STOCHASTIC MODELLING IN THE PETROCHEMICAL INDUSTRY (DISCRETE EVENT SIMULATION BASED)

MARLIZE MEYER Hons. B.Sc.

Mini-dissertation submitted for the degree Magister Scientiae in Operations Research at the Potchefstroomse Universiteit van Christelike Hoër Onderwys

Supervisor Dr. P.D. Pretorius

2004
Potchefstroom

ACKNOWLEDGEMENTS

Publishing of this study indicates a milestone in my life of something achieved not only by sheer determination but also through the grace of God and the unfailing support and trust of the people in my life. You may not have written the text but you have given me the courage to complete this study. Thank you!

To everybody who reads this study:

It is not important what we can achieve ourselves, it is important what God can achieve through us.

Therefore

Invest your time wisely.
Spend time with your family and your children.
Laugh more, pray more, care more.

Time passes so quickly. The world can wait a few more days for a new invention but the first steps of a baby, the total trust, the first smile, the first words and the first day of school will never be again. Like flowers need water the people in our life needs nourishment and that can only be measured in the time that we share with them.

OPSOMMING

Hierdie studie verskaf 'n gedetaileerde en belynde beskrywing van 'n simulasie proses wat die modeleerder sal help om 'n simulasie projek suksesvol te voltooi. Die studie lig die tekortkominge uit van baie van die prosesse wat in die literatuur bespreek word, en verminder so die risiko van 'n onsuksesvolle projek.

Een van die hoofredes waarom al die stappe nodig vir 'n simulasiestudie nie altyd in die literatuur genoem word nie is omdat alle modelle nie met 'n basismodel begin nie.

Die persoonlike dinamika van die rolspelers in die projek moet te alle tye gemonitor en bestuur word as gevolg van die impak wat dit mag hê op die totale projek sowel as die boodskap wat uitgedra mag word in forums waar die modeleerder nie verteenwoordig is nie.

Die Petrochemiese omgewing bied 'n groot uitdaging aan stogastiese modelering as gevolg van die kontinue omgewing, die kontinu – diskrete koppelvlakke en die interaktiewe prosesse en aanlegte wat deel daarvan vorm. Die Petrochemiese omgewing is ook 'n uitdagende modeleringsomgewing met programmatuur en tegnologie wat oor baie jare reeds opgebou is. Die aspekte van die werklike sisteem wat nie reeds ingesluit is in die historiese modelering nie, het 'n gaping gelaat waarin stogastiese modelering goed pas. Gedurende die studie word bewys dat stogastiese modelering hierdie gaping suksesvol kan vul. Voorbeelde van suksesvolle toegepaste stogastiese modelle in die Petrochemiese industrie word ook in die studie bespreek.

Een van die waardevolle bydraes van stogastiese modelering in 'n omgewing waar voerstrome gemeng word om 'n produk te maak, is die gevolgtrekking uit die resultate dat infrastruktuur beperkings gekombineer met 'n wanbalans in die volumes van die voerstrome op die tydstip wanneer 'n mengsel gemaak moet word mag uitloop op 'n groot verskil tussen verwagte en werklik behaalde volumes wat aan die mark verkoop kan word. Tydsberekening en beskikbaarheid van genoegsame volumes in die regte verhouding speel 'n kardinale rol in die volumes wat gemaak kan word. Die volgorde waarin die mengsels gemaak word het ook 'n impak.

ABSTRACT

This study provides a fully described and streamlined simulation process that will assist the modeller in successfully completing a simulation project. This study also highlights the shortcomings of many of the processes discussed in literature, and it reduces the risk of an unsuccessful project. One of the main reasons why all the steps are not usually mentioned in modelling environment is because most models do not require a base model.

People dynamics in the project should be monitored and managed carefully due to the impact it has on the overall project and the message that will be distributed in forums where the modeller is not present.

The Petrochemical industry poses a huge challenge for stochastic modelling due to its continuous nature, its discrete continuous interfaces and the highly interactive processes and plants. The Petrochemical industry is also a tough modelling environment with well established software tools and technologies. The aspects of the actual system, not covered by the historic software, have left a gap where stochastic modelling fits nicely. During this study it is proven that stochastic modelling can fill this gap with huge success. Some examples of stochastic models where they are applied in the Petrochemical Industry are discussed in this study.

One of the valuable contributions that stochastic modelling can make in the petrochemical industry is to show that infrastructure constraints combined with an imbalance in available blend volumes at the time of blending can often, surprisingly, create a huge gap between expected and actual volumes sold to market. Timing is critical and having the right volumes in the right balance is essential. Blend sequence can also have a huge impact on volumes achieved.

CONTENTS

CHAPTER 4 SUMMARY AND FURTHER WORK

List of flow diagrams

List of figures

List of distributions

List of graphs

List of tables

CHAPTER 1 INTRODUCTION

"*Simulation refers to a broad collection of methods and applications to mimic the behaviour of real systems, usually on computer with appropriate software*" (Kelton et al., 2002:3)

In the process of model building we are translating our real world problem into an equivalent modelling problem which we solve and then attempt to interpret. (Hangos & Cameron, 2001:10)

"*A simulation is a model that mimics reality ...*" (Robinson, 1994:3)

"*Discrete Event Simulation involves modelling of a system as it progresses through time ...*" (Robinson, 1994:3)

"*...the process of designing a model of a real system and conducting experiments with this model for the purpose of understanding the behaviour of the system and/or evaluating various strategies for the operation of the system.*" (Pegden et al., 1995:3)

Discrete simulation was originally used for queuing systems but it expanded so much that it can be used to solve most **quantitative** problems. This study discusses how stochastic modelling or stochastic simulation can be applied in the process industry and therefore also in the Petrochemical Industry.

"The term stochastic process is frequently used in connection with observations from a time orientated, physical process that is controlled by a random mechanism." (Hines & Montgomery., 1990:630)

"Stochastic programming deals with situations where some or all the parameters of the problem are described by random variables. Such cases seem typical of real-life problems where it is difficult to determine the values of the parameters exactly." (Taha, 1976:588)

In the rest of this document stochastic modelling is loosely used as the term to describe the use of discrete event simulation to model a process where all the parameters are described by random variables. The term also tries to capture the "time orientated, physical process." (Hines & Montgomery., 1990:630)

1.1 Research Goals

Simulation is a very old technique. Simulation, due to its nature, encompasses a little bit of several different specialised fields and when practiced becomes not only statistics but a form of art. Simulation can extract the essence of reality and capture it in a model. The constraints of simulation during its development stages were the steady state, the average and the time it took to develop a detailed model.

Today's technology made it possible to develop models that include the effect of the passage of time and the interaction between several interlinking processes, while including their intrinsic operation and deviations. This brought remarkable possibilities within grasp.

For many years stochastic modelling has explored and optimised the shop floor. Stochastic modelling became part of the modelling toolset of banks, restaurants, job shops, airports, major storage and grocery stores and many other logistics enterprises. Virgin soil still remain, being the chemical and petrochemical environments and within them, very dynamic, continuous environments.

The goal of this study is take stochastic modelling from its niche in the supply chain environment and advance it into a project environment and specifically into the more dynamic petrochemical environment and to show where stochastic modelling can add value. This goal is divided into four more detailed goals for the sake of the study.

Goal 1
This study is about exploring the entry into this dynamic field and it describes how Sasol has used stochastic modelling in the last ten years.

Goal 2
This study describes where stochastic modelling fits into the modelling framework and what value can be added with it.

Goal 3
This study outlines the management of a modelling project compared to the normal project environment and gives many model examples and applications.

Goal 4
This study attempt to illustrate that stochastic modelling can be extended successfully to the Petrochemical industry.

1.2 Background hypothesis

This study proposes a systematic approach to the development and analysis of models in the Process Industry using stochastic modelling techniques. The complexities can vary enormously but the approaches took on a "systems view" of the problem with regard to the components in the process, the inputs, the outputs of the system and the complex interactions which could occur due to the connected nature of the process.

There is a growing realization that significant benefits would be gained in the overall economics and performance of processes when a systems approach is adopted. This covered the design, control and operation of the process. (Hangos & Cameron, 2001:16)

In order to achieve this goal, there is a growing trend to reduce the complex behaviour of the process system to manageable forms – hence the use of models and simulation programs.

The idea, that one model serves all purposes, is fictitious. Models must be developed for a specific purpose or set of purposes and that purpose will direct the modelling task. Models with different purposes can be combined to cover a broad spectrum of services.

Modelling and simulation can be used in any area of model application, including process design, process control, trouble-shooting, risk reduction, operator training, environmental impact and logistics. Behind each of these areas there is a need for effective model development and documentation of the basis for these models which are developed.

A systematic approach is essential if reliability and continuity is required for the models.

A literature scan was done to find articles on Stochastic modelling or discrete simulation in the Petrochemical Industry. All the articles found were based on logistics and never Processes. Historically the Process industry belongs to the Process modelling environment and sometime a bit to LP (Linear programming) or other optimisation technologies. Until recently this was a new field for stochastic modelling. This study contains a few practical examples where stochastic modelling is needed and where it can provide superior value.

Basic Hypothesis:

The basic hypothesis of this study is therefore that stochastic modelling can be extended successfully to the Petrochemical industry. The value-add lies in the addition of passage of time and the exploration of the interaction of linked processes with their own internal variation.

1.3 Technologies and tools available for simulation

A mathematical model can be a simulation, a spreadsheet model in Excel® is often called a simulation, a match box design in 3 D is a simulation model as well and who of us are not fascinated by flight simulators. The field of simulation is extremely wide and basically any model can also be called a simulation. The type of simulation that is discussed in this book is very specific. It is called discrete event simulation or as it is referred to it in the operational environment, stochastic modelling.

Stochastic modelling refers to modelling that is time based and very close to reality. It is not deterministic modelling. Deterministic modelling can be time based but averages are used in the model. In a stochastic model distributions or actual data over time is used to describe input streams. Reactors, for example, have a distribution of rates for the incoming stream and actual data is used and fitted into distributions to depict the actual throughput through that plant over time. Where appropriate split factors are used to model the relationship between the incoming stream and the streams going out of the reactor after the reaction took place. The distributions are used because rates are not constant even for a specific set of reactor settings. All operation philosophies are also captured in the model and all rules that may have an influence on the real word system.

The number of rules and variables in these models will daunt the most capable mathematical program. Our powerful 20th century computers are not capable to handle it, yet. The reason why stochastic modelling does offer a sensible solution is because it includes the time component. The number of simultaneous actions is less. The model also runs sequentially. If the volume for a ten minute time interval (depending on the approach of the specific model) needs to be added to the model in ten minutes, the action gets added to the sequence of events that needs to happen in ten minutes. For the next 9.99 minutes the model takes no notice of it.

1.3.1 Solving problems in a purely mathematical way:

Mathematical modelling was the method relied upon when simulation as a tool in itself did not exist yet. Mathematical simulation was used in many fields and gave a workable, high quality solution to most problems. Mathematical techniques include linear programming, regression analysis and queuing theory.

Advantages of solving problems in a purely mathematical way

- Works well when optimisation is needed.
- Very sound mathematical background needed

Disadvantages of solving problems in a purely mathematical way

- Many assumptions need to be made to simulate a very involved system or when a lot of changes need to be made.
- Not the best option where time and variation plays a big role. ("interaction of random events" (Robinson, 1994:8))

1.3.2 Linear Programming

"Linear programming is a class of mathematical programming models concerned with the efficient allocation of limited resources to known activities with the objective of meeting the desired goal (such as maximising profit or minimising cost). The distinct characteristic of linear programming models is that the functions representing the objective and the constraints are linear", (Taha, 1976:15)

Linear models are used because they are usually easier to solve computationally, but the assumptions are stronger. The goal of the model is to adequately describe the existing system by making good assumptions. Many events that happen in an existing system can be approximated linearly without having a huge impact on the model outcome. Other events however are more difficult to approximate with a linear equation.

Applications

- Production planning
- Feed Mix Optimisation
- Stock cutting or splitting optimisation
- Water Quality Management
- Oil drilling and production
- Assembly balancing
- Inventory
- Economic optimisation

Advantages of linear programming

- Linear programming is an excellent optimisation tool
- The time necessary to build a model is not very extended in most cases.
- As economic data is often reported at many levels of an organisation it is not very difficult to acquire for a study.

Disadvantages of linear programming

- A lot of experience is necessary to understand the impact and constraints of using a linear approximation for non linear problems. The influence of the assumptions must be taken into consideration.
- Linear programming solutions divert from reality where there is lots of variability in the process.
- An optimum answer is provided but little indication can be given as to the stability of the process at that point
- Expertise in the technique is needed for analyses of the results

Like simulation Linear Programming can be used to solve a wide spectrum of problems. There are also other techniques like Non Linear Programming (NLP) and Mixed Integer Programming (MIP). These techniques are very powerful but very memory intensive and models often become unsolvable due to computing constraints. Some variation and time constraints can be catered for with the dawn of multi period modelling. Multi period modelling is one of the techniques offered in PIMS® (Process Industry Modelling System), a linear programming software tool, created and supported by AspenTech.

1.3.3 Monte Carlo Simulation

For more information on Monte Carlo sampling see Pegden *et al.* (1995:11)
In the Monte Carlo technique random numbers are generated using a random number generator. This numbers are then used to create artificial random data. The ability to generate random data forms the core of simulation originating from the atomic bomb project with the work of von Neumann and Ulam. In Los Alamos the security code "Monte-Carlo" was given to the mathematical technique. This became the term used for simulation or at least for processes that contains an element of chance". (Freund, 1984:195)

1.3.4 Spreadsheet modelling

Spreadsheets are often used in smaller projects especially where there is little variation or only a few variables. The Solver in Excel® increases the capacity of spreadsheet solutions to a great extent, but it still is limited. Various expert solvers and add-ins are available to spreadsheets and it has become one of the most often used and most powerful modelling tools. The reason for the preference is because people already use it for their normal spreadsheet needs. It is widely used as part of the Microsoft Office® toolkit and a common understanding already exist in most students when they graduate. Visual basic interfaces and macros also increase the capability of spreadsheets very much. All these reasons have made spreadsheets a big favorite and a tool worth considering.

The advantages of Spreadsheet solutions

- Spreadsheets are universally used, although for mostly non simulation purposes
- Spreadsheets is a nice way to model for small problems with not to much detail
- Spreadsheets provide a quick way to simulate
- Spreadsheets is user friendly in terms of ease of use and help files Various add-ons and Visual basic programming make Spreadsheets a very powerful and much preferred tool

The disadvantages of Spreadsheet solutions

- Models can become very involved and difficult to debug
- Spreadsheets cannot model variation without programming
- Spreadsheets cannot model extensive variation
- Continuity is difficult especially where macros or visual basic programming is used
- Spreadsheets is user unfriendly in terms of simulation solutions
- Spreadsheets are not easy to document and view, because of the sheer size of sheets.

1.3.5 Process Modelling

"...process modelling is a fundamental activity underlying the effective commercialisation of process ideas and the ongoing production of goods and services." (Hangos & Cameron, 2001:472)

" ... models offered in commercial simulation packages such as Aspen Plus®, Aspen Dynamics®, ... to generate solution to process engineering questions" (Hangos & Cameron, 2001:472)

Hangos & Cameron (2001) discuss the need and use of process modelling in the process industries and the development in modelling tools through Petrochemical and Petroleum Industries as well as in the Minerals Processing Sector.

Process modelling tools are mostly chemical and reaction based modelling. A reactor can be defined including its properties and the impact of reactor properties on yields in the various outgoing streams.

The current challenge for process models lies in industries with "large scale discrete-continuous operations" (Hangos & Cameron, 2001:16)

The advantages of Process models

- The fundamentals of the chemical processes are included
- Reliable data is obtained from the model

- Vectors can be generated
- Chemical properties and reactions are part of the toolset
- Most engineers have a basic process modelling background after they graduate
- Much is published on this topic that can be used as references

The disadvantages of Process models

- Modelling assumes an in depth chemical and engineering background
- Models are often rebuilt instead of reused (continuity problems)

1.3.6 Steady State Simulation

Steady state (stationary) modelling is often used in the chemical environment. "Steady state is assumed for this type of model." (Hangos & Cameron, 2001:21) It gives one an answer as to the workability of the system in a specific state. It would be wise to test more than one state to make sure that specific scenarios are addressed.

"Because a system is normally designed under steady state conditions, the initial transient state results of a simulation models must be bypassed before any output is recorded. " (Taha, 1976:524)

The advantages and disadvantages of Steady state models

- Steady state information is readily available
- Easier to build models in a steady state
- Relatively easy to obtain a mass balance in steady state

The disadvantages of Steady state models

- In a process environment even a steady state is not without some variation

1.3.7 Dynamic Simulation

Dynamic simulation is "the process model developed to represent changes in time..." (Hangos & Cameron, 2001:22) In dynamic models all the principles of steady state remains, but time intervals are added to get results usable under different conditions.

1.3.8 Simulators (software)

Simulators are very simple, solution specific simulation packages. They do not have the power of a full simulation package but are usually written for a very specific application and work well within the boundaries of the environment it was developed for. If one tries to use the simulators outside of these boundaries it will not work well or not at all. "However, the domains of many simulators are also rather restricted and are not generally as flexible as you might like in order to build valid models of your systems..." and simulators also "..traded away too much flexibility to achieve the ease-of-use goal." (Kelton *et al.*, 2002:12) For a prospective simulation user to buy a simulator instead of a full simulation package is a mistake easily made. It is wise to talk to people who have been using the software, as the vendors do not always advertise simulators as simulators. They are often advertised as simulation packages. If a simulator is all one needs it can be a significant saving to buy the suitable simulator instead of a full simulation package. The purpose of the software must be evaluated properly.

The main advantages of simulators are price and ease of use. The main disadvantage is the limited capability of the tool itself.

1.3.9 Discrete event simulation (using Arena®)

"Arena® combines the ease of use found in high level simulators with the flexibility of simulation languages, and even all the way down to general-purpose procedural languages like the Microsoft® Visual Basic® programming system or C®. It does this by providing alternative and interchangeable templates of graphical simulation modelling-and-analysis modules that you can combine to build a fairly wide variety of simulation models" (Kelton *et al.*, 2002:12) Arena® support both continuous and discrete processes and therefore also the capability to approximate continuous processes in a discrete way.

Arena® is hierarchical and the actual model looks much like a flow diagram with different parts of the model being modelled as sub models if required. Arena® also has many animation features to show the progress of entities through the system.

An entity can be a part or a person or whatever needs to move in or through the system. Entities can also be virtual for example where an entity is used to change a variable in the system at specific points in time.

The value of a variable in Arena® is global and similar to other programming software, can be seen throughout the model. The attribute of an entity however belongs only to that entity and will be part of it as the entity progresses through the model.

All information required to build and run simulation models is stored in the modules. *Flowchart modules* are placed in the model window and connected to each other to form a flowchart describing the logic of the process you are modelling. The basic flowchart modules and their functions are as follows:

- **Create** – the start of the process; the point at which entities, the items that move through the process – enter the simulation.
- **Dispose** – the end of the process, at which entities are removed from the simulation.
- **Process** – an activity, usually performed by one or more resources and requiring some time to complete. Resources can be machines or people (something that can be used for a time) or any object. The resources can be seized, used and then released. The process time allocated to the resource may be considered to be value added, non-value added, transfer, wait or other. The associated cost will be added to the appropriate category.
- **Decide** – a two-way or n-way branch in a process. One branch or multiple branches can be taken. A branch can be taken based on a logical expression or on a chance for something to happen.
- **Batch** – collection of a number of entities before they can continue processing.
- **Separate** – duplicate entities for concurrent or parallel processing, or separating a previously established batch of entities.
- **Assign** – change the value of some parameter during the process, such as the entity's type, a model variable, an animation feature or status.
- **Record** – collect a statistic or set of statistics, such as entity count or cycle time.

The flowchart view contains all of your model graphics, including the process flowchart, animation, and other drawing elements.

1.4 Useful terminology in discrete event simulation or stochastic modelling

Discrete time intervals

The straightforward meaning of **discrete** is not continuous or segregated. For time intervals it means that continuous time is broken up into set time intervals or time buckets as they are referred to by simulationists. This means that the system is modelled in exact time intervals, whether one second or one year. A continuous system can be modelled using very small time intervals. Both continuous and discrete actions can be modelled easily. The size of the time intervals is the choice of the modeller based on the requirements of the model, to describe the system adequately.

Discrete simulation is most often used when the problem is very involved, but its full potential is only realised in stochastic systems or in systems with a lot of variation or compounded variation.

Variation

"...the extent to which data are dispersed, or spread out..." (Freund, 1984:53) At a plant all the tankers arriving to load a product do not arrive at the same time, there is variation in the arrival times. All the tankers do not load the same product; they do not load the same volumes or stay for the same time. In every event or aspect of life there is inherent variation and variation imposed from outside of the specific situation.

Compounded Variation

Each task has variation and this in itself may pose no problem in a specific system, but the interaction of the variation of tasks in sequence may pose a real threat to that system because the tasks are interdependent. While a problem during one task may create a small backlog, the effect of this delay may be increased in further delays in the actual system. In his book "The Goal", Goldratt, (1992), explains very eloquently with his hiker example the impact of variation if it occurs in sequence. If one process excels, it does not necessarily mean that the others have the capacity to catch up, but if one reduces speed, it immediately impacts the systems upstream. Buffers reduce the direct impact.

Random number generator

All simulation packages (or most) make use of random number generators to introduce events into a system. The quality of these random number generators differs and their effect should be evaluated on the simulation model developed.

"An acceptable random number generator must be able to generate random numbers that are uniformly distributed between 0 and 1." (Pegden *et al.*, 1995:11)

Most of today's random numbers are generated by computer, using an algorithm. These numbers are usually pseudo random as the sequence that is generated repeats itself. These numbers are random enough to show no difference on a statistical basis between the generated random numbers and true random numbers. The impact of repeating the same sequence of numbers must be considered when interpreting results.

1.5 Layout of the rest of the study

Stochastic modelling projects, similar to other types of simulation projects, needs to be managed very closely to ensure timely and successful completion. A comparison is done in **Chapter 2** between the published methodology (Render *et al*, 2003:3), and the requirements and exceptions for stochastic modelling derived from building stochastic models.

The basis or framework is the same for any study or model but for stochastic modelling a lot of focus needs to be on training and explaining of the concepts and capacity. Detailed scoping is needed, thorough planning of the model and of scenarios is essential before starting the project and the implementation phase needs to be monitored closely. One key driver for success is often overlooked. This driver is choosing the right tool or software to best solve the problem.

Chapter 3 discusses stochastic modelling in the Petrochemical industry and where and how it is used. An example is discussed where stochastic modelling is used in fuel blending and distribution as well as an example in calculating plant throughput given variation, plant availability and interactions between several plants.

Chapter 4 gives the summary and suggests further work.

CHAPTER 2 MANAGEMENT OF STOCHASTIC MODELLING PROJECTS

2.1 Introduction

Stochastic modelling projects needs to be managed very closely to ensure timely and successful completion of an extended modelling project. A comparison is done between the published methodology, as in "Quantitative Analysis for management", Render et al (2003:3) and the requirements and exceptions for stochastic modelling as observed in practice. The basis or framework is the same for any study or model but for stochastic modelling a lot of focus needs to be on training and explaining of the concepts and building an understanding of the capacity of the tool. Detailed scoping of the project is needed, complete planning of scenarios is essential before starting the project and the implementation phase needs to be monitored closely. One key driver for success is often overlooked. This driver is choosing the right tool or software to best solve the problem and to not only use the known software in all solutions.

Many of the steps explained here can also be seen in "Discrete Event System simulation", Banks et al (1999:14) although it is discussed in little detail.

2.2 Background

Stochastic modelling is a tool often used in decision support and a stochastic modelling project like all other projects need to be managed well to be successful. Normal project planning does form a sound basis for project management for stochastic modelling but a whole process happens between the defining of the problem and the start of the model. During this process some of the decisions that need to be made are the tools to use, the level of detail required, what possible questions need to be answered, doing the conceptual design and identifying data needs.

Stochastic modelling projects due to their nature often span several parts of business. If it is not planned well the project may fail, due dates may not be met and the general feel of the project from the clients side will be negative.

In stochastic modelling projects one does scenarios and as the results are generated, new ideas which needs to be tested will be generated. System knowledge and understanding of interactions between parts of the system will increase. As learning takes place during the project scope changes also occur naturally. These changes need to be evaluated and a decision needs to be made if they will be part of current scope or not, because it has the capability to extend the project and therefore move out any results. Often from the modeller's point of view the scope needs to be changed to adequately describe the reality, but here much focus is necessary. The modeller need to focus on the intended outcome, the negotiated level of detail and the time schedule. Only after all factors are considered carefully can the decision be made to add the scope change or to move it out to the end of the project.

2.3 Defining the problem

"The first step in the quantitative approach is to develop a clear, concise statement of the problem" (Render et al., 2003:3). The first step in stochastic modelling is to understand the system or process. The problem as described by the client is often a result of several factors or constraints that may together create the perceived problem. The client often describes the "symptom" rather than the "illness". Using the "symptom" to describe the problem is not wrong because the "symptom" also needs to be solved, but the quest needs to be understanding of the system, its interactions and the implications of making changes in any part of the system. Render emphasizes the need to find the true causes when defining the problem. In many projects the causes are not that easy to find and building a stochastic model will assist the modeller in finding these causes. Some examples of defining the problem or setting objectives by using the "symptoms" are: Reduce queue lengths of tankers at a weighbridge or reduce process time in the system.

"Once we select the problem to be analyzed, the next step is to develop a model." (Render et al., 2003:3) Once one understands where the system show obvious problems it is then necessary to look at all the tools available to sufficiently describe the system that needs to be modelled, to assist the client in making a decision regarding the tool to use for solving the problem, to have a look at the level of detail required, identify the data needs and do a conceptual model design.

The proposed order of the additional steps before model building is:

2.3.1 Conducting training and information sessions

2.3.2 Choosing the appropriate tool

2.3.3 Deciding the level of detail

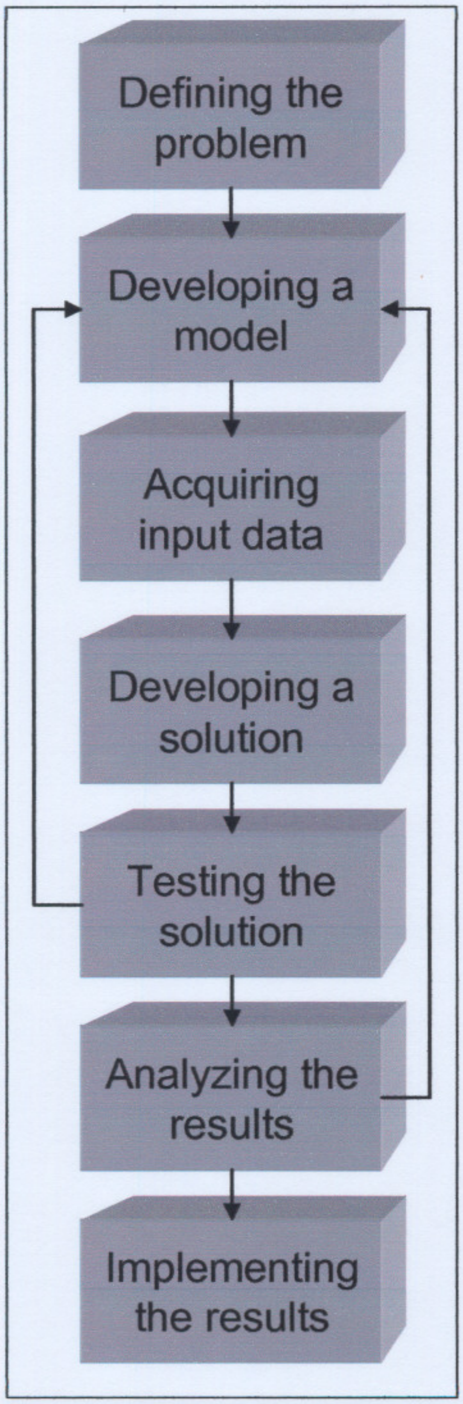
2.3.4 Conceptual design

2.3.5 Scenario planning

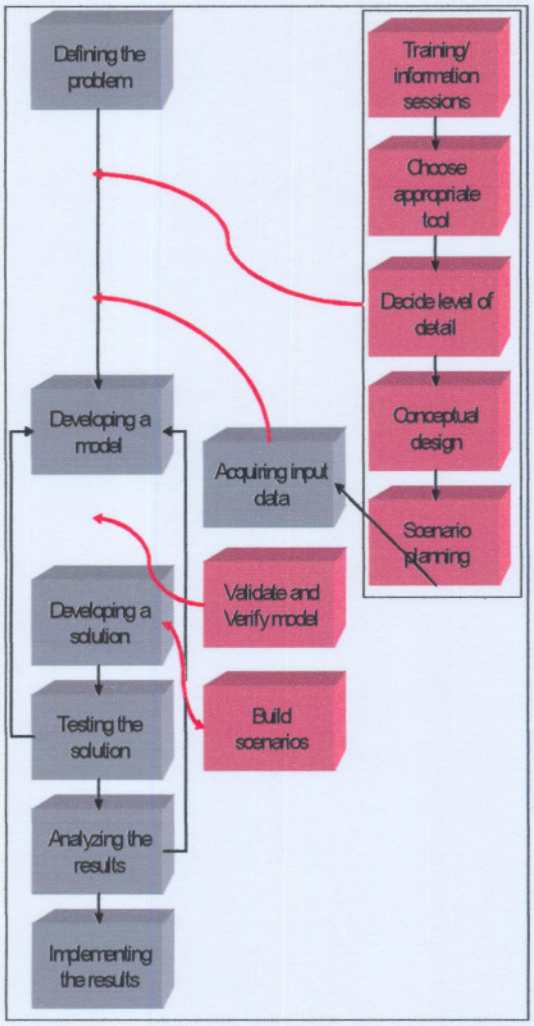
2.3.6 Optional: Study cost and schedule

The 6 steps are included in the following paragraphs to show their need and significance.

Flow diagram 2.1: Render et al (2003:3) – Proposed process



Flow diagram 2.2: Edited process



2.3.1 Conducting training and information sessions

For any model to be accepted by the client, the client needs to adequately understand what the capability of the model will be, what the typical questions are that can be answered and where and how the model can be further used in his environment. For example can he/she use the model again afterwards, can the model be used on a daily or weekly basis and what capability is necessary to fully use the model.

The best way to do this it to take the tools most often used in his environment and play them off against each other, not being biased, but showing the strong and weak points of each. In a petrochemical environment the tools most often used are process modelling and linear programming and sometimes advanced process control depending on the maturity of the business. The modeller also needs to understand the place of each tool to be able to enlighten the client and to compare them and make the right decision.

Stochastic modelling is usually used in job shop simulations or queuing problems e.g. banks and restaurants. It is not well known in the petrochemical environment and therefore emphasis needs to be placed on what value this tool can add in addition to the others. The client also needs to understand the essence of stochastic modelling. This whole process should only take one or two sessions and there has to be at least one example.

2.3.2 Choosing the appropriate tool

The next step in the process is to choose the correct tool. The value of building a linear program is less than that of building a stochastic model where storage and time interactions are important. On the contrary when economic optimisation is required a stochastic model can be of less value than a linear program. Process models can sort out a process beautifully but it will not assist fully in infrastructure sizing. Where day to day planning is required a combined model may be needed or software that is developed specifically for planning.

Through trail and error it is now obvious that the best solution is to sometimes combine technologies, to decide when point solutions are required and when the solution needs to answer questions on a daily basis. Modelling projects are expensive and doing them twice results in unnecessary costs. For the purpose of this study we assume that the choice of the specific project is stochastic modelling because of the effect of time and variation that plays a role in the results.

2.3.3 Deciding the level of detail

After deciding what tool to use, the next significant decision is the level of detail required for the model. This decision has a large impact on the duration of the project and a good rule of thumb is to always keep the level of detail as low as possible. This involves judging some if the possible system impacts, possibly reducing the number of answers that can be given and focussing on specific part of the system predefined by the involved parties. As is evident from this statement there is a level of risk and therefore a chance that the project may be expanded later. One way to make sure that the level of detail is not too low is to do scenario planning before the model is built. This will give the modeller an idea of what needs to be changed during the scenario testing phase and he can include this in the model from the beginning. Before this can be done the system must first be understood very well.

2.3.4 Conceptual design

The conceptual design can be done now, but it can also be done even earlier in the process, before the tool and level of detail is decided.

The first part of the conceptual design phase of the project is building a flow diagram of the process and writing down al the decision rules for how the system fit together. Part of this process is also conducting meetings with the plant or process operators and managers. It is necessary to understand the system from all the perspectives, the economical drivers, the market drivers and what rules drive and constrain operations. In bigger companies this information resides in different departments with very different focus areas and incentives. Understanding each lends the modeller a wide and unbiased perception of the whole operation and will assist in building a model that represents the reality reliably.

The second part of the conceptual design process is

- to plan the model
- to decide where to put more detail and what assumptions to make safely
- to thoroughly think through the methodology that has to be used for the particular model
- to assess the data needs and to make sure all software interfaces are considered
- to make sure the model caters for all the user requirements

2.3.5 Scenario planning

This step in the process looks at all the possible scenarios that can be evaluated to test various options for optimisation or solving of the problem. Because this step happens before the model is build no information on bottlenecks is available and the planned scenarios is based on brainstorming and coming up with possibilities and alternatives already considered in the business as possible solutions. This phase is necessary at this point so that the modeller understands what the variables are that needs to be flexible and that will probably have to change. This process ensures that the modeller do not fix parts of the system in assumptions that will require major changes later. Because of the nature of stochastic modelling and the growth that takes place throughout the project the scenarios will change. From the base model and through identification of bottlenecks some new scenarios will come up. It is therefore important to plan in the schedule and cost for at least five scenarios that are not part of the ones identified at this point.

2.3.6 Study cost and schedule

If it is possible for the specific client it is prudent to split the project here and bill for study cost for the first phase as that can take anything from one to four weeks to get to this point. Out of the steps up to this point a reliable time schedule and cost estimate can be calculated. Without the pre-work a stochastic modelling project may vary in cost and time between 20% and 50%. If the planning up to this point is done thoroughly the cost variation will be less than 10% and the project can be delivered on time. This box is not added to the proposed solution as this is dependant on who is doing the simulation. An in-house simulation does not necessarily require this step.

2.4 Identifying and Acquiring input data

2.4.1 Proposed change

The box for Identifying and Acquiring Input data and the box for Developing a model as they are arranged in Flow diagram 2.1 are swapped in Flow diagram 2.2, the proposed solution, as no simulation model can be built without the appropriate data and decision rules.

2.4.2 Getting the data together

From the conceptual model it is now easy to identify the data needs. From the perspective of the modeller it is important to give the client the responsibility for the data including the gathering thereof. The modeller can request that the data be delivered in a specific format electronically and the gathering of the data should be in the time schedule as a time component but not a cost component. Analysing the data however is the responsibility of the modeller.

Fitting curves and understanding correlation is an important component of this phase. Correlation that exists in the input data can have a major impact on stochastic modelling projects and where possible the correlation needs to be accounted for by using a regression function to calculate the correlated variable. Instead of using a calculation the correlated variable can be sampled from a distribution, without accounting for the correlation. There is an inherent risk present in that situations may occur that would not occur in the plant or that the variation in the model may be even more than in the plant. Where a risk factor in one area can lead to a direct risk in another area, the effect may be missed. These risks are less where the correlation pattern is inconsistent, which happens quite often in plant data as there are so many unrelated influences on plants and streams.

An important feature of a thorough data analysis is that some of the impacts, interactions and problems can already be identified before a model is built. It is therefore necessary to keep this in mind while preparing the data for the model. It is also good practice to give feedback to the client at this point.

2.5 Developing a model

When the data is available the model development can start. The data gathering and model development may be done interactively but it often results in a bit of rework and not in as much of a saving as expected. The decision to combine the steps depends on the model and data requirements and at this point the modeller will be able to assess the risk and make the decision. Building the base case model after a thorough conceptual design is a formality and it really only takes the time to build the logic. This part of the process is the heart of the study. Through good planning it can be short and smooth. The model development can be broken up in different phases or parts but the overall goal should always be kept in mind.

2.5.1 Proposed change

When the model is completed verification and validation should formally be done on the model to make sure that the model fits the current reality and also that the model logic is correct.

2.5.2 Verify and validate the model

This step can be combined with developing the model because without this step model development is not complete. Verification and validation entails making sure that the model adequately represents reality and also that the model itself is sound and that whatever is generated at the beginning of the model moves right through to the end following the correct routes. For stochastic models this step also means that the modeller evaluates the model with the client and gain acceptance for the model.

The client makes sure that he is happy that the model is in fact an accurate representation of the system. The way to present the model is for example in a chemical loading area is to write out the orders generated and the markets and volumes supplied for the specific products and compare these results to the input data. Plant and tank failures frequencies and durations can also be compared to the input data. Another test of the model accuracy is to check the sequences of products being made in the model using the rules to the actual sequence. From these results it will already be evident where queues form and where there are bottlenecks in the system. This step can therefore be combined to some extend with one of the following ones of analysing the results because it is often unsure where the one ends and the next one begins.

Animation is another tool that can be used to verify and validate a model. It makes any errors visible and it also helps the client to see the events happening in the model.

2.6 Developing a solution

Developing a solution in stochastic modelling means to take the base case model and test the identified scenarios and to deliver the model to the client in a usable way.

2.6.1 Proposed change

Build Scenarios is also added as a separate box to the proposed solution in Flow diagram 2.2 to build in all the possible options for improvement.

2.6.2 Build scenarios

The scenarios tested in the model may for example for chemical loading involve

- increasing storage space
- increasing number of tankers
- increasing workers
- increasing loading areas
- changing loading rules
- changing working hours
- increasing pump rates

Or any combination of the above.

Experimental design can help narrow the options for testing and may broaden the result space with some risk as the whole area is not explored. From the results of the scenario tests new options may arise and a lot of information is gained.

2.7 Testing the solution

From the scenario runs some new scenarios may be identified that involves specific combinations of changes. In contrast with other types of models in the case of stochastic modelling the model itself and the learning through exploration leads one to the solution and the solution needs to be tested and verified. The final test of the solution is a comparison between the model and the changed system.

2.8 Analysing the results

Results analysis is an integral part of the base model as well as each scenario and care must be taken to use the correct and clearest results. One possible constraint is the sheer volume of information that is available from a stochastic model and the larger the model the more daunting the task. The best way to approach this is to develop a set of specific questions to ask or areas to focus on. The sensitivity analysis for the stochastic model is as important for stochastic models as for other types of models "Because input data may not always be accurate or model assumptions may not be completely appropriate, sensitivity analysis can become an integral part of the quantitative analysis approach" (Pegden *et al.*, 1995:6). It is necessary to note that sensitivity analysis should be done before developing the final solution because it better describes the base model. The final results analysis can remain here as it combines all the hard work that was done into a final report or presentation making sense out of an immense amount of information and growth. The result is often only a recommended course of action but may in some cases be a model that can be used further.

2.9 Implementing the results

Implementing the results for a stochastic modeller is seldom part of the picture. The only implementation a modeller is sometimes involved in is implementing the model in the environment and training the new users. Client feedback can assist in evaluating the value of the implementation of the proposal.

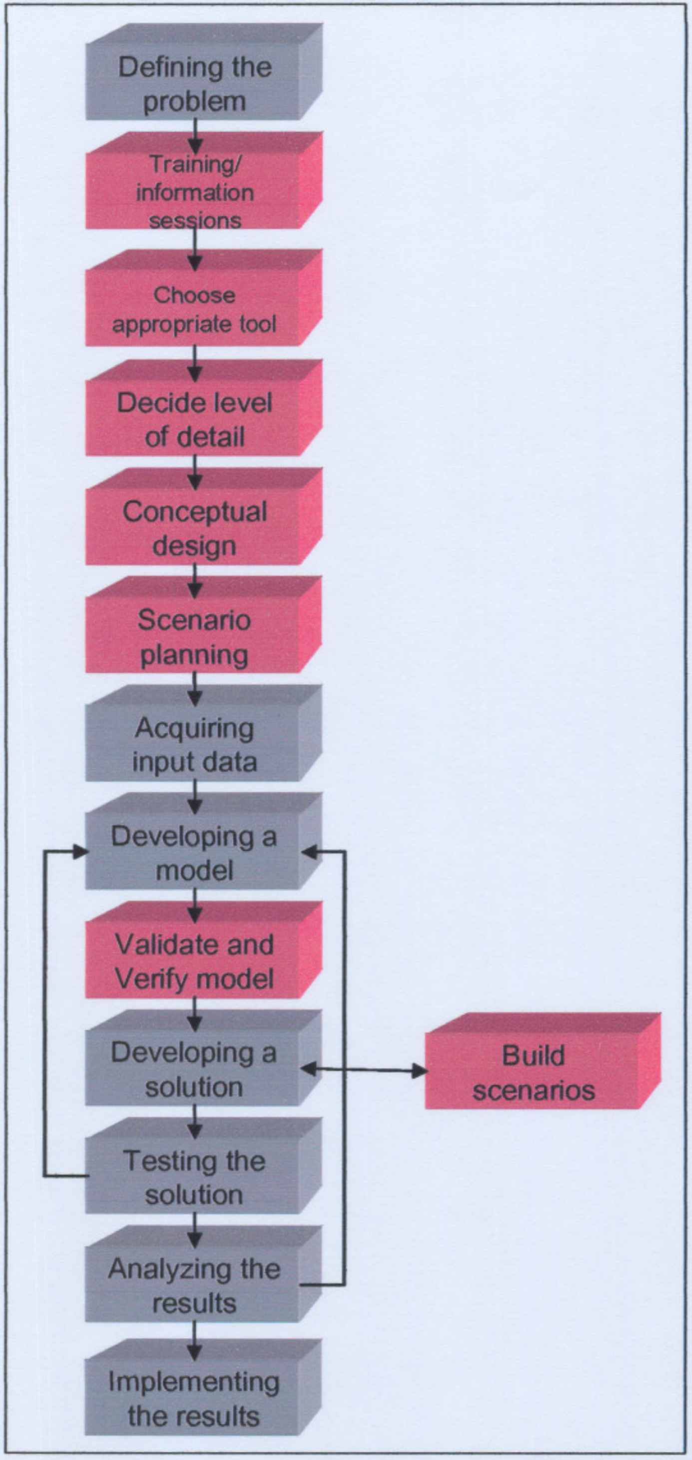
2.10 Cyclical nature of the process

For each of the steps in the process it is possible to backtrack and improve. After each step the success of the completed step needs to be evaluated and if it is not completed to satisfaction or to set measures it needs to be redone so that the next level can be build on top of it. Poor foundation will form a poor building; poor planning will results in a less than satisfying model.

2.11 Conclusion

The final proposal can be seen in Flow diagram 2.3. Each step plays a significant role in delivering models that are reliable and usable. The documentation required for these studies can be discussed but as far as the process for developing a model is concerned, the main ones are included. Other aspects for sustaining of models, like service level agreements, may form part of the process if required.

Flow Diagram 2.3: Final proposal



CHAPTER 3 CONTRIBUTIONS TO MODELLING IN SASOL

Stochastic modelling is a tool seldom used in continuous environments especially when new plant or processes are designed. It could however add a lot of value in this specific environment because of the dependencies between plants and the effects of one plant that is precipitated throughout the environment. When a plant is designed in isolation the risk is that the actual throughput when the plant is started up would be overestimated. Some of the reasons for this overestimation may be the interaction of various maintenance schedules or the interaction of the inherent variation in each plant with the surrounding plants. It may also be related to physical constraints of pumps or the supply and demand logistics.

This chapter discusses stochastic modelling in the Petrochemical industry and where and how it is used. An example is discussed where stochastic modelling is used in fuel blending and distribution as well as an example in calculating plant throughput given variation, plant availability and interactions between several plants.

3.1 Introduction

In the past few years stochastic modelling was often used in Sasol although it is not often used in continuous and therefore petrochemical environments. The value that can be added using stochastic modelling is easily underestimated because it is mostly seen as a logistics or discrete modelling tool only. In some of the projects done in Sasol the perceived boundaries of stochastic modelling were challenged and excellent results were obtained. Stochastic modelling has proven to be a tool worth considering in the modelling toolbox of petrochemical and other continuous industries.

A brief description of stochastic modelling and its capacity is given in this chapter. It is then followed by how stochastic modelling was and is currently used in Sasol with a few examples.

3.2 The Sasol picture

The Sasol vision as seen in Figure 3.1 emphasises the intention of Sasol to compete as a global enterprise and to excel in selected markets in a respected way.

Sasol is one of the only petrochemical companies in the world that derives fuel from coal, see figure 3.2. With its focus on higher value chemicals and changing markets, the supply streams to the fuel pool are changed frequently and the components available to make the different fuels are adjusted accordingly. A change in any part of a stream anywhere in the process inevitably affects the volume and properties of at least one component used in fuel blending and other plants or processes that may be supplied by other streams of the same process.

Sasol's market drive and the customer focus value ensure that the required volumes of high grade fuels and chemicals will be supplied to customers. The demand for increased levels of unleaded fuel as and when the market requires naturally causes a very dynamic environment.

This dynamic environment offers a challenge for a system that requires pumps, lines and tanks - physically fixed infrastructure that cannot easily be changed. The challenge is to ensure that, for every change in market requirements, the fuel pool is balanced and stable, the market needs can be met all the time and the infrastructure needed to sustain each specific scenario is indeed sufficient and available when needed.

This flexibility comes at a price. Through the work done on fuel blending at Sasol between 2001 and 2003 and by observing the effects of changing these blends continuously, it seems evident that the price is paid in infrastructure. It was observed that changes in plants or processes affect the whole system in many ways due to the interactions between plants and processes. This price can be minimised through holistic planning and by really understanding each process involved in fuel blending very well.

One potential difficulty is to have the right components available at the right time in sufficient volumes. Lost production may often never be recovered. For example, if a process is turned down in a continuous environment due to whatever reason, it cannot necessarily be turned up sufficiently again to make up for lost production. Tank constraints, process capacity constraints on both the receiving and supplying ends as well as physical plant constraints may decrease the capability to recover lost volume. Deterioration in properties may also impact negatively on total supply. If any blend is made to higher purity or specifications than what was required, the loss in purity of

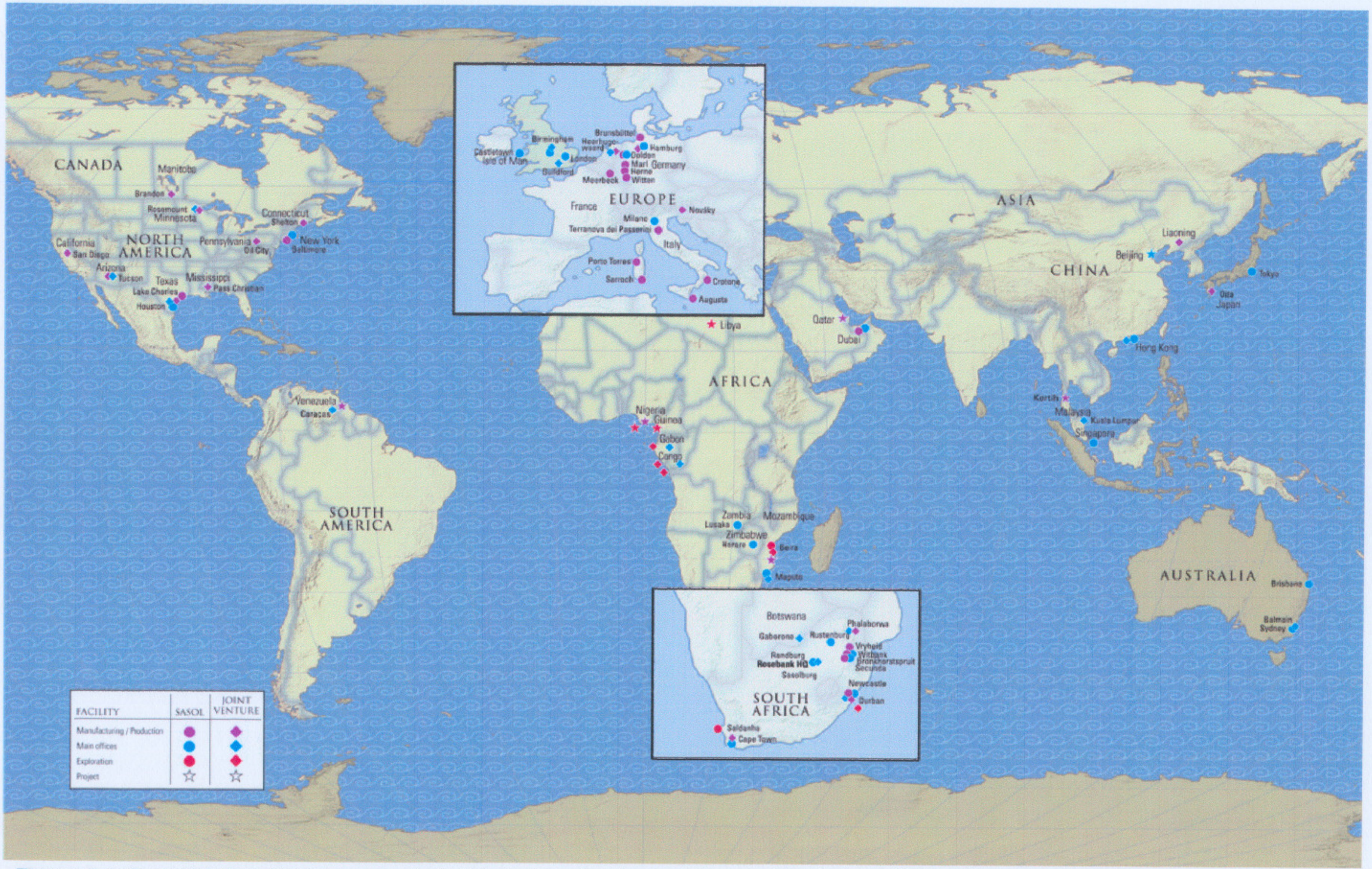
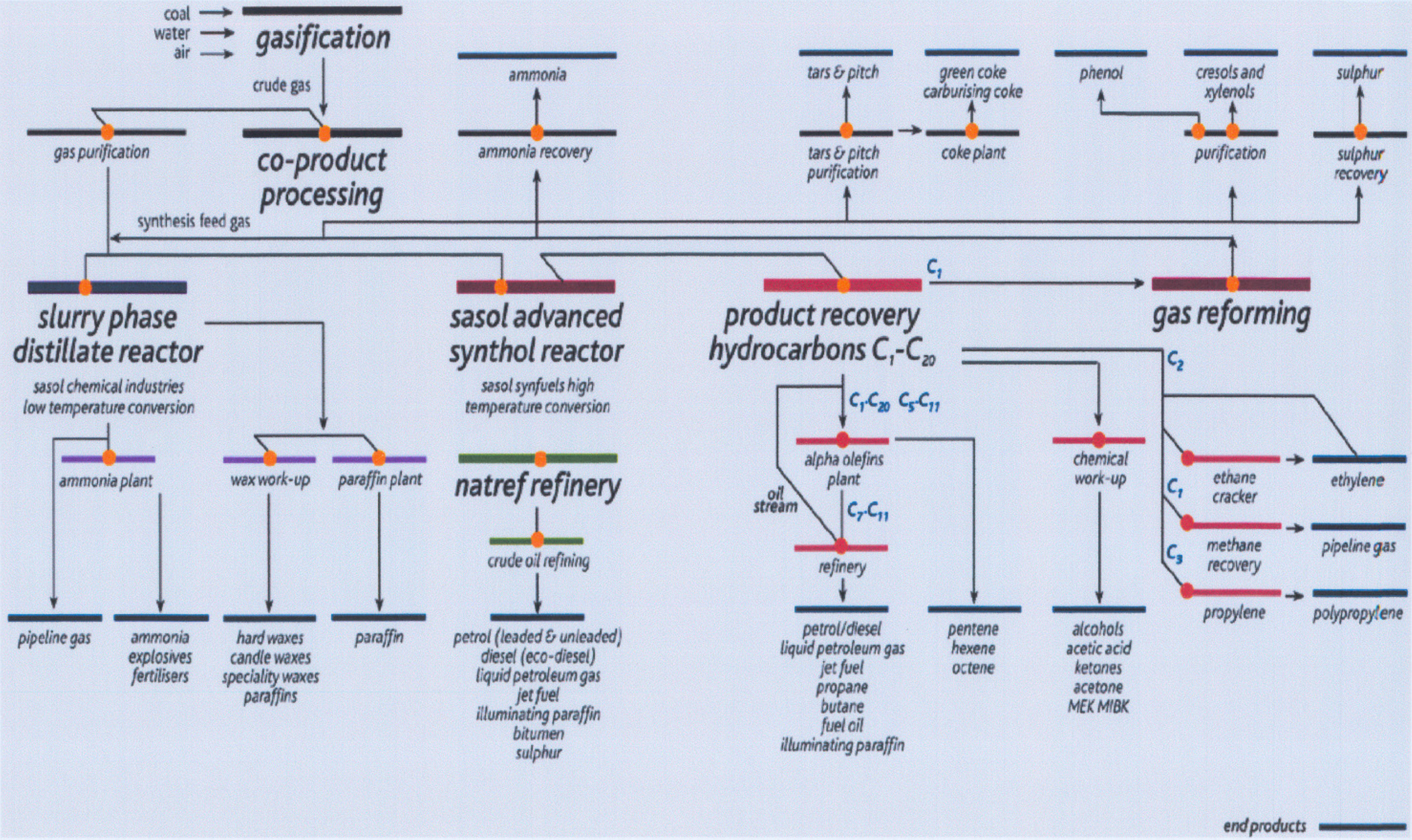


Figure 3.1: Sasol Vision: *To be a respected global enterprise, harnessing our talents in applying unique, innovative technologies to excel in selected markets in energy, fuels, chemicals and related sectors in Southern Africa and worldwide*

Figure 3.2

The Sasol Process



that property cannot be recovered to help improve the properties of a following batch. An example is when a blend needs to be made to 93 RON (Relative Octane Number) and it is made to 93.3 RON (Relative Octane Number). Octane is very expensive and difficult to come by. Losing .3 RON (Relative Octane Number) specs has a monetary effect and that .3 is lost for future blending.

3.3 Stochastic modelling

Stochastic modelling could be defined as follows:

"Generally, stochastic (pronounced stow-KAS-tik, from the Greek stochastikos, or "skilled at aiming," since stochos is a target) describes an approach to anything that is based on probability. In mathematics, a stochastic approach is one in which values are obtained from a corresponding sequence of jointly distributed random variables. "(Anon Whatis, 2001)

"Stochastic processes concern sequences of events governed by probabilistic laws" (Karlin & Taylor, 1975)

"[A] stochastic process is a variable that evolves over time in a way that is at least in part random"(Dixit & Pindyck, 2004:60)

"A stochastic model predicts a set of possible outcomes weighted by their likelihoods, or probabilities" (Taylor & Karlin, 1998:2)

These definitions provide a feeling of stochastic modelling, emphasizing the probabilistic basis thereof. The definitions imply that all states of a process do not have the same chance of happening and that an average may not in all cases be sufficient to describe a system or process. The second important principle derived from the quotes is that stochastic modelling models sequences of events through time. The power of stochastic modelling lies in these principles.

Stochastic modelling could be defined therefore as the modelling of a physical process taking into account the variability in volumes, properties, availability and anything else of a variable nature that is involved in that process as well as taking into account the timing of events and the effects of "bad" timing.

3.4 Stochastic modelling in the Petrochemical environment

Stochastic modelling is well known in discrete environments of which a supply and demand process is an excellent example. Logistic infrastructure can be discrete or discrete/continuous.

Often asked logistical questions from various environments include the following:

- What are infrastructure needs at the ports?
- How big must warehouses or tanks be?
- How many tankers, warehouses, tanks, front loaders etc is needed?
- Can the intended volumes be supplied to the ports in the intended time?
- Can the product be delivered to the client within the required time constraints?
- Must road tankers or rail transport be used?
- How does the shipping schedule influence availability of products or tank needs?
- How do production rates, maintenance and shortages influence the supply of products?
- Can the loading areas accommodate all the products?
- How do rail infrastructure constraints influence availability of product?
- What are the infrastructure needs for a specific business unit?

The logistics infrastructure could enable or damage any company because it represents the final and personal interface with the customer. How infrastructure can function optimally is therefore something that needs to be understood well and careful analyses and planning of these systems are required.

A good linear programming (LP) model can suggest an optimum solution that adheres to all requirements and optimise profits, but these suggestions might be difficult to apply given infrastructure constraints. One can measure to some extend how well the optimum solution from the LP performs and where problem areas can be expected in the daily running of the plants by testing the optimal solution in a model representing the logistics infrastructure where system constraints, interactions and variation are included. This is where stochastic models come in –

their edge being that they are able to represent the infrastructure and time effects and that they can identify where major problems can be expected. When the two techniques are used together, a better solution can be found that makes it possible to limit infrastructure changes and optimally use what is available.

3.5 Continuous environment

Stochastic modelling is very seldom used to model continuous processes, because in this environment processes are often sequential and if one part of the process is off-line, all other parts of the process are off-line too. When this is the case, stochastic modelling seems an over-investment.

The above however is only half the truth. Even continuous processes are highly variable and dependant on previous and subsequent processes. In practice it is often seen that plants or processes are seldom if ever capable of their published maximum capacities when placed in a total interactive system. Ramp-up and ramp-down rates of processes often prove to be less sustainable than expected. The reason for this usually can be found in the sequence of processes and the total picture that invariably displays a lower capacity. In all continuous processes there are a supply stream, coming from a variable feed product, and final or intermediate products going into tanks or other forms of storage.

The power of stochastic modelling is that it can "catch" the above variability. It can include the pumps and accessories around the process and their effects on the process and other processes around it. For example: A plant can be down for two to eight hours because of a failure of a pump. If it fails for less than three hours a hot start is required, but if it fails for longer than three hours the plant needs to be cold started. A cold start for the specific plant may take three days. A cold start means that the plant needs to cool down, maybe that the catalyst needs to be changed and then a temperature sequence needs to be followed for start-up. A hot start means that the plant can continue from the point that it was at, at the end of the failure. Due to how this plant operates, this type of incident may decrease the average capacity of the process.

With a stochastic model one can easily test different maintenance and failure scenarios, different ramp-up (the sequence of rates used to start up a reactor or plant) and ramp-down rates (the sequence of rates used to switch of a plant, sometimes referring to for example cooling down time), different operating procedures and the effect of new equipment installed in existing processes. It is possible to see if a planned throughput will be feasible or sustainable. With planning a new plant or process, stochastic modelling can assist with sizing of equipment to minimise risk without increasing the capacity of plants. Previously plants or tanks were often oversized to minimise risk or "for in case" because even though the cost implication of too much capacity is severe, the loss due to a plant or process being unable to meet its demands is much more serious. A stochastic model can model the new or changed plant or process as part of the existing system and show the effects of interactions not necessarily planned for. It is therefore not necessary in this case to build in extra capacity.

An example of a problem that can occur in a continuous process is as follows:
A plant has an average capacity of 10 000 tons per month and a turn-up capacity to a maximum of 15 000 tons per month. The average yearly capacity for this plant is therefore around 120 000 tons. This plant serves a seasonal market – that is the reason for the big difference between the average and maximum supply. During a specific year the required market need was on average 6 400 tons per month for about seven months. This adds up to 42 000 tons supplied during these seven months. This in effect means that if the total market commitment was 120 000 tons for the year that for the remaining five months a steady supply of 15 080 tons per month is required. This is not possible, except if some of the product was stored in tanks during the low seven months. It needs to be taken into account that even if the required production was 14 500 tons per month, it may still have been difficult to sustain given the sequence in the process and the fact that when variation is taken into account, no process sustains design capacity. Design capacity is often quoted in isolation with an assumed sustained supply of feed streams. When a plant attempts to run continuously at maximum capacity failures also increase due to scaling, catalyst failure or equipment failure and the intended capacity is not met.

3.6 Constraints of modelling technologies

It is important to note that no tool is perfect and that each tool has some advantages and disadvantages. Some constraints currently on many stochastic modelling packages are the following:

- Limitation on automatic optimisation
- Time needed to build the models

- Detailed information required
- Need for the modeller to learn and understand the system to be modelled
- Understanding of the results in a sensible way because of the volume of data available after a run is complete
- Danger of modelling causes and effects twice because of dependencies or correlation that are missed
- Continuity in models and modellers

Many of these constraints can be observed in most other modelling technologies as well. It still is important however to take note of them because of the effect they may have on the budget to build a model.

3.7 Blending and Distribution example

In the Sasol process, the changing market needs in terms of fuel was observed and the question arose as to how Sasol can meet these demands at all times and be able to supply their customers at the same rates and qualities or better.

Some of the questions that needed to be answered for each change in unleaded market demand to reach around 55 % Unleaded petrol where,

- What needs to be changed at Sasol?
- Can each new supply level be sustained?
- Will the infrastructure suffice for each step?
- And if not where are changes required?

Also in the case of short term interim scenarios where unleaded demand may be equal or slightly more than premium petrol demand, what will the effects be on the current infrastructure, with all the types of leaded petrol still supplied to the market. Lastly will it still be possible to adequately predict each scenario given our previous experience and knowledge?

Historically a blending solver was used to evaluate each scenario and judge the different fuel capacities in each scenario. This is an excellent tool and it proved itself well over the years. What was noted however was that for some of the scenarios it was difficult to sustain the proposed blends in the actual blending plant. The reason for this seemed to be infrastructure constraints not included in the model but it was difficult to prove because of the fact that the total volume sold stayed the same – only the component rundown slate changed or the demand for a specific fuel type increased somewhat. A method was needed to better understand the effect of each scenario on the actual blending plant. In 2000, stochastic modelling was identified as a possible solution to identify the bottlenecks and to make sure that the predicted unleaded petrol capacity for each scenario was truly in reality what the system could supply.

Stochastic modelling in 2000 was not widely proven in Sasol. It was already used in distribution and logistics models elsewhere in the company with a lot of value, but it was virtually unknown in the wider Sasol and basically seen as a tool used only to model job transport and logistics. An internet search provided no more information on stochastic modelling in the Petrochemical industry. It was widely used in the different mining areas as well as in logistics or supply chain modelling and a lot of information was available in those fields. No publications on using stochastic modelling and linear programming together could also be found. This provided an excellent opportunity with a fairly high level of risk.

The questions were:

- Can stochastic modelling fill the gap?
- Can stochastic modelling and optimisation techniques like a Solver® in Excel® be used together to give a value adding answer to the business.
- Would this approach be "the right thing to do"?

The answer for Sasol proofed to unequivocally that stochastic modelling could fill the gap and the two techniques together resulted in excellent benefits for the company.

3.7.1 The task / project

A request was received from a part of the Sasol business for a stochastic model of the petrol blending and distribution system to be able to test several future blending scenarios. The dynamic highly integrated nature of the system indicated stochastic modelling as an option. Difficulty obtaining reliable, repeatable results that mimics all the operational constraints when using other modelling tools also showed that stochastic modelling may be a feasible option. *Training of the client* to understand stochastic modelling in this case was therefore not necessary and the *tool was prescribed* by the business. Arena® is the stochastic modelling software currently used most in Sasol and the support structures were in place to support long term modelling projects.

The first step in the process was to adequately *understand the system* that needed to be modelled and the *level of detail* required. Some of the information needed was the operating procedures, rules and options in the system, the driving forces, the market demand and the interactions. The goal of the first step is to calculate the magnitude of the study and make sure that the objective of the study is thoroughly defined and understood.

One of the first decisions that were made was that only petrol will be included in the first phase of the study as the influence on other fuels were not expected to be as substantial. The second important boundary that was imposed on the study was the decision to not model each plant that supply product to the fuel pool in great detail. Each plant which runs down final products (as an example U1 – Unit 1 and U2 – Unit 2 as shown in figure 3.3) is seen as a black box with a few specific rules and a supply stream to the blending plant.

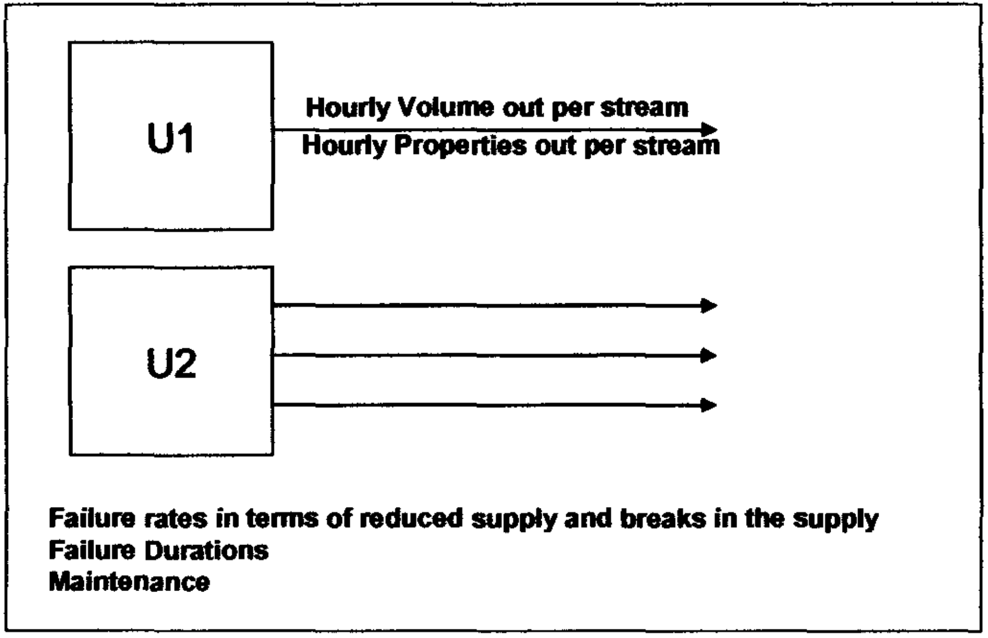
The idea was to capture short supply or decreased flow due to breakdowns or upstream upsets and also the normal routine maintenance like catalyst changes and pump checks. Shutdowns were initially included in the base model.

A consulting engineer drew up a realistic flow diagram with each and every pump, tank, line and connection in the whole system and then the tedious process of understanding the system started. The data needed in the study was identified during an interactive period working through the flow diagram and assessing the data needs with the relevant plant and modelling people.

While the data was gathered flow diagrams were put together on how to approach the model. The existing system was divided into more manageable parts by identifying the drivers and approaching the model from these different aspects. A lot of work went into writing up the logic flow and trying to understand how decisions are made at each point in time and exactly what decisions are made at these points. In chapter 2 the *conceptual design* and *scenario planning* were done as separate steps. In this project a lot of the design happened interactively as more data became available and the picture of the system became clearer. The scenarios were provided by the business and although the approach was followed that scenarios can go beyond those described it was also decided that at least the flexibility to cater for the specific scenarios required, needed to be included during the conceptual design of the model.

Volume data for each stream was collected from an in-house database with hourly measured data. Three months data for a perceived stable time period was used for the study. Properties that constrained the system were identified and all information on these properties available for the previous two years was also gathered. The actual orders supplied for 1999 was taken as a basis for the market demand. Here the orders for the other fuel types were also considered as they had an impact on the infrastructure needs on the product distribution side.

Figure 3.3 Units or plants and their component streams



3.7.2 The major parts of the model

It took around three months to gather all the data and put it together in a sensible way. The model was estimated to be developed within about three more months and then scenario testing would start for which another three months were scheduled. When the process was started it was based on the 80/20 idea. Twenty percent effort for eighty percent gain. Many assumptions were made and the model was broken up in three major parts.

a) Part 1: Component tank feeds

The supply streams from the different plants are to a large extent exhaustive. Everything that can be made is made and the goal is to mix all the streams into high quality fuels. Each plant supplying into the petrol pool is seen as a separate business with separate goals. When changes are made to these plants to sell more product or to change their product slate an impact study is done to make sure that the petrol pool are not adversely influenced. If the decision is then made that the pool will still be sound, then the changes are made to those plants and business continue as usual. In specific circumstances these decisions can be overturned again but the basic approach in most cases is that when this new way of working is approved the petrol pool will take whatever comes from this plant and use it. The impacts were evaluated with the Solver for the current system in each case. In the current scenario Sasol is able to sell all the petrol that is made. This means that there is no additional petrol storage except what is necessary to run the system and to keep within governmental constraints. There are also a few "levers" that can be pulled in the petrol system. One of these levers is Platformer severity. If needed the Platformer severity can be increased for short periods to increase RON (relative octane number) and reduce a bit of volume at a certain cost. This is also only true in scenarios where the Platformer is not run consistently at its highest severity.

In the model the system is run with no interventions – i.e. fatal make. If a scenario shows for example that there was too much Platformate or that the severity needs to be increased, that is run as a scenario. This can assist the blending operator in deciding what interventions may be required for a specific scenario and what the operating philosophy should be. In the model this supply streams exist as a separate model where each unit have a rundown stream or streams with specific volumes and properties based on the fitted distributions. Each unit also has mean time between failures (MTBF) and mean time to repair (MTTR) duration distributions for the normal plant upsets and failures and the smaller maintenance procedures. The different streams are then mixed as they run down to specific component tanks.

Each component tank or set of component tanks has a list of priority rules for use. If a tank is available the volume and properties are mixed into that tank linearly or in some cases with an offset to come nearer to the actual property impact. If these tanks are full and alternative tanks are available the component stream is rerouted to that tank. If not the stream is calculated as overflow and the product is lost for use. In the actual system an intervention will be done to limit the loss of these components, but the basis of this study is to run the system as it should be run with little or no interventions and to calculate the need for interventions in a specific scenario. For each of the streams a maximum overflow that can be handled with small interventions is calculated. This is seen as manageable overflow. The rest is then identified as part of the risk of running that specific scenario.

b) Part 2: Market demand and distribution

The main driver of the system and model is the market. The market drives what is required in terms of the product slate and the timing of blends. There are three modes of transport, road tankers, rail tankers and pipeline. The actual orders for 1999 was used as a basis for the market demand and throughout all the scenarios the distributions were kept constant but the supply was moved up or down depending on the specific scenarios. High volumes were needed at specific times for the pipeline and road and rail tanker orders were needed as they arrived. This was the easier part of the system in terms of modelling. Orders were known for the pipeline one week in advance. Average volumes were estimated for the following week for road and rail and the orders were filled as they arrived into the system.

At the point of arrival the order would queue if the loading points were not available yet. If a point was available, an available tank, with the correct product on specification and fully analysed, needed to be found. If the product was available and the pipe infrastructure to get from the tank to the loading point was available, the order was filled. If not the order waited until it could be filled. The procedure for the pipeline is the same as for the other transport modes. If an order needs to be pumped and the infrastructure and tank is available it starts to pump. If not it is delayed by ten minutes and it re-enters the queue as the first order to be filled. If another order was in the queue at that point in time and it had to be pumped, it would have tried to start the pumping process during the time that the previous order waited to rejoin the queue. In that way if a single product became unavailable it did not stop the whole queue, but it was postponed. The frequency that a pipeline order is postponed is an indication of system problems. The frequency of unavailability of final tanks available to pump is indicative of problems in supplying the fuel on time. The unavailability of loading points is indicative of infrastructure limitations on the distribution side.

c) Part 3: Choosing and making a blend

This part of the model is the interface between the supply of the feed streams and the supply to the market demand. This is where supply and demand is matched and blend size is calculated. This is also where interventions can be flagged and therefore this forms the heart of the model.

The first priority of the blend logic is to decide what blend and in what volume is needed to supply the market for the next week. The second priority, if there is already sufficient volume for one week, is to decide what blend needs to be made proactively so that component tanks do not overflow or the pool becomes imbalanced for future orders. The blending logic interfaces with a smaller version of the blend solver that will help the blending logic to optimise the next three blends.

3.7.3 Building the model with ARENA®

In Arena® the blending logic selects the next three blends that need to be made until the market is filled one week in advance. When components need to be pumped away or a blend can be made because the blender is available and there are sufficient components to make a blend, only one blend is calculated. The reason for choosing blends in advance is if one optimises a single blend one may compromise future blends in the petrol pool. If one looks at the current blend and the blends for the next week it is easier to optimise to a working blend. A working blend here is a blend that will adhere to all requirements, but also one that will not unnecessarily unbalance the petrol pool. If no acceptable blend can be found then an optimum suggested blend (default blend) is made. The frequency of the occurrence of this default blends flags that the pool is imbalanced and that it does not have the required components to blend correctly. However if one do not make use of the default blend the model sometimes choose a blend to far from the optimum and the pool cannot readjust. This problem has the potential to render the model results unusable.

a) Run time

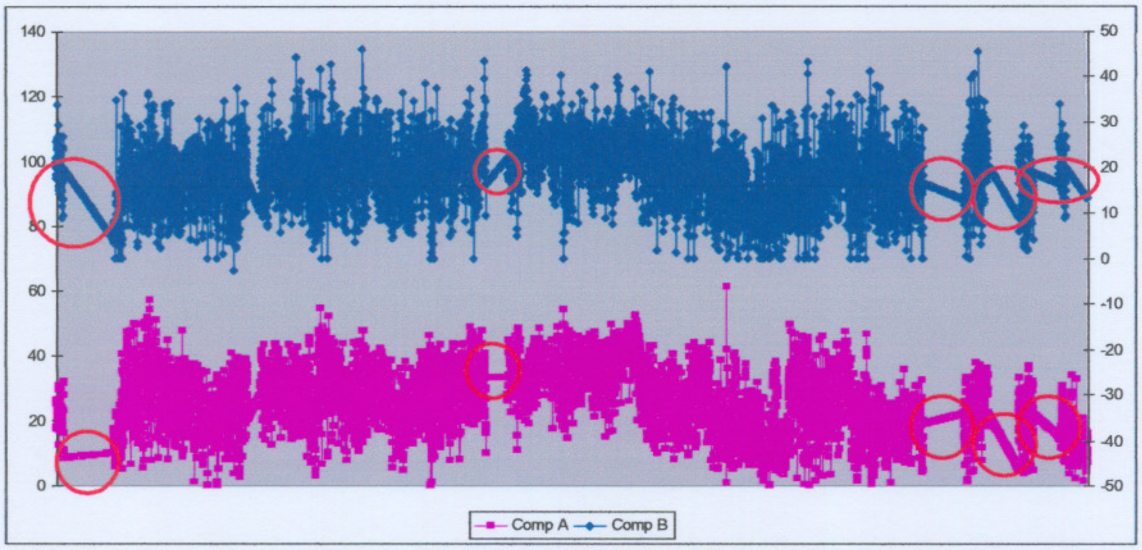
The model was planned to run for a yearly period at a time generating stream volumes in 20 minute time intervals. This is a discrete approximation of a continuous system with small enough time buckets so that the volume impact is negligible.

b) Analysing the incoming stream data

The data received from the in-house database was difficult to interpret. It was one thing to take the hourly data for three months and another to interpret it and fit sensible distributions for it. As an example two streams are drawn on a graph to explain the process used.

The red circles in Graph 3.1 are failures modelled separately, it may be plant upsets or catalyst changes or even shutdowns in a specific unit supplying both streams. It may at some points also be the database that was unavailable due to network problems. The first task was to identify what each failure was and the impact of each on the stream.

Graph 3.1 Time series graph of two component streams



Those failures that could be contributed to stream unavailabilities resulting from the units supplying the streams were taken out and modelled separately. The rest of the data was then used to fit distributions.

Distribution 3.1 shows one such distribution. For Distribution 3.1 a Normal distribution is a very good fit except for the one bar of the histogram that has a much higher chance of happening than the rest of the bars. One relatively simple solution is to remove some of the values in the indicated bar and add them as another distribution with a chance of occurrence based on the number of values removed.

Distribution 3.1 Histogram with a Normal distribution fit – full data

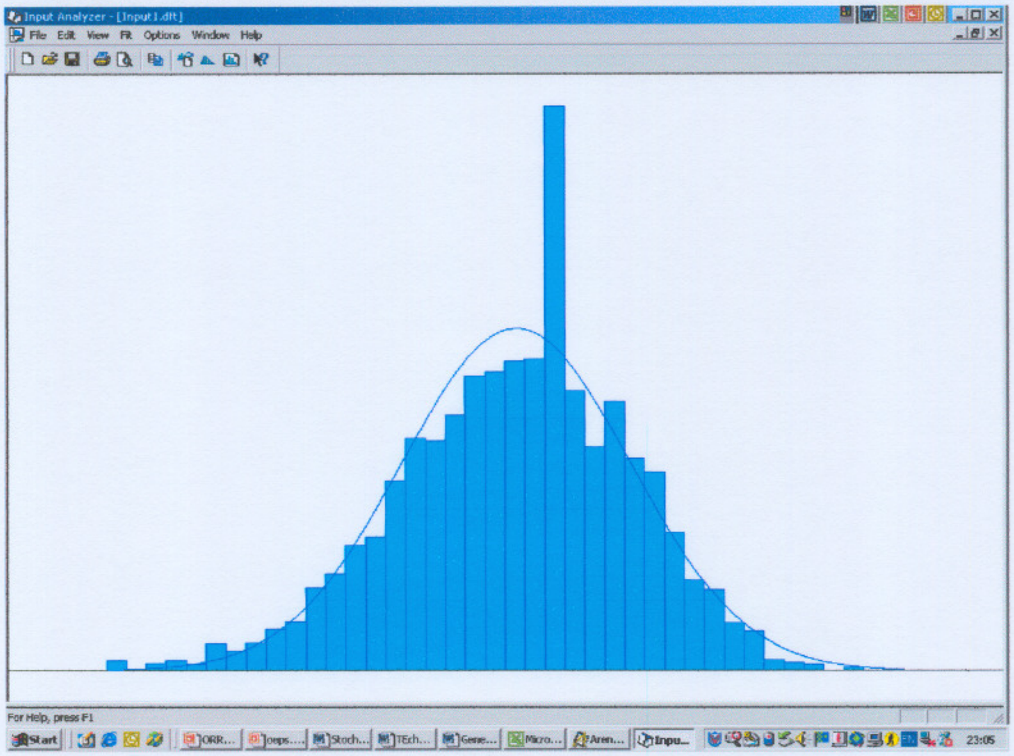


Table 3.1 Results from a Normal distribution fit in Arena

Distribution Summary

Distribution: Normal

Expression: NORM(29.8, 8.35)

Square Error: 0.002963

Chi Square Test

Number of intervals = 31

Degrees of freedom = 28

Test Statistic = 245

Corresponding p-value <.005

Kolmogorov-Smirnov Test

Test Statistic = 0.0434
Corresponding p-value < 0.01

A good indication of a good fit is in the first case a very small Square Error. But the statistically more reliable one is the p-value. It has to be in excess of 0.05. The nearer it is to 1 the better. "…as with any statistical test, a high p-value doesn't constitute "proof" of a good fit – just lack of evidence against the fit". The ***Chi-Square test*** and the ***Kolmogorov-Smirnov*** tests are explained in Simulation with Arena® (Kelton *et al.*, 2002:151). In distribution 3.1 the fit is not good.

Data Summary

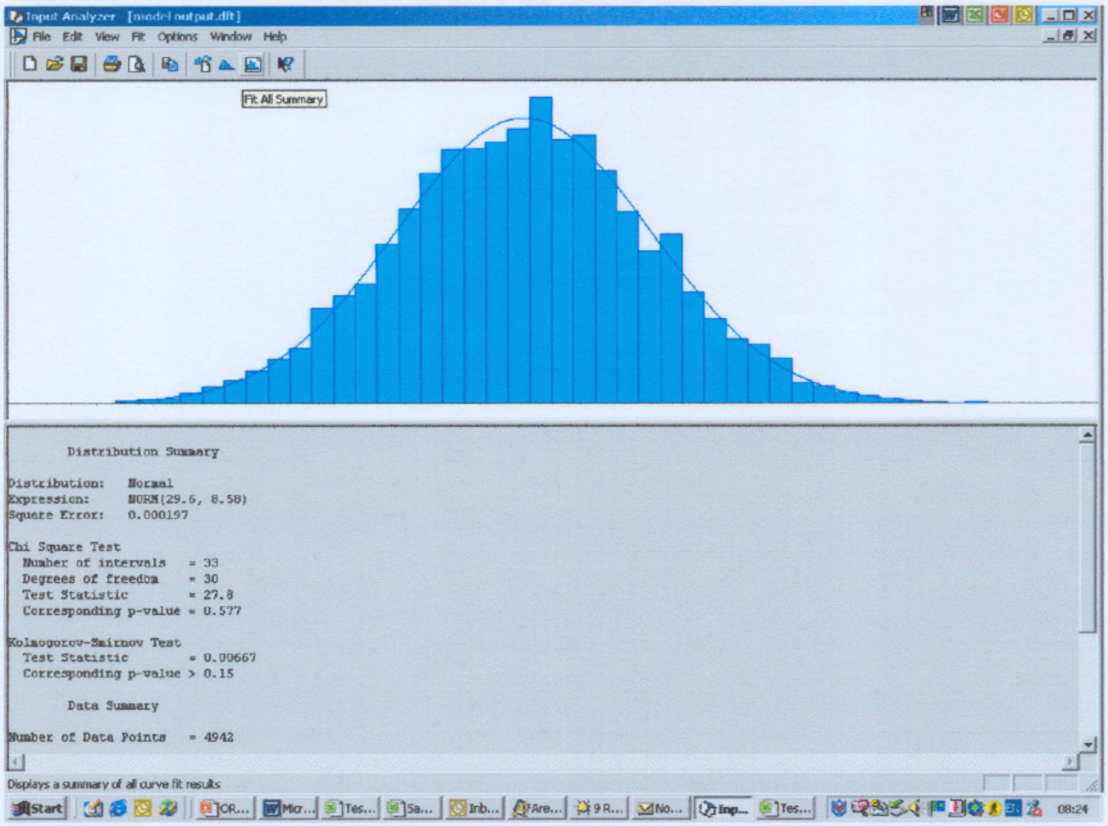
Number of Data Points = 3839
Min Data Value = 0
Max Data Value = 57.1
Sample Mean = 29.8
Sample Std Dev = 8.35

Histogram Summary

Histogram Range = 0 to 58
Number of Intervals = 40

The two new distributions are NORM(29.6, 8.58) and 32+2*BETA(1.24,1.84). There were 3839 data points for the normal distribution and 194 for the beta distribution resulting in a 4.8 percent chance for the Beta distribution to be used and a 95.2 percent chance for the Normal distribution to be used.

Distribution 3.2 Histogram with a Normal distribution fit – partial data



```
        Distribution Summary

Distribution:   Normal
Expression:     NORM(29.6, 8.58)
Square Error:   0.000197

Chi Square Test
  Number of intervals   = 33
  Degrees of freedom    = 30
  Test Statistic        = 27.8
  Corresponding p-value = 0.577

Kolmogorov-Smirnov Test
  Test Statistic        = 0.00667
  Corresponding p-value > 0.15

        Data Summary

Number of Data Points   = 4942
```

In distribution 3.2, the new Normal distribution has a p-value greater than 0.15 indicating a good fit. The generated data over time can be seen in Graph 3.3. The patterns and time dependency in the historical data are lost in the generated data, except where the maintenance will occur based on their own distributions at the required points in time. This technique is far from optimal, but may be a conservative approach to an extended problem.

Risks identified

1) When several streams and their properties are analyzed it is evident that there are several dependencies. There is correlation between different properties in a single stream, there is correlation between properties of different streams and there is correlation between the volumes and flows of some of the streams. In all cases the dependencies are not consistent.
2) When evaluating the existing system it became apparent that at certain points in time depending on several factors not included in the model, the rates at which streams will run down will be controlled by a global planning intervention.
3) There are also cases where properties were influenced by an intervention in a specific plant or unit by changing properties like temperature and pressure on a

unit.

4) Because of outside interventions like tanks becoming full or empty, that will influence these streams and which are considered to some extend in the model, the output from the model for that specific stream will be different from the generated distribution output.

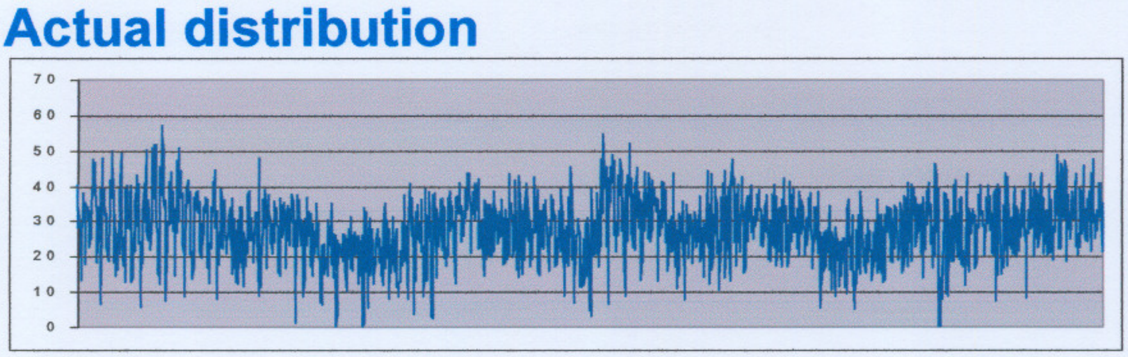
5) The source for the main feedstock for all the chemical plants and processes in the road to making petrol is coal, indicating a dependency very far back and out of bounds for this project.

Because of the risks of adding patterns to a single data stream that may adversely interact with another stream and totally skew the final results it was decided to make use of the distributions fitted and to ignore the patterns in this model. For the goal of the study the distributions used were therefore assumed to be sufficient, and the risks were communicated.

In literature a paper was published by Johnson(2002:457-467) where he discusses “Triangular approximations for continuous random variables in risk analysis” and the fact that where data is absent or difficult to find or interpret, it may be a safe option to use triangular distributions as it “is much more intuitive than ‘probability’ ....Thus the median, the quartiles, the 5% and 95% points and the 10% and 90% points can be described in terms that are likely to be meaningful to a manager” Johnson (2002:457)

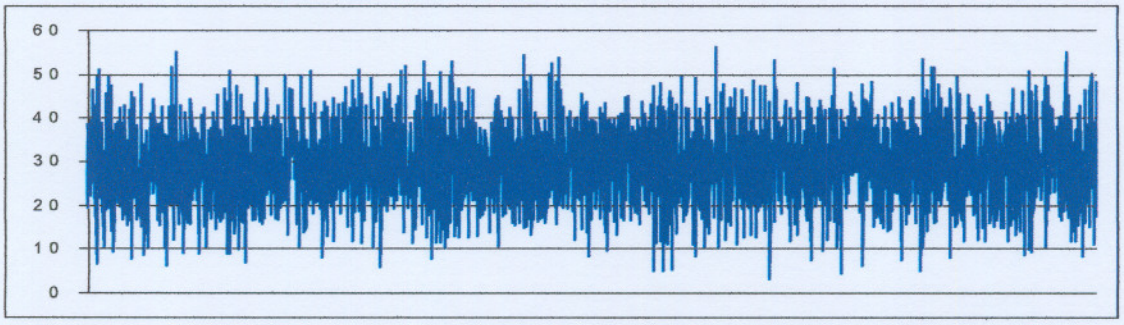
During the study some triangular distributions was used as well as the concept of bounding distributions at the points where the actual values of the distributions no longer make sense in reality, for example a Normal distribution that will give a flow of less than 0 may be bounded at 0. This technique is inherent in some of the later versions of Arena® as part of its input analyzer where the distributions are fitted.

Graph 3.2 Actual Distribution over time – hourly rates



Graph 3.3 Generated Distribution over time – hourly rates

Points	%	Distributions
194	0.048103	32 + 2 * BETA(1.24, 1.84)
3839	0.951897	NORM(29.6, 8.53)



c) Building a base case model

After about two months the base case model was developed, but the Visual Basic® interface with Excel gave problems. It was as though the software could not see each other running and error messages kept coming up. It needed to be sorted out and only then could the model continue. This was a serious problem and took about five to six months to solve. Neither the Arena® software vendor nor Microsoft could suggest a solution but the programming of the interface was sound. The problem was eventually solved by building the same logic in a roundabout way using Visual Basic programming so that the Excel and Arena® can run together without the one being aware of the other. A piece of code was added to the interface that "tricked" the two programs into waiting for a file instead of waiting for the other program to complete. The whole problem was solved in an hour, although the real software problem still remains and need to be sorted out by the software programmers. Problems like these had a significant impact on the project schedule.

d) Presenting the base case

From the presentation of the base case the feedback was very positive. For the first time a model was presented where the model tank graphs reacted in the same ways as the actual component tanks did. This was a milestone in the project because it was visual proof that the model worked and that had a great impact on the blending personnel. They started to see that stochastic modelling, being able to capture the time component, and especially the timing of blends, now also started to "see" that sometimes blends could not be made due to infrastructure constraints or component availability at specific times.

3.7.4 Base model problems

a) Too little detail – no 20/80

The model was still not good enough. The volumes made were not quite right, the mass balance was off and fewer blends were made than what was required. It was obvious from the tank graphs that insufficient volume was sold because the tanks kept overflowing. Many of the assumptions had to go. The model was at this point unable to use up all of the components and frequently made less than 95% of the intended scenario volume. This meant that using the 20/80 principle in this case was insufficient.

If there are too many assumptions it is easy to show that blends can be made – if the number of assumptions are too few the interaction existent in a system like this makes the model unusable because it frequently runs into constraints that the model itself cannot solve. Balance was necessary. Some interventions were made – one was to add a default base blend that was pre-tested for occasions where the solver was unable to find an optimum blend as discussed in 3.7.3b. Another intervention was to allow for the component tank to overflow, counting the overflow, to allow for some operator interventions that one does not want to build into a model – as discussed in 3.7.3a.

b) Run duration

A yearly model proved unrealistic. Too many interventions are needed in the running of the refinery to keep the petrol pool in balance and it would be an enormous task to build each of these into a model. The resulting model will then have many intervention options that can be used and it will not be run using the acceptable operating procedures. It is not sensible to build a model to run on exceptions. The system should usually be operated in a consistent way and interventions should only be used in exceptional cases. From these constraints the decision was made to run replications of three month time periods. It was possible to see over a three month period that the system was running empty or increasing volume over time. It was also possible to see if specific components give problems because thirty to fifty blends were made during each replication. Running more than one replication showed the effect of making a different sequence of blends in the same system over a three month period. If three replications were run and all three of them were unable to produce 95% or more of the intended volume it was fairly certain that the actual blend plant would also have problems to sustain the production of that specific slate of petrol. If however it was possible for two out of three to make the volumes and there were no obvious problems from the tank graphs, it could be deduced that it was possible to sustain blending for a scenario like that for three consecutive months. If the tank levels were very low at the end of this period, it could be deducted that it would not be possible to sustain this mode of operation for longer than three months.

At that point of the model development the duration of a three month run was seven hours. This was still very painful and the duration was due to the Excel interface that had to run about fifty times for each replication and nearly double that for a scenario that had difficulty to optimise blends. Several alternatives were evaluated to help solve the problem. One of them was to linearize the Solver so that calculation time could decrease and the other was to redevelop the model using PIMS (Process Industry Modelling System – a linear programming tool) and to use multi period modelling. The purpose of the LP (linear program) is to optimise the blend volumes required as well as the blend composition. A breakthrough was made during this period on one of the computing intensive properties used in the Solver and it could be linearised using a factor under certain conditions. This change decreased the runtime to around half an hour, a significant reduction. The PIMS (Process Industry Modelling System) model was not explored any further but still warrants a re-look, resources permitting.

c) Shutdown schedules

Shutdown schedules had to be removed for the factory units as well as for the final tanks. The component tank unavailabilities are the easiest ones to test, because the impact is basically on volume only as there are duplicate tanks in most cases. The only risk during this period is off spec product and then the slop (mixture of off specification materials) tanks can be used as an alternative. The duration of this maintenance can be anything between three to six weeks and although this is a long time the risk is very low.

Yearly tank maintenance on the final tanks is a different picture altogether. Because of the size of the tanks and the impact on the market the fuel slate may even change during statutory maintenance and short supply agreements may be agreed with government and other fuel companies. Modelling this type of contingencies is virtually impossible as it will change for each scenario.

In the case of the yearly shutdowns the picture becomes even more complex. When planning is done for a shutdown many of the rules change. The plants are run differently. The market supply gets negotiated. The transfers between Sasol and Natref changes and properties are renegotiated based on these scenarios. To build all of this into a running model would require a rebuild for every scenario with a very high risk of error. The decision was made to only include shutdown scenarios in the specific scenarios where they are expected to happen and to start there with a snapshot of the shutdown start time and then to try and go trough the shutdown with specific rules. Afterwards the system then needs to go back to "normal" in an acceptable period of time.

d) Model output constraints – so much data to be interpreted

Another problem was the sheer volume of output data. It was difficult to understand the results and make sense of it. It was difficult to decide when the run was feasible and when not. During this time many macros were developed to summarise output data into results that can be interpreted. Graphs are drawn for each of the component tanks and final tanks in the system. Incoming volume is analysed and compared to what was intended, blends made are analysed and property and volume problems are identified, risk areas are identified and evaluated carefully, market requirements and supply are summarised, tank overflows and shortages are calculated and flagged as well as any intervention. Information needs were identified several times by trail and error – if something was not reported on and a question was asked around that, that the modelling team could not answer, it was back to the drawing board.

e) How to continue

It is a painful realisation in any modelling project to come to, that the gain does not lie in the more obvious things but that small assumptions make the model not present reality and that the assumption needs to be discarded and greater detail needs to be built into the model. As assumptions were removed parts of the model was rebuild and new results were generated. Over a period of 4 – 6 months many of the assumptions were discarded and many parts of the model were rebuilt to better describe the reality. Output analysis and information were refined to be able to answer all the required questions and the learning curve was incredible.

3.7.5 First breakthrough

The first breakthrough came after about 18 months of the project when the first sensible understandable and logical results were put on the table and the model adequately represented the actual system for the first time. At this point most of the questions asked from the model could be answered. One problem remained. The model was a representative model of the plant a year before and the plant changed as well as many of the business rules. For this base case model some scenarios were already done and the scenario results differed from the expected results based on the Solver output. To really prove the model the gap needed to be filled.

For the next 4-6 months the base case was updated to the then current ± 3 months and the scenarios were re-run. Late in 2002 the first real results of the actual, then current system was available and the base case model adequately described the actual system and it was proven. That was a remarkable day in this project. At that point the model was nearly nine months overdue but the base case model and the plant operated the same. The trust in the model had grown by leaps and bound and many of the scenarios were very near to the results of the Solver.

3.7.6 Solving the mystery

The biggest question at that point in time was why were there such a big difference between expected Solver results and the Arena® results for some of the scenarios. Was it infrastructure? What percentage of the difference could be related to the addition of variation and distributions in streams and properties and what part to infrastructure and if one cannot attribute it to anything specific, what were the reasons for the difference.

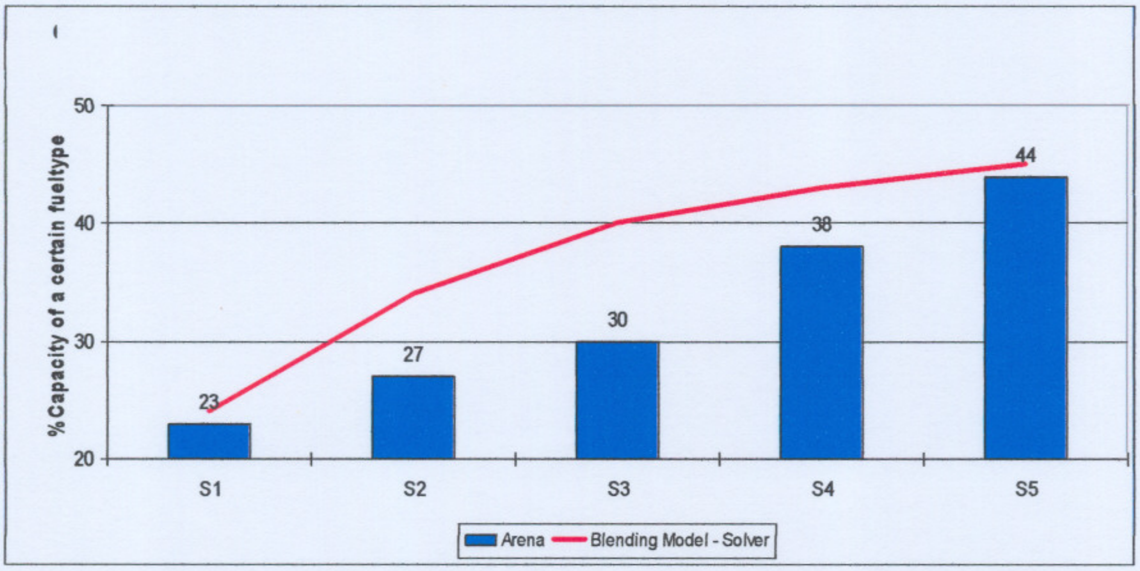
Options

- The first thing that could be done was to remove all the variation on stream ***volumes*** and replace it by the same averages used in the Solver
- The second was to remove all variation on stream ***properties*** and replace it by the same averages used in the Solver.
- The third was to remove all the failures.
- The fourth is to remove all tank constraints

Making the model deterministic should take the model back to represent the Solver more closely because it takes out most of the variation and it brings the model closer to a model based on averages. In the first scenario that this was tried, doing the first three interventions did bring the models closer in terms of the unleaded capacity of each. From these results it was possible to show that the Arena® model and Solver model were approximately the same. The smallest gap was due to infrastructure but taking away all tank constraints made the model infeasible. There are too many operating rules in the model and totally opening the tanks made a lot of decisions unnecessary. For example in the blending logic no tank ever overflowed in the new tank scenario and the tank priorities were rendered useless. It was adequate to only make the tanks bigger.

The surprising result from this part of the study was that it was not so much infrastructure that created the gap between reality and the solver. It was much more the effect of having the right balance of feed volumes available at the right properties exactly when they were required. It was about what was lost when a less optimal blend was made due to tanks overflowing or running empty. It was about timing the blends so that a high octane leaded blend did not follow or precede a high volume unleaded blend because both required high volumes of lower availability high octane streams and the more readily available lower octane streams could not be used in these blends in the sufficient volumes to not overflow their tanks. There were intermediate scenarios where high and low octane stream availability did not match well with the market. The properties of the streams at these points were such that small deviations from the optimum blend composition made it virtually impossible for consecutive blends to be made. These results were very difficult to communicate because the actual meaning of the results was camouflaged by the effect of reducing variation. It was only fully possible to communicate the impact when several specific scenarios modelled in the system were combined and the Solver and stochastic model results converged again for those combined scenarios. See S5 on the x-axis of Graph 3.4.

Graph 3.4 Sasol Synfuels Compounded Scenarios



3.7.7 Results

The stochastic model at this point was able to forecast which scenarios would make sense to use and which would be difficult to sustain. It also showed what infrastructure was needed for each scenario. The advantage of this knowledge could be put to good use by deciding how to plan and offset the unleaded capacity supplied by each factory. This model brought a lot of understanding from the shop floor through to management level. Where capital was required it could now be proven that it is really necessary.

3.7.8 Opportunities for further study

- Can Markov chains be a better assumption than fitted distributions for stream volumes?
- When there is correlation between streams, but the correlation cannot easily be translated into a mathematical equation for calculating the one stream from the other one, how can the interaction be modelled sensibly?
- If statistical significance and repeatability needs to be established for a highly interactive process with as much decision rules, options and results data as fuel blending, how can it be done in a sensible timeframe?

3.8 Other models at Sasol

The example in 3.7 is a true dynamic, continuous Petrochemical example but there are also various other models developed at Sasol.

Some of these are

Blending plant model
Supply chain model for a blending plant where there is a continuous supply of various chemical products into final tanks, blending of these chemicals and discrete off-take from these final blend tanks via road tankers. The goal of this model was to check various infrastructure scenarios and provide the business with workable alternatives to blend and distribute about 200 products.

Crude arrival, pipeline transfer and blending model
This model started from shipments and ended with blends in final products. The process was not modelled, but the supply chain infrastructure and the blending was included in the model scope. The goal was to increase production and evaluate infrastructure needs for the extended system.

Sasolburg road tanker loading models for various plants
In most cases these models were used to de-bottleneck the plants

Rail loading models for various plants
In most cases these models were used to assess increased throughput and to make sure that the infrastructure needs for new products were addressed.

Plant models
Various plant models including for example manufacture of explosives, waxes or fertilizers and other chemical products. These models included various processes, conveyors, packing mediums, pellitizers, forklifts, stores, silos, tanks, and the people required at each step. The objectives of these models were to build a base case model of the existing model, to implement known changes and to de-bottleneck the new scenario at the lowest cost. Various scenarios were proposed by the business and they were tested and compared. Some scenarios for optimisation were suggested by the modeller and the study results based on the learning that took place through the modelling process.

Coal Bunkers

Coal distribution and bunker models attempting to smooth out coal properties in huge bunkers

Plant models

Stochastic model of various Sasol processes including reactors, crackers, hydrotreaters, columns and other chemical units. The objectives of these models were mainly risk related, to make sure that the throughput promised can also be delivered.

3.9 Conclusion

For every change in one's business, a chain reaction is caused and other parts of the business is influenced either in purity of product, properties or volumes. These changes can adversely influence markets and the goal of every business is to minimise these influences or to warn or change the markets well in advance. Where interactions between parts of the business play a role, where these interactions have the highest risk and how they can be influenced needs to be known. Options in each business to handle changes need to be developed for maximum flexibility and sustainability. Stochastic modelling can play a leading role in achieving this.

3.10 Additional insights on people dynamics

A surprising conclusion from the study is the impact that the client and other project members can have on the study and the need for also managing the people component becomes clear.

3.10.1 Collect all the views of the interested parties

Claim 1:

When defining and understanding the environment it is necessary to collect all the views of the interested parties as they usually differ

Use the client expertise to define the relevant work groups and also the people who can be held accountable in each area. The first part of Conceptual design in Chapter 2.3.4

explains that it is critical to understand the problem perspective through the different levels of the company. It is necessary to understand that a marketer and operations manager has a very different world view and therefore a different perspective on the problem. Understanding the different viewpoints will broaden the perspective of the modeller and how the problem can be approached. Understanding each lends the modeller a wide and unbiased perception of the whole operation and will assist in building a model that adequately represents the reality.

3.10.2 Department incentives

Claim 2:

Department incentives may have conflicting goals

Understanding the incentives of a business for each department will illuminate out the focus of each department. It often happens in a project that these drivers are not geared to a common goal. Incentives are often internally focussed and biased towards internal optimisation. Small changes in department goals may result in great gains.

3.10.3 Conflicting goals and fear of sharing

Claim 3:

Conflicting goals and fear of sharing may make the project difficult

Possible problems during the data gathering phase are

- Information may be withheld.
- Incorrect information may be provided due to personal risk
- Information may be biased

Much of the risk in the data gathering phase is due to people who are not bought into the process. These people need to be managed carefully as they provide a serious threat to the process. The risk is reduced when more people are involved in this phase of the project. Most modellers are more comfortable with models than with people, but when one is required to manage the whole project it is also important to develop the skills of negotiating with people and of marketing the technology, the models and the results. The human element remains one of the most important elements in any study.

3.10.4 More than one modeller reduce risk

Claim 4:

More than one modeller is necessary to complete a successful project with low risk

Consider a buddy for the project. When there is only one modeller on a stochastic project, the risk of the project skyrockets. Anything can happen to a person and anybody can resign at any time. It is therefore important that there is always more than one person on the project that is comfortable with the model.

3.10.5 Check results analyses

Claim 5:

Results analyses must be checked with the client's perspective

The results need to be interpreted with consideration of all the information – not just blindly on the outcome of the simulation. Some things are just not possible due to interactions not always included in the scope of the project. Look at the bigger picture to enhance the interpretation – nobody knows the model as well as the modeller. Nobody also understands the reality as well as the client does, it is therefore critical to listen to what the client has to say about the results. He might see something that is missed by the model, but that became clear to him during the modelling process.

3.10.6 Presentation skills training and information transfer

Claim 6:

Presentation skills training is always helpful – transferring the information to the

client and project team successfully is essential

One important point to mention here though is that without good transparencies, a thorough discussion of the assumptions used, the results obtained from the model as well as the reasons for each recommendation and a sound knowledge of the model and concepts, the presentation may not be successful. All modelers should also be good presenters and a course to teach a modeler the finer skills of presentation is always helpful.

Even if the work done for the simulation was of excellent quality it needs to be transferred to the client in such a way that it is believable, and trusted as accurate and thorough. If an assumption is doubted at this stage – none of the recommendations will be taken seriously because the groundwork for the simulation will be questioned. The modeler could just as well not have done the project.

It may seems harsh to say "the modeler could just as well not have done the project" but in most organizations many people are included in these types of projects and the credibility of the results, is also the credibility of the modeler. If only one of the key people at the final presentation disbelieves the results, it means the model is seen as inadequate and **that** will be communicated. When the model is used to decide where to spend capital and how much the client really needs to spend and it amounts to millions one really needs to have a good case. If not, the modeler looses his/her value to the client and that is not easily regained.

3.11 Additional insight from the study – general

3.11.1 Over-design

Stochastic modelling reduces the need for additional capacity. Traditionally plants or tanks were often oversized to minimise risk or "for in case" because even though the cost implication of too much capacity is severe, the loss due to a plant or process being unable to meet its demands is much more serious. A stochastic model can model the new or changed plant or process as part of the existing system and show the effects of interactions not necessarily planned for. It is therefore not necessary in this case to build in extra capacity.

3.11.2 Volume balance versus time

Blend size is constrained by the balance of available volume at a given point in time in a blending plant. The surprising result from this part of the study was that it was not so much infrastructure that created the gap between reality and the Excel solver. It was much more the effect of having the right balance of feed volumes available at the right properties exactly when they were required. It was about what was lost when a less optimal blend was made due to tanks overflowing or running empty. It was about timing the blends so that a high octane leaded blend did not follow or precede a high volume unleaded blend because both required high volumes of lower availability high octane streams and the more readily available lower octane streams could not be used in these blends in the right volumes to not overflow their tanks.

3.11.3 Operability

Stochastic models can test various scenarios and show which ones are more operable given the infrastructure constraints

The stochastic model at this point was able to forecast which scenarios would make sense to use and which would be difficult to sustain. It also showed what infrastructure was needed for each scenario. The advantage of this knowledge could be put to good use by deciding how to plan and offset the unleaded capacity supplied by each factory. This model brought a lot of understanding from the shop floor through to management level. Where capital was required it could now be proven that it is really necessary.

Simulation proofs to be a very useful and powerful tool for studying continuous, dynamic, real world systems of an interactive nature. The simulation will pick up or mimic the bottlenecks and constraints of the real world system. Through scenarios one can explore alternative solutions to these problems without the cost involved in actually trying out the different scenarios. In a real world system the risk remains that when a solution is proposed and implemented, that the estimated benefit is far less than the intended one because of secondary constraints. A simulation study may show that a different and sometimes initially more expensive solution may be the one that reaps the most benefits. On the other hand such a study may also show that the proposed solution is overkill and that a much less expensive alternative may be sufficient.

3.11.4 Testing Reliability, Availability and Maintainability (RAM)

Stochastic models can test Reliability, Availability and Maintainability (RAM) during the design phase of a plant or reactor.

In any industry over time various mathematical methods have been developed to assist engineers during factory design to estimate plant and storage sizes to be able to give the required throughput of the total system. Often each specific part or item to be added to form the whole system is designed in isolation with the goal of estimating what that plant will be capable off. Historically average volumes are used to design tank size requirements and a volume percentage is added to reduce risk.

These historical design methods are effective, but often through risk minimisation tanks are significantly oversized. Plants parts on the other hand, by being treated in isolation are designed with little "catch – up" capacity. Catch-up capacity here refers to the capability of a plant to catch up on lost production due to a breakdown in the surrounding system. A RAM study is a simulation study method designed to bridge this gap by modelling for example the reactor to be sized with its peripherals that may negatively influence its total throughput over time.

Maintenance and breakdown profiles are added to the model and average volumes are used as input to these models to estimate the required throughput. The advantage of modelling reliability, availability and maintainability is that the design is more robust. The capacity of the plant is less of a guess and risk is mitigated to a large extend.

3.12 Further work

3.12.1 Stochastic models combined with process models

It is only reasonable to assume that stochastic models are not the only types of models needed in the Petrochemical industry. Rigorous process models are for the process industry what food is for people, but interactions between technologies can be exploited to get more comprehensive and reliable answers. One way that this can be done is to use experimental design to create a representative experiment and to then run the experiment with the process models. The solution matrix can then be used in a stochastic model to get improved stream properties and splits over a unit or reactor based on the chemical reactions that takes place. If enough points are added to the solution matrix one can possible assume a linear relationship between two consecutive points.

3.12.2 Stochastic models combined with linear programming or optimisation models

Using stochastic modelling with an optimiser or linear program can mean that the optimal linear program solution can now be tested for feasibility on the actual system and the infrastructure or interaction constraints can be driven out and solved. If an optimum is not "operable" another optimum or feasible solution with the same economic advantages may be tried possibly resulting in a better fit for the infrastructure constraints and therefore also easier to operate.

Exploring the synergies between technologies is therefore the quest for the next few years.

3.12.3 Analysing Risk

Analyzing the risk of a project to fail or not to meet the time schedule is quite a challenge. If this risk can be estimated well the risk can be mitigated and the appropriate measures taken when and if needed.

3.12.4 Correlation in input streams

Correlation in model arrival logic or in streams entering units or exiting from units is a threat for stochastic models. Not taking the relationships into account may introduce more variation in to the system than what actually exists and one may loose some information where two tanks may be full or empty at the same time or one may be overflowing and one empty at the same time point. The results may then be inaccurate. Some of these relationship problems may be solved by making use of process models that may estimate the appropriate splits and properties, but further work is needed here.

3.12.5 Patterns over time vs. drawing from a distribution

In actual plant data there are sequences in the data. For a sequence of time the volumes may go higher, for example after start-up as the plant progresses to its appropriate conditions, the volume will increase. For a planned shutdown flow may decrease over a few hours to reach zero. There is a lot of variation in plant data, but usually it will not go from the bottom to the top of the range in 5 or 20 minutes. In data drawn from a distribution immediate changes are possible. In this area a lot of work can be done on the potential of time series to improve the arrival stream volumes.

3.12.6 Statistical significance and repeatability in interactive processes

If statistical significance and repeatability needs to be established for a highly interactive process with as much decision rules, options and results data as fuel blending, how can it be done in a sensible timeframe? When a model runs for 15 hours to complete one replication is it possible to gain sufficient information? Can experimental design reduce number of runs?

CHAPTER 4 SUMMARY AND CONCLUSION

4.1 Summary

"Simulation refers to a broad collection of methods and applications to mimic the behaviour of real systems, usually on computer with appropriate software" (Kelton *et al.*, 2002:3)

In the process of model building we are translating our real world problem into an equivalent modelling problem which we solve and then attempt to interpret. (Hangos & Cameron, 2001:480)

Discrete simulation was originally used for queuing systems but it expanded so much that it can be used to solve most problems. Some of the traditional applications are

- As an alternative to playing with the system "Sometimes you can't and shouldn't play with the system" (Kelton *et al.*,2002:5)
- To evaluate new policies, operating procedures and decision or organisational rules without interrupting the operations.
- For queuing systems.
- To de-bottleneck systems.
- To understand the interactive variability in a system.
- To evaluate physical changes in a system.
- To evaluate a new system before it is built.
- The whole logistics chain of a system can be modelled and evaluated.
- Paper flow can be evaluated and changes recommended (business environment).
- Layout of storage facilities or systems can be evaluated.
- For job shop scheduling.

Some of the reasons why stochastic modelling are used in the petrochemical industry are

- At times deterministic mathematical models (based on averages) of the existing system are not trusted because of the gap between the model and reality.
- The real world system is integrated and complex.

- The impact of various planned expansions on an existing system needs to be estimated.
- To help drive out operational philosophies.
- To size infrastructure.
- For the design of new plants in existing systems and to estimate infrastructure impact and plant throughput.
- When there are certain known events in time that need to be accounted for like shutdowns.
- When discrete continuous interfaces exist. For example continuous volumes runs down into a tank and from there tankers come to load at discrete time intervals – or volume being added into a tank in batch mode but extracted via pipeline on a continuous basis.
- For capital approval.
- When a model is needed to tap the last bit of capacity or go above the 80:20 principle. The 80:20 refers to the 20% time investment for 80% gain. It is often necessary to invest the 80% time to gain the last 20%. With simulation the 80% time may be less than in other modelling applications and one can get a model that may be re-used in future, adding even more value than intended.
- To reduce risk.

A few examples of stochastic modelling applications are given in chapter 3. All of the simulation models, referred to in this study, are in use in the Petrochemical Industry and have already proven their value.

Stochastic modelling can add real value in the process industry because

- it is capable of modelling interactions between businesses
- it can handle variation
- problems occurring at specific time windows can be captured
- it provides an accurate picture or representation of the system
- logistic constrains can be included
- continuous discrete interfaces can be modelled,
- it can reduce risk in implementation
- it can show potential bottlenecks
- it is reliable in sizing infrastructure
- It can indicate where capital should be spent

- It can avert capital expenditure

For every change in one's business, a chain reaction is caused and other parts of the business are influenced either in purity of product, properties or volumes. These changes can adversely influence markets and the goal of every business is to minimise these influences or to warn or change the markets well in advance. Where interactions between subsystems of the business pose a risk, it is important to know and understand these risks and how they can potentially be reduced. Options in each business to handle interactions need to be developed for maximum flexibility and sustainability. Stochastic modelling can play a leading role in achieving this.

4.2 Goals – Review and conclusion:

Basic Hypothesis:

The basic hypothesis of this study is that stochastic modelling can be extended successfully to the Petrochemical industry. The value-add lies in the addition of passage of time and the exploration of the interaction of linked processes with their own internal variation.

Goal 1

This study is about exploring the entry into the Petrochemical field and it describes how Sasol has used this tool in the last ten years.

Goal 2

This study describes where stochastic modelling fits into the modelling framework and what value can be added with it.

Goal 3

Te study outlines the management of a modelling project compared to the normal project environment and gives model examples and applications.

Goal 4

The study attempt to proof that stochastic modelling can be extended successfully to the Petrochemical industry.

The models explained in this study illustrate clearly that Stochastic modelling can be extended to the Petrochemical Industry with significant value add. The projects done in this study was managed successfully according to the guidelines explained in Chapter 2. The continuous nature of a chemical process can be approximated adequately with a discrete one if the time gaps are small enough.

The challenge as explained in Hangos & Cameron (2001:16) is met.
The current challenge for process models lies in industries with "large scale discrete-continuous operations" (Hangos & Cameron, 2001:16)

The following recommendations are made in this study regarding the simulation process:

- It is important to get to know the system in question really well.
- The choice of software and being able to rationalise why specific software is used needs to be done before any modelling starts.
- Client training paves the way of an understanding relationship between client and modeller.
- A flow diagram and detail documentation on all decision rules are essential for success.
- People dynamics can play a major role in project success.
- A lot of forethought and preparation has to go into the model. It reduces modelling time and risk of rework.
- When preparation is done thoroughly, the actual model building is fast and it will run smoothly.
- A Base Case model must always be built.
- Scenarios can only be tested if the Base model is verified and validated and accepted by the client.
- Statistical analysis of results can enhance the results and clarify what the recommendations should be.
- Recommendations need to be made by the modeller, after client discussion.
- Continuity can only be assured through full documentation.
- Implementation is an indication of success.
- The credibility of the modeller is linked to the credibility of the model.

The capability of the modeller to sell the tool and the results will influence the outcome of the study.

Bibliography/References

ANON Whatis.com 2001

DIXIT, A.K., PINDYCK, R.S. 2004. Investment under uncertainty, Amazon.com

BANKS, J., CARSON, J.S.II. 1999. Discrete-Event System Simulation. Upper Saddle River, N.J. : Prentice Hall. 548 p.

BOX, G.E.P., HUNTER, W.G., HUNTER J.S. 1978. Statistics for Experimenters An introduction to Design, Data Analysis, and Model Building. New York : Wiley. 653 p.

FREUND, J.E. 1984. Modern Elementary Statistics. 6th ed. London : Prentice Hall. 561 p.

GOLDRATT, DR. E.M., COX, J., 1992, The Goal 2nd rev., South Africa: AGI. 333p

HANGOS, K.M., CAMERON, I.T. 2001., Process Modelling and Model Analysis. Process Systems Engineering. Vol.4. London : Academic Press. 543 p.

HINES, W.W., MONTGOMERY, D.C. 1990. Probability and Statistics in Engineering and Management Science. 3rd ed. New York : Wiley. 732 p.

KELTON, W.D., SADOWSKI R.P., SADOWSKI, D.A. 2002. Simulation with Arena® 2nd ed. New York : McGraw-Hill. 629 p.

PEGDEN, C.D., SHANNON, R.E., SADOWSKI, R.P. 1995. Introduction to Simulation using SIMAN. 2nd ed. New York : McGraw-Hill. 600 p.

RENDER, B., STAIR,R.M.JR., HANNA,M.E. 2003, Quantitative Analysis for Management – International Edition. 8th ed. New Jersey : Prentice Hall. 726 p.

ROBINSON, S. 1994. Successful Simulation A Practical approach to Simulation Projects. London : McGraw-Hill. 246 p.

Bibliography/References

ANON Whatis.com 2001

DIXIT, A.K., PINDYCK, R.S. 2004. Investment under uncertainty, Amazon.com

BANKS, J., CARSON, J.S.II. 1999. Discrete-Event System Simulation. Upper Saddle River, N.J. : Prentice Hall. 548 p.

BOX, G.E.P., HUNTER, W.G., HUNTER J.S. 1978. Statistics for Experimenters An introduction to Design, Data Analysis, and Model Building. New York : Wiley. 653 p.

FREUND, J.E. 1984. Modern Elementary Statistics. 6th ed. London : Prentice Hall. 561 p.

GOLDRATT, DR. E.M., COX, J., 1992, The Goal 2nd rev., South Africa: AGI. 333p

HANGOS, K.M., CAMERON, I.T. 2001., Process Modelling and Model Analysis. Process Systems Engineering. Vol.4. London : Academic Press. 543 p.

HINES, W.W., MONTGOMERY, D.C. 1990. Probability and Statistics in Engineering and Management Science. 3rd ed. New York : Wiley. 732 p.

KELTON, W.D., SADOWSKI R.P., SADOWSKI, D.A. 2002. Simulation with Arena® 2nd ed. New York : McGraw-Hill. 629 p.

PEGDEN, C.D., SHANNON, R.E., SADOWSKI, R.P. 1995. Introduction to Simulation using SIMAN. 2nd ed. New York : McGraw-Hill. 600 p.

RENDER, B., STAIR,R.M.JR., HANNA,M.E. 2003, Quantitative Analysis for Management – International Edition. 8th ed. New Jersey : Prentice Hall. 726 p.

ROBINSON, S. 1994. Successful Simulation A Practical approach to Simulation Projects. London : McGraw-Hill. 246 p.

KARLIN, S., TAYLOR, H.M. 1975. A First Course in Stochastic Processes, 2nd ed, Academic Press

TAHA, H.A. 1976., Operations Research an introduction. 2nd ed. New York : Macmillan. 648 p.

TAYLOR, H.M., KARLIN, S, 1998. An introduction to Stochastic modelling, 3rd ed, San Diego : Academic Press. 631p.

Articles

JOHNSON, D. 2002. Triangular approximations for random continuous variables in risk analysis. Journal of Operational Research Society, 53. Surrey : Unwin Brothers Ltd, The Gresham Press. p457-467