

**PREDICTION OF CEMENT CLINKER QUALITY FROM
THE RAW MATERIAL, LIMESTONE, QUALITY**

**PREDICTION OF CEMENT CLINKER QUALITY FROM THE
RAW MATERIAL, LIMESTONE, QUALITY**

By

CAH Taylor

THESIS

Submitted as partial fulfilment of the requirements for the degree :

MASTERS IN BUSINESS ADMINISTRATION

at the Graduate School of Management at the
Potchefstroom Universiteit vir Christilike Hoër Opvoeding

Study leader : Pieter Geldenhuys

Potchefstroom

1998

ACKNOWLEDGEMENTS

Without the support, assistance and patience offered to myself during the compiling of this report, the successful completion would never have been possible. A word of thanks must go out to the following people :

To Charmaine, for her patience and understanding

To my parents, for all the opportunities they provided for me

To my colleagues at Alpha Cement for their support and ideas

To my study group for their support

To my study leader, Pieter Geldenhuys

To Mike McDonald for assisting me with proof reading

To Stephan Olivier for his teaching and original ideas

OPSOMMING

TITEL :

VOORSPELLING VAN PORTLAND SEMENT KLINKER KWALITEIT GEBASEER OP KALKSTEEN (DIE HOOF ROU MATERIAAL) KWALITEIT

1. INLEIDING

Die sement proses, soos vandag bekend, is 'n hoë koste, kapitale intensiewe vervaardigings proses. Die primêre rou materiaal gebruik in die proses is kalksteen. Hierdie materiaal word met sekondêre materiaal fyngemaal tot 'n poeier. Daarna word dit in 'n oond gebrand teen 'n temperatuur van $\pm 1450^{\circ}\text{C}$. Dit is 'n hoë koste energie intensiewe proses en alle maatreëls moet gevolg word om die kostes te beheer.

Sement kwaliteit asook die kwaliteit van die rou materiaal wat gebruik word en die in proses produkte word in dié bedryf deur n paar verhoudinge gekwantifiseer. Die waardes van hierdie verhoudings asook die konstantheid daarvan is baie belangrik in die verbrandings proses. Om proses beheer te bemeester wil die produksie bestuurder graag ten minste twee van hierdie verhoudings goed beheer as ook een van die

oksiede self. Met hoë fluktuasies in kalksteen kwaliteit asook die kwaliteit van die bymiddels is dié behoefte baie moeilik om aan te beantwoord.

Slegs een van die verhoudings van Portland Sement klinker kan met akkuraatheid beheer word. Die ander parameters kan baie fluktueer met die gevolg dat die proses onstabiel raak met hoë koste konsekwensies.

In hierdie bespreking word 'n model om die tweede verhouding te voorspel (en dus te beheer, reeds in die kalksteen stapel bou proses) ontwerp

2. PROBLEEMSTELLING

Wisselvalligheid in klinkerkwaliteit lei tot onaanvaarbare fluktuasies in sement kwaliteit asook in sement prestasie in beton. Hierdie fluktuasies het ook 'n effek op die proses stabiliteit, met 'n substantiewe effek op kostes. Die probleem is dus tweeledig:

- I. Onaanvaarbare sement kwaliteit fluktuasies lei tot kliënte ontevredenheid, en
- II. Fluktuasie in die klinker proses veroorsaak dat kostes van produksie onaanvaarbaar steig.

3. DOELWIT

Die doelwit van hierdie studie is om 'n model om klinker kwaliteit te voorspel vanaf kalksteen kwaliteit, te ondersoek. 'n Bestaande model word gebruik as die begin punt, om 'n statistiese aanvaarbare verhouding tussen die twee materiale te kwantifiseer. Klem word gelê op die buigsaamheid van die verhouding as gevolg van die dinamiese natuur van die proses en die rou materiale wat gebruik word daarin.

4. METODE

Die volgende metodes is gebruik ten einde die probleem op te los:

- I. Die proses is in detail bestudeer, in 'n sisteems analiese, om 'n praktiese idee te kry van die kompleksiteit van die opdrag, en om te verseker dat 'n oplossing wat toepaslik asook prakties uitvoerbaar is, is gekry.
- II. Literatuur op sement chemie asook die wye sement proses is bestudeer om verdere agtergrond te kry.
- III. Literatuur oor statistiese metodes met klem op die bou van 'n model is deeglik nagegaan.
- IV. Die literatuur, saam met die kennis verkryg uit die sisteem analiese, is direk gebruik in die ontwikkeling van 'n model wat ook getoets en geverifieer is deur die literatuur direk toe te pas, met die hulp van 'n gevalle studie.

5. RESULTATE

'n Model is ontwerp om die klinker kwaliteit te voorspel van diè van die kalksteen gebruik in die proses. Die model is statisties getoets en bewys met resultate wat die beperkinge in die praktyk goed nakom. Die model is in Microsoft Excel ontwerp en bestaan uit twee dele:

- A. 'n Optimiserings gedeelte wat vir 'n paar doelwitte gebruik kan word, naamlik :
 1. Om genoeg data te genereer, deur middel van Liniêre Programmering, om 'n verhouding te kan bepaal tussen die klinker kwaliteit en diè van die kalksteen, of
 2. Om direk die klinker kwaliteit te voorspel deur Liniêre Programmering, of
 3. Om die sensitiwiteit van veranderinge in die kwaliteite van bymiddels en brandstowwe gebruik in die proses, te bepaal.
- B. 'n Verhouding bepaling gedeelte waar, deur middel van regressiemetodes, koëffisiënte van 'n formule wat gebruik word om die klinker kwaliteit te voorspel, bereken word.

6. SAMEVATING

'n Model om die verhouding tussen sement klinker kwaliteit en kalksteen kwaliteit is ontwikkel. Dit is 'n uitbreiding op 'n vroeë model ontwerp deur Dr Stephan Olivier, en word gebaseer op Liniêre Programmeringsteorië as ook Liniêre Regressiteorië. Die model kan gebruik word om die koëffisiënte van 'n stel eenvoudige formules te bereken, wat dan gebruik kan word om die klinker Silikaverhouding te bereken, of dit kan direk gebruik word om die Silikaverhouding te bepaal.

Die model wat gebaseer en getoets is op Alpha Sement Dudfield se proses, kan maklik vir enige droë klinker vervaardigings proses gebruik word, waar daar nie meer as drie bymiddels saam met die kalksteen gemeng word nie. Die model word in Excel sagteware verwerk.

TABLE OF CONTENTS

	Page
CHAPTER 1 : NATURE AND SCOPE OF STUDY	1
1.1. INTRODUCTION	1
1.2. PROBLEM STATEMENT	2
1.3. OBJECTIVE	4
1.4. SCOPE OF STUDY	5
1.5. RESEARCH METHODOLOGY	6
1.6. CONCLUSION	7
CHAPTER 2 : THEORETICAL LEGTIMACY	8
2.1. INTRODUCTION	8
2.2. CONSTRUCTING A MATHEMATICAL MODEL	9
2.3. LINEAR PROGRAMMING	14
2.4. CHOICE OF ANALYSIS	16
2.4.1. Multiple Linear Regression	17
2.5. MODEL TESTING AND VERIFICATION	19
2.5.1. Association versus Causality - Statistical Inference	19
2.5.2. Testing and verification of Multiple linear regression	20
2.5.3. Hypothesis testing : Two population means (large samples)	24
2.6. CONCLUSION	25

CHAPTER 3 : MODEL DEVELOPMENT USING A CASE STUDY	26
3.1. INTRODUCTION	26
3.2. FRAME OF REFERENCE	27
3.2.1. The Cement Production Process	27
3.2.2. Cement Chemistry	28
3.2.3. Overview of the Clinker Manufacturing Process	31
3.3. DESCRIPTION OF THE MODEL	33
3.3.1. The model flowsheet	33
3.3.2. System Analysis	34
3.3.3. Assumptions	36
3.3.4. Model Part 1 formulation - Optimisation Problem.	37
3.3.5. Model Part 2 - Prediction Formulation.	41
3.3.6. Overview of model Operation	42
3.4. MODEL TESTING AND VERIFICATION	43
3.4.1. Testing and verification of the Linear programming model	43
3.4.2. Testing and verification of formulae for derived clinker SR	46
3.5. Conclusion	48
CHAPTER 4 : CONCLUSION AND RECOMMENDATIONS	49
4.1. CONCLUSIONS	49
4.2. RECOMMENDATIONS	50
4.2.1. Use the Linear Program directly if practical	50
4.2.2. Determine the frequency with which the formulae should be updated.	51
4.2.3. Use the model for prospecting purposes	52

REFERENCES	53
-------------------	-----------

LIST OF DRAWINGS

Figure 2.1 - Producing a Mathematical model	11
Figure 2.2. - The definition of a mathematical model, in mathematical notation.	13
Figure 3.1 - Chemical ratios used in cement quality control.	30
Figure 3.2. - Model Flowsheet	34

LIST OF TABLES

Table 2.1. - Analysis techniques.	16
Table 3.1. - Verification of Optimisation Program in Practice - Stockpile A.	45
Table 3.2. - Verification of Optimisation Program in Practice - Stockpile B.	46
Table 3.3. - Verification of Formulae for derived clinker SR - Results of Linear regression.	46
Table 3.4. - Testing the model on independent data	47

TABLE OF APPENDICES

APPENDIX 1 - SYSTEMS ANALYSIS - SCATTER DIAGRAMS	56
APPENDIX 2 - TEST STATISTIC DECISION TREE	58
APPENDIX 3 - INPUT AND OUTPUT SCREENS	59
APPENDIX 4 - GLOSSARY OF TERMS	61

CHAPTER 1

NATURE AND SCOPE OF STUDY

1.1. Introduction

The purpose of this study is to provide the Production manager of a cement plant with a tool, in the form of a "model", to predict cement clinker quality from the quality of limestone, the primary raw material. The idea is to provide the manager with a set of formulae for this purpose, and enable him/her to test the validity of the formula at intervals of his/her choice. This should warn the manager in the event that a theoretically possible solution is not practically possible due to any plant or process limitations. The ultimate goal is to provide the Production Manager with a dynamic model, rather than a static formula only.

In this chapter a glossary of terms which will be used in the paper, will be provided. The problem to be addressed will be defined and explained and the purpose of this study, in addressing the problem, put in context. The reader will be provided with a frame of reference for the subject in order to ensure easy reading and understanding. The cement process will be briefly explained along with some basic cement chemistry relevant to the topic.

The scope of the study will be clearly stated and the research approach used explained. The chapter is rounded off with a short conclusion.

1.2. Problem statement

The formula used at Alpha Cement, Dudfield, to predict clinker Silica ratio from limestone quality does not provide consistent results.

This formula which was developed at the Alpha Cement, Dudfield plant by Dr Stephan Olivier, to predict clinker SR from limestone quality, worked very well for a few years, after which it started becoming more and more inaccurate. This original formula was a sizeable breakthrough in the industry and especially for this plant with a particularly variable raw material deposit, and allowed the production Manager to reduce variability in clinker quality substantially.

Variability in clinker quality has a negative impact on the clinkering process itself, the cement milling process and cement quality. In all three cases, additional costs are incurred by the company. The Clinkering process is the most expensive part of the cement manufacturing process due to high refractory and fuel costs, as is discussed under heading 3.2. Variation in clinker quality has a direct impact on the clinkering process itself, in that more fuel is used and refractory material lives are reduced resulting in additional costs.

The clinker silica ratio, SR, depends not only on the chemical compositions of the various raw materials, but on their relative amounts in the mix. Similarly, the quality and amount of fuel used also affects the clinker SR. A single formula can thus not be used for different kilns, unless they (and their processes) are identical.

If any one of the raw materials changes in composition it has an effect on the amount of each of the other raw materials and can change the relative amounts of all of them in the mix. Limestone being the main constituent in the mix has a significant impact on the mix quality, and is thus the focus of attention. The more invalid the formula becomes the more inconsistent becomes the clinker SR. This results in high variability in the clinker quality which has the following negative effects :

- I. The cement milling process becomes subject to higher variability resulting in inconsistent cement performance in concrete. This is very important in the construction industry.
- II. Changes in clinker SR results in unstable kiln operations which result in under and over burning of clinker, resulting in wasted or reworked material. The life of the refractory materials used to reline these kilns is reduced, resulting in more frequent kiln stops and increased costs. The growing inaccuracy of the original formula has thus got the Production and Quality managers looking for the following :

- I. A means of calculating the clinker SR for each kiln accurately from only the limestone composition, using a single formula or set of formulae.
- II. A means of updating these formulae periodically to ensure that variability in clinker quality is minimised.

1.3. Objective

To develop a model which will provide a formula to predict clinker silica ratio from limestone quality, and make it possible for the Production Manager to update the coefficients in the formula as and when the independent variables, on which the correlation is based, change.

To overcome the problem with inaccuracies in the formula mentioned in section 1.2 a new one needs to be developed. However, based on the experience we have, viz. that a formula does not hold true for all raw material and fuel qualities, a dynamic model needs to be investigated which will provide a formula to predict the final clinker SR, allow for plant limitations in the development of the formula, and provide a simple means of verifying and updating the formula.

The purpose of this study is to provide a single tool, to do all of the above and satisfy the requirements of the production manager, stated in section 1.2.

1.4. Scope of Study

This study will ultimately lead to a model as discussed in section 1.1. This model will be developed with the Alpha Cement kilns specifically in mind, however, with emphasis on flexibility. The model in its final form could be used for any cement kiln, with any raw materials, on condition that no more than four raw materials are used. The model is specific to the cement process, whereby the ash from the fuel forms part of the final product - clinker. The chemical ratios used to qualify the quality are specific to cement manufacture and thus the model cannot be used for any other process.

The model is based on already well known cement mix design theory and linear programming. These already well known and extensively used tools will be used in a structured fashion to provide the production manager with one of a number of options viz.,

- A prediction of clinker SR at the press of a button, where the SR will be directly calculated from the mix design using linear programming techniques. (There are already many software alternatives available to do this)
- a formula to calculate the clinker SIR
- the ability to verify and recalculate the coefficients of the formula developed at any time
- the ability to test the sensitivity of the formula to changes in the raw materials or fuel qualities.

1.5. Research Methodology

Research has been done as follows :

1. A literature survey was done on :
 - Literature pertaining to cement raw mix design and basic cement chemistry (used to identify the problem in section 3.2)
 - Literature on statistical methods pertaining to the development of, testing and verification of a mathematical model (chapter 2).
2. The Cement process was studied in depth at the cement works in order to enable the author to understand the literature, and thereby provide a model which is both practical and realistic.
3. Alternative statistical methods were studied and compared with respect to the requirements of the model. From this, a decision was made on which method would be used.
4. The accuracy required for the model was confirmed with management and cross checked with the theory to ensure statistical correctness and feasibility.
5. The model proposed was tested according to the literature studied.
6. The model was used on the limestone stockpiles which had already been used to manufacture clinker, and the results compared statistically with the clinker quality to establish accuracy of the predictions.

1.6. Conclusion

In this chapter the setting in which the problem needs to be solved has been outlined. With this together with the problem statement and a clear statement of the purpose of this paper, available literature can be studied and the appropriate methods to achieve the goal identified. The terms used in the paper have been defined, so the reader will find it easy to follow the coming chapters.

In chapter 2, a summary of the appropriate literature will be provided, with reference to the requirements of this paper. In chapter 3, the model will be developed using a case study. The literature from chapter 2 will be applied directly to achieve the end product required. In chapter 4 the final conclusions will be drawn and recommendations made.

CHAPTER 2

THEORETICAL LEGITIMACY

2.1. Introduction

In this chapter a summary of literature studied will be provided. Literature on the development of a mathematical model will be studied as well as literature pertaining specifically to linear programming and multiple regression analysis. Finally, methods of testing and verifying a mathematical model, will be studied.

The core interest is in predicting clinker SR from limestone quality, using a simple formula(e). In order to do this, one needs to find a correlation between the limestone quality and the quality of the clinker which was produced from that limestone. Due to the logistics of the process, it is not practical to carry out this last step repeatedly. This, however, is not a problem with a very good understanding of the process, i.e. If the chemical compositions of all the raw materials and fuel are known as well as what their proportions are in the mix and how much fuel is required, the exact clinker composition can be easily calculated (Holderbank cement course, 1996). All of the above, except the relative amounts of raw materials in the mix, are known to the production manager when building a limestone stockpile.

The production manager knows exactly what is required in terms of clinker composition. Thus, he needs to establish what the mix composition will be for a given grade of limestone, to achieve the

desired clinker quality. In effect, the manager needs the ability to simulate the clinker production process, chemically, in order to gather sufficient data. This data can then be analysed and a correlation sought between the limestone quality and the resultant clinker quality.

2.2. Constructing a mathematical model

According to Giordano & Weir [1,35], there are 6 steps in constructing a mathematical model, viz. :

step 1. - *Identify the problem* : this is often the most difficult step as the modeller has to sift through all the data and information and identify the exact aspect to be studied.

step 2. - *Make assumptions* : All factors cannot be taken into account in most cases, and thus to simplify the relationship, assumptions must be made. These assumptions fall into two categories, namely :

A. Classification of variables : The variables must be classified into two groups, viz. dependent or independent variables. If an independent variable :

1. has a relatively small effect compared to others, or
2. affects all alternatives reasonably equally,
it can be ignored.

B. Determination of interrelationships among the variables : Before a relationship between the variables is hypothesised, further simplification should be made by eliminating independent variables which are interrelated. If the problem is

sufficiently complex, it is possible to study sub-models. A sub-model is a separate study of one independent variable.

step 3. - *Solve or interpret the model* : Use the model and see what answer you get.

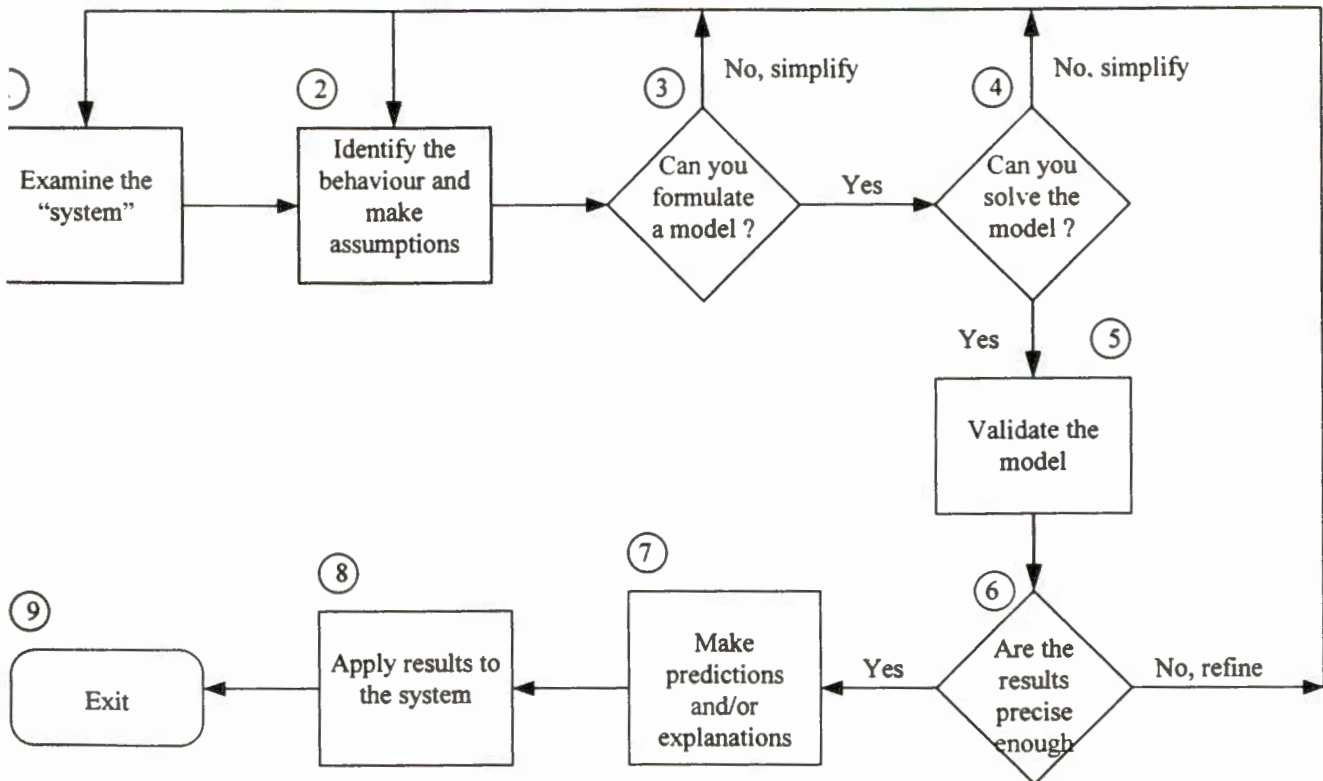
step 4. - *Verify the model* : test the results and ask yourself the following questions :

- A. Does the model answer the question in step 1 ?
- B. Is the model usable in a practical sense ?, i.e. can the data needed be gathered effectively?
- C. Does the model make common sense ?, i.e. **be careful not to test the model on the same data used to design it !** The assumptions made in step 2 could be valid for a restricted range of the independent variables only. Remember that a model does not become law just because it is verified repeatedly. Rather, the reasonableness of the model can be corroborated through the data collected.

step 5. - *Implement the model* : Make sure that the model is user friendly, otherwise it will not be used.

step 6. - *Maintain the model* : If the specific problem changes slightly, or any of the assumptions become invalid, the model must be reviewed.

Giordano & Weir [1,38], go further to describe the modelling process as having an iterative nature, as demonstrated in figure 2.1

Figure 2.1 - Producing a Mathematical model (Giordano, et al, 1985:39).

As was previously mentioned, the clinker silica ratio model has two phases, i.e.

- I. Data generation - here optimisation techniques are used to calculate clinker qualities from known limestone qualities,
- II. Establishment of an approximation formula based on the data generated in I, above.

This split makes it easy to focus on the appropriate literature, relevant to each phase. The first stage is an optimisation problem, where optimisation is defined as the process of minimising or maximising (Giordano, et al, 1985 : 120). An optimisation problem is said to be unconstrained if there are no constraints and constrained if one or more side conditions exist. Our problem most definitely does have constraints. These include process and quality related

constraints as well as logical constraints such as, for e.g., there cannot be a negative proportion of an additive in the mix and the total composition of the mix cannot be greater than 100%. We are thus dealing with a constrained optimisation problem.

An optimisation problem can be defined by the basic model described in figure 2.2. Optimisation problems can be linear or non-linear. If the criteria listed below are satisfied, then the problem is linear and called a **linear program** (Redelinghuis, et al, 1994 : 214) :

- I. There is one function (formula) to be solved. This formula is referred to as an **objective function**. If there is only one objective function to be solved, then there is spoken of a **unique** objective function.
- II. Decision variables (see figure 2.2. for definition) in the objective function and constraint functions must have a power of 1.
- III. No term in the objective function or any of the constraints can contain a product of the decision variables, i.e. $X_1 \times X_2$ is not allowed.
- IV. The coefficients of the decision variables in the objective function and each constraint are constants.
- V. The decision variables can be fractional or integer.

Figure 2.2. - The definition of a mathematical model, in mathematical notation.

$$\text{Solve } f_j(\mathbf{X}) \text{ for } j \text{ in } J$$

Subject to:

$$g_i(\mathbf{X}) \begin{matrix} \geq \\ = \\ \leq \end{matrix} b_i \text{ for } i \text{ in } I$$

Where :

- j describes the number of functions to optimise
- J is the total finite set of functions
- X is a vector = $\{x_1, x_2, x_3, \dots, x_n\}$, where each element is known as a **decision variable**
- $f_j(\mathbf{X})$ represents the function(s), called the **objective function(s)**, which must be optimised.
- $g_i(\mathbf{X})$ represents the **constraint functions** or **side conditions** which must be met.

If any one of the above criteria are **not** met, then the program is **non-linear**. Linear programs are the simplest to solve (Williams, 1978 : 5), and there are a number of techniques used to do this (Williams, 1978 : 36, 44, 54); (Thie, 1988 : 45, 176, 211). It would thus be appropriate at this point to ascertain whether our problem is linear or non-linear. This is to prevent studying theory on solving techniques which are not appropriate. The problem has been defined according to the notation explained above, in chapter 3, paragraph 3.3, and checked for the criteria described above. It turns out that all of the

criteria, are met, and thus the optimisation part of the problem is a linear program. Now it must be determined as to exactly how the linear program which has been defined in chapter 3, must be solved.

2.3. Linear Programming.

The definition of an optimisation problem can now be expanded for a linear program, which can now be written as follows (Vanderbei, 1996 : 7):

maximise	$C_1X_1 + C_2X_2 + \dots + C_nX_n$	(maximise $f(X)$)
subject to	$g_{11}X_1 + g_{12}X_2 + \dots + g_{1n}X_n \leq b_1$	\geq
	$g_{21}X_1 + g_{22}X_2 + \dots + g_{2n}X_n \leq b_2$	(subject to $g_i(X) = b_i$ for i in I)
	\vdots	\leq
	$g_{m1}X_1 + g_{m2}X_2 + \dots + g_{mn}X_n \leq b_m$	
	$X_1, X_2, \dots, X_n \geq 0$	

This notation is referred to as **standard form**, by statisticians. Note the following though, about the above standard form :

- It is an example - Although it says maximise, it could have stated minimise or equate.
- The less than signs in the constraints are examples, they can be greater than or equality signs as well, depending on the problem.
- $x_1, x_2, \dots, x_n \geq 0$, is an example of **logical constraints**. More logical constraints can be applied depending on the problem.

The solution to the problem can be **feasible** or infeasible. It is feasible if all the constraint requirements are met. However, an

Optimal solution is only obtained when a **feasible solution** provides the desired maximum, minimum or equality.

There are a number of methods available for solving linear programs, of which the following are most widely accepted (Redelinghuis, *et al*, 1989: 222, 255, 298, 342), (Thie, 1988 : 45, 176, 211):

- a) Simplex method.
- b) Algebraic methods.
- c) Graphical methods.
- d) Integer programming.
- e) Transportation methods.

In most cases, it is not practical to use any of the methods mentioned above, manually. A computer is usually used due to time constraints and problem complexity (Hicks, et al., 1995 : 12, 50). There are a number of software packages available to do this, including spreadsheets such as Lotus, Quattro-Pro and Excel. A computer will be used to solve the model, so none of the above methods will be discussed in this paper. Microsoft software is an extensively used package, and also the package preferred by the author. To use the built in functions in Excel will ensure easier maintenance of the model, as less computer programming code will be required.

2.4. Choice of Analysis

Any researcher faced with the need to analyse data requires a rationale for choosing a particular method of analysis (Kleinbaum, et al, 1998:11). There are four considerations in making this choice, i.e.

- I. the purpose of the investigation
- II. the mathematical characteristics of the variables involved
- III. the statistical assumptions made about these variables
- IV. how the data are collected

There are a number of analysis techniques available (Kleinbaum, et al, 1998 : 12), as described in the table 2.1 below :

Table 2.1. - Analysis techniques

Method	Dependent	Independent	General Purpose
Multiple regression	Continuous	Classically all continuous, but various types	To describe strength, direction and extent of relationship between several independent variables and a continuous dependent variable
Analysis of variance	Continuous	All nominal	To describe the relationship between a continuous dependent variable and one or more nominal independent variables
Analysis of covariance	Continuous	Mixture of nominal & continuous variables	To describe a relationship between a continuous dependent variable and one or more nominal independent variables, controlling for the effect of one or more continuous independent variables
Logistic regression	Dichotomous	Mixture of various types	To determine how one/more independent variables are related to the probability of the occurrence of one or two possible outcomes
Poisson regression	Discrete	Mixture of various types	To determine how one or more independent variables are related to the rate of occurrence of some outcome.

This table will be referred to later in this chapter.

2.4.1. Multiple Linear Regression Analysis

Clinker SR plotted versus various limestone components and ratios (scatter diagrams in appendix 1) showed evidence of a few good, but not perfect, linear relationships, the others showing very weak or no relationships. The problem has several continuous variables which, referring to table 2.1, implies that the appropriate technique to quantify the relationship is multiple regression analysis. The scatter diagrams indicate a linear relationship, and thus Multiple Linear regression would appear appropriate.

The multiple linear regression will provide a linear function with the following form :

$$Y = m_1X_1 + m_2X_2 + \dots + m_nX_n \quad (2.3.1.1)$$

where : X_i are the independent variables

m_i are coefficients (constants) determined from the Multiple regression analysis.

There are a number of statistical assumptions which are made when using multiple linear regression, and these must be studied and compared to the situation on hand. Consequently, it can be concluded whether or not the method is appropriate.

The assumptions are as follows (Kleinbaum, et al,1998:115) :

1. **Existence** : For each specific combination of values of the independent variables $X_1 - X_n$, Y is a random variable with a

certain probability distribution having finite mean and variance.

2. **Independence** : The Y values are statistically independent of one another.
3. **Linearity** : The mean value of Y for each specific combination of $X_1 - X_n$, is a linear function of $X_1..X_n$.
4. **Homoscedasticity** : The variance of Y is the same for any fixed combination of $X_1..X_n$.
5. **Normality** : For any fixed combination of $X_1..X_n$, the variable Y is normally distributed.

Multiple linear regression makes use of the **Least-squares approach**, which chooses as the best fitting model, the one that minimises the sum of the squares of the distances between the observed responses and those predicted by the fitted model.

Referring to chapter 3 of this paper, the reader will realise that all five of the above assumptions are satisfied by the problem on hand, making multiple linear regression an appropriate analysis tool.

2.5. Model testing and Verification

2.5.1. Association versus Causality - Statistical Inference

The finding of a “statistically significant” association in a particular study (no matter how well done) does not establish a causal relationship (Kleinbaum, et al, 1998 : 36), i.e. the relationship could be coincidental. There are two general categories of statistical inference - estimation and hypothesis testing (Kleinbaum, et al, 1998 : 21). In multiple regression, Hypothesis testing is generally used (Redelinghuis, et al, 1989 : 12).

Hypothesis testing is a procedure based on sample evidence and probability theory, used to determine whether the hypothesis is a reasonable statement and should not be rejected, or is unreasonable and should be rejected (Lind, et al, 1994 : 249). The procedure followed can be described as follows :

1. - The Null hypotheses, H_0 , is made. This is an assumption made about a population parameter.
 - The alternate hypothesis, H_1 , is made. This is a statement which will be accepted if the sample data provide enough evidence that the null hypothesis is false.
2. The probability of rejecting the null hypothesis when it is actually true, the level of significance, must be stated.

3. Find the test statistic - there are a number of possibilities available, including Student's t-test, the F distribution and the Chi square (Lind, et al, 1994 : 247, 281, 307, 327, 361). These will not be discussed in detail for the purposes of this paper.
4. Determine the decision rule, i.e. state the condition of rejection or non-rejection. For this the critical value must be determined. The critical value is the dividing point between the region where the null hypothesis is rejected and the region where it is not rejected. Later text will show that there are two appropriate test statistics for the purpose of this paper, viz. the F - distribution and the Normal distribution, Z. These will be explained in some more detail.

2.5.2. Testing and verification of Multiple linear regression

There are three basic types of tests to be done (Kleinbaum, et al, 1998 : 136). For each test the hypothesis testing procedure is followed. These tests are :

1. Overall test : Does the entire set of independent variables contribute significantly to the prediction of Y ?
2. Test for addition of a single variable : Does the addition of one independent variable add significantly to the prediction of Y achieved by other independent variables already present in the model ?
3. Test for addition of a group of variables : Does the addition of a group of independent variables add significantly to the

prediction of Y achieved by other independent variables already present in the model ?

These three steps were carried out already in the systems analyses to identify the best form of the correlation equation(s). They are statistical tests using hypothesis theory, as will be explained below for multiple regressions. More detail on these procedures can be found in Kleinbaum, et al (1, 136). How to interpret the results is important, and this will now be discussed.

The results from a linear regression analysis are provided in most cases, in a tabular form. This is called an ANOVA table, and can be found in slightly different forms (Kleinbaum, et al, 1998 : 118). ANOVA is short for "analysis of variance", and is a procedure used to determine whether the means from two or more samples are drawn from populations with the same mean.

There are two functions which need to be analysed in order to establish whether or not the correlation is accurate and whether or not it is by chance. First is the coefficient of determination (r^2). For example, if $r^2 = 0.965$, then it can be said that 96.5% of the variation in Y, is due to variation in the set of independent variables, $X_1...X_n$ (Lind, et al, 1994 : 333). Secondly, the F distribution and thus the F statistic is used to

test whether the coefficients from the regression analysis are valid. By carrying out a hypothesis test, using the steps presented earlier in this chapter, and using the F distribution as the test statistic, it can be determined whether or not the relationship is by chance or not. The steps are as follows (Lind, et al, 1994 : 373) :

Given that the coefficients from the regression are m_1 , m_2 , m_3 , and that the coefficients in the population are β_1 , β_2 , β_3 . We test if the net regression coefficients in the population are zero.

Step 1 : $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$

H_1 : Not all β 's are 0

If the null hypotheses is true, it implies that the regression coefficients are all zero and, logically, are of no use in predicting the dependent variable. If this should be the case, then another approach should be adopted to find a correlation.

Step 2 : Set significance level to 0.05

Step 3 : The appropriate test statistic is the F distribution

Step 4 : The decision rule is obtained from tables in most statistics text books (Redelinghuis, et al, 1989 : 12). A single tailed test is used.

The F critical value is obtained from the table as follows :
There are k degrees freedom in the numerator, where k is the number of independent variables. The degrees freedom in the denominator are calculated as $n - (k + 1)$, where n is the number of observations in the sample. $F_{critical}$ is obtained by looking up F from the table corresponding to the two definitions of degrees freedom explained above.

If the F critical value is smaller than the F value observed, then there is a relationship between the variables, i.e. the identified correlation is not incidental. Note, that the test is applied to the same data used to obtain the relationship. The regressions carried out in the model have $F \gg F_{critical}$. (F is very much larger than $F_{critical}$). Thus the relationship is causal and not incidental. The lowest value for r^2 obtained in the model, was 0.922. Thus, at worst, 92.2% of the variation in the clinker silica ratio can be accounted for in the variation of the LSF, Fe_2O_3 and SR of the limestone itself.

To be consistent with the theory of model development, for the purposes of this project, the formula will be tested on data which was not used in it's development. The hypothesis testing method to do this, is a test to determine whether two sets of samples come from the same population (Underhil, 1985 : 283). Refer to appendix 2 for a decision tree provided by Underhil. This test is only done to verify the procedure - it is

not necessary for the manager to do this. Once it has been proven, only the regression analysis needs to be analysed each time a new formula is developed. However, the second test can be used to check if the formula remains valid, at any time.

2.5.3. Hypothesis testing : Two population means (large samples)

For this, the hypothesis testing steps are again used but as follows (Lind, et al, 1994 : 270) :

Step 1 : $H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

where μ_1 and μ_2 are the means of the two sets of samples.

If the null hypotheses is true, it implies that the two samples come from the same population, and thus that the formula is suitable.

Step 2 : Set significance level to 0.01, and the test is **two tailed**.

Step 3 : The appropriate test statistic is the Z distribution, which is calculated as follows :

$$Z = (X_1 - X_2) / (S_1^2 / n_1 - S_2^2 / n_2)^{1/2} ;$$

Step 4 : The decision rule is obtained from the standard normal distribution. $Z_{critical}$ is obtained by looking up the calculated Z on the tables. For a significance level of 0.01 and a two tailed test, $Z_{critical}$ is 2.33. Thus the null hypothesis will be rejected if the Z factor is larger than 2.33.

2.6. Conclusion

The literature available on mathematical modelling is extensive, and makes very interesting reading. There are many techniques available to fit a model to any given data, and the methods available to verify this are just as numerous. We have only touched on a small area of this field of study. A clear message which one obtained from the literature, is that of making assumptions. Assumptions must be made when carrying out such an exercise. These assumptions must be understood by both the programmer or modeller and anyone using this model. If the user of such a model is not ware of the underlying assumptions, the model can be used incorrectly, or the model may become inaccurate and the user will not know why.

The model as well as the testing and verification thereof must also be practical, i.e. the model must be user friendly and the user must be able to, with some ease, use and test the model. In the following chapter the model to predict clinker SR from limestone analysis only, will be discussed in detail. The model is a direct application of the literature which was studied in this chapter.

CHAPTER 3

MODEL DEVELOPMENT USING A CASE STUDY

3.1. Introduction.

We are interested in predicting clinker quality from limestone quality, using a simple set of formulae. In order to do this, a correlation between the limestone quality and the quality of the clinker which was produced from that limestone, must be found. If we know what the chemical compositions of all the raw materials and fuel are, what their compositions are in the mix and how much fuel is required, the exact clinker composition can be easily calculated. All of the above, except the relative amounts of raw materials in the mix, are known to the production manager when building a limestone stockpile.

The production manager knows exactly what he requires in terms of clinker composition. Thus, he needs to establish what the mix composition will be for a given grade of limestone, to achieve the desired clinker quality. In effect, the manager needs the ability to simulate the clinker production process, chemically, in order to gather sufficient data.

This data can then be analysed and a correlation sought between the limestone quality and the resultant clinker quality. Thus, the problem is two fold, viz.

- I. the raw mix needs to be determined for a given limestone stockpile and the desired clinker quality, i.e. an optimisation exercise.
- II. a simplified relationship between the clinker quality (the solution from the optimisation problem) and the limestone quality must then be identified, such that an approximation formula can be developed.

A model to satisfy the production managers above requirements has been developed. It takes the form of a macro driven Excel spreadsheet. In this chapter, the model will be discussed in detail, tested and verified.

3.2. Frame Of reference

3.2.1. The Cement Production Process

This study is focused on the production Process of a multi-kiln cement plant, with highly variable limestone reserves, and / or highly variable additive materials and / or fuel. In order to create, test and verify the model, which could be used at any cement clinker manufacturing plant, the Alpha Cement Dudfield factory will be used.

A Cement Production Process can be divided into three main processes, viz.

- **Quarrying**, where limestone is mined via open cast mining methods, crushed and then stockpiled.
- **Clinker manufacture**, where stockpiled limestone is added to a number of additives and milled to a fine

powder. This is then clinkered (burnt) at 1400°C to form Portland cement clinker.

- **Cement milling**, where Portland cement clinker is milled with Gypsum to a fine powder, known as Portland cement.

The processes which this study will concentrate on are the first and second, i.e. mining and clinker burning. The results of this study are aimed at achieving consistency during the clinkering process, but will be applied during the first part of the process, i.e. during the building of limestone stockpiles.

3.2.2. Cement Chemistry

The main chemical components in limestone, cement and clinker can be quantified by the following oxides : CaO (calcium oxide), SiO_2 (Silica dioxide), Al_2O_3 (alumina) and Fe_2O_3 (iron oxide), MgO (magnesium oxide), K_2O (potassium oxide) and Na_2O (Sodium oxide).

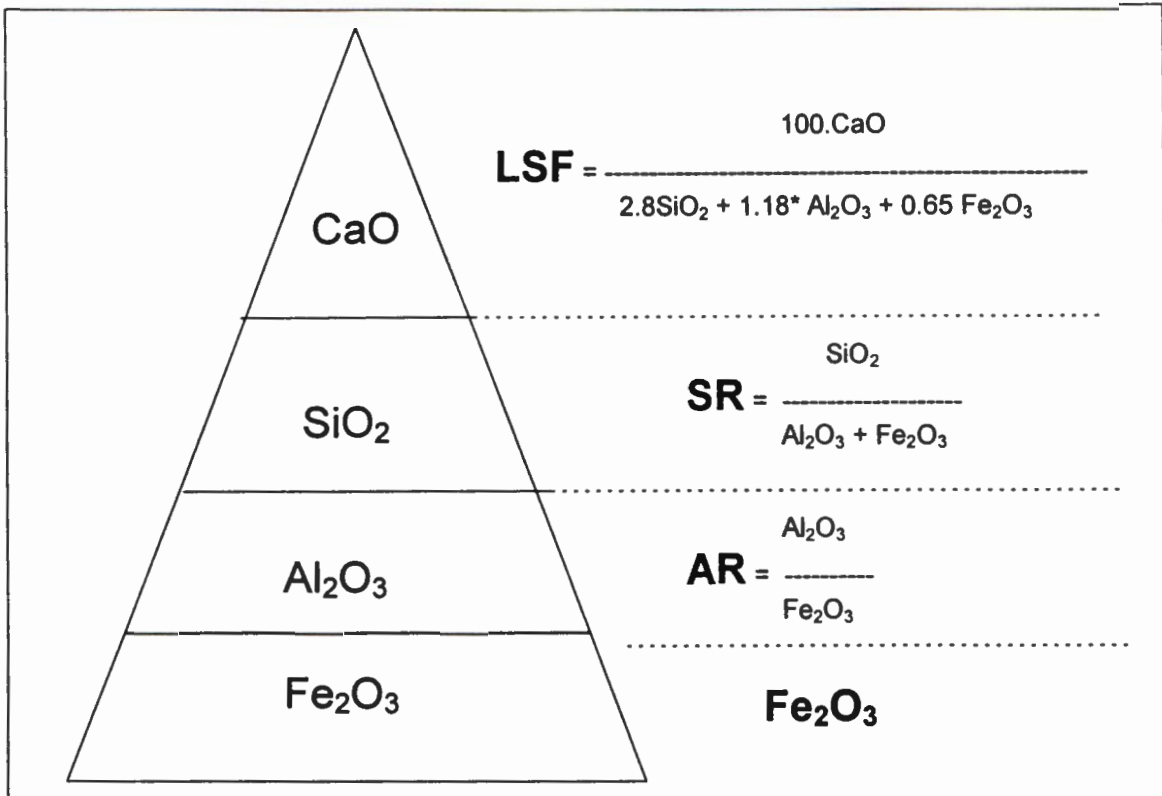
The quantities of all of these are important for cement performance. A minimum MgO content is desirable while a limited maximum alkali (K_2O and / or Na_2O) content is required. The first four oxides are very important in the cement chemistry.

The complex chemical and physical reactions which take place in a kiln require that the relative amounts of these components be within certain limits. Thus, for these oxides, it is not only the quantity which is important, but also the amount relative to the other three.

It is for this reason that the Lime saturation factor (LSF), Silica Modulus (SR) and Alumina Ratio (AR) are used to quantify cement, clinker and limestone quality. The hierarchy and definitions of these ratios are shown in figure 1.1. (See also Appendix 4.)

By controlling the limestone LSF one can ensure that a clinker LSF within certain limits will be obtained. Thus, it is ensured that the relative amounts of the four oxides are correct. However, there is no guarantee on where the SR will lie in the clinker (see the definitions in figure 1.1).

The SR (and the LSF) of the clinker is dependent on the Limestone quality, the amount of additives added to the limestone and the quality and amount of fuel used in the clinkering process. Thus, controlling SR of the limestone still does not guarantee SR of the clinker.

Figure 3.1 - Chemical ratios used in cement quality control.

The clinker SR plays a crucial role in the performance of the cement clinker rotary kiln and thus on the costs of manufacture. It is thus important that the SR in the clinker is estimated from the limestone quality in some way, to ensure that, through control :

- the SR is not too high or too low for quality and / or process reasons
- the SR variability is low enough to ensure stable, consistent kiln operation and clinker quality.

3.2.3. Overview of Clinker manufacturing Process

In limestone quality control, there are four basic focus issues, viz.

- I. Cement quality
- II. Cement consistency
- III. Kiln operation and optimisation
- IV. Optimisation of limestone reserves

Although the content of this study has an indirect effect on IV above, the work is being done to achieve III above, which in turn will lead to I and II. In order to increase the readers understanding, a short description of the clinkering process is important, viz. :

Limestone and additive materials (the likes of shale, clay, iron ore, magnetite, boiler ash and others), are milled to a fine powder characterised, typically, by a 10% retention on a 90 micron screen. This material, known as raw meal, is stored and blended in storage silos. From there it is extracted and fed to the rotary kiln from one end, known as the feed end. Pulverised coal is fired from the other end, known as the firing end. Air is sucked into the system via large exhaust fans and is used for combustion of the coal, providing the heat required to dry the raw material, drive off CO_2 from CaCO_3 and MgCO_3 and provide a clinker temperature of approximately 1400°C for the final chemical reaction to take place.

Due to the high temperatures, the kiln shell is lined with refractory bricks. There is a temperature gradient from the feed end to the firing end of the kiln. Each temperature zone has refractory bricks designed for those temperatures and the characteristics of the gases and solid materials in the kiln.

The final reaction takes place in the zone called the burning zone. A characteristic of the clinkering process is the formation of a solid coating on the refractory materials in this area. This is a desirable situation as the bricks are protected from the kiln atmosphere and temperatures. This results in longer life of the refractory materials. This situation is desirable under the assumption that the coating is stable. The coating penetrates the bricks and forms a bond on the surface. When the coating falls off, so does the top section of the brick. The more often this happens, the sooner the bricks fail.

The clinker LSF and SR are the key players in determining whether or not the coating will be stable or not. Provided these fall within certain limits a desirable coating will be formed. If they fall out of these limits the coating will not be formed. Most importantly, if they are variable and fluctuate around the desired limits, the coating forms and then breaks off and so on. This is the most undesirable situation.

In addition to the above, the fuel (coal) consumption also needs to be optimised. A high LSF, high SR mix is typically a hard to burn mix, which will require more fuel than a Low LSF, low SR mix which is very easy to burn.

The limits are set according to the quality requirements, but to optimise costs the variation between the limits should be minimised.

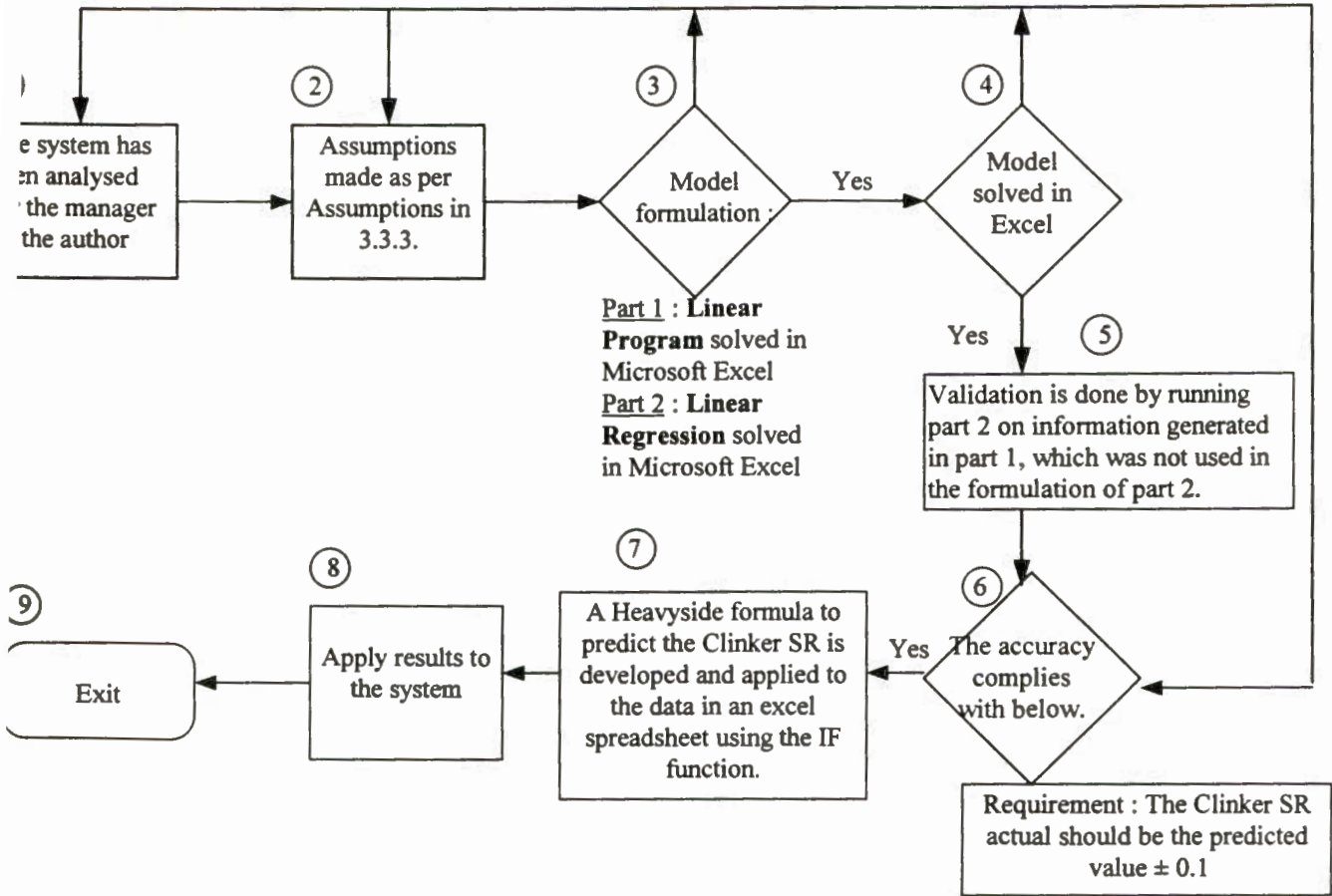
3.3. Description of the Model

3.3.1. The model flowsheet

The model can be described using Giordano & Weir's (1985:38) diagram (figure 2.1.) as a basis. The model can thus be described as per figure 3.1. The model, although, it is one entity, has been split into two parts for ease of explanation and to assist with the identification of literature requirements.

Part 2 of the model provides the manager with the ultimate goal, i.e. it provides a set of formulae for the manager to predict clinker SR from limestone data only. Part 1 is provided as a **tool** to generate the data needed for part 2 of the model.

Figure 3.2. - Model Flowsheet



3.3.2 System Analysis

The clinker composition, chemically, is a simple mix of the limestone, three additive materials and coal ash, all in ignited form¹. This can be described as follows :

$$x_1 + x_2 + x_3 + x_4 + x_5 = 1; \quad (3.3.2.1)$$

where :

- x_1 = fraction limestone
- x_2 = fraction additive 1
- x_3 = fraction additive 2
- x_4 = fraction additive 3
- x_5 = fraction coal ash

¹ When Carbonates such as CaCO_3 and MgCO_3 , are ignited, CO_2 is driven off leaving CaO and MgO . Thus the proportions after and prior to ignition differ. As clinker is produced at $\pm 1400^\circ\text{C}$, it is very safe to assume that all raw materials will be ignited in their final state.

All of these components have an effect on the clinker LSF and SR. Depending on the limestone quality, the mix composition will differ for the same clinker LSF and SR. It has been determined, from experience, that limestone with an LSF ranging from 117 to 133 can be used to produce clinker with an LSF in the range from 93 to 96. However, depending on the limestone quality and how the mix is proportioned, there is a wide range of possibilities for the clinker SR. Thus, an attempt must be made to predict the clinker SR to avoid fluctuations.

Requirements of the clinker composition are as follows :

1. LSF = Ideal target \pm 1.5
2. SR = ideal target \pm 0.15
3. K20 must be maximised, but to a limit, in order to maximise early cement strength development.
4. Fe₂O₃ = set point \pm 0.5, in order to ensure that there is optimum flux²

The following **limitations** have been identified :

1. The raw mill feeders have limitations which must be taken into account.
2. The heat consumption's differ for each kiln
3. All kilns do not use the same amount of different additives.

² flux is a measure of the physical state of the clinker in the burning zone of the kiln.

The following **relationships** have been identified from the data. (Scatter diagrams can be seen in appendix 1) :

1. As limestone LSF increases, Clinker SR decreases (strong tendency)
2. As limestone Fe₂O₃ content increases, clinker SR increases (fair tendency)
3. There is a tendency, although very weak by itself, for clinker SR to increase as limestone SR increases.
4. Clinker SR has a poor correlation with limestone AR

Using the three steps described in paragraph 2.5.2, the best form of the model was obtained, viz. :

$$\text{Clinker SR} = m_1 \text{LSF}_{\text{limestone}} + m_2 (1/\text{SR}_{\text{limestone}}) + m_3 (1/\text{Fe}_2\text{O}_{3,\text{limestone}}) + c \dots$$

3.3.1

where m_i are the coefficients and c is a constant.

3.3.3. Assumptions

1. The model assumes constant quality for additives and coal ash, for each exercise.
2. The heat consumption of the kiln is constant for each exercise.
3. The coal ash content of the coal is constant for each exercise.
4. The coal quality (and thus type or mix of types) is constant for each exercise.

5. Clinker SR is proportional to the limestone LSF, and is best demonstrated when applied as a Heavyside³ function. It is assumed that the Heavyside limits can be kept constant, regardless.
6. Variation in clinker SR follows a normal distribution.
7. No segregation of raw meal takes place in the raw meal silos.
8. The target clinker LSF will be achieved by the plant operating staff.

3.3.4. Model Part 1 formulation - Optimisation Problem.

This part of the model is used to generate data on clinker quality from various limestone stockpiles. The model can be described, based on the notation provided in chapter 2 for an optimisation problem, as follows :

There is one objective function $f(X)$, which requires that the Clinker LSF is equal to some desired value, L_d , the latter being stipulated by the production and/or quality manager, i.e.

$$LSF = L_d.$$

The clinker composition, chemically, is a simple mix of the limestone, three additive materials and coal ash, all in ignited form⁴, as per equation 3.3.2.1.

³ Heavyside function $H(x)$ is used when the relationship between the independent variable, x , and the dependent variable, H , is not constant for all ranges of x , but when a number of different functions can be identified for different ranges. Thus, when x lies in range 1, relationship 1 is applied, in range 2 function 2 is applied and so on. A Heavyside function is easy to apply in Microsoft Excel and other spreadsheet packages, by using the IF functions.

⁴ When Carbonates such as CaCO_3 and MgCO_3 , are ignited, CO_2 is driven off leaving CaO and MgO . Thus the proportions after and prior to ignition differ. As clinker is produced at $\pm 1400^\circ\text{C}$, it is very safe to assume that all raw materials will be ignited in their final state.

The coal ash fraction, x_5 (refer to eqn 3.3.2.1), is assumed constant for a given kiln for each exercise. This is based on the assumptions that the heat consumption remains constant, as does fuel used and the quality of that fuel.

Thus, the total amount of any chemical in the clinker, for example CaO, can be calculated as follows :

$$x_1C_1 + x_2C_2 + x_3C_3 + x_4C_4 + x_5C_5 = C_{\text{clinker}}$$

where $C_i = \text{CaO}$ (Calcium oxide) in each of the components.

The formula for LSF was provided in chapter 1. By substituting into the formula, the objective function can be obtained :

$$\frac{100(x_1C_1 + x_2C_2 + x_3C_3 + x_4C_4 + x_5C_5)}{2.8(x_1S_1 + x_2S_2 + x_3S_3 + x_4S_4 + x_5S_5) + 1.18(x_1A_1 + x_2A_2 + x_3A_3 + x_4A_4 + x_5A_5) + 0.65(x_1F_1 + x_2F_2 + x_3F_3 + x_4F_4 + x_5F_5)} = L_d$$

This non-linear formula can be manipulated into a linear form, with the new objective function being:

$$x_1Q_1 + x_2Q_2 + x_3Q_3 + x_4Q_4 + x_5Q_5 = 0;$$

where $Q_i = 100C_i + L_d(2.8)S_i - L_d(1.18)A_i - L_d(0.65)F_i$.

Subject to the following constraints :

1. Fe₂O₃

$$x_1F_1 + x_2F_2 + x_3F_3 + x_4F_4 + x_5F_5 \geq F_{dL}$$

$$x_1F_1 + x_2F_2 + x_3F_3 + x_4F_4 + x_5F_5 \leq F_{dH}$$

2. Silica Ratio

$$2.1. \quad \frac{x_1S_1 + x_2S_2 + x_3S_3 + x_4S_4 + x_5S_5}{x_1A_1 + x_2A_2 + x_3A_3 + x_4A_4 + x_5A_5 + x_1F_1 + x_2F_2 + x_3F_3 + x_4F_4 + x_5F_5} \geq S_{dL}$$

Or, this equation can be linearised as follows :

$$x_1\beta_1 + x_2\beta_2 + x_3\beta_3 + x_4\beta_4 + x_5\beta_5 \geq 0;$$

where $\beta_i = S_i - S_{dL}A_i - S_{dL}F_i$.

$$2.2. \quad \frac{x_1S_1 + x_2S_2 + x_3S_3 + x_4S_4 + x_5S_5}{x_1A_1 + x_2A_2 + x_3A_3 + x_4A_4 + x_5A_5 + x_1F_1 + x_2F_2 + x_3F_3 + x_4F_4 + x_5F_5} \leq S_{dH}$$

Or, this equation can be linearised as follows :

$$x_1\delta_1 + x_2\delta_2 + x_3\delta_3 + x_4\delta_4 + x_5\delta_5 + w = 0;$$

where $\delta_i = S_i - S_{dH}A_i - S_{dH}F_i$; w is a slack variable⁵

3. Alkalis

$$x_1K_1 + x_2K_2 + x_3K_3 + x_4K_4 + x_5K_5 \geq K_{dL}$$

4. Feeder limitations

$$\begin{aligned} x_1 &\geq x_{1L} ; x_1 \leq x_{1H} ; \\ x_2 &\geq x_{2L} ; x_2 \leq x_{2H} ; \\ x_3 &\geq x_{3L} ; x_3 \leq x_{3H} ; \\ x_4 &\geq x_{4L} ; x_4 \leq x_{4H} . \end{aligned}$$

5. Logical Constraints

$$\begin{aligned} x_1 &\geq 0; \\ x_2 &\geq 0; \\ x_3 &\geq 0; \\ x_4 &\geq 0. \\ x_1 + x_2 + x_3 + x_4 + x_5 &= 1; \end{aligned}$$

This model satisfies all five of the criteria stipulated for a linear program, in the literature, and is thus a linear programming problem. Microsoft Excel has an automatic linear program

⁵ A slack variable has been used to avoid using a smaller than sign. It is a constant.

function which satisfies the needs of our model. Fortunately, Excel does all the manipulations for the Manager. Thus, although it is important that he/she understands the model described above, the computer will do everything.

Examples of the input and output screens are presented in appendix 3. The primary goal of this part of the model is to produce the data required for part 2. In part 2, the data is used to obtain a formula, in the form of a Heavyside function, for the clinker SR as a function of limestone LSF, SR and Fe₂O₃. Part 1 can be used directly to calculate the clinker SR, and will give a more accurate answer. In fact, the author would urge the manager to use part 1 when possible as it is more accurate.

Part 1 also warns the manager when a solution is not possible due to feeder limitations. This is a useful advantage, as the corrections are made for the feeder limitations, and the corrected clinker SR can be calculated.

The limitations of the feeders are largely accounted for in the formula developed in part 2, due to the fact that the output from part 1 has been corrected for feeder capabilities. However, as the formula is an estimation, no alarming is possible.

Many software packages are available to perform an optimisation exercise like part 1. The principles used are also in no way new to the cement industry, but because it forms part of the total requirement for the model (and the verification thereof), it must be included, explained and understood. This is to ensure that the user is aware of the limitations and assumptions within which the model is used. It is also advantages to provide one tool for the whole process.

3.3.5. Model Part 2 - Prediction Formulation.

Multiple linear regression is used to find the causal relationship between the clinker SR and the limestone quality parameters. During the systems analysis (3.3.2.), these relationships were identified (scatter diagrams in appendix 1). If one studies the linear trend in the clinker SR versus limestone LSF scatter diagram in appendix 1, then it is apparent that if the graph is split up into smaller sections, the sections based on the limestone LSF, a number of more accurate approximations could be obtained, rather than merely one general approximation.

It was found, by experimentation by the author, that more accurate approximations are indeed obtained when the limestone is categorised according to its LSF values. From these relationships, using the theory briefly touched on in chapter 2, the model is defined as follows :

For : Limestone LSF < 110; SR clinker = $m_1X_1 + m_2X_2 + m_3X_3$;
 For : 110 < Limestone LSF < 120; SR clinker = $m'_1X_1 + m'_2X_2 + m'_3X_3$;
 For : 120 < Limestone LSF < 130; SR clinker = $m''_1X_1 + m''_2X_2 + m''_3X_3$;
 For : 130 < Limestone LSF < 140; SR clinker = $m'''_1X_1 + m'''_2X_2 + m'''_3X_3$;
 For : Limestone LSF > 140; SR clinker = $m''''_1X_1 + m''''_2X_2 + m''''_3X_3$;

Where X_1 = Limestone LSF;
 X_2 = 1/Limestone Fe₂O₃;
 X_3 = 1/Limestone SR;
 $m_i, m'_i, m''_i, m'''_i, m''''_i$ are the coefficients obtained from the multiple linear regressions.

The model has been tested for accuracy in each case and for causality. The coefficients of determination vary between 92.2% and 98.7%, meaning that between 92.2% and 98.7% of the variation in the clinker SR is a result of variation in the selected limestone components. Using the F-test in all cases the observed F value is very much higher than the critical value (see chapter 2), indicating a strong causal relationship. Thus, the model has not developed a formula based on a coincidental relationship.

3.3.6. Overview of model Operation

The model is run in Microsoft Excel, and has two parts or sections. In the first part, clinker quality data is generated from a mass of data on limestone qualities, using Linear programming techniques. This generated data is then used in part 2, where from multiple linear regressions, five formula for clinker SR are obtained. Each of these formulae is used for a specific range of limestone LSF, the total formula taking on the form of a Heavyside function. The combined formula is programmed in excel using a the IF function option.

Two models have been developed, one for 2 kilns combined, and the other for a single kiln line. The reason for combining two kilns, is that they may share common limestone stockpiles. Thus, a compromise is made on the accuracy of the predictions for these kilns, individually, but with the plant layout and logistics, this is often the best possible solution.

Each of these models can be used for any kiln or combination of two kilns sharing a common stockpile. However, to prevent the manager from having to re-enter the constants, the model has been saved as two separate models.

The linear programming model (part 1) can also be used directly to calculate the optimum clinker SR from a given limestone quality, and part 2 can be skipped. This is the more accurate method, as much of the error in the approximation is eliminated.

3.4. Model testing and verification

3.4.1. Testing and verification of the Linear programming model

The linear programming part of the model was tested by substituting the recommended mix proportions into formula 3.3.2.1, and then applying the LSF formula (formula 1.1.1), SR formula (formula 1.1.2) and AR formula (formula 1.1.3). This was done on 30 random limestone samples. The results were all identical, indicating that there are no errors in this part of

the excel spreadsheet itself. To test the suitability in practice, a number of completed limestone stockpiles were used to calculate the clinker SR. The clinker SR actually produced from these same stockpiles was then compared with the predictions.

This part of the testing must be done very carefully. The reason for this statement, is that the model cannot take into account abnormal problems or conditions experienced by the plant. For example, the raw meal in the raw meal silo can begin to segregate out if the stocks are too low or if there is a problem with the blending system. This can result in very high quality material entering the kiln system for a number of days, followed by some very low quality material, and so on. Thus, one has to ensure that these anomalies are not included in the tests for conformance.

The above example occurs very infrequently, but what does occur very frequently, is that an LSF exactly equal to that used in our model is not obtained. Thus, in order to verify the derived clinker SR linear program, the clinker analysis must be studied to ensure that only clinker with LSF very close to that used in the model is used for the verification purposes. The fact that any aims may not be obtained by plant personnel, is beyond the scope of this paper.

To ensure that clinker which is in range was used for the test, the clinker qualities were plotted graphically. Two periods were chosen based on their consistency (implying that there were no segregation problems) and because the clinker LSF's were very close to that chosen for the model testing, i.e. LSF = 95.

The hypothesis testing method for comparing population means, was used to verify whether or not these results come from the same population. The results two stockpiles, are presented in tables 3.1 and 3.2. It can be concluded from tables 3.1 and 3.2 that the model provides an accurate prediction of what will be achieved in practice subject, of course, to the assumptions made in paragraph 3.3.5.

Table 3.1. - Verification of Optimisation Program in Practice - Stockpile A.

	Derived clinker quality from the Optimisation model	Actual clinker quality achieved from the same limestone stockpile	Hypothesis testing
Period 1			
LSF	95.0	95.08	-
SR	3.33	3.33	-
AR	1.61	1.62	-
Ho	-	-	$\mu_1 = \mu_2$
H1	-	-	$\mu_1 \neq \mu_2$
Z	-	-	0.00
Z critical	-	-	2.33
Conclusion	-	-	$\mu_1 = \mu_2$

Table 3.2. - Verification of Optimisation Program in Practice - Stockpile B.

	Derived clinker quality from the Optimisation model	Actual clinker quality achieved from the same limestone stockpile	Hypothesis testing
Period 2			
LSF	95.0	95.02	-
SR	3.31	3.31	-
AR	1.64	1.63	-
Ho	-	-	$\mu_1 = \mu_2$
H1	-	-	$\mu_1 \neq \mu_2$
Z:	-	-	0.00
Z critical	-	-	2.33
Conclusion	-	-	$\mu_1 = \mu_2$

3.4.2. Testing and verification of formulae for derived clinker SR

The regression analysis were tested in three steps. In each case a statistical verification was done and the specific requirements as proposed by the management of the case study plant, have been presented and analysed.

Step 1 : Using the results from the regression table in excel, the coefficient of variance was used to test for dependency and the F - statistic was used to establish causality, i.e. whether or not the relationship is due to chance. The results are as follows :

Table 3.3. - Verification of Formulae for derived clinker SR - Results of Linear regression.

LSF	Statistical verification			Requirements in practice			
	r2	F	Mean	Minimum		Maximum	
				Model	Accepted	Model	Accepted
< 110	0.971	252.52	3.81	3.76	3.71	3.85	3.91
110 - 119.99	0.928	192.78	3.58	3.55	3.48	3.62	3.68
120 - 129.99	0.947	333.49	3.28	3.25	3.18	3.30	3.38
130 - 139.99	0.922	165.02	3.08	3.06	2.98	3.11	3.18
> 140	0.928	134.08	2.85	2.82	2.75	2.88	2.95

Fcritical = 26.5 in all cases.

From table 3.3, it is evident that the model provides acceptable results both statistically and from a practice point of view.

Step 2 : The formula for each LSF range, was applied to limestone samples with LSF's spanning the specific range, which were not used in the formulation of the relationship. The test for two population means was applied to confirm whether or not both sets come from one population. the results are as follows :

Table 3.4. - Testing the model on independent data

LSF	Z	Z _{critical}	Conclusion
< 110	0.42	2.33	Accept
110 - 119.99	1.97	2.33	Accept
120 - 129.99	1.28	2.33	Accept
130 - 139.99	1.40	2.33	Accept
> 140	0.17	2.33	Accept

Thus, the model provides acceptable results on independent data.

Step 3 : The third step is to test the formula in practice. The same procedure was followed as for the testing of the linear program, with the same results. Thus, the same results as per tables 3.1. and 3.2. were obtained.

3.5. Conclusion

The model developed to calculate clinker SR from limestone analysis appears to be suitable, both statistically and from a practice point of view. The practice point of view is an important factor as the limitations placed on the model by these requirements are more stringent than the statistical requirements. The reader must be cautioned at this point, again, of the numerous assumptions which had to be made in paragraph 3.3.3. The model has been tested and verified and found acceptable under these assumptions only.

The verification was done using data from the case study. It is imperative that the reader understands that the outputs of the model as tested cannot be used in any other situation. The model must be run and used to create the applicable formulae for each case. If it is found that the model does not provide acceptable results in another situation, then it can be said that one or more of the assumptions made in paragraph 3.3.3, are not being complied with in that situation.

A detailed conclusion on the contents of this chapter and those before it, will be provided in the next, and last, chapter of this paper. Recommendations will be made on how the author believes the model should be used and managed.

CHAPTER 4

CONCLUSION AND RECOMMENDATIONS

4.1. Conclusions

The purpose of this study has been to investigate a management tool, for a cement clinker manufacturing facility, to predict clinker silica ratio based only on limestone quality. The need for the tool was explained by means of a case study. This same case study was used for the investigation, refinement and verification of the tool.

The end product of this investigation, is an extension, and slight modification of an existing tool, developed in 1992 for the same purpose, by Dr Stephan Olivier. With the knowledge obtained through studying literature on cement chemistry and mathematical modelling extensively, it was possible to identify a relationship between the limestone quality parameters and the clinker silica ratio and to test this for statistical correctness. It was also possible to use this knowledge to set up a complete model, capable of :

- A. Calculating the resultant clinker silica ratio for a specified mix of raw materials and fuel, using linear programming techniques directly.
- B. To calculate the coefficients in the equations used to predict the clinker silica ratio from only the limestone quality by :
 - 1. generating clinker quality data for as many limestone analysis as the user likes, on which :
 - 2. a multiple linear regression is then conducted, providing the coefficients.

Based on the knowledge obtained from the literature study it can be concluded that :

- A. The model produces results which are both acceptable statistically and according to the requirements in practice.
- B. The model can only be used in the cement industry, and with a maximum of four clinker raw materials.
- C. The model is subject to assumptions, which the user must understand and check for before using the model.

This investigation has provided the author with a sound background on the subject, based on practice and theory. It is thus appropriate that a few recommendations are made on how and when the model should be used, for maximum benefit.

4.2. Recommendations

4.2.1. Use the Linear Program directly if practical

The linear program result will always be more accurate than the result obtained using the formulae. (Remember that the linear program is used to generate the data which is used to calculate the coefficients for the formulae.). This should be possible and practical where the derived clinker silica ratio is being repeatedly calculated. The advantages in doing this are:

1. accuracy, and

2. the ability to immediately see that plant (raw mill feeders) limitations are adhered to, and get a result which takes this into account, directly.

4.2.2. Determine the frequency with which the formulae (or raw material database) should be updated :

Depending on the consistency of the additive raw materials and fuel qualities, the formulae, and thus the raw material database, will have to be updated. Establish which frequency is necessary and practical to do this.

It is recommended that the raw material data base is linked to the plant's laboratory monthly reporting system, if possible, and that the formulae are updated once a month using the latest quality and consumption data. Alternatively, update the database manually. If this is not done and the qualities change, the model will still indicate that the results are statistically acceptable, but these will in fact be completely inaccurate.

Keep track of the clinker silica ratios which are realised and check that they remain in line with the predictions. If these start going out of line, use this as a warning to update the database with the newest raw material and fuel quality data. Also update the coal ash content and the heat consumption.

4.2.3. Use the model for prospecting purposes

Use the model to generate formulae for the plant or a technical services department which is investigating the suitability of limestone deposits. It may be somewhat simpler for these people to use a set of formulae, within in a spreadsheet than to run the linear program repeatedly. Alternatively, use the optimisation program directly.

REFERENCES

1. GIORDANI, F.R. & WEIR, M.R. 1985. *A first course in mathematical modelling*. Monterey, California : Brookes/Cole Publishing Co.
2. VANDERBEI, R.J. 1996. *Linear programming, Foundations and extensions*. Kluwer Academic Publishers.
3. GRANT, E.L. & LEAVENWORTH, R.S. 1996. *Statistic Quality Control*. McGraw-Hill.
4. Holderbank Management and Consulting, Cement Course Notes, Materials Handling volume 1. 1996.
5. Standard Handbook of Engineering Calculations 1995. Edited by TG Hicks, et al., New York : McGraw-Hill.
6. UNDERHILL, L.G. 1987. *Introstat*. Juta & Co.
7. THIE, P.R. 1988. *An introduction to Linear programming and Game theory*. John Wiley & sons.
8. WILLIAMS, H.P. 1978. *Model building in mathematical programming*. John Wiley % sons.

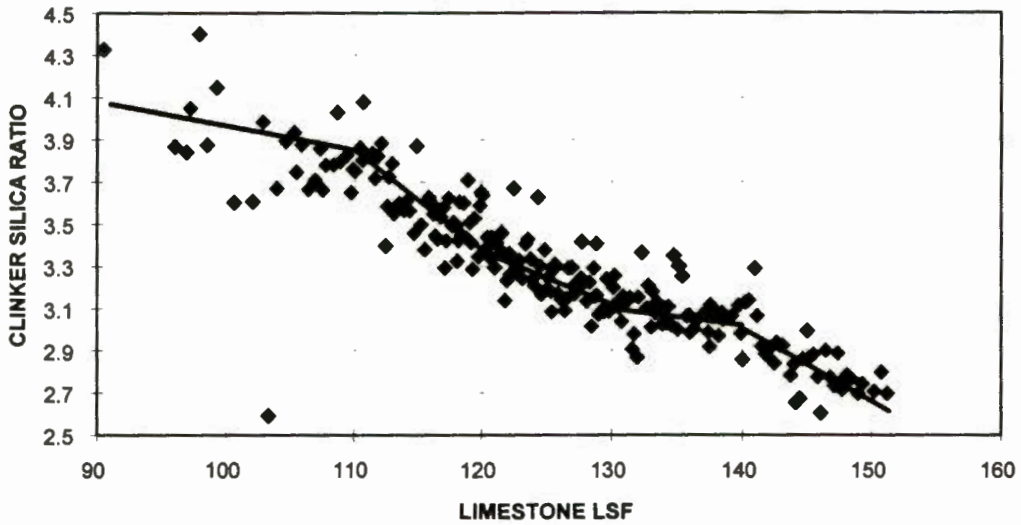
9. REDELINGHUIS, A; JULYAN, F.W; STEYN, B.L; BENADE, F.J.C. 1989. *Kwantitatiewe metodes vir bestuursbesluitneming*. Durban : Butterworth.
10. KLEINBAUM, D.G; KUPPER, L.L; MULLER, K.E & NIZAM, A. 1998. *Applied regression analysis and other multivariable methods*, Duxbury Press.
11. LIND, A.D & MASON, R.D. 1994. *Basic statistics for business and economics*. Irwin.
12. ATKINSON, A.C. 1987. *Plots, transformations and regression : an introduction to graphical methods of diagnostic regression analysis*. Oxford : Oxfordshire Clarendon Press.
13. DOWNIE, N.M. & HEATH, R.W. 1974. *Basic statistics*. New York , N.Y. : Harper & Row.
14. STEUER, R.E. 1985. *Multiple regression criteria optimisation : theory, computation and application*. New York, N.Y. : Wiley.
15. PHILIP, E.G; MURRAY, W & WRIGHT, M.H. 1991. *Numerical linear algebra and optimisation, Volume 1*. Redwood City, California : Addison-Wesley.
16. SIERKSMA, G, et al. 1996. *Linear and Integer programming : theory and practice*. New York : Dekker.

17. RAO, R.C. & TOUTENBURG, H. 1995. *Linear models : least squares and alternatives*. New York : Springer.
18. STAPLETON, J.H. 1995. *Linear statistical models*. New York : Wiley.
19. PADBERG, M.W. 1995. *Linear optimisation and extensions*. Berlin : Springer.
20. PANIK, M.J. 1996. *Linear programming : mathematics, theory and algorithms*. Dordrecht : Kluwer Academic.
21. NETER, J., et al. 1996. *Applied linear statistical models*. Chicago : Irwin.
22. KOUL, H.L. *Weighted empirical and linear models*. Hayward, California : Institute of Mathematical Statistics.

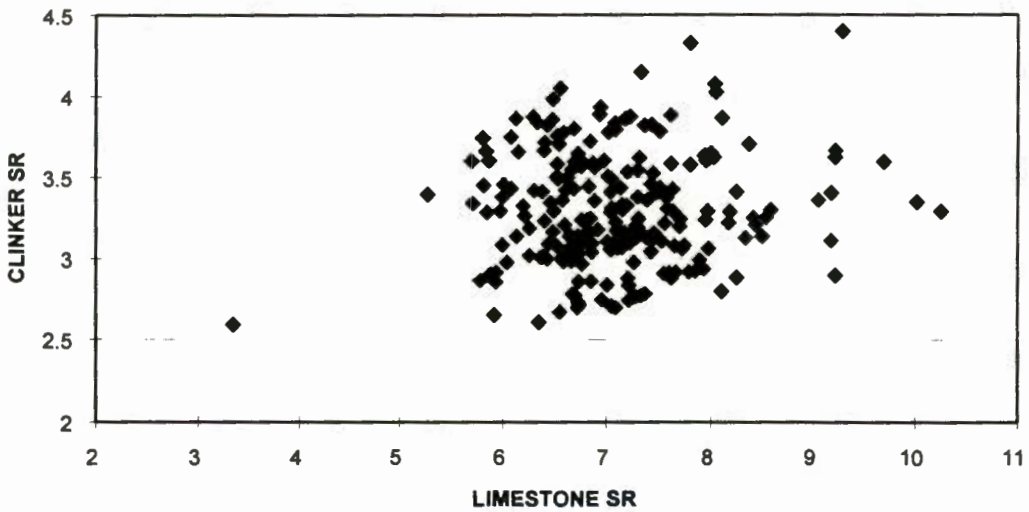
APPENDIX 1

SYSTEMS ANALYSIS - SCATTER DIAGRAMS

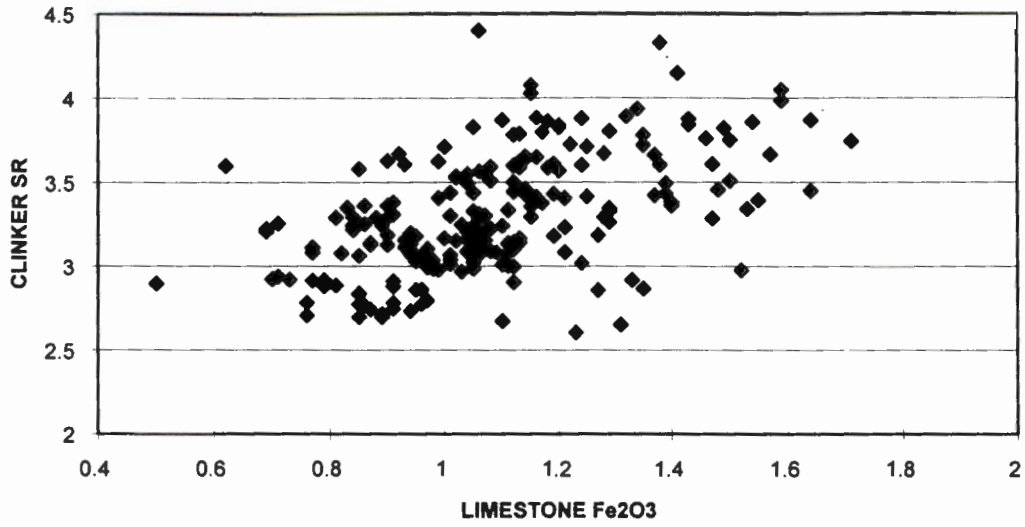
CLINKER SILICA RATIOS VERSUS LIMESTONE LSF



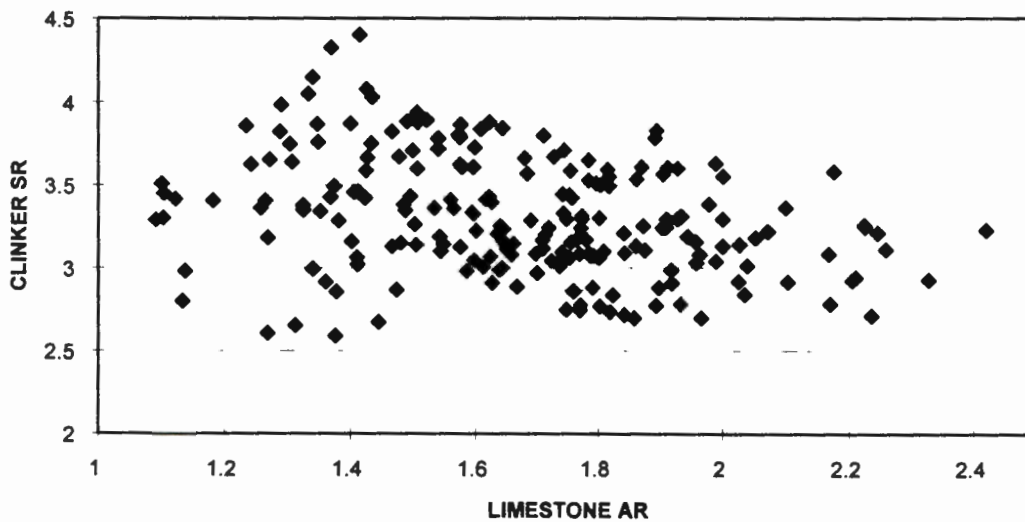
CLINKER SILICA RATIO VERSUS LIMESTONE SILICA RATIO

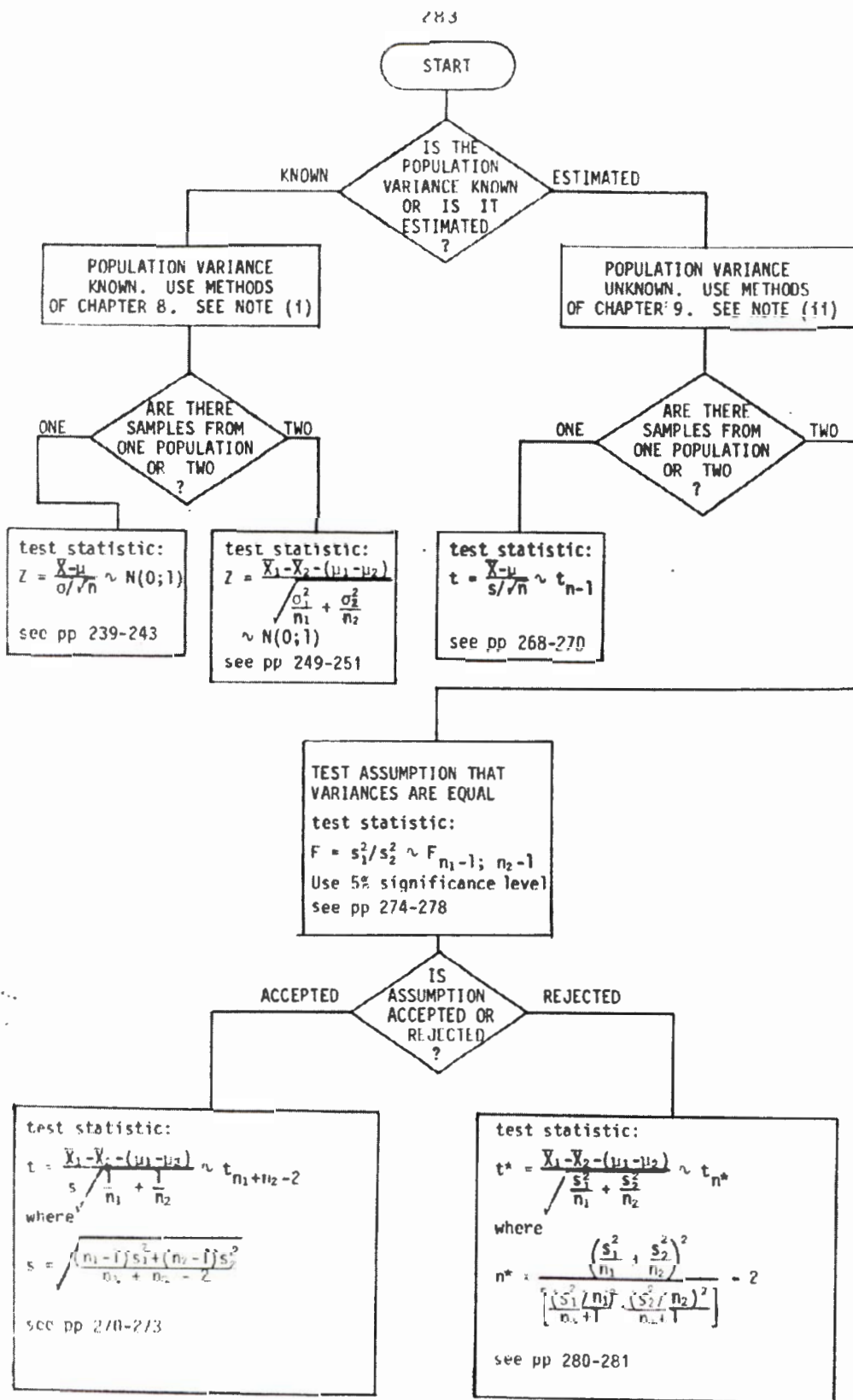


CLINKER SILICA RATIO VERSUS LIMESTONE Fe₂O₃ CONTENT



CLINKER SILICA RATIO VERSUS LIMESTONE ALUMINA RATIO





APPENDIX 3 - INPUT AND OUTPUT SCREENS

CLINKER SILICA RATIO MODEL - INPUT SCREEN

CLINKER SILICA RATIO MODEL - KILN ??????						CLINKER QUALITY		
MIX CALCULATION AND LINEAR PROGRAM						DESIRED	CALC	
MIX NUMBER :	1205					Heat Consumptions :	3.8900	MJ/t
					Coal LCV :	25.41	MJ/kg	
					Coal Ash content	16.73%		
					Coal Ash in clinker	2.56%		
LOSS FREE	LSTONE	ADD 1	ADD 2	ADD 3	ASH			
LOI	38.00	15.77	-2.84	3.86	0.00			
SiO2	20.20	42.93	4.67	49.79	40.62	22.81	22.25	
Al2O3	1.77	24.30	4.03	20.04	32.11	4.43	3.91	
Fe2O3	1.02	8.86	71.33	7.80	7.51	2.57	2.71	
CaO	72.21	8.44	0.74	3.44	7.33	67.03	65.09	
MgO	1.38	1.02	0.43	2.49	1.31	2.00	1.39	
K2O	0.09	0.41	0.02	6.27	0.52	0.11	0.34	
Total	96.67	85.96	81.22	89.83	89.40	98.95	95.69	
H2O	8.00	4.59	6.80	3.21	-			
LSF	121.75	5.46	1.15	2.05	4.68	94.72	94.79	
SIR	7.24	1.29	0.06	1.79	1.03	3.26	3.36	
ALR	1.74	2.74	0.06	2.57	4.28	1.73	1.45	

DERIVED CLINKER SR MODEL - OUPUT SCREEN

Date Run : 9/27/98

FORMULA

$$\text{Clinker SR} = m_1(\text{LSF}_{\text{limestone}}) + m_2(1/\text{Fe}_2\text{O}_3_{\text{limestone}}) + m_3(1/\text{SR}_{\text{limestone}}) + c$$

INPUT SUMMARY

Kiln No :					27685
Heat Consumption :					3.8900
Coal CV :					25.41
Coal Ash :					16.73%
	ADD 1	ADD 2	ADD 3	ASH	
LOI	15.77	-2.84	3.86		0.00
SiO ₂	42.93	4.67	49.79		40.62
Al ₂ O ₃	24.30	4.03	20.04		32.11
Fe ₂ O ₃	8.86	71.33	7.80		7.51
CaO	8.44	0.74	3.44		7.33
MgO	1.02	0.43	2.49		1.31
K ₂ O	0.41	0.02	6.27		0.52
Total	85.96	81.22	89.83		89.40
H ₂ O	4.59	6.80	3.21		-

RESULTS

Limestone LSF < 110				Limestone LSF 110 - 119.99			
m ₁	-1.518	r ²	0.9705	m ₁	-3.664	r ²	0.9278
m ₂	-10.996	F	252.52	m ₂	-11.483	F	192.76
m ₃	-0.750			m ₃	-0.603		
c	7.646			c	10.018		

Limestone LSF 120 - 129.99				Limestone LSF 130 - 139.99			
m ₁	-3.038	r ²	0.9470	m ₁	-1.358	r ²	0.9218
m ₂	-9.541	F	333.49	m ₂	-7.489	F	165.02
m ₃	-0.341			m ₃	-0.182		
c	8.747			c	6.153		

Limestone LSF > 140			
m ₁	-2.111	r ²	0.9284
m ₂	-8.505	F	134.08
m ₃	-0.152		
c	7.248		

APPENDIX 4

GLOSSARY OF TERMS

Limestone : A natural mineral consisting primarily of Calcium carbonate (CaCO_3). It is the main raw material used in cement clinker manufacture.

Stockpile(d) : The limestone used in the clinkering process must be as homogeneous as possible. This is to ensure minimum variation in additive addition, which in turn will ensure that the clinker produced is consistent. Limestone is mined, crushed and then stockpiled, the stockpile acting as a buffer between the mining and clinker manufacturing processes. The methods used to build and then reclaim material from these stockpiles are specially designed to ensure that the quality differential is blended, i.e. that the average quality is reclaimed consistently rather than random, large fluctuations in quality. Thus, the stockpile is used to homogenise the limestone.

Cement Clinker : The primary raw material in Portland cement. It is manufactured by burning a mixture of primarily limestone and some additives, at a temperature of approximately 1400°C.

Aluminite : An artificial mineral produced during the clinkering process. It can be expressed chemically as : $3\text{CaO}.\text{Al}_2\text{O}_3$. This is

known, in simplified terms, as C3A, in the cement industry. This mineral reacts very quickly with water, causing the cement to set.

Gypsum : Gypsum is a natural mineral with Calcium sulphate as it's main constituent. This is added to Portland cement clinker when milling Portland cement, as a set retarder. When water is added to the cement, the water reacts with the gypsum to form a compound called ettringite. Only when the gypsum has been used up does the water start to react with the Aluminite in the clinker. Thus, the gypsum retards the setting of the cement which is necessary in order to allow the necessary form work to be done before the concrete sets. Gypsum is chemically designated as : $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$.

Rotary kiln : For the purposes of this paper : A long cylindrical oven, at an angle of $\pm 2^\circ$ from the horizontal, which rotates about its axis, at a speed of ± 100 rph (revs per hour). The lengths vary between 60m and 200m, depending on the design and production output. Raw meal is fed from one end, called the feed end, while the heat required is obtained from fuel which enters on the opposite end. The primary fuel used in South Africa is coal. However, in many other countries, oil is used.

Raw Meal : The mixture of limestone ($\pm 95\%$) and additives which has been milled to a fine powder. The powder is characterised by a 10% residue on a 90 micron screen.

Additives : materials which are added to the limestone when producing raw meal. This is to ensure that the correct clinker mineralogy is obtained. These additives are used for their alumina content, iron oxide, and in some cases, for their calcium carbonate content. The latter is a high quality limestone which is used as a “sweetener” in some processes.

Mix design : this is the process of selecting the appropriate combination of a given quality limestone and additives, used to produce raw meal, in order to produce a clinker of desired quality *in a specific kiln system*. The latter must be specified as the type and amount of fuel used must be taken into account.

Refractory materials : Refractory materials are used to protect the shell of the rotary kiln from corrosion, and are made from materials with low thermal conductivity. Temperatures in the kiln go up to $\pm 1600^{\circ}\text{C}$. Refractory bricks are used to line most of the kiln shell. Different quality bricks, made up from different raw materials and using different processes, are used in the various zones of the kiln. (The kiln is split up into hypothetical zones, based on the temperature gradient in the kiln.) The reason for this is that, as the temperature changes so do the properties of the kiln atmosphere and the raw meal / clinker material. The bricks are exposed to thermal shock, mechanical wear and chemical attack, depending on where they are placed. The refractory bricks are selected according to their resistance to the conditions in a specific zone of the kiln.

LSF : The Lime Saturation Factor (LSF) is a mathematical ratio of Calcium oxide and the oxides : Iron oxide, Silica dioxide and Alumina as follows :

$$\text{LSF} = \frac{100.\text{CaO}}{2.8\text{SiO}_2 + 1.18* \text{Al}_2\text{O}_3 + 0.65 \text{Fe}_2\text{O}_3}$$

SIR : The Silica Ratio (also known as the silica modulus), SR, is defined as follows :

$$\text{SR} = \frac{\text{SiO}_2}{\text{Al}_2\text{O}_3 + \text{Fe}_2\text{O}_3}$$

AR : The alumina ratio (also known as the alumina modulus), AR, is defined as follows :

$$\text{AR} = \frac{\text{Al}_2\text{O}_3}{\text{Fe}_2\text{O}_3}$$