Implication of the Jensen's Inequality for System Dynamic Simulations: Application to a Plug & Play Integrated Refinery Supply Chain Model

By

Juan Esteban Montero

Master of Science in Finance, Universidad de Chile (2011)
Ingeniero Civil de Minas, Universidad de Santiago de Chile (2007)

SUBMITTED TO THE SYSTEM DESIGN AND MANAGEMENT PROGRAM IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN ENGINEERING AND MANAGEMENT
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
FEBRUARY 2014

Signature of Author ____

System Design and Management Program
January, 2014

Certified by ____

Professor Richard de Neufville
Thesis Supervisor
Professor of Engineering Systems and Civil and Environmental Engineering

Accepted by ____

Patrick Hale
Director
System Design and Management Program

(Blank page)

Implication of the Jensen's Inequality for System Dynamic Simulations: Application to a Plug & Play Integrated Refinery Supply Chain Model

By

Juan Esteban Montero

Submitted to the System Design and Management Program on February, 2014 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Engineering and Management.

ABSTRACT

This investigation studies how critical is the effect of considering uncertainty to a dynamic model because of Jensen's Inequality. This is done using as an example the supply chain of a refinery, which illustrates that the difference between probable and expected results can be significant, arguing that the distributions and probabilities can be dramatically different from the expected-planned value. Moreover, this research discusses that, from the perspective of the dynamics of the system, the mode of behavior can vary considerably as well, leading managers to dissimilar situations and contexts that will inevitably produce different decisions or strategies.

Supply chain management is a critical aspect of any business. The energy industry is a particularly relevant example of a global supply chain, representing a crucial challenge the management of complexity and relevance for the overall performance of the business.

The complexity of managing the supply chain of an energy company is produced by the physical size, diversity of operations and products and dynamics of the system, among many others causes. On top of the intrinsic complexity of the business itself, the manager of a supply chain should also consider the complexity of the models and methodologies used to make decisions about it. These models and methodologies are diverse and they serve different purposes under certain assumptions.

This study also discusses the complexity faced by supply chain managers, presenting a compilation of bibliographic research about different considerations and approaches. Managers often employ models and analytics to simplify the complexity and produce intuition by different means in order to form their decisions and strategies.

The analysis of the effects of uncertainty on the results and behavior of a dynamic simulation model is done by stochastically simulating an already-developed -plug & play- dynamic model of a refinery. This approach permits the exploration of different configurations, considering different definitions of uncertainty, analyzing and comparing their particular results.

Thesis Supervisor: Richard de Neufville
Title: Professor of Engineering Systems and Civil and Environmental Engineering

(Blank)

ACKNOWLEDGMENTS

Thank you for reading this thesis.

Thanks to my parents for inspiring me with their work ethic and strength to live their lives.

Joy Logan, David Cooke and Mike Neeley, from Chevron: Thanks for your support, mentoring and for setting an incredibly challenging date as a deadline for this project. Only challenging goals are worthy goals.

To Dr. Rajagopalan Srinivasan and Dr. Arief Adhitya for generously allowing me to use the model "IRIS" for this research and for guiding me to understand the fundamentals of its assumptions and mechanics.

To my Professor and Advisor, Richard de Neufville: Jensen's inequality is the least important thing that you taught me. I will be always inspired by how you care about people and how always, even during incredibly busy times for you, you found the way to make our project a priority and guide me wisely.

To my classmates and friends from MIT for the biggest lesson of humility and hard work. I admire you deeply and I feel motivated to work harder and smarter because I know you.

And to my wife, Claire Jarvis, for your patience and love.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Managers and decision makers often recur to models and analytics as tools to understand the real systems in which they work and define optimal decisions and goals. Very commonly, these models are complex in both combinatorial and dynamic ways, overwhelming the decision makers and making almost impossible to understand intuitively the behavior of the system and to verify that the results are actually the desired solution.

In most cases, models and simulations cannot be verified to prove accuracy. This is the main reason why the mathematical and physical laws have to be considered from the very beginning to construct (or at least try to) good models. One of the mathematical concepts that has to be considered by managers and decision makers is the Jensen's Inequality, which states that the "average of all the possible outcomes associated with uncertain parameters, generally does not equal the value obtained from using the average value of the parameters" (de Neufville, 2012).

This challenge evokes the famous computer science problem called P versus NP. This challenge states that "for all problems for which an algorithm can verify a given solution quickly (that is, in polynomial time), an algorithm can also find that solution quickly" (Aaronson, 2013). In practice, this has a profound implication for a decision maker, because it means that the models used to estimate results are, in practice, the only available tool to understand reality, but they are not verifiable (at least beforehand).

> *If P = NP, then the world would be a profoundly different place than we usually assume it to be. There would be no special value in 'creative leaps,' no fundamental gap between solving a problem and recognizing the solution once it's found. Everyone who could appreciate a symphony would be Mozart; everyone who could follow a step-by-step argument would be Gauss...* — Scott Aaronson, MIT

Systems Dynamics, founded at MIT Sloan in 1956 by Professor Jay W. Forrester, is a discipline that combines theory, methods, and philosophy to analyze the behavior of systems. "This is not only applied to management, but also in such other fields as environmental change, politics, economic behavior, medicine, and engineering" (MIT Sloan School of Management, 2013).

Dynamic simulations are highly non-linear (Sterman, 2000) and therefore the effects of Jensen's Inequality must be considered as a cause of producing the wrong results. This research evaluates the expected value of the result of a dynamic simulation, but also the different modes of behavior produced by those dynamic simulations.

The different effects of uncertainty can be relevant for the decision makers since they could certainly produce different behaviors due to different pressures and contextual factors. These differences could represent dissimilar strategies, decisions, risk perception and other factors that are certainly significant from the point of view of any manager.

The effects of Jensen's inequality on the expected value of a model have been studied in the past. The novelty of this research lays in the focus of the effects of Jensen's Inequality on the mode of behavior of a dynamic system.

SUMMARY

The exercise of exploring different configurations of uncertainty over the Integrated Refinery In-Silico (IRIS) plug & Play model produced the following observations:

- Considering uncertainty is necessary to obtain the correct result due to Jensen's Inequality, however, considering the fact that the dynamic complexity of a system like IRIS makes impossible to intuitively understand the behaviors of the model, it is also recommendable to explore the mode of behaviors and the dominance of the different feedback loop without uncertainty to simplify the analysis.

- IRIS served as a useful and didactic plug & play model allowing different configurations and simulations with and without stochastic uncertainty. This model also exemplifies that even simple structures are impacted severely by dynamic complexity. One example was observed by analyzing the behavior of the total profits over a period of 240 days without uncertainty (Appendix 9: The dynamics of not achieving demand).

- The configuration of uncertainty of the original setting specifies a random number seed, which determines the sequence of random numbers generated. This means that if the simulation is repeated using the same seeds, IRIS should get exactly the same sequence of random numbers and thus the same results. This representation of variability originally used by IRIS is useful and certainly more valuable than not considering uncertainty at all, but it doesn't consider the effects of Jensen's Inequality on the result of the simulation.

- The comparison between the original case and the configuration without uncertainty (Case 0), from the perspective of the total profits of the system for the first 120 days of simulation (Figure 10) suggests different mode of behaviors. This difference could be material from the point of view of the manager since the decisions and strategies derived from the results and trends will be significantly influenced.

- Even a very simple configuration of stochastic uncertainty, like the one proposed in Case 2, produced the following effects:
    - An expensive computation, that even with advanced implementations like parallel implementation, will still require time and memory to be computed.
    - There is significant evidence that the hypothesis that the value obtained from the original simulation (Base Case) and the expected value of the stochastic simulation are the same has to be rejected.
    - The range of probable results goes from almost zero profit after 120 days to almost $120 million, which is approximately twice the value estimated without stochasticity (Base Case). These differences are material from the point of view of the risk evaluation of the project.
    - It is suggested by the results of the stochastic simulation over the time horizon (120 days for Case 2) that the effect of the Jensen's Inequality over the variation of the result of the simulation grows over time. This effect seems to be a propagation of the effect of the Flaw of Averages over each time step of the dynamic simulation (Figure 20)

- Dynamic models are composed by variables connected in a particular structural relationship. These structures produce behaviors (Sterman, 2000) and therefore the effects of uncertainty on different parts of the structures are interesting for the decision maker because they will produce different behaviors. The comparison of Case 1a and Case 1b illustrated the different impacts that uncertainty can have in different modules of IRIS. In the first case (1a) uncertainty just produced an impact on the final result, but not in the mode of behavior. Case 1b, in contrast, produced both an impact on the final result and the mode of behavior of the system.

- The comparison of two different applications of uncertainty (case 1a and Case 1b) also showed that the expected value and the risk of the system can change in different directions. In this particular example, Case 1a produced a smaller expected profit than Case 1b, but also a smaller risk, represented by the standard deviation (Figure 15). The implication of this observation for the decision maker is that the decision of defining uncertainty in the modeling and analytic process is critical, since it will produce different effects depending where and how it is defined. The challenge from an architectural point of view is that it is not possible to define, at least beforehand, which one is better that the other one, since it will depend on the particular use required for the application.

- Different configurations of uncertainty could also have different impacts on the behavior of the system. The reason for this is that the dominance of each of the feedback loops of the system will depend on many different aspects, but mainly the values assigned for each variable. This can be added to the fact that the delays will behave in a different way depending on the inflow and outflow values of the stock variables representing the delay in the dynamic system. In IRIS, this is exemplified by the fact that the original configuration is not producing as much as demanded, which will be important for the behavior once the stocks are depleted (Appendix 9: The dynamics of not achieving demand).

- A sensitivity analysis suggested two different aspects:
    - The computation of the model can become a critical barrier in practice for the decision maker. Even the simple case with a parallel implementation of iterating the model for each variable required many hours that most of the time will not be available.
    - It is possible to identify the variables that are more relevant in impacting the result and behavior of the model when stochastically simulated (Figure 21). This information is useful for reducing the computational effort required, but also to define where to focus efforts.

In general, it is clear from the bibliographic research that there is a significant academic and industrial consensus on how important and challenging it is to manage supply chains. The detailed and dynamic complexity are both present in any supply chain system, and they are especially present in an energy company due to their global reach and diversity and scale of products.

The methods for modelling and analyzing a supply chain system are diverse and complex, creating a challenge just from the perspective of the techniques. The confusion between improvement and optimization can reduce the effectiveness of the efforts taken by the decision maker. Complexity is also expensive to compute and it could become a material barrier to overcome. Yet, it is important to keep in mind that even sophisticated and well developed models are limited.

Model relaxation could be a useful strategy if the appropriate non-relevant aspects are ignored. However, it is very important to consider that model scope is not just about relaxation but also about adding complexity when required. This is an architectural decision of any supply chain management endeavor.

Because of the mentioned reasons, dynamic simulations are highly non-linear (Sterman, 2000) and therefore the effects of Jensen's Inequality must be considered as a cause of producing the wrong results. IRIS served as a practical and flexible platform for testing the different effects. It is expected to find similar, and even bigger effects on more complex and realistic models. For this reason it is encouraged to research further the effects of Jensen's Inequality with more models and more specific applications.

RESEARCH OBJECTIVES

Dynamic systems are often simulated using deterministic inputs. These kinds of simulations are highly non-linear (Sterman, 2000) and therefore the effects of Jensen's Inequality, must be considered as a cause of producing the wrong results.

Main hypothesis for the thesis research:

Not considering uncertainty in the inputs of a dynamic simulation produces the wrong results and a different mode of behavior.

Research questions:

- Does Jensen's inequality alter the expected value of the result of a dynamic simulation?
- Does it also produce different modes of behavior of a dynamic simulation?
- How can these different results and modes of behavior be relevant for the decision makers?

BIBLIOGRAPHIC RESEARCH

THE IMPORTANCE OF OPTIMIZING THE VALUE CHAIN

The supply chain can be a significant aspect of a business, impacting how the value is transformed by a particular company. This chain of value involves the handling of raw materials, parts, processes and others that can together represent, conservatively estimated, up to 70% of the total cost of a product (Pal, et al., 2011). For this reason, supply chain management has been an attractive field for intense academic and industrial research. Several papers have approached this topic considering different themes, most commonly: general trends and issues in supply chain, dynamic modeling approaches, supply chain performance management issues, process maturity-supply chain performance relation, KPI prioritization and dependence, and human and organizational sides of supply chain performance management (Akyuz & Erkan, 2010). Actually, results show that if a company searches for supply chain cost reduction via coordinating their actions, superior performance is achievable (Disney, et al., 2006).

Supply Chain Management has become a field in itself, studying from the impact of managing the supply chain on the overall performance of the business and reducing the risks, to the efficiency of a supply chain management including "dimensions connected with real goods and services flows" (Lichocik & Sadowski, 2013). Since supply chains are commonly intricate international networks, even geopolitical considerations can have significant effects (Spillan, et al., 2013).

The considerations of planning, like "mindful planning processes help an organization avoid disruptions and be more resilient" (Ojha, et al., 2013) and the approach to treat the internal relationship to define value and performance are also included in the management of a supply chain. Some researchers argue that "it makes sense for managers to consider categorizing supply chain relationships similar to the way they categorize their end-user relationships" (Tokman, et al., 2013) which would eventually have a significant impact on the business.

It is common to see that practical results and real systems don't necessary show the level of improvements that are theoretically available, even with intense academic and industrial research on methodologies and technologies to optimize or improve value chains, these new ideas and technologies haven't necessary led to improved performance "because managers lack a framework for deciding which ones are best for their particular company's situation…" (Fisher, 1997).

YES, IT IS COMPLEX, AND NOT JUST BECAUSE OF THE DETAILS

Management of a supply chains is a complex task for several reasons: the physical size of the supply network and its inherent uncertainties (Papageorgiou, 2009) are commonly cited causes of complexity, still, a supply chain is a system within a bigger system, and therefore, it should recognize financial, social and environmental elements, in order to include sustainable development considerations for improving social and environmental impacts of production systems (Hall, et al., 2012). Similarly, supply chains, and especially in big industries like the energy industry, are influenced by regulatory development (Weijermars, 2012) that many times it is out of the direct control of the companies. Moreover, "in industry the focus has shifted from a pure logistics-oriented view towards the integration of pricing and revenue issues into cross-functional value chain planning models" (Kannegiesser & Gunther, 2011).

Furthermore, supply chains are not stationary systems from the point of view of their structure and behavior; they evolve according to the different strategies and plans of a particular company. An example of this evolving nature is the case of the clean energy transition experienced by the energy industry, where the effects of the commodity markets contract and spot demand can be distinguished. Contract demand is based on agreements between the company and customers that vary on time, with sales quantities and prices being fixed only for limited and defined periods (Kannegiesser & Gunther, 2011). In general, most of the current supply chains "involve numerous, heterogeneous, geographically distributed entities with varying dynamics, uncertainties, and complexity." (Suresh, et al., 2008).

Measuring the complexity of a supply chain helps to manage it. This is important when considering that globalization and its effects are also a relevant factor on the complexity of supply chains, especially on logistics activities (Isik, 2010). Some studies are also focused on understanding the level of integration and information sharing between partners and competitors. In this regard, "evidence shows that more extensive process integration and information sharing have favorable financial performance implications for supply chain partners" (Schloetzer, 2012).

Just the understanding, abstraction or modeling of a supply chain is not enough to successfully manage it. It is also necessary to realize that supply chain management goes together with the "software systems for supporting decisions at the strategic, tactical, and operational planning level" (Kannegiesser & Gunther, 2011). "Analysis and optimization of the SC requires consideration of numerous entities, each with its own dynamics and stochastic, participating in events occurring on various time lines, which make them significantly complex" (Suresh, et al., 2008) and can produce not just incredibly expensive computations but unpractical ones, since they would take years to run even with modern super computers.

Moreover, the complexity of a supply chain is not just significant in terms of the number of its components and interactions, which is known as combinatorial complexity or detailed complexity, it is also material because it is dynamic, and therefore the effects of its delays, for example, can produce missing intuition of even very simple problems in terms of details and combinatory (Sterman, 2000). When combining this challenge with optimization, on-line model adaptation has been proposed as required for appropriate prediction and re-optimization. "In most dynamic real-time optimization schemes, the available measurements are used to update the plant model, with uncertainty being lumped into selected uncertain plant parameters" (Bonvin & Srinivasan, 2013). This can be also incorporated in the modeling through decision rules and flexibility (de Neufville & Scholtes, 2011).

Real world systems are dynamic with several feedback loops and reactions from people operating them. These loops present many delays and nonlinearities that can produce a dynamic complexity even in small simple systems due to many reasons like self-organizing characteristics, adaptive, tightly coupled, among others reasons (Sterman, 2000).

YES, IT IS COMPLEX, ESPECIALLY FOR THE ENERGY INDUSTRY

Supply chains are important and complex in general, but in the energy industry this is more than ever true:

> "The energy value chain is one that is intrinsic to the existence of the world as we know it today. It is an industry that is more than 150 years old and is one of the most complexes in the world today. By the very nature of its product, this business is subject to risks such as geopolitical risk, business risk, financial risk, credit risk, market risk, currency risk – the list continues." (Kavi, 2009)

A global energy firm such as Chevron spans wide segments of the value chain in which it operates, from oil exploration to service stations, but it does not span the entire chain. Approximately "fifty percent of the crude oil it refines comes from other producers, and more than one third of the oil it refines is sold through other retail outlets" (Shank & Govindarajan, 1992). A global energy company is also detached from the consumption of its final product since it is, for example, "not in the auto business at all, the major user of gasoline. More narrowly, a firm such as Maxus Energy is only in the oil exploration and production business. The Limited Stores are big downstream in retail outlets but own no manufacturing facilities." (Shank & Govindarajan, 1992)

An energy supply chain as a whole is an incredibly huge and complex system. When zooming into a particular part of that system, even the small parts are complex. The offshore supply chain, for example, is incredibly complex; the "supply-chain related costs for an integrated oil company in the Gulf of Mexico can be $200 to $400 million annually. Even a modest reduction in these costs through optimization can produce substantial savings" (West & Lafferty, 2008). Understanding the systemic factors and conditions of these specifics parts of the business is a challenge in itself. The upstream oil & gas business is increasingly impacted by federal legislation, while midstream and downstream energy segments are more related to state regulators (Weijermars, 2010). Regulation, in general, has a significant impact on the performance of the mid and downstream oil and gas industries (Weijermars, 2012).

The complexities previously discussed are not just difficult to manage, but moreover to understand in terms of the causes and effects over time. Many studies have explored different aspects of these complexities in the oil and gas industry, for example, the link between industrial clusters and competitiveness couldn't be proven empirically, however, it has been suggested that clusters enhance and enable higher levels of agile practices (Yusuf, et al., 2013). These complexities are also evolving overtime due to changes in the structures of the companies themselves making even more difficult to understand what is going on. Recent evidence of this observable fact are "the mergers between BP-Amoco (1998), Exxon and Mobil (1998), Dow-Union Carbide (1999), and Chevron and Texaco (2000)" (Ross & Droge, 2002)

The intricate distinction of what is internal and external in an energy supply chain is as well a big source of complexity. "A typical refinery [supply chain] comprises oil suppliers, 3rd party logistics providers, shippers, jetty operators, and customers" (Suresh, et al., 2008), in which the refineries and other entities or subsystems can be owned or not by the same company that owns the supply chain. Each of the subsystems is also dependent of material complexity that arises from many aspects like "seasonal requirements, market competition, and geographical demand patterns" (Suresh, et al., 2008). The number of these subsystems is, in real energy supply chains, colossal, but finite, forming an intricate "process network by a set of processes that are interconnected in a finite number of ways" (Bok, et al., 2000). This, in a way, motivates the managers to describe in their models, those details of large, but limited aspects.

ALTERNATIVES METHODS FOR MODELING AND OPTIMIZATION

Selecting the appropriate tools or methodologies to understand and analyze a supply chain is a challenging problem by itself. Several researches have been conducted to tackle this problem, studying it from almost every possible perspective; modeling, analytics, architecture, design, systems aspects, human contributions, dynamic behavior, etc.

Some papers have considered the combined benefits of cost, time, and satisfaction level for customized services to derive a scheduling strategy for a supply chain formulating an approach of a multi-tier model and multiple objectives (Tang, et al., 2013). The effective control over each element of the supply chain has also been considered important: Inventories, facilities and transportation can impact significantly the efficiency of a supply chain, and the optimization of the system can be analyzed using many familiar optimization algorithms such as particle swarm optimization and simulated annealing optimization techniques, with performance and results from some of them more efficient than others. (Ahamed, et al., 2013).

Just the problem of defining the optimal design or configuration of a supply chain is not a discipline itself, but many bodies of knowledge combined together. From the mathematical modeling point of view, hundreds of alternatives of numeric and heuristic tools are available to use with different advantages and limitations (Papalambros & Wilde, 2000). If considering a multi-criterion optimization, different aspects, like maximization of the net present value and minimization of the expected lead time, can be modeled producing a Pareto-optimal curve "that reveals how the optimal net present value… change with different values of the expected lead time" (You & Grossmann, 2008). Other specific methodologies, like the Lagrangian-relaxation-based algorithm has been successfully applied in the literature with particular results and difficulties (Ozsen, et al., 2009). Figure 1 illustrates the different levels of optimization decisions that a manager optimizing a business system has to follow.

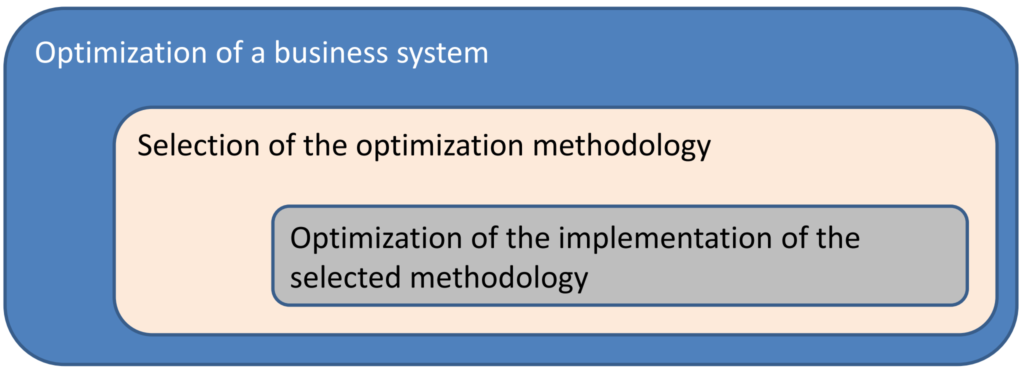

FIGURE 1: OPTIMIZATION LEVELS

The list is endless when talking about methodologies and approaches to optimize the supply chain; however, this is not the only aspect that is relevant in the analysis. The robustness of the supply chain under changing conditions is also a topic of study; in this regard, tools like multiclass queueing networks are an essential tool for modeling and analyzing complex supply chains stability subject to uncertainty (Schönlein, et al., 2013). On the other hand, and focusing on getting better results on expected value rather than having a system working under any condition, flexibility in the engineering design has been proposed instead of robustness to be a more effective strategy to tackle uncertainty (de Neufville & Scholtes, 2011).

A common example in the literature and the industry are the linear optimization models that have been presented to support tactical value chain planning decisions on sales, distribution, production, and procurement, specifically suited for precise applications in value chain network for chemical commodities (Kannegiesser & Gunther, 2011). More detailed examples can be found, like the multi-echelon supply chain networks of a multi-national corporation "combining both the strategic (facility planning) and tactical asset management (production–allocation–distribution) problems into an integrated asset management, capital budgeting, and supply chain redesign model as opposed to solving these problems individually" (Naraharisetti, et al., 2008).

"To optimize, a realistic embedded systems design requires efficient algorithms and tools that can handle multiple objectives simultaneously." (Qiu & Li, 2011). This is true for almost every real problem. Multiple objectives produce a different approach to optimization, which can go from maximizing-minimizing perspective to a trade-off proposition (de Weck, 2012). Many algorithms exist to deal with this challenge that can achieve general and specific purposes, like for example, "to go before reaching a goal of an algorithm capable of dealing with multi-parametric and multi-component (multiple locations) linear programs regardless of the location of the parameters." (Khalilpour & Karimi, 2014)

To make this challenge even more entertaining, and considering that the complexity of a supply chain is not just detail complexity but dynamic, the structure-behavior relationship of the supply chain can be studied from a systems dynamic point of view (Sterman, 2000). This approach could "allows the user to simulate and analyze different policies, configurations, uncertainties, etc." (Suresh, et al., 2008) in a way that is intuitive, even when the dynamic complexity, mentioned before, limits the understanding of the behavior of the system significantly.

OPTIMIZING DESIGN VERSUS IMPROVING PERFORMANCE OF CURRENT DESIGN

Managing a supply chain to obtain better results can be confusing for the decision maker even from the point of view of the objective. It is important to differentiate the strategies to improve a supply chain performance than the ones to optimize the design or configuration. The differences are significant and reside mainly in the set of tools and methodologies to use, but also in the limitations of each approach. Some researchers report that they "have not observed any significant direct relationship between supply chain management programs and practices and total quality management practices. SCM program has been observed to be directly linked with flexible systems practices and not by any other variable." (Siddiqui, et al., 2012). This flexibility focus is also another layer of analysis for the problem, since flexibility represents an alternative for the decision maker, incrementing the design space of available options (de Neufville & Scholtes, 2011).

The flexibility analysis is important since one of the considerations to keep in mind when defining the objective is to identify the ongoing changes that the supply chain is having during the time of improvement or optimization. These changes can be significant in a supply chain of a chemical multinational company, like most of the energy companies. "A facility once built or an existing facility may undergo multiple capacity expansions… Further, disinvestment can also be done in multiples of the same discrete capacity." (Naraharisetti & Karimi, 2010). Research has concluded that "mastering change and uncertainty correlates highest with turnover" (Yusuf, et al., 2012).

This problem of optimizing design versus improving performance opens the door to a whole world of improvement methodologies. References from traditional tools like Lean Manufacturing and Six Sigma can be used to guide this analysis. A pertinent factor to consider in this analysis is the maturity level of the system. A simple example is to understand that it would be unfair to say that a supply chain design or configuration needs to be optimized before the implementation of its design is completed. This is especially true in big and complex supply chains, where the implementation of the system could take years. It is important to differentiate between necessary corrective actions (Ulrich & Eppinger, 2012) and design limitations. Changing (or optimizing) the design or configuration when it is not working as supposed could lead to the wrong strategy, in other words: it is not necessary to fix what is not broken.

Another relevant consideration to have is the fact that a supply chain is a system composed, like most of the systems, by parts with conflicting interests, which can be, at least in theory, be solved by maximizing "the summed enterprise profits of the entire supply chain subject to various network constraints" (Gjerdrum, et al., 2001). These conflicts will produce a trade-off problem from the optimization point of view (de Weck, 2012), but also from the perspective of the human interactions that will require negotiations and leadership skills that are not related at all with the mathematical procedure (Katz, 2004).

In between the definitions of optimization and improvement we can find real time optimization, which is a class of methods that use measurements to bypass the effect of uncertainty on optimal performance and improve in real time the configuration of the system. Several real-time optimization schemes have been studied that "implement optimality not through numerical optimization but rather via the control of appropriate variables" (Francois, et al., 2012). Advanced real-time optimization, control, and estimation strategies are dependent on time-dependent data obtained at predefined sampling times (e.g., sensor measurements and model states)." (Zavala & Anitescu, 2010)

COMPUTATIONAL CHALLENGE

As if the challenge of selecting an appropriate methodology to analyze the improvement or optimization of a supply chain or getting intuition on such a complex system was not enough, the manager of a supply chain must also compute the model, facing another layer of difficulties and limitations for the project. The reality of energy industry is that the decision problems represented in a mathematical form often result in a very large scale optimization problems, with different types of variables, functions, etc. (Applequist, et al., 2000).

Computation time is relevant for big and complex systems with several elements and interactions, but is even more relevant when adding stochastic simulations to those models (de Neufville & Scholtes, 2011). Efforts to improve computational performance have been conducted. "Algorithms [have been] developed for reducing the dimension of large scale uncertain systems" (Russell & Braatz, 1998). Even optimization of the algorithm to optimize the supply chain should be considered, using methodologies that "permits a reduction in the required computation time as a result of the optimal nesting of [for example] the iteration loops" (Montagna & Iribarren, 1988), however, for real complex systems, even this will produce heavy systems to compute.

Computation of an optimization can be a huge problem, since the process of finding the solution (solving the optimization) "may result in a very complicated model which is computationally intractable even for small scale instance" (You & Grossmann, 2008). In the energy industry this is especially true, because we are facing a design problem for a complicated process supply chain network with multi-sourcing (Ozsen, et al., 2009).

Not just the scale of the optimization problem is a challenge from the computational point of view. The challenge for the decision maker also considers the fact that an energy supply chain is, most probably, a multi-objective optimization, and therefore compromise among the different goals should be considered (de Weck, 2012), for example, by Pareto optimality and using a subjective weighted-sum method. “Many alternatives are available to compare objectives, like scaling each of the objectives directly onto the range [0, 1], or to adopt the fuzzy set concept depending on the particularities of each analysis” (Chen, et al., 2003). Pareto optimality has a limitation related to the number of dimensions (or objectives) analyzed (Papalambros & Wilde, 2000).

In order to overcome the challenge derived by the computational cost of optimization, many methodologies have been proposed. One efficient strategy is to separate the problem in a two steps approaching the solution with a screening model first (de Neufville & Scholtes, 2011), or “bi-level decomposition algorithm that involves a relaxed problem and a sub problem for the original supply chain problem.” (Bok, et al., 2000). Meta-models, in general, are “commonly used as fast surrogates for the objective function to facilitate the optimization of simulation models… Empirical results indicate that [this approach] is effective in obtaining optimal solutions” (Quan, et al., 2013).

In order to reduce, or simplify the modeling, reduction strategies commonly takes two general approaches: model order reduction and data-driven model reduction (Biegler, et al., 2014). Regarding this approach, it is critical to keep in mind that “It helps you to ignore details, but you have to be careful to ignore details that are actually important” (Devadas, 2013). In order to develop a reduced model it is important to clarify several issues, like the properties needed for the reduced model-based optimization framework to converge to the optimum of the original system models, or the properties that govern the (re)construction of reduced models in order to balance model accuracy with computational cost during the optimization, or if the reduced model-based optimization can be performed efficiently without frequent recourse to the original models (Biegler, et al., 2014).

Since parameters can never be estimated perfectly and the decision maker “must always decide which to focus on and when to stop” studying it (Sterman, 2000), and the fact that most of energy supply chains exhibit decentralized characteristics, the manager of a complex supply chain could consider relaxation algorithms for optimal decision-making problems for such decentralized systems. “Such problems consist of a collection of interacting sub-systems, each one described by local properties and dynamics, joined together by the need to accomplish a common task which achieves overall optimal performance” (Androulakis & Reklaitis, 1999).

Techniques to reduce the size of the evaluated data by the algorithm have been developed. In dynamic optimization problems, for example, algorithms exist to refine the control grid iteratively "using a wavelet analysis of the previously obtained optimal solution. Additional grid points are only inserted where required and redundant grid points are eliminated" (Assassa & Marquardt, 2014). It has also been demonstrated that "incorporating measured variables that do not provide any additional information about faults degrades monitoring performance." (Ghosh, et al., 2014). In practice, "insisting on obtaining a high degree of accuracy can translate into long sampling times… limiting the application scope of real-time NLO to systems with slow dynamics." (Zavala & Anitescu, 2010).

The modeling strategy of simplifying the model is described by many authors using different terminology: meta-models (Quan, et al., 2013), screening models, (de Neufville & Scholtes, 2011), relaxation (Wikipedia, 2013), and many others. This bi-level decomposition, together with the Benders decomposition to solve very large linear programming problems that have a special block structure (Wikipedia, 2013)"are two major approaches that have been applied to multi-period optimization problems… bi-level decomposition is different from the Benders decomposition in that the master problem is given by a special purpose aggregation of the original problem which generally tends to predict tighter upper bounds." (Bok, et al., 2000)

LIMITATIONS OF MODELING AND OPTIMIZATION

> "Persons pretending to forecast the future shall be considered disorderly under subdivision 3, section 901 of the criminal code and liable to a fine of $250 and/or six months in prison." Section 889, New York State Code of Criminal Procedure (Webster & de Neufville, 2012)

The modeler and decision maker of a supply chain optimization should be aware of several limitations of the methodologies. Most of them come from the fact that the methods are constructed following some assumptions and those assumptions should hold during the analysis. Even well defined projects by very competent professionals have fallen in common mistakes. Studies, like for example, the analysis of the mixed-integer linear program proposed by Kannan et al. (Kannan, et al., 2009) that found inconsistencies in the model (Subramanian, et al., 2012) are common in the academic world. Just to mention some of the limitations and mistakes commonly cited:

- Integration of the information across the different subsystems of a supply chain is not always available for the analysis, for example, the integration of strategic management theory and IT knowledge could provide valuable information in these complex and complementary value chains (Drnevich & Croson, 2013), but is not commonly available for most of the companies. Currently, multi-scale process optimization still needs effective problem formulation and modeling environments, since "they cannot include detailed interactions with material design, complex fluid flow and transport effects with multiphase interactions." (Biegler, et al., 2014)

- Dynamic effects, like the bullwhip effect, are often ignored by the analysis. Some researchers have study quantitatively this effect on a discrete-event simulation model of a supply chain, (Bottani & Montanari, 2010). The bullwhip effect (or whiplash effect) is an observed phenomenon in forecast-driven distribution channels that "refers to a trend of larger and larger swings in inventory in response to changes in demand, as one looks at firms further back in the supply chain for a product" (Forrester, 1961). The design, retrofitting, expansion/shut-down, or the planning of the operation "to meet ever-changing market conditions can all be posed as a large scale dynamic decision problem." (Applequist, et al., 2000)

- Aggregation challenges have demonstrated to be a subtle limitation to simplifying optimization systems. "Several physical aggregation levels: globally dispersed entities, production lines within a manufacturing site, and even individual equipment items" (Applequist, et al., 2000) need to be considering when modeling

- Traditional optimization methodologies "provide logical guide lines or optimal solutions, [but] they are inadequate to deal with the diversity and heterogeneity of numerous configuration constraints" (Jiao, et al., 2009), which is an important consideration when facing constrained systems like the supply chain of an energy company.

- Unavoidable uncertainty (de Neufville & Scholtes, 2011), that even with "sophisticated methods such as time series… to improve the forecasting accuracy, uncertainties in demand are unavoidable due to ever changing market conditions" (You & Grossmann, 2008). Uncertainty can be considered and managed by means of the concept of financial risk, which is defined as the probability of not meeting a certain profit aspiration level (Guillén, et al., 2005), but it can also be defined from the perspective of other effects on the business. "The best remedial [flexible] actions can be identified and their adequacy determined if the dynamics of the supply chain are modeled and simulated." (Suresh, et al., 2008). "The high degree of uncertainty is due to the fact that [the] key factors [of a supply chain], all have significant stochastic components." (Applequist, et al., 2000)

The effects of the uncertainties in the product of a simulation and optimizations require further attention. Optimization with uncertainties on right-hand-side of the constraints has been addressed successfully in recent papers, but, "very little work exists on the same with uncertainties on the left-hand-side of the constraints or in the coefficients of the constraint matrix." (Khalilpour & Karimi, 2014). This could be an important cause of problems due to Jensen's Inequality.

THE JENSEN'S INEQUALITY

The limitations and considerations for the problem of optimizing or improving a supply chain are numerous. This thesis research is focused in one particular aspect: the effect that uncertainty creates on the dynamic simulation of a supply chain by the effect of the Jensen's inequality.

The Jensen's inequality, named after the Danish mathematician Johan Jensen, in its general expression is:

$$E(f(x)) \neq f(E(x))$$

EQUATION 1: JENSEN'S INEQUALITY GENERAL EXPRESSION

Equation 1 is valid for concave or convex functions. Specifically:

For convex f(x) functions:

$$E(f(x)) > f(E(x))$$

EQUATION 2: JENSEN'S INEQUALITY FOR CONVEX FUNCTIONS

For concave f(x) functions:

$$E(f(x)) < f(E(x))$$

EQUATION 3: JENSEN'S INEQUALITY FOR CONCAVE FUNCTIONS

The Jensen's Inequality produces the wrong results on non-linear models, and for this is reason is also called "Flaw" of Averages to contrast with the phrase referring to a "law" of averages. "The Flaw of Averages is a significant source of loss of potential value in the development of engineering systems in general" (de Neufville & Scholtes, 2011) and certainly in the specific case of a supply chain optimization, which is non-linear for almost every case. In other words: "Average of all the possible outcomes associated with uncertain parameters, generally does not equal the value obtained from using the average value of the parameters" (de Neufville, 2012). Non-linearity is a relevant aspect of most dynamic simulations (Sterman, 2000) and therefore the study of the implications of Jensen's Inequality, or the "Flaw of Averages" could be substantial for the decision makers of a supply chain.

A straightforward example of this inequality can be observed by the following function:

$$f(x) = x^2 - 3$$

EQUATION 4: EXAMPLE OF NONLINEAR FUNCTION

$$x = \{\, 2 \,, -5 \,, 0.4 \,, 6 \,\}$$

$$f\big(E(x)\big) = \left(\frac{2 - 5 + 0.4 + 6}{4}\right)^2 - 3 = -2.28$$

$$E\big(f(x)\big) = \frac{(2^2 - 3) + (-5^2 - 3) + (0.4^2 - 3) + (6^2 - 3)}{4} = 13.29$$

$$E\big(f(x)\big) = 13.29$$

$$f\big(E(x)\big) = -2.28$$

$$13.29 > -2.28$$

This is true for a convex function, like the one presented in Equation 4 and plotted in Figure 2 below.

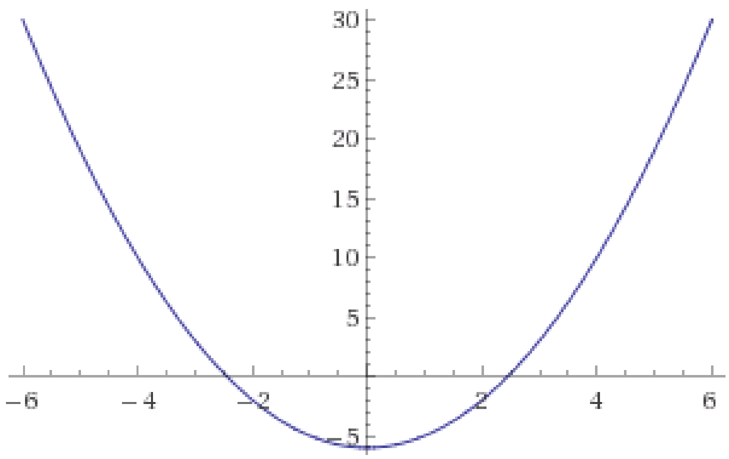

FIGURE 2: X^2-3 PLOT –EXAMPLE OF A CONVEX FUNCTION- (WOLFRAM ALPHA LLC, 2013)

Jensen's inequality holds under certain conditions and with certain optimality. (Guessab & Schmeisser, 2013), and it is materially important for any supply chain model, but particularly for complex, highly nonlinear systems like the supply chain of an energy company. The effect of this inequality has been tested in some financial context, proving that is statistically significant economically large, concluding that "Jensen's inequality, applied to finance, cannot be dismissed as insignificant, or as a theoretical and superfluous exercise in finance as some have advocated." (Azar, 2008)

SYSTEMS-THINKING

It is necessary to have a system-thinking perspective of the problem of managing a supply chain. The supply chain of an energy company is complex in both a detailed way and in a dynamic way (Sterman, 2000). The detailed complexity refers to the fact that "many aspects are important—for example, product life cycle, demand predictability, product variety, and market standards for lead times and service" (Fisher, 1997). This is not only important from the point of view of the structural architecture of the supply chain, but in addition because "some of these entities (such as the refinery units) operate continuously while others embody discrete events such as arrival of a VLCC, delivery of products, etc" (Suresh, et al., 2008).

A system-thinking perspective will also require considering the bigger picture in which the particular supply chain exists. Aspects like sustainability are increasingly complex with developing pressures and "ambiguous challenges that many current environmental management techniques cannot adequately address" (Matos & Hall, 2007). Still, the system-thinking approach will require as well to include detailed aspects of the systems, like "the internal low-level decisions (local scheduling, supervisory control and diagnosis, incidence handling) and the implications of incorporating these decisions for the dynamics of the entire supply chain (production switching between plants, dynamic product portfolios) have not yet been studied." (Puigjaner & Laınez, 2008). A systems-thinking management approach to a supply chain optimization should take in consideration the overlapping decision spaces and thus agreement among the different decision makers has to be achieved within the boundaries of the overlapped regions (Androulakis & Reklaitis, 1999).

The internal portion of a supply chain of an energy company require a systems-thinking approach, especially when considering, for example, that a model can include "external supply chain entities such as suppliers and customers, refinery functional departments such as procurement and sales, refinery units such as pipelines, crude tanks, CDUs, reformers, and crackers, and refinery economics." (Suresh, et al., 2008). The challenge for a high level decision maker of a supply chain is that, particularly in a major energy company, they don't have centralized control of all materials and information flows.

"Clearly the reality is that the entities operate in a decentralized and asynchronous manner. Entities (plants, suppliers, customers and shippers) have mainly their own but also some shared decision variables…Thus management of a supply chain really involves the coordination of the activities of semi-autonomous entities with overlapping decision spaces" (Applequist, et al., 2000). "The asynchrony of operation is due to the fact that these systems are essentially dynamic entities whose optimal policy has to be implemented continuously as soon as it has been identified." (Androulakis & Reklaitis, 1999)

In addition, the system shouldn't be only evaluated under regular and/or ideal conditions. The analysis should definitely consider special cases, like the potential impacts of emergencies or disruptive events like "man-made and natural disasters on chemical plants, complexes, and supply chains… to estimate the scope and duration of disruptive-event impacts, and overall system resilience" (Ehlen, et al., 2014). In this context, "the allocation of resources to properly train employees in disaster/disruption prevention activities" (Ojha, et al., 2013) is a critical aspect to consider that can be really challenging to model.

It was previously mentioned that, when facing computational performance challenges, decision makers try to simplify their models in order to reduce complexity and reduce CPU time. This is, in practice, limited by the fact that in many cases it is just not possible to simplify without losing meaning of the model for managerial purposes. Regarding this approach, it is important to remember the advice of "Everything Should Be Made as Simple as Possible, But Not Simpler" (attributed to Einstein and others).

In systems architecture, this is referred as the art of defining the appropriate level of complexity, which can be done by "scoping, aggregation, partitioning, and certification" (Maier & Rechtin, 2002). This means that, in a way, complexity is unavoidable in real problems, since "complexity arises in a system as more is asked of it (performance, functionality, robustness, etc.) and complexity manifests itself as the interfaces between elements or modules are defined" (Cameron & Crawley, 2012).

LIMITED FLEXIBILITY

So far, the discussion has only considered the complexity of modeling in general and the specific implications of some aspects like the uncertainty and computation time. On top of that, it is necessary to think about the systems as processes managed by people taking decisions and operating in an intelligent way to adapt to changing conditions (de Neufville & Scholtes, 2011). This is called adaptability or flexibility of systems, which it has been proved material for supply chains management since "supply chain portfolio flexibility is an important determinant for small-to-medium-sized firm satisfaction with supply chain portfolio performance." (Tokman, et al., 2013).

The dynamic capacity of the system to adapt its configurations to changing conditions is another layer of complexity for the system, moreover, is a factor that will certainly change the results of the simulations and optimization models.

PLUG & PLAY MODEL

INTRODUCTION TO THE PLUG&PLAY MODEL

This section describes a model used to explore the effects of considering uncertainty in the inputs of the model to evaluate the hypothesis of this research, which is that not considering the variability of the inputs of the model will produce the wrong result.

In order to achieve the objective of this research, different configuration of the inputs of a dynamic simulation-optimization will be tested to evaluate the effect of singular definitions of the inputs on the results and dynamics of the system. To focus on efficiently achieving this objective, the research methodology plugged & played an already developed model. This strategy stays away from getting into the discussion of the validity of the model itself. If the reader is interested in analyzing the internal mechanics of the model, it is suggested to study the original paper to find the information and academic discussion about it.

The plug & play model was taken from the research presented in 2008 in the Computers and Chemical Engineering Journal, volume 32 in the paper titled "Decision support for integrated refinery supply chains Part 1: Dynamic simulation". The authors of this paper are Suresh S. Pitty, Wenkai Li, Arief Adhitya, Rajagopalan Srinivasan and I.A. Karimi, at that time from the Department of Chemical and Biomolecular Engineering of the National University of Singapore and the Institute of Chemical and Engineering Sciences.

The model called Integrated Refinery In-Silico, or IRIS (Suresh, et al., 2008), was originally implemented for Matlab R2007a requiring the Matlab toolboxes Simulink and Signal Processing Blockset (Figure 3: IRIS block diagram). IRIS for Matlab was shared for this thesis research by Professors Rajagopalan Srinivasan [raj@iitgn.ac.in] and Arief Adhitya [arief_adhitya@ices.a-star.edu.sg] on Wednesday, September 25, 2013 8:33 AM via email. Together with the settings and results from the original paper (Appendix 1: IRIS settings and results from original paper).

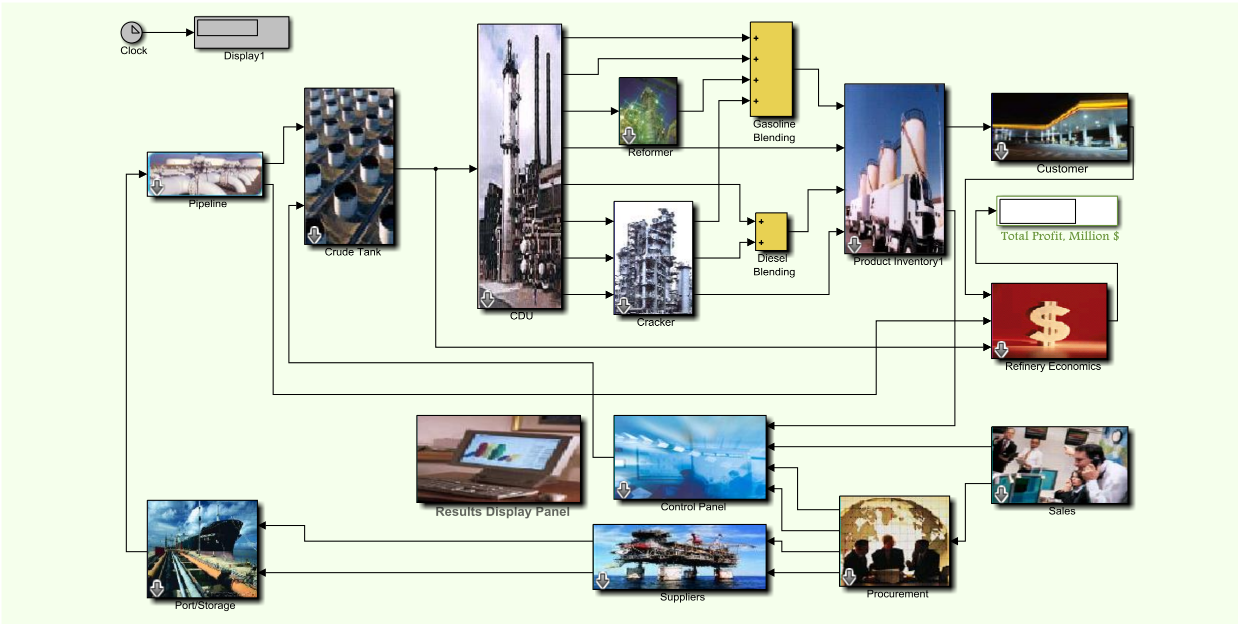

FIGURE 3: IRIS BLOCK DIAGRAM (SURESH, ET AL., 2008)

IRIS, as described by its authors in the original paper:

> "… is a dynamic model of an integrated refinery supply chain. The model explicitly considers the various supply chain activities such as crude oil supply and transportation, along with intra-refinery supply chain activities such as procurement planning, scheduling, and operations management. Discrete supply chain activities are integrated along with continuous production through bridging procurement, production, and demand management activities. Stochastic variations in transportation, yields, prices, and operational problems are considered in the proposed model. The economics of the refinery supply chain includes consideration of different crude slates, product prices, operation costs, transportation, etc… IRIS allows the user the flexibility to modify not only parameters, but also replace different policies and decision-making algorithms in a plug-and-play manner. It thus allows the user to simulate and analyze different policies, configurations, uncertainties, etc., through an easy-to-use graphical interface." (Suresh, et al., 2008)

In order to run the model on Matlab R2013a, an upgrade check was run on the model and some blocks were updated using the slupdate() Matlab function for automatic update of the blocks. Also, some blocks were considering continuous signals instead of the required discrete signals for the model. The blocks were set as discrete at a sample time of 0.01, which is the same sample time from the fundamental signal defined on the solver of the original model (Figure 4). After these adjustments, the model was ready to run on Matlab R2013a and all the settings and results from the original paper were replicated to ensure calibration.

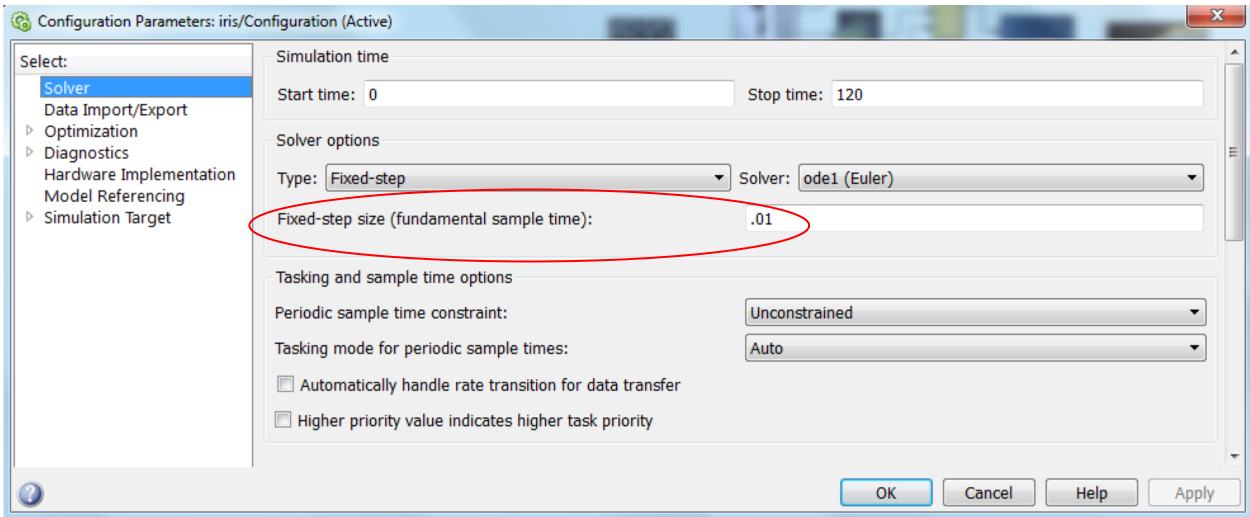

FIGURE 4: FUNDAMENTAL SAMPLE TIME

The simplicity that this model offers allows the perfect platform to explore the results of considering uncertainty into de inputs and assumptions of the model. As described by the authors, "IRIS can serve as an ideal test bed for [supply chain] analysis both for industry and academia" (Suresh, et al., 2008).

RELEVANCE OF THIS PLUG & PLAY MODEL FOR PRACTICAL VALUE CHAIN OPTIMIZATION

The IRIS model used, in a plug & play format, in this research, has been developed to support "decision making in a petroleum refinery should consider the overall supply chain performance" (Suresh, et al., 2008). The possibility to expand the analysis to get closer to crude to customer (C2C) perspective is useful for the energy industry because it allows bringing a systems-thinking perspective. An element like margin variability (which corresponds to crude price and product price variability) has high impact per the 80/20 rule in practice.

The original configuration proposed in the paper to present IRIS is valuable since it allows a generic assessment of the system from an academic point of view. Starting from this exercise, specific companies have the option to customize the configuration to it better to their unique realities. The dynamic model (IRIS) is useful as a tool for integrated decision support of a supply chain "reconfiguration studies such as change of demands, cycle times, or tank capacities" (Suresh, et al., 2008). This is possible through adjusting simple model parameters.

Moreover, a model like IRIS brings a broad systems approach with the possibility to explore different policies, such as "for demand forecasting, product pricing, production [plans], parcel sequencing, and others" (Suresh, et al., 2008). A model like IRIS can allow trying different scenarios and configurations of multiple systemic aspects of a supply chain of an energy company simultaneously. This is useful for analytic challenges like system optimization, and multi-objective improvement. Research has shown that optimization can be achieved using a diverse set of algorithms and simulations (Koo, et al., 2008).

A plug & play model like IRIS could be also useful for industrial training (Koo, et al., 2008). This is valuable since it allows having a systemic perspective of a supply chain system with the capability of being configured in a very straightforward way.

LIMITATIONS OF THE PLUG & PLAY MODEL

The limitations of the use of a plug & play model like IRIS will depend on the particular application that the user defines. For example, if the model is used for training purposes, the limitations will be related to the required competences to implement the model in software or, use already implemented software like Matlab.

Other limitations can be related to the lack of specificity or scope of the model for precise purposes, like optimizing the distribution of products or taking specific decision of a particular aspect of the system described by the model. The model, however, has the flexibility to add or expand some particular aspects on the interest of the use.

RUNNING THE MODEL

METHODOLOGY OVERVIEW

In order to evaluate the hypothesis of this research the following steps are followed:

1. A pre-developed –plug & play model called IRIS is obtained, implemented and tested. This is explained in the chapter "**Plug & Play Model**".
2. The original configuration of the model is deterministic emulating uncertainty, in other words, it doesn't rely on repeated random sampling to obtain numerical results, but incorporates a variation percentage and a variation seed that is fixed. In practice, this means that if the model runs several times the numerical results are going to be the same (Suresh, et al., 2008).
3. The original configuration is modified to eliminate the variation percentage. This is achieved by practically setting the variable to zero (0) and running the simulation again one more time (Case 0: Original setting without uncertainty (no variation %))
4. Uncertainty is incorporated in targeted ways. Different modules of the system are defined as uncertain with some realistic variation and the simulation is stochastically iterated 1,000 times. (Case 1: Focalized Uncertainty)
5. Stochasticity is incorporated in fourteen of the variables in a standard way (a random value with normal distribution). The dynamic simulation is stochastically iterated 1,000 times. (Case 2: standard uncertainty)
6. The distributions and mode o behaviors of the results of the simulations are compared statistically and graphically to analyze the effect of uncertainty on the dynamic simulation.

BASE CASE: ORIGINAL CONFIGURATION FROM THE ORIGINAL PAPER

The base case is the one presented in 2008 in the Computers and Chemical Engineering Journal, volume 32 in the paper titled "Decision support for integrated refinery supply chains Part 1: Dynamic simulation". In order to run different configurations and compared them with known base cases, the settings from the configuration in the original paper were considered. If the reader is interested in understanding the details of the configuration of this plug & play model, it is recommended to study the original paper (Suresh, et al., 2008).

The simulation from the original paper produced results on several metrics that are described in Figure 5. In order to compare different cases, this research will focus on the Total Profit metric:

Metric	Base case
Average crude inventory (kbbl)	328.0
	397.7
	619.7
	569.4
	318.0
Average product inventory (kbbl)	526.2
	301.3
	882.1
	267.2
Average CDU throughput (kbbl)	145.5
Product revenue (million $)	1065.0
Crude procurement cost (million $)	942.6
Crude inventory cost (million $)	13.4
Product inventory cost (million $)	11.9
Operating cost (million $)	35.1
Product deficit penalty (million $)	0.0
Demurrage cost (million $)	0.3
Profit (million $)	62.0

FIGURE 5: FROM THE TABLE "COMPARISON OF KPIS FROM THE DIFFERENT CASES" FROM THE ORIGINAL PAPER (SURESH, ET AL., 2008)

Other metrics are also useful for the analysis of this research. The crude inventory, for example, allows observing the internal dynamics of the system. Graphical results of the base case simulation including for crude inventory profile is shown in Figure 6.

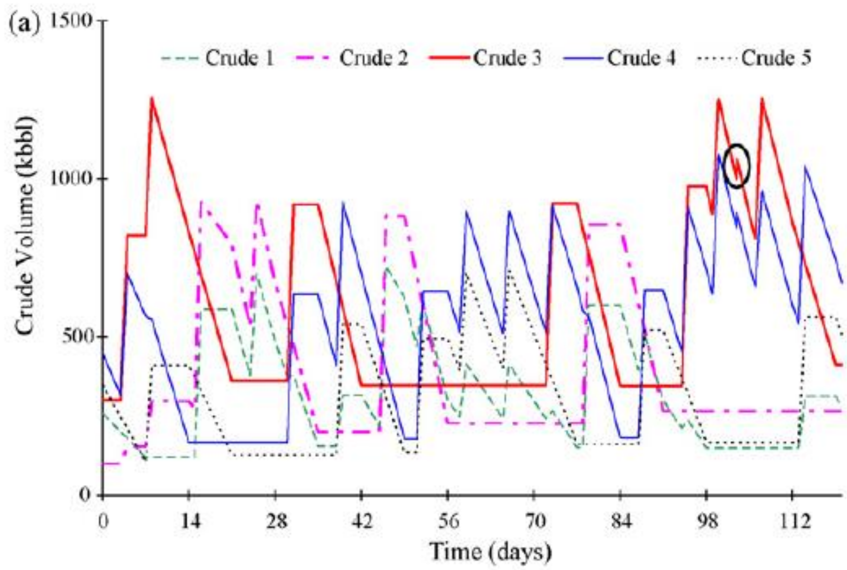

FIGURE 6: BASE CASE: CRUDE INVENTORY (SURESH, ET AL., 2008)

The total profit over time for the original configurations is represented in Figure 7 using the statistical software Minitab 16. This graph shows the effects of the variable Production cycle time (days) originally set as 7 days (represented through the moving averages of the time steps of 0.01 days).

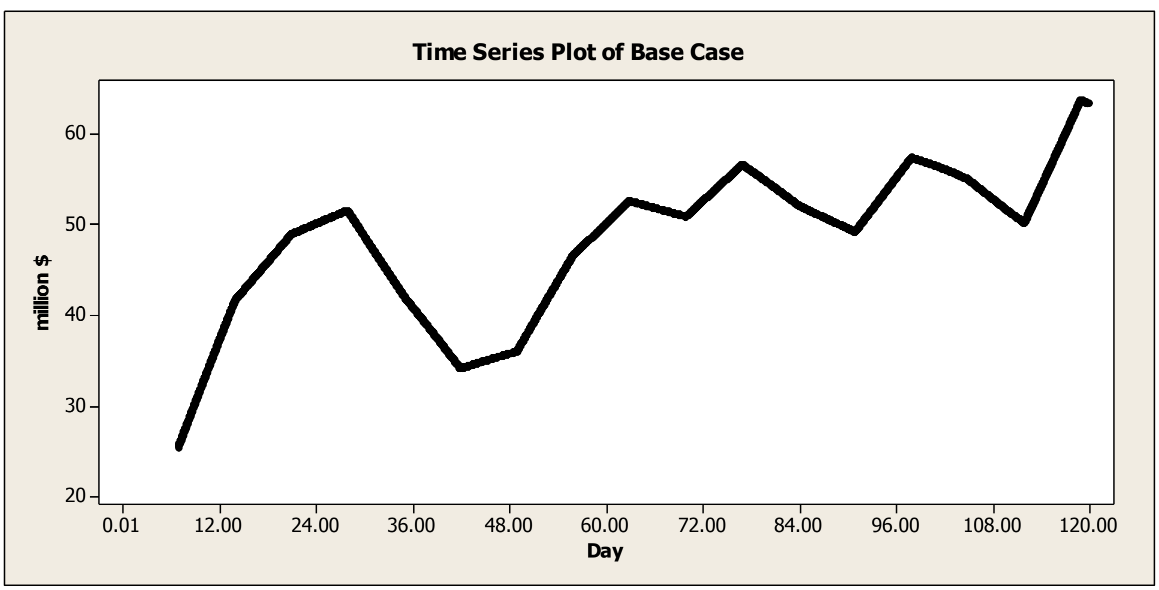

FIGURE 7: TOTAL PROFIT IN MILLIONS (ORIGINAL SIMULATION)

The crude inventory (Figure 6) was replicated with time horizon 120 days in Figure 8.

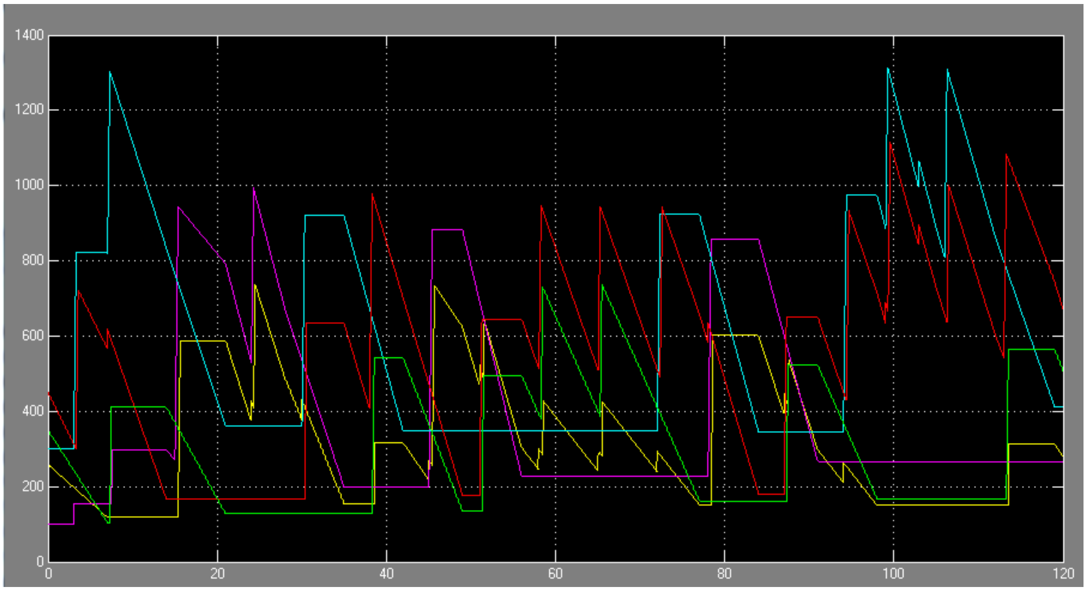

FIGURE 8: CRUDE INVENTORY WITH 60 DAYS TIME HORIZON (REPLICATING FIGURE 6)

CASE 0: ORIGINAL SETTING WITHOUT UNCERTAINTY (NO VARIATION %)

IRIS emulates stochastic variables using a variation percentage and a variation seed concept that is implemented in a single simulation. The original research used random number generators in Simulink to represent the uncertainty. This is achieved by specifying a random number seed, which determines the sequence of random numbers generated. This means that if the simulation is repeated using the same seeds, IRIS should get exactly the same sequence of random numbers and thus the same results, therefore, it is not necessary to run multiple times the model with this definition of uncertainty.

According to the authors of the original paper, it is possible to set the variation as a very small number, e.g. 0.00000000001, which is essentially 0, and the variation seeds can be any arbitrary numbers; this will represent different random number sequences. Following this approach, for the purpose of this research fourteen of the variables were redefined without variation, compared with Table 2, according to the following Table 1:

TABLE 1: VARIABLES VALUES REDEFINED TO ELIMINATE UNCERTAINTY

Subsystem	Variable	Definition in m.file
Refinery Economics	Product price variation percentage (%)	proprivarper=0.00000000001;
Sales	Magnitude of demand increase	magdeminc=0.00000000001;
	Demand variance percentage (%)	demvarper=0.00000000001;
Suppliers	Crude amount variation percentage (%)	cruamovarper=0.00000000001;
	Crude price variation percentage (%)	cruprivarper=0.00000000001;
	Disruption occurrence seed	disoccsee=0.00000000001;
	Disruption magnitude seed	dismagsee=0.00000000001;
Reformer	Reformer yield variation (%)	refyievar=0.00000000001;
Cracker	Cracker yield variation (%)	crayievar=0.00000000001;
CDU	CDU yield variation (%)	CDUyievar=0.00000000001;

The results on the metric of crude inventory of simulating IRIS for a time horizon of 120 days with the configuration described in Table 1 are the following:

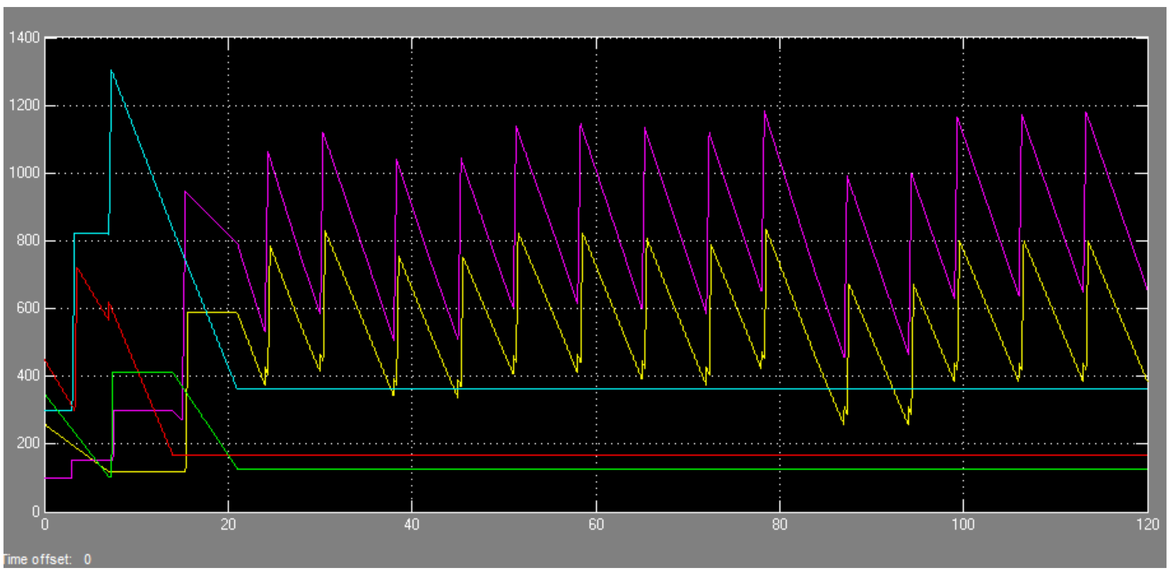

FIGURE 9: CRUDE INVENTORY FOR SIMULATION WITHOUT VARIATION %

Figure 9 shows a very different dynamic behavior than the one observed with the original configuration considering the variation % (Figure 8). This behavior is dynamically stable after a certain period (approximately day 25), with two crudes experiencing oscillation and three having a constant behavior over time. Oscillations are generally produced by feedback loops that are negative and processes that overshoot the goals (Sterman, 2000). It is quite possible that this oscillation is produced by the fact that the demand has cycles (7 days) in this case and therefore the goal is not achieved in a constant way.

The total profit at the end of the simulation without uncertainty is $60.47 million, which is arguably not very different from the $61.96 million produced by the simulation with variation %. This difference is not yet produced by the Jensen's inequality previously described (Equation 1), because both simulation consider deterministic inputs. The original paper emulates the stochasticity of the variables defining a variation % and a variation seed that is deterministic, in other words, it is an expected value of variation.

Even when not considering the numerical difference between the case with variation and without variation, it is important to analyze the mode behavior of the dynamic system. Figure 10 shows the comparison between the behavior, in a time series plot, of the case with variation % and without it. It is clear from the image that, even when the final result is close, the behavior of the system over time is completely different, in system dynamics words; the fundament mode of dynamic behavior (Sterman, 2000) is different.

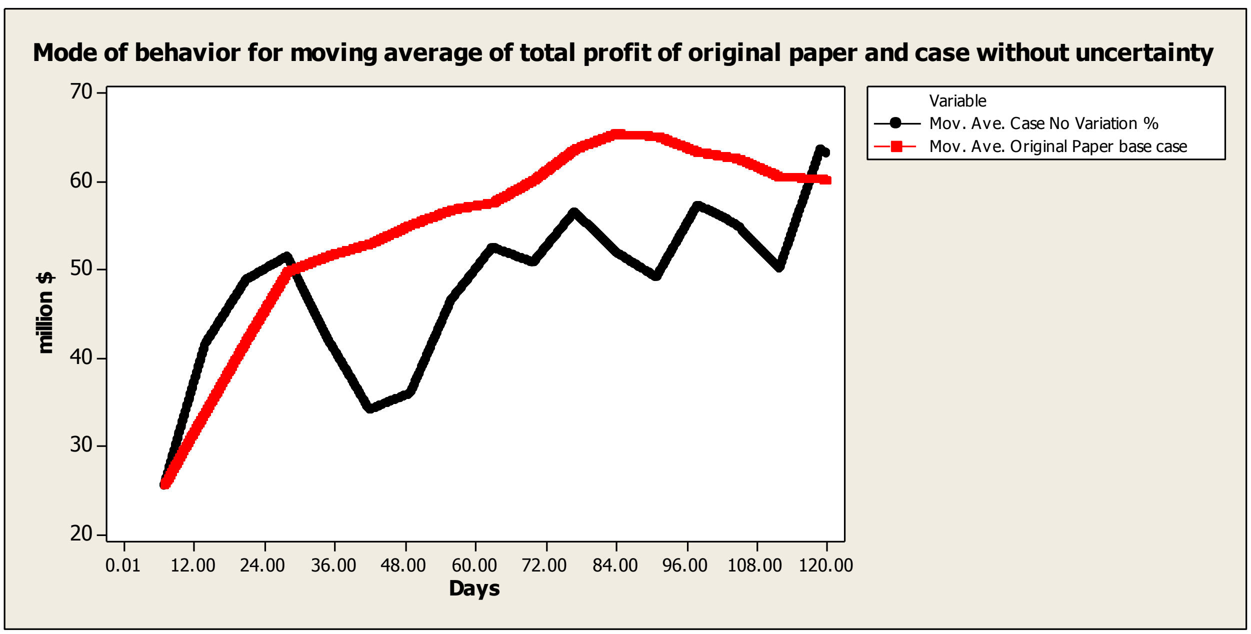

FIGURE 10: COMPARISON BETWEEN BASE CASE AND CASE 0 USING TIME SERIES OF THE MOVING AVERAGES

The similitude between the total profit over the 120 days with variation % and without variation performance will produce very different mode of behaviors. Both behaviors will require very different management styles, and since the performance of the systems should be relevant information for the decision makers (Katz, 2004), this difference is material.

CASE 1: FOCALIZED UNCERTAINTY

Case 1 analyzes uncertainty from specific, focused perspective, showing how, depending on where the uncertainty is defined, it will have a different impact on the final result and mode of behavior of the simulation.

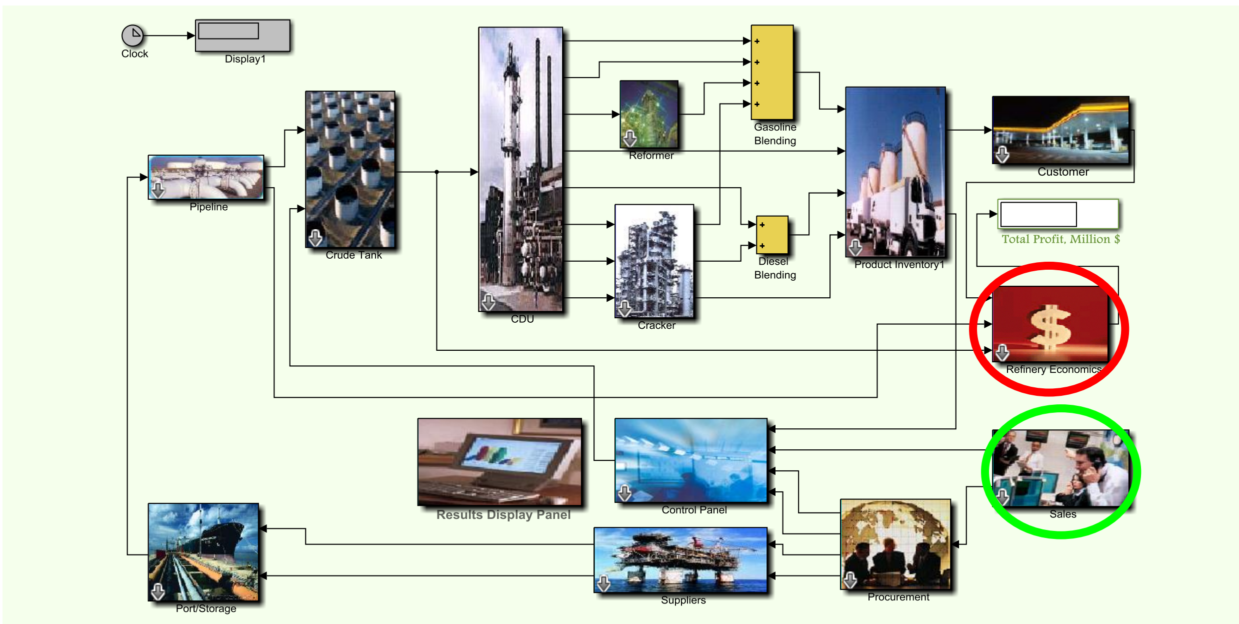

FIGURE 11: MODULES WITH STOCHASTIC VARIATION FOR CASE 1A AND 1B

Two modules of IRIS were selected for their position in the structure of the system. The module "Sales" is at the beginning of the structure of the system, and therefore it is expected that its variation will have an impact on the internal feedback loops. On the other hand, the "Refinery Economics" module is at the end of the system and therefore its variation wouldn't have much impact on the internal loops of the system.

CASE 1A: UNCERTAINTY OF REFINERY ECONOMICS

This variation is incorporated in the "Refinery Economics" module, marked by a red circle in Figure 11. Based on real prices of crude from NYMEX, U. S. Energy Information Administration for 120 days, the variation percentage of a crude oil can be 3.8% in a period of 120 days (New York

Mercantile Exchange (NYMEX), 2013). This simple calculation and the code to simulate are described in Appendix 6: Defining and Simulating specific variability. According to that information, the percentage of variation for a period of 120 days is defined as 3.8% in a normal distribution with 20% of variation.

Running 1,000 simulations[1] the following distribution of results:

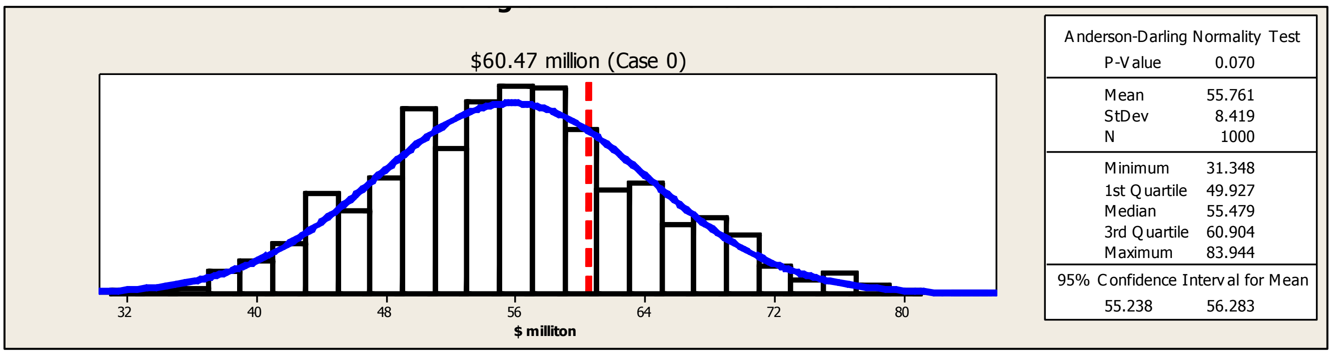

FIGURE 12: HISTOGRAM AND STATISTICS FOR CASE 1A

Figure 12 shows an expected performance of the process after the 120 days of simulation of $55.8 million total profit. This is compared with the Case 0 (no variation) of $60.5 million total profit after day 120. This is confirmed with a one-sample t-test that rejects the null hypothesis so they are not equal.

One-Sample T: Total Profit for Case1a versus Case 0

```
Test of mu = 60.47 vs not = 60.47

Variable           N    Mean  StDev  SE Mean        95% CI             T      P
t(120)_Case1a   1000  55.761  8.419    0.266  (55.238, 56.283)  -17.69  0.000
```

Figure 13 is clearly a concave function, and therefore it is a case of Jensen's Inequality as described in Equation 3:

[1]It took 22,140.2 seconds in Matlab R2013a with a computer with an Intel Core i5-3320M CPU@2.60GHz with 8.00GB of installed RAM with 2 cores (parallel computation).

EQUATION 5: CONCAVE FUNCTION SOWING JENSEN'S INEQUALITY

$$E\big(f(x)\big) < f\big(E(x)\big)$$

$$E\big(f(x)\big) = 55.71$$

$$f\big(E(x)\big) = 60.47$$

$$55.71 < 60.47$$

A time series plot is useful in order to understand the mode of behavior of the dynamic simulation. Figure 13 represents in different colors different statistics of the simulation, from the minimum (or worst case) to the maximum (or best case). These stochastic scenarios are interesting from a managerial perspective since they modify the results of the process, but not the mode of behavior of goal seeking pattern or growth and decline.

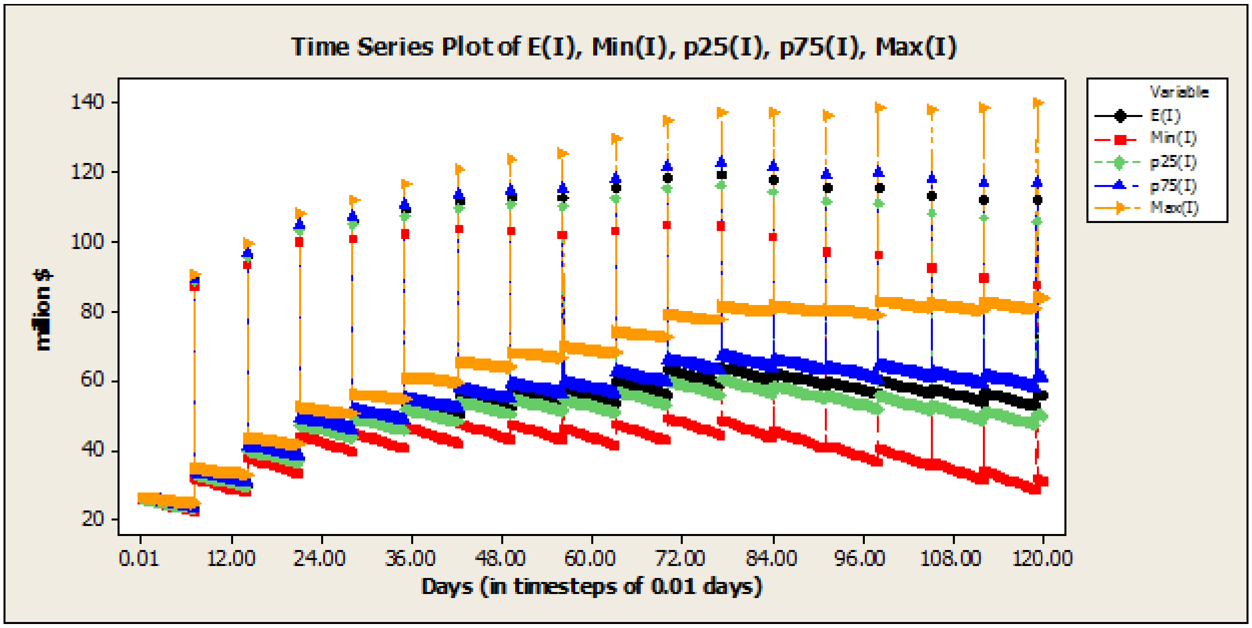

FIGURE 13: TIME SERIES PLOT FOR DIFFERENT STATISTICS FOR CASE 1A

This effect can have multiple causes. Some of them could be:

- The behavior of the dynamic simulation depends on the structure of the model dynamically simulated (Sterman, 2000). The way in which IRIS internal structure defines the dynamics of the subsystem Refinery Economics is important since this module is connected to the output and it almost doesn't impact the internal feedback loops of the operation of the supply chain.
- The economics of the refinery can be considered contextual factors, not necessarily controllable by the management. In most cases, those contextual factors should be considered and planned strategically to not affect the mode of behavior. In other words, the system should be robust enough to support external uncertainty without modifying its behavior.
- The influence of the uncertainty is not material enough to modify the dominance of different internal feedback loops of the system. This means that it could be possible with other values of uncertainty to observe tipping points were the behavior changes

As in Case1a, this case starts from the configuration of no variation % proposed in Case 0. Here, the only change is made in the subsystem called "Sales" in IRIS which is marked as a green circle in Figure 11, adding uncertainty in the magnitude of demand increase and demand variance. This change was made on a base of parallel computation (proposed in Equation 9) as described in Appendix 6: Defining and Simulating specific variability.

In this case, the results of total profits (million $) after 120 days of simulation[2]:

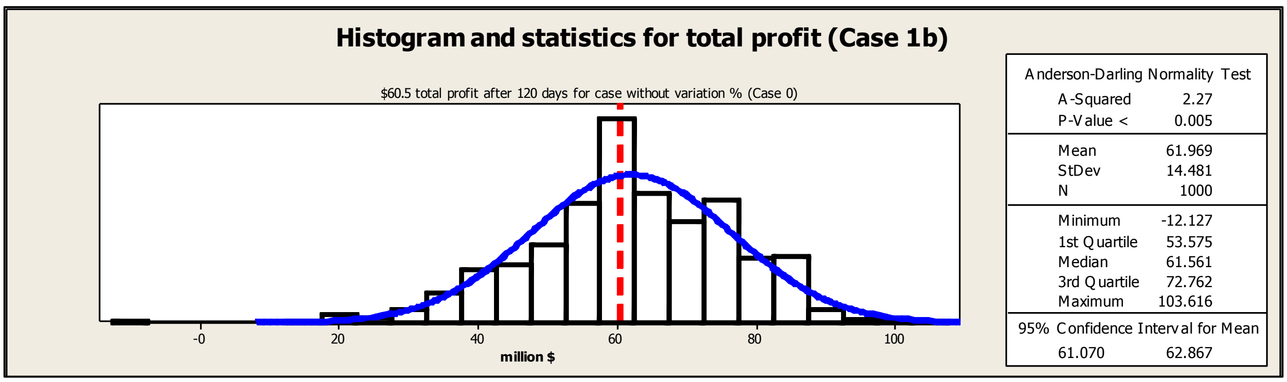

FIGURE 14: HISTOGRAM AND STATISTICS OF TOTAL PROFIT FOR CASE1B

This case gets a much closer value of total profit than the Case 0 expected result ($60.5 million), but it is still statistically not equal. This is probed with the following one-sample t-test:

One-Sample T: Total profit for Case 1b versus Total Profit for Case 0

```
Test of mu = 60.47 vs not = 60.47
Variable          N     Mean   StDev  SE Mean        95% CI            T      P
t(120)_Case1b  1000  61.969  14.481    0.458  (61.070, 62.867)  3.27  0.001
```

Two other comparisons of these results with the results produced by Case1a and Case1b are, first, that in Case 1b the minimum value of total profits (worst case) is negative ($-12.1 million), while in Case 1a the minimum value is just $31.3 million (still positive). Second, when comparing the distribution of the results of both Case 1a and Case 1b is that Case 1b has a bigger range of values (Figure 15).

[2] With a computational time of 22,195.5 seconds

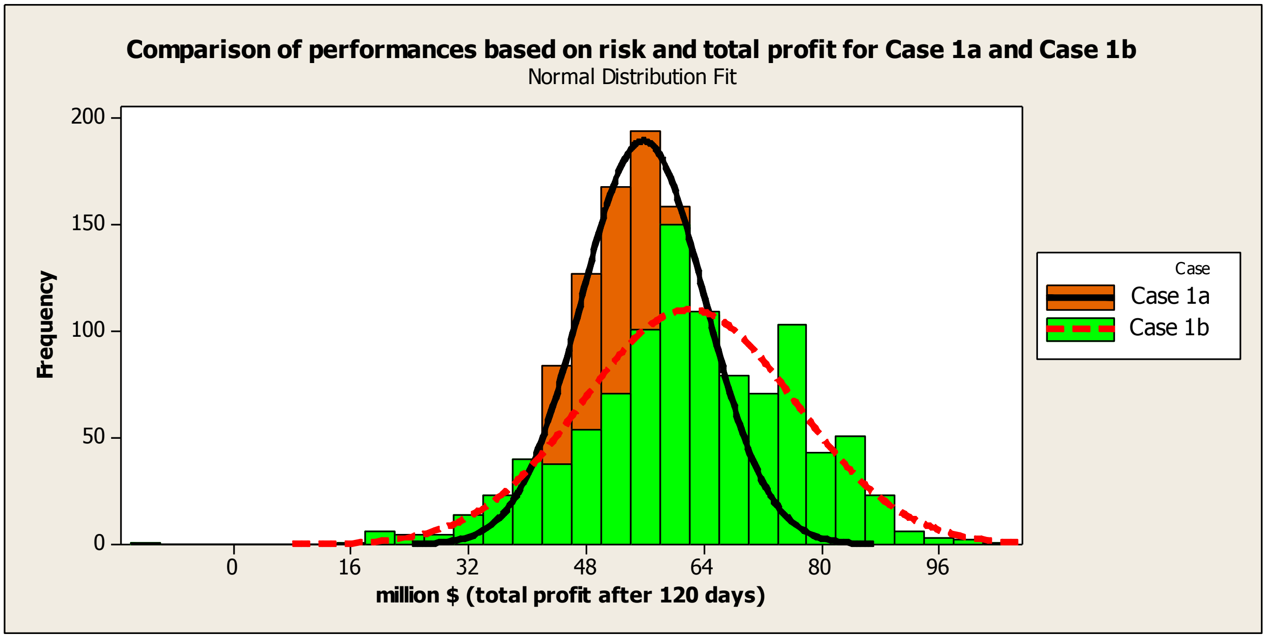

FIGURE 15: COMPARISON OF DISTRIBUTIONS FOR CASE1A AND CASE1B

The different distributions of Case 1a and Case 1b are relevant for managers because variability could be perceived as risks for the business. In other words, an equal variation % in the subsystem Sales represents a bigger risk than an equal variation % in the subsystem Refinery Economics. These results are statistically not equal as described in Appendix 7: Comparison between results of Case 1a and Case 1b.

Also relevant for the manager's point of view is the behavior of the dynamic system. Figure 16 shows a comparison of the mode behavior of both the expected value of Total Profit for Case 1a and Case 1b. The oscillation observed in Case 1b is not observed in Case 1a. The behavior of dynamic simulations depend on the structure of the dynamic model, and also of the dominance of the different internal feedback loops (Sterman, 2000). If the values, because of Jensen's Inequality, change with and without stochastic uncertainty, then the dominance of the different internal feedback loops will also change, producing a different mode of behavior.

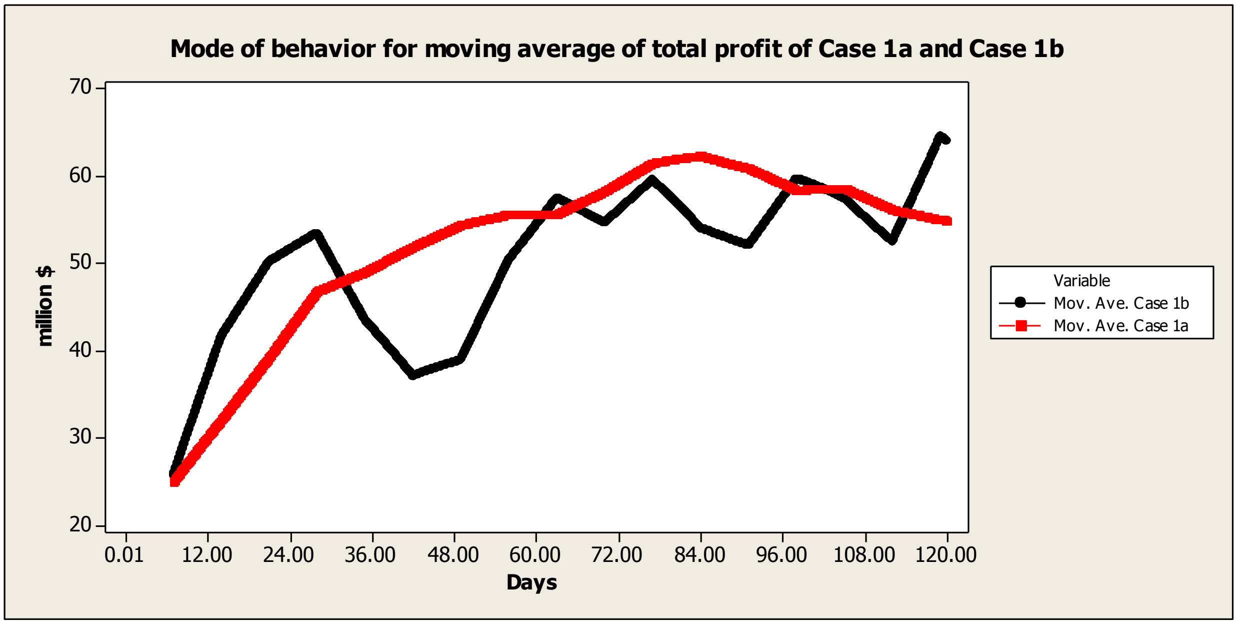

FIGURE 16: COMPARISON OF EXPECTED TOTAL PROFIT FOR CASE 1A AND CASE 1B

Uncertainty can also be incorporated in a standardized way into the model. This case explores the definition of uncertainty over fourteen variables and analyzes its consequences over the results and mode of behavior of the simulation. For this case, uncertainty is defined as a stochastic factor of some of the variables, as described in Appendix 3: Definition of stochastic uncertainty.

The result of simulating this configuration 1,000 times[3] for total profits after 120 days is $63.68 million[4]. It is arguable that this value is not materially different from the original simulation with the uncertainty incorporated in the variation % of the model with a result of $61.96 million (Base case: Original configuration From the original paper); however, when considering the stochastic simulation, the final distribution of the output (total profit at day 120th) shows a wide spectrum of possibilities to consider (Figure 17).

Even when the model is highly non-linear, the fact that the metric Total Profit is cumulative is smoothing the effect of Jensen's Inequality, since the behavior of the metric is quite close to a line. If the model was reconfigured to get closer to a line, the expected value of the stochastic simulation would not be different from the value obtained with the deterministic case, like shown in Appendix 10: Configuring IRIS for steady increment of Total Profits.

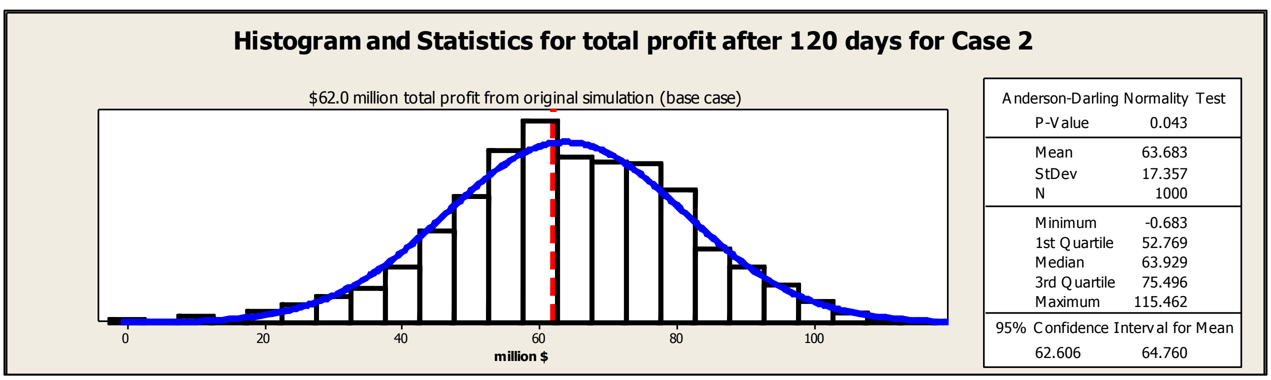

FIGURE 17: HISTOGRAM AND STATISTICS FOR CASE 2

[3] The number of iterations was validated in Appendix 5: Validation of the number of iterations for the stochastic simulation
[4] 22,937.1 seconds of CPU time

Statistically, a one-sample t test can probe that, even when very close. In this case null hypothesis that the value of Total Profit for the Base Case ($62 million) is equal to the expected value of the total profit computed with a stochastic simulation is rejected. This can be interpreted as that they are statistically not equal.

One-Sample T: Total Profit after 120 days for Case 2 vs Total Profit of Base Case

```
Test of mu = 62 vs not = 62

Variable      N     Mean    StDev  SE Mean         95% CI            T      P
T(120)     1000   63.683   17.357    0.549  (62.606, 64.760)  3.07  0.002
```

The actual distribution includes probable cases that are not visible with the deterministic approach. Figure 18 illustrates percentiles 25, 50 and 75 of the probability plot.

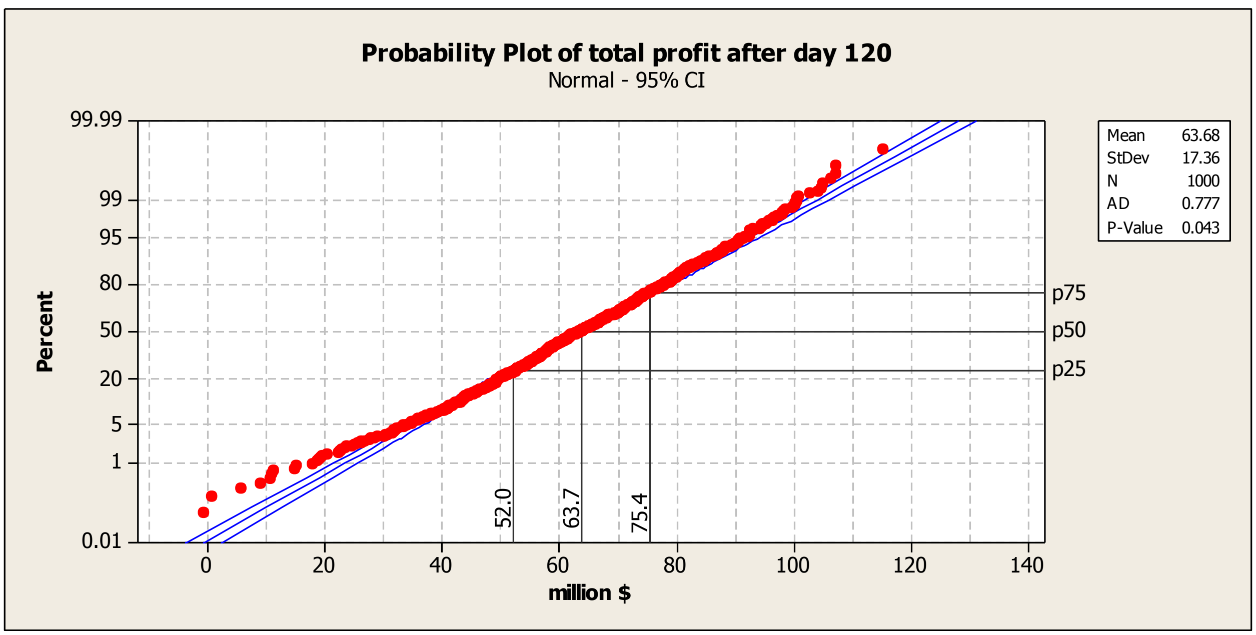

FIGURE 18: PROBABILISTIC PLOT FOR TOTATL PROFIT CASE 2

Figure 19 illustrates the results of total profit over the time of the stochastic simulation, showing the behavior of the expected value, max and min, and the percentiles 25 and 75. It is observable in this graph that the standard deviation grows over time.

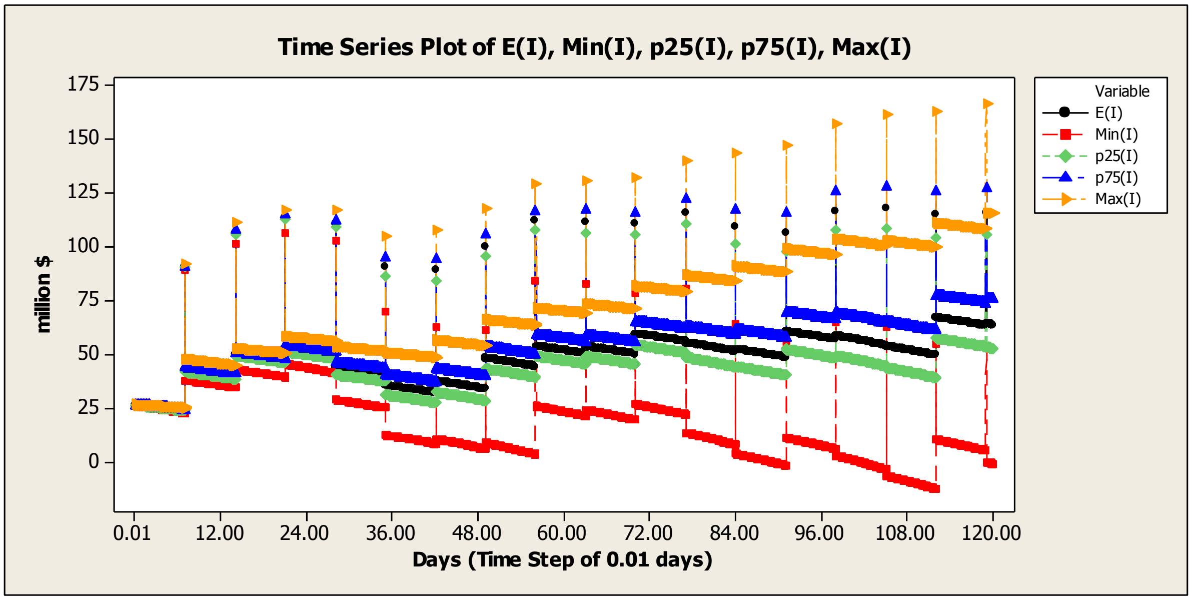

FIGURE 19: TIME SERIES PLOT OF DIFFERENT STATISTICS FOR CASE 2

This effect is a propagation of the probabilistic scenarios over the time horizon of a dynamic simulation. Figure 20 illustrates an almost steady expansion of the standard deviation over time. This is a considerable effect for managers since the longer the horizon of the simulation, the bigger the range of probable scenarios.

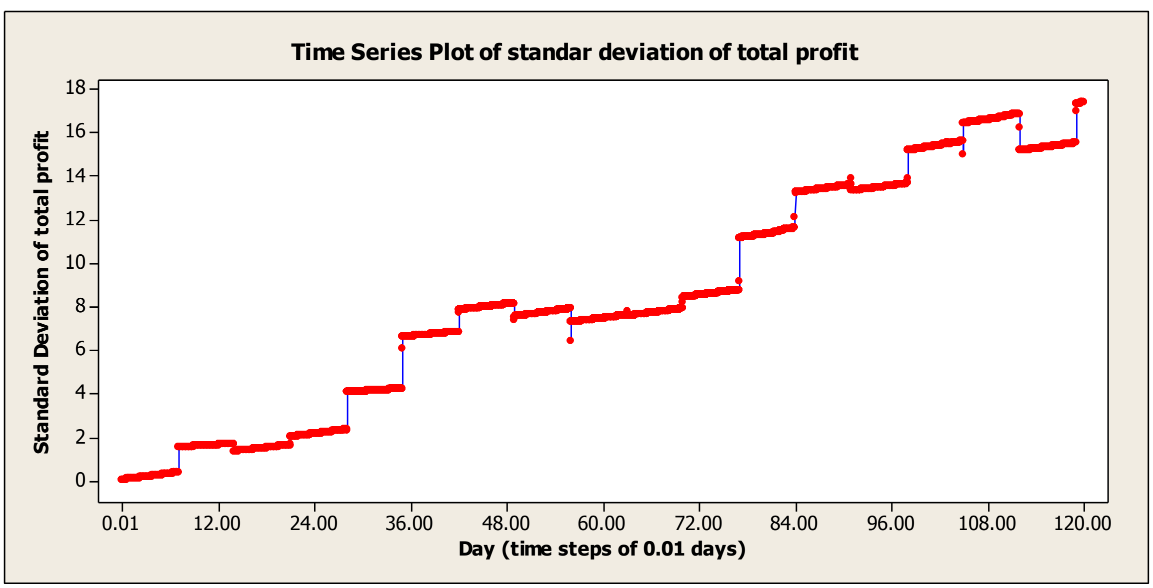

FIGURE 20: PROPAGATION OF STANDARD DEVIATION

Case 2 used a standard configuration of uncertainty for a defined set of variables for the IRIS model (Table 3). The problem with this approach is that, in practice, it is not possible to analyze the individual effect that uncertainty has on each variable. In order to isolate the effect of the uncertainty on different variables, separated stochastic simulations were conducted with stochastic variability in only one variables and everything else fixed.

The original configuration of Base Case was run multiple times using a parallel implementation in Matlab including the stochastic variability defined in Table 3 on at a time, which means 14 different simulations of 1,000 iterations each. The computation of this sensitivity analysis took 296,419.3 seconds (82.3 hours) in the computer used for this research. This time the iteration of stochastic variables of the inputs were also recorded. The correlation of the stochastic variation of each of the fourteen variables with the total profit results were tested to identify which variables have a significant impact on the result when they vary (Appendix 8: Uncertainty versus Sensitivity), together with a graphical comparison with a Marginal Plot.

The variables that have a significant impact on the total profit when they are stochastically varied are:

- Tank Volume (kbbl)
- Product price variation percentage (%)
- Operating cost (K$/kbbl)
- Product or crude inventory cost (K$/kbbl)
- Crude amount variation percentage (%)
- Demand variance percentage (%)
- Pumping rate (kbbl/hr)

The effects of the variation of this variables, as defined in Case 2 (Appendix 3: Definition of stochastic uncertainty), are presented in Figure 21.

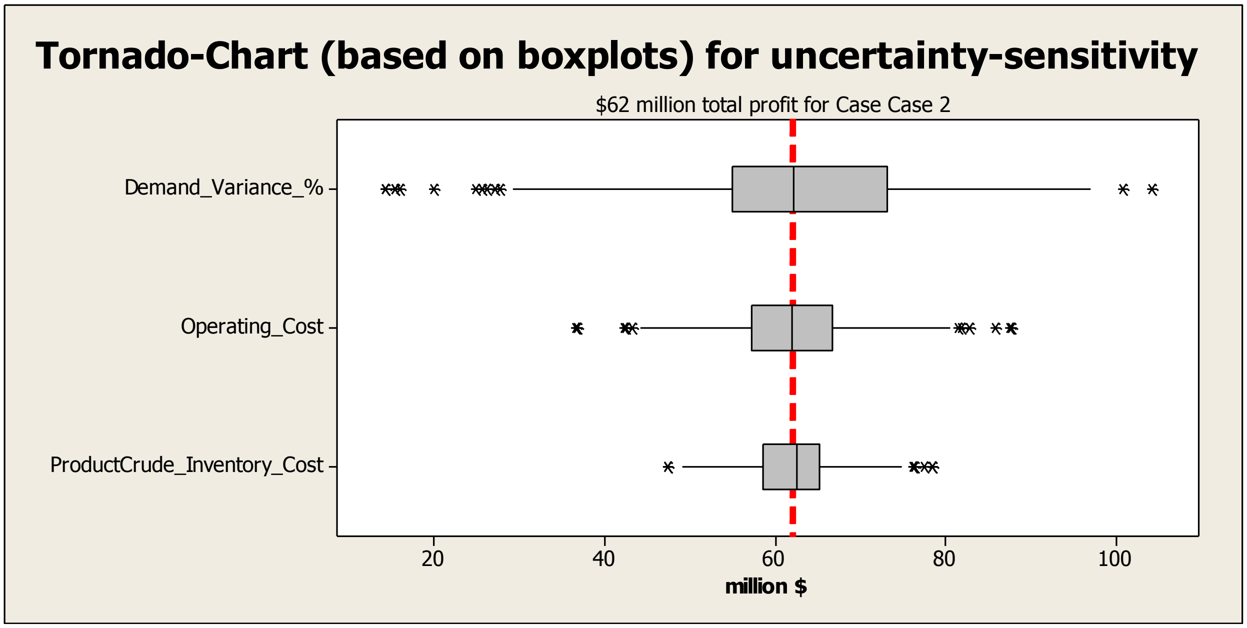

FIGURE 21: TORNADO CHART FOR SIGNIFICANT VARIABLES

From these variables, the only one that has a stochastic variation directly related to Total Profit result, in other words, the only one that produces more Total Profit when vary more, is the Crude amount variation percentage (%). This is relevant for managers since stochastic variation doesn't just produce different effects on the variation of whatever results defined as a metric for a system, but also it will vary in a different direction, which can be interpreted as that some variation can be good for the system. The correlation between this specific variable is described in details in Appendix 8: Uncertainty versus Sensitivity.

THE HUMAN SIDE OF THE EFFECTS OF JENSEN'S INEQUALITY ON THE DYNAMIC SIMULATION

Discontinuity is a source of non-linearity and management can creates discontinuity. This happens because managers or system operators decide to take some major decision about a project creating change in the function or model of behavior (de Neufville, 2012). The decisions will depend on many aspects: culture of the organization, particular context, results from previous periods, external pressures, personal goals, etc. Particularly important are the current mode of behaviors of the system. Different mode of behaviors will produce different managerial behaviors, for example, a declining metric of performance could produce different decisions than an increasing one.

In practice, the context is analyzed, most of the times, using a short time horizon and a limited set of data. This could produce a biased perception of the process performance that will, most probably, impact the behavior. In addition, disruption and special events will certainly influence the perception of status of performance, driving different behaviors or changing the function of performance producing non-linearity.

The practical consequences on the results of a system are that the estimated result will be wrong, and that the management would produce a personal resistance to recognizing and dealing with uncertainty. This is augmented when managers deal with many client relationships and has to respond to expectations of certainty (de Neufville, 2012).

CONCLUSIONS

The analysis of different configurations of uncertainty in the dynamic simulation allowed to observe the following conclusions:

1. Considering stochastic variability can have a major effect on the performance of the system. Since the dynamic model using in this research (IRIS) is non-linear, this is an empirical case of Jensen's Inequality. A specific case of this effect on a concave function was described in Case 1a: Uncertainty of refinery economics.

2. Even if the effect of uncertainty and non-linearity don't impact the expected value of the model, it will produce a distribution of results and a mode of behavior that could be significantly relevant for the decision maker. The mode of behavior of the dynamic results can foster completely different human behaviors from the managers and decision makers, even if the final simulated result is statistically not different.

3. The effects of stochastic variation and complexity on the average, risk or mode of behavior can be non-related, which means that, for example, a case with better expected value could also produce a worse mode of behavior or risk (Figure 15).

4. The stochastic simulation of a dynamic model like IRIS can produce contra-intuitive results due to dynamic complexity. Figure 36 represented a variable that increased total profits when it increased variability

5. The dynamic simulation in the case of IRIS considered a defined and limited time horizon. The effect of the stochastic simulation (Jensen's Inequality) seemed to be propagating over the standard deviation, showing a steady increment over the time horizon of the simulation (Figure 20).The relevance of this insight is that, since the standard deviation is growing over the time horizon of the dynamic simulation, the interval of confidence for the expected value will also grow, which could eventually trick the decision maker to think that the model is producing a similar result, when in fact is just the model losing the precision to make a significant comparison.

6. A sensitivity analysis (Appendix 8: Uncertainty versus Sensitivity) showed that stochastic variability is not important for every variable in the model. This is important since the decision maker should focus the efforts only on the variables that are relevant from the perspective of the sensitivity of the results of the system to their uncertainty.

BIBLIOGRAPHY

Wolfram Alpha LLC, 2013. *www.wolframalpha.com.* [Online]
Available at:
http://www.wolframalpha.com/input/?i=normal+distribution+mean+1+standard+deviation+0.2
[Accessed 18 10 2013].

Aaronson, S., 2013. *Quantum Computing and the Limits of Efficient Computation.* Cambridge: s.n.

Ahamed, S. S., Rangajanardhana, G. & Nagesh, E. L., 2013. Development of Supply Chain Tools Using Genetic Algorithm and Comparison with Particle Swarm Optimization and Simulated Annealing Optimization Algorithms. *IUP Journal of Supply Chain Management,* 10(2), pp. p33-43. 11p.

Akyuz, G. A. & Erkan, T. E., 2010. Supply chain performance measurement: a literature review. *International Journal of Production Research,* Volume 48, p. 5137–5155.

Androulakis, I. P. & Reklaitis, G. V., 1999. Approaches to asynchronous decentralized decision making. *Computers & Chemical Engineering,* 23(3), pp. 341-55.

Applequist, G. E., Pekny, J. F. & Reklaitis, G. V., 2000. Risk and uncertainty in managing chemical manufacturing supply chains. *Computers & Chemical Engineering,* 24(9-10), p. 2211–2222.

Assassa, F. & Marquardt, W., 2014. Dynamic Optimization Using Adaptive Direct Multiple Shooting. *Computers & Chemical Engineering,* 60(10), pp. 242-259.

Azar, S. A., 2008. Jensen's Inequality in Finance. *International Advances in Economic Research,* 07 August, 14(4), p. 433–440.

Biegler, L. T., Lang, Y.-d. & Lin, W., 2014. Multi-scale optimization for process systems engineering. *Computers & Chemical Engineering,* 60(10), p. 17–30.

Bok, J.-K., Grossmann, I. E. & Park, S., 2000. Supply Chain Optimization in Continuous Flexible Process Networks. *Industrial & Engineering Chemistry Research,* Volume 39, pp. 1279-1290.

Bonvin, D. & Srinivasan, B., 2013. Title:On the role of the necessary conditions of optimality in structuring dynamic real-time optimization schemes.. *Computers & Chemical Engineering,* 51(SI), pp. 172-180.

Bottani, E. & Montanari, R., 2010. Supply chain design and cost analysis through simulation. *International Journal of Production Research,* Volume 48, p. 2859–2886.

Cameron, B. & Crawley, E., 2012. *MIT ESD.34 System Architecture Class: Concept to Architecture": Managing Complexity".* Cambridge: s.n.

Chen, C.-L., Wang, B.-W. & Lee, W.-C., 2003. Multiobjective Optimization for a Multienterprise Supply Chain Network. *Industrial & Engineering Chemistry Research,,* Volume 42, pp. 1879-1889.

de Neufville, R., 2012. *MIT Class 1.146/16.861/ESD.71/ESD.710 Engineering Systems Analysis for Design: Flaw of Averages.* Cambridge: s.n.

de Neufville, R. & Scholtes, S., 2011. *Flexibility in Engineering Design.* Cambridge: The MIT Press.

de Weck, O., 2012. *MIT ESD.77 Class: Multiobjective Optimization.* Cambridge: s.n.

Devadas, S., 2013. *MIT Class 6.00 Introduction to Computer Science & Programming, Lecture 8.* Cambridge: s.n.

Disney, S. M., Lambrecht, M., Towill, D. R. & Van de Velde, W., 2006. The value of coordination in a two-echelon supply chain.

Drnevich, P. L. & Croson, D. C., 2013. Information Technology and Business-Level Strategy: Toward an Integrated Theoretical Perspective. *MIS Quarterly,* 37(2), pp. 483-509.

Ehlen, M. A. et al., 2014. Chemical supply chain modeling for analysis of homeland security events. *Computers & Chemical Engineering,* 60(10), p. 102–111.

Fisher, M. L., 1997. What Is the Right Supply Chain for Your Product?. *Harvard Business Review*, March, p. 105–116..

Forrester, J. W., 1961. *Industrial Dynamics.* Student Edition ed. s.l.:s.n.

Francois, G., Srinivasan, B. & Bonvin, D., 2012. Comparison of six implicit real-time optimization schemes. *APII-JESA Journal Europeen des Systemes Automatises,* 46(2), pp. 291-305.

Ghosh, K., Ramteke, M. & Srinivasan, R., 2014. Optimal Variable Selection for Effective Statistical Process Monitoring. *Computers & Chemical Engineering,* 60(10), pp. 260-276.

Gjerdrum, J., Shah, N. & Papageorgiou, L. G., 2001. Transfer Prices for Multienterprise Supply Chain Optimization. *Industrial&EngineeringChemistryResearch,* Volume 40, pp. 1650-1660.

Guessab, A. & Schmeisser, G., 2013. Necessary and sufficient conditions for the validity of Jensen's inequality. *Archiv der Mathematik,* 100(6), pp. 561-570.

Guillén, G. et al., 2005. Multiobjective supply chain design under uncertainty. *Chemical Engineering Science,* Volume 60, p. 1535 – 1553.

Hall, J., Matos, S. & Silvestre, B., 2012. Understanding why firms should invest in sustainable supply chains: a complexity approach. *International Journal of Production Research,* p. 1332–1348.

Isik, F., 2010. An entropy-based approach for measuring complexity in supply chains. *International Journal of Production Research,* Volume 48, p. 3681–3696.

Jiao, J. (., Xu, Q., Wu, Z. & Ng, N.-K., 2009. Coordinating product, process, and supply chain decisions: A constraint satisfaction approach. *Engineering Applications of Artificial Intelligence,* 22(7), pp. 992-1004.

Kannan, G., Noorul Haq, A. & Devika, M., 2009. Analysis of closed loop supply chain using genetic algorithm and particle swarm optimisation. *International Journal of Production Research,* 47(5), p. 1175–1200.

Kannegiesser, M. & Gunther, H.-O., 2011. An integrated optimization model for managing the global value chain of a chemical commodities manufacturer. *Journal of the Operational Research Society,* 62(4), p. 711.

Katz, R., 2004. *The Human Side of Managing Technological Innovation: A collection of readings.* Second Edition ed. New York: Oxford University Press.

Kavi, K., 2009. *The Upstream Energy Value Chain.* s.l.:s.n.

Khalilpour, R. & Karimi, I. A., 2014. Parametric optimization with uncertainty on the left hand side of linear programs. *Computers & Chemical Engineering,* 60(10), p. 31–40.

Koo, L. Y., Adhitya, A., Srinivasan, R. & Karimi, I. A., 2008. Decision support for integrated refinery supply chains Part 2. Design and operation. *Computers and Chemical Engineering,* Volume 32, p. 2787–2800.

Lichocik, G. & Sadowski, A., 2013. EFFICIENCY OF SUPPLY CHAIN MANAGEMENT. STRATEGIC AND OPERATIONAL APPROACH. *Scientific Journal of Logistics (LogForum),* 9(2), pp. p119-125. 7p.

LLamasoft, Inc., 2013. *www.llamasoft.com.* [Online]
Available at: http://www.llamasoft.com/news/llamasoft-multi-objective-optimization-gives-businesses-holistic-view-of-trade-offs-across-multiple-objectives/
[Accessed 18 September 2013].

Maier, M. W. & Rechtin, E., 2002. *The Art of Systems Architecting.* Second Edition ed. s.l.:CRC Press LLC.

MathWorks, 2013. *www.mathworks.com.* [Online]
Available at: http://www.mathworks.com/help/simulink/ug/types-of-sample-time.html#brrdmmw-6
[Accessed 03 10 2013].

Matos, S. & Hall, J., 2007. Integrating sustainable development in the supply chain: The case of life cycle assessment in oil and gas and agricultural biotechnology. *Journal of Operations Management,* 25(6), pp. 1083-1102.

MIT Sloan School of Management, 2013. *http://mitsloan.mit.edu.* [Online]
Available at: http://mitsloan.mit.edu/phd/system-dynamics.php
[Accessed 17 November 2013].

Montagna, J. M. & Iribarren, O. A., 1988. Optimal computation sequence in the simulation of chemical plants. *Computers & Chemical Engineering,* 12(1), p. 71–79.

Naraharisetti, P. & Karimi, I., 2010. Supply chain redesign and new process introduction in multi-purpose plants. *Chemical Engineering Science,* 65(8), p. 2596–2607.

Naraharisetti, P., Karimi, I. & Srinivasan, R., 2008. Supply chain redesign through optimal asset management and capital budgeting. *Computers & Chemical Engineering,* 32(12), pp. 3153-3169.

New York Mercantile Exchange (NYMEX), 2013. *www.eia.gov.* [Online]
Available at: http://www.eia.gov/dnav/pet/pet_pri_fut_s1_d.htm
[Accessed 30 10 2013].

Ojha, D., Gianiodis, P. T. & Manuj, I., 2013. Impact of logistical business continuity planning on operational capabilities and financial performance. *The International Journal of Logistics Management,* 24(2), pp. p180-209.

Ozsen, L., Daskin, M. S. & Coullard, C. R., 2009. Facility Location Modeling and Inventory Management with Multisourcing. *Transportation Science,* 43(4), p. 455–472.

Pal, A., Chan, F., Mahanty, B. & Tiwari, M., 2011. Aggregate procurement, production, and shipment planning decision problem for a three-echelon supply chain using swarm-based heuristics. *International Journal of Production Research,* Volume 49, p. 2873–2905.

Papageorgiou, L. G., 2009. Supply chain optimisation for the process industries: Advances and opportunities. *Computers & Chemical Engineering,* 33(12), p. 1931–1938.

Papalambros, P. Y. & Wilde, D. J., 2000. *Principles of Optimal Design: Modeling and Computation.* 2nd ed. Cambridge: Cambridge University Press.

Puigjaner, L. & Laınez, J. M., 2008. Capturing dynamics in integrated supply chain management. *Computers and Chemical Engineering,* Volume 32, p. 2582–2605.

Qiu, M. & Li, J., 2011. *Real-time embedded systems : optimization, synthesis, and networking.* s.l.:CRC Press.

Quan, N., Yin, J., Ng, S. H. & Lee, L. H., 2013. Simulation optimization via kriging: a sequential search using expected improvement with computing budget constraints. *IIE Transactions,* 45(7), pp. 763-780.

Ross, A. & Droge, C., 2002. An integrated benchmarking approach to distribution center performance using DEA modeling. *Journal of Operations Management,* 20(1), pp. 19-32.

Russell, E. L. & Braatz, R. D., 1998. Model reduction for the robustness margin computation of large scale uncertain systems. *Computers & Chemical Engineering,* 22(7-8), p. 913–926.

Schloetzer, J. D., 2012. *Process Integration and Information Sharing in Supply Chains.* s.l.:Georgetown University.

Schönlein, M., Makuschewitz, T., Wirth, F. & Scholz-Reiter, B., 2013. Measurement and optimization of robust stability of multiclass queueing networks: Applications in dynamic supply chains. *European Journal of Operational Research,* 229(1), pp. p179-189. 11p.

Shank, J. K. & Govindarajan, V., 1992. *Strategic Cost Management: The Value Chain Perspective.* s.l.:Dartmouth College.

Siddiqui, F., Haleem, A. & Sharma, C., 2012. The Impact of Supply Chain Management Practices in Total Quality Management Practices and Flexible System Practices Context: An Empirical Study in Oil and Gas Industry. *Global Journal of Flexible Systems Management,* p. 13(1):11–23.

Spillan, J. E., McGinnis, M. A., Kara, A. & Liu Yi, G., 2013. A comparison of the effect of logistic strategy and logistics integration on firm competitiveness in the USA and China. *The International Journal of Logistics Management,* 24(2), pp. p153-179.

Sterman, J. D., 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World.* Cambridge: Jeffrey J. Shelstad.

Subramanian, P., Ramkumar, N., Narendran, T. & Ganesh, K., 2012. A technical note on 'Analysis of closed loop supply chain using genetic algorithm and particle swarm optimisation'. *International Journal of Production Research,* Volume 50, p. 593–602.

Suresh, P. S. et al., 2008. Decision support for integrated refinery supply chains: Part 1. Dynamic simulation. *Computers & Chemical Engineering,* 32(11), p. 2767–2786.

Tang, L., Jing, K. & He, J., 2013. An improved ant colony optimisation algorithm for three-tier supply chain scheduling based on networked manufacturing. *International Journal of Production Research,* Volume 51, p. 3945–3962.

The MathWorks, Inc, 2013. *www.mathworks.com.* [Online]
Available at: http://www.mathworks.com/help/distcomp/advanced-topics.html
[Accessed 23 10 2013].

The MathWorks, Inc, 2013. *www.mathworks.com.* [Online]
Available at: http://www.mathworks.com/help/simulink/ug/types-of-sample-time.html
[Accessed 30 10 2013].

Tokman, M. et al., 2013. SME supply chain portfolios: firm satisfaction and organization resources. *The International Journal of Logistics Management,* 24(2), pp. pp. 271 - 300.

Ulrich, K. T. & Eppinger, S. D., 2012. *Product Design and Development.* Fifth Edition ed. Singapore: McGraw-Hill.

Webster, M. & de Neufville, R., 2012. *MIT ESD.71 Class: Uncertainty II: Probability and Risk Analysis.* Cambridge: s.n.

Weijermars, R., 2010. Value chain analysis of the natural gas industry-Lessons from the US regulatory success and opportunities for Europe. *Journal of Natural Gas Science and Engineering,* pp. p86-p104, 19p.

Weijermars, R., 2012. Regulatory reform options to revitalize the US natural gas value chain. *Utilities Policy,* June.pp. Vol. 21, p50-58. 9p.

Weijermars, R., 2012. Regulatory reform options to revitalize the US natural gas value chain. *Utilities Policy,* pp. 50-58.

West, B. & Lafferty, T., 2008. *Synchronizing the offshore supply chain creates new value.* s.l.:Deloitte Consulting LLP.

Wikipedia, 2013. *Wikipedia.* [Online]
Available at: http://en.wikipedia.org/wiki/Bullwhip_effect
[Accessed 09 11 2013].

Wikipedia, 2013. *Wikipedia.* [Online]
Available at: http://en.wikipedia.org/wiki/Relaxation_(approximation)
[Accessed 23 09 2013].

Wikipedia, 2013. *Wikipedia.* [Online]
Available at: http://en.wikipedia.org/wiki/Benders'_decomposition
[Accessed 23 09 2013].

Wolfram Alpha LLC, 2013. *www.wolframalpha.com.* [Online]
Available at: http://www.wolframalpha.com/input/?i=plot+x%5E2+-+6
[Accessed 06 10 2013].

You, F. & Grossmann, I. E., 2008. Design of responsive supply chains under demand uncertainty. *Computers & Chemical Engineering,* 32(12), pp. 3090-3111.

Yusuf, Y. Y. et al., 2012. *A relational study of supply chain agility, competitiveness and business performance in the oil and gas industry.* s.l.:s.n.

Yusuf, Y. Y. et al., 2013. A study of the diffusion of agility and cluster competitiveness in the oil and gas supply chains.

Zavala, V. M. & Anitescu, M., 2010. Real-Time Nonlinear Optimization as a Generalized Equation. *SIAM Journal on Control & Optimization,* 48(7/8), pp. 5444-5467. 24p.

APPENDIXES

APPENDIX 1: IRIS SETTINGS AND RESULTS FROM ORIGINAL PAPER

	IRIS	Control Panel				Sales		Customer	Suppliers	Procurement			
Case study	Simulation horizon	Procure-ment policy	Product-ion policy	Max throughput	Tank capacity	Mean product demand	Demand switch	Quality uncertainty seed	Disruption occurrence seed	Quality disruption emergency procurement	Supply disruption emergency procurement	Results shown in paper	Taken from IRIS
Base	120	1	1	250	[1250 1250 1250 1250 1250]	[405 100 205 165] / 7	off	[68 48 48 98]	78	off	off	Fig. 5a	Crude Inventory Profile (IRIS/Display)
												Fig. 5b	Planned Vs Actual Throughput (IRIS/Display)
												Table 12	IRIS/Display
1	120	2	1	250	[1250 1250 1250 1250 1250]	[405 100 205 165] / 7	off	[68 48 48 98]	78	off	off	Table 12	IRIS/Display
2	120	2	2	250	[1250 1250 1250 1250 1250]	[405 100 205 165] / 7	off	[68 48 48 98]	78	off	off	Table 12	IRIS/Display
3	120	2	2	250	[1250 1250 1000 1250 1250]	[405 100 205 165] / 7	off	[68 48 48 98]	78	off	off	Fig. 6a	Crude Inventory Profile (IRIS/Display)
												Fig. 6b	VLCC extra waiting period (IRIS/Display)
	120	2	2	250	[1250 1250 1250 1250 1000]	[405 100 205 165] / 7	off	[68 48 48 98]	78	off	off	Fig. 7	Crude Inventory Profile (IRIS/Display)
4	480	1	1	250	[12500 12500 12500 12500 12500]	[355 90 265 145] / 7	on	[68 48 48 98]	78	off	off	Fig. 8	Forecast Demand Profile (IRIS/Sales)
												Fig. 9	Actual throughput1 (IRIS/Display)
	480	1	1	400	[12500 12500 12500 12500 12500]	[355 90 265 145] / 7	on	[68 48 48 98]	78	off	off	Fig. 9	Actual throughput1 (IRIS/Display)
5	120	1	1	250	[1250 1250 1250 1250 1250]	[405 100 205 165] / 7	off	[68 48 48 98]	88	off	off	Fig. 10a	Crude Inventory Profile (IRIS/Display): crude 1 (yellow), crude 2 (purple)
	120	1	1	250	[1250 1250 1250 1250 1250]	[405 100 205 165] / 7	off	[68 48 48 98]	88	off	on	Fig. 10b	Crude Inventory Profile (IRIS/Display): crude 1 (yellow), crude 2 (purple)
6	120	2	2	250	[1250 1250 1250 1250 1250]	[365 125 205 150] / 7	off	[68 18 48 98]	78	off	off	Fig. 11a	Customer Satisfaction (IRIS/Display): jet fuel (second from top)
												Fig. 11b	Demand Deficit (IRIS/Display): jet fuel (second from top)
												Fig. 11c	Actual Demand Profile (IRIS/Sales): jet fuel (second from top)
												Fig. 12a	Product Inventory Profile (IRIS/Display): jet fuel (purple)
												Fig. 13	Actual throughput1 (IRIS/Display)
	120	2	2	250	[1250 1250 1250 1250 1250]	[365 125 205 150] / 7	off	[68 18 48 98]	78	on	off	Fig. 12b	Product Inventory Profile (IRIS/Display): jet fuel (purple)
												Fig. 13	Actual throughput1 (IRIS/Display)

FIGURE 22: IRIS SETTINGS AND RESULTS FROM ORIGINAL PAPER

Some of the variables considered by the original model are:

- Simulation horizon: Since the model is a dynamic simulation, the time horizon represents the simulated time that the dynamic model will run (Sterman, 2000), which is not the same than the running time of the model, for example, "IRIS simulation run for a 120-day horizon requires ~90 s on a Pentium IV, 2.8 GHz processor" (Suresh, et al., 2008). Interestingly, this occurs in a sequence of discrete time steps (or signals in Matlab) and not necessarily in a continuous time like the real world. (MathWorks, 2013). In the original base case the time horizon is defined as 120 days.

- Max throughput: IRIS considers that "the actual throughput is selected to also ensure operation within the minimum and maximum operable throughput limits" (Suresh, et al., 2008). This assumes that the maximum operable limit is a constant; however, in many cases the maximum capacities have a level of uncertainty due to reliability, operation strategy and many other factors that can limit the capacity.

- Tank capacity: The refinery modeled by IRIS segregates crudes in different tanks not allowing the mixing of crudes. The model represents five (5) tanks for each crude type, each having a fixed capacity making total storage capacity of 5 times the fixed capacity of each tank. In the base case of the original paper, the total tank capacity is 1250 kbbl formed by each tank having a capacity of 250 kbbl (Suresh, et al., 2008).

- Mean product demand: the refinery of the model from the point of view of a supply chain can operate either through a push or a pull-mode. In IRISI, the supply chain is managed by a production plan comprising the planned throughput and production mode, based on forecast product demands (Suresh, et al., 2008).

- Quality uncertainty seed: During product delivery, it may happen that a product fails quality tests and is rejected by the customer. This is modeled by IRIS using an stochastic, binary variable, whose value of 1 indicates product acceptance and 0 product rejection. This variable is generated randomly for all products at each due date and is dependent on the quality index random seed. (Suresh, et al., 2008)

- <u>Disruption occurrence seed:</u> Disruption scenarios can be defined by the user of the model to simulate different supply cases about the occurrence of the transportation delay. (Suresh, et al., 2008)

Specifically, the original IRIS model takes the values for the variables assigned directly from the Matlab Simulink subsystem masks, like exemplified in Figure 23.

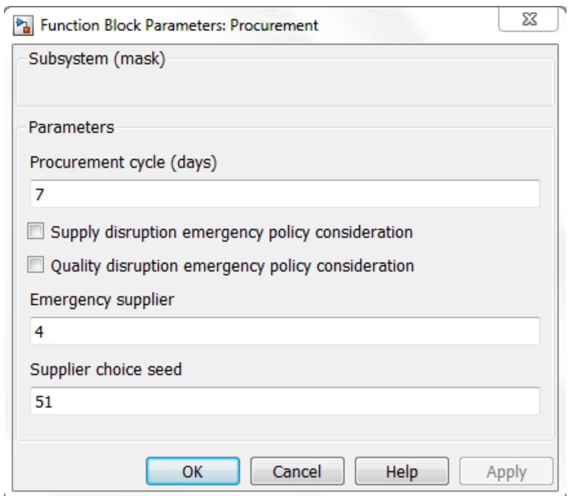

FIGURE 23: EXAMPLE OF VARIABLE VALUE ASSIGNMENT IN THE MATLAB SIMULINK SUBSYSTEM'S MASK

For this research, some of the variables were redefined from the same Matlab *.m that will run the simulations later with uncertainty. The variables were described in the *.m file in Table 2:

TABLE 2: VARIABLE DEFINITION AND VALUE ASSIGNMENT

Subsystem	Variable	Definition in m.file
Customer	Quality uncertainty seed	quauncsee=[68 48 48 98];
Control Panel	Number of crudes	numcru=5;
	Crude storage capacity limit (kbbl)	crustocaplim=[1250 1250 1250 1250 1250];
	Crude tank volume (kbbl)	crutanvol=250;
	Production cycletime (days)	procyc=7;
	Planning horizon (days)	plahor=28;
	Maximum throughput (kbbl/day)	maxthr=250;
	Minimum throughput (kbbl/day)	minthr=100;
	Crude safety stock (kbbl)	crusafsto=100;
	Product safety stock factor (0 to 1)	prosafstofac=0.2;
Refinery Economics	Product price variation seed	varprodprice=[10 12 14 16];
	Product price (K$/kbbl)	prodprice=[79 76 68 47];
	Product price variation percentage (%)	proprivarper=1;
	Demurrage charge (K$/day)	demcha=100;
	Operating cost (K$/kbbl)	opcost=2;
	Product or crude inventory cost (K$/kbbl)	procruinvcost=0.05;

	Penalty for order violation (K$/kbbl)	penordvio=[5 5 5 5];
Sales	Actual demand variation percentage (%):	actdemvarper=[5 5 5 5];
	Magnitude of demand increase	magdeminc=2;
	Mean product demand (kbbl/day)	meaprodem=[405 100 205 165]/7;
	Demand variation seed	demvarsee=[30 40 50 60];
	Forecast vs actual demand seed	forvsactdemsee=[50 51 51 51];
	Demand variance percentage (%)	demvarper=25;
Suppliers	Number of suppliers	numsup=4;
	Suppliers 1-3 crude amount upper bounds (kbbl)	supcruamoub=[680 680 680 680 680;700 700 700 700 700;720 720 720 720 720];
	Suppliers 1-3 crude amount variation seeds	supcruamovarsee=[10 12 14 16 18; 20 22 24 26 28; 30 32 34 36 38];
	Crude amount variation percentage (%)	cruamovarper=5;
	Suppliers 1-3 crude prices (K$/kbbl)	supcrupri=[55 56 53 50 52;55 56 53 50 52;55 56 53 50 52];
	Suppliers 1-3 crude price variation seeds	supcruprivarsee=[10 12 14 16 18;20 22 24 26 28;30 32 34 36 38];
	Crude price variation percentage (%)	cruprivarper=1;
	Emergency crude price (K$/kbbl)	emecrupri=[65 66 63 60 62];
	Disruption occurrence seed	disoccsee=78;
	Upper limit on disruption awareness (days)	uplimdirawa=17;
	Upper limit on disruption magnitude (days)	uplimdirmag=15;
	Disruption magnitude seed	dismagsee=100;
Product inventory	Initial product inventory level (kbbl)	iniproinvlev=[200 200 200 200];
Port / Storage	Pumping rate (kbbl/hr)	pumprate=75;
	VLCC allowable wait time before demurrage (days)	VLCCalldem=1;
	VLCC allowable idle time in proportion to volume (>= 0)	VLCCallvol=20;
Reformer	Reformer yield variation (%)	refyievar=0.1;
Cracker	Cracker yield variation (%)	crayievar=0.1;
CDU	CDU yield variation (%)	CDUyievar=0.01;
	Waste yield for each product (0-1)	wasyiepro=[0 0 0 0];

All the results from the original papers were replicated to ensure calibration of the model. The code used in the Matlab *.m in the file to run the model in Matlab R2013a considered a for loop in order to allow multiple runs to incorporate uncertainty. The construction was the following:

EQUATION 6: ORIGINAL CODE FOR SIMPLE LOOP OF THE SIMULINK SIMULATIONS

```
tic
open_system('iris')
for J = 1:1:1
%Here the definition of the variables (Table 2):
simOut=sim('iris','StopTime','120');
    if J==1
        so = simOut.get('simouttest'); % "so" is the variable exported by
simouttest
    else
        A = simOut.get('simouttest');
        so=cat(1,so,A);
    end
end
toc
```

The base case configuration with a time horizon of 120 days in the computer used for this research (Intel Core i5-3320M CPU@2.60GHz with 8.00GB of installed RAM) takes and elapsed time to run of 96.97 seconds.

Total Profit results are exported to the Matlab Workspace using a block to export the results in Simulink. This block was added to the original Simulink model with the name simouttest (red circle in Figure 24):

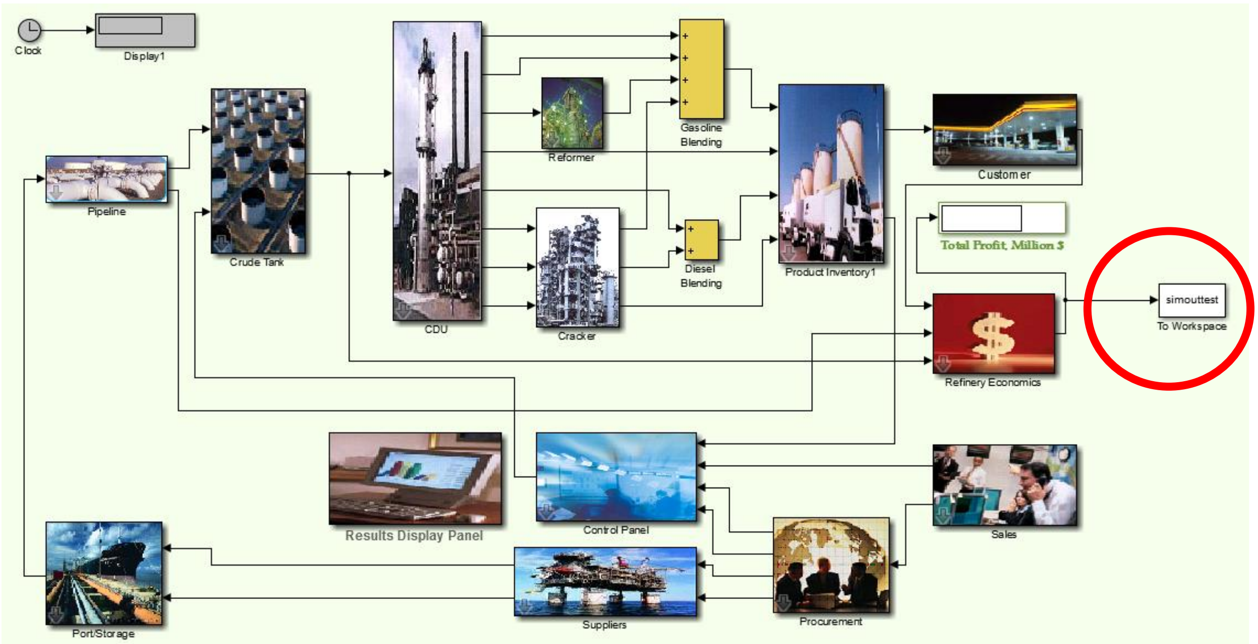

FIGURE 24: MATLAB SIMULINK MODEL WITH SIMOUTTEST BLOCK TO EXPORT RESULTS TO WORKSPACE

APPENDIX 3: DEFINITION OF STOCHASTIC UNCERTAINTY

The stochastic variability was incorporated through a factor of stochasticity as follows:

$$Vwu = Fu \cdot V$$

EQUATION 7: INCORPORATING UNCERTAINTY

Where:

- V: Deterministic (original) value of the variable
- Vwu: Stochastic variable
- Fu: Factor of uncertainty. This factor is defined as a normal distribution with μ=1 and σ=0.2 (Figure 25)

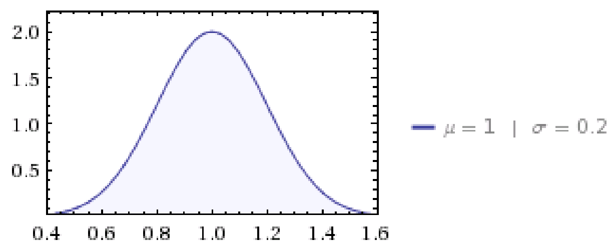

FIGURE 25: NORMAL DISTRIBUTION PLOT FOR μ=1 AND σ=0.2 (WOLFRAM ALPHA LLC, 2013)

This factor is incorporated in some of the variables of the model as described in Table 3, and the stochastic simulation was coded using a parallel computing loop (parfor loop in Matlab) as described in Appendix 4: Improving the Computational Performance through Parallel computation.

TABLE 3: DEFINITION OF UNCERTAINTIES IN THE *.M FILE

Subsystem	Variable	Definition in m.file
Control Panel	Crude tank volume (kbbl)	crutanvol= min(max(0,normrnd(1,0.2))*250,250);[5]
Refinery Economics	Product price variation percentage (%)	proprivarper= max(0,normrnd(1,0.2))*1;
	Operating cost (K$/kbbl)	opcost= max(0,normrnd(1,0.2))*2;
	Product or crude inventory cost (K$/kbbl)	procruinvcost= max(0,normrnd(1,0.2))*0.05;
Sales	Magnitude of demand increase	magdeminc= max(0,normrnd(1,0.2))*2;
	Demand variance percentage (%)	demvarper= max(0,normrnd(1,0.2))*25;
Suppliers	Crude amount variation percentage (%)	cruamovarper= max(0,normrnd(1,0.2))*5;
	Crude price variation percentage (%)	cruprivarper= max(0,normrnd(1,0.2))*1;
	Disruption occurrence seed	disoccsee= round(max(0,normrnd(1,0.2))*78);[6]
	Disruption magnitude seed	dismagsee= round(max(0,normrnd(1,0.2))*100);
Port / Storage	Pumping rate (kbbl/hr)	pumprate= max(0,normrnd(1,0.2))*75;
Reformer	Reformer yield variation (%)	refyievar= max(0,normrnd(1,0.2))*0.1;
Cracker	Cracker yield variation (%)	crayievar= max(0,normrnd(1,0.2))*0.1;
CDU	CDU yield variation (%)	CDUyievar= max(0,normrnd(1,0.2))*0.01;

[5] The maximum crude tank volume capacity is limited by the physical capacity of the tank. The uncertainty can be due to structural problems, operational considerations, among many others. The value is also non negative.

[6] The round() function is implemented because the model expects integers in that variable.

APPENDIX 4: IMPROVING THE COMPUTATIONAL PERFORMANCE THROUGH PARALLEL COMPUTATION

The code to run the plug & play model (IRIS) was original implemented in Matlab in a *.m file as described in Equation 8. The variables for the Simulink simulation were defined in a different way - using set_param()- than the code described in Equation 6. This is to consider the flexibility to move further to parallel computation if required (The MathWorks, Inc, 2013):

EQUATION 8: REGULAR FOR LOOP IMPLEMENTED WITH SET_PARAM() FUNCTION

```
%Regular loop:
tic;
output = cell(1,iterations);
load_system('iris') %Load system without viewing

for i = 1:1:iterations

    %DEFINING VARIABLES:
    %CONTROL PANEL
    set_param('iris/Control
Panel','tankvolume',num2str(min(max(0,normrnd(1,0.2))*250,250)));
    %REFINERY ECONOMICS:
    set_param('iris/Refinery
Economics','pdtpricevariationpercent',num2str(max(0,normrnd(1,0.2))*1),'opcos
t',num2str(max(0,normrnd(1,0.2))*2),'invcost',num2str(max(0,normrnd(1,0.2))*0
.05));
    %SALES:
set_param('iris/Sales','magnitudeofdemandincrease',num2str(max(0,normrnd(1,0.
2))*2),'demandvariance',num2str(max(0,normrnd(1,0.2))*25));
    %SUPPLIERS:
set_param('iris/Suppliers','crudeamountvariationpercent',num2str(max(0,normrn
d(1,0.2))*5),'crudepricevariationpercent',num2str(max(0,normrnd(1,0.2))*1),'s
upplierseed',num2str(round(max(0,normrnd(1,0.2))*78)),'magnitudeseed',num2str
(round(max(0,normrnd(1,0.2))*100)));
    %PORT/STORAGE
set_param('iris/PortStorage','pumprate',num2str(max(0,normrnd(1,0.2))*75));
    %REFORMER
set_param('iris/Reformer','variationpercentage',num2str(max(0,normrnd(1,0.2))
*0.1));
    %CRACKER
set_param('iris/Cracker','percentagevariation',num2str(max(0,normrnd(1,0.2))*
0.1));
    %CDU
set_param('iris/CDU','percentagevariation',num2str(max(0,normrnd(1,0.2))*0.01
));

    simOut=sim('iris','StopTime','60');
    output{i}=simOut.get('simouttest');
    i

end
trunr=toc
```

```
finaloutputr=squeeze(cell2mat(output)).';
```

The original computation with 1,000 iterations using the code of Equation 8 had a performance of 30,273.8 seconds of computation time. In order to improve this performance the code was rewritten to take advantage of the Parallel Computation capabilities available in Matlab R2013a. This was implemented simply changing the regular for loop for a parfor loop, which is a relatively straightforward change. The code developed for the parallel computation is described in Equation 9. Two relevant changes in this code compared with the one described in Equation 8 are that the Simulink model has to be loaded inside the parfor loop, and the variables have to be defined using the function set_param() in order to allow the system to distribute the Simulink simulations to parallel cores.

EQUATION 9: MATLAB CODE FOR PARALLEL COMPUTATION

```
%Parallel computation of the loop:
matlabpool
tic;
outputpc = cell(1,iterations);

parfor J = 1:1:iterations
    load_system('iris') %Load system without viewing

    %DEFINING VARIABLES:
    %CONTROL PANEL
    set_param('iris/Control
Panel','tankvolume',num2str(min(max(0,normrnd(1,0.2))*250,250)));
    %REFINERY ECONOMICS:
    set_param('iris/Refinery
Economics','pdtpricevariationpercent',num2str(max(0,normrnd(1,0.2))*1),'opcos
t',num2str(max(0,normrnd(1,0.2))*2),'invcost',num2str(max(0,normrnd(1,0.2))*0
.05));
    %SALES:
set_param('iris/Sales','magnitudeofdemandincrease',num2str(max(0,normrnd(1,0.
2))*2),'demandvariance',num2str(max(0,normrnd(1,0.2))*25));
    %SUPPLIERS:
set_param('iris/Suppliers','crudeamountvariationpercent',num2str(max(0,normrn
d(1,0.2))*5),'crudepricevariationpercent',num2str(max(0,normrnd(1,0.2))*1),'s
upplierseed',num2str(round(max(0,normrnd(1,0.2))*78)),'magnitudeseed',num2str
(round(max(0,normrnd(1,0.2))*100)));
    %PORT/STORAGE
set_param('iris/PortStorage','pumprate',num2str(max(0,normrnd(1,0.2))*75));
    %REFORMER
set_param('iris/Reformer','variationpercentage',num2str(max(0,normrnd(1,0.2))
*0.1));
    %CRACKER
set_param('iris/Cracker','percentagevariation',num2str(max(0,normrnd(1,0.2))*
0.1));
    %CDU
```

```
set_param('iris/CDU','percentagevariation',num2str(max(0,normrnd(1,0.2))*0.01
));

    simOut=sim('iris','StopTime','60');
    outputpc{J}=simOut.get('simouttest');
    J

end

trunpc=toc
matlabpool close
finaloutputpc=squeeze(cell2mat(outputpc)).';
```

The simulations were computed using an Intel Core i5-3320M CPU@2.60GHz with 8.00GB of installed RAM with 2 cores. The performance improvement of the parfor loop using the 2 available cores in the computer of this research produced a significant difference in performance for 1,000 iterations. The results are shown in Table 4.

TABLE 4: Comparison of Performance of Regular For Loop (EQUATION 8) And Parfor Loop (EQUATION 9) for 1,000 Iterations

Regular loop	Parfor loop	Units
30273.8	10663.9	Seconds
504.6	177.7	Minutes
8.4	3.0	Hours

In order to ensure that both the regular loop and the parallel loop (parfor) were producing results that are not statistically different, the histogram of both distributions were compared (Figure 26)

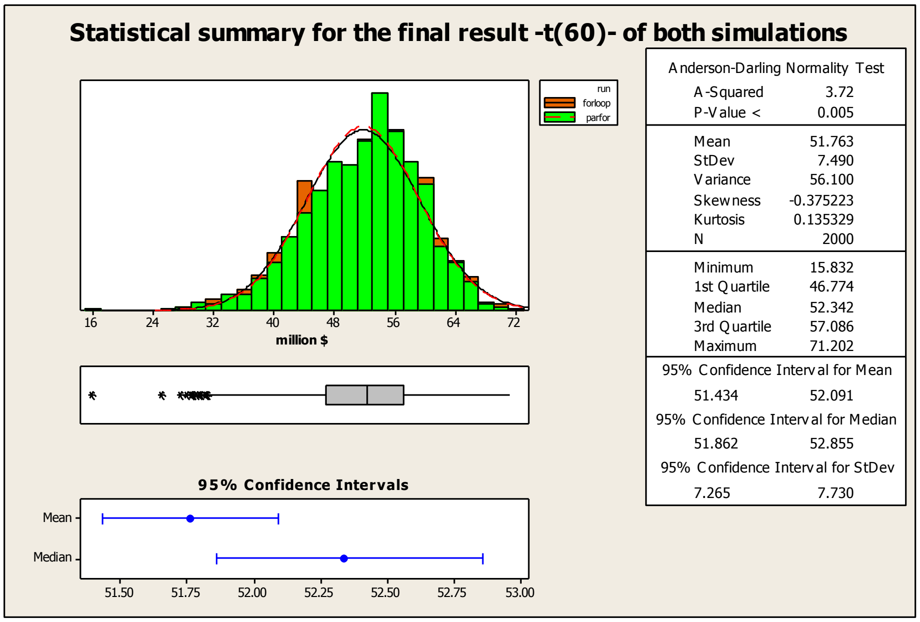

FIGURE 26: HISTOGRAMS OF BOTH SIMULATIONS (REGULAR FORLOOP AND PARFOR LOOP)

Also, the variance of both results is compared using the statistical software Minitab 16, and the conclusion is that they are not statistically different according to the following analysis of variance (ANOVA):

One-way ANOVA: t(60) versus run

```
Source    DF        SS    MS     F      P
run        1       0.1   0.1  0.00  0.969
Error   1998  112144.8  56.1
Total   1999  112144.9

S = 7.492   R-Sq = 0.00%   R-Sq(adj) = 0.00%


                              Individual 95% CIs For Mean Based on
                              Pooled StDev
Level      N    Mean  StDev  --------+---------+---------+---------+-
forloop  1000  51.756  7.562  (-----------------*------------------)
parfor   1000  51.769  7.421  (------------------*-----------------)
                              --------+---------+---------+---------+-
                                  51.50     51.75     52.00     52.25


Pooled StDev = 7.492
```

Finally, also using Minitab 16, the sample of results produced by the regular for loop and the sample produced by the parallel loop, were compared with the following two sample t-test, concluding that they are not statistically different (pvalue>α cannot reject null hypothesis):

Two-Sample T-Test and CI: t(60), run

```
Two-sample T for t(60)

run         N   Mean  StDev  SE Mean
forloop  1000  51.76   7.56     0.24
parfor   1000  51.77   7.42     0.23


Difference = mu (forloop) - mu (parfor)
Estimate for difference:  -0.013
95% CI for difference:  (-0.670, 0.644)
T-Test of difference = 0 (vs not =): T-Value = -0.04  P-Value = 0.969  DF =
     1998
Both use Pooled StDev = 7.4919
```

A box-plot (Figure 27) allows to graphically observe that both results are statistically not different.

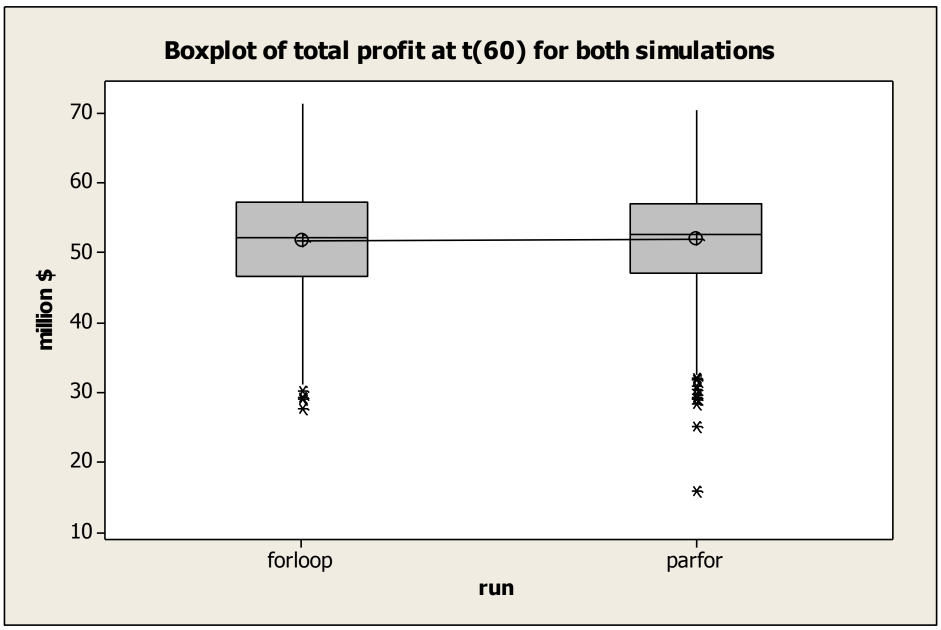

FIGURE 27: BOX PLOT OF BOTH REGULAR FOR LOOP AND PARALLEL LOOP

APPENDIX 5: VALIDATION OF THE NUMBER OF ITERATIONS FOR THE STOCHASTIC SIMULATION

One of the decisions of modeling the stochastic simulation is to define the appropriate number of iterations to minimize error and achieve a certain level of confidence. This is challenging in the case of the model represented by IRIS because the real population is unknown, and therefore μ and σ are unknown. This means that it is not possible to compare the result of the sample given by the simulation with a referential patron.

A workaround of this problem is to choose a certain sample size and compare that with a bigger sample size, ideally one order of magnitude bigger than the first one. If both samples produce results that are not statistically different, then the smaller on can be considered representative enough.

In this case, the number of iterations selected is 1,000. This is the sample size used in the chapter "Stochastic Simulations". The configuration of the simulation used to define the number of iterations is the configuration presented in Case 2: standard uncertainty. The second stochastic simulation considered 10,000 iterations. The code to run the simulations is the same one shown in Equation 9.

The simulation with 10,000 iterations ran for 108,524.2 seconds running in parallel the 2 cores of the Intel Core i5-3320M CPU@2.60GHz with 8.00GB of installed RAM. It produced the following results:

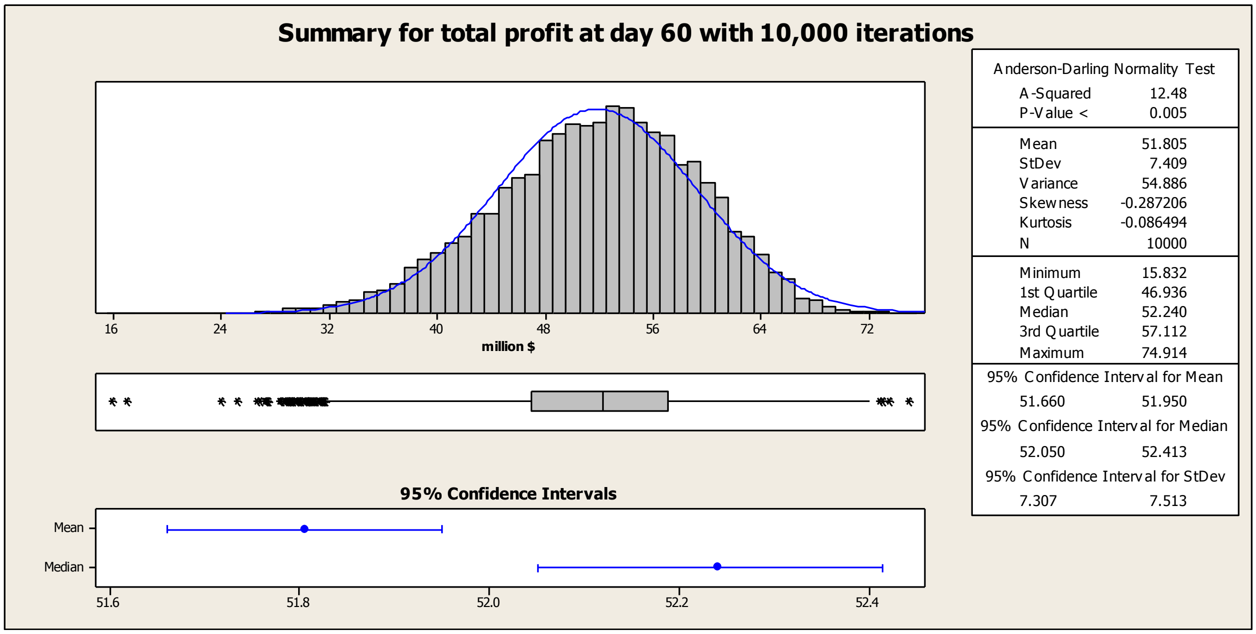

FIGURE 28: GRAPHICAL AND STATISTICAL SUMMARY FOR TOTAL PROFIT AFTER DAY 60 WITH 10,000 ITERATIONS

In order to compare both samples (10,000 iterations -Figure 28- and 1,000 iterations -Figure 17-) it valuable to compare their variances. This is possible through an analysis of variance (ANOVA) using Minitab 16:

One-way ANOVA: t(60)10,000, t(60)_parfor

```
Source      DF        SS    MS     F      P
Factor       1       1.2   1.2  0.02  0.883
Error    10998  603826.3  54.9
Total    10999  603827.5

S = 7.410   R-Sq = 0.00%   R-Sq(adj) = 0.00%


                                    Individual 95% CIs For Mean Based on
                                    Pooled StDev
Level              N    Mean  StDev  --------+---------+---------+---------+-
t(60)10,000    10000  51.805  7.409                (-----*-----)
t(60)_parfor    1000  51.769  7.421  (------------------*------------------)
                                     --------+---------+---------+---------+-
                                         51.50     51.75     52.00     52.25

Pooled StDev = 7.410


Grouping Information Using Tukey Method

                   N    Mean  Grouping
t(60)10,000    10000  51.805  A
```

```
t(60)_parfor   1000  51.769  A

Means that do not share a letter are significantly different.


Tukey 95% Simultaneous Confidence Intervals
All Pairwise Comparisons

Individual confidence level = 95.00%


t(60)10,000 subtracted from:

              Lower  Center  Upper  ---------+---------+---------+---------+
t(60)_parfor  -0.518  -0.036  0.446     (-------------*-------------)
                                    ---------+---------+---------+---------+
                                          -0.35      0.00      0.35      0.70


Grouping Information Using Fisher Method

                 N    Mean  Grouping
t(60)10,000   10000  51.805  A
t(60)_parfor   1000  51.769  A

Means that do not share a letter are significantly different.


Fisher 95% Individual Confidence Intervals
All Pairwise Comparisons

Simultaneous confidence level = 95.00%


t(60)10,000 subtracted from:

              Lower  Center  Upper  ---------+---------+---------+---------+
t(60)_parfor  -0.518  -0.036  0.446     (-------------*-------------)
                                    ---------+---------+---------+---------+
                                          -0.35      0.00      0.35      0.70
```

The conclusion from Tukey method and Fisher method (provided by Minitab 16) is that the variances from both samples are not significantly different. The analysis of variances shows a pvalue that is bigger than the defined α (0.05) therefore both variances cannot be considered different. Now that is established that both simulations produced variances that are not statistically different, it is possible to conduct the comparison using t-test.

Two-Sample T-Test and CI: t(60)_parfor, t(60)10,000

```
Two-sample T for t(60)_parfor vs t(60)10,000

                 N   Mean  StDev  SE Mean
t(60)_parfor   1000  51.77   7.42     0.23
t(60)10,000   10000  51.81   7.41    0.074
```

```
Difference = mu (t(60)_parfor) - mu (t(60)10,000)
Estimate for difference:  -0.036
95% CI for difference:  (-0.518, 0.446)
T-Test of difference = 0 (vs not =): T-Value = -0.15  P-Value = 0.883  DF =
     10998
Both use Pooled StDev = 7.4097
```

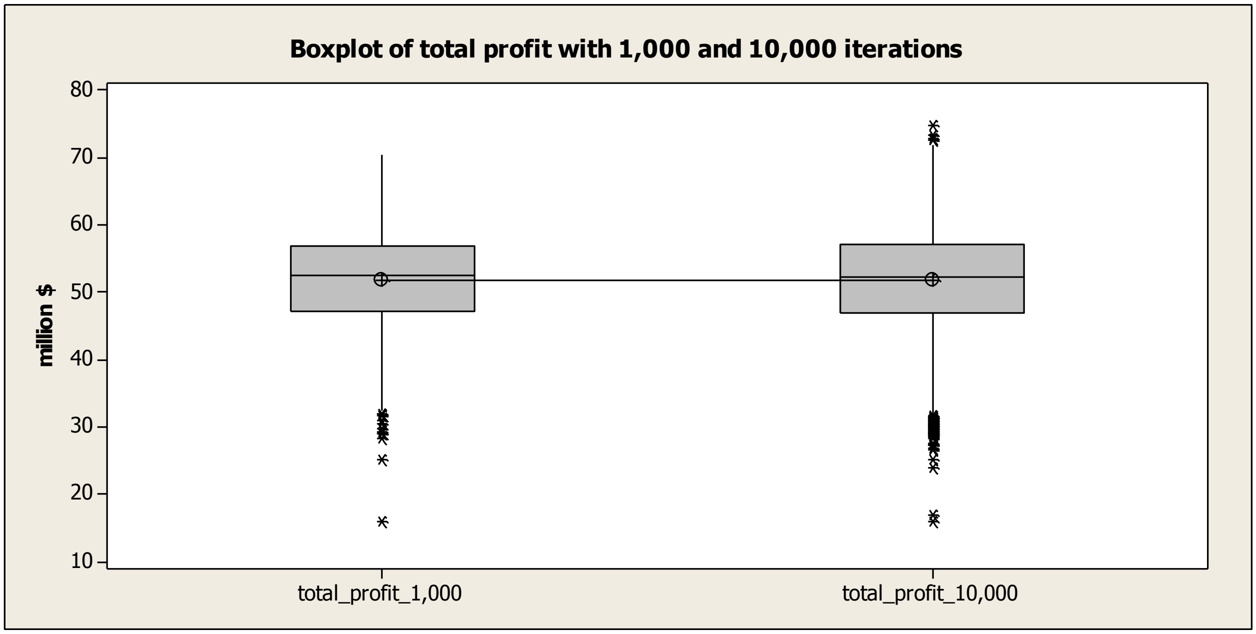

FIGURE 29: BOX PLOTS FOR 1,000 AND 10,000 ITERATIONS

Since the comparison from the t-test has a pvalue that is bigger than the defined α (0.05) it is possible to conclude that the null hypothesis that both samples are equal cannot be rejected, and therefore they cannot be considered statistically different.

Crude price variation: The variation of this variable was defined using a sample of information from the New York Mercantile Exchange (NYMEX). The original link is:

http://www.eia.gov/dnav/pet/pet_pri_fut_s1_d.htm.

From that link, the crude oil from Contract 1 was obtained from the following link:

http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=RCLC1&f=D

The sheet "Data 1" from that table gives the prices from different products. For this research, Cushing, OK Crude Oil Future Contract 4 (Dollars per Barrel) was used to calculate the standard deviation as a percentage (3.8%).

TABLE 5: CRUDE PRICE VARIATION SAMPLE

#	Date	Cushing, OK Crude Oil Future Contract 4 (Dollars per Barrel)
1	May 03, 2013	95.43
2	May 06, 2013	96.11
3	May 07, 2013	95.59
4	May 08, 2013	96.47
5	May 09, 2013	96.41
6	May 10, 2013	96.03
7	May 13, 2013	95.2
8	May 14, 2013	94.37
9	May 15, 2013	94.49
10	May 16, 2013	95.42
11	May 17, 2013	96.25
12	May 20, 2013	96.85
13	May 21, 2013	96.17
14	May 22, 2013	94.06
15	May 23, 2013	94
16	May 24, 2013	93.98
17	May 28, 2013	94.88
18	May 29, 2013	93.15
19	May 30, 2013	93.57
20	May 31, 2013	92.12
21	Jun 03, 2013	93.55
22	Jun 04, 2013	93.51
23	Jun 05, 2013	93.86
24	Jun 06, 2013	94.8
25	Jun 07, 2013	96.05
26	Jun 10, 2013	95.84
27	Jun 11, 2013	95.4

28	Jun 12, 2013	95.85
29	Jun 13, 2013	96.66
30	Jun 14, 2013	97.82
31	Jun 17, 2013	97.87
32	Jun 18, 2013	98.43
33	Jun 19, 2013	98.31
34	Jun 20, 2013	94.84
35	Jun 21, 2013	92.74
36	Jun 24, 2013	93.91
37	Jun 25, 2013	94.04
38	Jun 26, 2013	94.22
39	Jun 27, 2013	95.61
40	Jun 28, 2013	95.1
41	Jul 01, 2013	96.23
42	Jul 02, 2013	97.41
43	Jul 03, 2013	98.97
44	Jul 05, 2013	100.45
45	Jul 08, 2013	100.55
46	Jul 09, 2013	100.8
47	Jul 10, 2013	102.29
48	Jul 11, 2013	101.59
49	Jul 12, 2013	102.66
50	Jul 15, 2013	102.99
51	Jul 16, 2013	102.95
52	Jul 17, 2013	103.58
53	Jul 18, 2013	104.62
54	Jul 19, 2013	104.35
55	Jul 22, 2013	103.92
56	Jul 23, 2013	102.64
57	Jul 24, 2013	101.26
58	Jul 25, 2013	101.78
59	Jul 26, 2013	101.42
60	Jul 29, 2013	101.41
61	Jul 30, 2013	100.49
62	Jul 31, 2013	101.65
63	Aug 01, 2013	103.71
64	Aug 02, 2013	103.23
65	Aug 05, 2013	103.08
66	Aug 06, 2013	102.11
67	Aug 07, 2013	101.49
68	Aug 08, 2013	100.53
69	Aug 09, 2013	102.38
70	Aug 12, 2013	102.97
71	Aug 13, 2013	103.78
72	Aug 14, 2013	103.97
73	Aug 15, 2013	104.64
74	Aug 16, 2013	105.17
75	Aug 19, 2013	104.73
76	Aug 20, 2013	103.59
77	Aug 21, 2013	101.25
78	Aug 22, 2013	101.96
79	Aug 23, 2013	103.11
80	Aug 26, 2013	102.78
81	Aug 27, 2013	105.37
82	Aug 28, 2013	106.18

83	Aug 29, 2013	104.95
84	Aug 30, 2013	104.17
85	Sep 03, 2013	104.83
86	Sep 04, 2013	103.92
87	Sep 05, 2013	104.79
88	Sep 06, 2013	106.24
89	Sep 09, 2013	105.15
90	Sep 10, 2013	103.35
91	Sep 11, 2013	103.54
92	Sep 12, 2013	104.54
93	Sep 13, 2013	104.37
94	Sep 16, 2013	103.35
95	Sep 17, 2013	101.99
96	Sep 18, 2013	104.43
97	Sep 19, 2013	103.11
98	Sep 20, 2013	102.49
99	Sep 23, 2013	100.52
100	Sep 24, 2013	100.41
101	Sep 25, 2013	100.13
102	Sep 26, 2013	100.75
103	Sep 27, 2013	100.46
104	Sep 30, 2013	100.16
105	Oct 01, 2013	100.03
106	Oct 02, 2013	101.7
107	Oct 03, 2013	101.25
108	Oct 04, 2013	101.85
109	Oct 07, 2013	101.56
110	Oct 08, 2013	102.04
111	Oct 09, 2013	100.32
112	Oct 10, 2013	101.91
113	Oct 11, 2013	101.45
114	Oct 14, 2013	101.78
115	Oct 15, 2013	100.86
116	Oct 16, 2013	101.9
117	Oct 17, 2013	100.12
118	Oct 18, 2013	100.54
119	Oct 21, 2013	99.49
120	Oct 22, 2013	98.31

In order to simulate the effect of this specific variability, the Case 0: Original setting without uncertainty (no variation %) was considered only modifying the Refinery Economics subsystem and implemented in parallel as described in Equation 9: Matlab code for parallel computation. This is possible because IRIS defines these variable as Constant Sample Time, specifying a constant (Inf) sample time, which in Matlab means that “the block executes only once during model initialization” (The MathWorks, Inc, 2013). The base code was the one defined for Case 0 (no variation %), only modifying specifically the lines related to the Refinery Economics subsystem of IRIS:

Code modification for Case 1a:

```
%REFINERY ECONOMICS:
set_param('iris/Refinery Economics', 'pdtpricevariationpercent',
num2str(max(0,normrnd(1,0.2))*3.8), 'opcost',
num2str(max(0,normrnd(1,0.2))*2), 'invcost',
num2str(max(0,normrnd(1,0.2))*0.05));
```

Code modification for Case 1b:

```
%SALES:
set_param('iris/Sales', 'magnitudeofdemandincrease',
num2str(max(0,normrnd(1,0.2))*2), 'demandvariance',
num2str(max(0,normrnd(1,0.2))*25));
```

The comparison presented in Case 1: Focalized Uncertainty can be also observed in a box plot representation. When adding to the comparison the referential Case 0 (no variation %), it is visually clear that the Case 1a has worse expected performance than Case 1b, even when Case 1b has a lower minimum value (or more risk). This is an important difference that means that Case 1a is worse expected performance (in total profit after 120 days) but less risk and Case 1b is better expected performance but more risk.

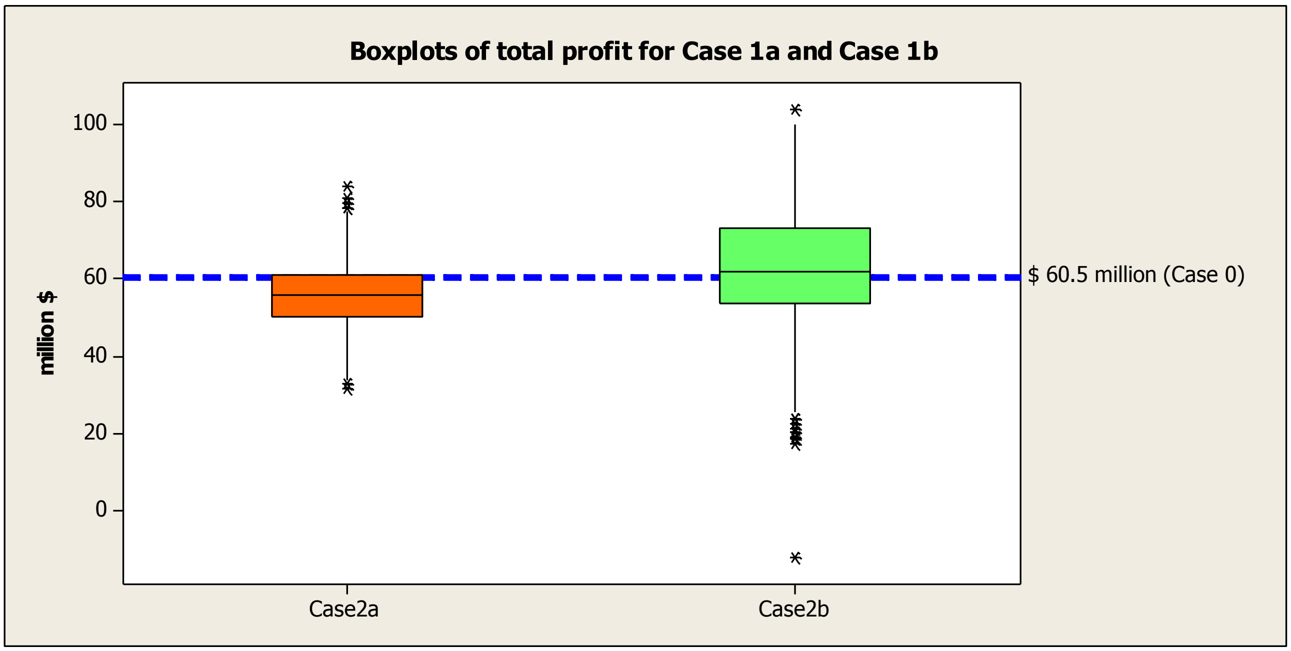

FIGURE 30: BOXPLOTS OF TOTAL PROFIT FOR CASE 1A AND CASE 1B

The variances of these two cases are also significantly different. An Analysis of Variance performed using the statistical software Minitab 16 using both Tukey method and Fisher method illustrates this conclusion:

One-way ANOVA: t(120)_Case1a, t(120)_Case1b

```
Source     DF      SS     MS       F      P
Factor      1   19270  19270  137.36  0.000
Error    1998  280293    140
Total    1999  299562
S = 11.84   R-Sq = 6.43%   R-Sq(adj) = 6.39%
```

```
                                   Individual 95% CIs For Mean Based on
                                   Pooled StDev
Level              N   Mean  StDev  -----+---------+---------+---------+----
t(120)_Case1a  1000  55.76   8.42  (---*--)
t(120)_Case1b  1000  61.97  14.48                               (---*---)
                                   -----+---------+---------+---------+----
                                     56.0      58.0      60.0      62.0

Pooled StDev = 11.84

Grouping Information Using Tukey Method

                  N   Mean  Grouping
t(120)_Case1b  1000  61.97  A
t(120)_Case1a  1000  55.76  B

Means that do not share a letter are significantly different.

Tukey 95% Simultaneous Confidence Intervals
All Pairwise Comparisons

Individual confidence level = 95.00%


t(120)_Case1a subtracted from:

               Lower  Center  Upper  ----+---------+---------+---------+-----
t(120)_Case1b   5.17    6.21   7.25                          (---*---)
                                     ----+---------+---------+---------+-----
                                       0.0       2.5       5.0       7.5

Grouping Information Using Fisher Method

                  N   Mean  Grouping
t(120)_Case1b  1000  61.97  A
t(120)_Case1a  1000  55.76  B

Means that do not share a letter are significantly different.

Fisher 95% Individual Confidence Intervals
All Pairwise Comparisons

Simultaneous confidence level = 95.01%


t(120)_Case1a subtracted from:

               Lower  Center  Upper  ----+---------+---------+---------+-----
t(120)_Case1b   5.17    6.21   7.25                          (---*---)
                                     ----+---------+---------+---------+-----
                                       0.0       2.5       5.0       7.5
```

Also, and now that is known that the variances are significantly different, a two sample test can be performed. This test, also conducted using Minitab 16, helps to conclude that both cases 1a and 1b are significantly different (pvalue<α).

Two-Sample T-Test and CI: t(120)_Case1a, t(120)_Case1b

```
Two-sample T for t(120)_Case1a vs t(120)_Case1b

                  N   Mean  StDev  SE Mean
t(120)_Case1a  1000  55.76   8.42     0.27
t(120)_Case1b  1000   62.0   14.5     0.46

Difference = mu (t(120)_Case1a) - mu (t(120)_Case1b)
Estimate for difference:  -6.208
95% CI for difference:  (-7.247, -5.169)
T-Test of difference = 0 (vs not =): T-Value = -11.72  P-Value = 0.000  DF =
     1605
```

APPENDIX 8: UNCERTAINTY VERSUS SENSITIVITY

In order to identify the specific effect that the variation of each of the 14 variables defined in Table 3 for the stochastic simulation, 14 simulations were conducted incorporating uncertainty one variable at a time. This simulation of 14 configurations took 296,419.3 seconds (82 hours) running with a parallel computation in the computer of this research. The variables with a significant correlation between their uncertainty and the Total Profit were compared in a marginal plot as follows:

Using the statistical software Minitab 16 the Pearson correlation values and the p-values were calculated for each of the 14 variable stochastic variations and the final result (Total Profit):

Correlations: a: Tank Volume, Ra: Total Profit varying Tank Volume

```
Pearson correlation of a and Ra = -1.000
P-Value = 0.000
```

Correlations: b: Product price variation percentage (%), Rb: Total Profit varying Product price variation percentage (%)

```
Pearson correlation of b and Rb = -1.000
P-Value = 0.000
```

Correlations: c: Operating cost (K$/kbbl), Rc: Total Profit varying Operating cost (K$/kbbl)

```
Pearson correlation of c and Rc = -1.000
P-Value = 0.000
```

Correlations: d: Product or crude inventory cost (K$/kbbl), Rd: Total Profit varying Product or crude inventory cost (K$/kbbl)

```
Pearson correlation of d and Rd = -1.000
P-Value = 0.000
```

Correlations: e: Magnitude of demand increase, Re: Total Profit varying Magnitude of demand increase

```
Pearson correlation of e and Re = 0.003
P-Value = 0.916
```

Correlations: f: Demand variance percentage (%), Rf: Total Profit varying Demand variance percentage (%)

```
Pearson correlation of f and Rf = -0.996
P-Value = 0.000
```

Correlations: g: Crude amount variation percentage (%), Rg: Total Profits varying Crude amount variation percentage (%)

```
Pearson correlation of g and Rg = 1.000
P-Value = 0.000
```

Correlations: h: Crude price variation percentage (%), Rh: Total Profit varying Crude price variation percentage (%)

```
Pearson correlation of h and Rh = -0.017
P-Value = 0.592
```

Correlations: k: Disruption occurrence seed, Rk: Total Profit varying Disruption occurrence seed

```
Pearson correlation of k and Rk = -0.007
P-Value = 0.834
```

Correlations: l: Disruption magnitude seed, Rl: Total Profit varying Disruption magnitude seed

```
Pearson correlation of l and Rl = 0.107
P-Value = 0.001
```

Correlations: m: Pumping rate (kbbl/hr), Rm: Total Profit varying Pumping rate (kbbl/hr)

```
Pearson correlation of m and Rm = -0.715
P-Value = 0.000
```

Correlations: n: Reformer yield variation (%), Rn: Total Profit varying Reformer yield variation (%)

```
Pearson correlation of n and Rn = -0.032
P-Value = 0.308
```

Correlations: o: Cracker yield variation (%), Ro: Total Profit varying Cracker yield variation (%)

```
Pearson correlation of o and Ro = -0.020
P-Value = 0.534
```

Correlations: p: CDU yield variation (%), Rp: Total Profit varying CDU yield variation (%)

```
Pearson correlation of p and Rp = 0.020
P-Value = 0.528
```

Only the variations of the following variables were considered significant for the final result of the dynamic simulation (Total Profit):

- Tank Volume (kbbl): Pearson correlation of a and Ra = -1.000, P-Value = 0.000
- Product price variation percentage: Pearson correlation of b and Rb = -1.000, P-Value = 0.000
- Operating cost (K$/kbbl): Pearson correlation of c and Rc = -1.000, P-Value = 0.000
- Product or crude inventory cost (K$/kbbl): Pearson correlation of d and Rd = -1.000, P-Value = 0.000
- Crude amount variation percentage (%): Pearson correlation of g and Rg = 1.000, P-Value = 0.000
- Demand variance percentage (%): Pearson correlation of f and Rf = -0.996, P-Value = 0.000
- Pumping rate (kbbl/hr): Pearson correlation of m and Rm = -0.715, P-Value = 0.000

The relationship between the uncertainty and the total profit result of the system after 120days is illustrated in the following graphs:

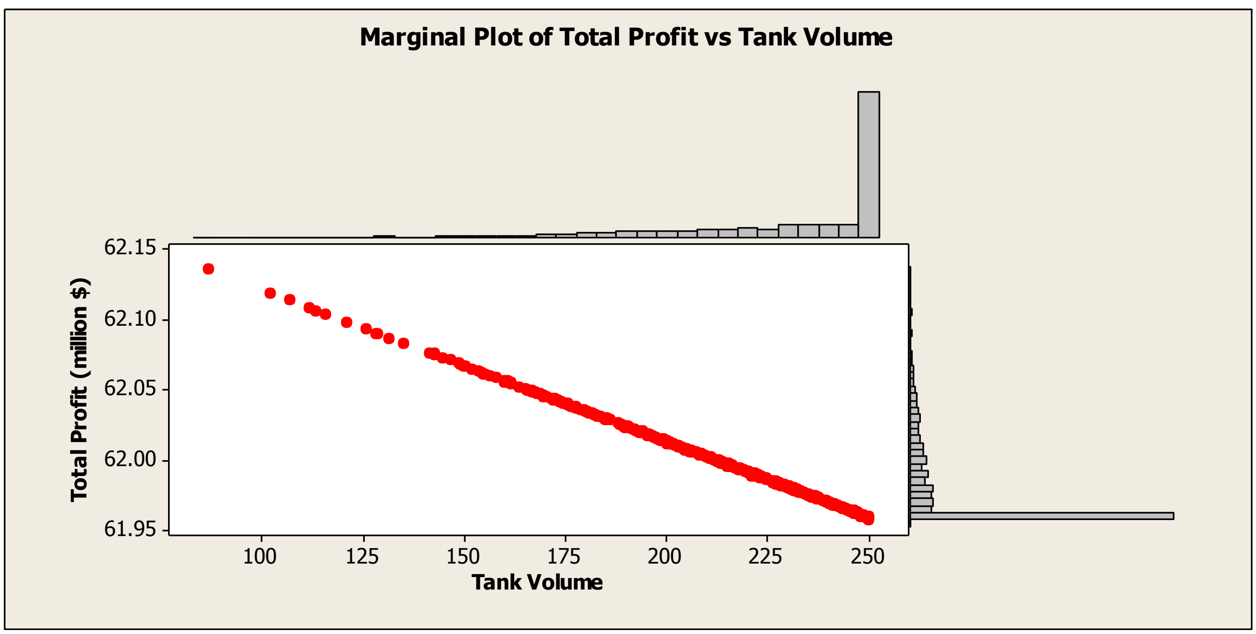

FIGURE 31: TANK VOLUME UNCERTAINTY VERSUS TOTAL PROFITS

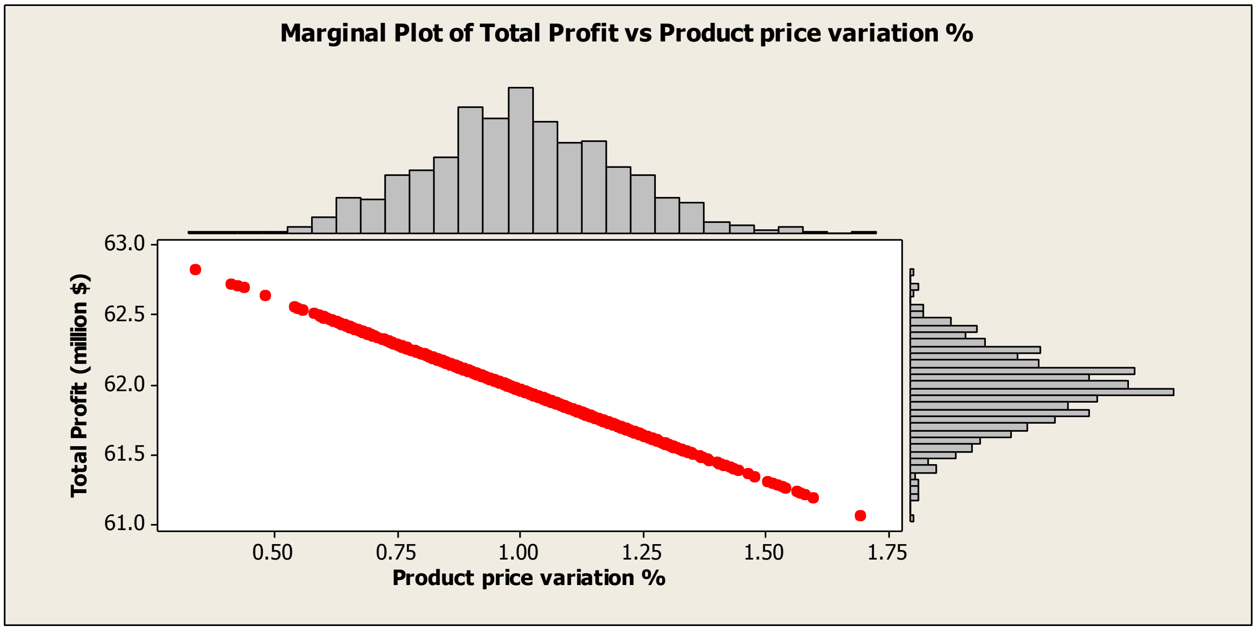

FIGURE 32: PRODUCT PRICE VARIATION % VERSUS TOTAL PROFITS

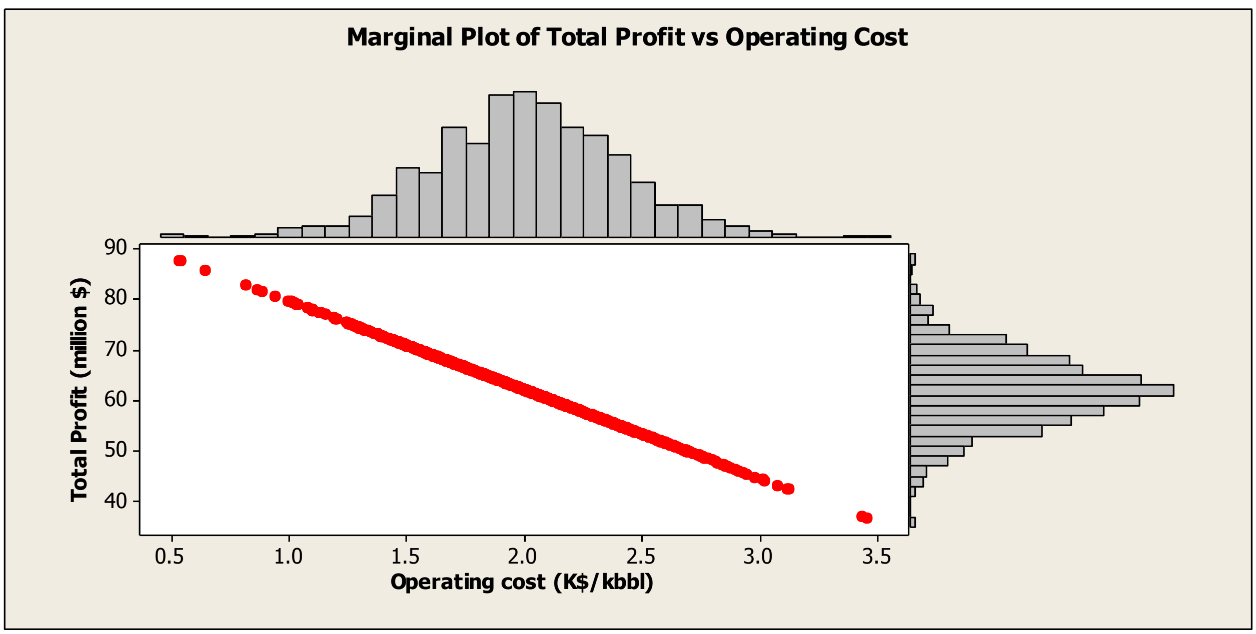

FIGURE 33: OPERATING COST VERSUS TOTAL PROFITS

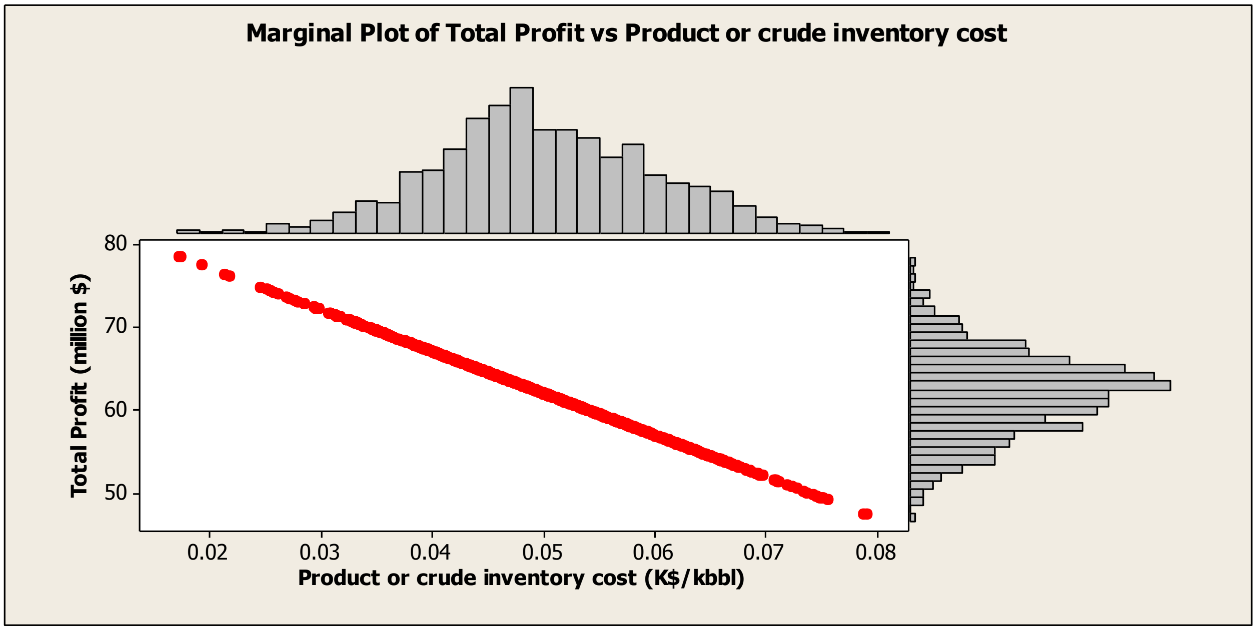

FIGURE 34: PRODUCT OR CRUDE INVENTORY COST VERSUS TOTAL PROFITS

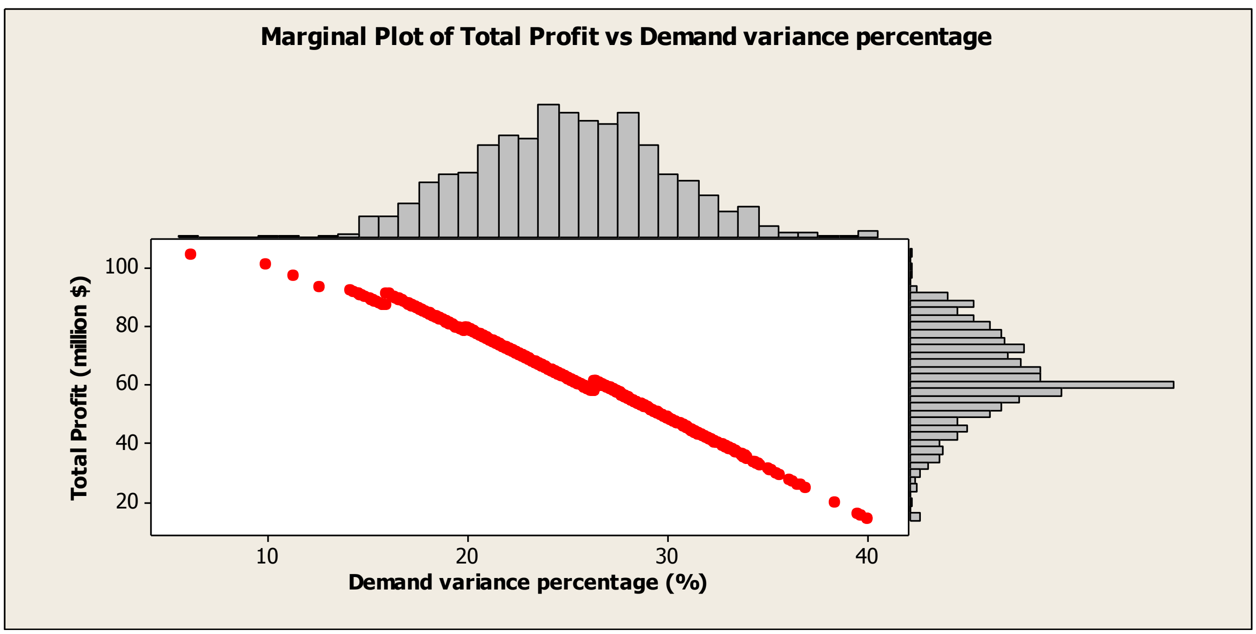

FIGURE 35: TOTAL PROFIT VS DEMAND VARIANCE %

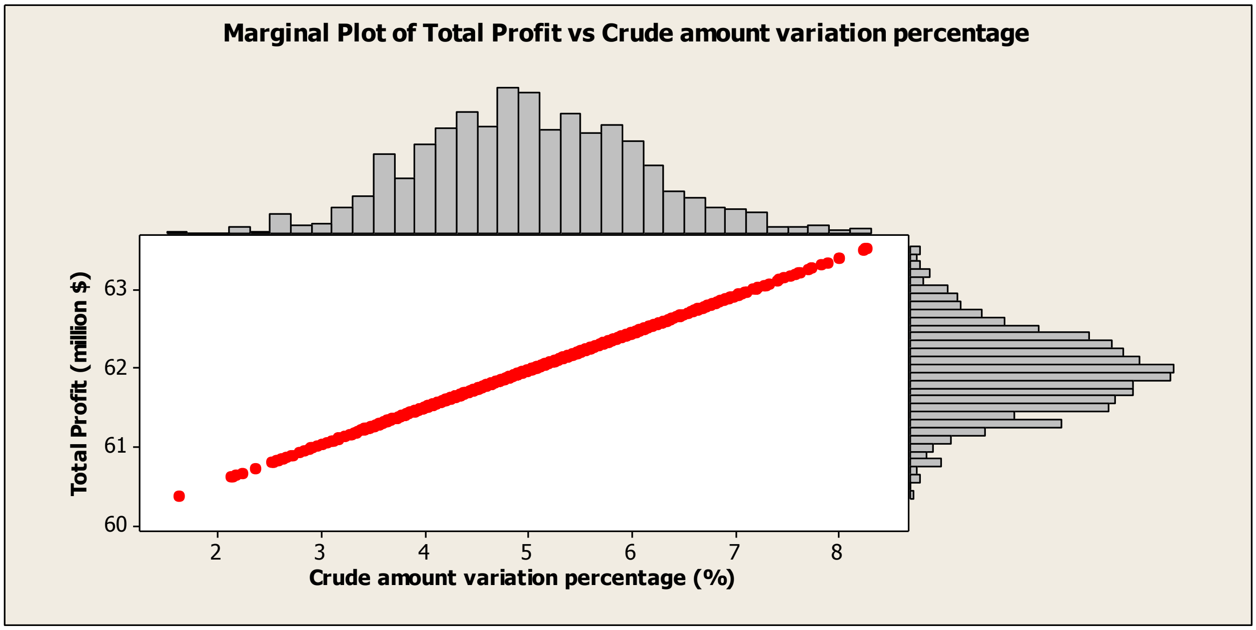

FIGURE 36: TOTAL PROFIT VS CRUDE AMOUNT VARIATION %

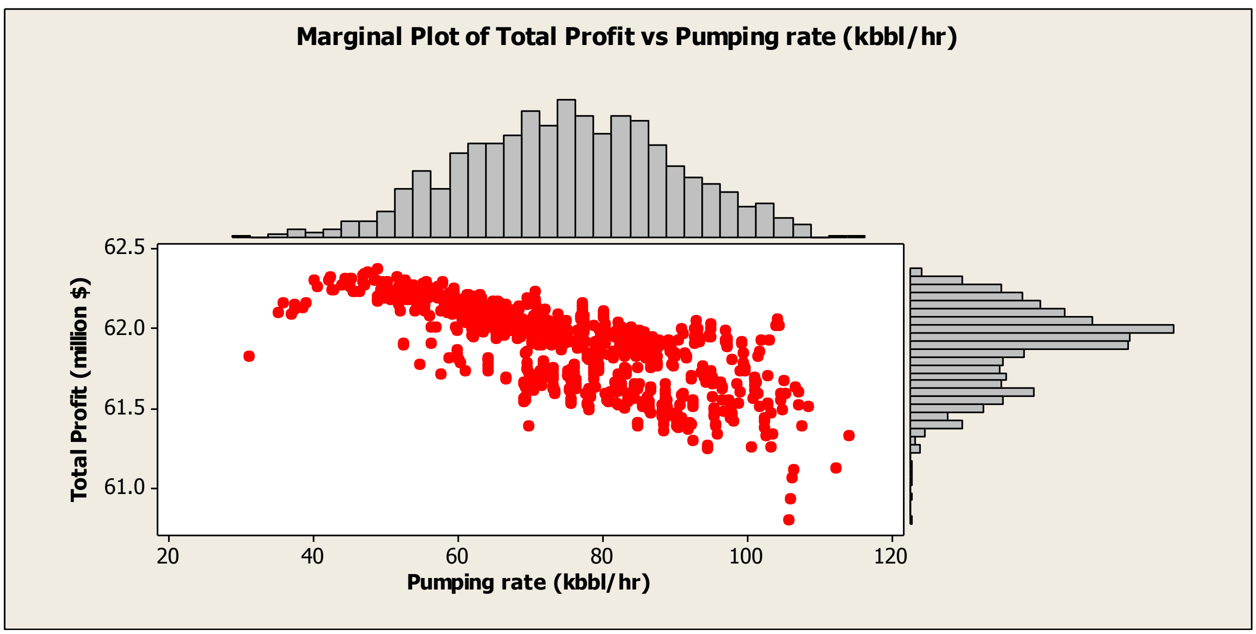

FIGURE 37: TOTAL PROFIT VS PUMPING RATE

Iris is a dynamic model and therefore it is impacted by the dominance of internal feedback loops. One example of these dominances is the effect produced by the fact that the original configuration (base case from original paper) doesn't have enough capacity to achieve the demand. The problem with not achieving demand is that in IRIS, this is heavily penalized.

To test this concept, the configuration without uncertainty (Case 0) was run for 240 days. The crude inventory is observable in Figure 38:

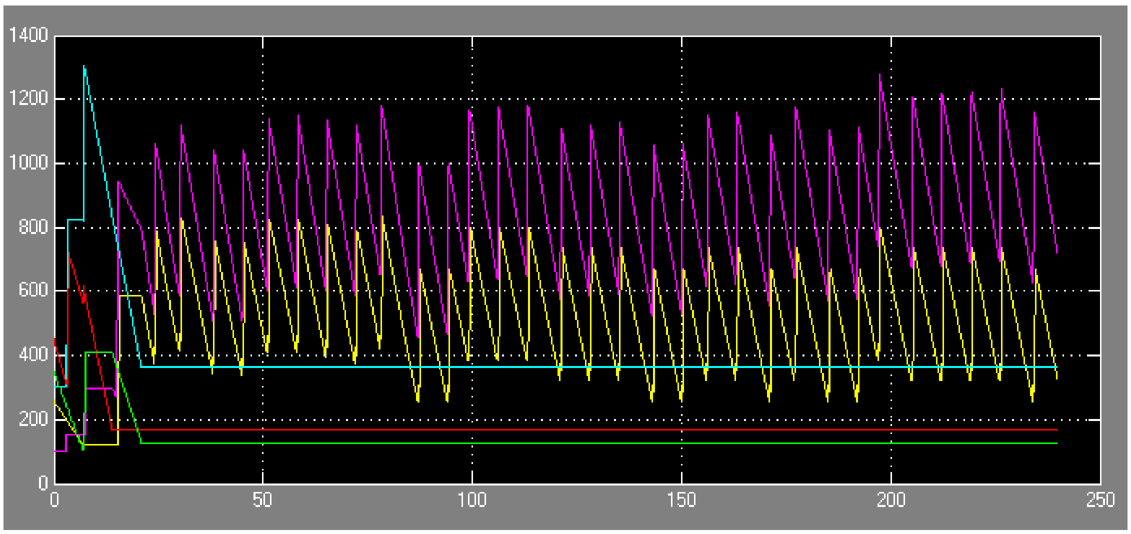

FIGURE 38: CRUDE INVENTORY OVER 240 DAYS

The plant quickly focuses on achieving the demand of two crudes (red and blue in Figure 39)

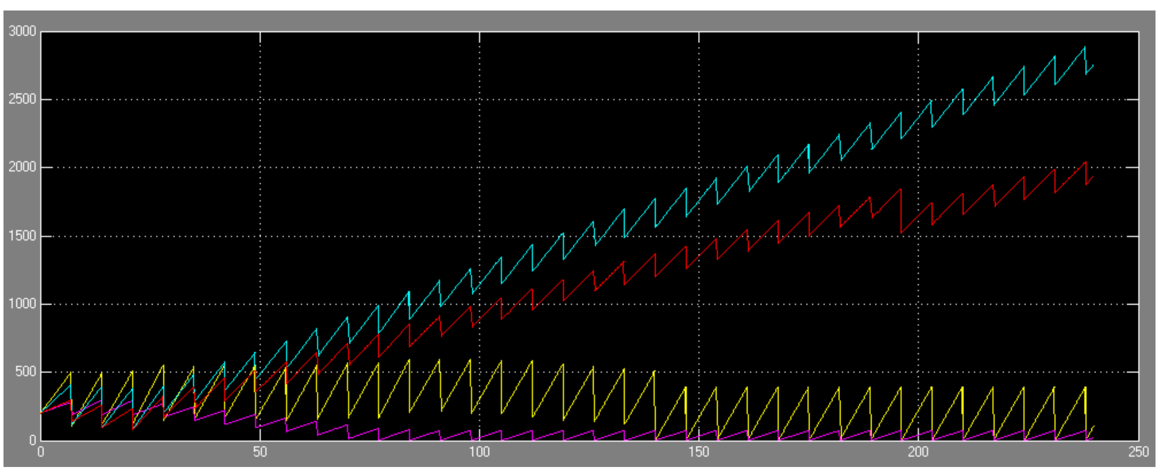

FIGURE 39: PRODUCT INVENTORY PROFILE FOR 240 DAYS

The customer satisfaction of the two products that are not supplied as demanded drops drastically (first two graphs in Figure 40.

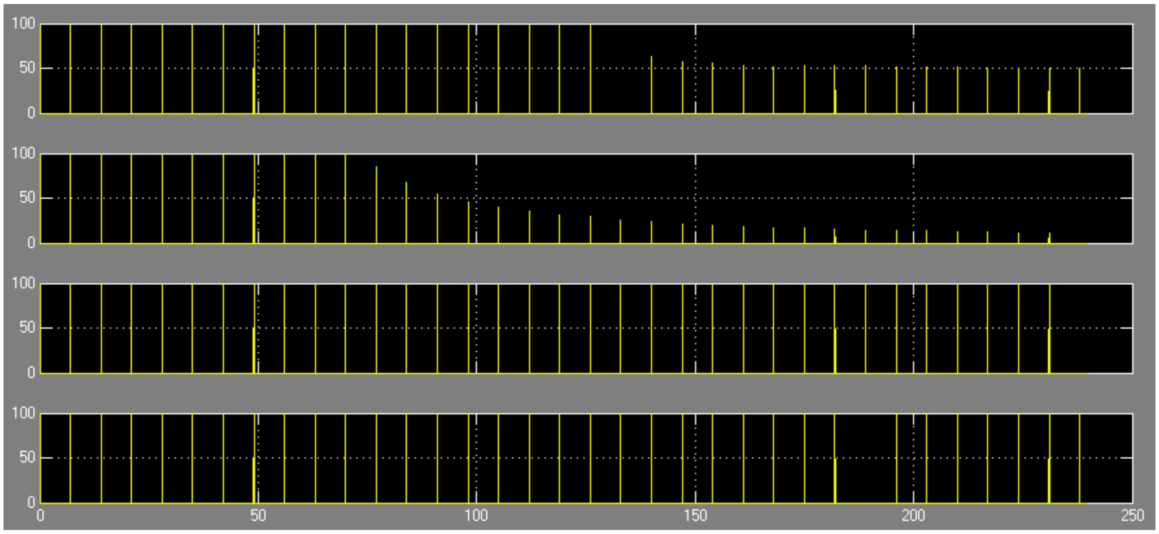

FIGURE 40: CUSTOMER SATISFACTION FOR CASE 0 OVER 240 DAYS

The loop of revenue produces a constant revenue over the whole period of the simulation:

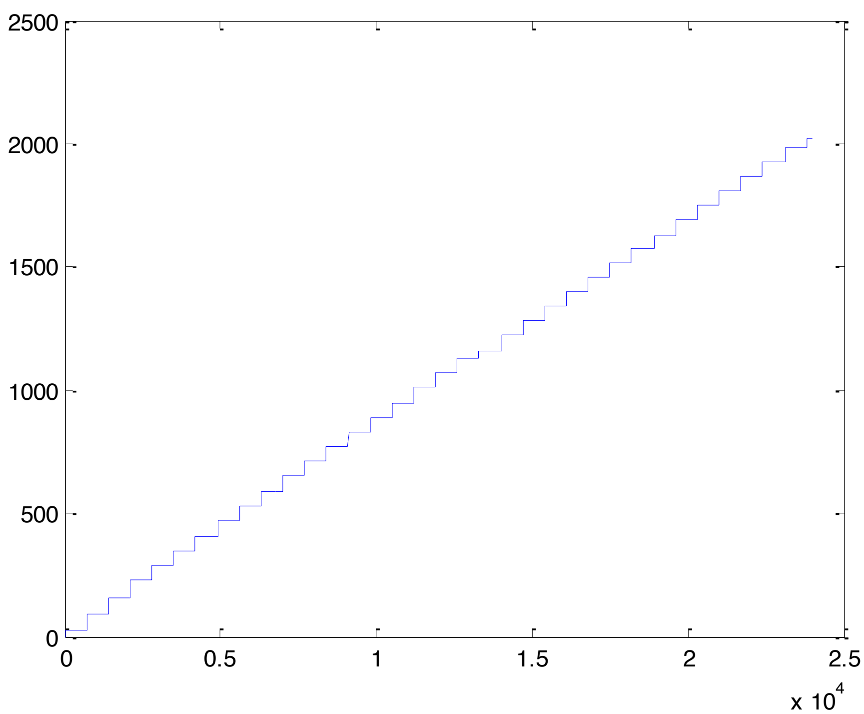

FIGURE 41: REVENUE FOR CASE 0 OVER 240 DAYS

The problem is that the penalty for not achieving the demand of two products grows exponentially:

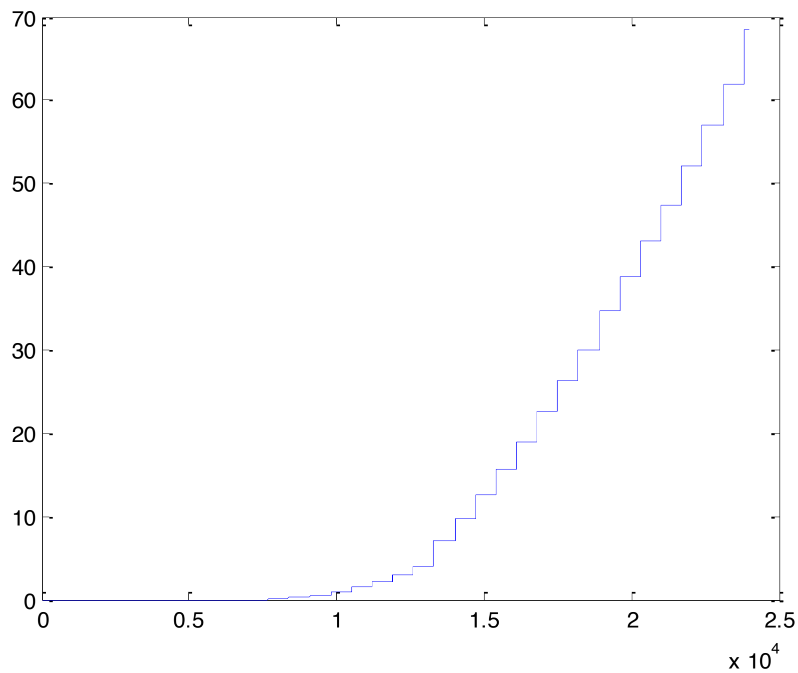

FIGURE 42: PENALTY FOR NOT ACHIEVING DEMAND

This exponential behavior produces a change in dominance of the feedback loops, moving the cash flow to negative around day 140:

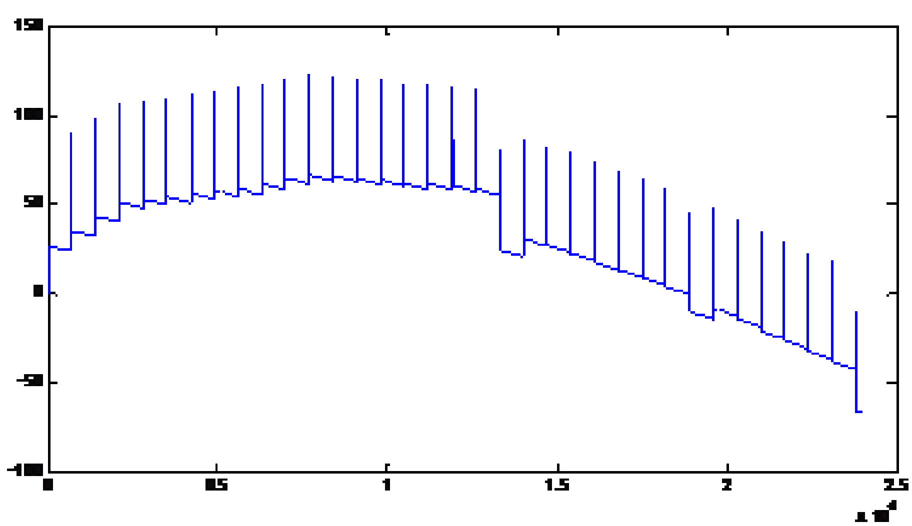

FIGURE 43: TOTAL PROFIT FOR CASE 0 OVER 240 DAYS

APPENDIX 10: CONFIGURING IRIS FOR STEADY INCREMENT OF TOTAL PROFITS

The original configuration of IRIS (Appendix 1: IRIS settings and results from original paper) considers several complex feedback loops and structures that produce non-linear behaviors like the one observed by the goal-seeking behavior of Total Profits in Figure 10 for case 0. In order to explore the effects of reducing that complexity diminishing the relevance of feedback loops like the one that is penalizing the system for not delivering what is demanded, the configuration was modified.

This modification considered a combination of aspects to diminish the effect of the negative feedback loop that is producing the goal-seeking behavior. The changes in the values of the settings were:

- Demand switch off
- Mean Product demand (kbbl/day) = [355 90 265 145]/7
- Procurement policy choice: 2
- Scheduling policy choice: 2
- Crude storage capacity limit (kbbl) = [12500 12500 12500 12500 12500]
- Maximum throughput (kbbl/day) = 400

Once the model was reconfigured with these new settings, the code used to run the simulation was:

EQUATION 10: CODE FOR SIMULATION WITH STEADY INCREMENT IN PROFITS

```
%Parallel computation of the loop:
matlabpool
tic;
iterations=1000;
outputpc = cell(1,iterations);

parfor J = 1:1:iterations
    load_system('iris') %Load system without viewing
    %DEFINING VARIABLES:
```

```
    %CONTROL PANEL
    a=normrnd(1,0.2);
      set_param('iris/Control Panel', 'tankvolume',
num2str(min(max(0,a)*250,250)), 'maxthruput',
num2str(max(0,normrnd(1,0.2))*400));
    %REFINERY ECONOMICS:
      set_param('iris/Refinery Economics', 'pdtpricevariationpercent',
num2str(max(0,normrnd(1,0.2))*1), 'opcost', num2str(max(0,normrnd(1,0.2))*2),
'invcost', num2str(max(0,normrnd(1,0.2))*0.05));
    %SALES:
      set_param('iris/Sales', 'magnitudeofdemandincrease',
num2str(max(0,normrnd(1,0.2))*2), 'demandvariance',
num2str(max(0,normrnd(1,0.2))*25));
    %SUPPLIERS:
      set_param('iris/Suppliers', 'crudeamountvariationpercent',
num2str(max(0,normrnd(1,0.2))*5),'crudepricevariationpercent',
num2str(max(0,normrnd(1,0.2))*1),
'supplierseed',num2str(round(max(0,normrnd(1,0.2))*78)), 'magnitudeseed',
num2str(round(max(0,normrnd(1,0.2))*100)));
    %PORT/STORAGE
      set_param('iris/PortStorage', 'pumprate',
num2str(max(0,normrnd(1,0.2))*75));
    %REFORMER
      set_param('iris/Reformer', 'variationpercentage',
num2str(max(0,normrnd(1,0.2))*0.1));
    %CRACKER
      set_param('iris/Cracker', 'percentagevariation',
num2str(max(0,normrnd(1,0.2))*0.1));
    %CDU
      set_param('iris/CDU', 'percentagevariation',
num2str(max(0,normrnd(1,0.2))*0.01));
      simOut=sim('iris','StopTime','480');
      outputpc{J}=simOut.get('simouttest');
      J
end
trunpc=toc
matlabpool close
```

After developing the new configuration with a simplified behavior, the model was tested with deterministic inputs for 3,000 days of time horizon:

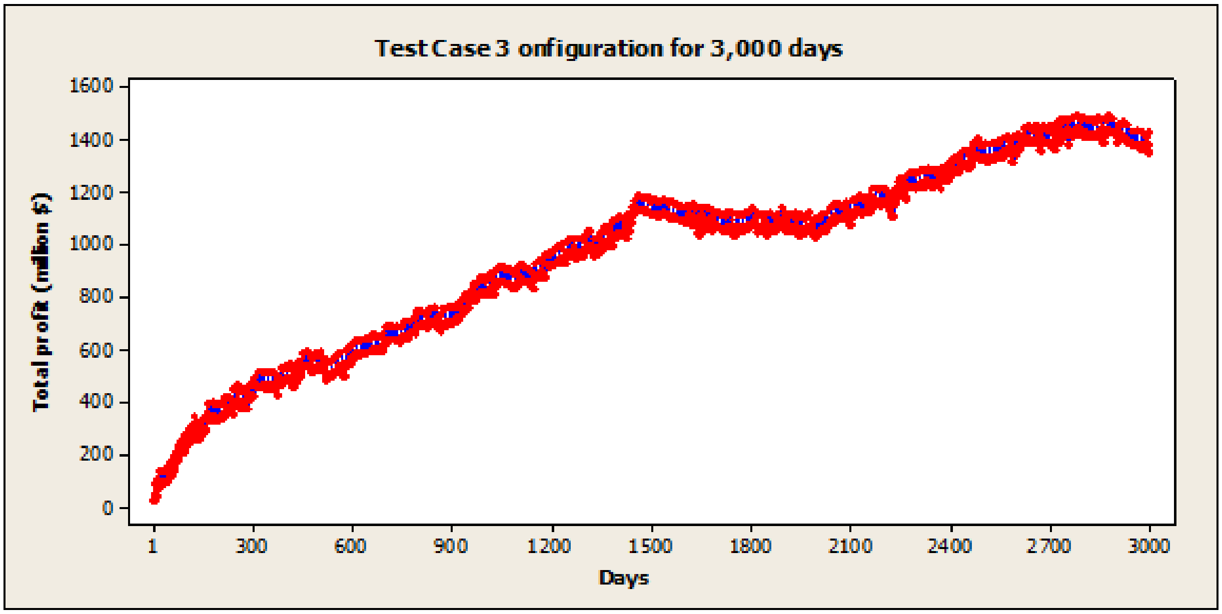

FIGURE 44: TIME SERIES PLOT FOR DETERMINISTIC TEST

The stochastic simulation of this configuration (Equation 10) gives the following behavior of the total profit over the 480 days of time horizon:

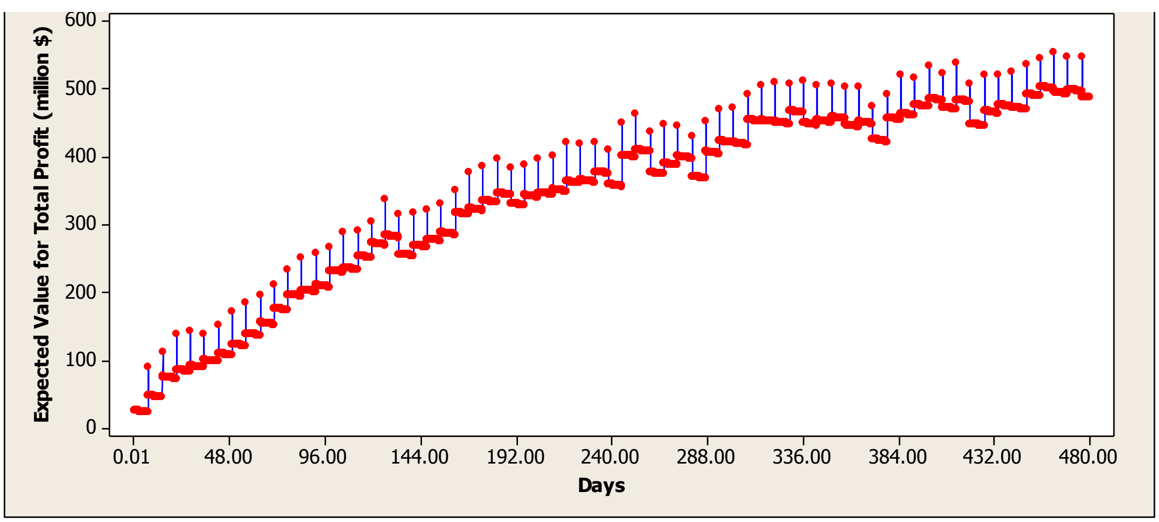

FIGURE 45: TIME SERIES PLOT OF EXPECTED TOTAL PROFIT

Since the behavior of total profit is quite similar to a linear function, the result of the stochastic simulation shouldn't be different from the result of the deterministic simulation. This is confirmed by a graphical and statistical comparison:

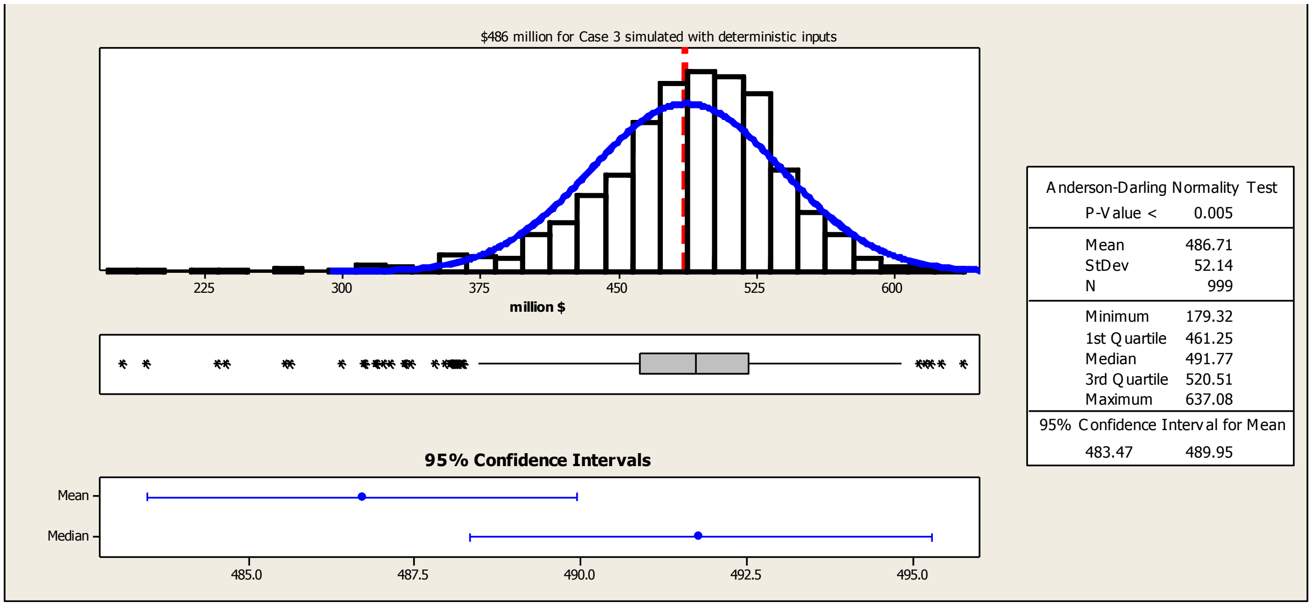

FIGURE 46: HISTOGRAM AND STATISTICS FOR TOTAL PROFIT (STEADY INCREMENT IN PROFIT)

The null hypothesis of a one sample t-test to compare the deterministic case with the expected value from the stochastic simulation cannot be rejected, therefore there is not enough statistical significance to say that they are not different.

One-Sample T: t(480)

```
Test of mu = 486 vs not = 486

Variable     N    Mean   StDev  SE Mean       95% CI            T      P
t(480)     999  486.71   52.14     1.65  (483.47, 489.95)  0.43  0.667
```

Similarly to Figure 20, Figure 47 is showing a propagation of the standard deviation over the time horizon of the simulation.

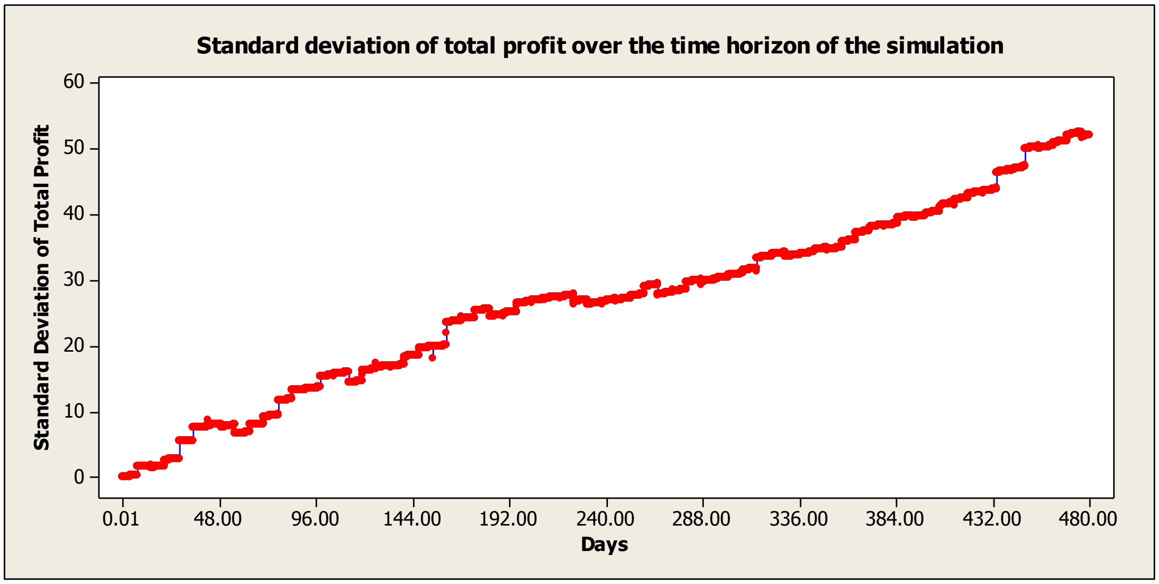

FIGURE 47: PROPAGATION OF THE STANDARD DEVIATION OVER TIME HORIZON

Complete time series plot of Figure 7:

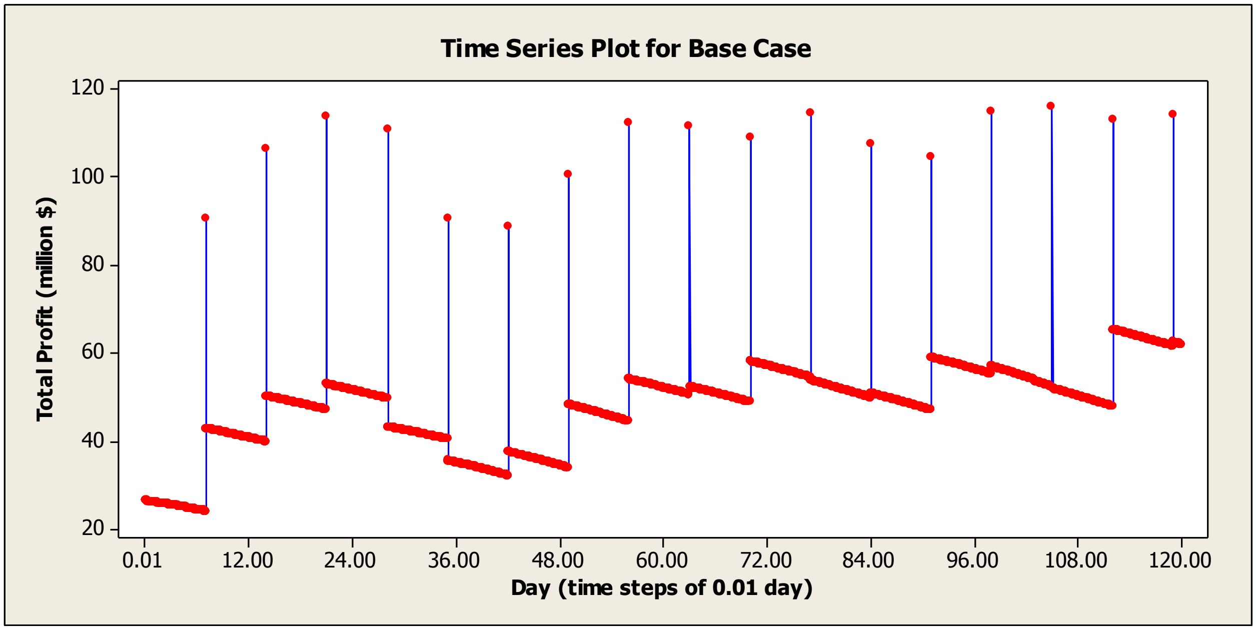

FIGURE 48: TIME SERIES PLOT FOR BASE CASE

Complete time series plot for Figure 10:

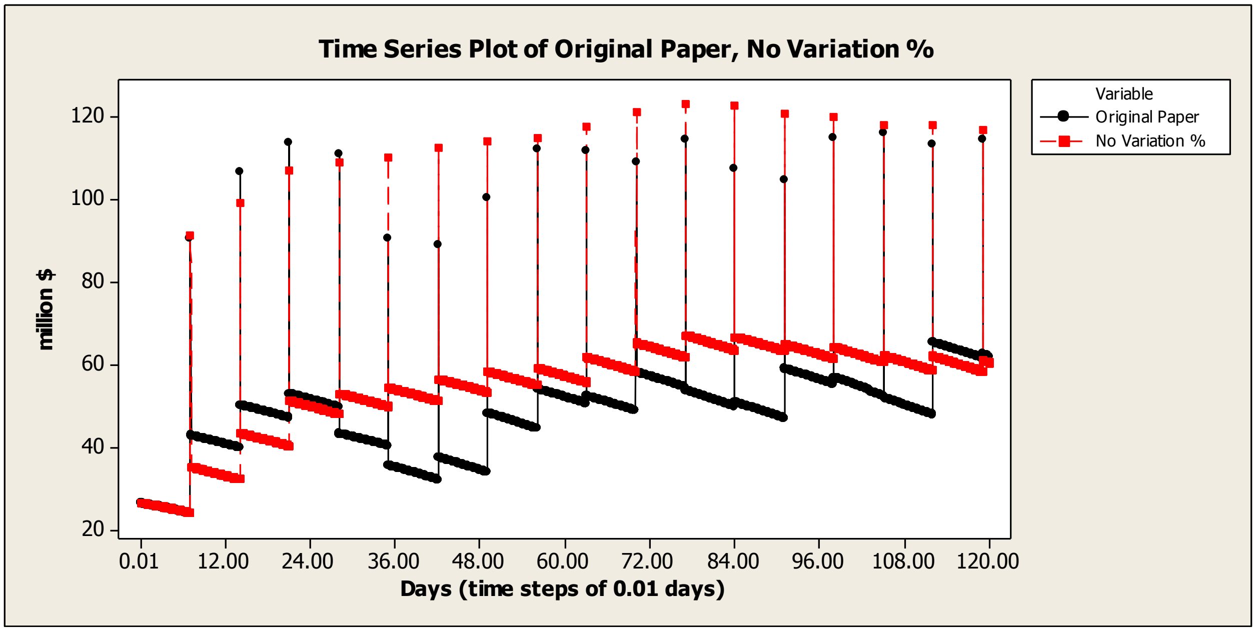

FIGURE 49: TIME SERIES PLOT FOR BASE CASE AND CASE 0

Complete time series plot for Figure 16:

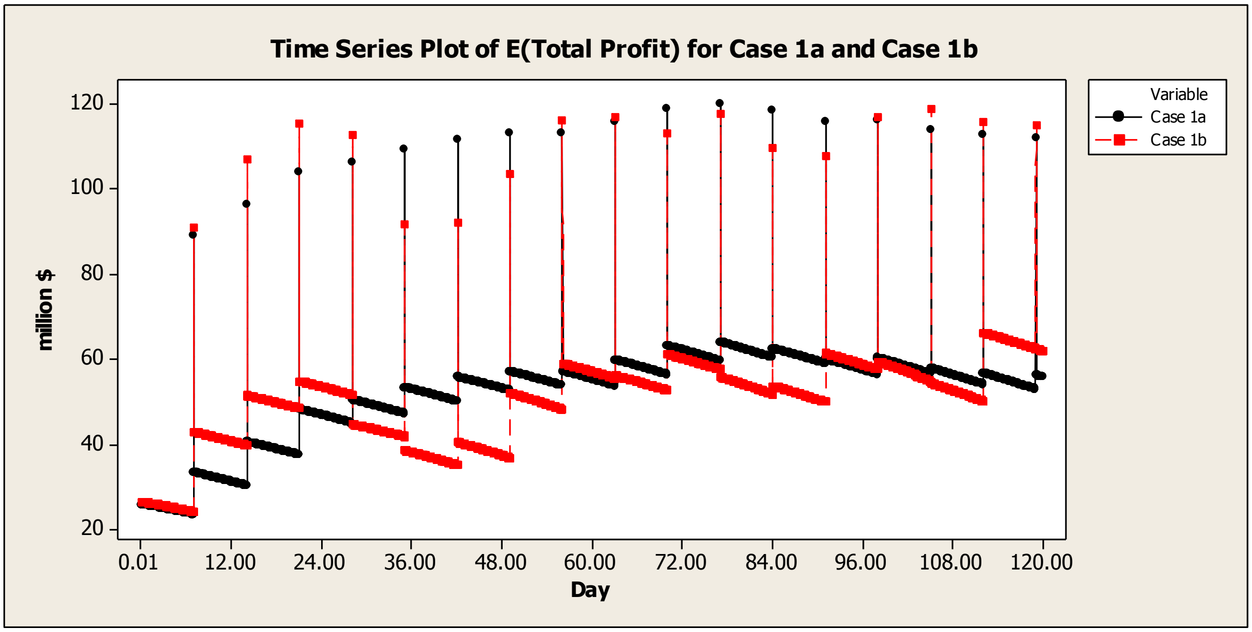

FIGURE 50: COMPLETE COMPARISON OF CASE 1A AND CASE 1B

All statistically significant variables for sensitivity analysis (Figure 21)

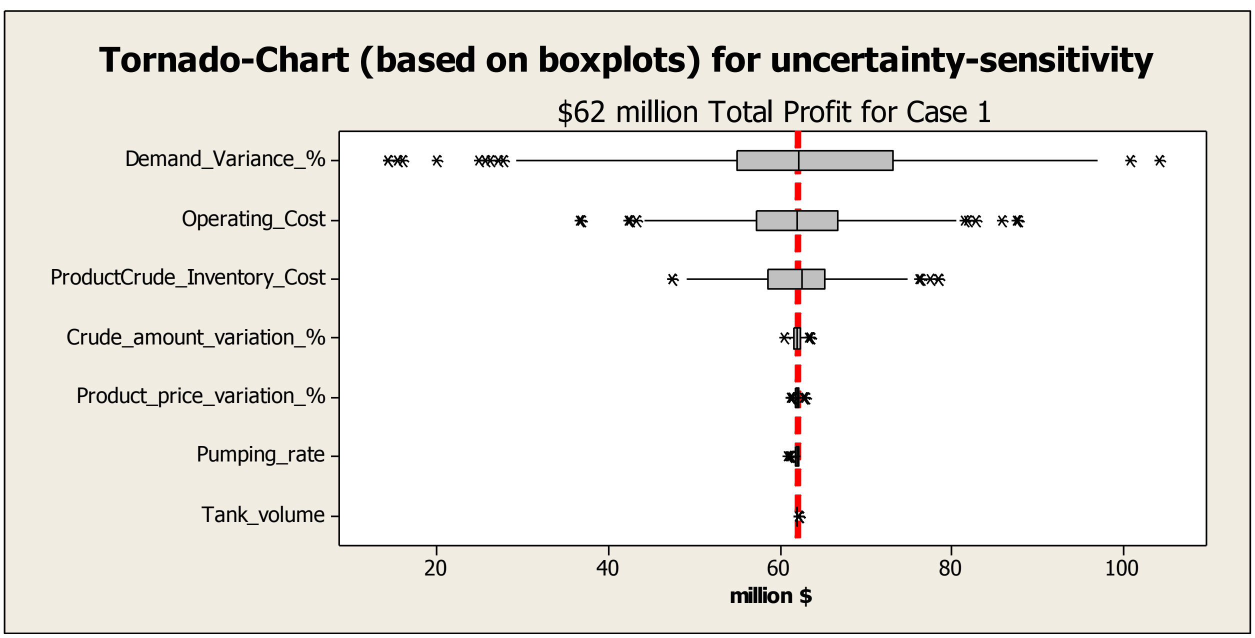

FIGURE 51: SENSISTIVITY ANALYSIS WITH ALL STATISTICALLY SIGNIFICANT VARIABLES

APPENDIX 12: EXECUTIVE SUMMARY

Implication of the Jensen's Inequality for System Dynamic Simulations: Application to a Plug & Play Integrated Refinery Supply Chain Model

EXECUTIVE SUMMARY

By
Juan Esteban Montero: Advanced Analytics, Chevron.
&
Richard de Neufville: MIT Professor of Engineering Systems and Civil and Environmental Engineering

Extracted From Thesis Research Submitted to the MIT System Design and Management Program on February, 2014 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Engineering and Management. **This document is intended for Chevron's internal distribution.**

ABSTRACT

This research studies how critical is the effect of considering uncertainty to a dynamic model because of Jensen's Inequality. This is done using as an example the supply chain of a refinery, which illustrates that the difference between probable and expected results can be significant, arguing that the distributions and probabilities can be dramatically different from the expected-planned value. Secondly, this research discusses that, from the perspective of the dynamics of the system, the mode of behavior can vary considerably as well, leading managers to dissimilar situations and contexts that will inevitably produce different decisions or strategies.

INTRODUCTION

Managers and decision makers often recur to models and analytics as tools to understand the real systems and define optimal decisions and goals. Very commonly these models are complex in almost any possible way, overwhelming the decision makers and making almost impossible to understand intuitively the behavior of the system and to verify that the results are actually the desired solution.

In most cases, models and simulations cannot be tested to prove accuracy. This is the main reason why the mathematical and physical laws have to be considered from the very beginning to construct (or at least try to) good models. One of the mathematical concepts that has to be considered by managers and decision makers is the Jensen's Inequality, which states that the "average of all the possible outcomes associated with uncertain parameters, generally does not equal the value obtained from using the average value of the parameters" (de Neufville, 2012).

Systems Dynamics, and dynamic simulations are highly non-linear (Sterman, 2000) and therefore the effects of Jensen's Inequality, must be considered as a cause of producing the wrong results. This research evaluates the expected value of the result of a dynamic simulation, but also the different modes of behavior produced by those dynamic simulations.

The different effects of uncertainty can be relevant for the decision makers since they could certainly produce different behaviors due to different pressures and contextual factors. These differences could represent dissimilar strategies, decisions, risk perception and other factors that are certainly significant from the point of view of any manager.

The effects of Jensen's inequality on the expected value of a model have been studied in the past. The novelty of this research lays on the focus of the effects of Jensen's Inequality on the mode of behavior of a dynamic system.

RESEARCH OBJECTIVES

Dynamic systems are often simulated using deterministic inputs. These kinds of simulations are highly non-linear (Sterman, 2000) and therefore the effects of Jensen's Inequality, must be considered as a cause of producing the wrong results.

Main hypothesis for the thesis research:

Not considering uncertainty in the inputs of a dynamic simulation produces the wrong results.

Research questions:

- Does Jensen's inequality alter the expected value of the result of a dynamic simulation?
- Does it also produce different modes of behavior of a dynamic simulation?
- Are the different results and mode of behaviors relevant for the decision makers?

PLUG & PLAY MODEL

INTRODUCTION TO THE PLUG&PLAY MODEL

The plug & play model was taken from the research presented in 2008 in the Computers and Chemical Engineering Journal, volume 32 in the paper titled "Decision support for integrated refinery supply chains Part 1: Dynamic simulation". The authors of this paper are Suresh S. Pitty, Wenkai Li, Arief Adhitya, Rajagopalan Srinivasan and I.A. Karimi, at that time from the Department of Chemical and Biomolecular Engineering of the National University of Singapore and the Institute of Chemical and Engineering Sciences.

RUNNING THE MODEL

METHODOLOGY OVERVIEW

In order to evaluate the hypothesis of this research the following steps are followed:

7. A pre-developed –plug & play model called IRIS is obtained, implemented and tested. This is explained in the chapter "**Plug & Play Model**".
8. The original configuration of the model is deterministic emulating uncertainty, in other words, it doesn't rely on repeated random sampling to obtain numerical results, but incorporates a variation percentage and a variation seed that is fixed. In practice, this means

that if the model runs several times the numerical results are going to be the same (Suresh, et al., 2008).

9. The original configuration is modified to eliminate the variation percentage. This is achieved by practically setting the variable to zero (0) and running the simulation again one more time.
10. Uncertainty is incorporated in targeted ways. Different modules of the system are defined as uncertain with some realistic variation and the simulation is stochastically iterated 1,000 times.
11. Stochasticity is incorporated in fourteen of the variables in a standard way (a random value with normal distribution). The dynamic simulation is stochastically iterated 1,000 times.
12. The distributions and mode o behaviors of the results of the simulations are compared statistically and graphically to analyze the effect of uncertainty on the dynamic simulation.

CONCLUSIONS

1. Considering stochastic variability can have a major effect on the performance of the system. Since the dynamic model used in this research (IRIS) is non-linear, this is an empirical case of Jensen's Inequality. Specifically of Jensen's inequality on a concave function, where the expected result from the deterministic simulation was compared with the distribution of results from the stochastic simulation giving a graphical and statistically significant difference (E(f(x))<f(E(x))):

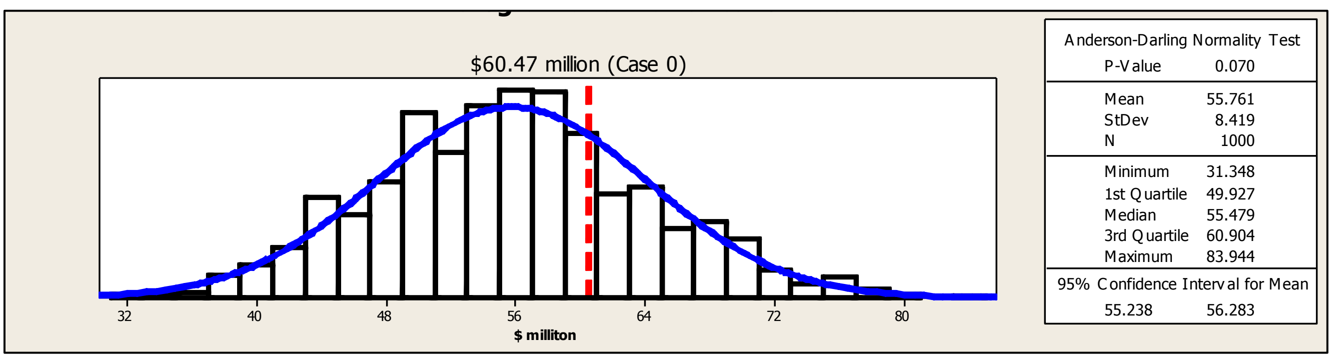

FIGURE 52: HISTOGRAM AND STATISTICS FOR CASE 1A

2. Even if the effect of uncertainty and non-linearity don't impact the expected value of the model, it will produce a distribution of results and a mode of behavior that could be significantly relevant for the decision maker. The mode of behavior of the dynamic results can foster completely different human behaviors from the managers and decision makers, even if the final simulated result is statistically not different. In the figure below, the oscillation produced by incorporating stochastic variability produces a completely different mode of behavior.

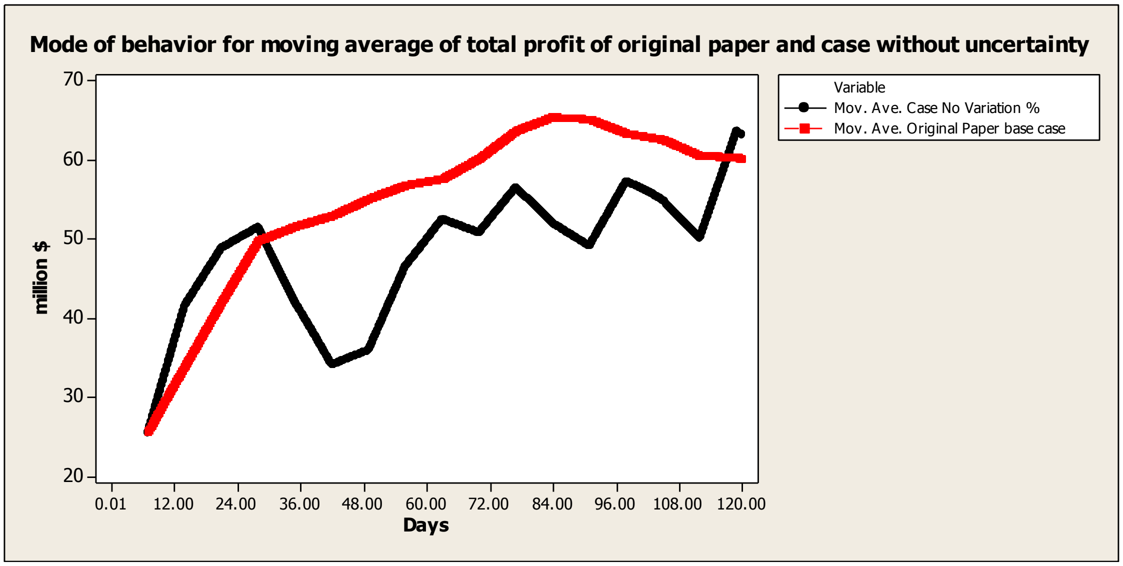

FIGURE 53: COMPARISON BETWEEN BASE CASE AND CASE 0 USING TIME SERIES OF THE MOVING AVERAGES

3. The effects of stochastic variation and complexity on the average, risk or mode of behavior can be non-related, which means that, for example, a case with better expected value could also produce a worse mode of behavior or risk.

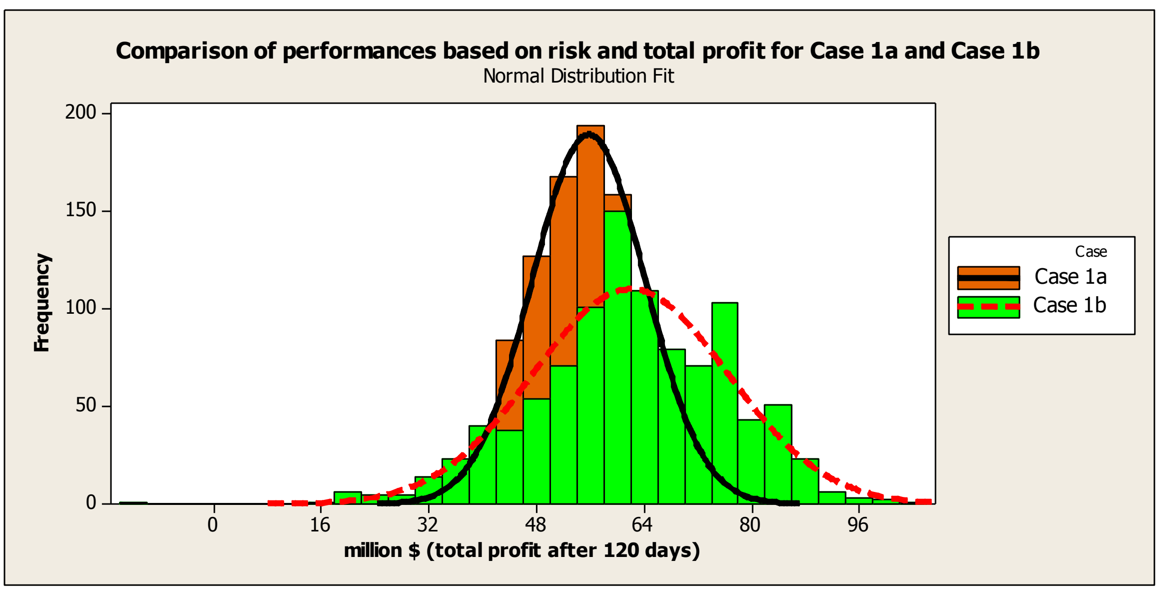

FIGURE 54: COMPARISON OF DISTRIBUTIONS FOR CASE1A AND CASE1B

4. The dynamic simulation in the case of IRIS considered a defined and limited time horizon. The effect of the stochastic simulation (Jensen's Inequality) seemed to be propagating over the standard deviation, showing a steady increment over the time horizon of the simulation. The relevance of this insight is that, since the standard deviation is growing over the time horizon of the dynamic simulation, the interval of confidence for the expected value will

also grow, which could eventually trick the decision maker to think that the model is producing a similar result, when in fact is just the model losing the precision to make a significant comparison.

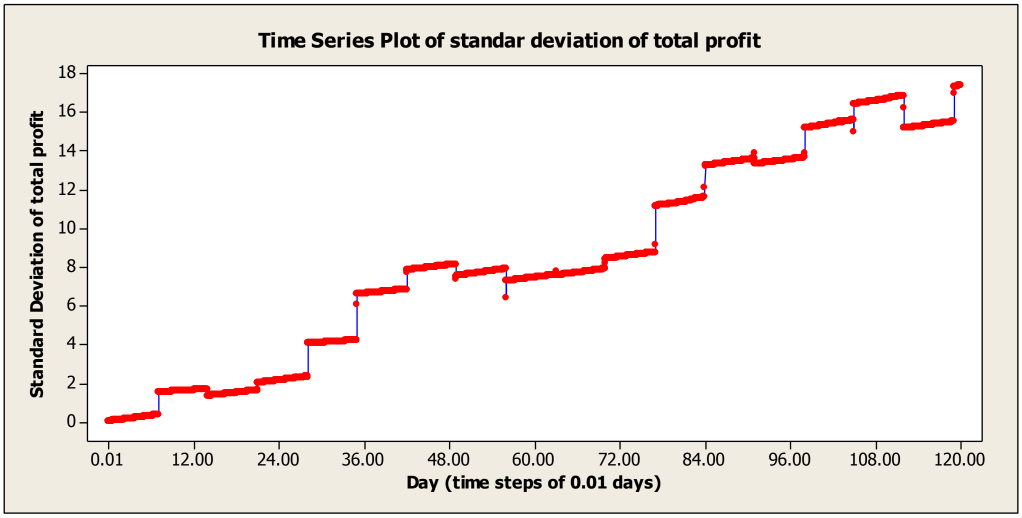

FIGURE 55: PROPAGATION OF STANDARD DEVIATION

5. A sensitivity analysis showed that stochastic variability is not important for every variable in the model. This is important since the decision maker should focus the efforts only on the variables that are relevant from the perspective of the sensitivity of the results of the system to their uncertainty.

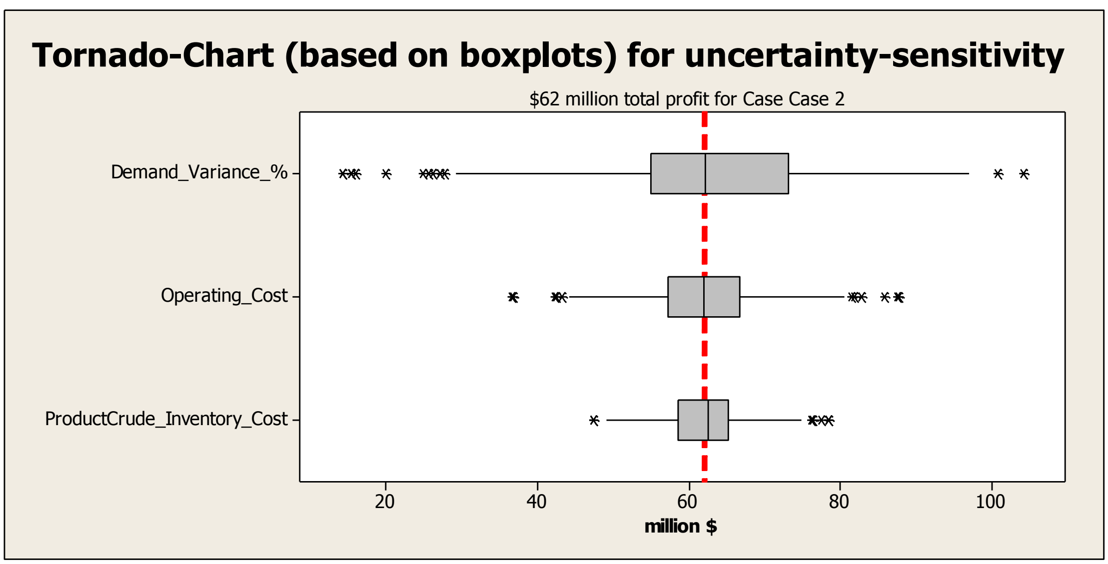

FIGURE 56: TORNADO CHART FOR SIGNIFICANT VARIABLES

6. The stochastic simulation of a dynamic model like IRIS can produce contra-intuitive results due to dynamic complexity.Figure 36 represented a variable that increased total profits when it increased variability

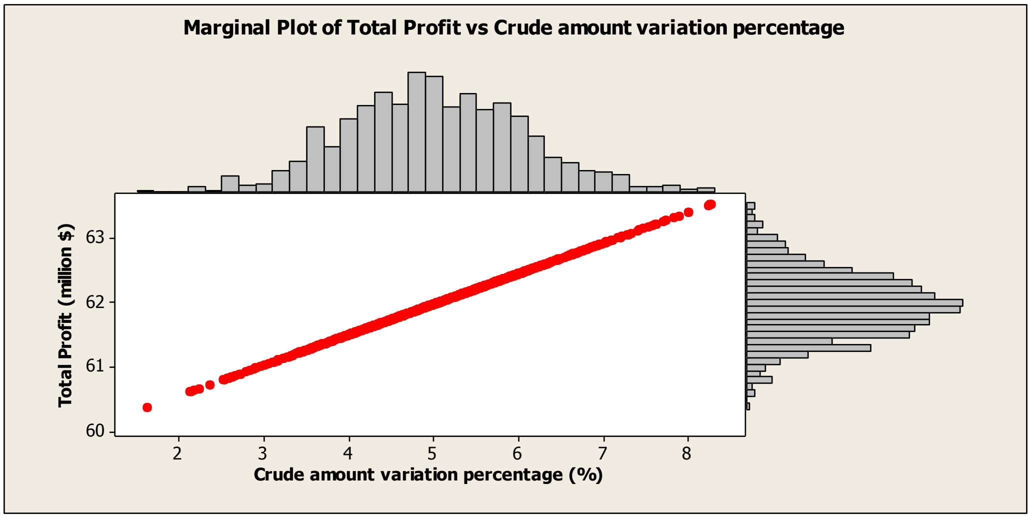

Figure 57: Total profit vs Crude Amount Variation %

DISCUSSION

In general, it is clear from the bibliographic research that there is a significant academic and industrial consensus on how importance and challenging is managing supply chains. The detailed and dynamic complexity are both present in any supply chain system, and they are especially present in an energy company due to their global reach and diversity and scale of products. Because of the mentioned reasons, dynamic simulations are highly non-linear (Sterman, 2000) and therefore the effects of Jensen's Inequality, must be considered as a cause of producing the wrong results. IRIS served as a practical and flexible platform for testing the different effects. It is expected to find similar, and even bigger, effects on more complex and realistic models. For this reason it is encouraged to research further the effects of Jensen's Inequality with more models and more specific applications.

Discontinuity is a source of non-linearity and management can creates discontinuity. This happens because managers or system operators decide to take some major decision about a project creating change in the function or model of behavior (de Neufville, 2012). The decisions will depend on many aspects: culture of the organization, particular context, results from previous periods, external pressures, personal goals, etc. Particularly important are the current mode of behaviors of the system. Different mode of behaviors will produce different managerial behaviors, for example, a declining metric of performance could produce different decisions than an increasing one.

The context is analyzed, most of the times, using a short time horizon and a limited set of data. This could produce a biased perception of the process performance that will, most probably, impact the behavior. In addition, disruption and special events will certainly influence the perception of status of performance, driving different behaviors or changing the function of performance producing nonlinearity. The practical consequences on the results of a system are that the estimated result will be wrong, and that the management would produce a personal resistance to recognizing and dealing with uncertainty. This is augmented when managers deal with many client relationships and has to respond to expectations of certainty (de Neufville, 2012).

(Blank)