanned disruptive events (Vin and Ierapetritou, 2001). For this, we propose the adoption of two metrics that have previously been used in the optimization literature, under similar and different contexts, to quantitatively measure and account for characteristics of planning under simultaneous uncertainty in three different parameters (namely commodities' prices, market demands, and product yields). The two metrics are: (1) the concepts of solution robustness and model robustness according to the pioneering idea of robust optimization by Mulvey and co-workers (1995) and (2) the coefficient of variation C_v.

7.1 SOLUTION ROBUSTNESS AND MODEL ROBUSTNESS

According to Mulvey et al. (1995), Bok et al. (1998) and Malcolm and Zenios (1994), *solution robustness* of an optimization model with respect to optimality is indicated by the optimal model solution that is almost optimal, or remains close to optimal, for any realization of the uncertain scenarios. This implies solutions that are less sensitive to changes in the data when different scenarios are considered. On the other hand, *model robustness* refers to solution robustness with respect to feasibility, with the optimal model solution that remains "almost" feasible for any realization of the scenarios. Thus, in general, for production planning problems, model robustness or model feasibility is represented by the optimal solution that has almost no shortfalls or surpluses in production as reflected by the expected total unmet demand and total excess production, respectively; both of which should be kept to a minimum. A trade-off exists between solution optimality and model-and-solution robustness. In order to investigate these trends, the following parameters are tabulated and analyzed from the raw computational results of the refinery production rates for the stochastic models:

- the expected deviation in profit as measured by variance $V(z_0)$;
- the expected total unmet demand or production shortfall;
- the expected total excess production or production surplus; and
- the expected recourse penalty costs E_s.

7.2 COEFFICIENT OF VARIATION

The concepts of the value of the stochastic solution VSS and the expected value of perfect information EVPI (Birge, 1982; Birge, 2005; Birge and Louveaux, 1997; Gupta and Maranas, 2000; Kall and Mayer, 2005; Uryasev, 2005) appear to be unsuitable in the case of nonlinear quadratic programming problems such as in the present work. Therefore, a different measure is sought to interpret the solutions obtained. One such approach is to investigate the coefficient of variation C_{v}. C_{v} for a set of values is defined as the ratio of the standard deviation to the expected value or mean, and is usually expressed in percentage. It is calculated as:

$$C_{\mathrm{v}} = \frac{\text{Standard Deviation}}{\text{Mean}} \times 100\% = \frac{\sigma}{\mu} \times 100\% = \frac{\sqrt{V}}{E} \times 100\% \qquad \textbf{(7.1)}$$

Statistically, C_{v} is a measure of reliability, or evaluated from the opposite but equivalent perspective, it is also indicative of a measure of uncertainty. It is alternatively interpreted as the inverse ratio of data to noise in the data in most conventional textbooks on statistics. Therefore, it is apparent that a small value of C_{v} is desirable as it signifies a small degree of noise or variability (in a data set, for instance) and hence, reflects low uncertainty.

It follows then that in the realm of stochastic optimization, coefficient of variation can be purposefully employed to investigate, denote, and compare and contrast the relative uncertainty in models being studied. In a risk minimization model, as the expected value of mean is reduced, the variability in the expected value (as typically measured by variance or standard deviation) is reduced as well. The ratio of this change can be captured and described by the coefficient of variation. Conversely, a comparison of the relative merit of models in terms of their robustness can also be represented by their respective values of coefficient of variation (that is, in the sense that a model with a lower coefficient of variation is favoured since there is less uncertainty associated with it, thus contributing to its reliability; this is in tandem with the original definition of C_v as a measure of reliability).

In addition to these arguments, in the seminal paper of his Nobel Prize-winning work (in the field of economics) of mean–variance model for optimization of investment portfolio selection, Markowitz (1952) remarked that the use of coefficient of variation as a measure of risk would equally ensure that the outcome of a decision-making process still lies in the set of efficient portfolios.

In a data set of normally distributed demands, if the coefficient of variation C_v of demand is given as a case problem parameter, the standard deviation is computed by the multiplication of C_v with the deterministic demand (Jung et al., 2004) Increasing values of C_v result in increasing fluctuations in the demand and this is again undesirable.

Computation of the coefficient of variation is based on the objective function of the formulated model. Table 7.1 displays the expressions to compute the coefficient of variation for the respective models developed in the preceding section. Note that the coefficient of variation for the corresponding deterministic case of each model is determined based on the expected (E) result of using the deterministic expected value (EV) solution, or EEV for short. In more elaborate terms, EEV is the solution obtained from solving the stochastic models using results from the deterministic expected value problem (that is, the deterministic model with the random parameters replaced by their expected values or means).

(*Please turn the page over for* Table 7.1.)

Table 7.1. Coefficient of variation for the deterministic and stochastic models developed

Approach	Model	Objective Function	Coefficient of Variation $C_v = \frac{\sigma}{\mu} = \frac{\sqrt{V}}{E}$
Deterministic		$c^T x$	[Given by the expected (E) result of using the deterministic expected value (EV) solution (EEV)]
1	Risk Model I	$\max\ z_1 = -V(z_0)$ $\max\ z_1 = E[z_0] - \theta_1 V(z_0)$	$C_v = \frac{\sqrt{V(z_0)}}{E[z_0]}$
2	Expectation Models I and II	I: $\max\ z_2 = E[z_0] - \theta_1 V(z_0) - E_s$ II: $\max\ z_2 = -V(z_0) - E_s$	$C_v = \frac{\sqrt{V(z_0)}}{E[z_0] - E_s}$
3	Risk Model II	$\max\ z_3 = E[z_0] - \theta_1 V(z_0) - E_s - \theta_2 V_s$	$C_v = \frac{\sqrt{V(z_0) + V_s}}{E[z_0] - E_s}$
4	Risk Model III (MAD)	$\max\ z_3 = E[z_0] - \theta_1 V(z_0) - E_s - \theta_3 W(p_s)$	$C_v = \frac{\sqrt{V(z_0) + W(p_s)}}{E[z_0] - E_s}$

CHAPTER 8

A Representative Numerical Example and Computational Results for Petroleum Refinery Planning under Uncertainty—I: The Base Case Deterministic Refinery Midterm/Medium-Term Production Planning Model

As a representative numerical example for the purpose of computational experimentation and testing, we consider the deterministic refinery production planning model proposed by Allen (1971) as the base case or core model, without loss of generality. The model is then reformulated with the addition of stochastic dimension according to the principles and approaches that have been extensively outlined in the previous section on general model development.

This model is also adopted in the work of Ravi and Reddy (1998), which employs the fuzzy programming technique to account for uncertainty. This provides a further avenue for us to analyze solutions from our work in light of the fuzzy approach solutions.

8.1 PROBLEM DESCRIPTION AND DESIGN OBJECTIVE

The base case models the planning of the operations of a petroleum refinery as an ordinary single objective linear programming (LP) problem of total daily profit maximization. Allen (1971) remarked that the LP approach is particularly useful in this context it provides considerable flexibility in the way a plant could be operated.

8.2 THE DETERMINISTIC REFINERY MIDTERM PRODUCTION PLANNING MODEL

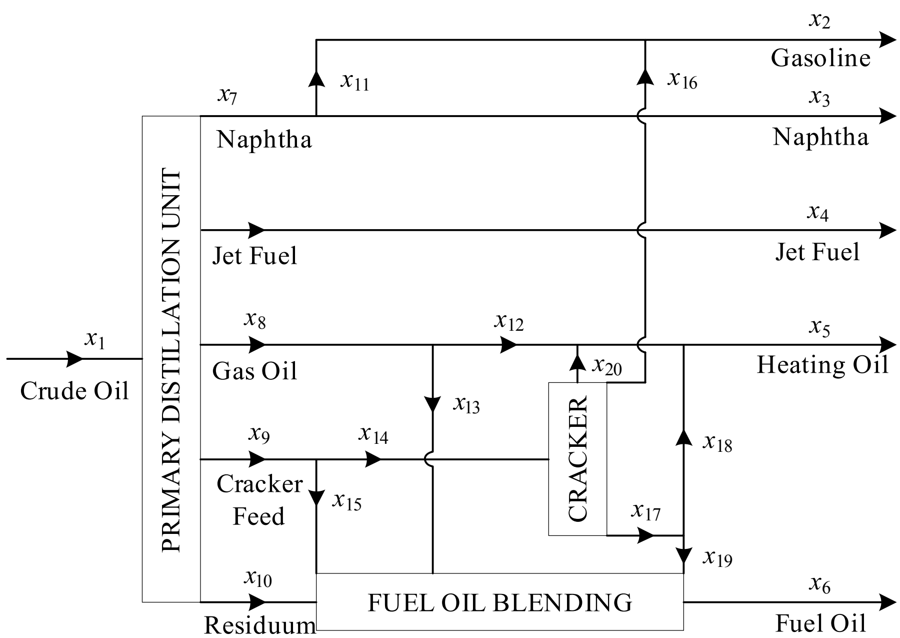


Figure 8.1. Simplified representation of a petroleum refinery for formulation of the deterministic linear program for midterm production planning

Figure 8.1 is a simplified representation of a refinery that is essentially made up of a primary distillation unit (or more commonly known nowadays as the crude distillation unit or CDU) and a middle distillates cracker (more widely known as the catalytic cracker in modern settings). The refinery processes crude oil to produce gasoline, naphtha (for gas making), jet fuel, heating oil, and fuel oil. The primary unit splits the crude into naphtha (13 weight percent, or 13 wt% yield), jet fuel (15 wt%), gas oil (22%), cracker feed (20%), and residue (30%). Gasoline is blended from naphtha and cracked blend stock in equal proportions. Naphtha and jet fuel products are straight run. Heating oil is a blend of 75% gas oil and 25% cracked oil. Fuel oil can be blended from primary residue, cracked feed, gas oil, and cracked oil in any proportions. Yields for the cracker (weight percent on feed) are flared gas (5%), gasoline blend stock (40%) and cracked oil (55%). This information along with the flow diagram of Figure 8.1 describes the physical system. All the variables that are in the same units of tonne/day, or denoted symbolically as t/d, are first assigned to process streams to represent the flow rate in each. Since in linear programming, decision variables cannot feasibly be negative, assigning a variable to a stream also defines its direction of flow and prevents any possibility of flow reversal, for instance, cracked blend stock going into naphtha product. In Figure 8.1, variables have been assigned to all the streams.

The minimum number needed to define a system fully should be identified. In this example, it is three since, for example, fixing x_1 determines x_7, x_4, x_8, x_9, and x_{10}; fixing x_2 then determines x_{11}, x_{16}, x_3, x_{14}, x_{17}, x_{20}, and x_{15};and finally fixing x_5 then determines x_{12}, x_{18}, x_{13}, x_{19}, and x_6. An LP model could be formulated using only these three structural variables or any other suitable three variables. In this case, the solution would only give values of the three variables and the remainder (if needed) would have to be calculated from them afterwards. It is usually more convenient to include some additional variables in the LP model over and above the minimum number. Each variable added needs an additional mass balance constraint to define it. Instead of calculating a variable separately from the solution, the means of finding its value is thus included in the model itself.

The next step is to construct linear constraints that describe the physical plant relationships and define the amount of flexibility existing in plant operation. These constraints are categorized as follows.

8.2.1 Limitations on Plant Capacity

In the example, the feed rates of crude oil to the primary unit and cracker, averaged over a period of time, can be anything from zero to the maximum plant capacity. The constraints are:

primary distillation unit: $x_1 \leq 15\ 000$ **(8.1)**

cracker: $x_{14} \leq 2500$ **(8.2)**

8.2.2 Mass Balances

Mass balance constraints are in the form of equalities. There are three types of such constraints: fixed plant yield, fixed blends or splits, and unrestricted balances. Except in some special situations such as planned shutdown of the plant or storage movements, the right hand-side of balance constraints is always zero. For the purpose of consistency, flow into the plant or stream junction has negative coefficients and flows out have positive coefficients. The constraints are as follows:

8.2.2.1 Fixed Yields

For the primary distillation unit:

$$-0.13x_1 + x_7 = 0 \tag{8.3}$$

$$-0.15x_1 + x_4 = 0 \tag{8.4}$$

$$-0.22x_1 + x_8 = 0 \tag{8.5}$$

$$-0.20x_1 + x_9 = 0 \tag{8.6}$$

$$-0.30x_1 + x_{10} = 0 \tag{8.7}$$

For the cracker:

$$-0.05x_{14} + x_{20} = 0 \tag{8.8}$$

$$-0.40x_{14} + x_{16} = 0 \tag{8.9}$$

$$-0.55x_{14} + x_{17} = 0 \quad \textbf{(8.10)}$$

8.2.2.2 Fixed Blends

For gasoline blending:

$$0.5x_2 - x_{11} = 0 \quad \textbf{(8.11)}$$

$$0.5x_2 - x_{16} = 0 \quad \textbf{(8.12)}$$

For heating oil blending:

$$0.75x_5 - x_{12} = 0 \quad \textbf{(8.13)}$$

$$0.25x_5 - x_{18} = 0 \quad \textbf{(8.14)}$$

8.2.2.3 Unrestricted Balances

Naphtha: $$-x_7 + x_3 + x_{11} = 0 \quad \textbf{(8.15)}$$

Gas oil: $$-x_8 + x_{12} + x_{13} = 0 \quad \textbf{(8.16)}$$

Cracker feed: $$-x_9 + x_{14} + x_{15} = 0 \quad \textbf{(8.17)}$$

Cracked oil: $$-x_{17} + x_{18} + x_{19} = 0 \quad \textbf{(8.18)}$$

Fuel oil: $$-x_{10} - x_{13} - x_{15} - x_{19} + x_6 = 0 \quad \textbf{(8.19)}$$

8.2.3 Raw Material Availabilities and Product Requirements

The constraints considered so far are concerned with the physical plant. Constraints are also needed relating to external factors such as the availability of raw materials and product requirements over a time period. For this example, there are no restrictions on crude oil availability or the minimum production required. The maximum production requirement constraints (in t/d) are as follows:

Gasoline: $$x_2 \leq 2700 \qquad \textbf{(8.20)}$$

Naphtha: $$x_3 \leq 1100 \qquad \textbf{(8.21)}$$

Jet fuel: $$x_4 \leq 2300 \qquad \textbf{(8.22)}$$

Heating oil: $$x_5 \leq 1700 \qquad \textbf{(8.23)}$$

Fuel oil: $$x_6 \leq 9500 \qquad \textbf{(8.24)}$$

8.2.4 Objective Function

Although optimization can be stated in many different ways, the common optimization to an industrial process is to maximize the profitability of the process, or to minimize the overall costs, in which the former is adopted in this work. In this model, the whole refinery is considered to be one process, where the process uses the given petroleum crude oil to produce various petroleum products in order to achieve specific economic objectives. Thus, the objective of the optimization at hand is to achieve maximum profitability given the type of crude oil and the refinery facilities. No major hardware change in the current facilities is considered in the optimization. The optimization tries to find the optimal operation modes of units and stream flows that maximize the overall profit of the whole refinery while observing all the possible process constraints.

Thus, the economic objective function considered is the net profit or net revenue to be maximized, in units of \$/day. The cost of acquiring the raw material crude oil and transforming it to finished products are subtracted from the gross revenues accruing from the sale of finished products. Note that there is no cost associated with blending as this cost is netted out of the unit sales price of the finished products. The sign convention denotes costs as negative and realization from sales as positive. Each element in it consists of the product of coefficient of unit cost or unit sales price (\$/ton) and a production flowrate variable (ton/day or t/d). Thus, the objective function is as follows (Shapiro, 1993):

$$\text{maximize } z = \underbrace{-7.5x_1}_{\substack{\text{crude}\\ \text{oil}}} \underbrace{-0.5x_1}_{\substack{\text{primary}\\ \text{unit}}} \underbrace{-1.5x_{14}}_{\text{cracker}} + \underbrace{18.5x_2}_{\text{gasoline}} + \underbrace{8.0x_3}_{\text{naphtha}} + \underbrace{12.5x_4}_{\text{jet fuel}} + \underbrace{14.5x_5}_{\text{heating oil}} + \underbrace{6.0x_6}_{\text{fuel oil}}$$

hence,

$$\underset{x_i}{\text{maximize}}\ z = -8.0x_1 + 18.5x_2 + 8.0x_3 + 12.5x_4 + 14.5x_5 + 6.0x_6 - 1.5x_{14} \qquad \textbf{(8.25)}$$

Since linear programming variables cannot feasibly be negative, an additional constraint to be specified is:

$$x_1, x_2, \ldots, x_{20} \geq 0 \quad \text{or} \quad x_i \geq 0, i = 1, 2, \ldots 20 \qquad \textbf{(8.26)}$$

It is noted that the constraints of production requirements is directly impacted by the market demand for the final refinery products. Therefore, in the stochastic version of the model, the random decision variables will be introduced into these constraints. Since the objective function is to maximize profit, the production requirements are expressed in terms of inequalities with an upper bound ("less than" inequalities) in order to ensure that the optimization problem is bounded (otherwise, the maximum profit can be solved to an infinite value). Conversely, if the objective function is to minimize production cost, then the production requirements would be expressed in the form of inequalities with a lower bound. This can be deduced logically, as a "less than" inequality would include the possibility of the solution of non-production (that is, the operations or production of the refinery is halted or the entire plant is shut down) as cost is equals to zero in this situation.

8.3 COMPUTATIONAL RESULTS FOR DETERMINISTIC MODEL

The deterministic LP model was set up on GAMS and solved using CPLEX version 9.0.2 (http://www.gams.com/dd/docs/solvers/cplex.pdf, accessed February 12, 2006). CPLEX has been proven to be a very stable LP solver, and the default settings are almost always sufficient to obtain an optimal solution within excellent solution times (http://www.gams.com/solvers/solvers.htm#CPLEX, accessed February 12, 2006). The solution computed was compared against the solution obtained using the computational software package LINDO (Linear Interactive and Discrete Optimizer) (Schrage, 1990), an easy-to-use engine for solving linear and integer optimization models. Both solutions have been verified to be consistent with each other (with values generated accurate to three decimal places for the computation with GAMS/CPLEX and to six decimal places with LINDO) and is tabulated in Table 8.1.

In CPLEX, a normal run performs an iterative procedure analogous to the LP primal simplex method until the optimal solution is reached. The optimal solution for this deterministic model are obtained by CPLEX after three (3) iterations in a trivial CPU time whereas nine (9) iterations are needed by using LINDO (also in within negligible CPU time). Table 8.2 details the computational statistics for solving Deterministic Model.

8.4 SENSITIVITY ANALYSIS FOR THE SOLUTION OF DETERMINISTIC MODEL

A sensitivity analysis for the objective function coefficients is performed using the available facility in LINDO to determine lower and upper limits of each coefficient with no change in the optimal solution. Table 8.3 displays the ensuing results of these limits in

Table 8.1. Computational results for Deterministic Model from GAMS/CPLEX and LINDO

Decision Variable	Value (ton/day) Lower Limit	Level	Upper Limit	Dual Price/ Marginal ($/ton)	Slack or Surplus Variable	Value (ton/day)	Dual Price/ Marginal ($/ton)
x_1	0	12 500	$+\infty$	0	s_1	2 500	0
x_2	0	2 000	$+\infty$	0	s_2	0	3.575
x_3	0	625	$+\infty$	0	s_3	700	0
x_4	0	1 875	$+\infty$	0	s_4	475	0
x_5	0	1 700	$+\infty$	0	s_5	425	0
x_6	0	6 175	$+\infty$	0	s_6	0	8.5
x_7	0	1 625	$+\infty$	0	s_7	3 325	0
x_8	0	2 750	$+\infty$	0	a_8	0	8.0
x_9	0	2 500	$+\infty$	0	a_9	0	12.5
x_{10}	0	3 750	$+\infty$	0	a_{10}	0	6.0
x_{11}	0	1 000	$+\infty$	0	a_{11}	0	9.825
x_{12}	0	1 275	$+\infty$	0	a_{12}	0	6.0
x_{13}	0	1 475	$+\infty$	0	a_{13}	0	0
x_{14}	0	2 500	$+\infty$	0	a_{14}	0	29.0
x_{15}	0	0	$+\infty$	–3.825	a_{15}	0	6.0
x_{16}	0	1 000	$+\infty$	0	a_{16}	0	8.0
x_{17}	0	1 375	$+\infty$	0	a_{17}	0	29.0
x_{18}	0	425	$+\infty$	0	a_{18}	0	6.0
x_{19}	0	950	$+\infty$	0	a_{19}	0	6.0
x_{20}	0	125	$+\infty$	0	a_{20}	0	8.0
					a_{21}	0	6.0
					a_{22}	0	9.825
					a_{23}	0	6.0
					a_{24}	0	6.0
z ($/day)	$-\infty$	23 387.50	$+\infty$	**(optimal objective function value = maximum profit)**			

Table 8.2. Computational statistics for Deterministic Model

Solver	Single continuous variables	Constraints	Resource usage/ CPU time (s)	Iterations
CPLEX	21	25	0.000 (trivial)	3

which the current solution or basis remains optimal. The allowable increase column section indicates the amount by which an objective function coefficient can be increased with the current basis remaining optimal, giving the lower limit value of the coefficient. Conversely, the allowable decrease column section is the amount by which the objective coefficient can be decreased with the current basis remaining optimal, thus determining the upper limit.

We observe that nine out of the 20 decision variables have positive infinity as an upper limit for their associated coefficients and this is deemed reasonable since all the decision variables are positive; moreover, the objective is to maximize profit. For example, the

Table 8.3. Sensitivity analysis for the objective function coefficients of Deterministic Model

Decision variable	Objective function coefficient ranges ($/ton)				
	Allowable decrease	Lower limit	Current value	Upper limit	Allowable increase
x_1	0.715 000	−8.715 000	−8.000 000	−7.235 000	0.765 000
x_2	4.468 750	14.031 25	18.500 000	+∞	+∞
x_3	14.300 000	−6.300 000	8.000 000	13.884 615	5.884 615
x_4	4.766 667	7.733 333	12.500 000	17.600 000	5.100 000
x_5	8.500 000	6.000 000	14.500 000	+∞	+∞
x_6	1.134 921	4.865079	6.000 000	7.062 500	1.062 500
x_7	5.500 000	−5.500 000	0.000 000	5.884 615	5.884 615
x_8	3.250 000	−3.250 000	0.000 000	3.477 273	3.477 273
x_9	3.575 000	−3.575 000	0.000 000	3.825 000	3.825 000
x_{10}	2.383 333	−2.383 333	0.000 000	2.550 000	2.550 000
x_{11}	8.937 500	−8.937 500	0.000 000	+∞	+∞
x_{12}	11.333 333	11.333 333	0.000 000	+∞	+∞
x_{13}	3.250 000	3.250 000	0.000 000	3.477 273	3.477 273
x_{14}	3.575 000	−5.075 000	−1.500 000	+∞	∞
x_{15}	+∞	−∞	0.000 000	3.825 000	3.825 000
x_{16}	8.937 500	−8.937 500	0.000 000	+∞	+∞
x_{17}	6.500 000	−6.500 000	0.000 000	+∞	+∞
x_{18}	34.000 000	−34.000 000	0.000 000	+∞	+∞
x_{19}	6.500 000	−6.500 000	0.000 000	34.000 000	34.000 000
x_{20}	71.500 000	−71.500 000	0.000 000	+∞	+∞

decision variable x_2 denoting the production mass flow rate of gasoline (in ton/day) has current objective function coefficient value of price of $18.50/ton for profit. The price of gasoline can be as low as $4.47/ton without altering the optimum profit. Therefore, if a customer wishes to enter into a purchasing agreement or contract for the commodities produced by the refinery, the trader or marketer, acting with the knowledge of the management, can negotiate the trading price down to as low as the extent of the lower limits of prices listed in Table 8.2 without affecting or "hurting" the company's profit, so to speak, as the optimal solution would not be changed. This boosts the refinery's flexibility to negotiate prices so long as it is within the bounds of each coefficient as determined in Table 8.2, especially in the volatile market of spot trading of crude oil and the commodities (Zayed and Minkarah, 2004).

As before, the sensitivity analysis for the right-hand side of constraints is executed as well by utilizing LINDO with the corresponding results displayed in Table 8.4. The upper limits for some of the constraints are positive infinity while the lower limits vary. Lower limits for most of the constraints are observed to be of negative value. Constraints whose upper limits are allowed to go to positive infinity imply that they are not critical to the production process. As an illustration, the right-hand side constraint for the production requirement or demand of heating oil can be as low as none without inflicting a change in the optimal solution. The other constraints can be analyzed in a similar manner (Zayed and Minkarah, 2004).

The values of the slack or surplus variables and the dual prices in Table 8.1 provide the most economical average operating plan for a 30-day period. For instance, it indicates that the primary distillation unit is not at full capacity as the solution generates a production mass of 12 500 tons/day against its maximum production capacity of 17 300 tons/day (as given by the summation of the right-hand-side values of constraints **(8.20)**–**(8.24)**). Another analytical observation reveals that the maximum production requirement is only met for heating oil.

Table 8.4. Sensitivity analysis for right-hand side of constraints of Deterministic Model

Constraints	Right-hand side of constraints ranges (ton/day)				
	Allowable decrease	Lower limit	Current value	Upper limit	Allowable increase
(8.1)	2 500.000 000	12 500.000 000	15 000.000 000	+∞	+∞
(8.2)	1340.909058	1 159.090 94	2 500.000 000	3 000.000 000	500.000 000
(8.3)	625.000 000	−625.000 000	0.000 000	475.000 000	475.000 000
(8.4)	1 875.000 000	−1 875.000 000	0.000 000	425.000 000	425.000 000
(8.5)	1 475.000 000	−1 475.000 000	0.000 000	3 325.000 000	3 325.000 000
(8.6)	500.000 000	−500.000 000	0.000 000	961.538 510	961.538 513
(8.7)	3 750.000 000	−3 750.000 000	0.000 000	3 325.000 000	3 325.000 000
(8.8)	125.000 000	−125.000 000	0.000 000	+∞	+∞
(8.9)	475.000 000	−475.000 000	0.000 000	350.000 000	350.000 000
(8.10)	950.000 000	−950.000 000	0.000 000	3 325.000 000	3 325.000 000
(8.11)	625.000 000	−625.000 000	0.000 000	475.000 000	475.000 000
(8.12)	475.000 000	−475.000 000	0.000 000	350.000 000	350.000 000
(8.13)	1 475.000 000	−1 475.000 000	0.000 000	1 275.000 000	1 275.000 000
(8.14)	950.000 000	−950.000 000	0.000 000	425.000 000	425.000 000
(8.15)	625.000 000	−625.000 000	0.000 000	475.000 000	475.000 000
(8.16)	1 475.000 000	−1 475.000 000	0.000 000	3 325.000 000	3 325.000 000
(8.17)	500.000 000	−500.000 000	0.000 000	961.538 510	961.538 513
(8.18)	950.000 000	−950.000 000	0.000 000	3 325.000 000	3 325.000 000
(8.19)	6 175.000 000	−6 175.000 000	0.000 000	3 325.000 000	3 325.000 000
(8.20)	700.000 000	2 000.000 000	2 700.000 000	+∞	+∞
(8.21)	475.000 000	625.000 000	1 100.000 000	+∞	+∞
(8.22)	425.000 000	1 875.000 000	2 300.000 000	+∞	+∞
(8.23)	1 700.000 000	0.000 000	1 700.000 000	3666.6666	1 966.666 626
(8.24)	3 325.000 000	6 175.000 000	9 500.000 000	+∞	+∞

By definition, the dual price or shadow price of a constraint of a linear programming model is the amount (or rate) by which the optimal value of the objective function is improved (increased in a maximization problem and decreased in a minimization problem) if the right-hand-side of a constraint is increased by one unit, with the current basis remaining optimal. A positive dual price means that increasing the right-hand side in question will improve the objective function value. A negative dual price means that increasing the right-hand side will have a reverse effect. Thus, the dual price of a slack variable corresponds to the effect of a marginal change in the right-hand-side of the appropriate constraint (Winston and Venkataramanan, 2003).

The dual prices of slacks on mass balance and product requirement rows can be interpreted more specifically. Consider a mass balance constraint:

$$-x_1 - x_2 + x_3 = 0 \tag{8.27}$$

where x_3 is the product stream. Introducing the artificial slack variable a_n and then rearranging, we obtain:

$$\begin{aligned} -x_1 - x_2 + x_3 + a_n &= 0 \\ x_1 + x_2 &= x_3 + a_n \end{aligned} \tag{8.28}$$

The product stream is increased by a_n and the feed streams x_1 and x_2 must increase correspondingly. The dual price of a_n indicates the effect of making marginally more products without taking into account its realization (which is on x_3), that is, it indicates the cost added by producing one extra item of the product, or in other words, the marginal cost of making the product. In addition to that, consider the product requirement constraint:

$$\begin{aligned} &x_3 \le 1100 \\ &\quad \Rightarrow x_3 + s_m = 1100 \\ &\quad \Rightarrow x_3 = 1100 - s_m \end{aligned} \tag{8.29}$$

The dual price of the slack variable s_m on this constraint indicates the effect of selling this product at the margin, that is, it indicates the marginal profit on the product. If the constraint is slack so that the slack variable is positive (basic), the profit at the margin must obviously be zero and this is in line with the zero dual price of all basic variables. Since cost + profit = realization for a product, the sum of the dual prices on its balance and requirement constraints equal its coefficient in the original objective function.

In this problem, there are two balance constraints on heating oil as given by equations **(8.13)** and **(8.14)**, the dual prices of which are both \$6.00/ton. This is the marginal cost of diverting gas oil and cracked oil from fuel oil to heating oil. The dual price for the constraint on heating oil production as given by inequality **(8.23)** is \$8.50/ton and this is the marginal profit on heating oil, in line with the realization of \$14.50/ton in the objective function as given by the coefficient of x_5 (Allen, 1971).

From the economic interpretation viewpoint, the dual prices can be seen as prices for the scarce resources that minimize the total accounting cost of these resources to the refinery, and yet involve a scarce-factor cost of producing a unit of each commodity that is no less than its unit profit yield. The dual prices indicate what proportion of its profits that the refinery owes to each such scarce factor (Baumol, 1958).

As stated earlier, the solution for the deterministic model was verified with the optimization software LINDO. Given that the numerical example presented as a case study addresses only the three primary units of a typical oil refinery, it should be kept in mind that while the example model is nonetheless representative, the results should be viewed as a (preliminary) proof of concept rather than a well-tested planning model for the operations of a refinery. The emphasis (and novelty) of this work lies chiefly in the five stochastic models, to be presented in subsequent sections, for planning in the downstream processing of the highly dynamic and uncertain hydrocarbon industry.

8.5 DISADVANTAGES OF THE SENSITIVITY ANALYSIS OF LINEAR PROGRAMMING AS MOTIVATION FOR STOCHASTIC PROGRAMMING

Sensitivity analysis is used with a pretext of an attempt to study the robustness of the solution to a linear programming (LP) model. If there is cause for concern regarding the accuracy of the data used, sensitivity analysis is undertaken to determine the manner in which the solution might change if the data were different. When the solution does not change (or when the nature of the solution does not change, as in when the basis remains optimal), it is believed that the proposed solution is appropriate. Unfortunately, such is not true if the solution is sensitive to the data. A question arises as to how to proceed if the solution, or the nature of the solution, varies when the data is changed. Therefore, although sensitivity analysis offers some sense of security, it is important to recognize that in many cases, this is really somewhat a false sense of security. If there is some uncertainty about the values of some data elements, it ought to be included in the model. This is precisely the situation for which the stochastic programming modelling approach are intended, that is, when we know that some of the data elements are difficult to predict or estimate (Higle, 2005).

Furthermore, in most cases, the output of SA is misleading when used to assess the impact of uncertainty. SA is most appropriate when the basic structure of the model is not altered by the presence of uncertainty—for example, when all uncertainties will be resolved before any decisions are made. When the decisions are to be made, a deterministic model will be appropriate, but as long as the available data and information remain uncertain, we will not know which deterministic model will be appropriate and suitable. SA is merely able to help us appreciate the impact of uncertainty without providing the measures to hedge against it. This is because sensitivity analysis based on the output of a model constructed on the presumption of deterministic data as in an LP will not reflect an ability to adopt to information that becomes available within a sequential decision process, thus rendering it ineffectual for decision making under uncertainty (Higle and Wallace, 2003).

This explains why Mulvey et al. (1995) argue that sensitivity analysis (SA) is a reactive approach to controlling uncertainty in justifying the adoption of the stochastic programming philosophy. As emphasized by Higle (2004), SA merely measures the sensitivity of a solution to changes in the input data. It provides no mechanism by which this sensitivity can be controlled. On the other hand, stochastic programming is a constructive approach that is superior to SA. With stochastic linear programming (SLP) models, the decision maker is afforded the flexibility of introducing recourse variables to take corrective actions.

Nevertheless, the SLP model optimizes only the first moment of the distribution of the objective value as it ignores higher moments of the distribution, in addition to the decision maker's preferences towards risk. These aspects are particularly important for asymmetric distributions and for risk-averse decision makers. Therefore, in this work the SLP formulation is extended by incorporating risk measures in the form of variance and mean-absolute deviation (MAD), as will be demonstrated in the stochastic models introduced in the following section.

In handling constraints, SLP models aim at finding the planning variable (x) such that for each realized scenario, an operating variable setting (y) is possible in satisfying the constraints. For systems with some redundancy, such a solution might always be possible. The SLP literature even allows for the notion of complete recourse, whereby a feasible solution y exists for all scenarios, and for any value of x that satisfies the recourse constraints.

CHAPTER 9

A Representative Numerical Example and Computational Results for Petroleum Refinery Planning under Uncertainty—II: The Stochastic Refinery Midterm/Medium-Term Production Planning Model

9.1 APPROACH 1: RISK MODEL I BASED ON THE MARKOWITZ'S MEAN–VARIANCE (*E*–*V*) APPROACH

The deterministic objective function is given by:

$$z = -8.0x_1 + 18.5x_2 + 8.0x_3 + 12.5x_4 + 14.5x_5 + 6.0x_6 - 1.5x_{14} \quad \textbf{(8.25)}$$

where in the following the coefficients, with the associated variables of amount of production in mass flowrate indicated in parentheses, denote the price of crude oil (x_1) as the raw material and the sales prices of the products or commodities, namely gasoline (x_2), naphtha (x_3), jet fuel (x_4), heating oil (x_5), fuel oil (x_6), and the feed to the cracking unit (x_{14}) (henceforth, referred to simply as the cracker feed), respectively. Therefore, if c is a row vector consisting of the price (or cost) coefficients as its elements and x is the column vector of production flowrate, then the objective function can simply be generally represented as:

$$z = c^T x \quad \textbf{(9.1)}$$

For the prioritized purpose of method demonstration of the validity of the mathematical programming tools proposed (that is, without claiming that the model captures all detailed aspects of the problem), three possible coefficients of variation (defined as the ratio of standard deviation to mean) depicting three different scenarios are considered to be representative of the uncertainty in the objective function coefficients of prices, based on the trends of the historical data presented in Sections 5.5.1 and 5.5.2. The three representative scenarios are constructed as (i) the "above average" scenario denoting a representative percentage of 10 percent positive deviation from the mean value; (ii) the "average" scenario that takes on the expected value or mean; and (iii) the "below average" scenario, correspondingly denoting a representative 10 percent negative deviation from the mean value.

Table 9.1 summarizes attributes of modelling uncertainty in the price of crude oil. Subsequently, Table 9.2 displays details of all three scenarios for all materials where price uncertainty is considered.

Table 9.1. Attributes of the scenario construction example for modelling crude oil price uncertainty

	Price Uncertainty: Objective Function Coefficient of Prices ($/ton)		
	Scenario 1 (s = 1)	Scenario 2 (s = 2)	Scenario 3 (s = 3)
	Above Average Price	Average Price	Below Average Price
Percentage of deviation from the expected value	+10%	0 (i.e., the expected value)	−10%
Price of crude oil ($/ton) P_s	P_1 = 8.8	P_2 = 8.0	P_3 = 7.2
Probability p_s	p_1 = 0.35	p_2 = 0.45	p_3 = 0.20

Table 9.2. Representative scenarios of price uncertainty in the refinery planning under uncertainty problem

	Price Uncertainty: Objective Function Coefficient of Prices ($/ton)		
	Scenario 1	Scenario 2	Scenario 3
Material/Product (i)	Above Average Price (+10%)	Average Price (Expected Value/Mean)	Below Average Price (−10%)
Crude oil (1)	−8.8	−8.0	−7.2
Gasoline (2)	20.35	18.5	16.65
Naphtha (3)	8.8	8.0	7.2
Jet fuel (4)	13.75	12.5	11.25
Heating oil (5)	15.95	14.5	13.05
Fuel oil (6)	6.6	6.0	5.4
Cracker feed (14)	−1.65	−1.5	−1.35
Probability p_s	0.35	0.45	0.2

As stressed in the general model development, since the objective function is linear, the expectation of the objective function value is given by the original objective function itself:

$$E[z_0] = E(-8.0x_1 + 18.5x_2 + 8.0x_3 + 12.5x_4 + 14.5x_5 + 6.0x_6 - 1.5x_{14})$$
$$= E(-8.0x_1) + E(18.5x_2) + E(8.0x_3) + E(12.5x_4) + E(14.5x_5) + E(6.0x_6) + E(-1.5x_{14})$$
$$E[z_0] = -8.0x_1 + 18.5x_2 + 8.0x_3 + 12.5x_4 + 14.5x_5 + 6.0x_6 - 1.5x_{14} \qquad \textbf{(9.2)}$$

To represent the three scenarios accounting for uncertainty in prices, the objective coefficients of the expected value of price in equation **(9.2)** are rewritten as follows, taking into account the probabilities of realization of each scenario:

$$\begin{aligned} E[z_0] = & E\{[(0.35)(-8.8) + (0.45)(-8.0) + (0.2)(-7.2)]x_1\} \\ & + E\{[(0.35)(20.35) + (0.45)(18.5) + (0.2)(16.65)]x_2\} \\ & + E\{[(0.35)(8.8) + (0.45)(8.0) + (0.2)(7.2)]x_3\} \\ & + E\{[(0.35)(13.75) + (0.45)(12.5) + (0.2)(11.25)]x_4\} \\ & + E\{[(0.35)(15.95) + (0.45)(14.5) + (0.2)(13.05)]x_5\} \\ & + E\{[(0.35)(6.6) + (0.45)(6.0) + (0.2)(5.4)]x_6\} \\ & + E\{[(0.35)(-1.65) + (0.45)(-1.5) + (0.2)(-1.35)]x_{14}\} \end{aligned} \qquad \textbf{(9.3)}$$

For a more explicit representation of the three scenarios considered, the terms in equation **(9.3)** are rearranged and clustered into three expressions, with each denoting a corresponding scenario:

$$\begin{aligned} E[z_0] = & (0.35)(-8.8x_1 + 20.35x_2 + 8.8x_3 + 13.75x_4 + 15.95x_5 + 6.6x_6 - 1.65x_{14}) \\ & + (0.45)(-8.0x_1 + 18.5x_2 + 8.0x_3 + 12.5x_4 + 14.5x_5 + 6.0x_6 - 1.5x_{14}) \\ & + (0.2)(-7.2x_1 + 16.65x_2 + 7.2x_3 + 11.25x_4 + 13.05x_5 + 5.4x_6 - 1.35x_{14}) \end{aligned} \qquad \textbf{(9.4)}$$

or in a general compact representation as given below:

$$E[z_0] = \sum_{i \in I} \sum_{s \in S} p_s C_i x_i, \quad i = \{1,2,3,4,5,6,14\} \in I_{\text{price}}^{\text{random}} \subseteq I, s = \{1,2,3\} \in S \qquad \textbf{(9.5)}$$

To formulate the variance for the given expected value of the objective function, note that it is the coefficients of the objective function that are random (as also emphasized earlier in the model development) and not the deterministic design variables x_1, x_2, x_3, x_4, x_5, x_6, x_{14}; thus, variance is expressed as:

$$\begin{aligned} V(z_0) &= V\left(-8.0x_1 + 18.5x_2 + 8.0x_3 + 12.5x_4 + 14.5x_5 + 6.0x_6 - 1.5x_{14}\right) \\ &= V\left(-8.0x_1\right) + V\left(18.5x_2\right) + V\left(8.0x_3\right) + V\left(12.5x_4\right) + V\left(14.5x_5\right) \\ &\quad + V\left(6.0x_6\right) + V\left(1.5x_{14}\right) \\ V(z_0) &= x_1^2 V\left(-8.0\right) + x_2^2 V\left(18.5\right) + x_3^2 V\left(8.0\right) + x_4^2 V\left(12.5\right) + x_5^2 V\left(14.5\right) \\ &\quad + x_6^2 V\left(6.0\right) + x_{14}^2 V\left(-1.5\right) \end{aligned} \tag{9.6}$$

or in a general compact representation:

$$V(z_0) = \sum_{i \in I} x_i^2 V(C_i), \quad i = \{1,2,3,4,5,6,14\} \in I_{\text{price}}^{\text{random}} \subseteq I \tag{9.7}$$

To evaluate the variance of the price coefficients, we use the formulation presented in equation **(5.7)** by substituting the objective functions for each of the three scenarios with its general form as given by equation **(9.1)**, as follows:

$$V\left(z_0\right) = p_{s_1}\left(c_{s_1}^T x - \overline{c}^T x\right)^2 + p_{s_2}\left(c_{s_2}^T x - \overline{c}^T x\right)^2 + p_{s_3}\left(c_{s_3}^T x - \overline{c}^T x\right)^2 \tag{9.8}$$

where $z_{s_i} = c_{s_i}^T x$ and $E[z_0] = \overline{c}^T x$.

Since Scenario 2 represents the average scenario, so $c_{s_2}^T x - \overline{c}^T x = 0$ and yields

$$\begin{aligned} V\left(z_0\right) &= p_{s_1}\left[\left(c_{s_1}^T - \overline{c}^T\right)x\right]^2 + p_{s_3}\left[\left(c_{s_3}^T - \overline{c}^T\right)x\right]^2 \\ V\left(z_0\right) &= p_{s_1}\left(c_{s_1}^T - \overline{c}^T\right)^2 x^2 + p_{s_3}\left(c_{s_3}^T - \overline{c}^T\right)^2 x^2 \end{aligned} \tag{9.9}$$

Comparison between equations **(9.6)** and **(9.9)** reveals that the variances of the price coefficients can be estimated as the variance for a sample consisting of the three considered scenarios. As an example, variance for the price coefficient of crude oil is given by:

$$\begin{aligned} V(-8.0) &= p_{s1}\left(c_{s1}^{T}-\bar{c}^{T}\right)^{2}+p_{s3}\left(c_{s3}^{T}-\bar{c}^{T}\right)^{2} \\ &= p_{s1}\left(c_{s1}^{T}-c_{s2}^{T}\right)^{2}+p_{s3}\left(c_{s3}^{T}-c_{s2}^{T}\right)^{2} \\ &= (0.35)\left[(-8.8)-(-8.0)\right]^{2}+(0.2)\left[(-7.2)-(-8.0)\right]^{2} \\ &= (0.35)[-0.8]^{2}+(0.2)[-0.8]^{2} \\ &= (0.35+0.2)(0.64) \\ V(-8.0) &= 0.352 \end{aligned} \tag{9.10}$$

Similar variance calculation procedure is carried out for the other variance terms to yield the results tabulated in Table 9.3.

Table 9.3. Variance of the random objective function coefficients of commodity prices

Product Type (i)	Variance of Price
Crude oil (1)	0.352
Gasoline (2)	1.882375
Naphtha (3)	0.352
Jet fuel (4)	0.859375
Heating oil (5)	1.156375
Fuel oil (6)	0.198
Cracker feed (14)	0.012375

Substituting the values of price variances calculated and tabulated in Table 9.3 into equation **(9.6)** yields:

$$\begin{aligned} V(z_0) = {} & (0.352)x_1^2+(1.882\,375)x_2^2+(0.352)x_3^2+(0.859\,375)x_4^2 \\ & +(1.156\,375)x_5^2+(0.198)x_6^2+(0.012\,375)x_{14}^2 \end{aligned} \tag{9.11}$$

Therefore, Risk Model I is formulated as:

maximize $z_1 = -V(z_0)$

$$= -\begin{bmatrix} (0.352)x_1^2 + (1.882\,375)x_2^2 + (0.352)x_3^2 + (0.859\,375)x_4^2 \\ +(1.156\,375)x_5^2 + (0.198)x_6^2 + (0.012\,375)x_{14}^2 \end{bmatrix}$$

s.t

$$E[z_0] = \begin{bmatrix} (0.35)(-8.8x_1 + 20.35x_2 + 8.8x_3 + 13.75x_4 + 15.95x_5 + 6.6x_6 - 1.65x_{14}) \\ +(0.45)(-8.0x_1 + 18.5x_2 + 8.0x_3 + 12.5x_4 + 14.5x_5 + 6.0x_6 - 1.5x_{14}) \\ +(0.2)(-7.2x_1 + 16.65x_2 + 7.2x_3 + 11.25x_4 + 13.05x_5 + 5.4x_6 - 1.35x_{14}) \end{bmatrix} \geq \begin{matrix}\text{Target} \\ \text{objective} \\ \text{function} \\ \text{value}\end{matrix}$$

constraints **(8.1)–(8.24)** and **(8.26)**

(9.12)

It is noted that the set of constraints for Risk Model I are the same as for the Deterministic Model.

9.1.1 Computational Results for Risk Model I

Table 9.4 tabulates the computational results for the implementation of Risk Model I on GAMS for a range of values of the target profit μ. Starting values of the first-stage deterministic decision variables have been initialized to the optimal solutions of the deterministic model presented earlier. From the raw computational results of Risk Model I, the standard deviation σ of profit is determined by taking the square root of the computed values of variance σ^2 of profit as given by the objective values. The main reason standard deviation is considered to be more representative for direct interpretation is by virtue of it having the same dimension as the expected value term. Note that σ is calculated by taking the square root of the absolute values of variance, that is, with the negative sign of variance disregarded. (Recall that the negative sign is present essentially because we are dealing with an optimization problem of profit maximization, in which it is desirable to minimize the effect of variation in profit by subtracting it from the profit- or cost-related terms.) Subsequently, Table 9.5 presents representative detailed results for two values of target profit: one that is equals to the profit computed by the Deterministic Model (that is, \$23 387.50/day) and the other, for target profit = \$23 500, with the intention of investigating and thereafter, inferring observable general behaviours of Risk Model I.

The problem size and the distribution of computational expense are noted in Table 9.6. Figure 9.1 then shows the efficient frontier (of Markowitz's mean–variance model) plot of expected maximum profit for different levels of risk as represented by the profit risk parameter θ_1 with standard deviation as the risk measure.

Table 9.4. Computational results for Risk Model I

Target profit μ (\$/day)	Optimal objective value $-V$	Standard deviation σ $(=\sqrt{V})$	Crude oil, x_1 (ton/day)	$C_v = \frac{\sigma}{\mu}$ Stochastic	Deterministic σ	Deterministic μ	Deterministic C_v
15 000	−14 410 603.5083	3 796.130 07	4 606.231	0.253 0753	10 002.129	23 738.312	0.421 349 631
16 000	−16 396 064.4361	4 049.205 408	4 913.313	0.253 0753	10 002.129	23 738.312	0.421 349 631
17 000	−18 509 619.6173	4 302.280 746	5 220.395	0.253 0753	10 002.129	23 738.312	0.421 349 631
18 000	−20 751 269.0519	4 555.356 084	5 527.477	0.253 0753	10 002.129	23 738.312	0.421 349 631
19 000	−24 054 832.6088	4 904.572 622	5 970.926	0.258 1354	10 002.129	23 738.312	0.421 349 631
20 000	−35 268 148.8569	5 938.699 256	7 348.858	0.296 935	10 002.129	23 738.312	0.421 349 631
21 000	−49 062 075.3823	7 004.432 552	8 726.790	0.333 5444	10 002.129	23 738.312	0.421 349 631
22 000	−65 436 612.1851	8 089.289 968	10 104.723	0.367 695	10 002.129	23 738.312	0.421 349 631
23 000	−84 391 759.2651	9 186.498 749	11 482.655	0.399 413	10 002.129	23 738.312	0.421 349 631
23 387.50	−92 430 619.3808	9 614.084 428	12 016.604	0.411 077 9	10 002.129	23 738.312	0.421 349 631
23 400	−92 696 388.9748	9 627.896 394	12 033.828	0.411 448 6	10 002.129	23 738.312	0.421 349 631
23 500	−94 837 061.6591	9 738.432 197	12 171.621	0.414 401 4	10 002.129	23 738.312	0.421 349 631
23 600	−97 003 540.4462	9 849.037 539	12 309.415	0.417 332 1	10 002.129	23 738.312	0.421 349 631
23 700	−99 195 825.3361	9 959.710 103	12 447.208	0.420 240 9	10 002.129	23 738.312	0.421 349 631
23 730	−99 858 542.9931	9 992.924 647	12 488.546	0.421 109 3	10 002.129	23 738.312	0.421 349 631
23 735	−99 969 221.7394	9 998.460 969	12 495.436	0.421 253 9	10 002.129	23 738.312	0.421 349 631
23 736	−99 991 365.2304	9 999.568 252	12 496.814	0.421 282 8	10 002.129	23 738.312	0.421 349 631
23 737	−100 013 511.3021	10 000.675 54	12 498.191	0.421 311 7	10 002.129	23 738.312	0.421 349 631
23 738	−100 035 659.9544	10 001.782 84	12 499.569	0.421 340 6	10 002.129	23 738.312	0.421 349 631
23 738.50		(infeasible solution)				(infeasible solution)	

Table 9.5. Detailed computational results for Risk Model I for (i) target profit = deterministic profit = $23 387.50/day and (ii) target profit = $23 500/day

First-Stage Variable	Stochastic Solution: Target Profit ($/day) 23 387.50	Stochastic Solution: Target Profit ($/day) 23 500.00
x_1	12 016.604	12 171.621
x_2	1 922.657	1 947.459
x_3	600.830	608.581
x_4	1 802.491	1 825.743
x_5	1 700.000	1 700.000
x_6	5 870.461	5 968.122
x_7	1 562.159	1 582.311
x_8	2 643.653	2 677.757
x_9	2 403.321	2 434.324
x_{10}	3 604.981	3 651.486
x_{11}	961.328	973.730
x_{12}	1 275.000	1 275.000
x_{13}	1 368.653	1 402.757
x_{14}	2 403.321	2 434.324
x_{15}	0	0
x_{16}	961.328	973.730
x_{17}	1 321.826	1 338.878
x_{18}	425.000	425.000
x_{19}	896.826	913.878
x_{20}	120.166	121.716

Table 9.6. Computational statistics for Risk Model I

Solver	Single continuous variables	Constraints	Resource usage/ CPU time (s)	Iterations
CONOPT 3	22	27	≈ (0.01–0.02)	3

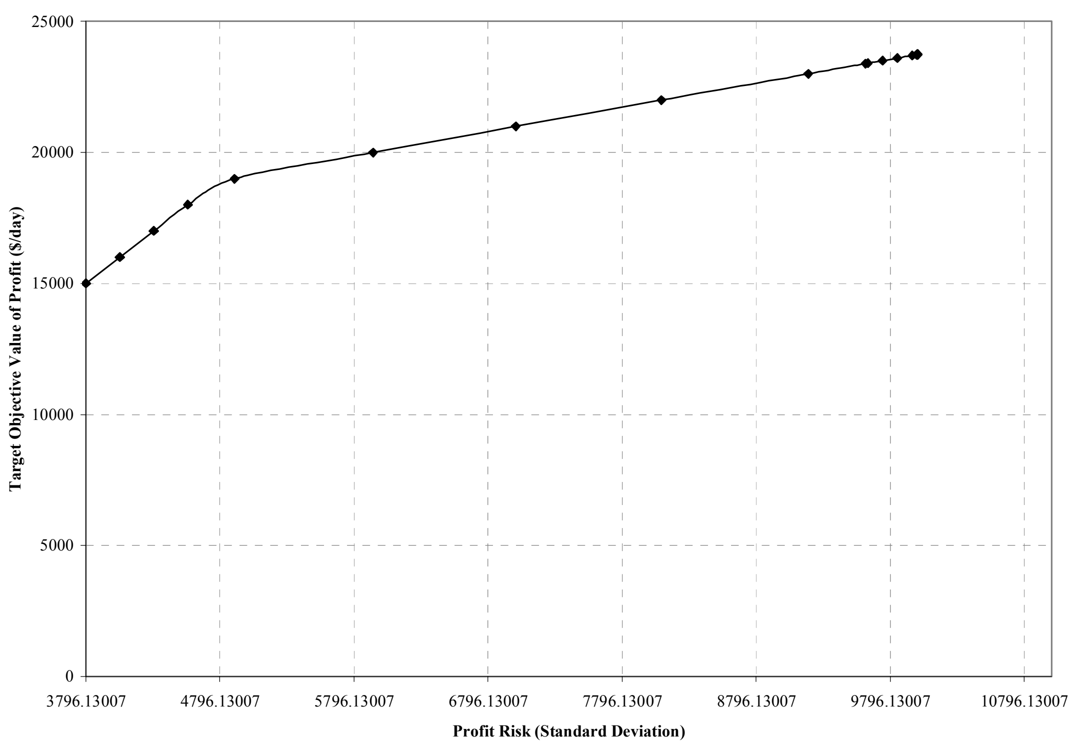


Figure 9.1. The efficient frontier plot of expected profit versus profit risk measured by standard deviation for Risk Model I

9.1.2 Analysis of Results for Risk Model I

The risk–return curve in Figure 8.2 provides a graphical representation of the relation between the expected profit and its associated risk as computed by standard deviation. It depicts the range of possible levels of solution robustness. The trend follows the shape of the efficient frontier proposed by the Markowitz's mean–variance model. Thus, as highlighted by Mulvey et al. (1995), the constructed efficient frontier provides an opportunity for the decision-maker to achieve a robust recommendation, which is not possible by means of traditional sensitivity analysis of the deterministic linear program presented previously.

Note also that it is of interest to know the amount of crude oil to be purchased by a refinery, as computed by the variable x_1, in order to achieve the targeted expected profit. Naturally, higher throughputs of crude oil commensurate with higher profits as this translates to higher production volume, so long as the refining capacity is not exceeded. With availability of information on the current price of crude oil, a decision-maker will be in a good position to assess the trade-off between the raw material cost of purchasing crude oil and the expected profit to be gained from sales of the production volume.

From Table 8.6, for a target profit equivalent to the deterministic profit, Risk Model I computes a crude oil flow rate of 12 016.604 ton/day that is lower than the deterministic model crude oil flow rate of 12 500 ton/day of crude oil. This exemplifies that for a lower raw material purchasing cost for crude oil, the production plan proposed by the stochastic Risk Model I is able to achieve the same amount of profit.

9.2 APPROACH 2: THE EXPECTATION MODELS I AND II

In this stochastic model, it is assumed that there is no alternative source of production and hence, if there is a shortfall in production, the demand is actually lost. Thus, the corresponding model considers the case where the in-house production of the refinery has to be anticipated at the beginning of the planning horizon, that is, the production variables x are fixed (which is essentially the underlying principle in adopting the two-stage stochastic programming framework.), no vendor production is allowed, and all unmet demand is lost.

9.2.1 Two-Stage Stochastic Programming with Fixed Recourse Framework to Model Uncertainty in Product Demand

For the purpose of utilizing the techniques of introducing slack variables and penalty functions in modelling randomness in the RHS coefficients of product demand constraints, consider the constraint for the production requirement of gasoline, x_2 as given by inequality **(8.20)** to be uncertain:

$$x_2 \leq 2700 \tag{8.20}$$

As in the case of price uncertainty, three possible realizations are also equivalently considered for the RHS coefficient random variable of inequality **(8.20)**, with each representing the demand scenario corresponding to the possibility of "average demand", "above average demand", and "below average demand". Details of the scenarios constructed to model uncertainty in market demand for gasoline is depicted in Table 8.8. A five (5) percent standard deviation from the mean value of market demand for gasoline is assumed to be reasonable based on preliminary investigation of available historical data.

Table 9.7. Attributes of the scenario construction example for modelling market demand uncertainty for gasoline

	Demand Uncertainty: Right-Hand-Side Coefficient of Constraints for Gasoline Demand (ton/day)		
	Scenario 1 ($s = 1$)	Scenario 2 ($s = 2$)	Scenario 3 ($s = 3$)
	Above Average Demand	Average Demand	Below Average Demand
Percentage of deviation from the expected value	+5%	0 (i.e., the expected value)	−5%
Gasoline demand (ton/day) $d_{i,s}$	$d_{2,1} = 2835$	$d_{2,2} = 2700$	$d_{2,3} = 2565$
Probability p_s	$p_1 = 0.35$	$p_2 = 0.45$	$p_3 = 0.20$

Again, for the purpose of illustrating the validity and capability of the mathematical programming methods involved and improvised (with less emphasis on the actual feasibility of the data to capture the detailed aspects of the problem), the similar three possible scenarios assumed for price uncertainty are applied to describe uncertainty in the demands of naphtha, jet fuel, heating oil, and fuel oil as given by the maximum production requirements inequalities of **(8.21)** to **(8.24)**, respectively. The resulting Table 8.9 displays the three scenarios constructed for demand uncertainty for the five products considered, with their corresponding probabilities equivalent to the probabilities for the three scenarios generated to model price uncertainty.

Assuming that it costs $c_i^+ = c_2^+$ = \$25 per unit of gasoline to purchase in the open market to meet the production requirement demand if there is a shortfall, and that it costs $c_i^- = c_2^-$ = \$20 per unit of gasoline to be stored in inventory if supply (production) exceeds demand, thus the expected recourse penalty for the second-stage cost due to uncertainty or randomness in gasoline demand is given by:

$$\begin{aligned} E_{s,\text{demand}}^{\text{gasoline}} &= \underbrace{p_1\left(c_2^+ z_{21}^+ + c_2^- z_{21}^-\right)}_{\text{Scenario 1}} + \underbrace{p_2\left(c_2^+ z_{22}^+ + c_2^- z_{22}^-\right)}_{\text{Scenario 2}} + \underbrace{p_3\left(c_2^+ z_{23}^+ + c_2^- z_{23}^-\right)}_{\text{Scenario 3}} \\ &= (0.35)\left(25z_{21}^+ + 20z_{21}^-\right) + (0.45)\left(25z_{22}^+ + 20z_{22}^-\right) + (0.2)\left(25z_{23}^+ + 20z_{23}^-\right) \end{aligned} \quad \textbf{(9.13)}$$

Table 9.8. Representative scenarios of market demand uncertainty in the refinery planning under uncertainty problem

	Demand Uncertainty: Right-Hand Side Coefficient of Constraints (ton/day)		
	Scenario 1	Scenario 2	Scenario 3
Product (type i)	Above Average Demand (+5%)	Average Demand (Expected Value/Mean)	Below Average Demand (−5%)
Gasoline (2)	2835	2700	2565
Naphtha (3)	1155	1100	1045
Jet fuel (4)	2415	2300	2185
Heating oil (5)	1785	1700	1615
Fuel oil (6)	9975	9500	9025
Probability p_s	0.35	0.45	0.2

The penalty costs incurred due to shortfalls and surpluses in production for demand uncertainty in the five products considered are listed in Table 8.10.

Table 9.9. Penalty costs incurred due to shortfalls and surpluses in production under market demand uncertainty

	Penalty cost incurred per unit (\$/ton)	
Product Type (i)	Shortfall in production (c_{is}^{+})	Surplus in production (c_{is}^{-})
Gasoline (2)	25	20
Naphtha (3)	17	13
Jet fuel (4)	5	4
Heating oil (5)	6	5
Fuel oil (6)	10	8

Therefore, the overall expected recourse penalty for the second-stage costs due to uncertainty in market demand as represented by randomness in the right-hand-side coefficients of the related constraints is given by:

$$
\begin{aligned}
E_{s,\text{demand}} = {} & p_1 \begin{bmatrix} \underbrace{\left(c_2^{+} z_{21}^{+} + c_2^{-} z_{21}^{-}\right)}_{\text{gasoline}} + \underbrace{\left(c_3^{+} z_{31}^{+} + c_3^{-} z_{31}^{-}\right)}_{\text{naphtha}} + \underbrace{\left(c_4^{+} z_{41}^{+} + c_4^{-} z_{41}^{-}\right)}_{\text{jet fuel}} \\ + \underbrace{\left(c_5^{+} z_{51}^{+} + c_5^{-} z_{51}^{-}\right)}_{\text{heating oil}} + \underbrace{\left(c_6^{+} z_{61}^{+} + c_6^{-} z_{61}^{-}\right)}_{\text{fuel oil}} \end{bmatrix} \\
& + p_2 \begin{bmatrix} \left(c_2^{+} z_{22}^{+} + c_2^{-} z_{22}^{-}\right) + \left(c_3^{+} z_{32}^{+} + c_3^{-} z_{32}^{-}\right) + \left(c_4^{+} z_{42}^{+} + c_4^{-} z_{42}^{-}\right) \\ + \left(c_5^{+} z_{52}^{+} + c_5^{-} z_{52}^{-}\right) + \left(c_6^{+} z_{62}^{+} + c_6^{-} z_{62}^{-}\right) \end{bmatrix} \\
& + p_3 \begin{bmatrix} \left(c_2^{+} z_{23}^{+} + c_2^{-} z_{23}^{-}\right) + \left(c_3^{+} z_{33}^{-} + c_3^{-} z_{33}^{-}\right) + \left(c_4^{+} z_{43}^{+} + c_4^{-} z_{43}^{-}\right) \\ + \left(c_5^{+} z_{53}^{+} + c_5^{-} z_{53}^{-}\right) + \left(c_6^{+} z_{63}^{+} + c_6^{-} z_{63}^{-}\right) \end{bmatrix}
\end{aligned} \tag{9.14}
$$

Substituting the probabilities and the penalty cost terms with their actual values give:

$$
\begin{aligned}
E_{s,\text{demand}} = {} & (0.35) \begin{bmatrix} \left(25 z_{21}^{+} + 20 z_{21}^{-}\right) + \left(17 z_{31}^{+} + 13 z_{31}^{-}\right) + \left(5 z_{41}^{+} + 4 z_{41}^{-}\right) \\ + \left(6 z_{51}^{+} + 5 z_{51}^{-}\right) + \left(10 z_{61}^{+} + 8 z_{61}^{-}\right) \end{bmatrix} \\
& + (0.45) \begin{bmatrix} \left(25 z_{22}^{+} + 20 z_{22}^{-}\right) + \left(17 z_{32}^{+} + 13 z_{32}^{-}\right) + \left(5 z_{42}^{+} + 4 z_{42}^{-}\right) \\ + \left(6 z_{52}^{+} + 5 z_{52}^{-}\right) + \left(10 z_{62}^{+} + 8 z_{62}^{-}\right) \end{bmatrix} \\
& + (0.2) \begin{bmatrix} \left(25 z_{23}^{+} + 20 z_{23}^{-}\right) + \left(17 z_{33}^{-} + 13 z_{33}^{-}\right) + \left(5 z_{43}^{+} + 4 z_{43}^{-}\right) \\ + \left(6 z_{53}^{+} + 5 z_{53}^{-}\right) + \left(10 z_{63}^{+} + 8 z_{63}^{-}\right) \end{bmatrix}
\end{aligned} \tag{9.15}
$$

The general compact representation for the above is given by:

$$E_{s,\text{demand}} = \sum_{i=2}^{6}\sum_{s=1}^{3} p_s \left(c_i^+ z_{is}^+ + c_i^- z_{is}^-\right) = \sum_{i\in I}\sum_{s\in S} p_s \left(c_i^+ z_{is}^+ + c_i^- z_{is}^-\right),$$
$$i = \{2,3,4,5,6\} \in I_{\text{demand}}^{\text{random}} \subseteq I, s = \{1,2,3\} \in S \quad \textbf{(9.16)}$$

As highlighted in the general stochastic mode development, to ensure that the original information structure associated with the decision process sequence is honoured, three new constraints to model the three scenarios generated for each product with uncertain demand are added to the stochastic model in place of the original deterministic constraint. It is noted that out of the three new constraints, the one representing the "average" scenario, is identical to the deterministic constraint as it models the mean-value constraint. Altogether, this sums up to 3 × 5 = 15 new constraints in place of the five constraints in the deterministic model for those five products. The general form of the new constraints is given by:

$$x_i + z_{is}^+ - z_{is}^- = d_{is}, \quad i = \{2,3,4,5,6\} \in I_{\text{demand}}^{\text{random}} \subseteq I, s = \{1,2,3\} \in S \quad \textbf{(9.17)}$$

For the sake of completeness, the 15 constraints are listed below:
for the demand uncertainty of gasoline:

$$x_2 + z_{21}^+ - z_{21}^- = d_{21} \quad \textbf{(9.18)}$$

$$x_2 + z_{22}^+ - z_{22}^- = d_{22} \quad \textbf{(9.19)}$$

$$x_2 + z_{23}^+ - z_{23}^- = d_{23} \quad \textbf{(9.20)}$$

for the demand uncertainty of naphtha:

$$x_3 + z_{31}^+ - z_{31}^- = d_{31} \quad \textbf{(9.21)}$$

$$x_3 + z_{32}^+ - z_{32}^- = d_{32} \quad \textbf{(9.22)}$$

$$x_3 + z_{33}^{+} - z_{33}^{-} = d_{33} \quad \textbf{(9.23)}$$

for the demand uncertainty of jet fuel:

$$x_4 + z_{41}^{+} - z_{41}^{-} = d_{41} \quad \textbf{(9.24)}$$

$$x_4 + z_{42}^{+} - z_{42}^{-} = d_{42} \quad \textbf{(9.25)}$$

$$x_4 + z_{43}^{+} - z_{43}^{-} = d_{43} \quad \textbf{(9.26)}$$

for the demand uncertainty of heating oil:

$$x_5 + z_{51}^{+} - z_{51}^{-} = d_{51} \quad \textbf{(9.27)}$$

$$x_5 + z_{52}^{+} - z_{52}^{-} = d_{52} \quad \textbf{(9.28)}$$

$$x_5 + z_{53}^{+} - z_{53}^{-} = d_{53} \quad \textbf{(9.29)}$$

for the demand uncertainty of fuel oil:

$$x_6 + z_{61}^{+} - z_{61}^{-} = d_{61} \quad \textbf{(9.30)}$$

$$x_6 + z_{62}^{+} - z_{62}^{-} = d_{62} \quad \textbf{(9.31)}$$

$$x_6 + z_{63}^{+} - z_{63}^{-} = d_{63} \quad \textbf{(9.32)}$$

9.2.2 Two-Stage Stochastic Programming with Fixed Recourse Framework to Model Uncertainty in Product Yields

For the purpose of utilizing the techniques of introducing slack variables and penalty functions in modelling randomness in the LHS technological coefficients of product yields, consider the mass balance given by equation **(3)** for the fixed yield of naphtha from crude oil in the primary distillation unit (PDU):

$$-0.13x_1 + x_7 = 0 \quad \textbf{(8.3)}$$

To be consistent with the case of price and demand uncertainty, three possible scenarios are also considered for the LHS coefficient random variable of x_1 in equation **(8.3)**, that is, 0.13 (the minus sign is used to indicate inlet flow as pointed out in the presentation of the deterministic model), with each scenario corresponding to the depiction of "average product yield", "above average product yield", and "below average product yield". The three scenarios constructed to model uncertainty in yield of naphtha from crude oil in the PDU are detailed in Table 9.10. A five (5) percent deviation from the mean value of naphtha yield is assumed to be reasonable based on preliminary investigation of available historical data. Then, by using a similar approach, all three scenarios accounting for uncertainty in the other product yields from crude oil (besides naphtha), comprising yields of naphtha, jet fuel, gas oil, and cracker feed in the PDU, are tabulated in Table 9.11. As stressed in the previous section general model development, in ensuring that the material balances are satisfied, yields for residuum is determined by subtracting the summation of yields for the other four products from unity. As emphasized also, this does not mispresent the physical meaning of the problem as the yield of the residuum (or residual) is relatively negligible anyway in a typical atmospheric distillation unit.

Table 9.10. Attributes of the scenario construction example for modelling uncertainty in yield of naphtha from crude oil in the primary distillation unit

	Yield Uncertainty: Left-Hand Side Coefficient of Mass Balance for Naphtha Yield from Crude Oil (unitless)		
	Scenario 1 ($s = 1$)	Scenario 2 ($s = 2$)	Scenario 3 ($s = 3$)
	Above Average Yield	Average Yield	Below Average Yield
Percentage of deviation from the expected value	+10%	0 (i.e., the expected value or mean itself)	–10%
Yield of naphtha from crude oil	–0.143	–0.13	–0.117
Probability p_s	$p_1 = 0.35$	$p_2 = 0.45$	$p_3 = 0.20$

Table 9.11. Representative scenarios of uncertainty in product yields from crude oil in the primary distillation unit for the refinery planning under uncertainty problem

	Yield Uncertainty: Left-Hand Side Coefficient of Mass Balances for Fixed Yields (unitless)		
	Scenario 1 ($s = 1$)	Scenario 2 ($s = 2$)	Scenario 3 ($s = 3$)
Product Type (i)	Above Average Yield (+1%)	Average Yield (Expected Value/Mean)	Below Average Yield (–1%)
Naphtha (x_7)	–0.1365	–0.13	–0.1235
Jet fuel (x_4)	–0.1575	–0.15	–0.1425
Gas oil (x_8)	–0.231	–0.22	–0.209
Cracker feed (x_9)	–0.21	–0.20	–0.19
Residuum (x_{10})	–0.265	–0.30	–0.335
Probability p_s	0.35	0.45	0.20

It is further assumed that a penalty cost of q_i^+ is incurred per unit of shortage of crude oil yields $y_{i,s}^+$ and a penalty cost of q_i^- for excess of crude oil yields $y_{i,s}^-$. Thus, the expected recourse penalty for the second-stage cost due to uncertainty or randomness in crude oil yield to naphtha is given by:

$$E_{s,\text{yield}}^{\text{naphtha}} = \underbrace{p_1\left(q_2^+ y_{21}^+ + q_2^- y_{21}^-\right)}_{\text{Scenario 1}} + \underbrace{p_2\left(q_2^+ y_{22}^+ + q_2^- y_{22}^-\right)}_{\text{Scenario 2}} + \underbrace{p_3\left(q_2^+ y_{23}^+ + q_2^- y_{23}^-\right)}_{\text{Scenario 3}} \quad \textbf{(9.33)}$$

The associated penalty costs incurred due to deviations in product yields from crude oil are assumed to be as depicted in Table 9.12.

Table 9.12. Penalty costs incurred due to uncertainty in product yields from crude oil

	Cost incurred per unit deviation ($/unit)	
Product Type (i)	Yield decrement (q_{is}^+)	Yield increment (q_{is}^-)
Naphtha (3)	5	3
Jet fuel (4)	5	4
Gas oil (x_8)	5	3
Cracker feed (x_9)	5	3
Residuum (x_{10})	5	3

Therefore, the expected recourse penalty for the second-stage costs due to uncertainty in product yields from crude oil as represented by randomness in the left-hand-side coefficients of the mass balances for fixed yields is given by:

$$E_{s,\text{yield}} = p_1 \begin{bmatrix} \underbrace{\left(q_3^+ y_{31}^+ + q_3^- y_{31}^-\right)}_{\text{naphtha}} + \underbrace{\left(q_4^+ y_{41}^+ + q_4^- y_{41}^-\right)}_{\text{jet fuel}} + \underbrace{\left(q_8^+ y_{81}^+ + q_8^- y_{81}^-\right)}_{\text{gas oil}} \\ + \underbrace{\left(q_9^+ y_{91}^+ + q_9^- y_{91}^-\right)}_{\text{cracker feed}} + \underbrace{\left(q_{10}^+ y_{10,1}^+ + q_{10}^- y_{10,1}^-\right)}_{\text{residuum}} \end{bmatrix}$$

$$+ p_2 \begin{bmatrix} \left(q_3^+ y_{32}^+ + q_3^- y_{32}^-\right) + \left(q_4^+ y_{42}^+ + q_4^- y_{42}^-\right) + \left(q_8^+ y_{82}^+ + q_8^- y_{82}^-\right) \\ + \left(q_9^+ y_{92}^+ + q_9^- y_{92}^-\right) + \left(q_{10}^+ y_{10,2}^+ + q_{10}^- y_{10,2}^-\right) \end{bmatrix}$$

$$+ p_3 \begin{bmatrix} \left(q_3^+ y_{33}^+ + q_3^- y_{33}^-\right) + \left(q_4^+ y_{43}^+ + q_4^- y_{43}^-\right) + \left(q_8^+ y_{83}^+ + q_8^- y_{83}^-\right) \\ + \left(q_9^+ y_{93}^+ + q_9^- y_{93}^-\right) + \left(q_{10}^+ y_{10,3}^+ + q_{10}^- y_{10,3}^-\right) \end{bmatrix} \quad \textbf{(9.34)}$$

Substituting the probabilities and the penalty cost terms with their actual values gives:

$$E_{s,\text{yield}} = (0.35) \begin{bmatrix} \left(5y_{21}^+ + 3y_{21}^-\right) + \left(5_{31}^+ + 4y_{31}^-\right) + \left(5y_{41}^+ + 3y_{41}^-\right) \\ + \left(5y_{51}^+ + 3y_{51}^-\right) + \left(5y_{61}^+ + 3y_{61}^-\right) \end{bmatrix}$$

$$+ (0.45) \begin{bmatrix} \left(5y_{22}^+ + 3y_{22}^-\right) + \left(5y_{32}^+ + 4y_{32}^-\right) + \left(5y_{42}^+ + 3y_{42}^-\right) \\ + \left(5y_{52}^+ + 3y_{52}^-\right) + \left(5y_{62}^+ + 3y_{62}^-\right) \end{bmatrix} \quad \textbf{(9.35)}$$

$$+ (0.2) \begin{bmatrix} \left(5y_{23}^+ + 3y_{23}^-\right) + \left(5y_{33}^- + 4y_{33}^-\right) + \left(5y_{43}^+ + 3y_{43}^-\right) \\ + \left(5y_{53}^+ + 3y_{53}^-\right) + \left(5y_{63}^+ + 3y_{63}^-\right) \end{bmatrix}$$

The general compact representation for the above is thus given by:

$$E_{s,\text{yield}} = \sum_{i=3}^{10} \sum_{s=1}^{3} p_s \left(q_i^+ y_{i,s}^+ + q_i^- y_{i,s}^-\right) = \sum_{i \in I} \sum_{s \in S} p_s \left(q_i^+ y_{i,s}^+ + q_i^- y_{i,s}^-\right),$$
$$i = \{3,4,8,9,10\} \in I_{\text{yield}}^{\text{random}} \subseteq I, s = \{1,2,3\} \in S \quad \textbf{(9.36)}$$

To ensure that the original information structure associated with the decision process sequence is honoured, three new constraints to account for the three scenarios dealing with product yields uncertainty from crude oil are introduced for each affected product. The general form of the constraints is:

$$T_i x_1 + x_i + y_{i,s}^+ - y_{is}^- = 0, \quad i = \{3,4,8,9,10\} \in I_{\text{yield}}^{\text{random}} \subseteq I, s = \{1,2,3\} \in S \qquad \textbf{(9.37)}$$

For example, the three new constraints for the uncertainty in yield of naphtha from crude oil is given by:

$$-0.143x_1 + x_7 + y_{31}^+ - y_{31}^- = 0 \qquad \textbf{(9.38)}$$

$$-0.13x_1 + x_7 + y_{32}^+ - y_{32}^- = 0 \qquad \textbf{(9.39)}$$

$$-0.117x_1 + x_7 + y_{33}^+ - y_{33}^- = 0 \qquad \textbf{(9.40)}$$

Again, for the purpose of completeness, the entire set of three new constraints for every product yield uncertainty from crude oil for the other products comprising jet fuel, gas oil, cracker feed, and residuum is listed below:
for the randomness in jet fuel yield from crude oil:

$$-0.165x_1 + x_4 + y_{41}^+ - y_{41}^- = 0 \qquad \textbf{(9.41)}$$

$$-0.15x_1 + x_4 + y_{42}^+ - y_{42}^- = 0 \qquad \textbf{(9.42)}$$

$$-0.135x_1 + x_4 + y_{43}^+ - y_{43}^- = 0 \qquad \textbf{(9.43)}$$

for the randomness in gas oil yield from crude oil:

$$-0.242x_1 + x_8 + y_{81}^+ - y_{81}^- = 0 \qquad \textbf{(9.44)}$$

$$-0.22x_1 + x_8 + y_{82}^+ - y_{82}^- = 0 \qquad \textbf{(9.45)}$$

$$-0.198x_1 + x_8 + y_{83}^+ - y_{83}^- = 0 \qquad \textbf{(9.46)}$$

for the randomness in cracker feed yield from crude oil:

$$-0.22x_1 + x_9 + y_{91}^+ - y_{91}^- = 0 \qquad \textbf{(9.47)}$$

$$-0.20x_1 + x_9 + y_{92}^+ - y_{92}^- = 0 \qquad \textbf{(9.48)}$$

$$-0.18x_1 + x_9 + y_{93}^+ - y_{93}^- = 0 \qquad \textbf{(9.49)}$$

for the randomness in residuum yield from crude oil:

$$-0.33x_1 + x_{10} + y_{10,1}^+ - y_{10,1}^- = 0 \qquad \textbf{(9.50)}$$

$$-0.30x_1 + x_{10} + y_{10,2}^+ - y_{10,2}^- = 0 \qquad \textbf{(9.51)}$$

$$-0.27x_1 + x_{10} + y_{10,3}^+ - y_{10,3}^- = 0 \qquad \textbf{(9.52)}$$

The complete scenario formulation to simultaneously handle uncertainties in commodity prices, product demand, and product yields is given by the combination of Tables 9.2, 9.8, and 9.11. As stressed in the general stochastic model development, the major assumption that enables the combination of the sub-scenarios is that demands, yields, and prices in each sub-scenario are highly-correlated. This means that for instance, the possibility of the scenario where prices are "average" with demands being "above average" and yields being "below average" (or any other combination of sub-scenarios) is not considered. The combined tables are presented as Table 9.13.

Table 9.13. Complete scenario formulation for the refinery production planning under uncertainty in commodity prices, market demands for products, and product yields problem

Product Type (i)	Scenario 1 Above Average	Scenario 2 Average (Expected Value/Mean)	Scenario 3 Below Average
Price Uncertainty: Objective Function Coefficient of Prices ($/day)			
Crude oil (1)	–8.8	–8.0	–7.2
Gasoline (2)	20.35	18.5	16.65
Naphtha (3)	8.8	8.0	7.2
Jet fuel (4)	13.75	12.5	11.25
Heating oil (5)	15.95	14.5	13.05
Fuel oil (6)	6.6	6.0	5.4
Cracker feed (14)	–1.65	–1.5	–1.35
Demand Uncertainty: Right-Hand-Side Coefficient of Constraints for Production Requirement (ton/day)			
Gasoline (2)	2835	2700	2565
Naphtha (3)	1155	1100	1045
Jet fuel (4)	2415	2300	2185
Heating oil (5)	1785	1700	1615
Fuel oil (6)	9975	9500	9025
Yield Uncertainty: Left-Hand Side Coefficient of Mass Balances for Fixed Yields (unitless)			
Naphtha (7)	–0.1365	–0.13	–0.1235
Jet fuel (4)	–0.1575	–0.15	–0.1425
Gas oil (8)	–0.231	–0.22	–0.209
Cracker feed (9)	–0.21	–0.20	–0.19
Residuum (10)	–0.265	–0.30	–0.335
Probability p_s	0.35	0.45	0.20

The corresponding expected recourse penalty for the second-stage costs due to uncertainties in both demands and yields is:

$$E_{s'} = E_{s,\text{demand}} + E_{s,\text{yield}} = p_1\xi_1 + p_2\xi_2 + p_3\xi_3$$

$$\begin{aligned}
E_{s'} = {} & (0.35)\left\{ \begin{array}{l} \left[\begin{array}{l} \left(25z_{21}^{+} + 20z_{21}^{-}\right) + \left(17z_{31}^{+} + 13z_{31}^{-}\right) + \left(5z_{41}^{+} + 4z_{41}^{-}\right) \\ + \left(6z_{51}^{+} + 5z_{51}^{-}\right) + \left(10z_{61}^{+} + 8z_{61}^{-}\right) \end{array} \right] \\ + \left[\begin{array}{l} \left(5y_{31}^{+} + 3y_{31}^{-}\right) + \left(5_{41}^{+} + 4y_{41}^{-}\right) + \left(5y_{81}^{+} + 3y_{81}^{-}\right) \\ + \left(5y_{91}^{+} + 3y_{91}^{-}\right) + \left(5y_{10,1}^{+} + 3y_{10,1}^{-}\right) \end{array} \right] \end{array} \right\} \\
& + (0.45)\left\{ \begin{array}{l} \left[\begin{array}{l} \left(25z_{22}^{+} + 20z_{22}^{-}\right) + \left(17z_{32}^{+} + 13z_{32}^{-}\right) + \left(5z_{42}^{+} + 4z_{42}^{-}\right) \\ + \left(6z_{52}^{+} + 5z_{52}^{-}\right) + \left(10z_{62}^{+} + 8z_{62}^{-}\right) \end{array} \right] \\ + \left[\begin{array}{l} \left(5y_{32}^{+} + 3y_{32}^{-}\right) + \left(5y_{42}^{+} + 4y_{42}^{-}\right) + \left(5y_{82}^{+} + 3y_{82}^{-}\right) \\ + \left(5y_{92}^{+} + 3y_{92}^{-}\right) + \left(5y_{10,2}^{+} + 3y_{10,2}^{-}\right) \end{array} \right] \end{array} \right\} \\
& + (0.2)\left\{ \begin{array}{l} \left[\begin{array}{l} \left(25z_{23}^{+} + 20z_{23}^{-}\right) + \left(17z_{33}^{-} + 13z_{33}^{-}\right) + \left(5z_{43}^{+} + 4z_{43}^{-}\right) \\ + \left(6z_{53}^{+} + 5z_{53}^{-}\right) + \left(10z_{63}^{+} + 8z_{63}^{-}\right) \end{array} \right] \\ + \left[\begin{array}{l} \left(5y_{33}^{+} + 3y_{33}^{-}\right) + \left(5y_{43}^{-} + 4y_{43}^{-}\right) + \left(5y_{83}^{+} + 3y_{83}^{-}\right) \\ + \left(5y_{93}^{+} + 3y_{93}^{-}\right) + \left(5y_{10,3}^{+} + 3y_{10,3}^{-}\right) \end{array} \right] \end{array} \right\}
\end{aligned} \quad \textbf{(9.53)}$$

The general compact representation for the above is given by:

$$E_{s'} = \sum_{i \in I} \sum_{s \in S} p_s \left[\left(c_i^{+} z_{i,s}^{+} + c_i^{-} z_{i,s}^{-} \right) + \left(q_i^{+} y_{i,s}^{+} + q_i^{-} y_{i,s}^{-} \right) \right] \quad \textbf{(9.54)}$$

Thus, **Expectation Model I** is represented as the following:

$$\text{maximize } z_2 = E[z_0] - \theta_1 V(z_0) - E_{s'}$$

$$= \sum_{i \in I} \sum_{s \in S} p_s C_i x_i - \theta_1 \sum_{i \in I} x_i^2 V(C_i) - \sum_{i \in I} \sum_{s \in S} p_s \left[\left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^- \right) + \left(q_i^+ y_{i,s}^+ + q_i^- y_{i,s}^- \right) \right]$$

$$\text{maximize } z_2 = \begin{bmatrix} (0.35)(-8.8x_1 + 20.35x_2 + 8.8x_3 + 13.75x_4 + 15.95x_5 + 6.6x_6 - 1.65x_{14}) \\ +(0.45)(-8.0x_1 + 18.5x_2 + 8.0x_3 + 12.5x_4 + 14.5x_5 + 6.0x_6 - 1.5x_{14}) \\ +(0.2)(-7.2x_1 + 16.65x_2 + 7.2x_3 + 11.25x_4 + 13.05x_5 + 5.4x_6 - 1.35x_{14}) \end{bmatrix}$$

$$-\theta_1 \begin{bmatrix} \left(\frac{32}{75}\right) x_1^2 + \left(\frac{1369}{600}\right) x_2^2 + \left(\frac{32}{75}\right) x_3^2 + \left(\frac{25}{24}\right) x_4^2 + \left(\frac{841}{600}\right) x_5^2 + (0.24)\, x_6^2 \\ +(0.015)\, x_{14}^2 \end{bmatrix}$$

$$-\begin{bmatrix} (0.35)\left\{ \begin{matrix} \begin{bmatrix} \left(25z_{21}^+ + 20z_{21}^-\right) + \left(17z_{31}^+ + 13z_{31}^-\right) + \left(5z_{41}^+ + 4z_{41}^-\right) \\ +\left(6z_{51}^+ + 5z_{51}^-\right) + \left(10z_{61}^+ + 8z_{61}^-\right) \end{bmatrix} \\ + \begin{bmatrix} \left(5y_{31}^+ + 3y_{31}^-\right) + \left(5_{41}^+ + 4y_{41}^-\right) + \left(5y_{81}^+ + 3y_{81}^-\right) \\ +\left(5y_{91}^+ + 3y_{91}^-\right) + \left(5y_{10,1}^+ + 3y_{10,1}^-\right) \end{bmatrix} \end{matrix} \right\} \\ +(0.45)\left\{ \begin{matrix} \begin{bmatrix} \left(25z_{22}^+ + 20z_{22}^-\right) + \left(17z_{32}^+ + 13z_{32}^-\right) + \left(5z_{42}^+ + 4z_{42}^-\right) \\ +\left(6z_{52}^+ + 5z_{52}^-\right) + \left(10z_{62}^+ + 8z_{62}^-\right) \end{bmatrix} \\ + \begin{bmatrix} \left(5y_{32}^+ + 3y_{32}^-\right) + \left(5y_{42}^+ + 4y_{42}^-\right) + \left(5y_{82}^+ + 3y_{82}^-\right) \\ +\left(5y_{92}^+ + 3y_{92}^-\right) + \left(5y_{10,2}^+ + 3y_{10,2}^-\right) \end{bmatrix} \end{matrix} \right\} \\ +(0.2)\left\{ \begin{matrix} \begin{bmatrix} \left(25z_{23}^+ + 20z_{23}^-\right) + \left(17z_{33}^- + 13z_{33}^-\right) + \left(5z_{43}^+ + 4z_{43}^-\right) \\ +\left(6z_{53}^+ + 5z_{53}^-\right) + \left(10z_{63}^+ + 8z_{63}^-\right) \end{bmatrix} \\ + \begin{bmatrix} \left(5y_{33}^+ + 3y_{33}^-\right) + \left(5y_{43}^- + 4y_{43}^-\right) + \left(5y_{83}^+ + 3y_{83}^-\right) \\ +\left(5y_{93}^+ + 3y_{93}^-\right) + \left(5y_{10,3}^+ + 3y_{10,3}^-\right) \end{bmatrix} \end{matrix} \right\} \end{bmatrix}$$

(9.55)

s.t. deterministic constraints (first stage) **(8.1)**, **(8.2)**, **(8.8)–(8.19)**, and **(8.26)**,
stochastic constraints (second stage): **(9.17)–(9.32)** and **(9.37)–(9.52)**.

The alternative **Expectation Model II** is given as follows:

$$\text{maximize } z_2 = -V(z_0) - E_{s'}$$

$$= -\theta_1 \sum_{i\in I} x_i^2 V(C_i) - \sum_{i\in I}\sum_{s\in S} p_s \left[\left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^- \right) + \left(q_i^+ y_{i,s}^+ + q_i^- y_{i,s}^- \right) \right]$$

$$\text{maximize } z_2 = -\begin{bmatrix} \left(\frac{32}{75}\right)x_1^2 + \left(\frac{1369}{600}\right)x_2^2 + \left(\frac{32}{75}\right)x_3^2 + \left(\frac{25}{24}\right)x_4^2 + \left(\frac{841}{600}\right)x_5^2 + (0.24)x_6^2 \\ +(0.015)x_{14}^2 \end{bmatrix}$$

$$-\begin{bmatrix} (0.35)\left\{ \begin{bmatrix} \left(25z_{21}^+ + 20z_{21}^-\right) + \left(17z_{31}^+ + 13z_{31}^-\right) + \left(5z_{41}^+ + 4z_{41}^-\right) \\ +\left(6z_{51}^+ + 5z_{51}^-\right) + \left(10z_{61}^+ + 8z_{61}^-\right) \end{bmatrix} + \begin{bmatrix} \left(5y_{31}^+ + 3y_{31}^-\right) + \left(5_{41}^+ + 4y_{41}^-\right) + \left(5y_{81}^+ + 3y_{81}^-\right) \\ +\left(5y_{91}^+ + 3y_{91}^-\right) + \left(5y_{10,1}^+ + 3y_{10,1}^-\right) \end{bmatrix} \right\} \\ +(0.45)\left\{ \begin{bmatrix} \left(25z_{22}^+ + 20z_{22}^-\right) + \left(17z_{32}^+ + 13z_{32}^-\right) + \left(5z_{42}^+ + 4z_{42}^-\right) \\ +\left(6z_{52}^+ + 5z_{52}^-\right) + \left(10z_{62}^+ + 8z_{62}^-\right) \end{bmatrix} + \begin{bmatrix} \left(5y_{32}^+ + 3y_{32}^-\right) + \left(5y_{42}^+ + 4y_{42}^-\right) + \left(5y_{82}^+ + 3y_{82}^-\right) \\ +\left(5y_{92}^+ + 3y_{92}^-\right) + \left(5y_{10,2}^+ + 3y_{10,2}^-\right) \end{bmatrix} \right\} \\ +(0.2)\left\{ \begin{bmatrix} \left(25z_{23}^+ + 20z_{23}^-\right) + \left(17z_{33}^- + 13z_{33}^-\right) + \left(5z_{43}^+ + 4z_{43}^-\right) \\ +\left(6z_{53}^+ + 5z_{53}^-\right) + \left(10z_{63}^+ + 8z_{63}^-\right) \end{bmatrix} + \begin{bmatrix} \left(5y_{33}^+ + 3y_{33}^-\right) + \left(5y_{43}^- + 4y_{43}^-\right) + \left(5y_{83}^+ + 3y_{83}^-\right) \\ +\left(5y_{93}^+ + 3y_{93}^-\right) + \left(5y_{10,3}^+ + 3y_{10,3}^-\right) \end{bmatrix} \right\} \end{bmatrix} \quad \textbf{(9.56)}$$

s.t

$$E[z_0] = \begin{bmatrix} (0.35)\left(-8.8x_1 + 20.35x_2 + 8.8x_3 + 13.75x_4 + 15.95x_5 + 6.6x_6 - 1.65x_{14}\right) \\ +(0.45)\left(-8.0x_1 + 18.5x_2 + 8.0x_3 + 12.5x_4 + 14.5x_5 + 6.0x_6 - 1.5x_{14}\right) \\ +(0.2)\left(-7.2x_1 + 16.65x_2 + 7.2x_3 + 11.25x_4 + 13.05x_5 + 5.4x_6 - 1.35x_{14}\right) \end{bmatrix} \geq \begin{matrix}\text{Target} \\ \text{objective} \\ \text{function value}\end{matrix}$$

deterministic constraints (first stage) **(8.1)**, **(8.2)**, **(8.8)–(8.19)**, and **(8.26)**,

stochastic constraints (second stage): **(9.17)–(9.32)** and **(9.37)–(9.52)**.

9.2.3 Computational Results for Expectation Model I

Table 9.14 tabulates the computational results and some analytical outputs for the implementation of Expectation Model I on GAMS for a range of values of the profit risk parameter θ_1. Starting values of the first-stage deterministic decision variables have been initialized to the optimal solutions of the deterministic model. A representative detailed results for profit risk factor $\theta_1 = 0.000\ 03$ is presented Table 9.15. The problem size and the distribution of computational expense are noted in the ensuing Table 9.16. Figure 9.2 then depicts the efficient frontier plot of expected profit versus profit risk as measured by variance while an alternative representation of the computed results by plotting expected profit against the profit risk factor θ_1 (also with variance as the risk measure) is shown in Figure 9.3.

Note that the actual true expected profit that is of interest in a stochastic model is still the original equation or expression for the deterministic profit as given by equation **(9.4)**. This fact extends to all other stochastic models as well.

(*Please turn the page over for* Table 9.14.)

Table 9.14. Computational results for Expectation Model I

Profit risk factor θ_1	Optimal objective value	Expected deviation in profit $V(z_0)$	Expected total unmet demand/ production shortfall	Expected total excess production/ production surplus	Expected recourse penalty costs E_s	$\sigma = \sqrt{V(z_0)}$	Expected profit $E[z_0]$	$\mu = \lvert E[z_0] - E_s \rvert$	$C_v = \frac{\sigma}{\mu}$ Stochastic	$C_v = \frac{\sigma}{\mu}$ Deterministic
0	27 717.869	1.3352E+8	2 575.000	26 815.000	54 056.875	11 555.247	81 774.744	27 717.869	0.416 888	0.368 830 774
0.000 000 000 01	27 717.867	1.3352E+8	2 575.000	26 815.000	54 056.875	11 555.247	81 774.744	27 717.869	0.416 888	0.368 830 774
0.000 000 000 1	27 717.855	1.3352E+8	2 575.000	26 815.000	54 056.875	11 555.247	81 774.744	27 717.869	0.416 888	0.368 830 774
0.000 000 001	27 717.735	1.3352E+8	2 575.000	26 815.000	54 056.875	11 555.247	81 774.744	27 717.869	0.416 888	0.368 830 774
0.000 000 01	27 716.534	1.3352E+8	2 575.000	26 815.000	54 056.875	11 555.247	81 774.744	27 717.869	0.416 888	0.368 830 774
0.000 000 1	27 704.516	1.3352E+8	2 575.000	26 815.000	54 056.875	11 555.247	81 774.744	27 717.869	0.416 888	0.368 830 774
0.000 001	27 584.345	1.3352E+8	2 575.000	26 815.000	54 056.875	11 555.247	81 774.744	27 717.869	0.416 888	0.368 830 774
0.000 01	26 382.631	1.3352E+8	2 575.000	26 815.000	54 056.875	11 555.247	81 774.744	27 717.869	0.416 888	0.368 830 774
0.000 015	25 715.013	1.3352E+8	2 575.000	26 815.000	54 056.875	11 555.247	81 774.744	27 717.869	0.416 888	0.368 830 774
0.000 02	25 047.394	1.3352E+8	2 575.000	26 815.000	54 056.875	11 555.247	81 774.744	27 717.869	0.416 888	0.368 830 774
0.000 025	24 379.776	1.3352E+8	2 575.000	26 815.000	54 056.875	11 555.247	81 774.744	27 717.869	0.416 888	0.368 830 774
0.000 03	23 712.157	1.3352E+8	2 575.000	26 815.000	54 056.875	11 555.247	81 774.744	27 717.869	0.416 888	0.368 830 774
0.000 05	21 041.682	1.3352E+8	2 575.000	26 815.000	54 056.875	11 555.247	81 774.744	27 717.869	0.416 888	0.368 830 774
0.000 1	14 365.496	1.3352E+8	2 575.000	26 815.000	54 056.875	11 555.247	81 774.744	27 717.869	0.416 888	0.368 830 774
0.000 2	2 828.089	9.6341E+7	2 575.000	17 800.529	40232.400	9 815.345	62 328.689	22 096.289	0.444 207 8	0.368 830 774
0.000 3	–6 806.011	9.6341E+7	2 575.000	17 800.529	40232.400	9 815.345	62 328.689	22 096.289	0.444 207 8	0.368 830 774
0.000 5	–2.527E+4	8.8371E+7	3 610.000	16 373.039	42198.208	9 400.606	61 114.658	18 916.450	0.496 954	0.368 830 774
0.001	–6.172E+4	5.5505E+7	11 880.928	12 529.639	64716.008	7 450.150	58 505.228	6 210.780	1.199 551 4	0.368 830 774
0.002	–1.131E+5	4.4503E+7	16 299.595	10 814.676	78476.371	6 671.022	54 345.834	–2.413E+4	0.276 461 7	0.368 830 774
0.003	–1.434E+5	2.0328E+7	27 971.940	7 268.856	1.2045E+5	4 508.663	37 998.995	8.245E+4	0.054 683 6	0.368 830 774
0.004	–1.587E+5	1.1435E+7	33 953.955	5 451.642	1.4144E+5	3 381.497	28 499.246	–1.129E+5	0.029 951 3	0.368 830 774
0.004 5	–1.638E+5	9.0347E+6	35 947.960	4 845.904	1.4844E+5	3 005.775	25 332.663	–1.231E+5	0.024 417 3	0.368 830 774
0.005	–1.678E+5	7.3181E+6	37 543.164	4 361.314	1.5404E+5	2 705.198	22 799.397	–1.312E+5	0.020 618 9	0.368 830 774
0.01	–1.861E+5	1.8295E+6	44 721.582	2 180.657	1.7923E+5	1 352.599	11 399.698	1.678E+5	0.008 060 8	0.368 830 774

Table 9.15. Detailed computational results for Expectation Model I for θ_1 = 0.000 03

First-Stage Variable	Stochastic Solution	Product (i)	Production Shortfall z_{ij}^+ or Surplus z_{ij}^- (t/d) Scenario 1 z_{i1}^+	z_{i1}^-	Scenario 2 z_{i2}^+	z_{i2}^-	Scenario 3 z_{i3}^+	z_{i3}^-
x_1	15 000.000	Demands RHS Coefficients Randomness						
x_2	2 000.000	Gasoline (2)	835.000	0	700.000	0	565.000	0
x_3	1 155.000	Naphtha (3)	0	0	0	55.000	0	110.000
x_4	3 637.500	Jet Fuel (4)	0	1222.500	0	1337.500	0	1452.500
x_5	3 835.000	Heating Oil (5)	0	2050.000	0.00	2135.000	0	2220.000
x_6	9 500.000	Fuel Oil (6)	475.000	0	0	0	0	475.000
x_7	2 155.000							
x_8	4 635.000	Production Yields LHS Coefficients Randomness						
x_9	4 350.000	Naphtha (7)	0	107.500	0	205.000	0	302.500
x_{10}	5 475.000	Jet Fuel (4)	0	1275.000	0	1387.500	0	1500.000
x_{11}	1 000.000	Gas Oil (8)	0	1170.000	0	1335.000	0	1500.000
x_{12}	2 876.250	Cracker Feed (9)	0	1200.000	0	1350.000	0	1500.000
x_{13}	1 758.750	Residuum (10)	0	1500.000	0	975.000	0	450.000
x_{14}	2 500.000							
x_{15}	1 850.000	E(Penalty Costs)		20 229.125		23 123.250		10 704.500
x_{16}	1 000.000	E_{total}	54 056.875					
x_{17}	1 375.000							
x_{18}	958.750							
x_{19}	416.250							
x_{20}	125.000							
Expected Profit z_0 (\$/day)	81 774.744							

Table 9.16. Computational statistics for Expectation Model I

Solver	Single continuous variables	Constraints	Resource usage/ CPU time (s)	Iterations
CONOPT 3	91	85	≈ (0.03–0.11)	10

9.2.4 Analysis of Results for Expectation Model I

Note that since this is a profit maximization problem, larger values of the risk factor θ_1 correspond to lower profits, in contrast with the general notion that higher profits are associated with higher risks, which is typically the case in cost minimization problems. The difference arise since a profit maximization problem is the negative of a cost minimization problem, hence the observed reverse in the trend of the relationship between risk (as computed by variance for this model) and expected profit.

It is observed that reducing values of θ_1 translate to increment in the expected profit, since this generally leads to a reduction in expected production shortfalls but with increasing production surpluses, in which the fixed penalty cost for the former is higher than the latter. Thus, this reflects high model feasibility although not in the absolute sense since there is increase in excess production. Nonetheless, this shows that a suitable operating range of θ_1 values ought to be selected in order to achieve optimality between expected profit and expected production feasibility. However, this observation appears to be somewhat contradictory to Mulvey et al. (1991) who reported that the more robustness desired, the higher is the cost or the lower is the profit.

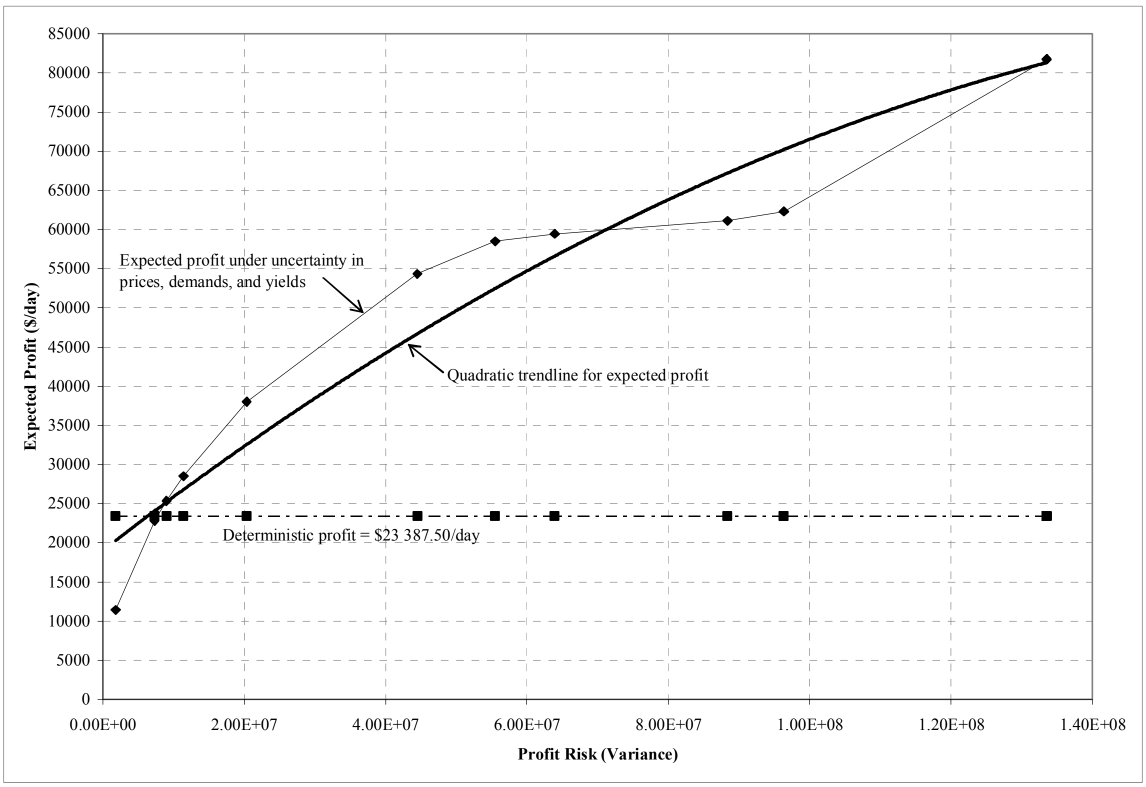


Figure 9.2. The efficient frontier plot of expected profit versus profit risk measured by variance for Expectation Model I

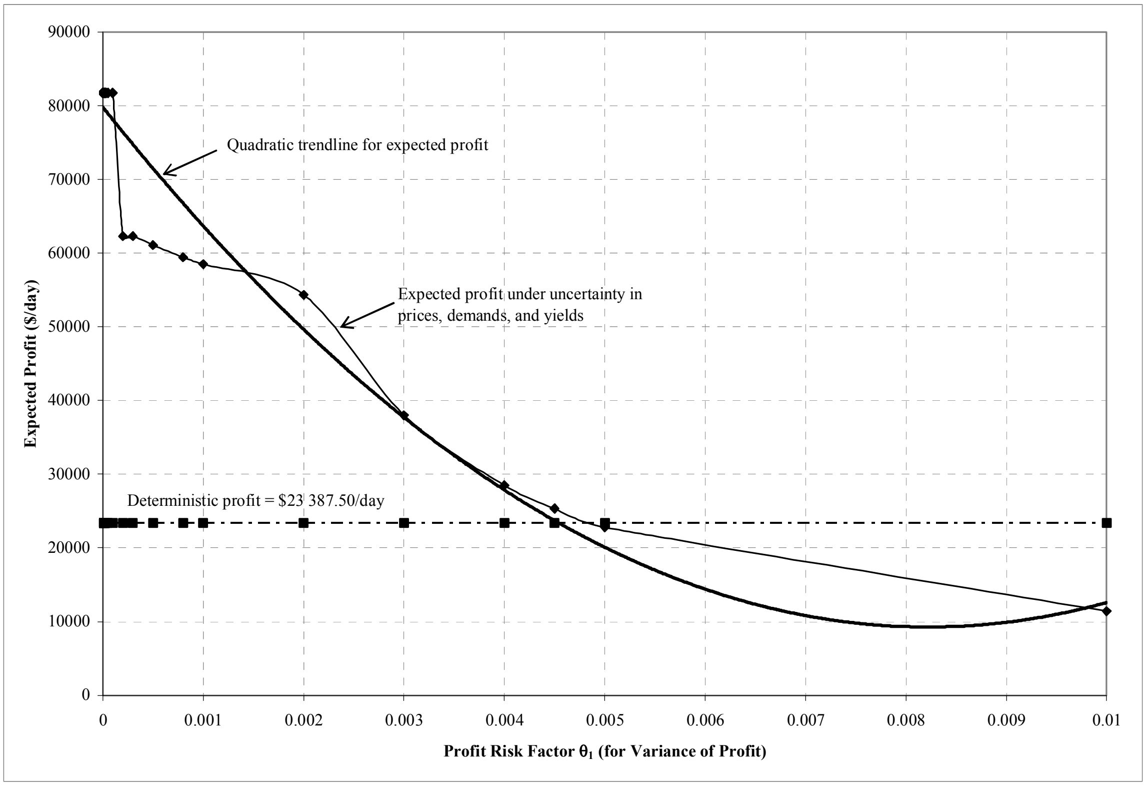


Figure 9.3. Plot of expected profit for different levels of risk as represented by the profit risk factor θ_1 with variance as the risk measure for Expectation Model I

9.2.5 Computational Results for Expectation Model II

Table 9.17 tabulates the computational results for the implementation of Expectation Model II on GAMS for a range of values of the target profit μ. Starting values of the first-stage deterministic decision variables have been initialized to the optimal solutions of the deterministic model. As in Risk Model I, a similar analytical procedure is adopted for Expectation Model II, in which the standard deviation σ of profit is determined by computing the square root of the absolute values of variance, obviating the negative sign.

Representative detailed results for the target profit equals to the deterministic profit of $23 387.50/day is shown in Table 9.18 that immediately follows. Table 9.19 then displays the associated problem size and the distribution of computational expense. Figure 9.4 is plotted to show the efficient frontier plot of expected maximum profit for different levels of risk as represented by the profit risk parameter θ_1 with standard deviation as the risk measure.

Table 9.17. Computational results for Expectation Model II

Target profit $E[z_0]$ (\$/day)	Optimal objective value $-V$	Standard deviation σ ($\sqrt{V}$)	Crude oil, x_1 (ton/day)	Expected total unmet demand/ production shortfall	Expected total excess production/ production surplus	Expected recourse penalty costs E_s	Expected profit $E[z_0]^a$	$\mu = \lvert E[z_0] - E_s \rvert$	$C_v = \frac{\sigma}{\mu}$ Stochastic	Deterministic
20 000	–5 478 239.9155	2 340.564 017	2 388.054	41 759.127	3 080.590	1.7064E+5	20 000.000	150 640	0.015 537	0.244790235
21 000	–6 020 579.7549	2 453.686 972	2 507.457	41 252.084	3 234.620	1.6895E+5	21 000.000	147 950	0.016 585	0.244790235
22 000	–6.589E+6	2 534.205 000	2 626.860	40 745.040	3 388.649	1.6726E+5	22 000.000	145 260	0.017 672	0.244790235
23 000	–7 184 873.4139	2 680.461 418	2 746.263	40 237.996	3 542.679	1.6557E+5	23 000.000	142 570	0.018 801	0.244790235
23 387.50	–7 422 731.2055	2 694.033 000	2 792.531	40 041.517	3 602.365	1.6492E+5	23 387.500	141 532.5	0.019 035	0.244790235
24 000	–7 806 827.2335	2 794.070 012	2 865.665	39 730.953	3 696.708	1.6389E+5	24 000.000	139 890	0.019 973	(infeasible)
25 000	–8 455 319.0466	2 907.803 131	2 985.068	39 223.909	3 850.738	1.6220E+5	25 000.000	137 200	0.021 194	(infeasible)
30 000	–12 095 848.0135	3 477.908 569	3 582.082	36 688.691	4 620.885	1.5375E+5	30 000.000	123 750	0.028 104	(infeasible)
40 000	–21 370 060.5296	4 622.776 279	4 776.109	32 421.525	6 964.451	1.3967E+5	40 000.000	99 670	0.046 381	(infeasible)
50 000	–33 300 678.6457	5 770.674 020	5 970.136	28 826.906	9 980.563	1.2819E+5	50 000.000	78 190	0.073 803	(infeasible)
75 000	–75 956 886.1800	8 715.324 789	9 072.786	19 481.863	17 833.008	99 527.049	75 000.000	24 527.05	0.355 335	(infeasible)
100 000	–151 547 526.7028	12 310.464 11	13 107.107	12 269.630	33 132.289	1.0328E+5	1.0000E+5	3 280	3.753 19	(infeasible)
103 000	–167 799 266.8795	12 953.735 63	13 916.386	10 228.811	35 607.246	1.0123E+5	1.0300E+5	1 770	7.318 495	(infeasible)
104 000	–175 333 797.2295	13 241.366 89	14 294.845	9 134.309	36 624.542	99 572.659	1.0400E+5	4 427.341	2.990 817	(infeasible)
105 000	–183 127 805.4033	13 532.472 26	14 673.303	8 039.808	37 641.838	97 913.829	1.0500E+5	7 086.171	1.909 702	(infeasible)
105 500	–187 122 113.6741	13 679.258 52	14 862.532	7 492.557	38 150.486	97 084.414	1.0550E+5	8 415.586	1.625 467	(infeasible)
105 800	–189 549 835.9754	13 767.709 90	14 976.070	7 164.207	38 455.675	96 586.765	1.0580E+5	9 213.235	1.494 34	(infeasible)
105 900	(infeasible solution)									

Table 9.18. Detailed computational results for Expectation Model II for target profit = deterministic profit = \$23 387.50

First-Stage Variable	Stochastic Solution	Product (i)	Production Shortfall z_{ij}^{+} or Surplus z_{ij}^{-} (t/d)					
			Scenario 1		Scenario 2		Scenario 3	
			z_{i1}^{+}	z_{i1}^{-}	z_{i2}^{+}	z_{i2}^{-}	z_{i3}^{+}	z_{i3}^{-}
x_1	2792.531	Demands RHS Coefficients Randomness						
x_2	647.867	Gasoline (2)	2187.133	0	2052.133	0	1917.133	0
x_3	300.197	Naphtha (3)	854.803	0	799.803	0	744.803	0
x_4	677.189	Jet Fuel (4)F	1737.811	0	1622.811	0	1507.811	0
x_5	1150.523	Heating Oil (5)	634.477	0	549.477	0	464.477	0
x_6	177.052	Fuel Oil (6)	8797.948	0	8322.948	0	7847.948	0
x_7	624.131							
x_8	862.892	Production Yields LHS Coefficients Randomness						
x_9	809.834	Naphtha (7)	0	242.950	0	261.102	0	279.253
x_{10}	1019.274	Jet Fuel (4)	0	237.365	0	258.309	0	279.253
x_{11}	323.934	Gas Oil (8)	0	217.817	0	248.535	0	279.253
x_{12}	862.892	Cracker Feed (9)	0	223.402	0	251.328	0	279.253
x_{13}	0	Residuum (10)	0	279.253	0	181.515	0	83.776
x_{14}	809.834							
x_{15}	0	E(Penalty Costs)		60 733.785		73 530.470		30 655.398
x_{16}	323.934	E_{total}	1.6492E+5					
x_{17}	445.409							
x_{18}	287.631							
x_{19}	157.778							
x_{20}	40.492							
Expected Profit Z (\$/day)	23 387.50							

Table 9.19. Computational statistics for Expectation Model II

Solver	Single continuous variables	Constraints	Resource usage/ CPU time (s)	Iterations
CONOPT 3	92	87	≈ (0.07–0.081)	10–13

9.2.6 Analysis of Results for Expectation Model II

Table 9.17 shows that the maximum expected profit, with a corresponding (very) high risk taken, is proposed to be approximately \$105 800/day. For a target profit equivalent to the deterministic profit of \$23 387.5, the detailed computational results displayed in Table 9.18 for Expectation Model II propose a marginally lower raw material flow rate of 2792.531 ton/day of crude oil compared to the flow rate of 12 500 ton/day proposed by the deterministic model. This augurs well for the stochastic model since for a lower purchasing cost of crude oil, the same amount of profit can be achieved by executing the production plan proposed by Expectation Model II.

The value of the computed coefficient of variation is lower for Expectation Model II compared to the value for the corresponding deterministic case for a same target profit, indicating lower degree of uncertainty in the stochastic model, which is exactly what Expectation Model II is intended to demonstrate.

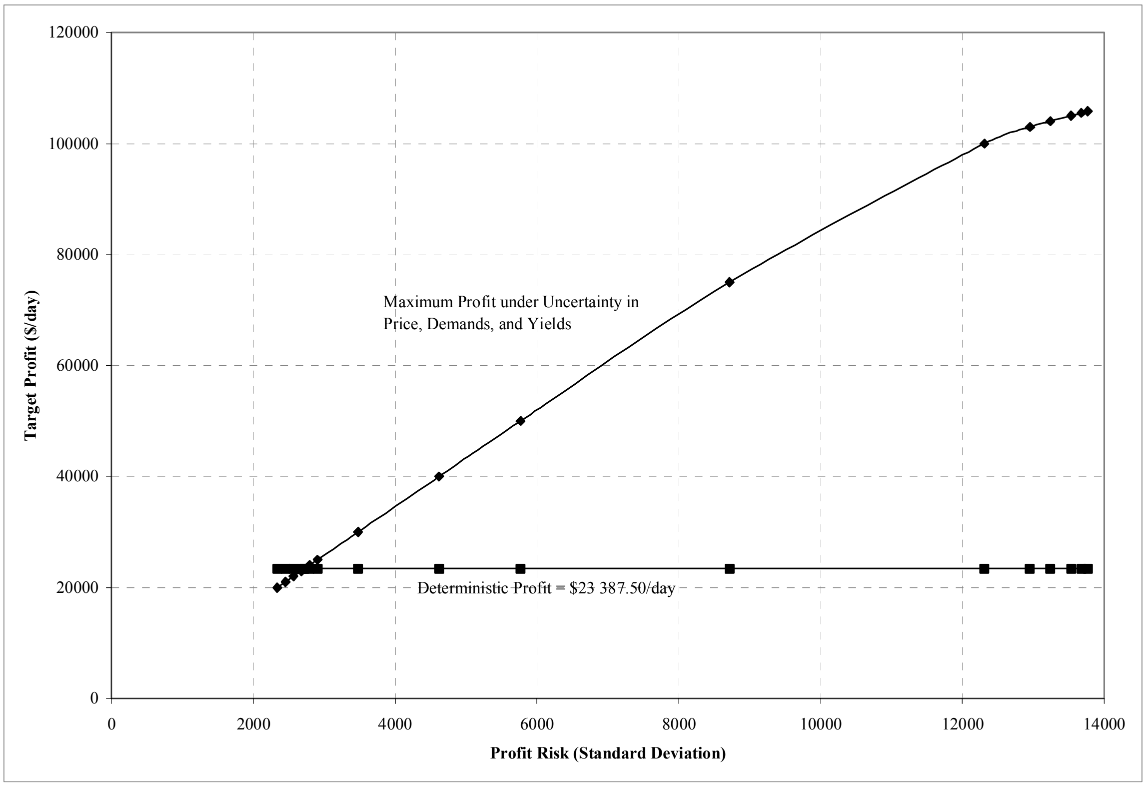


Figure 9.4. The efficient frontier plot of expected profit versus profit risk measured by standard deviation for Expectation Model II

(*Please turn the page over for the next section of* 9.3: Approach 3: Risk Model II.)

9.3 APPROACH 3: RISK MODEL II

9.3.1 Two-Stage Stochastic Programming with Fixed Recourse of Minimization of the Expected Value and the Variance of the Recourse Penalty Costs

Variance for the expected recourse penalty for the second-stage costs V_s is given by:

$$
\begin{aligned}
V_s &= p_1\left(\xi_1 - E_{s'}\right)^2 + p_2\left(\xi_2 - E_{s'}\right)^2 + p_3\left(\xi_3 - E_{s'}\right)^2 \\
V_s &= (0.35)\left(\xi_1 - E_{s'}\right)^2 + (0.45)\left(\xi_2 - E_{s'}\right)^2 + (0.2)\left(\xi_3 - E_{s'}\right)^2
\end{aligned}
\tag{9.57}
$$

or in general representation:

$$
V_s = \sum_{i\in I}\sum_{s\in S} p_s \left\{ \begin{bmatrix} \left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^-\right) \\ +\left(q_i^+ y_{i,s}^+ + q_i^- y_{i,s}^-\right) \end{bmatrix} - \sum_{i\in I}\sum_{s\in S} p_s \begin{bmatrix} \left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^-\right) \\ +\left(q_i^+ y_{i,s}^+ + q_i^- y_{i,s}^-\right) \end{bmatrix} \right\}^2 \tag{9.58}
$$

where ξ_1, ξ_2, ξ_3, and $E_{s'}$ is given by equation **(9.53)**.

Therefore, Risk Model II is as follows:

$$
\begin{aligned}
\text{maximize } z_3 &= z_2 - V_s \\
&= E\left[z_0\right] - \theta_1 \mathrm{Var}(z_0) - E_s - V_s \\
&= \sum_{i\in I}\sum_{s\in S} p_s C_i x_i - \theta_1 \sum_{i\in I} x_i^2 V(C_i) - \sum_{i\in I}\sum_{s\in S} p_s\left[\left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^-\right) + \left(q_i^+ y_{i,s}^+ + q_i^- y_{i,s}^-\right)\right] - \theta_2 V_s \\
\text{maximize } z_3 &= \begin{bmatrix} (0.35)\left(-8.8x_1 + 20.35x_2 + 8.8x_3 + 13.75x_4 + 15.95x_5 + 6.6x_6 - 1.65x_{14}\right) \\ +(0.45)\left(-8.0x_1 + 18.5x_2 + 8.0x_3 + 12.5x_4 + 14.5x_5 + 6.0x_6 - 1.5x_{14}\right) \\ +(0.2)\left(-7.2x_1 + 16.65x_2 + 7.2x_3 + 11.25x_4 + 13.05x_5 + 5.4x_6 - 1.35x_{14}\right) \end{bmatrix} \\
&\quad - \theta_1 \begin{bmatrix} (0.352)x_1^2 + (1.882\ 375)x_2^2 + (0.352)x_3^2 + (0.859\ 375)x_4^2 + (1.156\ 375)x_5^2 \\ +(0.198)x_6^2 + (0.012\ 375)x_{14}^2 \end{bmatrix} \\
&\quad - E_s - V_s
\end{aligned}
\tag{9.59}
$$

s.t. deterministic constraints (first stage) **(8.1)**, **(8.2)**, **(8.8)–(8.19)**, and **(8.26)**, stochastic constraints (second stage): **(9.17)–(9.32)** and **(9.37)–(9.52)**.

9.3.2 Computational Results for Risk Model II

Tables 9.20, 9.22, and 9.24 tabulate the computational results for the implementation of Risk Model II on GAMS for a range of values of the recourse penalty costs risk parameter θ_2 for three distinct cases of the value of the profit risk parameter fixed at θ_1 = 0.000 000 000 1, 0.000 000 1, and 0.000 015 5, respectively. Starting values of the first-stage deterministic decision variables have been initialized to the optimal solutions of the deterministic model. Representative detailed results are presented in Tables 9.21, 9.23, and 9.25 that immediately follow each of the three cases for suitable (or particular meaningful) values of θ_2.A number of different parameters are of interest in observing the trends and patterns that contribute to robustness in the model and robustness in the computed solution, as extensively analyzed in the ensuing discussion.

(*Please turn the page over for* Table 9.20.)

Table 9.20. Computational results for Risk Model II for $\theta_1 = 0.000\ 000\ 000\ 1$

Recourse penalty costs risk factor θ_2	Optimal objective value	Expected variation in profit $V(z_0)$	Expected total unmet demand/ production shortfall	Expected total excess production/ production surplus	Expected recourse penalty costs E_s	Expected variation in recourse penalty costs V_s	$\sigma = \sqrt{V(z_0)+V_s}$	Expected profit $E[z_0]$	$\mu = E[z_0]-E_s$	$C_v = \frac{\sigma}{\mu}$ Stochastic	Deterministic
0.000 001	27 709.6840	1.7871E+8	2 575.000	26 815.000	54 056.875	8.1669E+6	13 670.119	81 774.744	27 717.869	0.493 187 95	(infeasible)
0.000 01	27 636.1817	1.7871E+8	2 575.000	26 815.000	54 056.875	8.1669E+6	13 670.119	81 774.744	27 717.869	0.493 187 95	(infeasible)
0.000 1	26 901.1589	1.7871E+8	2 575.000	26 815.000	54 056.875	8.1669E+6	13 670.119	81 774.744	27 717.869	0.493 187 95	(infeasible)
0.001	25 695.4215	1.7818E+8	2 466.605	26 706.605	53 843.639	2.0473E+5	13 356.232	79 743.805	25 900.166	0.515 681 328	(infeasible)
0.01	25 511.1674	1.7818E+8	2 539.118	26 779.118	54 195.864	2047.268	13 348.516	79 727.522	25 531.658	0.522 822 137	(infeasible)
0.1	25 492.7420	1.7818E+8	2 518.476	26 758.476	54 231.086	20.473	13 348.428	79 725.893	25 494.807	0.523 574 389	(infeasible)
1	25 490.8994	1.7818E+8	2 385.328	26 625.328	54 234.609	0.205	13 348.426	79 725.731	25 491.122	0.523 649 999	(infeasible)
2	25 490.7970	1.7818E+8	2 508.325	26 748.325	54 234.804	0.051	13 348.426	79 725.722	25 490.917	0.523 654 21	(infeasible)
5	25 490.7356	1.7818E+8	2 508.330	26 748.330	54 234.922	0.008	13 348.426	79 725.716	25 490.794	0.523 656 737	(infeasible)
10	25 490.7152	1.7818E+8	2 565.618	26 805.618	54 234.961	0.002	13 348.426	79 725.714	25 490.753	0.523 657 579	(infeasible)
50	25 490.6988	1.7818E+8	2 556.034	26 796.034	54 234.992	8.1891E−5	13 348.426	79 725.713	25 490.721	0.523 658 236	(infeasible)
100	25 490.6967	1.7818E+8	2 518.750	26 758.750	54 234.996	2.0473E−5	13 348.426	79 725.713	25 490.717	0.523 658 318	(infeasible)
1 000	25 490.6949	1.7818E+8	2 518.750	26 758.750	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
2 000	25 490.6948	1.7818E+8	2 543.023	26 783.023	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
3 000	25 490.6948	1.7818E+8	2 495.782	26 735.782	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
4 000	25 490.6947	1.7818E+8	2 518.750	26 758.750	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
5 000	25 490.6947	1.7818E+8	2 385.000	26 625.000	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
6 000	25 490.6947	1.7818E+8	2 456.349	26 696.349	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
7 000	25 490.6947	1.7818E+8	2 555.208	26 795.208	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
8 000	25 490.6947	1.7818E+8	2 518.750	26 758.750	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
9 000	25 490.6947	1.7818E+8	2 508.333	26 748.333	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 000	25 490.6947	1.7818E+8	2 542.312	26 782.312	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 500	25 490.6947	1.7818E+8	2 408.750	26 648.750	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 600	25 490.6947	1.7818E+8	2 591.667	26 831.667	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 700	25 490.6947	1.7818E+8	2 589.658	26 829.658	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 710	25 490.6947	1.7818E+8	2 591.667	26 831.667	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 715	25 490.6947	1.7818E+8	2 543.944	26 783.944	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 716	25 490.6947	1.7818E+8	2 508.333	26 748.333	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 717	(GAMS output: `Infeasible solution. There are no superbasic variables.`)										

Note: Trial solutions for $\theta_1 < 0.000\ 001$ are not considered since improvement in expected profit is not anticipated based on trends in computed values.

Table 9.21. Detailed computational results for Risk Model II for $\theta_1 = 0.000\,000\,000\,1$, $\theta_2 = 50$

First-Stage Variable	Stochastic Solution	Product (i)	Production Shortfall z_{ij}^{+} or Surplus z_{ij}^{-} (t/d)					
			Scenario 1		Scenario 2		Scenario 3	
			z_{i1}^{+}	z_{i1}^{-}	z_{i2}^{+}	z_{i2}^{-}	z_{i3}^{+}	z_{i3}^{-}
x_1	15 000.000	Demands RHS Coefficients Randomness						
x_2	2 000.000	Gasoline (2)	835.000	0	700.000	0	565.000	0
x_3	1 155.000	Naphtha (3)	0	0	0	55.000	0	110.000
x_4	3 637.500	Jet Fuel (4)	0	1222.500	0	1337.500	0	1452.500
x_5	3 597.500	Heating Oil (5)	0	1812.500	0.00	1897.50	0	1982.500
x_6	9 737.500	Fuel Oil (6)	237.500	0	0	237.500	0	712.500
x_7	2 155.000							
x_8	4 635.000	Production Yields LHS Coefficients Randomness						
x_9	4 350.000	Naphtha (7)	0	107.500	0	205.000	0	302.500
x_{10}	5 475.000	Jet Fuel (4)	0	1275.000	0	1387.500	0	1500.000
x_{11}	1 000.000	Gas Oil (8)	0	1170.000	0	1335.000	0	1500.000
x_{12}	2 698.125	Cracker Feed (9)	0	1200.000	0	1350.000	0	1500.000
x_{13}	1 936.875	Residuum (10)	0	1500.000	0	975.000	0	450.000
x_{14}	2 500.000							
x_{15}	1 850.000	E(Penalty Costs)		18 982.250		24 405.742		10 847.000
x_{16}	1 000.000	E_{total}	54 234.992					
x_{17}	1 375.000							
x_{18}	899.375							
x_{19}	475.625							
x_{20}	125.000							
Expected Profit z ($/day)	79 725.713							

Table 9.22. Computational results for Risk Model II for $\theta_1 = 0.000\,000\,1$

Recourse penalty cost risk factor θ_2	Optimal objective value	Expected variation in profit $V(z_0)$	Expected total unmet demand/ production shortfall	Expected total excess production/ production surplus	Expected recourse penalty costs E_s	Expected variation in recourse penalty costs V_s	$\sigma = \sqrt{V(z_0)+V_s}$	Expected profit $E[z_0]$	$\mu = E[z_0] - E_s$	$C_v = \frac{\sigma}{\mu}$ Stochastic	Deterministic
0.000 001	27 691.8313	1.7871E+8	2 575.000	26 815.000	54 056.875	8.1669E+6	13 670.119	81 774.744	27 717.869	0.493 187 95	(infeasible)
0.000 01	27 618.3290	1.7871E+8	2 575.000	26 815.000	54 056.875	8.1669E+6	13 670.119	81 774.744	27 717.869	0.493 187 95	(infeasible)
0.000 1	26 883.3062	1.7871E+8	2 575.000	26 815.000	54 056.875	8.1669E+6	13 670.119	81 774.744	27 717.869	0.493 187 95	(infeasible)
0.001	25 677.6209	1.7818E+8	2 467.445	26 707.445	53 843.621	2.0473E+5	13 356.232	79 743.787	25 900.166	0.515 681 328	(infeasible)
0.01	25 493.3671	1.7818E+8	2 564.890	26 804.890	54 195.862	2047.264	13 348.516	79 727.520	25 531.658	0.522 822 137	(infeasible)
0.1	25 474.9417	1.7818E+8	2 384.824	26 624.824	54 231.086	20.473	13 348.428	79 725.893	25 494.807	0.523 574 389	(infeasible)
1	25 473.0992	1.7818E+8	2 456.203	26 696.203	54 234.609	0.205	13 348.426	79 725.731	25 491.122	0.523 649 999	(infeasible)
2	25 472.9968	17.818E+8	2 572.688	26 812.688	54 234.804	0.051	13 348.426	79 725.722	25 490.917	0.523 654 21	(infeasible)
3	25 472.9627	1.7818E+8	2 572.701	26 812.701	54 234.870	0.023	13 348.426	79 725.719	25 490.849	0.523 655 607	(infeasible)
4	25 472.9456	1.7818E+8	2 384.996	26 624.996	54 234.902	0.013	13 348.426	79 725.717	25 490.815	0.523 656 305	(infeasible)
5	25 472.9354	1.7818E+8	2 518.745	26 758.745	54 234.922	0.008	13 348.426	79 725.716	25 490.794	0.523 656 737	(infeasible)
6	25 472.9286	1.7818E+8	2 508.330	26 748.330	54 234.935	0.006	13 348.426	79 725.716	25 490.781	0.523 657 004	(infeasible)
7	25 472.9237	1.7818E+8	2 508.331	26 748.331	54 234.944	0.004	13 348.426	79 725.715	25 490.771	0.523 657 209	(infeasible)
8	25 472.9200	1.7818E+8	2 508.331	26 748.331	54 234.951	0.003	13 348.426	79 725.715	25 490.764	0.523 657 353	(infeasible)
9	25 472.9172	1.7818E+8	2 508.331	26 748.331	54 234.957	0.003	13 348.426	79 725.715	25 490.758	0.523 657 476	(infeasible)
10	25 472.9149	17.818E+8	2 508.332	26 748.332	54 234.961	0.002	13 348.426	79 725.714	25 490.753	0.523 657 579	(infeasible)
50	25 472.8985	17.818E+8	2 385.000	26 625.000	54 234.992	8.1891E−5	13 348.426	79 725.713	25 490.721	0.523 658 236	(infeasible)
100	25 472.8965	17.818E+8	2 385.000	26 625.000	54 234.996	2.0473E−5	13 348.426	79 725.713	25 490.717	0.523 658 318	(infeasible)
1 000	25 472.8947	17.818E+8	2 508.333	26 748.333	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
2 000	25 472.8946	17.818E+8	2 556.003	26 796.003	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
3 000	25 472.8945	17.818E+8	2 456.250	26 696.250	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
4 000	25 472.8945	17.818E+8	2 556.960	26 796.960	54 234.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
5 000	25 472.8945	17.818E+8	2 456.250	26 696.250	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
6 000	25 472.8945	17.818E+8	2 456.250	26 696.250	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
7 000	25 472.8945	17.818E+8	2 508.333	26 748.333	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
8 000	25 472.8945	17.818E+8	2 518.750	26 758.750	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
9 000	25 472.8945	17.818E+8	2 556.150	26 796.150	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 000	25 472.8945	17.818E+8	2 475.000	26 715.000	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 500	25 472.8945	17.818E+8	2 408.750	26 648.750	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 600	25 472.8945	17.818E+8	2 589.655	26 829.655	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 700	25 472.8945	17.818E+8	2 591.146	26 831.750	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 710	25 472.8945	17.818E+8	2 518.750	26 758.750	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 715	25 472.8945	17.818E+8	2 591.667	26 831.667	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 716	25 472.8945	17.818E+8	2 584.141	26 824.141	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 717	(GAMS output: `Infeasible solution. There are no superbasic variables.`)										

Note: Trial solutions for $\theta_1 < 0.000\,001$ are not considered since improvement in expected profit is not anticipated based on trends in computed values.

Table 9.23. Detailed computational results for Risk Model II for $\theta_1 = 0.000\,000\,1$, $\theta_2 = 50$

First-Stage Variable	Stochastic Solution	Product (i)	Production Shortfall z_{ij}^{+} or Surplus z_{ij}^{-} (t/d)					
			Scenario 1		Scenario 2		Scenario 3	
			z_{i1}^{+}	z_{i1}^{-}	z_{i2}^{+}	z_{i2}^{-}	z_{i3}^{+}	z_{i3}^{-}
x_1	15 000.000		Demands RHS Coefficients Randomness					
x_2	2 000.000	Gasoline (2)	835.000	0	700.000	0	565.000	0
x_3	1 155.000	Naphtha (3)	0	0	0	55.000	0	110.000
x_4	3 637.500	Jet Fuel (4)	0	1222.500	237.494	1574.994	0	1452.500
x_5	3 597.500	Heating Oil (5)	0	1812.500	0	1897.500	0	1982.500
x_6	9 737.500	Fuel Oil (6)	237.500	0	0	237.500	0	712.500
x_7	2 155.000							
x_8	4 635.000		Production Yields LHS Coefficients Randomness					
x_9	4 350.000	Naphtha (7)	0	107.500	0	205.000	0	302.500
x_{10}	5 475.000	Jet Fuel (4)	0	1275.000	0	1387.500	0	1500.000
x_{11}	1 000.000	Gas Oil (8)	0	1170.000	0	1335.000	0	1500.000
x_{12}	2 698.125	Cracker Feed (9)	0	1200.000	0	1350.000	0	1500.000
x_{13}	1 936.875	Residuum (10)	0	1500.000	0	975.000	0	450.000
x_{14}	2 500.000							
x_{15}	1 850.000	E(Penalty Costs)		18 982.250		24 405.742		10 847.000
x_{16}	1 000.000	E_{total}	54 234.992					
x_{17}	1 375.000							
x_{18}	899.375							
x_{19}	475.625							
x_{20}	125.000							
Expected Profit $E[z_0]$ (\$/day)	79 725.713							

Table 9.24. Computational results for Risk Model II for $\theta_1 = 0.000\,015\,5$

Recourse penalty cost risk factor θ_2	Optimal objective value	Expected variation in profit $V(z_0)$	Expected total unmet demand/production shortfall	Expected total excess production/production surplus	Expected recourse penalty costs E_s	Expected variation in recourse penalty costs V_s	$\sigma = \sqrt{V(z_0)+V_s}$	Expected profit $E[z_0]$	$\mu = E[z_0] - E_s$	$C_v = \frac{\sigma}{\mu}$ Stochastic	Deterministic
0.000 001	24 939.7708	1.7871E+8	2 575.000	26 815.000	54 056.875	8.1669E+6	13 670.119	81 774.744	27 717.869	0.493 187 95	(infeasible)
0.000 01	24 866.2685	1.7871E+8	2 575.000	26 815.000	54 056.875	8.1669E+6	13 670.119	81 774.744	27 717.869	0.493 187 95	(infeasible)
0.000 1	24 131.2457	1.7871E+8	2 575.000	26 815.000	54 056.875	8.1669E+6	13 670.119	81 774.744	27 717.869	0.493 187 95	(infeasible)
0.000 2	23 604.6348	1.7837E+8	2 434.859	26 674.859	53 594.411	3.0094E+6	13 467.742	80 565.679	26 971.269	0.499 336 609	(infeasible)
0.000 25	23 477.8943	1.7830E+8	2 399.000	26 639.000	53 476.076	2.1550E+6	13 433.225	80 256.306	26 780.230	0.501 609 77	(infeasible)
0.000 26	23 456.9294	1.7829E+8	2 393.481	26 633.481	53 457.862	2.0403E+6	13 428.546	80 208.689	26 750.826	0.501 986 219	(infeasible)
0.000 27	23 437.0473	1.7828E+8	2 388.370	26 628.370	53 440.996	1.9381E+6	13 424.365	80 164.594	26 723.598	0.502 341 227	(infeasible)
0.000 28	23 418.1321	1.7827E+8	2 383.624	26 623.624	53 425.333	1.8466E+6	13 420.613	80 123.644	26 698.311	0.502 676 48	(infeasible)
0.000 29	23 400.0840	1.7826E+8	2 379.204	26 619.204	53 410.749	1.7644E+6	13 417.233	80 085.515	26 674.766	0.502 993 466	(infeasible)
0.000 3	23 382.8164	1.7825E+8	2 375.079	26 615.079	53 397.135	1.6903E+6	13 414.175	80 049.925	26 652.789	0.503 293 483	(infeasible)
0.000 5	23 138.2622	1.7819E+8	2 362.668	26 602.668	53 446.596	8.1869E+5	13 379.291	79 756.099	26 309.503	0.508 534 54	(infeasible)
0.001	22 933.5888	1.7818E+8	2 492.502	26 732.502	53 840.808	2.0467E+5	13 356.208	79 740.916	25 900.108	0.515 681 556	(infeasible)
0.01	22 749.3826	1.7818E+8	2 523.914	26 763.914	54 195.582	2046.737	13 348.514	79 727.234	25 531.652	0.522 822 182	(infeasible)
0.1	22 730.9619	1.7818E+8	2 531.028	26 771.028	54 231.058	20.467	13 348.428	79 725.865	25 494.806	0.523 574 41	(infeasible)
1	22 729.1199	1.7818E+8	2 578.902	26 818.902	54 234.606	0.205	13 348.426	79 725.728	25 491.122	0.523 649 999	(infeasible)
10	22 728.9357	1.7818E+8	2 580.388	26 820.388	54 234.961	0.002	13 348.426	79 725.714	25 490.753	0.523 657 579	(infeasible)
50	22 728.9193	1.7818E+8	2 408.755	26 648.755	54 234.992	8.1869E−5	13 348.426	79 725.713	25 490.721	0.523 658 236	(infeasible)
100	22 728.9172	1.7818E+8	2 590.624	26 830.624	54 234.996	2.0467E−5	13 348.426	79 725.713	25 490.717	0.523 658 318	(infeasible)
500	22 728.9156	1.7818E+8	2 550.749	26 790.749	54 234.999	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
1 000	22 728.9154	1.7818E+8	2 531.818	26 771.818	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
5 000	22 728.9152	1.7818E+8	2 408.750	26 648.750	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
10 000	22 728.9152	1.7818E+8	2 560.606	26 800.606	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
12 000	22 728.9152	1.7818E+8	2 597.388	26 837.388	54 235.000	0	13 348.426	79 725.713	25 490.713	0.523 658 401	(infeasible)
12 500	(GAMS output: `Infeasible solution. A free variable exceeds the allowable range.`)										

Note: Trial solutions for $\theta_1 < 0.000\,001$ are not considered since improvement in expected profit is not anticipated based on trends in computed values.

Table 9.25. Detailed computational results for Risk Model II for θ_1 = 0.000 015 5, θ_2 = 0.001

First-Stage Variable	Stochastic Solution	Product (i)	Production Shortfall z_{ij}^+ or Surplus z_{ij}^- (t/d)					
			Scenario 1		Scenario 2		Scenario 3	
			z_{i1}^+	z_{i1}^-	z_{i2}^+	z_{i2}^-	z_{i3}^+	z_{i3}^-
x_1	15 000.000		Demands RHS Coefficients Randomness					
x_2	2 000.000	Gasoline (2)	835.000	0	700.000	0	565.000	0
x_3	1 155.000	Naphtha (3)	0	0	0	55.000	0	110.000
x_4	3 637.500	Jet Fuel (4)	0	1222.500	0	1337.500	0	1452.500
x_5	3 599.262	Heating Oil (5)	0	1814.262	0	1899.263	0	1984.262
x_6	9735.738	Fuel Oil (6)	239.262	0	0	235.738	0	710.738
x_7	2 155.000							
x_8	4 635.000		Production Yields LHS Coefficients Randomness					
x_9	4 350.000	Naphtha (7)	0	107.500	0	205.000	0	302.500
x_{10}	5 475.000	Jet Fuel (4)	0	1275.000	0	1387.500	0	1500.000
x_{11}	1 000.000	Gas Oil (8)	0	1170.000	0	1335.000	0	1500.000
x_{12}	2 699.447	Cracker Feed (9)	0	1200.000	0	1350.000	0	1500.000
x_{13}	1 935.553	Residuum (10)	0	1500.000	0	975.000	0	450.000
x_{14}	2 500.000							
x_{15}	1 850.000	E(Penalty Costs)		18 991.502		24 003.364		10 845.943
x_{16}	1 000.000	E_{total}	53 840.808					
x_{17}	1 375.000							
x_{18}	899.816							
x_{19}	475.184							
x_{20}	125.000							
Expected Profit Z (\$/day)	79 740.916							

The problem size and the efficient distribution of computational expense are depicted through the computational statistics shown in Table 8.30. Figure 8.6 then illustrates the relationship between expected profit and the various levels of risk as dependent on the tradeoff dictated by the profit risk factor θ_1 and the recourse penalty costs risk factor θ_2 with variance as the risk measure.

Table 9.26. Computational statistics for Risk Model II

Solver	Single continuous variables	Constraints	Resource usage/ CPU time (s)	Iterations
CONOPT 3	96	90	≈ (0.07–0.081)	10–13

9.3.3 Analysis of Results for Risk Model II

The values of θ_1 and θ_2 denotes the importance of risk in the model as contributed by variation in profit and variation in recourse penalty costs, respectively, in comparison with the corresponding expected values of the model's objective.

Similar to the expected value models, smaller values of θ_1 correspond to higher expected profit, as shown in Figure 9.5. With increasingly larger θ_1, the reduction in expected profit becomes almost constant as demonstrated for the cases of $\theta_1 = 0.000\ 000\ 1$ and $\theta_1 = 0.000\ 000\ 000\ 1$; the converse is true as well, that is, with increasingly smaller θ_1, the increment in expected profit becomes roughly constant.

Although the pair of increasing θ_2 with fixed value of θ_1 corresponds to reduction expected profit, it generally leads to a reduction of expected production shortfalls and surpluses as well. Based on the conceptual definition of model robustness presented earlier for Expectation Model I, this reflects high model feasibility. Therefore, a suitable operating range of θ_2 values ought to be selected in order to achieve optimality between expected profit and expected production feasibility. Increasing θ_2 also reduces the expected variation or deviation in the recourse penalty costs under different realized scenarios. This in turn translates to increased solution robustness. It thus depends on the policy adopted by the decision maker, as characterized by the values of the factors θ_1 and θ_2 chosen, in reflecting whether these tradeoffs are acceptable based on the desired degree of model robustness and solution robustness, as reported by Bok et al. (1998).

In general, the coefficients of variation decrease with larger values of θ_2. This is definitely a desirable behaviour since for higher expected profits, there is diminising uncertainty in the model, thus signifying model and solution robustness. It is also observed that for larger values of θ_2 until approximately greater than 100, the coefficient of variation remain at a static value of 0.5237 (correct to four significant figures), therefore indicating stability and minimal degree of uncertainty in the model.

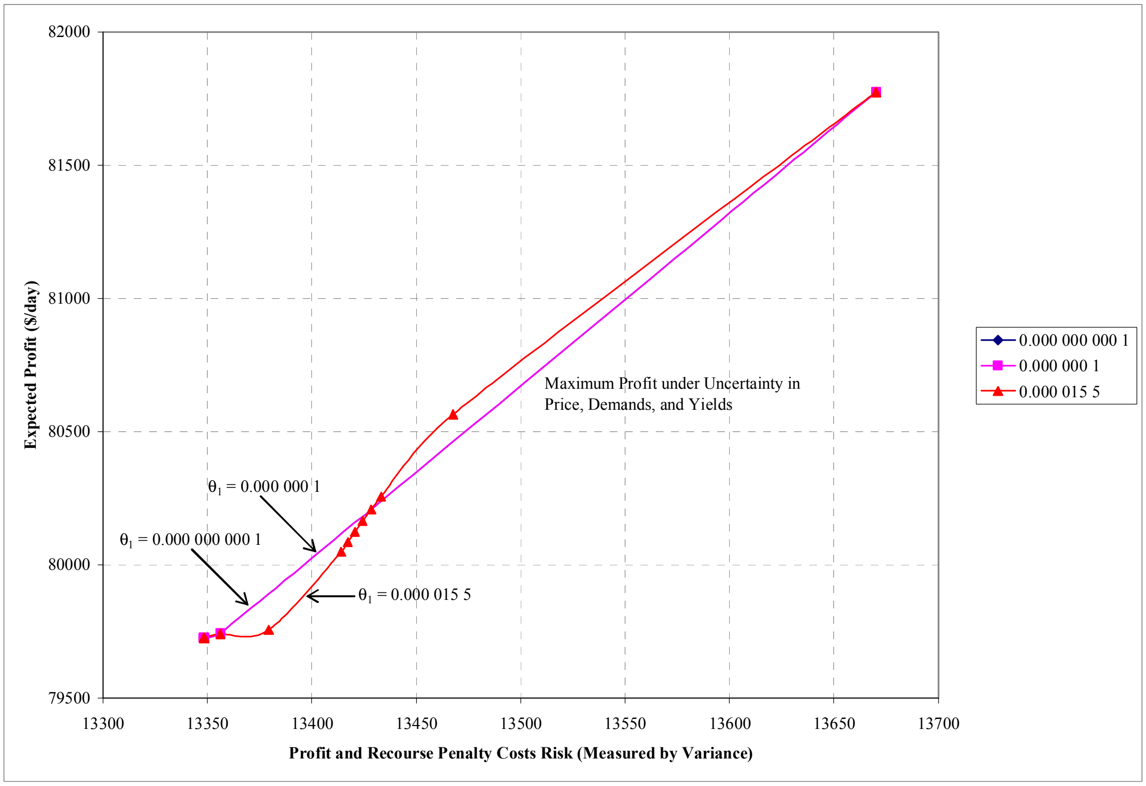


Figure 9.5. The efficient frontier plot of expected profit versus risk imposed by variations in both profit and the recourse penalty costs as measured by variance for Risk Model II. Note that the plot for θ_1 = 0.000 000 000 1 overlaps with the plot for θ_1 = 0.000 000 1.

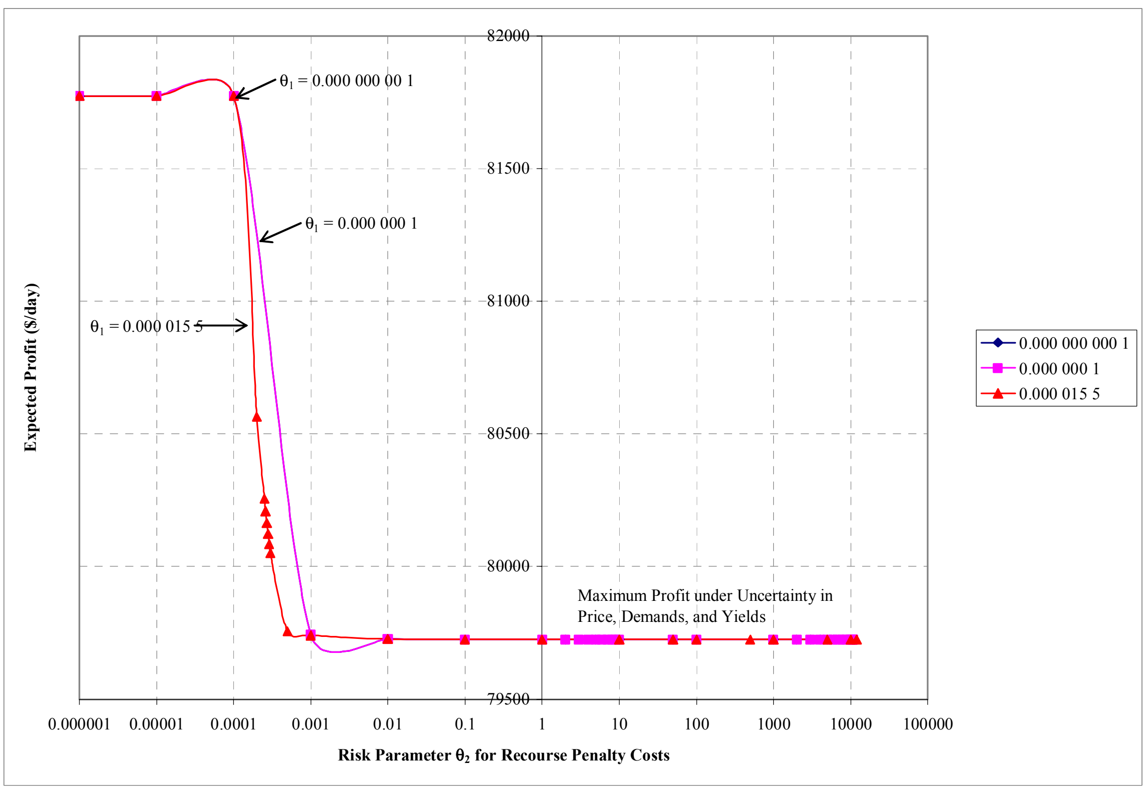


Figure 9.6. Plot of expected profit for different levels of risk as represented by the profit risk factor θ_1 and the recourse penalty costs risk factor θ_2 (with θ_1 and θ_2 in logarithmic scales due to wide range of values) with variance as the risk measure for Risk Model II. Note that the plot for θ_1 = 0.000 000 000 1 overlaps with the plot for θ_1 = 0.000 000 1.

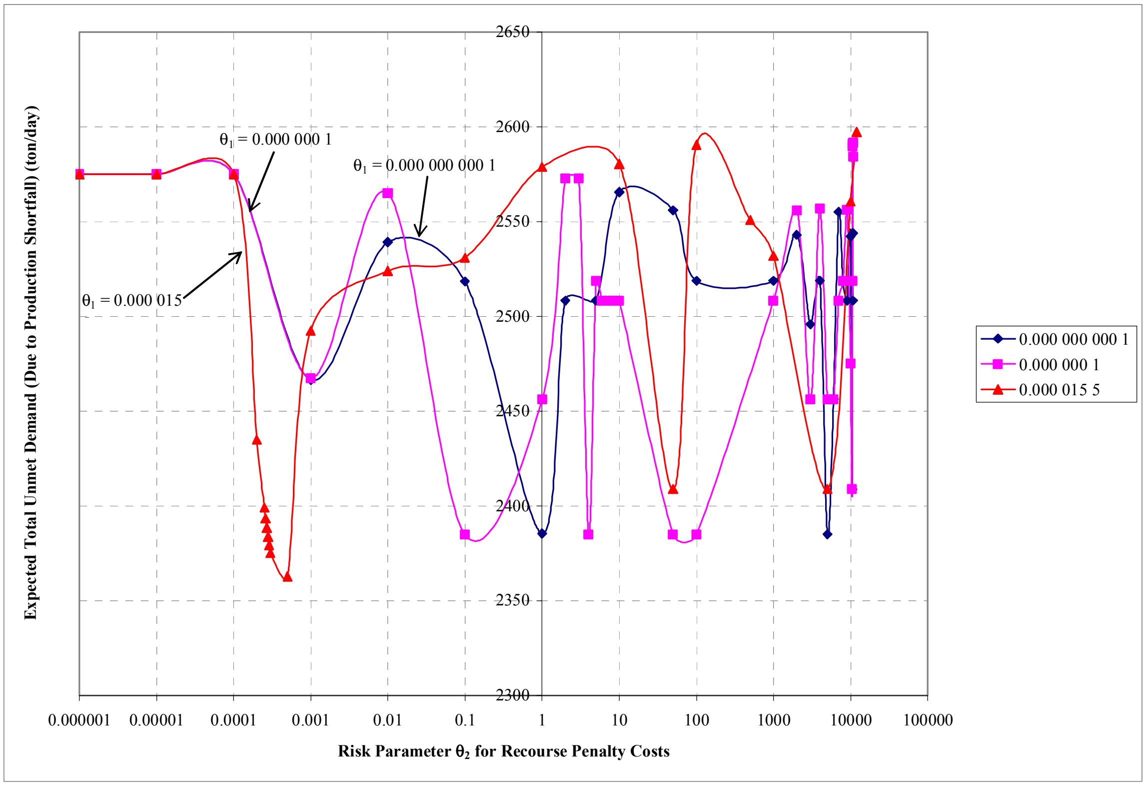


Figure 9.7. Investigating model robustness via the plot of expected total unmet demand (due to production shortfall) versus the recourse penalty costs risk factor θ_2

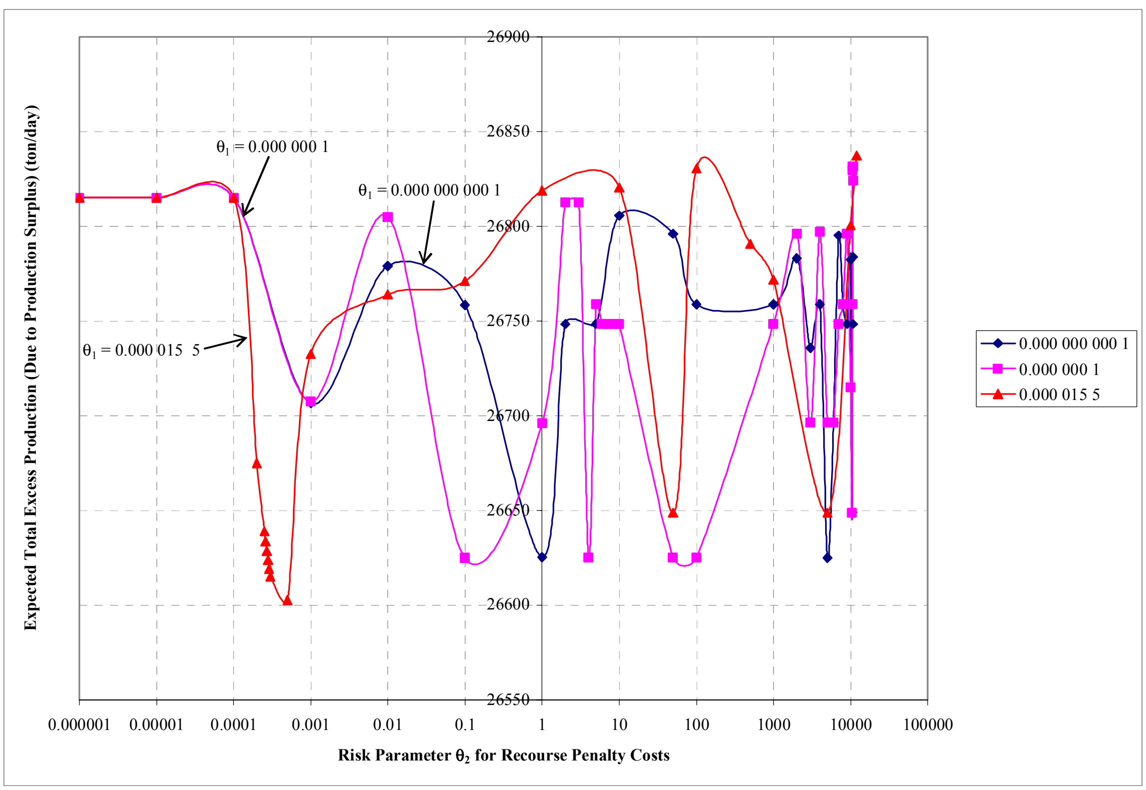


Figure 9.8. Investigating model robustness via the plot of expected total excess production (due to production surplus) versus the recourse penalty costs risk factor θ_2

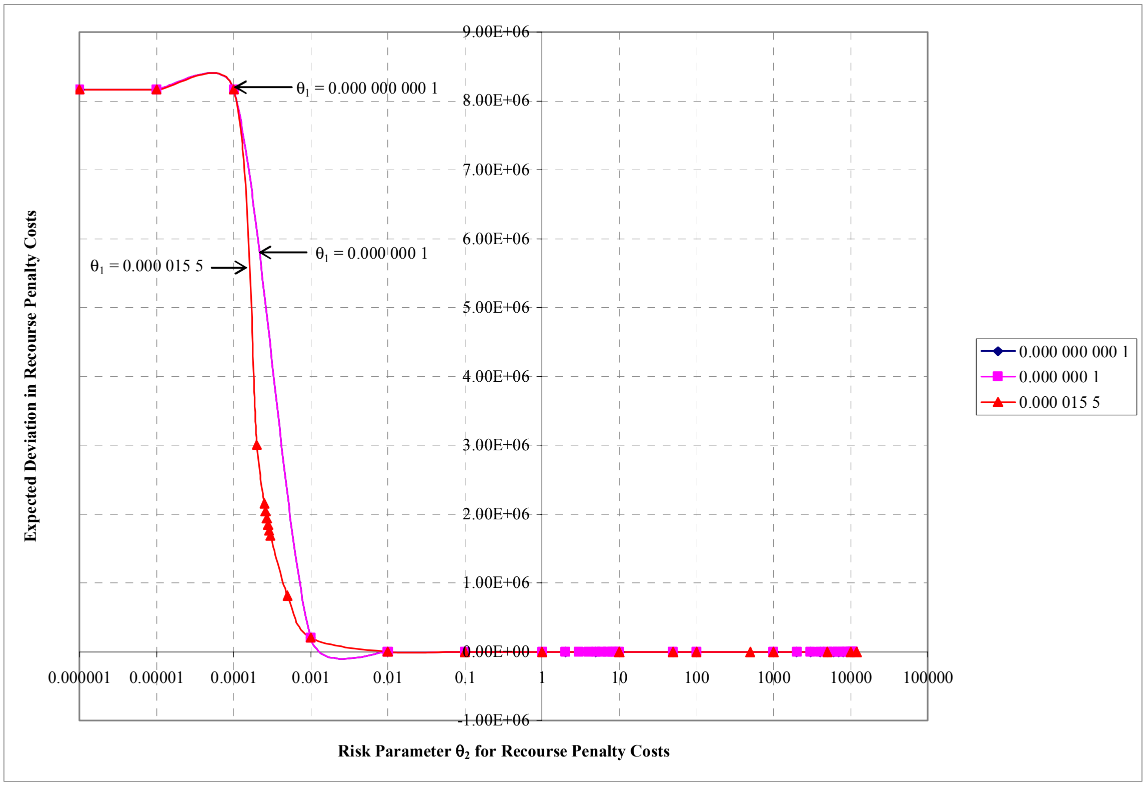


Figure 9.9. Investigating solution robustness via the plot of expected variation in the recourse penalty costs versus the recourse penalty costs risk factor θ_2. Note that the plot for θ_1 = 0.000 000 000 1 overlaps with the plot for θ_1 = 0.000 000 1.

9.3.4 Comparison of Performance between Expectation Model I and Risk Model II

As emphasized, the motivation for employing Risk Model II is to account for the presence of risk in decision-making that is not considered by the risk-neutral Expectation Model I. Although a robust mathematical (or statistical) approach for direct comparison between the expected profit obtained by the proposed models of the two approach may not be conventionally available, from the general trend computed, it is apparent that Risk Model II consistently registers a higher expected profit, thus testifying to its superior robustness in the face of multitude uncertainties. In fact, the average expected profit registered by Risk Model II is a commendable \$80 000/day for feasible pair of values of (θ_1, θ_2) whereas for Expectation Model I, the expected profit even dipped below the deterministic profit (of \$23 387.50/day, for profit risk factor θ_1 that are approximately larger than 0.0045), as evidenced from Table 9.14.

9.4 APPROACH 4: RISK MODEL III

9.4.1 Two-Stage Stochastic Programming with Fixed Recourse for Minimization of the Expected Value and the Mean-Absolute Deviation (MAD) of the Variation in Recourse Penalty Costs

The mean-absolute deviation (MAD) for the expected variation in the recourse penalty for the second-stage costs $W(p_s)$ is given by:

$$\begin{aligned} W\left(p_s\right) &= p_1\left|\xi_1 - E_{s'}\right| + p_2\left|\xi_2 - E_{s'}\right| + p_3\left|\xi_3 - E_{s'}\right| \\ W\left(p_s\right) &= \left(0.35\right)\left|\xi_1 - E_{s'}\right| + \left(0.45\right)\left|\xi_2 - E_{s'}\right| + \left(0.2\right)\left|\xi_3 - E_{s'}\right| \end{aligned} \quad \textbf{(9.60)}$$

where ξ_1, ξ_2, ξ_3, and $E_{s'}$ is given by equation **(9.53)**.

Therefore, Risk Model III is presented as follows:

$$\begin{aligned} \text{maximize } z_4 &= z_2 - \theta_3 W(p_s) \\ &= E(z_0) - \theta_1 V(z_0) - E_s - \theta_3 W(p_s) \\ &= E(z_0) - \theta_1 V(z_0) - \sum_{i\in I}\sum_{s\in S} p_s\left[\left(c_i^+ z_{is}^+ + c_i^- z_{is}^-\right) + \left(q_i^+ y_{is}^+ + q_i^- y_{is}^-\right)\right] \\ &\quad - \theta_3 \sum_{i\in I}\sum_{s\in S} p_s \left| \left(c_i^+ z_{is}^+ + c_i^- z_{is}^-\right) + \left(q_i^+ y_{is}^+ + q_i^- y_{is}^-\right) - \sum_{i\in I}\sum_{s\in S} p_s \begin{bmatrix} \left(c_i^+ z_{is}^+ + c_i^- z_{is}^-\right) \\ + \left(q_i^+ y_{is}^+ + q_i^- y_{is}^-\right) \end{bmatrix} \right| \end{aligned}$$

$$\begin{aligned} \text{maximize } z_4 &= \begin{bmatrix} \left(0.35\right)\left(-8.8x_1 + 20.35x_2 + 8.8x_3 + 13.75x_4 + 15.95x_5 + 6.6x_6 - 1.65x_{14}\right) \\ + \left(0.45\right)\left(-8.0x_1 + 18.5x_2 + 8.0x_3 + 12.5x_4 + 14.5x_5 + 6.0x_6 - 1.5x_{14}\right) \\ + \left(0.2\right)\left(-7.2x_1 + 16.65x_2 + 7.2x_3 + 11.25x_4 + 13.05x_5 + 5.4x_6 - 1.35x_{14}\right) \end{bmatrix} \\ &\quad - \theta_1 \begin{bmatrix} \left(0.352\right)x_1^2 + \left(1.882\ 375\right)x_2^2 + \left(0.352\right)x_3^2 + \left(0.859\ 375\right)x_4^2 \\ + \left(1.156\ 375\right)x_5^2 + \left(0.198\right)x_6^2 + \left(0.012\ 375\right)x_{14}^2 \end{bmatrix} \\ &\quad - \sum_{i\in I}\sum_{s\in S} p_s\left[\left(c_i^+ z_{is}^+ + c_i^- z_{is}^-\right) + \left(q_i^+ y_{is}^+ + q_i^- y_{is}^-\right)\right] \\ &\quad - \theta_3 \sum_{i\in I}\sum_{s\in S} p_s \left| \left(c_i^+ z_{is}^+ + c_i^- z_{is}^-\right) + \left(q_i^+ y_{is}^+ + q_i^- y_{is}^-\right) - \sum_{i\in I}\sum_{s\in S} p_s \begin{bmatrix} \left(c_i^+ z_{is}^+ + c_i^- z_{is}^-\right) \\ + \left(q_i^+ y_{is}^+ + q_i^- y_{is}^-\right) \end{bmatrix} \right| \end{aligned}$$

(9.61)

s.t. deterministic constraints (first stage) **(8.1)**, **(8.2)**, **(8.8)–(8.19)**, and **(8.26)**,

stochastic constraints (second stage): **(9.17)–(9.32)** and **(9.37)–(9.52)**.

where $\theta_1 < \theta_3$.

9.4.2 Comment on the Implementation of Risk Model III on GAMS

Since the absolute deviation is not differentiable at the singularity occurring at the inflection point, this calls for the use of a solver within the GAMS environment that is able to execute local optimization of a nonlinear program (NLP) with nonsmooth functions (Rardin, accessed on September 30, 2005). The default solver stipulated in GAMS for this class of problem is CONOPT 3 (Drud, 1996, http://www.gams.com/solvers/conopt.pdf, accessed on January 10, 2006), which is also, incidentally, the GAMS default solver for NLP. Hence, the default solve statement in GAMS can be used without the need to specify the type of solver to be CONOPT 3.

9.4.3 Computational Results for Risk Model III

Table 9.27 records the computational results for the implementation of Risk Model III on GAMS for a range of values for θ_1, the tradeoff factor for variability in profit measured by variance and for θ_3, the tradeoff factor for variability in the recourse penalty costs measured by the mean-absolute deviation (MAD). This is followed by a set of detailed results for which the case of $\theta_1 = 0.000\,8$ and $\theta_3 = 0.01$ is considered to be representative, with a number of different parameters enumerated to investigate particular trends and patterns that potentially contribute to robustness in the proposed model and solution. The associated computational statistics describing the problem size and the efficient distribution of computational expense is summarized in Table 9.28. Figure 9.6 then depicts the Markowitz's efficient frontier plot of expected profit versus risk imposed by deviations in both profit and the recourse penalty costs. Finally, the results are analyzed and discussed.

Table 9.27. Computational results for Risk Model III for a selected representative range of values of θ_1 and θ_3

Risk factors θ_1, θ_3 ($\theta_1 < \theta_3$)	Optimal objective value	Expected deviation between profit $V(z_0)$	Expected total unmet demand/ production shortfall	Expected total excess production/ production surplus	Expected recourse penalty costs E_s	Deviation between recourse penalty costs $W(p_s)$	$\sigma = \sqrt{V(z_0) + W(p_s)}$	Expected profit $E[z_0]^a$	$\mu = \lvert E[z_0] - E_s \rvert$	$C_v = \frac{\sigma}{\mu}$ Stochastic	Deterministic
0, 0	27 717.8688	1.3352E+8	2575.000	26 815.000	54 056.875	2 618.437	11 555.360	81 774.744	27 717.869	0.416 892 078	(infeasible)
0.000 001, 0.000 1	27 584.0832	1.3352E+8	2575.000	26 815.000	54 056.875	2 618.437	11 555.360	81 774.744	27 717.869	0.416 892 078	(infeasible)
0.000 01, 0.001	26 380.0130	1.3352E+8	2575.000	26 815.000	54 056.875	2 618.437	11 555.360	81 774.744	27 717.869	0.416 892 078	(infeasible)
0.000 01, 0.00316	26 374.3572	1.3352E+8	2575.000	26 815.000	54 056.875	2 618.437	11 555.360	81 774.744	27 717.869	0.416 892 078	(infeasible)
0.000 1, 0.001	14 362.8776	1.3352E+8	2575.000	26 815.000	54 056.875	2 618.437	11 555.360	81 774.744	27 717.869	0.416 892 078	(infeasible)
0.000 1, 0.01	14 339.3116	1.3352E+8	2575.000	26 815.000	54 056.875	2 618.437	11 555.360	81 774.744	27 717.869	0.416 892 078	(infeasible)
0.000 2, 0.01	2 801.8097	9.6341E+7	2575.000	17 800.529	40 232.400	2 627.927	9 815.479	62 328.689	22 096.289	0.444 213 913	(infeasible)
0.000 3, 0.01	−6 832.2901	9.6341E+7	2575.000	17 800.529	40 232.400	2 627.927	9 815.479	62 328.689	22 096.289	0.444 213 913	(infeasible)
0.000 4, 0.01	−16 397.3532	8.9590E+7	3525.000	16 661.810	42 444.168	3 827.126	9 465.385	61 920.958	19 476.790	0.485 982 803	(infeasible)
0.000 5, 0.01	−25 311.7774	8.8371E+7	3610.000	16 373.039	42 198.208	4 252.945	9 400.832	61 114.658	18 916.450	0.496 965 974	(infeasible)
0.000 6, 0.01	−34 148.9170	8.8371E+7	3610.000	16 373.039	42 198.208	4 252.945	9 400.832	61 114.658	18 916.450	0.496 965 974	(infeasible)
0.000 7, 0.01	−42 800.1921	7.9660E+7	5555.807	15 466.745	47 553.299	4 255.947	8 925.503	60 557.911	13 004.612	0.686 333 664	(infeasible)
0.000 8, 0.001	−49 888.4004	6.3920E+7	9424.122	13 665.010	58 199.357	4 261.916	7 995.254	59 451.082	1 251.725	6.387 388 604	(infeasible)
0.000 8, 0.01	−49 926.7577	6.3920E+7	9423.998	13 665.067	58 199.017	4 261.916	7 995.284	59 451.117	1 252.101	6.385 494 461	(infeasible)
0.000 9, 0.01	−56 172.9937	6.2367E+7	9836.111	13 473.118	59 333.201	4 262.552	7 897.553	59 333.201	0	$\to \infty$	(infeasible)
0.001, 0.01	−62 409.7013	6.2367E+7	9836.111	13 473.118	59 333.201	4 262.552	7 897.553	59 333.201	0	$\to \infty$	(infeasible)
0.001, 0.0316	−62 501.7724	6.2367E+7	9836.111	13 473.118	59 333.201	4 262.552	7 897.553	59 333.201	0	$\to \infty$	(infeasible)
0.01, 0.1	−62 502.0179	6.2367E+7	9836.111	13 473.118	59 333.201	4 262.552	7 897.553	59 333.201	0	$\to \infty$	(infeasible)
0.02, 0.1	−1 247 767.7704	6.2367E+7	9836.111	13 473.118	59 333.201	4 262.552	7 897.553	59 333.201	0	$\to \infty$	(infeasible)
0.1, 0.3162	−6 238 055.3951	6.2367E+7	9836.111	13 473.118	59 333.201	4 262.552	7 897.553	59 333.201	0	$\to \infty$	(infeasible)
0.5, 0.7071	−31 186 551.9317	6.2367E+7	9836.111	13 473.118	59 333.201	4 262.552	7 897.553	59 333.201	0	$\to \infty$	(infeasible)
1, 1	−62 371 338.3145	6.2367E+7	9836.111	13 473.118	59 333.201	4 262.552	7 897.553	59 333.201	0	$\to \infty$	(infeasible)
40, 20	−2 494 768 281.5457	6.2367E+7	9836.111	13 473.118	59 333.201	4 262.552	7 897.553	59 333.201	0	$\to \infty$	(infeasible)
100, 10	−6 236 750 201.7930	6.2367E+7	9836.111	13 473.118	59 333.201	4 262.552	7 897.553	59 333.201	0	$\to \infty$	(infeasible)
1000, 100	−62 367 502 017.9260	6.2367E+7	9836.111	13 473.118	59 333.201	4 262.552	7 897.553	59 333.201	0	$\to \infty$	(infeasible)
10 000, 100	(GAMS output: `Infeasible solution. A free variable exceeds the allowable range.`)										(infeasible)

Table 9.28. Detailed computational results for Risk Model III for $\theta_1 = 0.000\ 8$, $\theta_3 = 0.01$

First-Stage Variable	Stochastic Solution	Product (i)	Production Shortfall z_{ij}^{+} or Surplus z_{ij}^{-} (t/d)					
			Scenario 1		Scenario 2		Scenario 3	
			z_{i1}^{+}	z_{i1}^{-}	z_{i2}^{+}	z_{i2}^{-}	z_{i3}^{+}	z_{i3}^{-}
x_1	10 282.158	Demands RHS Coefficients Randomness						
x_2	2 000.000	Gasoline (2)	835.000	0	700.000	0	565.000	0
x_3	1 155.000	Naphtha (3)	0	0	0	55.000	0	110.000
x_4	2 493.423	Jet Fuel (4)	0	78.423	0	193.423	0	308.423
x_5	1 700.000	Heating Oil (5)	85.000	0	0	0	0	85.000
x_6	7 087.001	Fuel Oil (6)	2887.999	0	2412.999	0	1937.999	0
x_7	2 155.000							
x_8	3 177.187	Production Yields LHS Coefficients Randomness						
x_9	2 981.826	Naphtha (7)	0	751.485	0	818.319	0	885.153
x_{10}	3 752.988	Jet Fuel (4)	0	873.983	0	951.100	0	1028.216
x_{11}	1 000.000	Gas Oil (8)	0	802.008	0	915.112	0	1028.216
x_{12}	1 275.000	Cracker Feed (9)	0	822.573	0	925.394	0	1028.216
x_{13}	1 902.187	Residuum (10)	0	1028.216	0	668.340	0	308.465
x_{14}	2 500.000							
x_{15}	481.826	E(Penalty Costs)		22 500.614		25 607.063		10 091.340
x_{16}	1 000.000	E_{total}	58 199.017					
x_{17}	1 375.000							
x_{18}	425.000							
x_{19}	950.000							
x_{20}	125.000							
Expected Profit Z ($/day)	59 451.117							

Table 9.29. Computational statistics for Risk Model III

Solver	Single continuous variables	Constraints	Resource usage/ CPU time (s)	Iterations
CONOPT 3	96	91	≈ (0.049–0.100)	14

9.4.4 Analysis of Results for Risk Model III

As is the case with Risk Model II, the values of θ_1 and θ_3 denotes the importance of risk in the model as contributed by variation in profit and variation in recourse penalty costs, respectively, in comparison with the corresponding expected values of the model's objective.

From Table 8.31, similar trends with the expected value models are observed in which reducing values of θ_1 implicates higher expected profit. With increasingly smaller θ_1, the increment in expected profit becomes constant; the same constancy trend is also observed in the initially declining expected profit for increasing values of θ_1.

One of the reasons the pair of reducing values of θ_1 with fixed value of θ_3 leads to increasing expected profit is attributable to the general decrement in production shortfalls but increasing production surpluses, in which the fixed penalty cost for the latter is lower than the former. Based on its conceptual definition (as mentioned earlier), it certainly augurs well to select a higher operating value of θ_1 to achieve both high model feasibility as well as increased profit. Moreover, higher values of θ_1 alto decreasing variation in the recourse penalty costs, which implies solution robustness. This argument is further strengthened by the decreasing values of the coefficient of variation, which indicates less uncertainty in the model on a whole. This again demonstrates that a proper selection of the operating range of θ_1 and θ_3 is crucial in varying the tradeoffs between the desired degree of model robustness and solution robustness, to ultimately obtain optimality between expected profit and expected production feasibility.

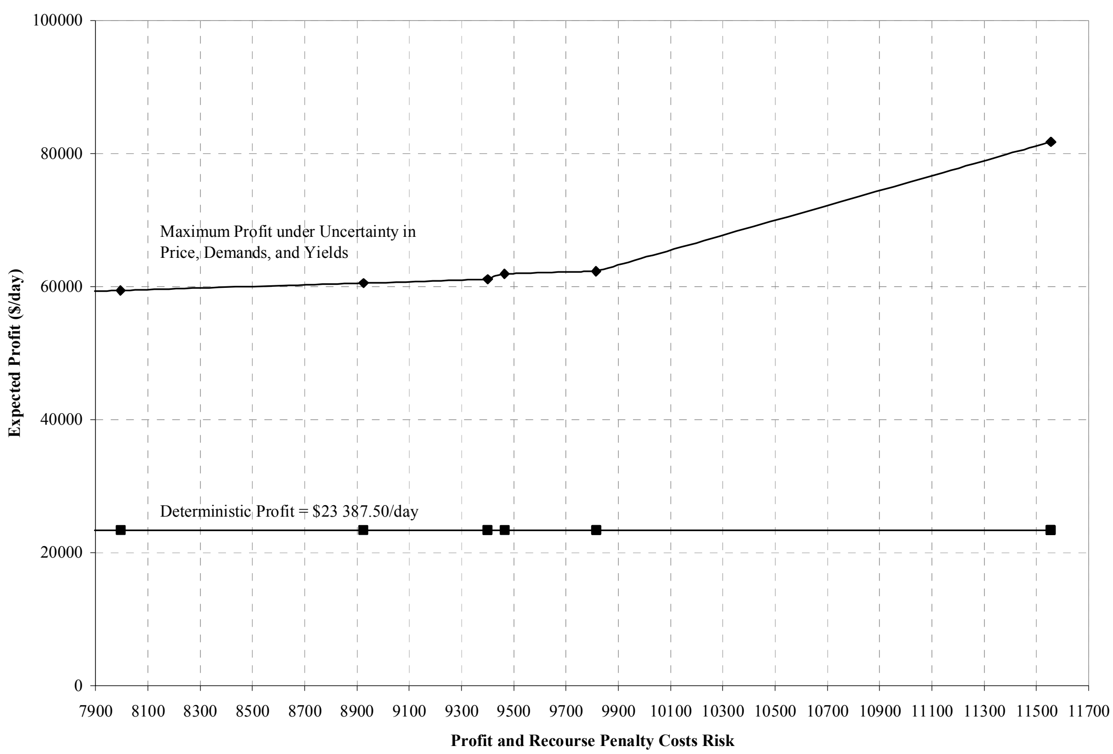


Figure 9.10. The efficient frontier plot of expected profit versus risk imposed by variations in both profit (measured by variance) and the recourse penalty costs (measured by mean-absolute deviation) for Risk Model III

9.5 SUMMARY OF RESULTS AND COMPARISON AGAINST RESULTS FROM THE FUZZY LINEAR FRACTIONAL GOAL PROGRAMMING APPROACH

Ravi and Reddy (1998) adopted the fuzzy linear fractional goal programming approach to the same deterministic linear program for refinery operations planning by Allen (1971) that is used for the present numerical example. Based on their arguments that a decision maker is very often more interested in the optimization of ratios, they identified the ratios or the fractional goals to be treated as fuzzy goals for optimization as the following:

$$f_1 = \frac{\text{profit}}{\text{capacity of the primary unit } (x_1)} \left(\Rightarrow f_1 = \frac{Z}{x_1} \right) \quad \textbf{(9.62)}$$

$$f_2 = \frac{\text{profit}}{\text{capacity of the cracker } (x_{14})} \left(\Rightarrow f_1 = \frac{Z}{x_{14}} \right) \quad \textbf{(9.63)}$$

for which the deterministic model of Allen (1971) is then reformulated accordingly. Table 9.30 presents a summary of the results obtained from the base case deterministic model and the five different approaches adopted in applying the methods of stochastic optimization, to be compared against results from the fuzzy linear programming approach.

Ravi and Reddy (1998) advocated that their results, although registering 1.3 percent less profit than the linear programming approach of Allen (1971), yielded higher optimal fractional goal values, thus translating to better ratio values of profit/capacity of primary unit capacity f_1 and profit/capacity of the cracker unit f_2 simultaneously. Following the same premise, we conclude that the stochastic models, in addition to proposing midterm plans that consistently register higher expected profits than both the deterministic and the fuzzy programming models, assure the decision maker of good ratios for both f_1 and f_2. Maximization of these ratios lead to maximum or near maximum profit, with minimum or near minimum capacities of the respective units simultaneously. As petroleum refiners use the technology of fluidized bed catalytic crackers (FCC) to convert more crude oil to blending stocks for producing gasoline, which is unarguably the most commercially attractive end product of the refining activity, it certainly augurs well for the stochastic models to have a good ratio of maximum profit with regards to a minimum capacity of the highly capital-intensive FCC unit. Moreover, about 45 percent of worldwide gasoline production is contributed by FCC processes and its ancillary unit. Thus, the overall economic performance of a refinery considerably hinges on a prudent investment in FCC due to its large throughput, high product-feed upgrade, and mercantile importance (Alvarez-Ramirez et al., 2004).

Table 9.30. Comparison of results obtained from the deterministic model, the stochastic models, and the fuzzy linear fractional goal program by Ravi and Reddy (1998)

First-Stage Variable	Deterministic Linear Program	Stochastic Risk Model I	Stochastic Expectation Model I (for $\theta_1 = 0.000\ 03$)	Stochastic Expectation Model II	Stochastic Risk Model II (for $\theta_1 = 0.000\ 000\ 1$, $\theta_2 = 50$)	Stochastic Risk Model III (for $\theta_1 = 0.000\ 8$, $\theta_3 = 0.01$)	Fuzzy Linear Fractional Goal Program*
x_1	12 500	12 171.621	15 000.000	2792.531	15 000.000	10 282.158	12 054.59
x_2	2 000	1 947.459	2 000.000	647.867	2 000.000	2 000.000	1 928.74
x_3	625	608.581	1 155.000	300.197	1 155.000	1 155.000	602.73
x_4	1 875	1 825.743	3 637.500	677.189	3 637.500	2 493.423	1 808.19
x_5	1 700	1 700.000	3 835.000	1150.523	3 597.500	1 700.000	1 700.00
x_6	6 175	5 968.122	9 500.000	177.052	9737.500	7 087.001	5 894.39
x_7	1 625	1 582.311	2 155.000	624.131	2 155.000	2 155.000	1 567.09
x_8	2 750	2 677.757	4 635.000	862.892	4 635.000	3 177.187	2 652.01
x_9	2 500	2 434.324	4 350.000	809.834	4 350.000	2 981.826	2 410.92
x_{10}	3 750	3 651.486	5 475.000	1019.274	5 475.000	3 752.988	3 616.38
x_{11}	1 000	973.730	1 000.000	323.934	1 000.000	1 000.000	964.37
x_{12}	1 275	1 275.000	2 876.250	862.892	2 698.125	1 275.000	1 275.00
x_{13}	1 475	1 402.757	1 758.750	0	1 936.875	1 902.187	1 377.01
x_{14}	2 500	2 434.324	2 500.000	809.834	2 500.000	2 500.000	2 410.92
x_{15}	0	0	1 850.000	0	1 850.000	481.826	0.0
x_{16}	1 000	973.730	1 000.000	323.934	1 000.000	1 000.000	964.38
x_{17}	1 375	1 338.878	1 375.000	445.409	1 375.000	1 375.000	1 326.01
x_{18}	425	425.000	958.750	287.631	899.375	425.000	425.00
x_{19}	950	913.878	416.250	157.778	475.625	950.000	901.01
x_{20}	125	121.716	125.000	40.492	125.000	125.000	120.55
Profit ($/day)	23 387.50	23 500.000	81 774.744	23 387.50	79 725.713	59 451.117	23 069.06
f_1	1.871	1.931	5.452	8.375	5.315	5.782	1.914
f_2	9.355	9.654	32.710	28.879	31.890	23.780	9.569

* Taken from the work of Ravi and Reddy (1998)

9.6 ADDITIONAL REMARKS

In retrospect, we mildly caution the respected reader that this work primarily intends to demonstrate the validity of the stochastic concepts reviewed along with the methods developed and improvised by using a typical and realistic refinery planning problem under uncertainty. However, it is certainly not our claim that the model captures *all* detailed aspects of the problem but rather, it demonstrates the capabilities of the proposed tools.

CHAPTER 10

Conclusions

10.1 SUMMARY OF WORK

This thesis research focuses on the methodology of developing effective yet straightforward stochastic optimization models for the midterm production planning of a petroleum refinery by accounting for three major factors of uncertainty simultaneously, namely commodities' prices, market demands, and product yields. In addition, we consider the importance of risk in decision-making under uncertainty in the proposed stochastic planning models by explicitly accounting for the tradeoffs between expected profit and variation in both profit and the recourse penalty costs. These measures of risk (in the form of variance for the price coefficients and variance or Mean-Absolute Deviation (MAD) for the recourse penalty costs, adopted in different models) are incorporated within the general framework of the proposed models with the aim of achieving robustness in decision-making activities especially in view of the highly volatile hydrocarbon industry in which the petroleum refineries operate. The following four approaches are implemented, resulting correspondingly in four decision-making models under uncertainty:

1. the Markowitz's mean–variance (MV) model to handle randomness in the objective function coefficients of prices by minimizing the variance of the expected value or mean of the random coefficients, subject to a target profit constraint;
2. the two-stage stochastic programming with fixed recourse via scenario analysis approach to model randomness in the right-hand side and the left-hand side (technological) coefficients by minimizing the expected recourse penalty costs due to constraints' violations;
3. incorporation of the Markowitz's MV approach within the bilevel decision-making framework developed in the preceding approach to minimize both the expectation and the variance of the recourse penalty costs; and

4. reformulation of the model developed in the third approach by utilizing the Mean–Absolute Deviation (MAD) as the measure of risk imposed by the recourse penalty costs.

An exposition for justifying the rationale for adopting each of the four stochastic modelling methods is undertaken in this work by comprehensively surveying past successes in employing such approaches in the open literature, but in the light of the limitations of the approaches.

As emphasized throughout, the novelty of these approaches lies in the explicit modelling and formulation of uncertainties considered for the large-scale optimization problem of petroleum refinery production planning. This has been accomplished through the utilization of slack variables to account for violations of the stochastic constraints of possible scenarios of product demands and product yields, primarily within a bilevel decision-making framework of two-stage stochastic programming that incorporates the Markowitz's mean–variance model of portfolio selection optimization as a hedging tool against the presence of risk that arises due to variation in profit and the penalty costs of recourse actions.

10.2 MAJOR CONTRIBUTIONS OF THIS RESEARCH

The major contributions of this work are threefold. First, we formulate a slate of highly tractable stochastic optimization models through an explicit modelling of the presence of uncertainties for application in the production planning of petroleum refineries, primarily within a two-stage stochastic programming framework. This is accomplished via the systematic adoption of a hybrid of effective yet clear-cut stochastic optimization techniques that obviates the use of the conventional brute force approach of Monte Carlo-based methods.

Second, we consider the incorporation of the concept of Mean-Absolute Deviation (MAD) as a measure of risk for the petroleum refining activities, instead of the traditional measure using variance. The numerous benefits of doing so are elucidated in Section 5.8.1, with the most significant being the ability to linearize the MAD expression in the objective function, thus producing an equivalent linear programming problem that can be solved accurately and competently by harnessing the combination of efficient algorithms with the power of today's modern computing technology.

Additionally, in an effort to establish the effectiveness of the proposed stochastic models while simultaneously hedging against the various forms of uncertainties in commodity prices, market demands, and product yields, we carry out a series of extensive and rigorous computational experiments to investigate and interpret the behaviour and the overall robustness of the multiobjective optimization models. This is executed in light of similar methodologies that have been employed in previous works, notably the applications of Mulvey et al. (1995)'s robust optimization approach by Bok et al. (1998) for the capacity expansion of a petrochemicals processing network and by Malcolm and Zenios (1994) for the capacity expansion of power systems. Oftentimes, we also take into account many of the useful suggestions and guidelines that are offered by the now classical paper by Rardin and Lin (1982) on issues and techniques concerning test problems for computational experiments. Two performance metrics are thus considered, namely: (1) the concepts of solution robustness and model robustness as introduced by Mulvey et al. (1995) and (2) the coefficient of variation C_v, in order to gauge the performances of the four proposed stochastic models against each other, as well as against the deterministic model and the fuzzy programming approach of Ravi and Reddy (1998).

CHAPTER 11

Recommendations for Future Work and Way Forward

11.1 SUMMARY OF RECOMMENDATIONS FOR FUTURE WORK

In this chapter, we intend to bring together and consolidate under one roof, the numerous recommendations scattered over the main content of the dissertation. We are also inclined to offer some personal thoughts and opinions on a promising research agenda for future undertakings in this exciting area of developing stochastic programming tools and applications, for petroleum refinery operations management in specific and the wider spectrum of business decision-making in general. To provide a compact discussion, the essence of the recommendations and suggestions for future work are enumerated as follows:

1. to implement extensions to the present general deterministic midterm production planning model for petroleum refineries by considering the the following features:

- rigorous modelling of advanced petroleum refining process unit such as the hydrotreating and hydrocracking units that are instrumental especially in the current drive towards clean fuels production;
- capacity expansion through the installation of multiple number of processing units; an example with apparent economic value would be the installation of an additional piece of the fluidized bed catalytic cracker (FCCU) to increase the production of gasoline that is arguably the most profitable of refining end products;
- modelling nonlinearities in blending operations as constraints for production of fuels in satisfying the following operating objectives, which are typically prioritized in this order as advocated by Bodington (19950: (i) quality specifications; (ii) shipment schedule; (iii) quality giveaway; (iv) blending cost; and () inventory targets;

- multiferinery modelling of a network of refineries, with consideration for inter-refinery transfer or transportation of intermediates and products under subcontracting agreements, with the addition of production (supply) and distribution modelling; interactions with potential spin-off industries particularly petrochemicals processing will also be examined;
- incorporation of complex economic instruments such as royalties, taxes, and tariffs in the objective function of net present value (van den Heever, 2001; van den Heever et al., 2000; 2001);

2. improvements in methodology for dealing with uncertainty:

- implementation of multi-stage stochastic programming framework with the stages corresponding to time periods, hence effectively formulating multiperiod models, which is an established alternative strategy in dealing with uncertainty;
- incorporation of uncertainties in less-commonly considered but equally paramount and planning-sensitive factors such as (i) process variations as indicated by variable process parameters, for example, flow rates and temperatures; (ii) cancelled/rushed orders; (iii) equipment failure; (iv) technology obsolescence; and (v) sales uncertainty;
- improvement in effective procedures of scenario selection and generation, which is a highly essential (if not the most essential) key component towards developing stochastic programming models that are truly robust in the face of uncertainties; in particular, we intend to pursue the incorporation of the scenario planning paradigm (van der Heijden, 2005; Godet, 1987; Schoemaker and van der Heijden, 1992; Lindgren and Bandhold, 2003; Schwartz, 1991; Stokke et al., 1990; Huss and Honton, 1987) that has been so successfully practised at Royal Dutch/Shell, which is not only the leading oil-and-gas company worldwide but is even arguably, one of the most successful corporations of modern times,
- improvement in the theory and procedure for specifying the weights for multiobjective optimization to account for the tradeoffs between expected profit and risk;

3. incorporation of advanced metrics for risk measurement and modelling since the conventional metric of variance is a symmetric risk measure that desirably penalizes profits below the target levels but also undesirably penalizes profits above the target levels, other than the mean-absolute deviation (MAD) that is considered in this work with much resulting promise; the alternative risk measures to be considered will include the expected downside risk (Eppen et al., 1989), the Value-at-Risk (VaR) (Barbaro and Bagajewicz, 2004a; Mulvey, 2001), and the Conditional Value-at-Risk (CVaR) (Rockafellar and Uryasev, 2002).

11.2 FINAL REMARKS

We are convinced that there is still a plethora of research opportunities in this field with directions of research problems themselves muddled with uncertainty (ironically) and multitude of questions unanswered, as widely acknowledged by both academics and practitioners of the field. In advancing this stand, we have the privilege of support from none other than George (Bernard) Dantzig himself, fondly dubbed as the father of linear programming (primarily for devising the simplex method) and one of the pioneers of stochastic programming, with the following truly inspiring quote from Dantzig, which has and will continue to serve to motivate researchers, both current and potential, in this remarkably rich and fascinating field of stochastic optimization:

"Planning under uncertainty. This, I feel, is the real field we should all be working on." (http://www.e-optimization.com/directory/trailblazers/dantzig/interview_planning.cfm, accessed on March 2, 2006; http://www2.informs.org/History/dantzig/in_interview_irv10.htm, accessed on March 9, 2006).

Incidentally, to cap the wonderful journey of working on this thesis research, we take this opportunity to pay our utmost respect and homage to Dantzig (November 8, 1914–May 13, 2005), whose recent demise last year would certainly carve a legacy that will long survive him in continuing to spur the area of mathematical programming/optimization/operations research towards attaining greater heights and meaning. (As an aside, "optimization" is the term strongly preferred and even advocated by George L. Nemhauser, truly one of the giants of the field and co-author with Laurence A. Wolsey of the biblical encyclopaedia of the discrete sub-area of this field entitled *Integer and Combinatorial Optimization* (Nemhauser, 1994).)

References and Literature Cited

To borrow a similar principle from perhaps one of the most highly readable advanced level text in mathematics as *Lectures on Polytopes* (Springer-Verlag, 1995), beautifully penned by Gunter M. Ziegler (which addresses a very useful theory for the strong formulations of integer programming problems), I have attempted to compile a bibliography on production planning focussing on problems under uncertainty with emphasis on chemical process systems engineering (PSE)-related problems, with REAL NAMES, to illustrate that behind all of these mathematics (some of it spectacularly beautiful yet functional), there are actual REAL PEOPLE. In the few cases where I could not find more than initials, perhaps just assume that that is all that they have (with this to be taken in a light tone).

Acevedo, J. and Efstratios N. Pistikopoulos. Stochastic optimization based algorithms for process synthesis under uncertainty. *Computers & Chemical Engineering*. 22 (1998): 647–671.

Ahmed, Shabbir and Nikolaos Vasili Sahinidis. Robust process planning under uncertainty. *Industrial & Engineering Chemistry Research* 37 (1998): 1883–1892.

Allen, D. H. Linear programming models for plant operations planning. *British Chemical Engineering*.16 (1971): 685–691.

Alvarez-Ramirez, Jose, Jesus Valencia, and Hector Puebla. Multivariable control configurations for composition regulation in a fluid catalytic cracking unit. *Chemical Engineering Journal* 99 (2004): 187–201.

Applequist, George Einar. Economic risk management for chemical manufacturing supply chain planning. *PhD Thesis*. West Lafayette, Indiana, USA: Purdue University, 2002.

Applequist, George Einar, Joseph F. Pekny, and Gintaras V. Reklaitis. Risk and uncertainty in managing chemical manufacturing supply chains. *Computers & Chemical Engineering* 24 (2000): 2211–2222.

Applequist, George Einar, O. Samikoglu, Joseph Pekny, and Gintaras V. Reklaitis. Issues in the use, design, and evolution of process scheduling and planning systems. *ISA Transactions* 36 (1997), 2: 81–121.

Aseeri, Ahmed and Miguel J. Bagajewicz. New measures and procedures to manage financial risk with applications to the planning of gas commercialization in Asia. *Computers & Chemical Engineering* 28 (2004): 2791–2821.

Aseeri, Ahmed, Patrick Gorman, and Miguel J. Bagajewicz. Financial risk management in offshore oil infrastructure planning and scheduling. *Industrial & Engineering Chemistry Research* 43 (2004), 12: 3063–3072.

Balasubramanian, J. and Ignacio E. Grossmann. Approximation to multistage stochastic optimization in multiperiod batch plant scheduling under demand uncertainty. *Industrial & Engineering Chemistry Research*. 43 (2004): 3695–3713.

Barbaro, Andres and Miguel J. Bagajewicz. Managing financial risk in planning under uncertainty. *AIChE Journal* 50 (May 2004a), 5: 963–989.

Barbaro, Andres and Miguel J. Bagajewicz. Use of inventory and option contracts to hedge financial risk in planning under uncertainty. *AIChE Journal* 50 (May 2004b), 5: 963–989.

Bassett, M. H., P. Dave, Francis J. Doyle III, G. K. Kudva, Joseph F. Pekny. Gintaras V. Reklaitis, S. Subrahmanyam, D. L. Miller, and M. G. Zentner. Perspectives on model based integration of process operations. *Computers & Chemical Engineering* 20 (1996), 6/7: 821–844.

Bassett, M. H., Joseph F. Pekny, and Gintaras V. Reklaitis. Using detailed scheduling to obtain realistic operating policies for a batch processing facility. *Industrial & Engineering Chemistry Research* 36 (1997): 1717–1726.

Baumol, William J. Activity analysis in one lesson. *American Economic Review*. 48 (1958), 5: 837 – 873.

Beale, E. On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society, Series B (Methodological*) 17 (1955), 2: 173–184.

Bellman, R. and Lotfi Asker Zadeh. Decision-making in a fuzzy environment. *Management Science* 17 (1970): 141–161.

Benders, J. F. Partitioning procedures for solving mixed-variables programming problems. *Numerische, Mathematik* 4 (1962): 238–252.

Ben-Tal, Aharon. and Arkadi Nemirovski. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming* 88 (2000): 411–424.

Biegler, Lorenz T. From nonlinear programming theory to practical optimization algorithms: a process engineering viewpoint. *Computers & Chemical Engineering* 17 (1993), Supplement 1: 63–80.

Birewar, Deepak B. and Ignacio E. Grossmann. Simultaneous production planning and scheduling in multiproduct plants. *Industrial & Engineering Chemistry Research* 29 (1990), 4: 570–580.

Birge, John R. and Francois Louveaux. *Introduction to Stochastic Programming*. New York: Springer-Verlag New York, Inc., 1997.

Birge, John R. and Roger J.-B. Wets. (editors). Preface for Proceedings of the 5th International Conference on Stochastic Programming. *Annals of Operations Research* 30 & 31 (1991).

Birge, John R. The value of the stochastic solution in stochastic linear programs with fixed recourse. *Mathematical Programming* 24 (1982): 314–325.

Birge, John R. Stochastic programming computation and applications. *INFORMS Journal on Computing* 9 (1997), 2: 111–133.

Birge, John R. *Personal Communication*, 2005.

Bitran, Gabriel. R. and Hax, Arnaldo C. On the design of the hierarchical production planning systems. *Decision Sciences* 8 (1977): 28 – 55.

Bodington, C. Edward. Management. In *Planning, Scheduling, and Control Integration in the Process Industries*. Edited by C. Edward Bodington. New York: McGraw-Hill, 1995.

Bodington, C. Edward. *Planning, Scheduling, and Control Integration in the Process Industries*. New York: McGraw-Hill, 1995.

Bok, Jin-Kwang, Heeman Lee, and Sunwon Park. Robust investment model for long-range capacity expansion of chemical processing networks under uncertain demand forecast scenarios. *Computers & Chemical Engineering* 22 (1998): 1037–1049.

Bok, Jin-Kwang., Ignacio E. Grossmann, and Sunwon Park. Supply chain optimization in continuous flexible processes. *Industrial & Engineering Chemistry Research* 39 (2000): 1279–1290.

Bonfill, Anna, Miguel Bagajewicz, Antonio Espuna, and Luis Puigjaner. Risk management in the scheduling of batch plants under uncertain market demand. *Industrial & Engineering Chemistry Research* 43 (2004): 741–750.

Bopp, A. E., V. R. Kannan, S. W. Palocsay, and S. P. Stevens. An optimization model for planning natural gas purchases, transportation, storage, and deliverability. *Omega: International Journal of Management Science* 24 (1996), 5: 511–522.

Brauers, J. and M. Weber. A new method of scenario analysis for strategic planning. *Journal of Forecasting* 7 (1988): 31–47.

Brooke, A., D. Kendrick, A. Meeraus, and Ramesh Raman. *GAMS—A User's Guide: Tutorial by Richard E. Rosenthal*. http://www.gams.com/docs/gams/GAMSUsersGuide.pdf (accessed December 9, 2005) Washington DC, USA: GAMS Development Corporation, December 1998.

Cabrini, Silvina M., Brian G. Stark, Hayri Onal, Scott H. Irwin, Darrel L. Good, and Joao Martines-Filho. Efficiency analysis of agricultural market advisory services: a nonlinear mixed-integer programming approach. *Manufacturing & Service Operations Management* 6 (Summer 2004), 3: 237–252.

ChemLOCUS Chem Journal: Korea Chemical Market Information. *Naphtha Remained Firm at $340–350*. http://www.chemlocus.com/news/daily_read.htm?menu=D1&Sequence=6538&cpage=14&sub= (accessed on January 24, 2006).

Cheng, Lifei, Eswaran Subrahmanian, and Arthur W. Westerberg. Multiobjective decision processes under uncertainty: applications, problem formulations, and solution strategies. *Industrial & Engineering Chemistry Research* 44 (2005): 2405–2415.

Cheng, L., Eswaran Subrahmanian, and Arthur W. Westerberg. Design and planning under uncertainty: issues on problem formulation and solution. *Computers & Chemical Engineering* 27 (2003): 781–801.

Clay, R. L. and Ignacio E. Grossmann. Optimization of stochastic planning models. *Transaction of the Institution of Chemical Engineers (IChemE)* 72 (May 1994), Part A: 415–419.

Clay, R. L. and Ignacio E. Grossmann. A disaggregation algorithm for the optimization of stochastic planning models. *Computers & Chemical Engineering* 21 (1997): 751–774.

Charnes, Abraham. and William. W. Cooper. Chance-constrained programming. *Management Science* 6 (October 1959), 1: 73–79.

Dantus, Mauricio. M. and Karen A. High. Evaluation of waste minimization alternatives under uncertainty: a multiobjective optimization approach. *Computers & Chemical Engineering* 23 (1999): 1493–1508.

Dantzig George Bernard. Linear programming under uncertainty. *Management Science* 1 (1955), 3 & 4: 197–206.

Dantzig George Bernard and Gerd Infanger. Multi-stage stochastic linear programs for portfolio optimization. *Annals of Operations Research* 45 (1993): 59–76.

Dantzig, George Bernard and P. Wolfe. The decomposition principle for linear programs. *Operations Research* 8 (1960): 101–111.

Dempster, Michael Alan Howarth. Introduction to stochastic programming. In *Stochastic Programming*. Edited by M. A. H. Dempster. New York: Academic Press, 1980, pp. 3–59.

Dempster, Michael Alan Howarth, N. Hicks Pedron, E. A. Medova, J. E. Scott, and A. Sembos. Planning logistics operations in the oil industry. *Journal of the Operational Research Society* 51 (2000), 11: 1271–1288.

Didziulis, Vytis. Linkages between the markets for crude oil and the markets for refined products. *Ph.D Thesis (Advisor: Eugene Kroch)*. University of Pennsylvania, USA, 1990.

Dimitriadis, A. D., Nilay Shah, and C. C. Pantelides. RTN-based rolling horizon algorithms for medium term scheduling of multipurpose plants. *Computers & Chemical Engineering* S21 (1997): S1061–S1066.

Diwekar, U. M. and Kalagnanam, J. R. Efficient sampling technique for optimization under uncertainty. *AIChE Journal* 43 (February 1997), 2: 440–447.

Diwekar, Urmilla M. Synthesis approach to the determination of optimal waste blends under uncertainty.

Drud, Arne. *CONOPT: A System for Large Scale Nonlinear Optimization, Reference Manual for CONOPT Subroutine Library*. Bagsvaerd, Denmark: ARKI Consulting and Development A/S, 1996. http://www.gams.com/solvers/conopt.pdf (accessed on January 10, 2006).

Dupacova, Jitka. Applications of stochastic programming: achievements and questions. *European Journal of Operational Research* 140 (2002): 281–290.

Dutta, D., R. N. Tiwari, and J. R. Rao. Multiple objective linear fractional programming—a fuzzy set theoretic approach. *Fuzzy Sets & Systems* 52 (1992), 1: 39–45.

Dyer, Martin and Leen Stougie. Computational complexity of stochastic programming problems. *Mathematical Programming, Series A* 106 (2006): 423–432.

Edgar, Thomas F., David Mautner Himmelblau, and Leon S. Lasdon. *Optimization of Chemical Processes*. Second edition. New York, USA: McGraw-Hill, 2001.

Energy Information Administration, Department of Energy, United States of America. *World Oil Market and Oil Price Chronologies: 1970–2004.* http://www.eia.doe.gov/emeu/cabs/chron.html, accessed December 27, 2005.

ENSPM Formation Industrie. *Training Course Notes on Refining Processes*. France: ENSPM Formation Industrie, IFP Training, 1993.

Eppen, Gary D., R. Kipp Martin, and Linus Schrage. A scenario approach to capacity planning. *Operations Research* 37 (1989), 4: 517–527.

Ermoliev, Yuri M. and Wets, Roger J. B.. *Numerical Techniques for Stochastic Optimization*. Virginia, United States of America: Springer-verlag Berlin Heidelberg New York, 1988.

Escudero, Laureano F., Pasumarti V. Kamesam, Alan J. King, and Roger J.-B. Wets. Production planning via scenario modelling. *Annals of Operations Research* 43 (1993): 311–335.

Escudero, Laureano F., F. J. Quintana, and J. Salmeron. CORO, a modeling and an algorithmic framework for oil supply, transformation, and distribution optimization under uncertainty. *European Journal of Operational Research* 114 (1999), 3, 638–656.

Evers, W. H. A new model for stochastic linear programming. *Management Science* 13 (1967): 680.

Favennec, Jean-Pierre and Alain Pigeyre. Refining: a technical summary, investments, margins, costs, probable future developments. In *Refinery Operation and Management. Volume 5 of Petroleum Refining Series*. Edited by Jean-Pierre Favennec. Paris: Editions Technip, 2001.

Favennec, Jean-Pierre (editor). *Refinery Operation and Management. Volume 5 of Petroleum Refining Series*. Paris: Editions Technip, 2001.

Fisher, James N. and James W. Zellhart. Planning. In *Planning, Scheduling, and Control Integration in the Process Industries*. Edited by C. Edward Bodington. New York: McGraw-Hill, 1995.

Fourer, Robert. *Linear Programming: Software Survey*. http://www.lionhrtpub.com/orms/orms-12-03/frlp.html (accessed on December 15, 2005).

Friedman, Y. and Gintaras V. Reklaitis. Flexible solutions to linear programs under uncertainty: inequality constraints. *AIChE Journal*. 21 (1975): 77–83.

G., A. Jimenez and Dale F. Rudd. Use of a recursive mixed-integer programming model to detect an optimal integration sequence for the Mexican petrochemical industry. *Computers & Chemical Engineering* 11 (1987), 3: 291–301.

GAMS Corporation. *Solver Descriptions for GAMS/CPLEX*. http://www.gams.com/solvers/solvers.htm#CPLEX, accessed February 12, 2006.

Gary, James H. and Glenn E. Handwerk. *Petroleum Refining: Technology and Economics*. Third edition. New York: Marcel Dekker, Inc., 1994.

Gassmann, Horand I. Modelling support for stochastic programs. *Annals of Operations Research* 82 (1998): 107–137.

Geoffrion, A. M. Generalized Bender's decomposition. *Journal of Optimization Theory and Applications* 10 (1972): 237–260.

Glynn, Peter W. and Donald L. Iglehart. Importance sampling for stochastic simulations. *Management Science* 35 (1989): 1367–1391.

Godet, Michel. *Scenarios and Strategic Management*. London: Butterworth Scientific Ltd, 1987.

Goel, Vikas and Grossmann, Ignacio E. A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. *Computers & Chemical Engineering* 28 (2004): 1409–1429.

Grossmann, Ignacio E. Introduction to GAMS. *Chemical Engineering Optimization Models with GAMS, CACHE Design Case Studies Series, Case Study No. 6* (1991). http://www.che.boun.edu.tr/che477/gms-mod.html (accessed on September 30, 2005).

Grossmann, Ignacio E. Enterprise-wide optimization: a new frontier in process systems engineering. *AIChE Journal* 51 (2005), 7: 1846–1857.

Grossmann, Ignacio E. and R. W. H. Sargent. Optimum design of chemical plants with uncertain parameters. *AIChE Journal* 24 (1978), 6: 1021–1028.

Grossmann, Ignacio E., and Roger W. H. Sargent. Optimum design of multipurpose chemical plants. *Industrial & Engineering Chemistry Process Design & Development* 18 (1979): 343 – 348.

Grossmann, Ignacio E., K. P. Halemane, and Ross E. Swaney. Optimization strategies for flexible chemical processes. *Computers & Chemical Engineering*. 7 (1983), 4: 439–462.

Grossmann, Ignacio E. and Jorge Santibanez. Application of mixed-integer linear programming in process synthesis. *Computers & Chemical Engineering* 4 (1980): 205–214.

Grossmann, Ignacio E., Susara A. van den Heever, and Iiro Harjunkoski. Discrete optimization methods and their role in the integration of planning and scheduling. *Proceedings of Chemical Process Control Conference 6 at Tucson,* 2001. Grossmann Research Group at http://egon.cheme.cmu.edu/papers.html (accessed on December 10, 2005)

Guillen, G., Miguel Bagajewicz, S. E. Sequeira, A. Espuna, and Luis Puigjaner. Management of pricing policies and financial risk as a key element for short term scheduling optimization. *Industrial & Engineering Chemistry Research* 44 (2005), 3: 557–575.

Guldmann, J. M. and F. Wang. Optimizing the natural gas supply mix of local distribution utilities. *European Journal of Operational Research* 112 (1999): 598–612.

Gupta, Anshuman and Costas D. Maranas. A hierarchical Lagrangean relaxation procedure for solving midterm planning problems. *Industrial & Engineering Chemistry Research* 38 (1999): 1937–1947.

Gupta, Anshuman and Costas D. Maranas. A two-stage modeling and solution framework for multisite midterm planning under demand uncertainty. *Industrial & Engineering Chemistry Research* 39 (2000): 3799–3813.

Gupta, Anshuman and Costas D. Maranas. Managing demand uncertainty in supply chain planning. *Computers & Chemical Engineering* 27 (2003): 1219–1227.

Gupta, Anshuman, Costas D. Maranas, and Conor M. McDonald. Mid-term supply chain planning under demand uncertainty: customer demand satisfaction and inventory management. *Computers & Chemical Engineering* 24 (2000): 2613–2621.

Hartmann, J. C. M. Distinguish between scheduling and planning models. *Hydrocarbon Processing* 77 (1998), 7: 5–9.

Heijden, Kees van der. *Scenarios: The Art of Strategic Conversation*. Second edition. John Wiley & Sons, 2005.

Henrion, Rene, Pu Li, Andris Moller, Marc C. Steinbach, Moritz Wendt, and Gunter Wozny. Stochastic optimization for operating chemical processes under uncertainty. In *Online Optimization of Large Scale Systems*. Edited by Martin Grötschel, Sven O. Krumke, Jörg Rambau. Berlin, New York: Springer, 2001.

Higle, Julia L. Variance reduction and objective function evaluation in stochastic linear programs. *INFORMS Journal on Computing* 10 (Spring 1998), 2: 236 – 247.

Higle, Julia L. Stochastic programming: optimization when uncertainty matters. *Chapter 1, The INFORMS Tutorials in Operations Research (OR) Series*. tucson.sie.arizona.edu/faculty/higle/images/pdf/ HigleInforms05.pdf (accessed on November 11, 2005). Presented at the Institute for Operations Research and the Management Science (INFORMS) Annual Meeting at New Orleans and San Francisco, November 13–16, 2005.

Higle, Julia L. and Suvrajeet Sen. *Stochastic Decomposition: A Statistical Method for Large Scale Stochastic Linear Programming*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1996.

Higle, Julia L. and Stein W. Wallace. Sensitivity analysis and uncertainty in linear programming. *Interfaces* 33 (2003), 4: 53–60.

Himmelblau, David M. and T. C. Bickel. Optimal expansion of a hydrodesulfurization process. *Computers & Chemical Engineering* 4 (1980): 101–112.

Hirsch, Robert L., Roger Bezdek, and Robert Wendling. Peaking of world oil production and its mitigation. *AIChE Journal* 52 (January 2006), 1: 2–8.

Horner, Peter. Planning under uncertainty: questions and answers with George Dantzig. *OR/MS Today*. 26 (1999): 26–30.

Hsieh, S. and C. C. Chiang. Manufacturing-to-sale planning model for fuel oil production. *International Journal of Advanced Manufacturing Technology* 18 (2001): 303–311.

Huss, William R. and Edward J. Honton. Scenario planning—what style should you use? *Long Range Planning* 20 (1987), 4: 21–29.

Ierapetritou, M. G., J. Acevedo, and E. N. Pistikopoulos. An optimization approach for process engineering problems under uncertainty. *Computers & Chemical Engineering* 20 (1996), 6/7: 703–709.

Ierapetritou, Marianthi G, and Efstratios N. Pistikopoulos. Novel optimization approach of stochastic planning models. *Industrial & Engineering Chemistry Research* 33 (1994a): 1930–1942.

Ierapetritou, Marianthi G. and Efstratios N. Pistikopoulos. Simultaneous incorporation of flexibility and economic risk in operational planning under uncertainty. *Computers & Chemical Engineering* 18 (1994b), 3: 163–189.

Ierapetritou, Marianthi G, and Efstratios N. Pistikopoulos. Batch plant design and operations under uncertainty. *Industrial & Engineering Chemistry Research* 35 (1996), 3: 772–787.

Ierapetritou, Marianthi G., Efstratios N. Pistikopoulos, and Christodoulos A. Floudas. Operational planning under uncertainty. *Computers & Chemical Engineering* 18 (1994), Supplement: 163–189.

Ierapetritou, Marianthi G., Efstratios N. Pistikopoulos, and Christodoulos A. Floudas. Operational planning under uncertainty. *Computers & Chemical Engineering* 20 (1996), 12: 1499–1516.

ILOG CPLEX Division. *CPLEX 9 User's Manual*. Incline Village, NV: ILOG CPLEX Division, 2000. http://www.gams.com/dd/docs/solvers/cplex.pdf (accessed February 12, 2006).

Infanger, G. *Planning under Uncertainty: Solving Large Scale Stochastic Linear Programs*. Danvers, Massachusetts: Boyd and Fraser Publishing Co, 1994.

Iyer, R. R. and Ignacio E. Grossmann. A bilevel decomposition algorithm for long-range planning of process networks. *Industrial & Engineering Chemistry Research* 37 (1998): 474–481.

Janjira, Saran, Rathanawan Magaraphan, and Miguel J. Bagajewicz. Simultaneous treatment of environmental and financial risk in process design. *International Journal of Environement and Pollution* (in press).

Jechura, John. *Course Notes for ChEN409: Refining Processes*. Golden, Colorado, USA: Colorado School of Mines. http://jechura.com/ChEN409/, accessed on October 17, 2005

Jia, Zhenya and Marianthi G. Ierapetritou. Mixed-integer linear programming model for gasoline blending and distribution scheduling. *Industrial & Engineering Chemistry Research* 42 (2003): 825–835.

Jia, Zhenya and Marianthi G. Ierapetritou. Short-term scheduling under uncertainty using MILP sensitivity analysis. *Industrial & Engineering Chemistry Research* 43 (2004), 14: 3782–3791.

Jia, Zhenya and Marianthi G. Ierapetritou. Efficient short-term scheduling of refinery operations based on a continuous time formulation. *Computers & Chemical Engineering* 28 (2004), 6–7: 1001–1019.

Jobson, D. and B. Korkie. Putting Markowitz theory to work. *Journal of Portfolio Management* 7 (1981), 4: 70–74.

Johns W. R., G. Marketos, and D. W. T. Rippin. The optimal design of chemical plants to meet time-varying demands in the presence of technical and commercial uncertainty. *Transactions of the Institution of Chemical Engineers* 56 (1978): 249–257.

Jones, D. S. J. *Elements of Petroleum Processing*. Chichester, West Sussex, England: John Wiley & Sons. 1995.

Jonsbraten, T. W. Oil-field optimization under price uncertainty. *Journal of the Operational Research Society* 49 (1998), 8: 811–818.

Julka, Nirupam, Iftekhar Karimi, and Rajagopalan Srinivasan. Agent-based supply chain management—2: a refinery application. *Computers & Chemical Engineering* 26 (2002): 1771–1781.

Jung, June Young, Gary Blau, Joseph F. Pekny, Gintaras V. Reklaitis, and David Eversdyk. A simulation-based optimization approach to supply chain management under demand uncertainty. *Computers & Chemical Engineering* 28 (2004): 2087–2106.

Kall, Peter and Janos Mayer. *Stochastic Linear Programming: Model, Theory, and Computation*. New York: Springer, 2005.

Kall, Peter and S. W. Wallace. *Stochastic Programming*. New York: John Wiley & Sons, 1994.

Karimi, Iftekhar A. and Conor M. McDonald. and. Planning and scheduling of parallel semicontinuous processes. 2. Short term scheduling. *Industrial & Engineering Chemistry Research* 36 (1997): 2701–2714.

Karmarkar, Narendra. A new polynomial-time algorithm for linear programming. *Combinatorica* 4 (1984), 4: 373–395.

Kira, D., M. Kusy, and I. Rakita. A stochastic linear programming approach to hierarchical production planning. *Journal of the Operational Research Society* 48 (1997): 207 – 211.

Konno, Hiroshi and Tomoyuki Koshizuka. Mean-absolute deviation model. *IIE Transaction*. 37 (2005): 893–900.

Konno, Hiroshi and Annista Wijayanayake. Portfolio optimization under D.C. transaction costs and minimal transaction unit constraints. *Journal of Global Optimization* 22 (2002): 137–154.

Konno, Hiroshi and Hiroaki Yamazaki. Mean-absolute deviation portfolio optimization model and its applications to Tokyo Stock Market. *Management Science* 37 (May 1991), 5: 519–531.

Kristoffersen, Trine Krogh. Deviation measures in linear two-stage stochastic programming. *Mathematical Methods of Operations Research* 62 (November 2005), 2: 255 – 274.

Lababidi, Haitham M. S., Mohamed A. Ahmed, Imad M. Alatiqi, and Adel F. Al-Enzi. Optimizing the supply chain of a petrochemical company under uncertain operating and economic conditions. *Industrial & Engineering Chemistry Research* 43 (2004): 63–73.

Lai K, K, Stephen C. H. Leung, and Yue Wu. A two-stage recourse model for production planning with stochastic demand. *Lecture Notes in Computer Science*, v 3483, n IV, *Computational Science and Its Applications – ICCSA 2005: International Conference, Proceedings*. 250260, 2005.

Lindgren, Mats and Hans Bandhold. *Scenario Planning: The Link between Future and Strategy*. New York, USA: Palgrave Macmillan.

Lin, Xiaoxia, Stacy L. Janak, Christodoulos A. Floudas. A new robust optimization approach for scheduling under uncertainty: I. Bounded uncertainty. *Computers & Chemical Engineering* 28 (2004): 1069–1085.

Liou, Jiunn-Shyan, Vemuri Balakotaiah, Dan Luss. Dispersion and diffusion influences on yield in complex reaction networks. *AIChE Journal* 35 (September 1989), 9: 1509–1520.

Liu, Ming Long and Nikolaos Vasili Sahinidis. Optimization in process planning under uncertainty. *Industrial & Engineering Chemistry Research* 35 (1996): 4154–xxxx.

Liu, Ming Long and Nikolaos Vasili Sahinidis. Process planning in a fuzzy environment. *European Journal of OperationalResearch* 100 (1997), 1: 142–169.

Li, Wenkai. *Modeling Oil Refinery for Production Planning, Scheduling, and Economic Analysis*. Hong Kong University of Science and Technology: PhD Thesis, 2004.

Li, Wenkai, Chi-Wai Hui, and An-Xue Li. Integrating CDU, FCC, and product blending models into refinery planning. *Computers & Chemical Engineering* 29 (2005): 2010–2028.

Li, Wenkai, Chi-Wai Hui, Pu Li, and An-Xue Li. Refinery planning under uncertainty. *Industrial & Engineering Chemistry Research* 43 (2004): 6742–6755.

Luenberger, D. G. *Investment Science*. New York: Oxford University Press, 1998.

Luss, Hanan. Operations research and capacity expansion problems: a survey. *Operations Research* 30 (1982), 5: 907–947.

Lustig, Irvin J. Interview with George Dantzig by Irvin J. Lustig on planning under uncertainty. *George Dantzig Memorial Site* <http://www2.informs.org/History/dantzig/in_interview_irv10.htm>, accessed on March 9, 2006 and *E –Optimization* <http://www.e-optimization.com/directory/trailblazers/dantzig/interview_planning.cfm>, accessed on March 2, 2006.

Maiti, S. N., J, Eberhardt, S. Kundu, P. J. Cadenhouse-Beaty, and D. J. Adams. How to efficiently plan a grassroots refinery. *Hydrocarbon Processing*. (2001)

Malcolm, Scott A. and Stavros A. Zenios. Robust optimization of power systems capacity expansion under uncertainty. *Journal of the Operational Research Society* 45 (1994), 9: 1040–1049.

Manandhar, S., A. Tarim, and T. Walsh. Scenario-based stochastic constraint programming. *International Joint Conferences on Artificial Intelligence*. Acapulco, Mexico, August 9–15, paper 214.

Maples, Robert E. *Petroleum Refinery Process Economics*. Tulsa, Oklahoma, USA: PennWell Publishing Company, 1993.

Markowitz, Harry Max. Portfolio selection. *Journal of Finance* 7 (1952), 1: 77–91.

Markowitz, Harry M. *Portfolio Selection: Efficient Diversification of Investments*, Coyles Foundation Monograph. New Haven and London: Yale University Press, 1959. This has more recently been reprinted as: Markowitz, Harry M. *Porfolio Selection*. Oxford, United Kingdom: Blackwell Publishers Inc., 1991.

Marti, Kurt. *Stochastic Optimization Methods*. Germany: Spring Berlin Heidelberg, 2005.

McDonald, Conor M. Synthesizing enterprise-wide optimization with global information technologies: harmony or discord? *AIChE Symposium Series: Proceedings of the Third International Conference of the Foundations of Computer-Aided Process Operations*, Snowbird, Utah, USA, July 5–10, 1998, 94 (1998), 320: 62–74.

McDonald, Conor M. and Iftekhar A. Karimi. Planning and scheduling of parallel semicontinuous processes. 1. Production planning. *Industrial & Engineering Chemistry Research* 36 (1997): 2691–2700.

Michaud, Robert. *Efficient Asset Management*. First published by Harvard Business School Pres, 1998. New York: Oxford University Press, 2001.

Michaud, Richard and Robert Michaud. Issues in estimation error and portfolio optimization. *Presented at the Eighth Annual Financial Econometrics Conference, "Portfolio Selection, Estimation, and Optimization—New Methods" at the Centre for Advanced Studies in Finance and the Institute for Quantitative Finance and Insurance, University of Waterloo*, March 3, 2006.

Mulvey, John M. Introduction to financial optimization: Mathematical Programming Special Issue. *Mathematical Programming Series B* 89 (2001): 205–216.

Mulvey, John M., Daniel. P. Rosenbaum, and Bala Shetty. Strategic financial risk management and operations research. *European Journal of Operations Research* 97 (1997): 1–16.

Mulvey, John M., Robert J. Vanderbei, and Stavros A. Zenios. Robust optimization of large-scale systems. *Operations Research* 43 (1995), 2: 264–281.

Murtagh, Bruce A. Nonlinear programming and risk management. 37th Annual ORSNZ Conference, November 29–30, 2002, University of Auckland, New Zealand. http://www.esc.auckland.ac.n/
Organisations/ORSNZ/conf37/Papers/Murtagh.pdf (accessed November 12, 2005).

Neiro, Sergio M. S. and Jose M. Pinto. Multiperiod optimization for production planning of petroleum refineries. *Chemical Engineering Communications* 192 (2005): 62–88.

Neiro, Sergio M. S. and Jose M. Pinto. A general modeling framework for the operational planning of petroleum supply chains. *Computers & Chemical Engineering*. 28 (2004): 871–896.

Nelissen, Franz. Optimization under uncertainty using GAMS: success stories and some frustrations. *Presented at the Gesellschaft fur Operations Research (GOR) Workshop "Optimization under Uncertainty" at Bad Honnef, Germany on October 20 – 21, 2005*. http://www.gams.com/presentations/
present_uncertainty.pdf (accessed December 17, 2005).

Nemhauser, George L. The age of optimization: solving large-scale real-world problems. *Operations Research* 42 (1994), 1: 5–13.

Nemhauser, George L. and Laurence A. Wolsey. *Integer and Combinatorial Optimization*. New York, New York, USA: John Wiley & Sons, 1988.

Norton, Leon C. and Ignacio E. Grossmann. Strategic planning model for complete process flexibility. *Industrial & Engineering Chemistry Research* 33 (1994): 69–76.

Occupational Safety and Health Administration of the United States of America (USA) Government. *OSHA Technical Manual, Section IV, Chapter 2: Petroleum Refining Processes* (March 13, 2003). http://www.osha.gov/dts/osta/otm/otm_iv/otm_iv_2.html, accessed on September 30, 2005.

Organization of Economic Co-operation and Development. http://www.oecd.org (accessed January 3, 2006).

Pai, C. -C. David and R. R. Hughes. Strategies for formulating and solving two-stage problems for process design under uncertainty. *Computers & Chemical Engineering* 12 (1987), 6: 695–706.

Papahristodoulou, C. and E. Dotzauer. Optimal portfolios using linear programming models. *Journal of the Operational Research Society* 55 (2004): 1169–1177.

Petersen, Eugene E. and Alexis T. Bell. *Catalyst Deactivation*. New York, United States of America: Marcel Dekker, Inc.

Petkov, Spas B. and Costas D. Maranas. Design of single-product campaign batch plants under demand uncertainty. *AIChE Journal* 44 (April 1998), 4: 896–911.

Pflug, Georg Ch. *Optimization of Stochastic Models: The Interface between Simulation and Optimization*. Bostob, Massachusetts: Kluwer Academic Publishers, 1996.

Pintaric, Zorka Novak and Zdravko Kravanja. The two-level strategy for MINLP synthesis of process flowsheets under uncertainty. *Computers & Chemical Engineering* 24 (2000): 195–201.

Pinto, J. M., M. Joly, and L. F. L. Moro. Planning and scheduling models for refinery operations. *Computers & Chemical Engineering*. 24 (2000): 2259–2276.

Pinto, Jose. M. and L. F. L. Moro. A planning model for petroleum refineries. *Brazilian Journal of Chemical Engineering* 14 (2000), 4–7, 575–586.

Pistikopouslos, Efstratios N. and Grossmann, Ignacio E. Optimal retrofit design for improving process flexibility in linear systems. *Computers & Chemical Engineering* 12 (1988), 7: 719–731.

Pistikopouslos, Efstratios N. and Grossmann, Ignacio E. Optimal retrofit design for improving process flexibility in nonlinear systems. I. Fixed degree of flexibility. *Computers & Chemical Engineering* 13 (1989), 9: 1003–1016.

Pistikopouslos, Efstratios N. and Grossmann, Ignacio E. Optimal retrofit design for improving process flexibility in nonlinear systems. II. Optimal level of flexibility. *Computers & Chemical Engineering* 13 (1989), 10: 1087–1096.

Pistikopoulos, Efstratios N. and T. A. Mazzuchi. A novel flexibility analysis approach for processes with stochastic parameters. *Computers & Chemical Engineering* 14 (1990): 991–1000.

Pongsakdi, Arkadej, Pramoch Rangsunvigit, Kitipat Siemanond, and Miguel J. Bagajewicz. Financial risk management in the planning of refinery operations. *International Journal of Production Economics* 103 (2006): 64–86.

Ponnambalam, K. *Lecture Notes for Probabilistic Design Course*. http://epoch.uwaterloo.ca/~ponnu/syde511f05.html (accessed September 22, 2005). Waterloo, Ontario, Canada: University of Waterloo, 2005.

Prekopa, Andras. *Stochastic Programming*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1995.

Rardin, Ronald L. and Benjamin W. Lin. Test problems for computational experiments—issues and techniques. In *Evaluating Mathematical Programming Techniques: Proceedings of a Conference Held at the National Bureau of Standards at Boulder, Colorado on January 5 – 6, 1981*. Edited by John M. Mulvey. Berlin, Heidelberg, New York: Springer–Verlag, 1982.

Rardin, Ronald L. *Notes on GAMS for Optimization*. http://gilbreth.ecn.purdue.edu/~rardin/gams/
notes.html (May 21 1999) (accessed September 30, 2005).

Rardin, Ronald L. *Optimization in Operations Research*. Upper Saddle River, New Jersey, USA: Prentice Hall, 1998.

Ravi, V. and P. J. Reddy. Fuzzy linear fractional goal programming applied to refinery operations planning. *Fuzzy Sets and Systems* 96 (1998): 173–182.

Reklaitis, Gintaras V. Review of scheduling of process operations. *AIChE Symposium Series: Selected Topics on Computer-Aided Process Design and Analysis* 78 (1982), 214: 119–133.

Roberts, S. M. *Dynamic Programming in Chemical Engineering and Process Control*. New York, USA: Academic Press, 1964.

Rockafellar, R. Tyrrell and Stanislav Uryasev. Conditional Value-at-Risk for general loss distributions. *Journal of Banking and Finance* 26 (2002), 7: 1443–1471.

Romero, J., M. Badell, Miguel Bagajewicz, and L. Puigjaner. Integrating budgeting models into scheduling and planning models for the chemical batch industry. *Industrial & Engineering Chemistry Research* 42 (2003), 24: 6125–6134.

Rooney, W. C. and Lorenz T. Biegler. Optimal process design with model parameter uncertainty and process variability. *AIChE Journal* 49 (2003): 438–449.

Ross, Sheldon M. *Introduction to Probability and Statistics for Engineers and Scientists*. Burlington, Massachusetts, United States of America: Elsevier Academic Press, 2004.

Ruszczynski, Andrzej. Decomposition methods in stochastic programming. *Mathematical Programming* 79 (1997), 1–3, 333–353.

Ruszczynski, Andrzej and Alexander Shapiro. *Stochastic Programming, Handbooks in Operations Research and Management Science, Volume 10*. Amsterdam, The Netherlands: Elsevier Science, B.V., 2003.

Rudolf, M., H. J. Wolter, and H. Zimmermann. A linear model for tracking error minimization. *Journal of Banking & Finance* 23 (1999): 85–103.

Sabri, Ehap H. and Benita M. Beamon. A multi-objective approach to simultaneous strategic and operational planning in supply chain design. *Omega: International Journal of Management Science* 28 (2000): 581–598.

Sahdev, M. K., K. K. Jain, and Pankaj Srivastava. Petroleum refinery planning and optimization using linear programming, 2004. www.cheresources.com/refinery_planning_optimization.shtml, accessed February 22, 2006.

Sahinidis, Nikolaos Vasili. Optimization under uncertainty: state-of-the-art and opportunities. *Computers & Chemical Engineering* 28 (2004), 9: 971–983.

Sahinidis, Nikolaos Vasili, Ignacio E. Grossmann, R. E. Fornari, and M. Chathrathi. Optimization model for long range planning in chemical industry. *Computers & Chemical Engineering* 13 (1989), 9: 1049–1063.

Sahinidis, Nikolaos Vasili and Ignacio E. Grossmann. Reformulation of multiperiod MILP models for planning and scheduling of chemical processes. *Computers & Chemical Engineering* 15 (1991), 4: 255–272.

Samsatli, Nouri J., Lazaros G. Papageorgiou, and Nilay Shah. Robustness metrics for dynamic optimization models under parameter uncertainty. *AIChE Journal* 44 (1998), 9: 1993–2006.

Schoemaker, P. and van der Heijden, Kees. Integrating scenarios into strategic planning at Royal Dutch/Shell. *Planning Review* 20 (May–June 1992), 3: 41–46.

Schrage, Linus. *User's Manual for LINDO.* Palo Alto, California: Scientific Press, 1990.

Schwartz, P. *The Art of the Long View*. New York: Doubleday Currency, 1991.

Schwartz, P. Composing a plot for your scenario. *Planning Review* 20 (1992), 3: 4–9.

Sen, Suvrajeet. Stochastic programming: computational issues and challenges. In *Encyclopedia of Operations Research and Management Science*. Edited by Saul I. Gass and Carl M. Harris. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2001, pp. 784–789.

Sen, Suvrajeet and Julia L. Higle. An introductory tutorial on stochastic linear programming models. *Interfaces* 29 (1999): 33–61.

Sermon, P. A., M. S. W. Vong, and M. Matheson. Catalyst coking, activation, and deactivation. In *Deactivation and Testing of Hydrocarbon Processing Catalysts*. Edited by Paul O' Connor, Toru Takatsuka, and Geoffrey L. Woolery. Developed from a symposium sponsored by the Division of Petroleum Chemistry, Inc. at the 210th National Meeting of the American Chemical Society, Chicago, Illinois, August 20–25, 1995. Washington DC, United States of America: American Chemical Society, 1996.

Shah, Nilay. Single- and multisite planning and scheduling: current status and future challenges. *AIChE Symposium Series: Proceedings of the Third International Conference of the Foundations of Computer-Aided Process Operations*, Snowbird, Utah, USA, July 5–10, 1998, 94 (1998), 320: 75–90.

Shah, Nilay, Costas C. Pantelides, and Roger W. H. Sargent. A general algorithm for short-term scheduling of batch operations—II. Computational issues. *Computers & Chemical Engineering* 17 (1993), 2: 229–244.

Shapiro, A. and T. A. Homem-de-Mello. A simulation-based approach to two-stage stochastic programming with recourse. *Mathematical Programming* 81 (1998): 301–325.

Shapiro, Jeremy F. Mathematical programming models and methods for production planning and scheduling. In *Handbooks in Operations Research and Management Science, Volume 4: Logistics of Production and Inventory*. Edited by Stephen C. Graves Alexander H. G. Rinnooy Kan, and P. H. Zipkin. Amsterdam: Elsevier Science Publishers B. V., 1993, pp. 371–443.

Shapiro, Jeremy F. On the connections among activity-based costing, mathematical programming models for analyzing strategic decisions, and the resource-based view of the firm. *European Journal of Operational Research* 118 (1999): 295–314.

Shapiro, Jeremy F. Challenges of strategic supply chain planning and modeling. *Computers & Chemical Engineering* 28 (2004): 855–861.

Shimizu, Y. and T. Takamatsu. Application of mixed-integer linear programming in multiterm expansion planning under multiobjectives. *Computers & Chemical Engineering* 9 (1985), 4: 367–377.

Shobrys, Donald E. and Douglas C. White. Planning, scheduling and control systems: why can they not work together. *Computers & Chemical Engineering* 24 (2000): 163–173.

Shobrys, Donald E. Scheduling. In *Planning, Scheduling, and Control Integration in the Process Industries*. Edited by C. Edward Bodington. New York: McGraw-Hill, 1995, p. 123–156.

Simaan, Yusif. Estimation of risk in portfolio selection: the mean-variance model and the mean-absolute deviation. *Management Science* 43 (1997): 1437–1446.

Speranza, M. Grazia. A heuristic algorithm for a portfolio optimization model applied to the Milan stock market. *Computers & Operations Research* 23 (1996), 5: 433–441.

Subrahmanyam, Sriram, Joseph F, Pekny, and Gintaras V. Reklaitis. Design of batch chemical plants under market uncertainty. *Industrial & Engineering Chemistry Research* 33 (1994): 2688–2701.

Speight, James G. *Petroleum Chemistry and Refining*. Washington, D. C., USA: Taylor & Francis, 1998.

Straub, David Anthony and Ignacio E. Grossmann. Design optimization of stochastic flexibility. *Computers & Chemical Engineering* 17 (1993): 339–354.

Stokke, P. R., W. K. Ralston, T. A. Boyce, and I. H. Wilson. Scenario planning for Norwegian oil and gas. *Long Range Planning* 23 (1990), 2: 17–26.

Tarhan, Bora and Ignacio E. Grossmann. *A Multistage Stochastic Programming Approach with Strategies for Uncertainty Reduction in the Planning of Process Networks with Uncertain Yields*. http://egon.cheme.cmu.edu/Papers/Escape2006Bora.pdf, accessed December 10, 2005.

Terwiesch, P., Mukul Agarwal, and David W. T. Rippin. Batch unit optimization with imperfect modelling: a survey. *Journal of Process Control* 4 (1994), 238–257.

Tucker, Michael A. LP modeling—past, present, and future. www.kbcat.com/pdfs/tech/tp_002.pdf, accessed February 22, 2006.

Uryasev, Stanislav. *Course Notes for ESI 6912: Introduction to Stochastic Optimization*. Industrial and Systems Engineering Department, University of Florida. http://www.ise.ufl.edu/esi6912/FALL2005/
index.htm (accessed on November 11, 2005).

Uryasev, Stanislav and Panos M. Pardalos. *Stochastic Optimization: Algorithms and Applications*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2001.

van den Heever, Susara A., Ignacio E. Grossmann, Sriram Vasantharajan, and Krisanne Edwards. Integrating complex economic objectives with the design and planning of offshore oilfield infrastructures. *Computers & Chemical Engineering* 24 (2000), 1049–1055.

van den Heever, Susara A., Ignacio E. Grossmann, Sriram Vasantharajan, and Krisanne Edwards. A Lagrangean decomposition heuristic for the design and planning of offshore hydrocarbon field infrastructures with complex economic objectives. *Industrial & Engineering Chemistry Research* 40 (2001), 13: 2857–2875.

van den Heever, Susara A.. Multiperiod mixed-integer nonlinear programming (MINLP) models and methods for the optimal design and planning of process and supply networks. *PhD Thesis*.

Vidal, Carlos J. and Marc Goetschalckx. Strategic production–distribution models: a critical review with emphasis on global supply chain models. *European Journal of Operational Research* 98 (1997): 1–18.

Vin, Jeetmanyu P. and Marianthi G. Ierapetritou. Robust short-term scheduling of multiproduct batch plants under demand uncertainty. *Industral & Engineering Chemistry Research* 40 (2001): 4543–4554.

von Neumann, John and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton, New Jersey, USA: Princeton University Press, 1953.

Walkup, David W. and Roger J.-B. Wets. Stochastic programs with recourse. *SIAM Journal of Applied Mathematics* 15 (1967), 6: 1299–1314.

Wellons, H. S. and Gintaras, V. Reklaitis. The design of multiproduct batch plants under uncertainty with staged expansion. *Computers & Chemical Engineering* 13 (1989), 1/2: 115–126.

Wendt, Moritz, Pu Li, and Gunter Wozny. Nonlinear chance-constrained process optimization under uncertainty. *Industral & Engineering Chemistry Research* 41 (2002), 15: 3621–3629.

Westerberg, Arthur W. Optimization. In *Batch Processing Systems Engineering: Fundamentals and Applications for Chemical Engineering*. Edited by Gintaras V. Reklaitis, Aydin K. Sunol, David W. T. Rippin, and Oner Hortacsu, Berlin, Germany: Springer–Verlag, 1996.

Wets, Roger J-B. Stochastic programs with fixed recourse: the equivalent deterministic program. *SIAM Review* 16 (July 1974), 3: 309–339.

Wets, Roger J.-B. Solving stochastic programs with simple recourse. *Stochastics* 10 (1983), 3–4, 219–242.

Wets, Roger J.-B. Stochastic programming. Chapter VIII in *Handbooks in Operations Research & Management Science, Volume I: Optimization*. Edited by G. L. Nemhauser, Alexander H. G. Rinnooy Kan and M. J. Todd. Amsterdam, The Netherlands: Elsevier Science Publsihers B. V., 1989.

Wets, Roger J.-B. Challenges in stochastic programming. *Mathematical Programming* 75 (1996): 115–135.

Wets, Roger J.-B. *Annotated Bibliography for my Banff IRS Lecture, May 2005, on "Equilibrium in a Stochastic Environment: A 'Constructive' Approach*. http://www.math.ucdavis.edu/~rjbw/ARTICLES/ref2Circ.pdf, accessed on April 9, 2006

Winston, Wayne L. and Munirpallam Venkataramanan. *Introduction to Mathematical Programming*. California, USA: Brooks/Cole—Thomson Learning, 2003.

Xueya, Zhang and R. W. H. Sargent. The optimal operation of mixed production facilities—extensions and improvements. *Computers & Chemical Engineering* 22 (1998), 9: 1287–1295.

Yang, Yueh Jiun, William R. Stockwell, and Jana B. Milford. Effect of chemical product yield uncertainties on reactivities of VOCs and emissions from reformulated gasolines and methanol fuels. *Environmental Science & Technology* 30 (April 1996), 4: 1392–1397.

Yin, Karen K., G. George Yin, and Hu Liu. Stochastic modeling for inventory and production planning in the paper industry. *AIChE Journal* 50 (2004), 11: 2877–2890.

Zayed, Tarek M. and Issam Minkarah. Resource allocation for concrete batch plant operation: case study. *Journal of Construction Engineering and Management* 130 (2004), 4: 560–569.

Zenios, Stavros A. and Pan Kang. Mean-absolute deviation portfolio optimization for mortgage-backed securities. *Annals of Operations Research* 45 (1993), 433–450.

Zhang, N. and X. X. Zhu. A novel modelling and decomposition strategy for overall refinery optimisation. *Computers & Chemical Engineering* 24 (2000): 1543–1548.

Zimmermann, H.-J. *Fuzzy set theory and its applications*. 2nd ed. Boston, Massachusetts, United States of America: Kluwer Academic Publishers.

Appendices

APPENDIX A

A Review of the Markowitz's Mean–Variance or Expected Returns–Variance of Returns (*E*−*V*) Rule Approach

The following is a review of the widely-used Markowitz's mean–variance model as found in Konno and Yamazaki (1991). Let R_j be a random variable representing the rate of return (per period) of the asset S_j, $j = 1, \ldots, n$. Also let x_j be the amount of money to be invested in S_j out of the total fund M_0.

The expected return (per period) of this investment is given by

$$r(x_1,\ldots,x_n) = E\left[\sum_{j=1}^{n} R_j x_j\right] = \sum_{j=1}^{n} E\left[R_j\right] x_j \qquad \textbf{(A.1)}$$

where $E[\cdot]$ represents the expected value of the random variable in the bracket. An investor prefers to have $r(x_1,\ldots,x_n)$ as large as possible. At the same time, he wants to make the risk as small as possible.

Markowitz, in his seminal work (1959), employed the standard deviation of the (per period) return:

$$\sigma(x_1,\ldots,x_n) = \sqrt{E\left[\left\{\sum_{j=1}^{n} R_j x_j - E\left[\sum_{j=1}^{n} R_j x_j\right]\right\}\right]} \qquad \textbf{(A.2)}$$

as the measure of risk and formulated the portfolio optimization problem as a parametric quadratic programming problem, as presented in the following:

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}x_i x_j \\
\text{subject to} \quad & \sum_{j=1}^{n} r_j x_j \geq \rho M_0, \\
& \sum_{j=1}^{n} x_j = M_0 \\
& 0 \leq x_j \leq u_j, \quad j = 1,\ldots,n
\end{aligned} \qquad \textbf{(A.3)}$$

where $r_j = E[R_j]$ and $\sigma_{ij} = E\left[\left(R_i - r_i\right)\left(R_j - r_j\right)\right]$ and ρ is a parameter representing the minimal rate of return required by an investor. Also, u_j is the maximum amount of money that can be invested into S_j. This model is known to be valid if (i) R_j's are multivariate normally distributed and/or (ii) an investor is risk averse in the sense that he prefers less standard deviation of the portfolio to more.

APPENDIX B1, Part I

Daily crude oil price data for the period January 5, 2004–December 30, 2005 (Energy Information Administration (EIA), http://www.eia.doe.gov/pub/oil_gas/petroleum/analysis_publications/oil_market_basics/Price_links.htm (accessed December 27, 2005), http://www.eia.doe.gov/neic/historic/hpetroleum.htm (accessed December 28, 2005)) (*Note*: n.a. = not available)

	Price (USD/barrel)			Price (USD/barrel)			Price (USD/barrel)			Price (USD/barrel)			Price (USD/barrel)	
Date	Brent	WTI at Cushing	Date	Brent	WTI at Cushing	Date	Brent	WTI at Cushing	Date	Brent	WTI at Cushing	Date	Brent	WTI at Cushing
01/01/04	n.a.	n.a.	11/02/04	30.37	34.08	23/03/04	n.a.	37.70	03/05/04	n.a.	38.23	11/06/04	35.64	38.45
02/01/04	n.a.	n.a.	12/02/04	30.15	33.93	24/03/04	34.36	37.58	04/05/04	n.a.	38.98	14/06/04	35.65	37.58
05/01/04	31.02	33.78	13/02/04	30.18	33.93	25/03/04	n.a.	35.53	05/05/04	37.09	39.58	15/06/04	34.55	37.18
06/01/04	31.16	33.73	16/02/04	31.04	34.44	26/03/04	32.82	35.73	06/05/04	37.31	39.38	16/06/04	34.50	37.33
07/01/04	n.a.	33.63	17/02/04	30.96	35.18	29/03/04	32.38	35.45	07/05/04	n.a.	39.93	17/06/04	35.14	38.48
08/01/04	31.15	33.98	18/02/04	31.61	35.45	30/03/04	33.17	36.25	10/05/04	37.04	38.93	18/06/04	35.72	38.75
09/01/04	31.92	34.32	19/02/04	31.43	36.03	31/03/04	32.41	35.78	11/05/04	37.34	40.08	21/06/04	34.99	37.63
12/01/04	31.95	34.73	20/02/04	31.57	35.63	01/04/04	32.19	34.28	12/05/04	38.52	40.78	22/06/04	34.62	38.13
13/01/04	32.45	34.45	23/02/04	31.42	35.78	02/04/04	30.84	34.38	13/05/04	38.30	41.08	23/06/04	34.60	37.68
14/01/04	31.55	34.53	24/02/04	n.a.	36.08	05/04/04	31.16	34.38	14/05/04	39.07	41.38	24/06/04	34.85	37.75
15/01/04	31.94	33.65	25/02/04	32.40	37.18	06/04/04	31.43	34.98	17/05/04	38.90	41.53	25/06/04	34.40	37.15
16/01/04	31.23	35.08	26/02/04	32.20	35.53	07/04/04	33.08	36.15	18/05/04	38.15	40.53	28/06/04	33.25	36.34
19/01/04	31.61	35.08	27/02/04	32.53	36.18	08/04/04	33.31	37.13	19/05/04	38.69	41.50	29/06/04	32.73	35.68
20/01/04	32.18	36.23	01/03/04	33.63	36.33	09/04/04	n.a.	n.a.	20/05/04	38.90	40.93	30/06/04	33.04	37.05
21/01/04	31.41	34.98	02/03/04	34.27	36.68	12/04/04	n.a.	37.83	21/05/04	37.40	39.98	01/07/04	33.64	38.73
22/01/04	31.51	34.43	03/03/04	n.a.	35.83	13/04/04	33.71	37.23	24/05/04	38.85	41.88	02/07/04	35.12	38.38
23/01/04	n.a.	35.03	04/03/04	n.a.	36.63	14/04/04	33.00	36.73	25/05/04	38.27	41.13	05/07/04	n.a.	n.a.
26/01/04	n.a.	34.48	05/03/04	n.a.	37.28	15/04/04	33.65	37.05	26/05/04	38.05	40.70	06/07/04	36.53	39.65
27/01/04	30.67	34.43	08/03/04	33.69	37.08	16/04/04	33.89	37.73	27/05/04	36.96	39.40	07/07/04	36.07	39.08
28/01/04	n.a.	33.75	09/03/04	n.a.	36.28	19/04/04	33.63	37.43	28/05/04	37.04	39.90	08/07/04	36.93	40.33
29/01/04	29.48	32.68	10/03/04	n.a.	36.13	20/04/04	33.15	37.60	31/05/04	n.a.	n.a.	09/07/04	37.64	39.98
30/01/04	29.52	33.15	11/03/04	n.a.	36.78	21/04/04	32.75	36.53	01/06/04	39.00	42.35	12/07/04	37.85	39.50
02/02/04	29.42	34.98	12/03/04	n.a.	36.18	22/04/04	32.87	37.43	02/06/04	38.42	39.95	13/07/04	36.53	39.43
03/02/04	30.05	34.13	15/03/04	n.a.	36.93	23/04/04	33.28	37.13	03/06/04	36.65	39.30	14/07/04	37.38	40.98
04/02/04	n.a.	33.13	16/03/04	n.a.	37.48	26/04/04	33.00	36.98	04/06/04	35.86	38.60	15/07/04	38.02	40.78
05/02/04	29.39	33.08	17/03/04	n.a.	38.18	27/04/04	34.33	37.53	07/06/04	35.30	38.65	16/07/04	38.81	41.03
06/02/04	n.a.	32.48	18/03/04	34.44	37.93	28/04/04	34.69	37.48	08/06/04	35.72	37.28	19/07/04	39.08	41.45
09/02/04	29.07	32.83	19/03/04	n.a.	38.08	29/04/04	33.98	37.33	09/06/04	34.74	37.55	20/07/04	38.94	40.86
10/02/04	n.a.	33.88	22/03/04	n.a.	37.13	30/04/04	35.02	37.38	10/06/04	35.80	38.45	21/07/04	38.63	40.83

APPENDIX B1, Part II (*continued*) (*Note*: n.a. = not available)

	Price (USD/barrel)			Price (USD/barrel)			Price (USD/barrel)			Price (USD/barrel)			Price (USD/barrel)	
Date	Brent	WTI at Cushing	Date	Brent	WTI at Cushing	Date	Brent	WTI at Cushing	Date	Brent	WTI at Cushing	Date	Brent	WTI at Cushing
22/07/04	39.33	41.46	01/09/04	41.12	44.01	12/10/04	51.47	52.52	22/11/04	42.67	48.50	31/12/04		
23/07/04	39.48	41.53	02/09/04	42.10	44.07	13/10/04	49.82	53.65	23/11/04	43.05	48.75	03/01/05		42.13
26/07/04	39.32	41.44	03/09/04	41.33	44.00	14/10/04		54.77	24/11/04	42.42	49.27	04/01/05	40.08	43.92
27/07/04	40.08	41.84	06/09/04			15/10/04	51.15	54.94	25/11/04			05/01/05	41.03	43.40
28/07/04	41.16	42.90	07/09/04	40.53	43.35	18/10/04	49.39	53.67	26/11/04	42.99		06/01/05	43.07	45.57
29/07/04	40.88	42.75	08/09/04	40.94	42.78	19/10/04	48.84	53.29	29/11/04	43.57	49.77	07/01/05	43.54	45.44
30/07/04	41.61	43.80	09/09/04	41.76	44.62	20/10/04	49.43	54.92	30/11/04	44.04	49.14	10/01/05	44.63	45.54
02/08/04	41.68	43.83	10/09/04	42.48	42.82	21/10/04	50.73	54.47	01/12/04	41.47	47.50	11/01/05	43.16	45.69
03/08/04	42.51	44.15	13/09/04		43.88	22/10/04	52.09	55.17	02/12/04	38.25	43.26	12/01/05	42.76	46.38
04/08/04	41.94	42.83	14/09/04	41.68	44.40	25/10/04	52.13	55.23	03/12/04		42.55	13/01/05	45.12	48.05
05/08/04	43.50	44.40	15/09/04	42.15	43.59	26/10/04	51.43	55.18	06/12/04	38.61	42.99	14/01/05	45.66	48.39
06/08/04	43.00	43.95	16/09/04	41.16	43.90	27/10/04	51.97	52.47	07/12/04	37.41	41.47	17/01/05		
09/08/04	44.00	44.85	17/09/04	42.68	45.60	28/10/04	49.79	50.93	08/12/04	37.52	41.95	18/01/05	44.86	48.39
10/08/04	43.32	44.55	20/09/04	43.49	46.36	29/10/04	48.02	51.77	09/12/04	38.36	42.54	19/01/05	45.05	47.56
11/08/04	43.30	44.80	21/09/04	44.36	47.10	01/11/04	46.54	50.14	10/12/04	38.07	40.72	20/01/05	44.10	46.92
12/08/04	44.04	45.50	22/09/04	45.93	48.56	02/11/04	45.92	49.63	13/12/04	36.86	41.02	21/01/05	45.75	48.29
13/08/04	44.96	46.58	23/09/04		48.37	03/11/04	45.17	50.89	14/12/04	37.25	41.83	24/01/05	45.65	48.62
16/08/04	45.13	46.05	24/09/04		48.81	04/11/04	45.59	48.83	15/12/04		44.20	25/01/05	46.08	49.45
17/08/04	44.60	46.75	27/09/04	47.00	49.65	05/11/04	44.54	49.62	16/12/04	40.67	44.19	26/01/05	46.23	48.79
18/08/04	44.26	47.28	28/09/04	47.09	49.91	08/11/04	44.27	49.09	17/12/04		46.29	27/01/05	46.75	48.85
19/08/04	44.80	48.70	29/09/04	46.81	49.51	09/11/04	42.46	47.38	20/12/04	43.02	45.65	28/01/05	44.98	47.19
20/08/04	44.77	47.86	30/09/04	47.26	49.65	10/11/04	42.16	48.87	21/12/04	42.70	45.56	31/01/05		48.21
23/08/04	44.04	46.70	01/10/04	47.29	50.13	11/11/04	43.47	47.73	22/12/04	42.81	44.05	01/02/05	45.05	47.13
24/08/04	43.09	45.61	04/10/04	47.00	49.91	12/11/04	41.58	47.33	23/12/04	41.14	43.99	02/02/05	44.09	46.70
25/08/04	42.37	43.55	05/10/04	47.16	51.10	15/11/04		46.88	24/12/04			03/02/05	43.01	46.46
26/08/04	40.73	43.10	06/10/04	47.42	52.03	16/11/04	40.69	46.12	27/12/04		41.33	04/02/05	43.62	46.49
27/08/04	40.63	43.18	07/10/04	48.44	52.68	17/11/04	39.83	46.85	28/12/04		41.78	07/02/05	42.68	45.29
30/08/04		42.28	08/10/04	48.99	53.32	18/11/04	40.03	46.23	29/12/04	38.96	43.65	08/02/05		45.41
31/08/04	39.75	42.12	11/10/04	50.76	53.65	19/11/04	42.45	48.45	30/12/04	39.90	43.46	09/02/05	43.37	45.47

APPENDIX B1, Part III (*continued*) (*Note*: n.a. = not available)

	Price (USD/barrel)			Price (USD/barrel)			Price (USD/barrel)			Price (USD/barrel)			Price (USD/barrel)	
Date	Brent	WTI at Cushing	Date	Brent	WTI at Cushing	Date	Brent	WTI at Cushing	Date	Brent	WTI at Cushing	Date	Brent	WTI at Cushing
10/02/05	43.90	47.11	23/03/05		50.36	03/05/05	49.83	49.51	14/06/05	53.29	55.01	25/07/05	57.44	56.01
11/02/05	44.34	47.17	24/03/05	52.36	51.10	04/05/05	50.07	50.14	15/06/05	53.50	55.57	26/07/05	58.52	59.21
14/02/05	44.00	47.45	25/03/05			05/05/05	48.79	50.84	16/06/05	54.74	56.58	27/07/05	58.34	59.12
15/02/05	44.67	47.27	28/03/05		54.06	06/05/05	50.21	50.97	17/06/05	56.15	58.48	28/07/05	58.44	59.95
16/02/05	44.83	48.34	29/03/05	51.28	54.24	09/05/05	49.45	52.04	20/06/05	57.17	59.37	29/07/05	59.79	60.57
17/02/05	45.99	47.55	30/03/05	50.54	54.00	10/05/05	50.26	52.08	21/06/05	57.32	58.91	01/08/05	61.01	61.58
18/02/05	45.74	48.36	31/03/05	52.90	55.41	11/05/05	48.94	50.46	22/06/05	55.92	57.85	02/08/05	60.39	61.90
21/02/05			01/04/05	53.89	57.28	12/05/05	46.63	48.55	23/06/05	57.10	59.23	03/08/05	60.21	60.87
22/02/05	47.27	51.16	04/04/05	55.54	57.01	13/05/05	46.32	48.67	24/06/05	57.03	59.65	04/08/05	61.03	61.39
23/02/05	47.76	50.63	05/04/05	54.44	56.05	16/05/05	46.35	48.62	27/06/05	58.49	60.54	05/08/05	63.04	62.32
24/02/05	49.01	50.75	06/04/05	54.18	55.86	17/05/05	47.29	48.98	28/06/05	57.68	58.20	08/08/05	62.68	63.95
25/02/05	49.33	50.95	07/04/05	54.08	54.12	18/05/05	48.19	47.26	29/06/05	55.48	57.26	09/08/05	62.57	63.08
28/02/05		51.76	08/04/05	51.76	53.33	19/05/05	48.09	46.93	30/06/05	55.35	56.50	10/08/05	62.74	64.90
01/03/05	49.92	51.69	11/04/05	52.11	53.72	20/05/05	48.22	46.81	01/07/05	57.29	58.76	11/08/05	65.64	65.81
02/03/05	50.63	53.06	12/04/05	52.61	51.87	23/05/05	46.63	48.62	04/07/05			12/08/05	67.09	66.87
03/03/05	53.08	53.58	13/04/05	50.93	50.23	24/05/05	46.98	49.38	05/07/05	57.69	59.60	15/08/05	66.70	66.28
04/03/05	52.23	53.79	14/04/05	49.79	51.14	25/05/05	48.95	50.80	06/07/05	58.50	61.29	16/08/05	66.06	66.09
07/03/05	52.46	53.90	15/04/05		50.50	26/05/05	48.99	51.02	07/07/05	57.43	60.74	17/08/05	63.71	63.26
08/03/05	53.03	54.60	18/04/05	49.48	50.38	27/05/05	49.71	51.85	08/07/05	58.95	59.64	18/08/05	61.71	63.28
09/03/05	53.85	54.77	19/04/05	50.20	52.30	31/05/05	50.52	51.98	11/07/05	56.04	58.93	19/08/05	63.93	65.36
10/03/05		53.55	20/04/05	51.20	52.45	01/06/05	51.82	54.61	12/07/05	58.15	60.63	22/08/05	65.64	65.46
11/03/05	52.63	54.43	21/04/05	51.64	52.21	02/06/05	52.80	53.64	13/07/05	58.22	60.02	23/08/05	64.97	65.37
14/03/05	53.24	54.95	22/04/05	53.28	54.40	03/06/05	53.41	55.04	14/07/05	56.78	57.81	24/08/05	66.59	67.08
15/03/05	53.49	55.05	25/04/05	53.11	53.58	06/06/05	53.71	54.50	15/07/05	57.41	58.10	25/08/05	66.02	67.30
16/03/05	54.90	56.47	26/04/05	51.73	54.21	07/06/05	53.18	53.77	18/07/05	56.93	57.33	26/08/05	66.47	66.14
17/03/05	55.48	56.40	27/04/05	51.05	51.62	08/06/05	53.25	52.55	19/07/05	56.63	57.47	29/08/05		67.21
18/03/05	55.68	56.73	28/04/05	50.24	51.78	09/06/05	52.75	54.29	20/07/05	56.28	56.73	30/08/05	66.04	69.82
21/03/05	55.49	56.62	29/04/05	50.74	49.73	10/06/05	53.25	53.55	21/07/05	55.54	55.49	31/08/05	66.81	68.95
22/03/05	55.70	54.04	02/05/05		50.93	13/06/05	53.07	55.63	22/07/05	56.95	56.66	01/09/05	66.99	69.48

APPENDIX B1, Part IV (*continued*) (*Note*: n.a. = not available)

	Price (USD/barrel)			Price (USD/barrel)			Price (USD/barrel)	
Date	Brent	WTI at Cushing	Date	Brent	WTI at Cushing	Date	Brent	WTI at Cushing
02/09/05	66.51	67.58	13/10/05	60.18	63.09	23/11/05	54.52	58.47
05/09/05			14/10/05	58.24	62.64	24/11/05		
06/09/05	63.96	65.97	17/10/05		64.37	25/11/05	53.04	
07/09/05			18/10/05	58.87	63.21	28/11/05	52.99	57.37
08/09/05	63.79	64.50	19/10/05	58.24	62.41	29/11/05	53.44	56.51
09/09/05	63.25	64.09	20/10/05	56.60	61.04	30/11/05	53.24	57.33
12/09/05	60.75	63.35	21/10/05	56.91	61.19	01/12/05	53.43	58.48
13/09/05	61.70	63.12	24/10/05	57.45	61.13	02/12/05	54.98	59.33
14/09/05	61.25	65.10	25/10/05	58.90	63.05	05/12/05	57.64	59.92
15/09/05	61.25	64.76	26/10/05	60.07	60.67	06/12/05	55.78	59.95
16/09/05	60.15	63.01	27/10/05	58.18	61.10	07/12/05	55.27	59.22
19/09/05	63.31	67.40	28/10/05		61.23	08/12/05	56.28	60.67
20/09/05	62.61	66.24	31/10/05	58.50	59.77	09/12/05	56.73	59.40
21/09/05	64.49	66.76	01/11/05	58.38	59.86	12/12/05	57.42	61.31
22/09/05		66.47	02/11/05	58.71	59.76	13/12/05	58.85	61.38
23/09/05	62.02	64.37	03/11/05	59.44	61.79	14/12/05	59.47	60.86
26/09/05	61.62	65.83	04/11/05	60.29	60.59	15/12/05	59.89	60.00
27/09/05	62.79	65.07	07/11/05	58.31	59.48	16/12/05	58.21	58.07
28/09/05	62.88	66.36	08/11/05	57.60	59.72	19/12/05	56.19	57.35
29/09/05	63.04	66.80	09/11/05	57.46	58.94	20/12/05	56.10	57.99
30/09/05	61.88	66.25	10/11/05	54.90	57.81	21/12/05	56.46	58.37
03/10/05	62.00	65.48	11/11/05	53.78	57.54	22/12/05	57.13	58.24
04/10/05	60.10	63.91	14/11/05	54.20	57.70	23/12/05		
05/10/05	61.07	62.80	15/11/05	54.09	56.99	26/12/05		
06/10/05	58.25	61.37	16/11/05	54.10	57.89	27/12/05		58.16
07/10/05	58.63	61.84	17/11/05	53.77	56.35	28/12/05	57.15	59.82
10/10/05	56.25	61.81	18/11/05	52.66	56.15	29/12/05	57.35	60.32
11/10/05	58.53	63.54	21/11/05	53.13	57.21	30/12/05		
12/10/05	59.28	64.13	22/11/05	53.98	58.45			

APPENDIX B2

Weekly USA retail gasoline price (cents per gallon) for all grades and all formulations for the period of January 5, 2004–December 26, 2005 (Energy Information Administration (EIA), *Retail Gasoline Historical Prices*, http://www.eia.doe.gov/oil_gas/petroleum/data_publications/wrgp/mogas_history.html, accessed on January 23, 2006) (*Note*: n.a. = not available)

Date	Gasoline price (cent/gallon)	Date	Gasoline price (cent/gallon)	Date	Gasoline price (cent/gallon)
05/01/04	155.2	06/09/04	189.3	09/05/05	223.1
12/01/04	160.3	13/09/04	188.9	16/05/05	220.6
19/01/04	163.7	20/09/04	190.8	23/05/05	216.9
26/01/04	166.4	27/09/04	195.9	30/05/05	214.1
02/02/04	166	04/10/04	198	06/06/05	215.9
09/02/04	168.1	11/10/04	203.5	13/06/05	217.3
16/02/04	169	18/10/04	207.7	20/06/05	220.4
23/02/04	173	25/10/04	207.4	27/06/05	225.7
01/03/04	175.8	01/11/04	207.6	04/07/05	226.8
08/03/04	178	08/11/04	204.5	11/07/05	236.9
15/03/04	176.7	15/11/04	201.4	18/07/05	236
22/03/04	178.5	22/11/04	199.2	25/07/05	233.3
29/03/04	180	29/11/04	198.9	01/08/05	233.5
05/04/04	182.2	06/12/04	195.6	08/08/05	241
12/04/04	182.7	13/12/04	189.3	15/08/05	259.2
19/04/04	185.3	20/12/04	186.1	22/08/05	265.4
26/04/04	185.3	27/12/04	183.8	29/08/05	265.3
03/05/04	188.4	03/01/05	182.4	05/09/05	311.7
10/05/04	197.9	10/01/05	183.7	12/09/05	300.2
17/05/04	205.5	17/01/05	186.3	19/09/05	283.5
24/05/04	210.4	24/01/05	189.6	26/09/05	285.1
31/05/04	209.2	31/01/05	195.3	03/10/05	297.5
07/06/04	207.5	07/02/05	195.2	10/10/05	289.6
14/06/04	202.9	14/02/05	194.1	17/10/05	277.5
21/06/04	198.1	21/02/05	194.8	24/10/05	265.2
28/06/04	196.5	28/02/05	196.9	31/10/05	252.8
05/07/04	193.9	07/03/05	204	07/11/05	242.4
12/07/04	195.9	14/03/05	209.8	14/11/05	234.2
19/07/04	197.1	21/03/05	214.9	21/11/05	224.7
26/07/04	194.8	28/03/05	219.4	28/11/05	220
02/08/04	193	04/04/05	225.8	05/12/05	219.1
09/08/04	192	11/04/05	232.1	12/12/05	222.8
16/08/04	191.7	18/04/05	228	19/12/05	225.5
23/08/04	192.6	25/04/05	227.9	26/12/05	224.1
30/08/04	190.9	02/05/05	227.7		

APPENDIX B3, Part I

Daily USA Gulf Coast kerosene-type jet fuel spot price FOB (free-on-board) for the period of January 5, 2004–December 23, 2005 (Energy Information Administration (EIA), *Historical Petroleum Price Data—Other Product Prices*, http://www.eia.doe.gov/neic/historic/hpetroleum2.htm#Other, accessed on January 23, 2006) (*Note*: n.a. = not available)

Date	Jet fuel price (cent/gallon)	Date	Jet fuel price (cent/gallon)	Date	Jet fuel price (cent/gallon)	Date	Jet fuel price (cent/gallon)	Date	Jet fuel price (cent/gallon)	Date	Jet fuel price (cent/gallon)
05/01/04	96.1	16/02/04		29/03/04	91.79	11/05/04	114.6	23/06/04	104.28	04/08/04	118.7
06/01/04	98.2	17/02/04	98.15	30/03/04	93.86	12/05/04	114.9	24/06/04	106.61	05/08/04	122.18
07/01/04	96.95	18/02/04	99	31/03/04	93.63	13/05/04	116.88	25/06/04	106.28	06/08/04	120.68
08/01/04	99.12	19/02/04	94.7	01/04/04	88.13	14/05/04	118.5	28/06/04	104.3	09/08/04	122.38
09/01/04	101.83	20/02/04	90.5	02/04/04	88.38	17/05/04	115.55	29/06/04	102.75	10/08/04	119.78
12/01/04	102.66	23/02/04	90.85	05/04/04	88.15	18/05/04	111.21	30/06/04	106.3	11/08/04	118.88
13/01/04	102.66	24/02/04	91.4	06/04/04	90.8	19/05/04	110.63	01/07/04	112	12/08/04	122.37
14/01/04	98.18	25/02/04	96.31	07/04/04	93.6	20/05/04	108.02	02/07/04	112.7	13/08/04	125.43
15/01/04	94.23	26/02/04	99.75	08/04/04	96.25	21/05/04	105.43	05/07/04		16/08/04	123.85
16/01/04	98.83	27/02/04	98.25	12/04/04	99.5	24/05/04	109.39	06/07/04	113.89	17/08/04	125.23
19/01/04		01/03/04	100.45	13/04/04	97.88	25/05/04	105.98	07/07/04	112.43	18/08/04	125.9
20/01/04	103.49	02/03/04	98.85	14/04/04	99.08	26/05/04	102.03	08/07/04	114.21	19/08/04	130.9
21/01/04	106.25	03/03/04	94.35	15/04/04	102.51	27/05/04	98.94	09/07/04	112.1	20/08/04	127.1
22/01/04	106.63	04/03/04	97	16/04/04	101.65	28/05/04	101.3	12/07/04	108.8	23/08/04	124.95
23/01/04	103.15	05/03/04	96	19/04/04	99.28	01/06/04	107.05	13/07/04	109.75	24/08/04	125.21
26/01/04	103.15	08/03/04	85.6	20/04/04	97.88	02/06/04	101.85	14/07/04	112.6	25/08/04	120.9
27/01/04	99.48	09/03/04	92.6	21/04/04	97.53	03/06/04	102.25	15/07/04	112.78	26/08/04	120
28/01/04	97.35	10/03/04	91.34	22/04/04	99.4	04/06/04	100.75	16/07/04	113.2	27/08/04	121.98
29/01/04	95	11/03/04	92.1	23/04/04	98.44	07/06/04	102.38	19/07/04	114.18	30/08/04	119.96
30/01/04	93.45	12/03/04	91.38	26/04/04	100.8	08/06/04	99.15	20/07/04	111.48	31/08/04	119.53
02/02/04	93.65	15/03/04	94.55	27/04/04	99.35	09/06/04	100.71	21/07/04	112.55	01/09/04	125.18
03/02/04	95.03	16/03/04	94.32	28/04/04	99.93	10/06/04	102.53	22/07/04	116.99	02/09/04	126.48
04/02/04	90.8	17/03/04	98.5	29/04/04	101.93	11/06/04		23/07/04	118.7	03/09/04	125.68
05/02/04	89.93	18/03/04	97.65	30/04/04	102.35	14/06/04	100.7	26/07/04	117.35	06/09/04	
06/02/04	87.93	19/03/04	97.2	03/05/04	104.98	15/06/04	101.35	27/07/04	117.93	07/09/04	125.48
09/02/04	88.58	22/03/04	96.2	04/05/04	105.8	16/06/04	101.23	28/07/04	119.39	08/09/04	125.25
10/02/04	90.25	23/03/04	97.48	05/05/04	108.83	17/06/04	105.57	29/07/04	119.03	09/09/04	130.1
11/02/04	91.63	24/03/04	94.97	06/05/04	108.05	18/06/04	104.7	30/07/04	121.78	10/09/04	123.25
12/02/04	91.63	25/03/04	93.61	07/05/04	112.05	21/06/04	101.79	02/08/04	120.8	13/09/04	128.23
13/02/04	95.25	26/03/04	93.83	10/05/04	110.55	22/06/04	104.4	03/08/04	122.53	14/09/04	133.4

APPENDIX B3, Part II (*continued*) (*Note*: n.a. = not available)

Date	Jet fuel price (cent/gallon)	Date	Jet fuel price (cent/gallon)	Date	Jet fuel price (cent/gallon)	Date	Jet fuel price (cent/gallon)	Date	Jet fuel price (cent/gallon)	Date	Jet fuel price (cent/gallon)
15/09/04	134.25	29/10/04	141.95	14/12/04	126.34	27/01/05	139.38	14/03/05	154.8	28/04/05	153.8
16/09/04	135.26	01/11/04	135.72	15/12/04	136	28/01/05	132.95	15/03/05	156.2	29/04/05	146.9
17/09/04	138.55	02/11/04	133.75	16/12/04	131.68	31/01/05	134.5	16/03/05	161.92	02/05/05	151.5
20/09/04	137.05	03/11/04	137.9	17/12/04	134.63	01/02/05	130.5	17/03/05	159.84	03/05/05	150.75
21/09/04	139.46	04/11/04	131.91	20/12/04	129.45	02/02/05	128.33	18/03/05	161.68	04/05/05	153.16
22/09/04	143.11	05/11/04	132.95	21/12/04	130.19	03/02/05	124.63	21/03/05	162.58	05/05/05	151.75
23/09/04	144.3	08/11/04	132.05	22/12/04	127.16	04/02/05	125.4	22/03/05	159.38	06/05/05	150.38
24/09/04	145.55	09/11/04	130.93	23/12/04	128.29	07/02/05	120.53	23/03/05	160.75	09/05/05	150.13
27/09/04	149.39	10/11/04	138.46	24/12/04		08/02/05	123.17	24/03/05	159.25	10/05/05	153.1
28/09/04	150.4	11/11/04	131.71	27/12/04	114.91	09/02/05	126.88	28/03/05	158.43	11/05/05	148.44
29/09/04	151	12/11/04	132.15	28/12/04	116.1	10/02/05	130.75	29/03/05	155.05	12/05/05	144.96
30/09/04	149.05	15/11/04	129.7	29/12/04	122.75	11/02/05	130.25	30/03/05	160.3	13/05/05	143.4
01/10/04	151.3	16/11/04	128.05	30/12/04	120.95	14/02/05	129.05	31/03/05	166.9	16/05/05	141.05
04/10/04	149.24	17/11/04	135.58	31/12/04		15/02/05	129.38	01/04/05	173.85	17/05/05	142.65
05/10/04	148.98	18/11/04	137.24	03/01/05	114.99	16/02/05	136.73	04/04/05	169.63	18/05/05	141.2
06/10/04	148.83	19/11/04	142.68	04/01/05	119.75	17/02/05	133.38	05/04/05	167.88	19/05/05	143.8
07/10/04	147.28	22/11/04	138	05/01/05	121.1	18/02/05	136	06/04/05	164.1	20/05/05	141.4
08/10/04	149.58	23/11/04	136.35	06/01/05	129	21/02/05		07/04/05	158.33	23/05/05	140.78
11/10/04	150.7	24/11/04	139.35	07/01/05	129.18	22/02/05	143.63	08/04/05	156.18	24/05/05	143.35
12/10/04	148.39	25/11/04		10/01/05	129.88	23/02/05	146.68	11/04/05	154.91	25/05/05	147.1
13/10/04	155.83	26/11/04		11/01/05	131.73	24/02/05	145.73	12/04/05	151.66	26/05/05	149.33
14/10/04	160.3	29/11/04	138.45	12/01/05	132.7	25/02/05	146.25	13/04/05	151.96	27/05/05	149.98
15/10/04	160.55	30/11/04	131.95	13/01/05	135.9	28/02/05	147.73	14/04/05	155.18	30/05/05	
18/10/04	154.99	01/12/04	122.63	14/01/05	136.95	01/03/05	148.79	15/04/05	153.6	31/05/05	151.69
19/10/04	152.08	02/12/04	119.23	17/01/05		02/03/05	151.85	18/04/05	150.51	01/06/05	160.2
20/10/04	156.63	03/12/04	114.85	18/01/05	136.5	03/03/05	148.7	19/04/05	153.75	02/06/05	161.4
21/10/04	158.53	06/12/04	113.4	19/01/05	135.95	04/03/05	148.4	20/04/05	155.73	03/06/05	167.6
22/10/04	159.18	07/12/04	111.29	20/01/05	136.38	07/03/05	148.83	21/04/05	159.2	06/06/05	165.66
25/10/04	155.13	08/12/04	114.79	21/01/05	141.2	08/03/05	153.24	22/04/05	160.75	07/06/05	164.44
26/10/04	156.03	09/12/04	120.18	24/01/05	142.9	09/03/05	151.93	25/04/05	157.18	08/06/05	159.25
27/10/04	146.28	10/12/04	113.35	25/01/05	144.3	10/03/05	151.88	26/04/05	156.2	09/06/05	166.46
28/10/04	140.9	13/12/04	120.3	26/01/05	142.9	11/03/05	155.88	27/04/05	151.5	10/06/05	165.15

APPENDIX B3, Part III (*continued*) (*Note*: n.a. = not available)

Date	Jet fuel price (cent/gallon)	Date	Jet fuel price (cent/gallon)	Date	Jet fuel price (cent/gallon)	Date	Jet fuel price (cent/gallon)	Date	Jet fuel price (cent/gallon)
13/06/05	170.65	27/07/05	163.65	09/09/05	194.88	25/10/05	195.57	08/12/05	178.1
14/06/05	166.59	28/07/05	165.8	12/09/05	184.5	26/10/05	192.74	09/12/05	172.18
15/06/05	164.83	29/07/05	166.45	13/09/05	188.26	27/10/05	193.75	12/12/05	174.15
16/06/05	165.89	01/08/05	170.43	14/09/05	200.44	28/10/05	196.4	13/12/05	180.55
17/06/05	169.1	02/08/05	170.05	15/09/05	196.5	31/10/05	189.38	14/12/05	180.7
20/06/05	169.35	03/08/05	166.4	16/09/05	190.5	01/11/05	182.4	15/12/05	176
21/06/05	165.4	04/08/05	170	19/09/05	216.5	02/11/05	178	16/12/05	171.98
22/06/05	164.28	05/08/05	172.3	20/09/05	215.29	03/11/05	182.75	19/12/05	169.38
23/06/05	169.39	08/08/05	179.55	21/09/05	219.79	04/11/05	176.55	20/12/05	165.7
24/06/05	167.55	09/08/05	177.55	22/09/05	225.3	07/11/05	175.95	21/12/05	169.5
27/06/05	168.63	10/08/05	185.6	23/09/05	217.5	08/11/05	174.55	22/12/05	171.85
28/06/05	162.78	11/08/05	192.83	26/09/05	239.48	09/11/05	177.88	23/12/05	170.5
29/06/05	161.63	12/08/05	194.5	27/09/05	252.2	10/11/05	171.5		
30/06/05	161.93	15/08/05	190.59	28/09/05	295	11/11/05	169		
01/07/05	170.6	16/08/05	189.6	29/09/05	302.5	14/11/05	169.53		
04/07/05		17/08/05	179.55	30/09/05	264.5	15/11/05	163.09		
05/07/05	172.57	18/08/05	182.88	03/10/05	290	16/11/05	167.65		
06/07/05	179.49	19/08/05	186.75	04/10/05	309.97	17/11/05	164.45		
07/07/05	174.88	22/08/05	184.2	05/10/05	313.03	18/11/05	164.7		
08/07/05	170.05	23/08/05	187.2	06/10/05	274.5	21/11/05	164.95		
11/07/05	166.3	24/08/05	194.21	07/10/05	284	22/11/05	168.42		
12/07/05	171.9	25/08/05	192.1	10/10/05	278.2	23/11/05	169.25		
13/07/05	167.83	26/08/05	189.25	11/10/05	277.7	24/11/05			
14/07/05	163.35	29/08/05	197.23	12/10/05	280	25/11/05			
15/07/05	164.5	30/08/05	229.5	13/10/05	237.85	28/11/05	158.53		
18/07/05	160.75	31/08/05	228.84	14/10/05	230.5	29/11/05	156.4		
19/07/05	162.37	01/09/05	242.25	17/10/05	256.8	30/11/05	160.65		
20/07/05	159.8	02/09/05	220.5	18/10/05	252.72	01/12/05	165.2		
21/07/05	157.15	05/09/05		19/10/05	213.75	02/12/05	168		
22/07/05	163.5	06/09/05	215.25	20/10/05	192.15	05/12/05	172.25		
25/07/05	164.08	07/09/05	208.85	21/10/05	193.25	06/12/05	174.65		
26/07/05	164.35	08/09/05	197.78	24/10/05	183.77	07/12/05	170.5		

APPENDIX B4

Weekly USA No. 2 heating oil residential price (cents per gallon excluding taxes) for the period of January 5, 2004–December 26, 2005 (Energy Information Administration (EIA), *Heating Oil and Propane Update* at http://tonto.eia.doe.gov/oog/info/hopu/hopu.asp, accessed on January 23, 2006) [*Note*: n.a. = not available. Also, there is no data available for the following periods: (i) between March 16, 2004 and October 3, 2004 and (ii) between March 15, 2005 and October 2, 2005.]

Date	Heating oil price (cent/gallon)	Date	Heating oil price (cent/gallon)
05/01/04	149.797	03/01/05	195.116
12/01/04	156.176	10/01/05	194.56
19/01/04	158.444	17/01/05	196.386
26/01/04	162.162	24/01/05	198.994
02/02/04	162.521	31/01/05	201.789
09/02/04	161.531	07/02/05	199.034
16/02/04	161.116	14/02/05	198.119
23/02/04	160.947	21/02/05	198.377
01/03/04	160.256	28/02/05	204.27
08/03/04	160.13	07/03/05	208.85
15/03/04	159.122	14/03/05	211.85
04/10/04	182.787	03/10/05	269.159
11/10/04	190.849	10/10/05	264.83
18/10/04	199.156	17/10/05	265.007
25/10/04	206.028	24/10/05	262.298
01/11/04	205.961	31/10/05	257.725
08/11/04	202.824	07/11/05	250.824
15/11/04	201.673	14/11/05	246.556
22/11/04	202.522	21/11/05	243.124
29/11/04	202.964	28/11/05	241.72
06/12/04	197.013	05/12/05	241.035
13/12/04	194.709	12/12/05	241.403
20/12/04	199.344	19/12/05	243.803
27/12/04	197.757	26/12/05	243.3

APPENDIX B5

Monthly USA residual fuel oil retail sales by all sellers (cents per gallon) for the period of January 5, 2004–November 30, 2005 (Energy Information Administration (EIA), *Residual Fuel Oil Prices by Sales Type*, http://tonto.eia.doe.gov/dnav/pet/pet_pri_resid_dcu_nus_m.htm, accessed on January 24, 2006) (*Note*: n.a. = not available)

Date	Fuel oil price (cent/gallon)
Jan-2004	70.6
Feb-2004	69.1
Mar-2004	65.8
Apr-2004	67.6
May-2004	72.6
Jun-2004	73.4
Jul-2004	70.2
Aug-2004	72
Sep-2004	74.1
Oct-2004	81.3
Nov-2004	80.3
Dec-2004	74.4
Jan-2005	77.3
Feb-2005	81.4
Mar-2005	88.1
Apr-2005	96.5
May-2005	99.6
Jun-2005	99.5
Jul-2005	103.2
Aug-2005	109.6
Sep-2005	122.9
Oct-2005	126.7
Nov-2005	120.5

APPENDIX C

The Mean-Absolute Deviation (MAD) Model for Portfolio Optimization (Konno and Yamazaki,1991; Konno and Wijayanayake, 2002)

Let R_j be the rate of return of jth asset ($j = 1, \ldots, n$) and let $\boldsymbol{x} = (x_1, \ldots, x_n)$ be a portfolio, a vector of proportion of investments into each asset. Let X be an investable set, i.e., a set of feasible portfolios. For simplicity, it is assumed that X is a set defined below:

$$X = \{\boldsymbol{x} = (x_1, \cdots, x_n) \left| \sum_{j=1}^{n} x_j = 1,\ 0 \le x_j \le \alpha_j,\ j = 1, \cdots, n \right.\} \qquad \textbf{(C.1)}$$

The rate of return $R(\boldsymbol{x})$ of the portfolio $\boldsymbol{x}$ is given by

$$R(\boldsymbol{x}) = \sum_{j=1}^{n} R_j x_j \qquad \textbf{(C.2)}$$

Let r_j be the expected value of the rate of return R_j of the jth asset. The absolute deviation $W(\boldsymbol{x})$ of the rate of return $R(\boldsymbol{x})$ of the portfolio $\boldsymbol{x}$ is given by

$$W(\boldsymbol{x}) = E[|R(\boldsymbol{x}) - E[R(\boldsymbol{x})]|] \qquad \textbf{(C.3)}$$

It is assumed that $\mathbf{R} = (R_1, \ldots, R_n)$ is distributed over a finite set of points $\{(r_{1t}, \cdots, r_{nt}),$ $t = 1, \cdots, T\}$ and that the probability of occurrence of $(r_{1t}, \cdots, x_{nt})$ is given by p_t, $t = 1, \cdots, T$. Then:

$$r_j = \sum_{t=1}^{T} p_t r_{jt} \qquad \textbf{(C.4)}$$

and

$$W(\boldsymbol{x}) = \sum_{t=1}^{T} p_t \left| \sum_{j=1}^{n} \left(r_{jt} - r_j \right) x_j \right. \quad \textbf{(C.5)}$$

The mean-absolute deviation (MAD) portfolio optimization model is defined as follows:

$$\begin{aligned} &\text{minimize} \quad W(x) \equiv \sum_{t=1}^{T} p_t \left| \sum_{j=1}^{n} \left(r_{jt} - r_j \right) x_j \right| \\ &\text{subject to} \quad \sum_{j=1}^{n} r_j x_j = \rho, \\ &\qquad\qquad\quad x \in X, \end{aligned} \quad \textbf{(C.6)}$$

where ρ is a given constant representing the expected rate of return of the portfolio. The MAD model can be formulated in an alternative way:

$$\begin{aligned} &\text{maximize} \quad \sum_{j=1}^{n} r_j x_j \\ &\text{subject to} \quad \sum_{t=1}^{T} p_t \left| \sum_{j=1}^{n} \left(r_{jt} - r_j \right) x_j \right| \le w \\ &\qquad\qquad\quad x \in X \end{aligned} \quad \textbf{(C.7)}$$

where w is a given constant representing the tolerable level of risk. Both **(C.6)** and **(C.7)** can be used interchangeably to generate an efficient frontier.

By standard results in linear programming, the problem **(C.7)** can be converted to a linear system of inequalities as follows:

$$
\begin{aligned}
&\text{maximize} \quad \sum_{j=1}^{n} r_j x_j \\
&\text{subject to} \quad \sum_{t=1}^{T} y_t \le \frac{w}{2} \\
&\qquad y_t \ge p_t \sum_{j=1}^{n} \left(r_{jt} - r_j \right) x_j, \quad y_t \ge 0, \quad t = 1, \cdots, T, \\
&\qquad \sum_{j=1}^{n} x_j = 1, \quad x_j \in U_j, \qquad\qquad j = 1, \cdots, n.
\end{aligned}
\tag{C.8}
$$

APPENDIX D

GAMS Program Codes for the Numerical Example

Appendix D1: The Deterministic Midterm Refinery Production Planning Model

```
$TITLE Deterministic Model


Variables

Z;


Positive Variables

x1, x2, x3, x4, x5, x6, x7, x8, x9, x10, x11, x12, x13, x14, x15, x16, x17, x18, x19, x20;


Equations

OBJ, CON1, CON2, EQN3, EQN4, EQN5, EQN6, EQN7, EQN8, EQN9, EQN10, EQN11, EQN12, EQN13,
EQN14, EQN15, EQN16, EQN17, EQN18, EQN19, CON20, CON21, CON22, CON23, CON24;
OBJ.. Z =E= -8.0*x1 + 18.5*x2 + 8.0*x3 + 12.5*x4 + 14.5*x5 + 6.0*x6 - 1.5*x14;
CON1.. x1 =L= 15000;
CON2.. x14 =L= 2500;
EQN3.. -0.13*x1 + x7 =E= 0;
EQN4.. -0.15*x1 + x4 =E= 0;
EQN5.. -0.22*x1 + x8 =E= 0;
EQN6.. -0.20*x1 + x9 =E= 0;
EQN7.. -0.30*x1 + x10 =E= 0;
EQN8.. -0.05*x14 + x20 =E= 0;
EQN9.. -0.40*x14 + x16 =E= 0;
EQN10.. -0.55*x14 + x17 =E= 0;
EQN11.. 0.5*x2 - x11 =E= 0;
EQN12.. 0.5*x2 - x16 =E= 0;
EQN13.. 0.75*x5 - x12 =E= 0;
EQN14.. 0.25*x5 - x18 =E= 0;
EQN15.. -x7 + x3 +x11 =E= 0;
EQN16.. -x8 + x12 +x13 =E= 0;
EQN17.. -x9 + x14 +x15 =E= 0;
EQN18.. -x17 + x18 +x19 =E= 0;
EQN19.. -x10 - x13 - x15 - x19 + x6 =E= 0;
CON20.. x2 =L= 2700;
CON21.. x3 =L= 1100;
CON22.. x4 =L= 2300;
CON23.. x5 =L= 1700;
CON24.. x6 =L= 9500;

Model Refinery / all /;

Solve Refinery using LP maximizing Z;
```

Appendix D2: Approach 1—Risk Model I Based on the Markowitz's Mean–Variance (*E–V* or MV) Approach to Handle Randomness in the Objective Function Coefficients of Prices

```
$TITLE Approach 1: Risk Model I Based on the Markowitz's Mean-Variance Approach to Handle
Randomness in the Objective Function Coefficients of Prices


Sets

i product type /1*21/
s scenarios /1*3/


Table pc(i,s) price of product type i per realization s
         1         2        3
1        8.8       8.0      7.2
2        20.35     18.5     16.65
3        8.8       8.0      7.2
4        13.75     12.5     11.25
5        15.95     14.5     13.05
6        6.6       6        5.4
14       1.65      1.5      1.35;


Variables

Z1;


Positive Variables

x1, x2, x3, x4, x5, x6, x7, x8, x9, x10, x11, x12, x13, x14, x15, x16, x17, x18, x19, x20,
Ep;


Scalars

p1 "probability for scenario 1" /0.35/
p2 "probability for scenario 2" /0.45/
p3 "probability for scenario 3" /0.2/

v1 variance of price of crude oil /0.352/
v2 variance of price of gasoline /1.882375/
v3 variance of price of naphtha /0.352/
v4 variance of price of jet fuel /0.859375/
v5 variance of price of heating oil /1.156375/
v6 variance of price of crude oil /0.198/
v14 variance of price of cracker feed /0.012375/


Equations

OBJ, CON0, CON1, CON2, EQN3, EQN4, EQN5, EQN6, EQN7, EQN8, EQN9, EQN10, EQN11, EQN12,
EQN13, EQN14, EQN15, EQN16, EQN17, EQN18, EQN19, CON20, CON21, CON22, CON23, CON24,
Eprofit;

OBJ..
Z1 =E= -((v1*SQR(x1)) + (v2*SQR(x2)) + (v3*SQR(x3)) + (v4*SQR(x4)) + (v5*SQR(x5)) +
(v6*SQR(x6)) + (v14*SQR(x14)));

CON0..
[p1*(-pc('1','1')*x1 + pc('2','1')*x2 + pc('3','1')*x3 + pc('4','1')*x4 + pc('5','1')*x5 +
pc('6','1')*x6 - pc('14','1')*x14)]
+ [p2*(-pc('1','2')*x1 + pc('2','2')*x2 + pc('3','2')*x3 + pc('4','2')*x4 + pc('5','2')*x5
+ pc('6','2')*x6 - pc('14','2')*x14)]
```

```
+ [p3*(-pc('1','3')*x1 + pc('2','3')*x2 + pc('3','3')*x3 + pc('4','3')*x4 + pc('5','3')*x5
+ pc('6','3')*x6 - pc('14','3')*x14)] =G= 23500;

CON1.. x1 =L= 15000;
CON2.. x14 =L= 2500;
EQN3.. -0.13*x1 + x7 =E= 0;
EQN4.. -0.15*x1 + x4 =E= 0;
EQN5.. -0.22*x1 + x8 =E= 0;
EQN6.. -0.20*x1 + x9 =E= 0;
EQN7.. -0.30*x1 + x10 =E= 0;
EQN8.. -0.05*x14 + x20 =E= 0;
EQN9.. -0.40*x14 + x16 =E= 0;
EQN10.. -0.55*x14 + x17 =E= 0;
EQN11.. 0.5*x2 - x11 =E= 0;
EQN12.. 0.5*x2 - x16 =E= 0;
EQN13.. 0.75*x5 - x12 =E= 0;
EQN14.. 0.25*x5 - x18 =E= 0;
EQN15.. -x7 + x3 +x11 =E= 0;
EQN16.. -x8 + x12 +x13 =E= 0;
EQN17.. -x9 + x14 +x15 =E= 0;
EQN18.. -x17 + x18 +x19 =E= 0;
EQN19.. -x10 - x13 - x15 - x19 + x6 =E= 0;
CON20.. x2 =L= 2700;
CON21.. x3 =L= 1100;
CON22.. x4 =L= 2300;
CON23.. x5 =L= 1700;
CON24.. x6 =L= 9500;

Eprofit.. Ep =E= [p1*(-pc('1','1')*x1 + pc('2','1')*x2 + pc('3','1')*x3 + pc('4','1')*x4 +
pc('5','1')*x5 + pc('6','1')*x6 - pc('14','1')*x14)]
+ [p2*(-pc('1','2')*x1 + pc('2','2')*x2 + pc('3','2')*x3 + pc('4','2')*x4 + pc('5','2')*x5
+ pc('6','2')*x6 - pc('14','2')*x14)]
+ [p3*(-pc('1','3')*x1 + pc('2','3')*x2 + pc('3','3')*x3 + pc('4','3')*x4 + pc('5','3')*x5
+ pc('6','3')*x6 - pc('14','3')*x14)];


Model Refinery / all /;

Solve Refinery Using NLP Maximizing Z1;
```

Appendix D3: Approach 2.1—Expectation Model I as a Combination of the Markowitz's Mean–Variance Approach and the Two-Stage Stochastic Programming with Fixed Recourse Framework

```
$TITLE Approach 2: Expectation Model I as a Combination of the Markowitz's Mean-Variance
Approach and the Two-Stage Stochastic Programming with Fixed Recourse Framework


Sets
i product type /1*21/
s scenarios /1*3/
k production shortfall and surplus or yield decrement or increment /1, 2/


Table pc(i,s) price of product type i per realization j
         1         2        3
1        8.8       8.0      7.2
2        20.35     18.5     16.65
3        8.8       8.0      7.2
4        13.75     12.5     11.25
5        15.95     14.5     13.05
6        6.6       6        5.4
14       1.65      1.5      1.35;


Table d(i,s) demand of product type i per realization j
      1        2        3
2     2835     2700     2565
3     1155     1100     1045
4     2415     2300     2185
5     1785     1700     1615
6     9975     9500     9025;


Table y(i,s) yield of product type i per realization j
      1           2         3
3     -0.1365     -0.13     -0.1235
4     -0.1575     -0.15     -0.1425
8     -0.231      -0.22     -0.209
9     -0.21       -0.20     -0.19
10    -0.265      -0.30     -0.335;


Table c(i,k) penalty cost for product type i due to production shortfall or surplus
      1     2
1     55    50
2     25    20
3     17    13
4     5     4
5     6     5
6     10    8;


Table q(i,k) penalty cost for product type i due to yield decrement or increment
      1     2
3     5     3
4     5     4
8     5     3
9     5     3
10    5     3


Positive Variables

x1, x2, x3, x4, x5, x6, x7, x8, x9, x10, x11, x12, x13, x14, x15, x16, x17, x18, x19, x20,
z211, z212, z221, z222, z231, z232,
```

```
z311, z312, z321, z322, z331, z332,
z411, z412, z421, z422, z431, z432,
z511, z512, z521, z522, z531, z532,
z611, z612, z621, z622, z631, z632,
y311, y312, y321, y322, y331, y332,
y411, y412, y421, y422, y431, y432,
y811, y812, y821, y822, y831, y832,
y911, y912, y921, y922, y931, y932,
y1011, y1012, y1021, y1022, y1031, y1032
;


Variables

Es1, Es2, Es3,
Z2,
Vp, Tshortfall, Tsurplus, Es, Vpsq, Ep, Ecv
;


Scalars

v1 "variance of price of crude oil" /0.352/
v2 "variance of price of gasoline" /1.882375/
v3 "variance of price of naphtha" /0.352/
v4 "variance of price of jet fuel" /0.859375/
v5 "variance of price of heating oil" /1.156375/
v6 "variance of price of crude oil" /0.198/
v14 "variance of price of cracker feed" /0.012375/

p1 "probability for scenario 1" /0.35/
p2 "probability for scenario 2" /0.45/
p3 "probability for scenario 3" /0.2/


Equations

OBJ "maximize profit",
CON1, CON2, EQN8, EQN9, EQN10, EQN11, EQN12, EQN13, EQN14, EQN15, EQN16, EQN17, EQN18,
EQN19,
Escenario, Escenario1, Escenario2, Escenario3,
CONgas1, CONgas2, CONgas3,
CONnap1, CONnap2, CONnap3,
CONjf1, CONjf2, CONjf3,
CONho1, CONho2, CONho3,
CONfo1, CONfo2, CONfo3,
CONlhsnap1, CONlhsnap11, CONlhsnap12,
CONlhsnap2, CONlhsnap21, CONlhsnap22,
CONlhsnap3, CONlhsnap31, CONlhsnap32,
CONlhsjf1, CONlhsjf11, CONlhsjf12,
CONlhsjf2, CONlhsjf21, CONlhsjf22,
CONlhsjf3, CONlhsjf31, CONlhsjf32,
CONlhsgo1, CONlhsgo11, CONlhsgo12,
CONlhsgo2, CONlhsgo21, CONlhsgo22,
CONlhsgo3, CONlhsgo31, CONlhsgo32,
CONlhscf1, CONlhscf11, CONlhscf12,
CONlhscf2, CONlhscf21, CONlhscf22,
CONlhscf3, CONlhscf31, CONlhscf32,
CONlhsr1, CONlhsr11, CONlhsr12,
CONlhsr2, CONlhsr21, CONlhsr22,
CONlhsr3, CONlhsr31, CONlhsr32,
Eprofit, Vprofit, Totalshortfall, Totalsurplus, Vpsqrt, Ecvar;


OBJ.. Z2 =E=
[p1*(-pc('1','1')*x1 + pc('2','1')*x2 + pc('3','1')*x3 + pc('4','1')*x4 + pc('5','1')*x5 +
pc('6','1')*x6 - pc('14','1')*x14)]
+ [p2*(-pc('1','2')*x1 + pc('2','2')*x2 + pc('3','2')*x3 + pc('4','2')*x4 + pc('5','2')*x5
+ pc('6','2')*x6 - pc('14','2')*x14)]
+ [p3*(-pc('1','3')*x1 + pc('2','3')*x2 + pc('3','3')*x3 + pc('4','3')*x4 + pc('5','3')*x5
+ pc('6','3')*x6 - pc('14','3')*x14)]
```

```
- 0.0000000001*Vp
- Es;

Vprofit.. Vp =E= (v1*SQR(x1)) + (v2*SQR(x2)) + (v3*SQR(x3)) + (v4*SQR(x4)) + (v5*SQR(x5))
+ (v6*SQR(x6)) + (v14*SQR(x14));

CON1.. x1 =L= 15000;
CON2.. x14 =L= 2500;
*EQN3.. -0.13*x1 + x7 =E= 0;
*EQN4.. -0.15*x1 + x4 =E= 0;
*EQN5.. -0.22*x1 + x8 =E= 0;
*EQN6.. -0.20*x1 + x9 =E= 0;
*EQN7.. -0.30*x1 + x10 =E= 0;
EQN8.. -0.05*x14 + x20 =E= 0;
EQN9.. -0.40*x14 + x16 =E= 0;
EQN10.. -0.55*x14 + x17 =E= 0;
EQN11.. 0.5*x2 - x11 =E= 0;
EQN12.. 0.5*x2 - x16 =E= 0;
EQN13.. 0.75*x5 - x12 =E= 0;
EQN14.. 0.25*x5 - x18 =E= 0;
EQN15.. -x7 + x3 +x11 =E= 0;
EQN16.. -x8 + x12 +x13 =E= 0;
EQN17.. -x9 + x14 +x15 =E= 0;
EQN18.. -x17 + x18 +x19 =E= 0;
EQN19.. -x10 - x13 - x15 - x19 + x6 =E= 0;
*CON20.. x2 =L= 2700;
*CON21.. x3 =L= 1100;
*CON22.. x4 =L= 2300;
*CON23.. x5 =L= 1700;
*CON24.. x6 =L= 9500;

Escenario1.. Es1 =E= p1*(((c('2','1')*z211 + c('2','2')*z212) + (c('3','1')*z311 +
c('3','2')*z312) + (c('4','1')*z411 + c('4','2')*z412) + (c('5','1')*z511 +
c('5','2')*z512) + (c('6','1')*z611 + c('6','2')*z612))
+ ((q('3','1')*y311 + q('3','2')*y312) + (q('4','1')*y411 + q('4','2')*y412) +
(q('8','1')*y811 + q('8','2')*y812) + (q('9','1')*y911 + q('9','2')*y912) +
(q('10','1')*y1011 + q('10','2')*y1012)));

CONgas1.. x2 + z211 - z212 =E= d('2','1');
CONnap1.. x3 + z311 - z312 =E= d('3','1');
CONjf1.. x4 + z411 - z412 =E= d('4','1');
CONho1.. x5 + z511 - z512 =E= d('5','1');
CONfo1.. x6 + z611 - z612 =E= d('6','1');
CONlhsnap1.. y('3','1')*x1 + x7 + y311 - y312 =E= 0;
         CONlhsnap11.. y311 =L= 0.1*x1;
         CONlhsnap12.. y312 =L= 0.1*x1;
CONlhsjf1.. y('4','1')*x1 + x4 + y411 - y412 =E= 0;
         CONlhsjf11.. y411 =L= 0.1*x1;
         CONlhsjf12.. y412 =L= 0.1*x1;
CONlhsgo1.. y('8','1')*x1 + x8 + y811 - y812 =E= 0;
         CONlhsgo11.. y811 =L= 0.1*x1;
         CONlhsgo12.. y812 =L= 0.1*x1;
CONlhscf1.. y('9','1')*x1 + x9 + y911 - y912 =E= 0;
         CONlhscf11.. y911 =L= 0.1*x1;
         CONlhscf12.. y912 =L= 0.1*x1;
CONlhsr1.. y('10','1')*x1 + x10 + y1011 - y1012 =E= 0;
         CONlhsr11.. y1011 =L= 0.1*x1;
         CONlhsr12.. y1012 =L= 0.1*x1;

Escenario2.. Es2 =E= p2*(((c('2','1')*z221 + c('2','2')*z222) + (c('3','1')*z321 +
c('3','2')*z322) + (c('4','1')*z421 + c('4','2')*z422) + (c('5','1')*z521 +
c('5','2')*z522) + (c('6','1')*z621 + c('6','2')*z622))
+ ((q('3','1')*y321 + q('3','2')*y322) + (q('4','1')*y421 + q('4','2')*y422) +
(q('8','1')*y821 + q('8','2')*y822) + (q('9','1')*y921 + q('9','2')*y922) +
(q('10','1')*y1021 + q('10','2')*y1022)));
CONgas2.. x2 + z221 - z222 =E= d('2','2');
CONnap2.. x3 + z321 - z322 =E= d('3','2');
CONjf2.. x4 + z421 - z422 =E= d('4','2');
CONho2.. x5 + z521 - z522 =E= d('5','2');
CONfo2.. x6 + z621 - z622 =E= d('6','2');
CONlhsnap2.. y('3','2')*x1 + x7 + y321 - y322 =E= 0;
```

```
         CONlhsnap21..y321 =L= 0.1*x1;
         CONlhsnap22.. y322 =L= 0.1*x1;
CONlhsjf2.. y('4','2')*x1 + x4 + y421 - y422 =E= 0;
         CONlhsjf21.. y421 =L= 0.1*x1;
         CONlhsjf22.. y422 =L= 0.1*x1;
CONlhsgo2.. y('8','2')*x1 + x8 + y821 - y822 =E= 0;
         CONlhsgo21.. y821 =L= 0.1*x1;
         CONlhsgo22.. y822 =L= 0.1*x1;
CONlhscf2.. y('9','2')*x1 + x9 + y921 - y922 =E= 0;
         CONlhscf21.. y921 =L= 0.1*x1;
         CONlhscf22.. y922 =L= 0.1*x1;
CONlhsr2.. y('10','2')*x1 + x10 + y1021 - y1022 =E= 0;
         CONlhsr21.. y1011 =L= 0.1*x1;
         CONlhsr22.. y1012 =L= 0.1*x1;

Escenario3.. Es3 =E= p3*(((c('2','1')*z231 + c('2','2')*z232) + (c('3','1')*z331 +
c('3','2')*z332) + (c('4','1')*z431 + c('4','2')*z432) + (c('5','1')*z531 +
c('5','2')*z532) + (c('6','1')*z631 + c('6','2')*z632))
+ ((q('3','1')*y331 + q('3','2')*y332) + (q('4','1')*y431 + q('4','2')*y432) +
(q('8','1')*y831 + q('8','2')*y832) + (q('9','1')*y931 + q('9','2')*y932) +
(q('10','1')*y1031 + q('10','2')*y1032)));
CONgas3.. x2 + z231 - z232 =E= d('2','3');
CONnap3.. x3 + z331 - z332 =E= d('3','3');
CONjf3.. x4 + z431 - z432 =E= d('4','3');
CONho3.. x5 + z531 - z532 =E= d('5','3');
CONfo3.. x6 + z631 - z632 =E= d('6','3');
CONlhsnap3.. y('3','3')*x1 + x7 + y331 - y332 =E= 0;
         CONlhsnap31..y331 =L= 0.1*x1;
         CONlhsnap32.. y332 =L= 0.1*x1;
CONlhsjf3.. y('4','3')*x1 + x4 + y431 - y432 =E= 0;
         CONlhsjf31.. y431 =L= 0.1*x1;
         CONlhsjf32.. y432 =L= 0.1*x1;
CONlhsgo3.. y('8','3')*x1 + x8 + y831 - y832 =E= 0;
         CONlhsgo31.. y831 =L= 0.1*x1;
         CONlhsgo32.. y832 =L= 0.1*x1;
CONlhscf3.. y('9','3')*x1 + x9 + y931 - y932 =E= 0;
         CONlhscf31.. y931 =L= 0.1*x1;
         CONlhscf32.. y932 =L= 0.1*x1;
CONlhsr3.. y('10','3')*x1 + x10 + y1031 - y1032 =E= 0;
         CONlhsr31.. y1031 =L= 0.1*x1;
         CONlhsr32.. y1032 =L= 0.1*x1;

Escenario.. Es =E= Es1 + Es2 + Es3;

Eprofit.. Ep =E=
[p1*(-pc('1','1')*x1 + pc('2','1')*x2 + pc('3','1')*x3 + pc('4','1')*x4 + pc('5','1')*x5 +
pc('6','1')*x6 - pc('14','1')*x14)]
+ [p2*(-pc('1','2')*x1 + pc('2','2')*x2 + pc('3','2')*x3 + pc('4','2')*x4 + pc('5','2')*x5
+ pc('6','2')*x6 - pc('14','2')*x14)]
+ [p3*(-pc('1','3')*x1 + pc('2','3')*x2 + pc('3','3')*x3 + pc('4','3')*x4 + pc('5','3')*x5
+ pc('6','3')*x6 - pc('14','3')*x14)];


Totalshortfall.. Tshortfall =E= z211 + z221 + z231 + z311 + z321 + z331 + z411 + z421 +
z431 + z511 + z521 + z531 + z611 + z621 + z631
+ y311 + y321 + y331 + y411 + y421 + y431 + y811 + y821 + y831 + y911 + y921 + y931 +
y1011 + y1021 + y1031;

Totalsurplus.. Tsurplus =E= z212 + z222 + z232 + z312 + z322 + z332 + z412 + z422 + z432 +
z512 + z522 + z532 + z612 + z622 + z632
+ y312 + y322 + y332 + y412 + y422 + y432 + y812 + y822 + y832 + y912 + y922 + y932 +
y1012 + y1022 + y1032;


*Vprofit2.. Vp2 =E= (v1*SQR(x1)) + (v2*SQR(x2)) + (v3*SQR(x3)) + (v4*SQR(x4)) +
(v5*SQR(x5)) + (v6*SQR(x6)) + (v14*SQR(x14));

Vpsqrt.. Vpsq =E= SQRT((v1*SQR(x1)) + (v2*SQR(x2)) + (v3*SQR(x3)) + (v4*SQR(x4)) +
(v5*SQR(x5)) + (v6*SQR(x6)) + (v14*SQR(x14)));
Ecvar.. Ecv =E= Ep - Es;
```

```
*Covar.. Cv =E= Vpsq/(Ep - Es);


Model REFINERY / ALL /;

* Starting values
x1.l = 12500;
x2.l = 2000;
x3.l = 625;
x4.l = 1875;
x5.l = 1700;
x6.l = 6175;
x7.l = 1625;
x8.l = 2750;
x9.l = 2500;
x10.l = 3750;
x11.l = 1000;
x12.l = 1275;
x13.l = 1475;
x14.l = 2500;
x15.l = 0;
x16.l = 1000;
x17.l = 1375;
x18.l = 425;
x19.l = 950;
x20.l = 125;

Option NLP = conopt3;

Solve Refinery Using NLP Maximizing Z2;
```

Appendix D4: Approach 2.2—Expectation Model II as a Combination of the Markowitz's Mean–Variance Approach and the Two-Stage Stochastic Programming with Fixed Recourse Framework

```
$TITLE Approach 3: Expectation Model II as a Combination of the Markowitz's Mean-Variance
Approach and the Two-Stage Stochastic Programming with Fixed Recourse Framework


Sets

i product type /1*21/
s scenarios /1*3/
k production shortfall and surplus or yield decrement or increment /1, 2/


Table pc(i,s) price of product type i per realization s
         1         2        3
1        8.8       8.0      7.2
2        20.35     18.5     16.65
3        8.8       8.0      7.2
4        13.75     12.5     11.25
5        15.95     14.5     13.05
6        6.6       6        5.4
14       1.65      1.5      1.35;


Table d(i,s) demand of product type i per realization s
      1        2        3
2     2835     2700     2565
3     1155     1100     1045
4     2415     2300     2185
5     1785     1700     1615
6     9975     9500     9025;


Table y(i,s) yield of product type i per realization s
      1           2        3
3     -0.1365     -0.13     -0.1235
4     -0.1575     -0.15     -0.1425
8     -0.231      -0.22     -0.209
9     -0.21       -0.20     -0.19
10    -0.265      -0.30     -0.335;


Table c(i,k) penalty cost for product type i due to production shortfall or surplus
      1     2
1     55    50
2     25    20
3     17    13
4     5     4
5     6     5
6     10    8;


Table q(i,k) penalty cost for product type i due to yield decrement or increment
      1     2
3     5     3
4     5     4
8     5     3
9     5     3
10    5     3


Variables

Z2,
```

```
x2, x3, x4, x5, x6, x7, x8, x9, x10, x11, x12, x13, x14, x15, x16, x17, x18, x19, x20,
z211, z212, z221, z222, z231, z232,
z311, z312, z321, z322, z331, z332,
z411, z412, z421, z422, z431, z432,
z511, z512, z521, z522, z531, z532,
z611, z612, z621, z622, z631, z632,
y311, y312, y321, y322, y331, y332,
y411, y412, y421, y422, y431, y432,
y811, y812, y821, y822, y831, y832,
y911, y912, y921, y922, y931, y932,
y1011, y1012, y1021, y1022, y1031, y1032,
Es1, Es11, Es2, Es21, Es3, Es31,
Vpsq, x1, Tshortfall, Tsurplus, Es, Ep, Ecv;


Positive Variables

x2, x3, x4, x5, x6, x7, x8, x9, x10, x11, x12, x13, x14, x15, x16, x17, x18, x19, x20,
z211, z212, z221, z222, z231, z232,
z311, z312, z321, z322, z331, z332,
z411, z412, z421, z422, z431, z432,
z511, z512, z521, z522, z531, z532,
z611, z612, z621, z622, z631, z632,
y311, y312, y321, y322, y331, y332,
y411, y412, y421, y422, y431, y432,
y811, y812, y821, y822, y831, y832,
y911, y912, y921, y922, y931, y932,
y1011, y1012, y1021, y1022, y1031, y1032,
Es1, Es11, Es2, Es21, Es3, Es31,
Vpsq, x1, Tshortfall, Tsurplus, Es, Ep;


Scalars

v1 variance of price of crude oil /0.352/
v2 variance of price of gasoline /1.882375/
v3 variance of price of naphtha" /0.352/
v4 variance of price of jet fuel" /0.859375/
v5 variance of price of heating oil" /1.156375/
v6 variance of price of crude oil" /0.198/
v14 variance of price of cracker feed" /0.012375/

p1 probability for scenario 1 /0.35/
p2 probability for scenario 2 /0.45/
p3 probability for scenario 3 /0.2/


Equations

OBJ "maximize profit",
CON0, CON1, CON2, EQN8, EQN9, EQN10, EQN11, EQN12, EQN13, EQN14, EQN15, EQN16, EQN17,
EQN18, EQN19,
Escenario, Escenario1, Escenario11,
Escenario2, Escenario21,
Escenario3, Escenario31,
CONgas1, CONgas2, CONgas3,
CONnap1, CONnap2, CONnap3,
CONjf1, CONjf2, CONjf3,
CONho1, CONho2, CONho3,
CONfo1, CONfo2, CONfo3,
CONlhsnap1, CONlhsnap11, CONlhsnap12,
CONlhsnap2, CONlhsnap21, CONlhsnap22,
CONlhsnap3, CONlhsnap31, CONlhsnap32,
CONlhsjf1, CONlhsjf11, CONlhsjf12,
CONlhsjf2, CONlhsjf21, CONlhsjf22,
CONlhsjf3, CONlhsjf31, CONlhsjf32,
CONlhsgo1, CONlhsgo11, CONlhsgo12,
CONlhsgo2, CONlhsgo21, CONlhsgo22,
CONlhsgo3, CONlhsgo31, CONlhsgo32,
CONlhscf1, CONlhscf11, CONlhscf12,
CONlhscf2, CONlhscf21, CONlhscf22,
```

```
CONlhscf3, CONlhscf31, CONlhscf32,
CONlhsr1, CONlhsr11, CONlhsr12,
CONlhsr2, CONlhsr21, CONlhsr22,
CONlhsr3, CONlhsr31, CONlhsr32,
Totalshortfall, Totalsurplus, Eprofit, Vpsqrt, Ecvar;
*Ecvar;

OBJ..
Z2 =E= -((v1*SQR(x1)) + (v2*SQR(x2)) + (v3*SQR(x3)) + (v4*SQR(x4)) + (v5*SQR(x5)) +
(v6*SQR(x6)) + (v14*SQR(x14))) - Es;

Vpsqrt.. Vpsq =e= SQRT((v1*SQR(x1)) + (v2*SQR(x2)) + (v3*SQR(x3)) + (v4*SQR(x4)) +
(v5*SQR(x5)) + (v6*SQR(x6)) + (v14*SQR(x14)));

CON0..
[p1*(-pc('1','1')*x1 + pc('2','1')*x2 + pc('3','1')*x3 + pc('4','1')*x4 + pc('5','1')*x5 +
pc('6','1')*x6 - pc('14','1')*x14)]
+ [p2*(-pc('1','2')*x1 + pc('2','2')*x2 + pc('3','2')*x3 + pc('4','2')*x4 + pc('5','2')*x5
+ pc('6','2')*x6 - pc('14','2')*x14)]
+ [p3*(-pc('1','3')*x1 + pc('2','3')*x2 + pc('3','3')*x3 + pc('4','3')*x4 + pc('5','3')*x5
+ pc('6','3')*x6 - pc('14','3')*x14)] =G= 23387.50;

*Ecvar.. Ecv =E= Ep - Es;

CON1.. x1 =L= 15000;
CON2.. x14 =L= 2500;
*EQN3.. -0.13*x1 + x7 =E= 0;
*EQN4.. -0.15*x1 + x4 =E= 0;
*EQN5.. -0.22*x1 + x8 =E= 0;
*EQN6.. -0.20*x1 + x9 =E= 0;
*EQN7.. -0.30*x1 + x10 =E= 0;
EQN8.. -0.05*x14 + x20 =E= 0;
EQN9.. -0.40*x14 + x16 =E= 0;
EQN10.. -0.55*x14 + x17 =E= 0;
EQN11.. 0.5*x2 - x11 =E= 0;
EQN12.. 0.5*x2 - x16 =E= 0;
EQN13.. 0.75*x5 - x12 =E= 0;
EQN14.. 0.25*x5 - x18 =E= 0;
EQN15.. -x7 + x3 +x11 =E= 0;
EQN16.. -x8 + x12 +x13 =E= 0;
EQN17.. -x9 + x14 +x15 =E= 0;
EQN18.. -x17 + x18 +x19 =E= 0;
EQN19.. -x10 - x13 - x15 - x19 + x6 =E= 0;
*CON20.. x2 =L= 2700;
*CON21.. x3 =L= 1100;
*CON22.. x4 =L= 2300;
*CON23.. x5 =L= 1700;
*CON24.. x6 =L= 9500;

Escenario1.. Es1 =E= p1*Es11;
Escenario11.. Es11 =E= (((c('2','1')*z211 + c('2','2')*z212) + (c('3','1')*z311 +
c('3','2')*z312) + (c('4','1')*z411 + c('4','2')*z412) + (c('5','1')*z511 +
c('5','2')*z512) + (c('6','1')*z611 + c('6','2')*z612))
+ ((q('3','1')*y311 + q('3','2')*y312) + (q('4','1')*y411 + q('4','2')*y412) +
(q('8','1')*y811 + q('8','2')*y812) + (q('9','1')*y911 + q('9','2')*y912) +
(q('10','1')*y1011 + q('10','2')*y1012)));

CONgas1.. x2 + z211 - z212 =E= d('2','1');
CONnap1.. x3 + z311 - z312 =E= d('3','1');
CONjf1.. x4 + z411 - z412 =E= d('4','1');
CONho1.. x5 + z511 - z512 =E= d('5','1');
CONfo1.. x6 + z611 - z612 =E= d('6','1');
CONlhsnap1.. y('3','1')*x1 + x7 + y311 - y312 =E= 0;
         CONlhsnap11.. y311 =L= 0.1*x1;
         CONlhsnap12.. y312 =L= 0.1*x1;
CONlhsjf1.. y('4','1')*x1 + x4 + y411 - y412 =E= 0;
         CONlhsjf11.. y411 =L= 0.1*x1;
         CONlhsjf12.. y412 =L= 0.1*x1;
CONlhsgo1.. y('8','1')*x1 + x8 + y811 - y812 =E= 0;
         CONlhsgo11.. y811 =L= 0.1*x1;
         CONlhsgo12.. y812 =L= 0.1*x1;
```

```
CONlhscf1.. y('9','1')*x1 + x9 + y911 - y912 =E= 0;
         CONlhscf11.. y911 =L= 0.1*x1;
         CONlhscf12.. y912 =L= 0.1*x1;
CONlhsr1.. y('10','1')*x1 + x10 + y1011 - y1012 =E= 0;
         CONlhsr11.. y1011 =L= 0.1*x1;
         CONlhsr12.. y1012 =L= 0.1*x1;

Escenario2.. Es2 =E= p2*Es21;
Escenario21.. Es21 =E= (((c('2','1')*z221 + c('2','2')*z222) + (c('3','1')*z321 +
c('3','2')*z322) + (c('4','1')*z421 + c('4','2')*z422) + (c('5','1')*z521 +
c('5','2')*z522) + (c('6','1')*z621 + c('6','2')*z622))
+ ((q('3','1')*y321 + q('3','2')*y322) + (q('4','1')*y421 + q('4','2')*y422) +
(q('8','1')*y821 + q('8','2')*y822) + (q('9','1')*y921 + q('9','2')*y922) +
(q('10','1')*y1021 + q('10','2')*y1022)));
CONgas2.. x2 + z221 - z222 =E= d('2','2');
CONnap2.. x3 + z321 - z322 =E= d('3','2');
CONjf2.. x4 + z421 - z422 =E= d('4','2');
CONho2.. x5 + z521 - z522 =E= d('5','2');
CONfo2.. x6 + z621 - z622 =E= d('6','2');
CONlhsnap2.. y('3','2')*x1 + x7 + y321 - y322 =E= 0;
         CONlhsnap21..y321 =L= 0.1*x1;
         CONlhsnap22.. y322 =L= 0.1*x1;
CONlhsjf2.. y('4','2')*x1 + x4 + y421 - y422 =E= 0;
         CONlhsjf21.. y421 =L= 0.1*x1;
         CONlhsjf22.. y422 =L= 0.1*x1;
CONlhsgo2.. y('8','2')*x1 + x8 + y821 - y822 =E= 0;
         CONlhsgo21.. y821 =L= 0.1*x1;
         CONlhsgo22.. y822 =L= 0.1*x1;
CONlhscf2.. y('9','2')*x1 + x9 + y921 - y922 =E= 0;
         CONlhscf21.. y921 =L= 0.1*x1;
         CONlhscf22.. y922 =L= 0.1*x1;
CONlhsr2.. y('10','2')*x1 + x10 + y1021 - y1022 =E= 0;
         CONlhsr21.. y1011 =L= 0.1*x1;
         CONlhsr22.. y1012 =L= 0.1*x1;

Escenario3.. Es3 =E= p3*Es31;
Escenario31.. Es31 =E= (((c('2','1')*z231 + c('2','2')*z232) + (c('3','1')*z331 +
c('3','2')*z332) + (c('4','1')*z431 + c('4','2')*z432) + (c('5','1')*z531 +
c('5','2')*z532) + (c('6','1')*z631 + c('6','2')*z632))
+ ((q('3','1')*y331 + q('3','2')*y332) + (q('4','1')*y431 + q('4','2')*y432) +
(q('8','1')*y831 + q('8','2')*y832) + (q('9','1')*y931 + q('9','2')*y932) +
(q('10','1')*y1031 + q('10','2')*y1032)));
CONgas3.. x2 + z231 - z232 =E= d('2','3');
CONnap3.. x3 + z331 - z332 =E= d('3','3');
CONjf3.. x4 + z431 - z432 =E= d('4','3');
CONho3.. x5 + z531 - z532 =E= d('5','3');
CONfo3.. x6 + z631 - z632 =E= d('6','3');
CONlhsnap3.. y('3','3')*x1 + x7 + y331 - y332 =E= 0;
         CONlhsnap31..y331 =L= 0.1*x1;
         CONlhsnap32.. y332 =L= 0.1*x1;
CONlhsjf3.. y('4','3')*x1 + x4 + y431 - y432 =E= 0;
         CONlhsjf31.. y431 =L= 0.1*x1;
         CONlhsjf32.. y432 =L= 0.1*x1;
CONlhsgo3.. y('8','3')*x1 + x8 + y831 - y832 =E= 0;
         CONlhsgo31.. y831 =L= 0.1*x1;
         CONlhsgo32.. y832 =L= 0.1*x1;
CONlhscf3.. y('9','3')*x1 + x9 + y931 - y932 =E= 0;
         CONlhscf31.. y931 =L= 0.1*x1;
         CONlhscf32.. y932 =L= 0.1*x1;
CONlhsr3.. y('10','3')*x1 + x10 + y1031 - y1032 =E= 0;
         CONlhsr31.. y1031 =L= 0.1*x1;
         CONlhsr32.. y1032 =L= 0.1*x1;

Escenario.. Es =E= Es1 + Es2 + Es3;

Eprofit.. Ep =E=
[p1*(-pc('1','1')*x1 + pc('2','1')*x2 + pc('3','1')*x3 + pc('4','1')*x4 + pc('5','1')*x5 +
pc('6','1')*x6 - pc('14','1')*x14)]
+ [p2*(-pc('1','2')*x1 + pc('2','2')*x2 + pc('3','2')*x3 + pc('4','2')*x4 + pc('5','2')*x5
+ pc('6','2')*x6 - pc('14','2')*x14)]
+ [p3*(-pc('1','3')*x1 + pc('2','3')*x2 + pc('3','3')*x3 + pc('4','3')*x4 + pc('5','3')*x5
```

```
+ pc('6','3')*x6 - pc('14','3')*x14)];

Totalshortfall.. Tshortfall =E= z211 + z221 + z231 + z311 + z321 + z331 + z411 + z421 +
z431 + z511 + z521 + z531 + z611 + z621 + z631
+ y311 + y321 + y331 + y411 + y421 + y431 + y811 + y821 + y831 + y911 + y921 + y931 +
y1011 + y1021 + y1031;

Totalsurplus.. Tsurplus =E= z212 + z222 + z232 + z312 + z322 + z332 + z412 + z422 + z432 +
z512 + z522 + z532 + z612 + z622 + z632
+ y312 + y322 + y332 + y412 + y422 + y432 + y812 + y822 + y832 + y912 + y922 + y932 +
y1012 + y1022 + y1032;

Ecvar.. Ecv =E= Ep - Es;

*display Eprofit;
*display Escenario;
*display x1.l;


Model Refinery / all /;

Option NLP = conopt3;

Solve Refinery Using NLP Maximizing Z2;
```

Appendix D5: Approach 3—Risk Model II with Variance as the Measure of Risk of the Recourse Penalty Costs

```
$TITLE Approach 3, Risk Model II for Two-Stage Stochastic Programming with Fixed Recourse
of Minimization of the Expected Value and the Variance of the Recourse Penalty Costs


Sets

i "product type" /1*21/
s "scenarios" /1*3/
k "production shortfall and surplus or yield decrement or increment" /1, 2/


Table pc(i,s) "price of product type i per realization s"
         1         2        3
1        8.8       8.0      7.2
2        20.35     18.5     16.65
3        8.8       8.0      7.2
4        13.75     12.5     11.25
5        15.95     14.5     13.05
6        6.6       6        5.4
14       1.65      1.5      1.35;


Table d(i,s) "demand of product type i per realization s"
      1        2        3
2     2835     2700     2565
3     1155     1100     1045
4     2415     2300     2185
5     1785     1700     1615
6     9975     9500     9025;


Table y(i,s) "yield of product type i per realization s"
      1           2          3
3     -0.1365     -0.13      -0.1235
4     -0.1575     -0.15      -0.1425
8     -0.231      -0.22      -0.209
9     -0.21       -0.20      -0.19
10    -0.265      -0.30      -0.335;


Table c(i,k) "penalty cost for product type i due to production requirement shortfall or
surplus compared against market demand"
      1     2
1     55    50
2     25    20
3     17    13
4     5     4
5     6     5
6     10    8;

Table q(i,k) "penalty cost for product type i due to yield decrement or increment"
      1     2
3     5     3
4     5     4
8     5     3
9     5     3
10    5     3


Variables

Z2, Ecv;

Positive Variables
```

```
x1, x2, x3, x4, x5, x6, x7, x8, x9, x10, x11, x12, x13, x14, x15, x16, x17, x18, x19, x20,
z211, z212, z221, z222, z231, z232,
z311, z312, z321, z322, z331, z332,
z411, z412, z421, z422, z431, z432,
z511, z512, z521, z522, z531, z532,
z611, z612, z621, z622, z631, z632,
y311, y312, y321, y322, y331, y332,
y411, y412, y421, y422, y431, y432,
y811, y812, y821, y822, y831, y832,
y911, y912, y921, y922, y931, y932,
y1011, y1012, y1021, y1022, y1031, y1032,
Es1, Es11, Es2, Es21, Es3, Es31,
Vp, Tshortfall, Tsurplus, Es, Vs, Vpsq, Ep;


Parameters

p(s)  probability of the realization of scenario  /1 0.35, 2 0.45, 3 0.2/


Scalars

V1 variance of price of crude oil /0.352/
V2 variance of price of gasoline /1.882375/
V3 variance of price of naphtha /0.352/
V4 variance of price of jet fuel /0.859375/
V5 variance of price of heating oil /1.156375/
V6 variance of price of crude oil /0.198/
V14 variance of price of cracker feed /0.012375/


Equations

OBJ "maximize profit",
CON1, CON2, EQN8, EQN9, EQN10, EQN11, EQN12, EQN13, EQN14, EQN15, EQN16, EQN17, EQN18,
EQN19,
Escenario, Escenario1, Escenario11, Escenario2, Escenario21, Escenario3, Escenario31,
Vscenario,
CONgas1, CONgas2, CONgas3,
CONnap1, CONnap2, CONnap3,
CONjf1, CONjf2, CONjf3,
CONho1, CONho2, CONho3,
CONfo1, CONfo2, CONfo3,
CONlhsnap1, CONlhsnap11, CONlhsnap12,
CONlhsnap2, CONlhsnap21, CONlhsnap22,
CONlhsnap3, CONlhsnap31, CONlhsnap32,
CONlhsjf1, CONlhsjf11, CONlhsjf12,
CONlhsjf2, CONlhsjf21, CONlhsjf22,
CONlhsjf3, CONlhsjf31, CONlhsjf32,
CONlhsgo1, CONlhsgo11, CONlhsgo12,
CONlhsgo2, CONlhsgo21, CONlhsgo22,
CONlhsgo3, CONlhsgo31, CONlhsgo32,
CONlhscf1, CONlhscf11, CONlhscf12,
CONlhscf2, CONlhscf21, CONlhscf22,
CONlhscf3, CONlhscf31, CONlhscf32,
CONlhsr1, CONlhsr11, CONlhsr12,
CONlhsr2, CONlhsr21, CONlhsr22,
CONlhsr3, CONlhsr31, CONlhsr32,
Eprofit, Vprofit, Totalshortfall, Totalsurplus, Ecvar;

OBJ..
Z2 =E=
[p('1')*(-pc('1','1')*x1   +   pc('2','1')*x2   +   pc('3','1')*x3   +   pc('4','1')*x4   +
pc('5','1')*x5 + pc('6','1')*x6 - pc('14','1')*x14)]
+  [p('2')*(-pc('1','2')*x1  +  pc('2','2')*x2  +  pc('3','2')*x3  +  pc('4','2')*x4  +
pc('5','2')*x5 + pc('6','2')*x6 - pc('14','2')*x14)]
+  [p('3')*(-pc('1','3')*x1  +  pc('2','3')*x2  +  pc('3','3')*x3  +  pc('4','3')*x4  +
pc('5','3')*x5 + pc('6','3')*x6 - pc('14','3')*x14)]
- 0.0000000001*Vp
- Es - 50*Vs;
```

```
Vprofit.. Vp =E= ((V1*SQR(x1)) + (V2*SQR(x2)) + (V3*SQR(x3)) + (V4*SQR(x4)) + (V5*SQR(x5))
+ (V6*SQR(x6)) + (V14*SQR(x14)));

Vscenario.. Vs =E= ((p('1')*SQR(Es11 - Es)) + (p('2')*SQR(Es21 - Es)) + (p('3')*SQR(Es31 -
Es)));

CON1.. x1 =L= 15000;
CON2.. x14 =L= 2500;
*EQN3.. -0.13*x1 + x7 =E= 0;
*EQN4.. -0.15*x1 + x4 =E= 0;
*EQN5.. -0.22*x1 + x8 =E= 0;
*EQN6.. -0.20*x1 + x9 =E= 0;
*EQN7.. -0.30*x1 + x10 =E= 0;
EQN8.. -0.05*x14 + x20 =E= 0;
EQN9.. -0.40*x14 + x16 =E= 0;
EQN10.. -0.55*x14 + x17 =E= 0;
EQN11.. 0.5*x2 - x11 =E= 0;
EQN12.. 0.5*x2 - x16 =E= 0;
EQN13.. 0.75*x5 - x12 =E= 0;
EQN14.. 0.25*x5 - x18 =E= 0;
EQN15.. -x7 + x3 +x11 =E= 0;
EQN16.. -x8 + x12 +x13 =E= 0;
EQN17.. -x9 + x14 +x15 =E= 0;
EQN18.. -x17 + x18 +x19 =E= 0;
EQN19.. -x10 - x13 - x15 - x19 + x6 =E= 0;
*CON20.. x2 =L= 2700;
*CON21.. x3 =L= 1100;
*CON22.. x4 =L= 2300;
*CON23.. x5 =L= 1700;
*CON24.. x6 =L= 9500;

*Scenario 1: High demand
Escenario1.. Es1 =E= p('1')*Es11;;
Escenario11.. Es11 =E= (((c('2','1')*z211 + c('2','2')*z212) + (c('3','1')*z311 +
c('3','2')*z312) + (c('4','1')*z411 + c('4','2')*z412) + (c('5','1')*z511 +
c('5','2')*z512) + (c('6','1')*z611 + c('6','2')*z612))
+ ((q('3','1')*y311 + q('3','2')*y312) + (q('4','1')*y411 + q('4','2')*y412) +
(q('8','1')*y811 + q('8','2')*y812) + (q('9','1')*y911 + q('9','2')*y912) +
(q('10','1')*y1011 + q('10','2')*y1012)));

CONgas1.. x2 + z211 - z212 =E= d('2','1');
CONnap1.. x3 + z311 - z312 =E= d('3','1');
CONjf1.. x4 + z411 - z412 =E= d('4','1');
CONho1.. x5 + z511 - z512 =E= d('5','1');
CONfo1.. x6 + z611 - z612 =E= d('6','1');
CONlhsnap1.. y('3','1')*x1 + x7 + y311 - y312 =E= 0;
         CONlhsnap11.. y311 =L= 0.1*x1;
         CONlhsnap12.. y312 =L= 0.1*x1;
CONlhsjf1.. y('4','1')*x1 + x4 + y411 - y412 =E= 0;
         CONlhsjf11.. y411 =L= 0.1*x1;
         CONlhsjf12.. y412 =L= 0.1*x1;
CONlhsgo1.. y('8','1')*x1 + x8 + y811 - y812 =E= 0;
         CONlhsgo11.. y811 =L= 0.1*x1;
         CONlhsgo12.. y812 =L= 0.1*x1;
CONlhscf1.. y('9','1')*x1 + x9 + y911 - y912 =E= 0;
         CONlhscf11.. y911 =L= 0.1*x1;
         CONlhscf12.. y912 =L= 0.1*x1;
CONlhsr1.. y('10','1')*x1 + x10 + y1011 - y1012 =E= 0;
         CONlhsr11.. y1011 =L= 0.1*x1;
         CONlhsr12.. y1012 =L= 0.1*x1;

*Scenario 2: Medium demand
Escenario2.. Es2 =E= p('2')*Es21;
Escenario21.. Es21 =E= (((c('2','1')*z221 + c('2','2')*z222) + (c('3','1')*z321 +
c('3','2')*z322) + (c('4','1')*z421 + c('4','2')*z422) + (c('5','1')*z521 +
c('5','2')*z522) + (c('6','1')*z621 + c('6','2')*z622))
+ ((q('3','1')*y321 + q('3','2')*y322) + (q('4','1')*y421 + q('4','2')*y422) +
(q('8','1')*y821 + q('8','2')*y822) + (q('9','1')*y921 + q('9','2')*y922) +
(q('10','1')*y1021 + q('10','2')*y1022)));
CONgas2.. x2 + z221 - z222 =E= d('2','2');
```

```
CONnap2.. x3 + z321 - z322 =E= d('3','2');
CONjf2.. x4 + z421 - z422 =E= d('4','2');
CONho2.. x5 + z521 - z522 =E= d('5','2');
CONfo2.. x6 + z621 - z622 =E= d('6','2');
CONlhsnap2.. y('3','2')*x1 + x7 + y321 - y322 =E= 0;
         CONlhsnap21..y321 =L= 0.1*x1;
         CONlhsnap22.. y322 =L= 0.1*x1;
CONlhsjf2.. y('4','2')*x1 + x4 + y421 - y422 =E= 0;
         CONlhsjf21.. y421 =L= 0.1*x1;
         CONlhsjf22.. y422 =L= 0.1*x1;
CONlhsgo2.. y('8','2')*x1 + x8 + y821 - y822 =E= 0;
         CONlhsgo21.. y821 =L= 0.1*x1;
         CONlhsgo22.. y822 =L= 0.1*x1;
CONlhscf2.. y('9','2')*x1 + x9 + y921 - y922 =E= 0;
         CONlhscf21.. y921 =L= 0.1*x1;
         CONlhscf22.. y922 =L= 0.1*x1;
CONlhsr2.. y('10','2')*x1 + x10 + y1021 - y1022 =E= 0;
         CONlhsr21.. y1011 =L= 0.1*x1;
         CONlhsr22.. y1012 =L= 0.1*x1;

*Scenario 3: Low demand
Escenario3.. Es3 =E= p('3')*Es31;
Escenario31..  Es31  =E=  (((c('2','1')*z231  +  c('2','2')*z232)  +  (c('3','1')*z331  +
c('3','2')*z332)   +   (c('4','1')*z431   +   c('4','2')*z432)   +   (c('5','1')*z531   +
c('5','2')*z532) + (c('6','1')*z631 + c('6','2')*z632))
+  ((q('3','1')*y331  +  q('3','2')*y332)  +  (q('4','1')*y431  +  q('4','2')*y432)  +
(q('8','1')*y831   +   q('8','2')*y832)   +   (q('9','1')*y931   +   q('9','2')*y932)   +
(q('10','1')*y1031 + q('10','2')*y1032)));
CONgas3.. x2 + z231 - z232 =E= d('2','3');
CONnap3.. x3 + z331 - z332 =E= d('3','3');
CONjf3.. x4 + z431 - z432 =E= d('4','3');
CONho3.. x5 + z531 - z532 =E= d('5','3');
CONfo3.. x6 + z631 - z632 =E= d('6','3');
CONlhsnap3.. y('3','3')*x1 + x7 + y331 - y332 =E= 0;
         CONlhsnap31..y331 =L= 0.1*x1;
         CONlhsnap32.. y332 =L= 0.1*x1;
CONlhsjf3.. y('4','3')*x1 + x4 + y431 - y432 =E= 0;
         CONlhsjf31.. y431 =L= 0.1*x1;
         CONlhsjf32.. y432 =L= 0.1*x1;
CONlhsgo3.. y('8','3')*x1 + x8 + y831 - y832 =E= 0;
         CONlhsgo31.. y831 =L= 0.1*x1;
         CONlhsgo32.. y832 =L= 0.1*x1;
CONlhscf3.. y('9','3')*x1 + x9 + y931 - y932 =E= 0;
         CONlhscf31.. y931 =L= 0.1*x1;
         CONlhscf32.. y932 =L= 0.1*x1;
CONlhsr3.. y('10','3')*x1 + x10 + y1031 - y1032 =E= 0;
         CONlhsr31.. y1031 =L= 0.1*x1;
         CONlhsr32.. y1032 =L= 0.1*x1;

Escenario.. Es =E= Es1 + Es2 + Es3;


Eprofit.. Ep =E=
[p('1')*(-pc('1','1')*x1   +   pc('2','1')*x2   +   pc('3','1')*x3   +   pc('4','1')*x4   +
pc('5','1')*x5 + pc('6','1')*x6 - pc('14','1')*x14)]
+  [p('2')*(-pc('1','2')*x1  +  pc('2','2')*x2  +  pc('3','2')*x3  +  pc('4','2')*x4  +
pc('5','2')*x5 + pc('6','2')*x6 - pc('14','2')*x14)]
+  [p('3')*(-pc('1','3')*x1  +  pc('2','3')*x2  +  pc('3','3')*x3  +  pc('4','3')*x4  +
pc('5','3')*x5 + pc('6','3')*x6 - pc('14','3')*x14)];

Totalshortfall.. Tshortfall =E= z211 + z221 + z231 + z311 + z321 + z331 + z411 + z421 +
z431 + z511 + z521 + z531 + z611 + z621 + z631
+ y311 + y321 + y331 + y411 + y421 + y431 + y811 + y821 + y831 + y911 + y921 + y931 +
y1011 + y1021 + y1031;

Totalsurplus.. Tsurplus =E= z212 + z222 + z232 + z312 + z322 + z332 + z412 + z422 + z432 +
z512 + z522 + z532 + z612 + z622 + z632
+ y312 + y322 + y332 + y412 + y422 + y432 + y812 + y822 + y832 + y912 + y922 + y932 +
y1012 + y1022 + y1032;

*Vprofit2..  Vp2  =E=  (v1*SQR(x1))  +  (v2*SQR(x2))  +  (v3*SQR(x3))  +  (v4*SQR(x4))  +
```

```
(v5*SQR(x5)) + (v6*SQR(x6)) + (v14*SQR(x14));

*Vpsqrt.. Vpsq =E= SQRT(Vp + Vs);

Ecvar.. Ecv =E= Ep - Es;
*Covar.. Cv =E= Vpsq/(Ep - Es);


Model Refinery / all /;

* Starting values
x1.up = 15000;          x1.l = 12500;
x2.up = 2700;           x2.l = 2000;
x3.up = 1100;           x3.l = 625;
x4.up = 2300;           x4.l = 1875;
x5.up = 1700;           x5.l = 1700;
x6.up = 9500;           x6.l = 6175;
x7.up = 1950;           x7.l = 1625;
x8.up = 3300;           x8.l = 2750;
x9.up = 3000;           x9.l = 2500;
x10.up = 3000;          x10.l = 3750;
x11.up = 1350;          x11.l = 1000;
x12.up = 1275;          x12.l = 1275;
x13.up = 3300;          x13.l = 1475;
x14.up = 3000;          x14.l = 2500;
x15.up = 3000;          x15.l = 0;
x16.up = 1200;          x16.l = 1000;
x17.up = 1650;          x17.l = 1375;
x18.up = 425;           x18.l = 425;
x19.up = 1650;          x19.l = 950;
x20.up = 150;           x20.l = 125;

Options nlp = conopt3;

Solve REFINERY USING NLP MAXIMIZING Z2;
```

Appendix D6: Approach 4—Risk Model III with Mean-Absolute Deviation (MAD) as the Measure of Risk Imposed by the Recourse Penalty Costs

```
$TITLE Approach 4: Risk Model III of Two-Stage Stochastic Programming with Fixed Recourse
for Minimization of the Expected Value and the Mean-Absolute Deviation (MAD) of the
Variation in Recourse Penalty Costs


Sets

i "product type" /1*21/
s "scenarios" /1*3/
k "production shortfall and surplus or yield decrement or increment" /1, 2/


Table pc(i,s) "price of product type i per realization j"
         1         2         3
1        8.8       8.0       7.2
2        20.35     18.5      16.65
3        8.8       8.0       7.2
4        13.75     12.5      11.25
5        15.95     14.5      13.05
6        6.6       6         5.4
14       1.65      1.5       1.35;


Table d(i,s) "demand of product type i per realization j"
      1        2        3
2     2835     2700     2565
3     1155     1100     1045
4     2415     2300     2185
5     1785     1700     1615
6     9975     9500     9025;


Table y(i,s) "yield of product type i per realization j"
      1            2         3
3     -0.1365      -0.13     -0.1235
4     -0.1575      -0.15     -0.1425
8     -0.231       -0.22     -0.209
9     -0.21        -0.20     -0.19
10    -0.265       -0.30     -0.335;


Table c(i,k) "penalty cost for product type i due to production shortfall or surplus"
      1     2
1     55    50
2     25    20
3     17    13
4     5     4
5     6     5
6     10    8;


Table q(i,k) "penalty cost for product type i due to yield decrement or increment"
      1     2
3     5     3
4     5     4
8     5     3
9     5     3
10    5     3


Variables

Z2, Ecv;
```

```
Positive variables

x1, x2, x3, x4, x5, x6, x7, x8, x9, x10, x11, x12, x13, x14, x15, x16, x17, x18, x19, x20,
z211, z212, z221, z222, z231, z232,
z311, z312, z321, z322, z331, z332,
z411, z412, z421, z422, z431, z432,
z511, z512, z521, z522, z531, z532,
z611, z612, z621, z622, z631, z632,
y311, y312, y321, y322, y331, y332,
y411, y412, y421, y422, y431, y432,
y811, y812, y821, y822, y831, y832,
y911, y912, y921, y922, y931, y932,
y1011, y1012, y1021, y1022, y1031, y1032,
Es1, Es11, Es2, Es21, Es3, Es31,
MADs, MADs1,
Vp, Tshortfall, Tsurplus, Es, Ep;


Parameters

p(s)   probability of the realization of scenario   /1 0.35, 2 0.45, 3 0.2/


Scalars

V1 variance of price of crude oil /0.352/
V2 variance of price of gasoline /1.882375/
V3 variance of price of naphtha /0.352/
V4 variance of price of jet fuel /0.859375/
V5 variance of price of heating oil /1.156375/
V6 variance of price of crude oil /0.198/
V14 variance of price of cracker feed /0.012375/


Equations

OBJ "maximize profit",
CON1, CON2, EQN8, EQN9, EQN10, EQN11, EQN12, EQN13, EQN14, EQN15, EQN16, EQN17, EQN18,
EQN19,
Escenario, Escenario1, Escenario11,
Escenario2, Escenario21,
Escenario3, Escenario31,
CONgas1, CONgas2, CONgas3,
CONnap1, CONnap2, CONnap3,
CONjf1, CONjf2, CONjf3,
CONho1, CONho2, CONho3,
CONfo1, CONfo2, CONfo3,
CONlhsnap1, CONlhsnap11, CONlhsnap12,
CONlhsnap2, CONlhsnap21, CONlhsnap22,
CONlhsnap3, CONlhsnap31, CONlhsnap32,
CONlhsjf1, CONlhsjf11, CONlhsjf12,
CONlhsjf2, CONlhsjf21, CONlhsjf22,
CONlhsjf3, CONlhsjf31, CONlhsjf32,
CONlhsgo1, CONlhsgo11, CONlhsgo12,
CONlhsgo2, CONlhsgo21, CONlhsgo22,
CONlhsgo3, CONlhsgo31, CONlhsgo32,
CONlhscf1, CONlhscf11, CONlhscf12,
CONlhscf2, CONlhscf21, CONlhscf22,
CONlhscf3, CONlhscf31, CONlhscf32,
CONlhsr1, CONlhsr11, CONlhsr12,
CONlhsr2, CONlhsr21, CONlhsr22,
CONlhsr3, CONlhsr31, CONlhsr32,
MADscenario, MADcon1, MADcon2, MADcon3,
Eprofit, Vprofit, Totalshortfall, Totalsurplus, Ecvar;

OBJ..
Z2 =e=
[p('1')*(-pc('1','1')*x1   +   pc('2','1')*x2   +   pc('3','1')*x3   +   pc('4','1')*x4   +
pc('5','1')*x5 + pc('6','1')*x6 - pc('14','1')*x14)]
+  [p('2')*(-pc('1','2')*x1  +  pc('2','2')*x2  +  pc('3','2')*x3  +  pc('4','2')*x4  +
```

```
pc('5','2')*x5 + pc('6','2')*x6 - pc('14','2')*x14)]
+  [p('3')*(-pc('1','3')*x1  +  pc('2','3')*x2  +  pc('3','3')*x3  +  pc('4','3')*x4  +
pc('5','3')*x5 + pc('6','3')*x6 - pc('14','3')*x14)]
- 0.0008*Vp
- Es - 0.01*MADs;

Vprofit.. Vp =E= ((V1*SQR(x1)) + (V2*SQR(x2)) + (V3*SQR(x3)) + (V4*SQR(x4)) + (V5*SQR(x5))
+ (V6*SQR(x6)) + (V14*SQR(x14)));

CON1.. x1 =L= 15000;
CON2.. x14 =L= 2500;
*EQN3.. -0.13*x1 + x7 =E= 0;
*EQN4.. -0.15*x1 + x4 =E= 0;
*EQN5.. -0.22*x1 + x8 =E= 0;
*EQN6.. -0.20*x1 + x9 =E= 0;
*EQN7.. -0.30*x1 + x10 =E= 0;
EQN8.. -0.05*x14 + x20 =E= 0;
EQN9.. -0.40*x14 + x16 =E= 0;
EQN10.. -0.55*x14 + x17 =E= 0;
EQN11.. 0.5*x2 - x11 =E= 0;
EQN12.. 0.5*x2 - x16 =E= 0;
EQN13.. 0.75*x5 - x12 =E= 0;
EQN14.. 0.25*x5 - x18 =E= 0;
EQN15.. -x7 + x3 +x11 =E= 0;
EQN16.. -x8 + x12 +x13 =E= 0;
EQN17.. -x9 + x14 +x15 =E= 0;
EQN18.. -x17 + x18 +x19 =E= 0;
EQN19.. -x10 - x13 - x15 - x19 + x6 =E= 0;
*CON20.. x2 =L= 2700;
*CON21.. x3 =L= 1100;
*CON22.. x4 =L= 2300;
*CON23.. x5 =L= 1700;
*CON24.. x6 =L= 9500;

*Scenario 1: High demand
Escenario1.. Es1 =E= p('1')*Es11;
Escenario11..  Es11  =E=  (((c('2','1')*z211  +  c('2','2')*z212)  +  (c('3','1')*z311  +
c('3','2')*z312)   +   (c('4','1')*z411   +   c('4','2')*z412)   +   (c('5','1')*z511   +
c('5','2')*z512) + (c('6','1')*z611 + c('6','2')*z612))
+  ((q('3','1')*y311  +  q('3','2')*y312)  +  (q('4','1')*y411  +  q('4','2')*y412)  +
(q('8','1')*y811   +   q('8','2')*y812)   +   (q('9','1')*y911   +   q('9','2')*y912)   +
(q('10','1')*y1011 + q('10','2')*y1012)));

CONgas1.. x2 + z211 - z212 =E= d('2','1');
CONnap1.. x3 + z311 - z312 =E= d('3','1');
CONjf1.. x4 + z411 - z412 =E= d('4','1');
CONho1.. x5 + z511 - z512 =E= d('5','1');
CONfo1.. x6 + z611 - z612 =E= d('6','1');
CONlhsnap1.. y('3','1')*x1 + x7 + y311 - y312 =E= 0;
         CONlhsnap11.. y311 =L= 0.1*x1;
         CONlhsnap12.. y312 =L= 0.1*x1;
CONlhsjf1.. y('4','1')*x1 + x4 + y411 - y412 =E= 0;
         CONlhsjf11.. y411 =L= 0.1*x1;
         CONlhsjf12.. y412 =L= 0.1*x1;
CONlhsgo1.. y('8','1')*x1 + x8 + y811 - y812 =E= 0;
         CONlhsgo11.. y811 =L= 0.1*x1;
         CONlhsgo12.. y812 =L= 0.1*x1;
CONlhscf1.. y('9','1')*x1 + x9 + y911 - y912 =E= 0;
         CONlhscf11.. y911 =L= 0.1*x1;
         CONlhscf12.. y912 =L= 0.1*x1;
CONlhsr1.. y('10','1')*x1 + x10 + y1011 - y1012 =E= 0;
         CONlhsr11.. y1011 =L= 0.1*x1;
         CONlhsr12.. y1012 =L= 0.1*x1;

*Scenario 2: Medium demand
Escenario2.. Es2 =E= p('2')*Es21;
Escenario21..  Es21  =E=  (((c('2','1')*z221  +  c('2','2')*z222)  +  (c('3','1')*z321  +
c('3','2')*z322)   +   (c('4','1')*z421   +   c('4','2')*z422)   +   (c('5','1')*z521   +
c('5','2')*z522) + (c('6','1')*z621 + c('6','2')*z622))
+  ((q('3','1')*y321  +  q('3','2')*y322)  +  (q('4','1')*y421  +  q('4','2')*y422)  +
(q('8','1')*y821   +   q('8','2')*y822)   +   (q('9','1')*y921   +   q('9','2')*y922)   +
```

```
(q('10','1')*y1021 + q('10','2')*y1022)));
CONgas2.. x2 + z221 - z222 =E= d('2','2');
CONnap2.. x3 + z321 - z322 =E= d('3','2');
CONjf2.. x4 + z421 - z422 =E= d('4','2');
CONho2.. x5 + z521 - z522 =E= d('5','2');
CONfo2.. x6 + z621 - z622 =E= d('6','2');
CONlhsnap2.. y('3','2')*x1 + x7 + y321 - y322 =E= 0;
         CONlhsnap21..y321 =L= 0.1*x1;
         CONlhsnap22.. y322 =L= 0.1*x1;
CONlhsjf2.. y('4','2')*x1 + x4 + y421 - y422 =E= 0;
         CONlhsjf21.. y421 =L= 0.1*x1;
         CONlhsjf22.. y422 =L= 0.1*x1;
CONlhsgo2.. y('8','2')*x1 + x8 + y821 - y822 =E= 0;
         CONlhsgo21.. y821 =L= 0.1*x1;
         CONlhsgo22.. y822 =L= 0.1*x1;
CONlhscf2.. y('9','2')*x1 + x9 + y921 - y922 =E= 0;
         CONlhscf21.. y921 =L= 0.1*x1;
         CONlhscf22.. y922 =L= 0.1*x1;
CONlhsr2.. y('10','2')*x1 + x10 + y1021 - y1022 =E= 0;
         CONlhsr21.. y1011 =L= 0.1*x1;
         CONlhsr22.. y1012 =L= 0.1*x1;

*Scenario 3: Low demand
Escenario3.. Es3 =E= p('3')*Es31;
Escenario31.. Es31 =E= (((c('2','1')*z231 + c('2','2')*z232) + (c('3','1')*z331 +
c('3','2')*z332) + (c('4','1')*z431 + c('4','2')*z432) + (c('5','1')*z531 +
c('5','2')*z532) + (c('6','1')*z631 + c('6','2')*z632))
+ ((q('3','1')*y331 + q('3','2')*y332) + (q('4','1')*y431 + q('4','2')*y432) +
(q('8','1')*y831 + q('8','2')*y832) + (q('9','1')*y931 + q('9','2')*y932) +
(q('10','1')*y1031 + q('10','2')*y1032)));
CONgas3.. x2 + z231 - z232 =E= d('2','3');
CONnap3.. x3 + z331 - z332 =E= d('3','3');
CONjf3.. x4 + z431 - z432 =E= d('4','3');
CONho3.. x5 + z531 - z532 =E= d('5','3');
CONfo3.. x6 + z631 - z632 =E= d('6','3');
CONlhsnap3.. y('3','3')*x1 + x7 + y331 - y332 =E= 0;
         CONlhsnap31..y331 =L= 0.1*x1;
         CONlhsnap32.. y332 =L= 0.1*x1;
CONlhsjf3.. y('4','3')*x1 + x4 + y431 - y432 =E= 0;
         CONlhsjf31.. y431 =L= 0.1*x1;
         CONlhsjf32.. y432 =L= 0.1*x1;
CONlhsgo3.. y('8','3')*x1 + x8 + y831 - y832 =E= 0;
         CONlhsgo31.. y831 =L= 0.1*x1;
         CONlhsgo32.. y832 =L= 0.1*x1;
CONlhscf3.. y('9','3')*x1 + x9 + y931 - y932 =E= 0;
         CONlhscf31.. y931 =L= 0.1*x1;
         CONlhscf32.. y932 =L= 0.1*x1;
CONlhsr3.. y('10','3')*x1 + x10 + y1031 - y1032 =E= 0;
         CONlhsr31.. y1031 =L= 0.1*x1;
         CONlhsr32.. y1032 =L= 0.1*x1;

Escenario.. Es =E= Es1 + Es2 + Es3;
MADscenario.. MADs =E= (p('1')*abs(Es11 - Es)) + (p('2')*abs(Es21 - Es)) +
(p('3')*abs(Es31 - Es));
MADcon1.. MADs1 =G= -MADs;
MADcon2.. MADs1 =G= MADs;
MADcon3.. MADs1 =G= 0;
*Vscenario.. Vs =E= (p1*SQR(Es11 - Es)) + (p2*SQR(Es21 - Es)) + (p3*SQR(Es31 - Es));


Eprofit.. Ep =E=
[p('1')*(-pc('1','1')*x1 + pc('2','1')*x2 + pc('3','1')*x3 + pc('4','1')*x4 +
pc('5','1')*x5 + pc('6','1')*x6 - pc('14','1')*x14)]
+ [p('2')*(-pc('1','2')*x1 + pc('2','2')*x2 + pc('3','2')*x3 + pc('4','2')*x4 +
pc('5','2')*x5 + pc('6','2')*x6 - pc('14','2')*x14)]
+ [p('3')*(-pc('1','3')*x1 + pc('2','3')*x2 + pc('3','3')*x3 + pc('4','3')*x4 +
pc('5','3')*x5 + pc('6','3')*x6 - pc('14','3')*x14)];

Totalshortfall.. Tshortfall =E= z211 + z221 + z231 + z311 + z321 + z331 + z411 + z421 +
z431 + z511 + z521 + z531 + z611 + z621 + z631
+ y311 + y321 + y331 + y411 + y421 + y431 + y811 + y821 + y831 + y911 + y921 + y931 +
```

```
y1011 + y1021 + y1031;

Totalsurplus.. Tsurplus =E= z212 + z222 + z232 + z312 + z322 + z332 + z412 + z422 + z432 +
z512 + z522 + z532 + z612 + z622 + z632
+ y312 + y322 + y332 + y412 + y422 + y432 + y812 + y822 + y832 + y912 + y922 + y932 +
y1012 + y1022 + y1032;

*Vpsqrt.. Vpsq =E= SQRT(Vp + MADs);

Ecvar.. Ecv =E= Ep - Es;

Model Refinery / all /;

* Starting values
x1.up = 15000;          x1.l = 12500;
x2.up = 2700;           x2.l = 2000;
x3.up = 1100;           x3.l = 625;
x4.up = 2300;           x4.l = 1875;
x5.up = 1700;           x5.l = 1700;
x6.up = 9500;           x6.l = 6175;
x7.up = 1950;           x7.l = 1625;
x8.up = 3300;           x8.l = 2750;
x9.up = 3000;           x9.l = 2500;
x10.up = 3000;          x10.l = 3750;
x11.up = 1350;          x11.l = 1000;
x12.up = 1275;          x12.l = 1275;
x13.up = 3300;          x13.l = 1475;
x14.up = 3000;          x14.l = 2500;
x15.up = 3000;          x15.l = 0;
x16.up = 1200;          x16.l = 1000;
x17.up = 1650;          x17.l = 1375;
x18.up = 425;           x18.l = 425;
x19.up = 1650;          x19.l = 950;
x20.up = 150;           x20.l = 125;

Option dnlp = conopt3;

Solve REFINERY USING DNLP MAXIMIZING Z2;
```

Final Remarks

Manuscript on parts of this dissertation have been submitted for publication as indicated in the following:

Khor, C. S., A. Elkamel, and P. L. Douglas. Stochastic Refinery Planning with Risk Management. To appear in *Petroleum Science and Technology* (accepted October 21, 2006), 16 pages.

Khor, C. S., A. Elkamel, K. Ponnambalam, and P. L. Douglas. Two-stage stochastic programming with fixed recourse via scenario planning with financial and operational risk management for petroleum refinery planning under uncertainty. *Chemical Engineering and Process*ing (submitted for publication and under review), 63 pages.

Brief Biographical Notes on the Author

Cheng Seong Khor is currently on study leave from a teaching position in the PETRONAS University of Technology, the academic institution set up by PETRONAS, the Malaysian state-owned national oil corporation, to pursue his graduate studies in the Department of Chemical Engineering at the University of Waterloo, Ontario, Canada. His general research interests concern the theory and application of the optimization approach to chemical process systems engineering (PSE) problems, particularly in the hydrocarbon industry. He holds a Bachelor in Chemical Engineering (Honours) degree from the National University of Malaysia (Universiti Kebangsaan Malaysia).