DYNAMIC SIMULATION OF RED MUD WASHERS USED IN ALUMINUM INDUSTRIES

By

Alexander Ardila Labiosa
B.Eng Chemical

Submitted in total fulfilment of the requirements of the degree of Master in Chemical Engineering

SCHOOL OF CIVIL, ENVIRONMENTAL AND CHEMICAL ENGINEERING
RMIT UNIVERSITY
VICTORIA, 3000, AUSTRALIA
March, 2010

To my parents and my nieces

DECLARATION

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work, paid or unpaid, carried out by a third party is acknowledged; and, ethics procedures and guidelines have been followed.

Alexander Ardila Labiosa
08-09-2010

ACKNOWLEDGMENTS

This project was undertaken within the School of Civil, Environmental and Chemical Engineering at Royal Melbourne Institute of Technology (RMIT) in agreement with Rio Tinto Yarwum refinery (Gladstone, QLD).

I wish to thank Mr. Luis Alban who was my supervisor for his guidance, advice, encouragement and for giving me the opportunity to explore another field of engineering. I am grateful for his support and dedication in helping me in every possible manner. Also, I will like to point out the dedication of my supervisor towards the profession and the students who work under him.

I would like to acknowledge the assistance of Berni Bennett, because through her I could make all the arrangements for my stay in Gladstone during my visits.

I also express gratitude to all the people from Rio Tinto who collaborated with any part of this thesis:

- Dr. Terry Snow (Superintendent Process Control) for allowing me to conduct my research on Rio Tinto and for the opportunity to work in a very friendly environment.
- Dr. Rudy Tio (Specialist – Advanced Control) for the advice in any possible manner and because he was the key for the research subject, and for making me feel part of his team. Also, I thank him for the weekly safety meetings, because those meetings show how each decision has an impact on the company.
- Dr. Ainul Azzah Abd Aziz (Senior Research Scientist) for her guidance, direction and patience when I made mistakes during the experiments and for teaching me useful concepts about rheology.
- Dr. Lachlan Boddington (Research Engineer) for the supply of flocculant and samples from the washers.

I wish to thank all personnel in charge of coordination of safety inductions, orientations and use of laboratory facilities.

I want to thank all the personnel of the RMPC group at RMIT because, although I did not belong to the group, they made me feel like one of them.

To all the people of Royal Melbourne Institute of Technology for allowing an unknown foreigner, from an unknown neighbourhood to conduct his research in this University.

ABSTRACT

Clarifier-Thickener equipment is used in a wide range of continuous sedimentation and sludge thickening processes where solid particles from continuous inflow mixtures are separated from the liquid. In this operation, the concentration of solids increases due to settling, so that the formation of a thicker bed is inevitable with time. Under optimal operating conditions, it is always possible to obtain two discharges from these vessels: a highly concentrated suspension at the bottom (underflow), and a clarified liquid stream at the top of the equipment (overflow or effluent).

In the Bayer Process an insoluble sub-product is formed as a result of the digestion of the bauxite ore with caustic soda. This product is called “red mud” and it has to be continuously removed by settlers or thickeners/clarifiers.

This project proposes the simulation of a continuous thickener/clarifier in order to predict the concentration profile and the height of the mud level (process controlled variable) as alternative to current measurement system, that contains long delay discrete sampling time(15 minutes each measurement). The simulation also, can provide an option of creating a knowledge base for off-line control. The project essentially involves two methods of simulation, namely mathematical modelling and neural networks. The model based applies a kinematic model to approximate the process behaviour using both, equipment and suspension characteristics. On the other hand, due to the large amount of historical data, neural network is proposed for system identification.

The first method is based on the solution of a highly nonlinear model, based on kinematic modelling of sedimentation extended to flocculated suspensions. This approach uses a conservative finite difference scheme of the upwind type for solving an initial boundary value problem (IBVP). The successful simulation of the equipment and further validation of the mathematical model are then achieved once the properties of the suspensions have been determined. These properties are related to flux batch settling and solid stress functions, whose parameters can be obtained experimentally.

On-site testing of the characteristics of red mud was conducted at the Rio Tinto Yarwum Alumina Refinery in Queensland. The settling properties of the suspension were determined via batch settling. For measurement of the rheology properties, the vane technique was used employing a Haake VT 550 rheometer.

The results of the simulation showed that the concentration profile and height of the heavy mud level can be determined via a steady state model for a given underflow concentration. These results, however, were not in good agreement with measured data.

The second method of simulation involved the use of Rio Tinto Yarwun historical data to develop a neural model in order to obtain a relationship between process variables. This approach has the advantage that no mathematical model is needed.

With this method, historical data (continuous data) are obtained and analysed, and daily averages of the variables involved in the process are calculated. Different network architectures are tested according to the washer process. Ultimately, two networks were developed to describe washer dynamics.

KEYWORDS

Sedimentation, flocculated suspension, conservation laws, finite difference schemes, yield stress, settling function, red mud, neural networks, backpropagation, learning algorithms, network training, network weights.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

f_{bk}	:	Flux density function
t	:	Time
z	:	Height of the vessel (variable height)
L	:	Feed level
φ_C	:	Critical volumetric concentration
$A(.)$	:	Diffusive function
g	:	Acceleration of gravity
$q(.)$	:	Volume average velocity
C	:	Courant-Friedrichs-Lewy number
u	:	Constant (Equation (4.12))
w	:	Weight (intensity of interaction between neurons)
$f(.)$	:	Activation function (Equation (5.4))
$f_{bk}^{EO}(.)$	:	Engquist-Osher numerical flux
$E(.)$	:	Sum of the square error (Equation (5.12))
u_C	:	Sediment velocity
u_O	:	Settling velocity of a single particle
$h(.)$	:	Settling function (Equation (6.2))
n	:	Richardson-Zaki parameter (Equation (6.2))
C_1	:	Initial concentration (Equation (6.6))
V_1	:	Initial volume (Equation (6.6))
V_2	:	Target volume (Equation (6.6))
C_2	:	Target concentration (Equation (6.6))
K	:	Power law constant (Equation (6.9))
H	:	Vane height (Equation (6.11))
D_V	:	Vane diameter (Equation (6.11))
T_m	:	Maximum torque (Equation (6.11))
MM	:	Minmod limiter function
A	:	Transversal area of the vessel

Q_F : Volume feed rate

P_y : Compressive yield stress

Greek letters:

∇ : Gradient

φ_D : Discharge volumetric concentration

Δt : Time step

Δz : Spatial mesh width

$\Delta\rho$: Solid-liquid density

τ_y : Yield stress

γ : Shear rate

α : Parameter (Equation (4.11))

β : Parameter (Equation (4.11))

ε : Signal error (Equation (5.11))

θ : Threshold (neural network)

θ_L : Limiter

σ_e : Effective stress function

φ : Volumetric fraction of solids

φ_{MAX} : Maximum volumetric fraction of solids

SECTION A:
INTRODUCTION AND THEORY

Chapter I

INTRODUCTION

Sedimentation is a unit operation involving the separation of solid and liquid phases of a dilute suspension by the action of gravity. The aim is to obtain a concentrated suspension and a clear fluid.

This process is carried out in industries such as mining and waste-water treatment, in thickeners that produce a residue containing separate components due to addition of flocculants.

In the alumina production process (the Bayer process), thickeners are used in order to clarify the alumina-rich liquor from insoluble red mud residue. Additional thickeners are used in the washing process for caustic liquor recovery, resulting in a more concentrated residue.

During this process, a heavy mud level is maintained in the bottom of the washer in order to achieve design underflow densities.

This project describes two methodologies for simulation of settling vessels. Essentially, it identifies how a settling system would respond to changes in the input variables and the effect of these changes on control variables such as heavy mud level.

I.1. RESEARCH OBJECTIVES

This work involves the dynamic simulation of a washer vessel in which sedimentation phenomena are involved. It examines two simulation techniques: finite difference and neural networks.

The objectives of this work are:

- To review fundamental theories of sedimentation
- To identify mathematical methodologies to solve partial differential equations.

- To mathematically model the sedimentation process, extending it to flocculated suspensions.
- To design and conduct on-site experiments on-site to obtain red mud settling and rheological characteristics.
- To create a database with historical data obtained from Rio Tinto Yarwun for neural network application.
- To develop a neural network model able to simulate the dynamic behaviour of the red mud washing process.

I.2. THESIS OUTLINE

This thesis is oriented towards the understanding of the red mud washing process and the subsequent development of a tool to simulate this process. The simulation is focused on the prediction of the heavy mud level in the bottom of the vessel.

The thesis content is structured in four sections. Section A contains an introduction and purpose of the study, along with a background review of the alumina refinery process, including the tailings and washing process. Section B presents the mathematical model of the washing process and the theory underlying neural networks. Section C provides an analysis of the research methodology and the results of the simulation. Finally, in section D, a summary of results, conclusions and suggestions for future work are developed.

I.2.1. SECTION A: INTRODUCTION AND THEORY

Chapter I introduces the thesis, its objectives and an outline of the project.

In chapter II, wastes produced in the mining industry are described, together with the background and explanation of the Bayer process and the tailings produced by this process.

In chapter III, different theories and the phenomenon of sedimentation are presented.

I.2.2. SECTION B: SIMULATION PRINCIPLES

In chapter IV, the mathematical approach to the sedimentation process is stated, followed by the methodology applied in solving the settling model.

In chapter V, a second simulation approach is presented. This approach focuses on the development of a neural networks model for system identification. Some explanation follows regarding how this tool can be applied in the process industry.

I.2.3. SECTION C: MODELING AND SIMULATION RESULTS

Chapter VI contains an explanation of the workings of the simulation model, along with the methodology employed for red mud characterization.

Chapter VII presents the design of the neural network model and simulation.

I.2.4. SECTION D: SUMMARY OF RESULTS, CONCLUSIONS, AND FURTHER WORK

Finally, in chapter VIII an overall summary of results is provided, together with conclusions and recommendations for further work.

Chapter II

WASTES FROM MINERAL RESOURCES

Wastes industries based on mineral resources – e.g., mining, milling and metallurgy – can be characterized as follows:

Table 2.1: Wastes from mineral resources (Lottermoser, 2007; Bell & Donnelly, 2006)

Wastes	Mining Industry	Milling Industry	Metallurgical Industry
Solid	▪ Unwanted materials in the overburden (rocks) and gangue. ▪ Minimize the effects by revegetation and landscaping.	▪ Remaining host rock (tailings) after mineral remotion. ▪ Vegetative stabilization	▪ From relatively innocuous slag (blast furnaces) to unstable solids (washing of gaseous wastes and metallic products).
Liquid	▪ Acid mine drainage (coal mining industry). ▪ Limited chemical stabilization, revegetation and landscaping.	▪ Water decanted from tailings ponds (suspended solid, low concentration of Cyanide) ▪ Remotion of suspended solids to 20-30 mg/L	▪ From wash water to effluent from wet scrubbers.
Gaseous	▪ Insignificant production of dust. Production of methane (coal mines).	▪ Negligible pollution problems.	▪ Sulphur dioxide (copper and lead smelting). ▪ Conversion into sulphuric acid.

II.1. TAILINGS DISPOSAL

Tailings are the solid waste products of milling and possible subsequent processing of ore; however, in some mineral industries tailings have some other name. Notable examples are the “slimes” produced by the phosphate industry and the “red mud” of the aluminium industry (Lottermoser, 2007). Considerable differences exist among tailings from different industries but considerable similarities also exist. Tailings may range from acidic to basic, but all are transported to their final resting place as slurry. Drop boxes are used all over to keep grades below one percent to minimize pipe abrasion. Final water content ranges from approximately 15 percent for oil shale tailings to as much as 85 percent for the red mud of the aluminium industry, but all contain some water at equilibrium with gravity. Some contain a percentage of

material with respectable shear strength but most contain a percentage of material that is essentially liquid. All contain some chemically inert material but many contain sufficient pyrite to produce the equivalent of an acid mine drainage problem upon exposure to the atmosphere (Lottermoser, 2007).

The most characteristic difference among tailings is particle-size distribution (Lottermoser, 2007). This factor is most significant because it ultimately determines the method that must be used for safe containment if the tailings are to be discarded terrestrially. If the tailings contain a sufficient percentage of coarse materials (usually called “sands”), the tailings can be segregated during disposal and the coarse sands used as a confining embankment. The finer materials can be separated by gravity in a method called “upstream construction” or they can be segregated by centrifuging (“cycloning”) in the method called downstream construction (Lottermoser, 2007).

If the tailings do not contain sufficient sands to form a well-drained embankment of acceptable density, permeability, shear strength, and factor of safety, or if the sands are used to backfill underground openings (Wills, 1997), then the embankment must be constructed with imported materials, which may be either natural rock or mine waste rock.

II.2. THE BAYER PROCESS

II.2.1. BACKGROUND

In 1889, the Austrian chemist Karl Bayer patented the process to obtain alumina from bauxite ore using a solution of sodium hydroxide. Karl was the son of Friedrich Bayer, founder of the Bayer chemical and pharmaceutical company. The first industrial plants for alumina production based on Bayer process were installed in France and Ireland in the 1890s (Williams, 1975).

Previously, alumina was produced using the Le Chatelier process (1969). In this process, bauxite is mixed with sodium carbonate for further calcination in a kiln between 1000 -1100^{o}C, which is called a pyrogenic process. The reaction product is sodium aluminate, which is then leached out at 80^{o}C; an aluminium solution is obtained for further precipitation using carbon dioxide from the oven. These kinds of

installations were built in Europe and United States and provided great amounts of alumina. The Le Chatelier process lost competitiveness with the appearance of the Bayer process and practically disappeared in the 1940s. Since the 1960s, the entire alumina production worldwide has been obtained using the Bayer process.

II.2.2. THE PROCESS

As mentioned, the worldwide production of bauxite predominantly uses the Bayer process. The alumina produced is then processed by the Hall-Heroult electrolytic method to generate aluminium. From the Bayer process perspective, the chemical and mineral compositions of the bauxite are very important; in fact, the operation used will depend on these properties. Gibbsite bauxites, like surinam, trombets and worsley, are the most difficult to digest, followed by boehmitic bauxites (Pryor, 1965). The most difficult to digest are diasporic bauxites, which require higher caustic concentration, higher digestion temperatures (280-300^{o}C) and the use of lime for co-digestion for alumina extraction (Pryor, 1965). Fortunately, most of the commercial bauxites are the relatively easy to process gibbsite, boehmite or gibbsite-boehmite.

Generally, aluminium refineries are suited to treat a particular form of bauxite, so bauxite exchange between different refineries is uncommon. Another limiting aspect is the amount of solid residue (red mud) produced during the Bayer process (Pryor, 1965). If any refinery is designed to process high quality bauxite with a low “red mud factor”, the refinery should have a proper capacity to process that type of bauxite.

The Bayer process consists of bauxite washing and pulverization, followed by high pressure and temperature caustic digestion (sodium hydroxide) (Hudson, 1987). The liquor obtained contains a solution of sodium aluminate and non-dissolved solid residue from the bauxite, which contains iron, silica and titanium. This residue is called “red mud”. The clear solution of sodium aluminate is pumped into a tank called a “precipitator”. In this tank, fine particles of alumina are added to promote precipitation of alumina particles. The settling particles in the bottom of the tank are extracted and go through a rotary kiln at 1100^{o}C for water removal (Hudson, 1987). The product is a white powder which is pure alumina. The caustic soda is recovered and re-used in the process. This process, which is used to obtain pure alumina from

bauxite ore, has undergone few changes since the opening of the first refinery in 1893.

The Bayer process includes four stages (Hudson, 1987; Rio Tinto, 2008): digestion, liquor clarification, precipitation of the hydrate and calcination of the alumina (Figure (2.1)).

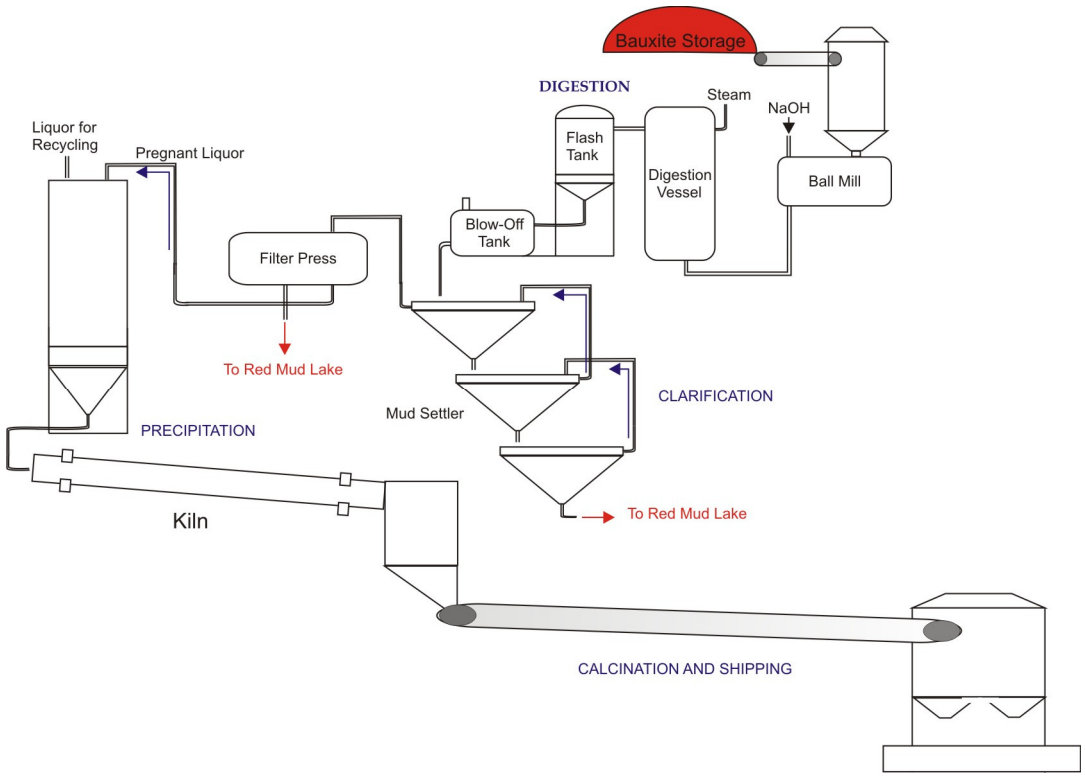


Figure 2.1: The Bayer Process (Hudson, 1987; Rio Tinto 2008)

II.2.2.1. DIGESTION

The process of digestion of the bauxite includes several stages (Hudson, 1987): grinding, desilication and digestion.

II.5.2.1.1. GRINDING

In this stage, fragments of the washed bauxite of particle size of 20 mm are ground inside of a mill (ball or rolling mill) to improve the liquid-solid contact during digestion. A recycle solution of caustic is added to produce a slurry capable of being pumped; lime may also be added for phosphate control (Hudson, 1987).

II.2.2.1.2. DESILICATION

The silica compounds in the bauxite are chemically attacked by caustic soda, causing alumina and soda losses by combination, to form a solid product called "desilication product (DSP)". In order to remove this silica from the slurry, it is heated and maintained at atmospheric pressure in a pre-treatment area. The greater part of the desilication products then become part of the residual slurry as aluminium-silica compounds. The DSP is not a simple compound, but is composed of a series of compounds of the zeolite type. The reaction of kaolinite during the digestion process shows the desilication process, according with (Wills, 1997):

$$Al_2O_3 \cdot 2SiO_2 \cdot 2H_2O + 6H_2O \rightarrow 2NaAlO_2 + 2Na_2SiO_3 + 5H_2O \quad (2.1)$$

$$2Na_2SiO_3 + 2NaAlO_2 + (2+x)H_2O \rightarrow Na_2O \cdot Al_2O_3 \cdot 2SiO_2 \cdot xH_2O + 4NaOH \quad (2.2)$$

II.2.2.1.3. DIGESTION

The slurry containing $Al_2O_3 \cdot xH_2O$ is pumped by high pressure pumps to digesters operating in series, which are equipped with stirring systems. Once mixed with steam and caustic solution, the alumina from the bauxite forms a sodium aluminate solution (Hudson, 1987). A number of non-dissolved impurities remain, primarily iron, titanium and silica compounds.

Modern refineries work with temperatures between 200 and 240°C and with pressures of approximately 30 atm. The conditions used in the digesters (concentration, temperature and pressure) may vary depending on the bauxite properties (Hudson, 1987). Theoretically, higher temperatures are more favourable for this process, but other disadvantages such as corrosion and dissolution of other oxides may be presented (Hudson, 1987).

Due to the high pressure and temperatures, the follow reactions are produced quickly (Hudson, 1987; Rio Tinto 2008):

$$2NaOH + Al_2O_3 \cdot 3H_2O \rightarrow 2NaAlO_2 + 4H_2O \quad (2.3)$$

$$2NaOH + Al_2O_3 \cdot H_2O \rightarrow 2NaAlO_2 + 2H_2O \quad (2.4)$$

Depending of the treatment times in the digester, recovery percentage as high as 97% of total alumina can be achieved.

After digestion, approximately 30% (by weight) of the bauxite remains in suspension, forming a reddish slurry containing red mud (a mixture of iron and titanium oxides and silica). The red mud is extracted from the digesters and cooled in a series of reactors with successively lower pressures. The generated heat is used to ore-heat the caustic liquor (Hudson, 1987).

II.2.2.2. CLARIFICATION

After the extraction stage, the liquor (which contains dissolved alumina) is separated from insoluble residue, and then is purified and filtered. The red mud is washed in order to extract as much caustic as possible and recycled (Hudson, 1987; Rio Tinto, 2008).

The clarification stage includes the following operations: sedimentation, washing and filtration.

The majority of the residual solids of the red mud are removed from the liquor by sedimentation using 40 meter diameter settling tanks (Hudson, 1987; Rio Tinto, 2008). Flocculants are added to improve settling velocity and to obtain a clearer supernatant (Hudson, 1987; Rio Tinto, 2008).

Washing process

Red mud is washed with water in a counter-current washer train to allow recovery of the caustic soda and any residual alumina that may have remained in the slurry, before pumping it to a storage area. The washing process involves several washers (between four and seven) operating in series (Figure (2.2)) (Hudson, 1987; Rio Tinto, 2008). The inputs at each stage involve the underflow of the previous thickener and the overflow of the next one (Rio Tinto, 2008). The outputs of the process include a concentrated product from the first washer (overflow) and a high solids concentration from the last thickener (underflow).

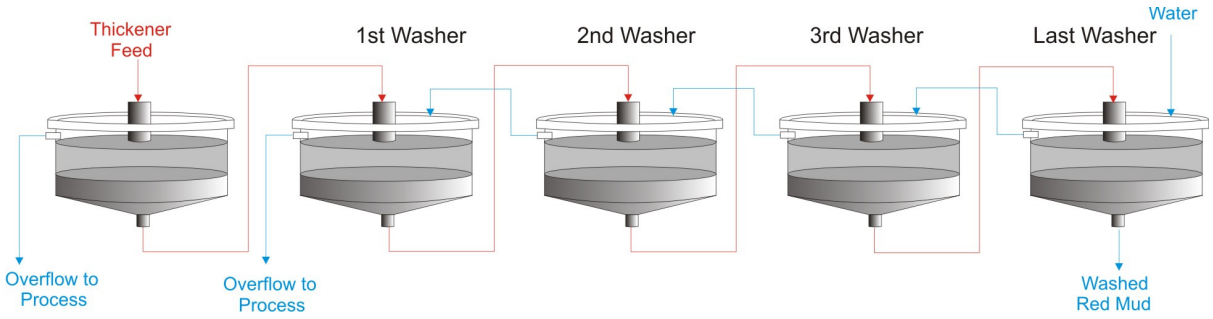


Figure 2.2: Washer train (Rio Tinto, 2008)

Lime is added to the caustic liquor to remove carbonates (Na_2CO_3), which are formed due to reaction with bauxite compounds and with other compounds from the air (Hudson, 1987). These carbonates reduce the effectiveness of the caustic liquor to dissolve the alumina. Lime regenerates the caustic soda, and the calcium carbonate that forms is removed with the red mud (Hudson, 1987):

$$Na_2CO_3 + Ca(OH)_2 \rightarrow CaCO_3 + 2NaOH \quad (2.5)$$

The supernatant preceding the sedimentation stage contains small particles of fine slurry, so it is filtered using constant pressure filters with polypropylene membranes (Hudson, 1987).

Once all solids have been removed, the liquor that comes from this area contains a supersaturated alumina solution, which is cooled (Hudson, 1987). The heat produced in this process is used to heat the liquor in the digestion stage.

II.2.2.3. PRECIPITATION OF THE HYDRATE

At this stage, the alumina is recovered from the liquor by crystal precipitation. The alumina precipitates as $Al_2O_3 \cdot 3H_2O$ due to the inverse reaction of the extraction process (Hudson, 1987), but in this case, the product can be controlled according to the operational condition of the refinery (temperature of precipitation, cooling rate, etc.) (Rio Tinto, 2008):

$$2NaAlO_2 + 4H_2O \rightarrow Al_2O_3 \cdot 3H_2O + 2NaOH \quad (2.6)$$

The alumina liquor is carried to precipitation tanks, and crystals of trihydrate alumina are added, generally with a fine particle size, to promote crystal growth (Hudson, 1987; Rio Tinto, 2008). The mixture is agitated inside the tanks for approximately 3 hours. During this process, several crystals of different sizes are formed. The inlet temperature, rate of alumina crystal addition and the caustic concentration are control variables used to regulate the particle size of the product. The particle size is an important parameter during the smelting process, so control of this variable is important (Hudson, 1987; Rio Tinto, 2008).

The mixture of crystals of different particle sizes is separated from the liquor and classified by size inside so-called gravity classification tanks (Hudson, 1987). The primary classifiers are used for coarse particles, which is the hydrate product. The crystals of intermediate and fine size from secondary and tertiary classifiers are washed and returned to the precipitation tanks to act as promoters for this operation (Hudson, 1987).

The solids-free caustic liquor (supernatant) from the tertiary classifier is recovered by evaporation and is concentrated, heated and recycled in order to dissolve the alumina in the digesters (Hudson, 1987). Finally, fresh caustic soda is added to compensate for losses during the process (Hudson, 1987).

II.2.2.4. CALCINATION

The washed hydrate is dried and heated to temperatures around 900-1200^{o}C (Hudson, 1987; Rio Tinto, 2008). This heating process, or calcination, is carried out in a rotary kiln. During this process, the alumina is obtained as follows (Hudson, 1987):

$$2Al(OH)_3 \rightarrow Al_2O_3 + 3H_2O \qquad (2.7)$$

II.3. RED MUD

The red mud is a residue from the Bayer process of alumina refining (Hudson, 1987). The amount of residue obtained (per ton of alumina produced) varies, according with the type of bauxite used, from 0.5 ton for high quality bauxites to 2.5 ton for low

quality bauxites. Worldwide, the annual production of red mud exceeds 60 millions of tons (Paradis, 1992).

The physical and chemical properties depend of the bauxite used, and how the ore is processed. In general, bauxite is mainly composed of aluminium oxide (monohydrate and trihydrate alumina, in different proportions). The main impurities are iron, silica and titanium oxide, as well as zinc, phosphorus, nickel or chrome, but in small traces (Lottermoser, 2007). The residue remaining after alumina production contains undissolved impurities, as well as any alumina not extracted during the process.

The Table (2.2) below shows the variability in the chemical composition that red mud may present:

Table 2.2: Variation in red mud chemical composition (IAI, 2002)

Compound	Amount (%)
Fe_2O_3	30-60
Al_2O_3	10-20
SiO_2	3-50
Na_2O	2-10
CaO	2-8
TiO_2	Trace-10

Worldwide, the aluminium industry has been under considerable pressure to minimize the environmental impact of its activities. However, it has been recognized that practical solutions should be applied. As a consequence, pond storage has been accepted by governments as “the best available technology”. The key focus of legislation is on preventing the release of the alkali, contained in the red mud, to the environment, as well as restoration of sites after mine closures. However, in the near future, policies like “taxes for waste” and “enduring responsibility”, which affect the treatment and final destination of the red mud, will be considered (Lottermoser, 2007).

In addition, while most of the minerals in red mud are not hazardous, special attention is needed due to its high alkalinity. The aluminium industry has recognized

the need to develop economically viable solutions in order to have a minimum impact on the environment by the wastes produced by the process and the sub-products obtained during the operation of the refinery (IAI, 2002). As a consequence, recently research has been conducted to obtain alternative treatments for neutralization of the alkali released by the red mud, for alkali recovery or for red mud neutralization (Menzies, Fulton, Morrell, 2004). Some of these treatments are shown in the table below:

Table 2.3: Possible treatments for red mud (Parekh & Goldberger, 1976; Whittaker *et al*., 1955)

<table>
<tr><th>Disposal</th><th>Neutralization</th></tr>
<tr><td>Sea disposal
Wet storage in ponds
Dry Storage</td><td rowspan="3">Mineral acids
Gases produced by storage
Biological activity
Sea water
Alkali recovery
Washing
Lime addition</td></tr>
<tr><th>Usage</th></tr>
<tr><td>Dye brick
Cement additive
Road construction
Iron process
Steal production</td></tr>
</table>

Nevertheless, in the industry, besides the use of sea water to decrease the alkalinity (Menzies *et al*, 2004), chemical processes are not generally used. Most effort is focused on improving the waste storage area (e.g. drainage control, stability studies of ponds, dry storage, etc) and recovery of the area after mine closure.

For these reasons, the high content of sodium, as a pollutant, is one of the most important issues concerning disposal of red mud. The sodium in red mud may be present in different forms such as sodium (Na) – an excess of alkali from the digestion stage – and sodium aluminium-silicates (sodalite), which are related structurally to zeolites (Williams, 1975).

The high concentration of sodium in red mud has several consequences as follows (Bell & Donnelly, 2006):

- Red mud drying: sodium (a small cation) attracts water molecules by ionic force. The sodium presence also increases water electric conductivity and inhibits the possibility of electric-drying of red mud.
- Handling and storage: sodium affects rheological properties of the red mud, which makes handling and storage more difficult.
- Superficial runoff: the alkalinity of the runoff (pH ~12) is due to the sodium present in the red mud, which needs to be controlled over large periods of time.
- Groundwater pollution
- Red mud usage: the sodium content in the red mud is one of the most serious problems when considering material reuse.

The alkalinity of the red mud causes many difficulties, not only because of the handling, but also because it represents caustic losses in the production process. The alkalinity, which is added by digestion liquor, may be recovered by separation and recycling of the liquor. Part of the alkalinity associated with solids of the red mud may be recovered by chemical interaction between solids and the liquor during storage.

The amount of caustic consumed during production of each ton of alumina is a fundamental issue from a production cost point of view. The majority of the caustic product goes to the residue. If those residues are use in agricultural activities, for example, a neutralization process is first required. This leads to additional expense as additional products are needed for neutralization, such as, mineral acids or exchange resins (Bell & Donnelly, 2006).

Chapter III

SEDIMENTATION

III.1. INTRODUCTION

The process of sedimentation can be defined as partial separation or concentration of suspended particles from a liquid by gravity settling (Perry, 1997). The process involves functional operations of thickening and clarification. The primary purpose of thickening is to increase the concentration of suspended solids in a feed stream, while that of clarification is to remove a relatively small quantity of suspended particles, thus producing a clear effluent (Foust, 1960). These two functions are similar and occur simultaneously. The name of the process depends on the desired result. Generally, a thickening operation is designed for the heavier-duty requirements imposed by a large quantity of relatively concentrated pulp, while clarifiers usually will include features that ensure essentially complete suspended solids removal, such as greater depth, special provision for coagulation or flocculation of the feed suspension, and greater overflow-weir length (Perry, 1976).

The problems of interest for concentration in processes such as mine tailings and ceramic suspensions are the selection, design and operation of the consolidation equipment. To achieve high concentrations, the equipment must compress the suspensions (Pryor, 1965).

Mine tailings are typically concentrated in large thickeners, as filtration or centrifugation is usually impractical due to the large treatment volumes. The most common thickener is the circular basin type (McCabe & Smith, 2005). The flocculant-treated feed slurry enters through the central feed well, which disperses the feed gently into the thickener. The feed suspension falls until it reaches a height where its density matches the density of the surrounding suspension and it spreads at that level. Solids concentration increases in the downward direction and this gives stability to the thickening process. The overflow is collected in a trough around the periphery of the basin. Raking mechanisms, slowly turning around the centre column, promote solids consolidation in the compression zone and aid solids discharge through the bottom central opening (Perry, 1976).

The settling operation is affected by several factors including size and shape of particles, size distribution, density differences (solid and liquid), slurry concentration and surface properties such as chemical content, chemical additives and the suspension medium (Perry, 1976). Because of these features, the solids experience particulate or hindered (aggregate moving) settling.

The ratio of water to solids in a slurry usually has a significant impact on the techniques and economics of the transport operations. Dilute slurries tend to behave in a fashion that is closer to that of Newtonian fluids, while concentrated slurries can exhibit strong non-Newtonian behaviour, which consequently affect the energy that is required to pump the material at the required rate. Generally speaking, there are usually advantages in reducing the amount of carrier fluid relative to the amount of solids to improve the energy efficiency as measured by the energy required to transport 1 kg of solids. Thus dewatering slurries must always be considered in practice (Pryor, 1965).

The natural tendency of the solids to settle under the influence of gravity is exploited to remove some of the water of the slurry. When the particles that make up the slurry are small, the settling is quite slow and special techniques are required to achieve a separation. Because of slow rates of settling that are commonly encountered, comparatively large equipment is required (Pryor, 1965).

III.2. THEORY OF SEDIMENTATION

The separation of a dilute slurry by gravity settling into a clear fluid and slurry of higher solids content is called sedimentation. The mechanism of sedimentation may be best described by observation of what occurs during a batch settling test as solids settle from slurry in a glass cylinder. Figure (3.1.a) shows newly prepared slurry of a uniform concentration of uniform solid particles through out the cylinder. As soon as the process starts, all particles begin to settle and are assumed to rapidly approach the terminal velocities under hindered-settling conditions. Several zones of concentration will be established (Figure (3.1.b)). Zone *D* of settled solids will predominantly include the heavier faster-settling particles. A poorly defined transition zone above the settled material contains channels through which fluid must rise. This

fluid is forced from zone *D* as it compresses. Zone *C* is a region of variable size distribution and non-uniform concentration. Zone *B* is a uniform-concentration zone, of approximately the same concentration and distribution as initially. At the top of region *B* is a boundary above which is clear liquid, region *A*. If the original slurry is closely sized with respect to smallest particles, the line between *A* and *B* will be sharp (Foust, 1960).

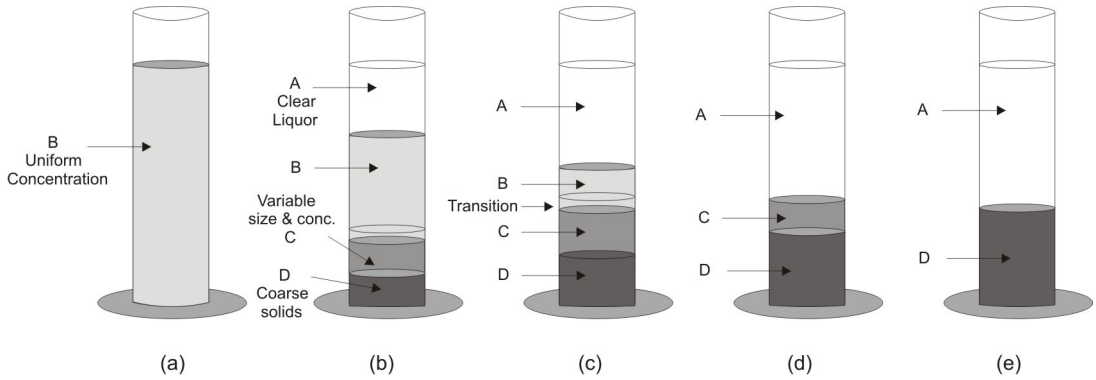


Figure 3.1: Batch sedimentation (Foust, 1960)

As sedimentation continues, the heights of each zone vary, as indicated in Figure (3.1.b,c,d). Note that both *A* and *D* grow larger at the expense of *B*. Eventually, a point is reached where *B* and *C* disappear and all the solids appear in *D*; this is referred to as the *critical settling point* (Figure (3.1.e)) – that is, the point at which a single distinct interface forms between clear liquid and sediment. The sedimentation process from this point on consists of a slow compression of the solids, with liquid from the boundary layer of each particle being forced upward through the solids into the clear zone. Settling rates are very slow in this dense slurry. The final phase is an extreme case of hindered settling (Foust, 1960).

In a batch settling test, the position of the zones varies with time. The same zones will be present in continuously operating equipment. However, once steady state has been reached (i.e., the slurry fed per unit time to the thickener is equal to the rate of sludge and clear liquor removal), the heights of each zone will be constant.

The zones are shown in Figure (3.2) for a continuous sedimentation (Foust, 1960).

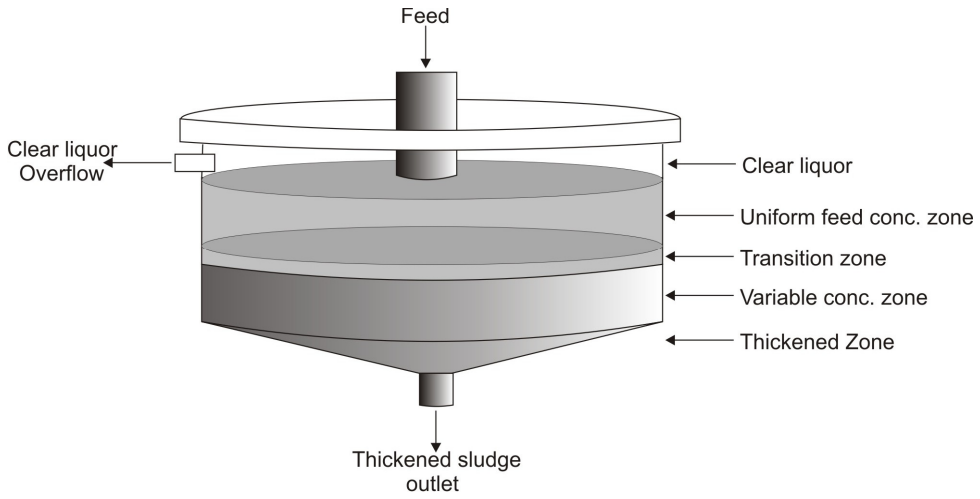


Figure 3.2: Settling zones in continuous thickeners (Foust, 1960)

The method used to obtain information from a simple batch settling curve is known as the Kynch construction. The Kynch construction is applied to the batch settling curve and this establishes the relationship between the rate of settling of a slurry and the local solid content (Foust, 1960).

III.2.1. THEORIES FOR THICKENER DESIGN

The main methods for thickener design can be represented by three categories based on macroscopic, kinematic and dynamic balance equations.

Macroscopic balances:

The first equation based on macroscopic balances was established by Mishler in 1912, who performed a simple mass balance on the equipment.

The method consists of experimental measurements of initial sedimentation velocity of a particular suspension using the feed concentration of a specific thickener. The area of the thickener is then calculated (Mishler, 1912).

The Coe and Clevenger (1916) design method determines the initial settling velocity of a suspension between feed and critical concentration, then solids handling rate is calculated. These scientists were the first to use the data obtained from a batch settling experiment using a laboratory column to design an industrial thickener. They assumed that solids settling velocity is a function of concentration only. However, in

the compression zone, this assumption does not hold because the solid volume fraction also varies with depth.

Kinematics models:

The Kynch theory of sedimentation (Kynch, 1952) allowed a new approach that was fast and reliable. A single batch test was developed to measure flux-concentration relationships and the settling behaviour could then be predicted. However, the theory assumes that wall effects and interactions between particles are not present. Therefore, settling velocity is only a function of local slurry concentration. The theory is also only applicable to the hindered settling zone, because it does not consider the compression zone. Both of these assumptions limit the accuracy if applied to the Bayer process.

Several research studies, such as those by Talmage and Fitch (1955), Wilhelm and Nadie (1979) and Oltmann, Hasset and Yoshioka (1957) have extended the Kynch method. Talmage and Fitch (1955), for example, proposed an extension to the Kynch theory. They used settling plots to provide data for thickener design. Data derived included solid flux, initial slurry height and concentration at the liquid-slurry interface. The slope of the height versus time plot provided the settling rate of the suspension. The slope at different times represented the settling velocity at different concentrations.

Fitch and Stevenson (1976) further developed the Talmage and Fitch method. They indentified the critical concentration, which could then be used to calculate the unit area (AU), that is, the thickener area needed to treat a certain daily flow rate of dry solids.

Dynamic models:

The main proponent of the dynamic methods is Adorjan (1975, 1976). However, this method is not often used in the mining industry. Other researchers also further developed these methods (Damasceno *et al*, 1992).

The main limitation of the dynamic methods is the experimental identification of the parameters of the suspension. Adorjan used compression cells for slurry compressibility determination. However, compression cells require a pressure that is significantly higher than currently present inside thickeners.

This case provides a more convenient measurement of different variables to calculate the effective stress of solids. For example, a concentration gradient can be obtained using batch or continuous test modelling, by gamma ray absorption, X ray absorption, ultrasound absorption while the pressure gradient can be measured using a manometer. Based on this information, effective solid stress can be calculated as well as suspension permeability (Huang, 1989).

III.2.2. KYNCH THEORY OF SEDIMENTATION

In 1952, the mathematician G.J. Kynch from Birmingham University (Great Britain) presented his article: “A theory of sedimentation”, a kinematic theory of sedimentation based on the propagation of concentration waves in a suspension. The suspension is considered to be continuous, i.e., an *ideal suspension,* and the sedimentation process is represented by a continuous equation of the solid phase:

$$\frac{\partial \varphi}{\partial t} + \frac{\partial f_{bk}}{\partial z} = 0 \quad ; \quad \begin{cases} 0 \le z \le L \\ t > 0 \end{cases} \tag{3.1}$$

where φ is the volumetric fraction of solids, t is time and z is the height of the interface, f_{bk} is the flux density function which satisfies: $f_{bk}(\varphi) = f_{bk}(\varphi_{\max}) = 0$ and $f_{bk}(\varphi) < 0$ for $0 < \varphi < \varphi_{\max}$.

The main limitation of an ideal suspension, as defined by the Kynch theory, is that the solid phase is neglected.

Kynch made the following assumptions (Kynch, 1952):

a. Differential settling due to differences in shape, size or composition of mineral particles does not take place,

b. Both solid and liquid are incompressible,
c. There is no mass transfer between components,
d. The sedimentation velocity is a function of concentration and tends to zero at a concentration equivalent to the sediment layer at the bottom of the container,
e. The concentration of particles in any horizontal plane is uniform,
f. The wall effects are negligible.

The ideal suspension concept is very useful for modelling purposes and is similar to the concept of the ideal gas used in thermodynamics. The theory of mixtures is used to obtain a mathematical model that predicts that an ideal suspension behaves similar to a glass sphere suspension (Shannon and Tory, 1996; Davies *et al.,* 1988).

In reality, the volumetric concentration of solids φ depends on three dimensions of the vessel, as well as space and time. However, Kynch assumes that the concentration is uniform, making this a one-dimensional phenomenon that can be simply described by $\varphi = \varphi(z,t)$ (Kynch, 1952).

The single batch test involves the suspension of a slurry in a transparent cylinder. The height of the clear liquor interface with the slurry is measured over time until the level falls to a minimum height; that is, until particles settle at the bottom as sludge. Where the sedimentation rate is very slow or the supernatant liquid remains turbid and unclear, flocculants are added (Foust, 1960).

SECTION B:
SIMULATION PRINCIPLES

Chapter IV

BASICS AND MATHEMATICAL APPROACH

IV.1. INTRODUCTION

Mathematical models are an essential element for process simulation. A chemical process is modelled with mathematical relations based on conservation and thermodynamic laws, as well as control and design restrictions. These models are represented by a system of algebraic and differential equations which describe specific equipment in a process.

Thus, a model can be defined as the mathematical representation of a system or phenomenon by application of fundamental principles. The principles of a mathematical model are physical and chemical laws (mass, energy and momentum conservation laws) and the transport equations.

In this chapter, a conservation law that describes the behaviour of the sedimentation process in industrial thickeners will be presented, as well as a numerical method that can be used to solve this model.

IV.2. CONTINUOUS THICKENER/CLARIFIER

A thickener can be defined as an industrial unit in which the concentration of a suspension is increased by sedimentation, with the formation of a clear liquid (Coulson and Richardson, 2002).

The operation of this type of equipment may be carried out batchwise or continuously (Foust, 1976). The equipment is typically a cylindrical tank with openings for a slurry feed and product draw-off. When steady-state is achieved, the solids are withdrawn continuously in the underflow at the rate they are supplied in the feed.

Normally, an inventory of pulp is maintained in order to achieve the desired concentration. This volume will vary somewhat as operating conditions change;

sometimes this inventory can be used for storage of solids when feed and underflow rates are reduced or temporarily suspended (McCabe & Smith, 2005).

Thickeners are usually constituted by some components: a tank to contain the slurry, feed piping and a feedwell to allow the feed stream to enter the tank, a rake mechanism to assist in moving the concentrated solids to the withdrawal points, an underflow solids-withdrawal system, and an overflow launder (McCabe & Smith, 2005). A schematic of a cylinder thickener that operates at steady-state is shown in Figure (4.1).

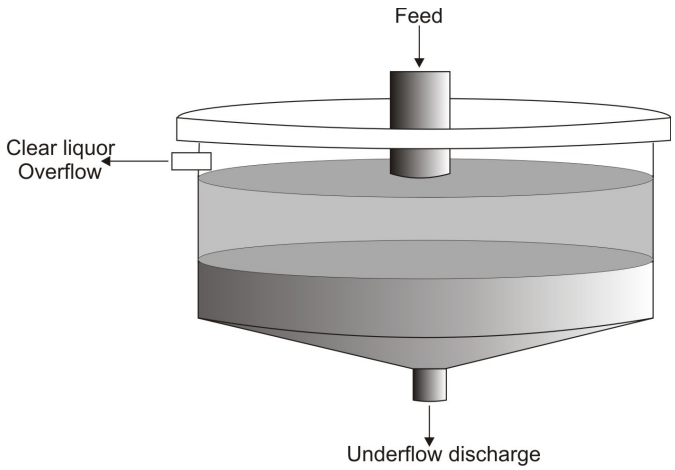


Figure 4.1: Schematic representation of an ideal thickener operating at steady-state (Foust, 1960)

The suspension is fed in at the centre, at depth of from 0.3 to 1 m below the surface of the liquid, with as little disturbance as possible. The thickened liquor is continuously removed through an outlet at the bottom, and any solids which are deposited on the floor of the tank may be directed towards the outlet by means of a slowly rotating rake mechanism that incorporates scrapers. The rakes are often hinged so that the arms fold up automatically if the torque exceeds a certain value; this prevents it from being damaged if it is overloaded. The raking action can increase the degree of thickening achieved in a thickener of given size. The clarified liquid is continuously removed from an overflow which runs round the whole of the upper edge of the tank. The solids are therefore moving continuously downwards, and then inwards towards the thickened liquor outlet; the liquid is moving upwards and radially outwards (Perry, 1976).

In order to maximize throughput, the rate of settling is increased by adding small amounts of an electrolyte, which causes precipitation of particles and the formation of flocs. This practice has resulted in thickener classification as either *conventional* or *high-rate* (Perry, 1976).

Conventional thickeners are characterized by having the feed-well at the top of the equipment and, once entering the thickener, the feed flow is mixed with part of the recovery liquor and diluted to a *conjugate concentration*. The diluted suspension settles at constant rate, forming a blanket of variable height, to become sediment in the bottom of the vessel.

On the other hand, high rate thickeners are characterized by a very deep feed-well that discharges the feed flow below the sediment. When the feed and sediment are mixed, it forms a suspension greater than the feed and greater or equal to the critical concentration. For this reason, no sediment zone is found in high rate thickeners. Part of the discharge flow is usually recycled to increase the feed concentration before being mixed with the sediment. Generally, these thickeners have greater capacity than do conventional ones. However, high rate thickeners, whose residence time is by the order of minutes, instead of hours for conventional, have been found to be unstable and, therefore difficult to operate and control (Perry, 1976).

IV.2.1. MATHEMATICAL MODEL

The description of the physical phenomenon of sedimentation can be represented by a strongly degenerate conservation law (Bürger, 1999). The compression effects can be described by a strongly degenerate diffusion term, while flux discontinuity is represented by a convection term of the partial differential equation (see Equation (4.1)).

Consider the case of a flocculated suspension in an ICT (*Ideal Continuous Thickener*), as shown in Figure (4.2). An ICT is a thickener in which the Kynch assumptions are present; for instance, where there are no wall effects and variables (concentration) depend only on the height of the vessel (z) and time (t). In $z=L$, feed enters to the vessel and in $z=0$ a discharge surface is presented which indicates a

continuous operation of the process. This model is appropriate to indicate, in a simple manner, the behaviour of the sedimentation process. The left side of Figure (4.2) shows the case of batch sedimentation which differs from the continuous process because of the closed vessel.

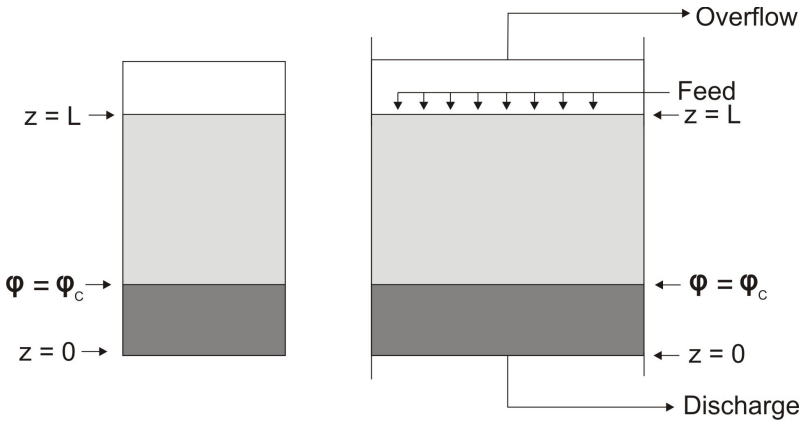


Figure 4.2: Left: Batch settling column. Right: ICT (Ideal Continuous Thickener) (Burger et al, 2000)

Concerning the one-dimensional case, the sedimentation theory presents equilibrium equations of mass and linear momentum which can be manipulated until a strongly degenerate parabolic equation is obtained (Bustos et al, 1999):

$$\frac{\partial \varphi}{\partial t} + \frac{\partial f(\varphi)}{dz} = \frac{\partial^2 A(\varphi)}{dz^2} \tag{4.1}$$

with $(z,t) \in]0,1[\times [0,T[$ and the integrated diffusive coefficient is defined as (Bürger, 1999):

$$A(\varphi) := \int_0^{\varphi} a(s) \cdot ds\,, \qquad a(\varphi) \geq 0 \tag{4.2}$$

In general, the diffusive coefficient is equal to zero (0) in intervals of φ. In such cases, Equation (4.1) becomes a hyperbolic equation (Bustos et al, 1999).

Solutions of (4.1) have discontinuities because of nonlinearities of the flux density function $f(\varphi)$ and because of the degenerate form of the diffusion coefficient. For this reason, entropic solutions are taken into account to obtain a well-posed problem. When Equation (4.1) is entirely hyperbolic, the values of the solution are propagated

through characteristics that may intercept the boundaries of the space-time domain from the interior; consequently, Dirichlet conditions should be treated as entropic conditions (Burger et al, 2000).

The constitutive equations for this process imply that $a(\varphi)$ has a degenerate behaviour, which means $a(\varphi)=0$ for $\varphi \le \varphi_C$ and $a(\varphi)>0$ for $\varphi=\varphi_C$, where φ_C is a constant called the *critical concentration*. Therefore, due to the degenerate nature of the diffusion coefficient $a(\varphi)$, Equation (4.1) is a partial differential equation of the hyperbolic-parabolic type (Burger et al, 2000).

The phenomenological problem of sedimentation-consolidation model can be represented as an initial-boundary value problem (IBVP) for a strongly degenerate hyperbolic-parabolic differential equation, as follows (Bustos et al, 1999):

$$\frac{\partial \varphi}{\partial t}+\frac{\partial(q(t)\cdot\varphi+f(\varphi))}{\partial z}=\frac{\partial^2 A(\varphi)}{\partial z^2}, \qquad (z,t)\in\ ]0,L[\times[0,T[\tag{4.2}$$

$$\varphi(z,0)=\varphi_0(z), \qquad z\in[0,L] \tag{4.3}$$

$$\varphi(L,t)=0, \qquad t\in\ ]0,T] \tag{4.4}$$

$$f(\varphi(0,t))-\frac{\partial A(\varphi(0,t))}{\partial z}=0, \qquad t\in\ ]0,T] \tag{4.5}$$

This problem can have an alternative representation (Bustos et al, 1999):

$$\frac{\partial \varphi}{\partial t}+\frac{\partial(q(t)\cdot\varphi+f(\varphi))}{\partial z}=\frac{\partial^2 A(\varphi)}{\partial z^2}, \qquad (z,t)\in\ ]0,L[\times[0,T[\tag{4.6}$$

$$\varphi(z,0)=\varphi_0(z), \qquad z\in[0,L] \tag{4.7}$$

$$q(t)\cdot\varphi(L,t)-\frac{\partial A(\varphi(L,t))}{\partial z}=\Psi(t), \qquad t\in\ ]0,T] \tag{4.8}$$

$$f(\varphi(0,t))-\frac{\partial A(\varphi(0,t))}{\partial z}=0, \qquad t\in\ ]0,T] \tag{4.9}$$

For both problems, *f* is continuous and differentiable for traces:

$f\le 0,\ sop(f)\subset[0,\varphi_{\max}],\ \|f'\|_\infty\le\infty,\ a(\varphi)\ge 0,\ sop(a)\subset sop(f),\ a(\varphi)=0$ for $\varphi\le\varphi_C$,

$0<\varphi_C<\varphi_{\max}$, $q(t)\le 0$, $\forall t\in[0,T],\ TV(q)<\infty$, $TV(q')<\infty$

In sedimentation-consolidation models for flocculated suspensions, the *z* value increases vertically; $\varphi = \varphi(z,t)$ represents the solids volumetric concentration, $q(t).\varphi \leq 0$ is the convective flux, $f(\varphi)$ is a function that relates the local relative velocity solid-liquid with local solids concentration which, in this model, is denoted by Kynch batch flux density, and:

$$a(u) = -\frac{f(\varphi)\cdot \sigma'_e(\varphi)}{\Delta\rho \cdot g \cdot \varphi} \tag{4.9}$$

where $\Delta\rho > 0$ denotes the difference of density between solid and liquid, g is the acceleration of gravity, and $\sigma'_e \geq 0$ is the derivative of the effective stress function (Bustos et al, 1999).

The effective stress function $\sigma_e(\varphi)$ can be represented by this behaviour (Bustos et al, 1999):

$$\sigma_e(\varphi) \begin{cases} = const., & if\ \varphi \leq \varphi_C \\ > 0, & if\ \varphi > \varphi_C \end{cases} \qquad \sigma'_e(\varphi) := \frac{d\sigma_e}{d\varphi} \begin{cases} = 0., & if\ \varphi \leq \varphi_C \\ > 0, & if\ \varphi > \varphi_C \end{cases} \tag{4.10}$$

Note that the degenerate property of the partial differential Equation (4.1) is due to this behaviour (Bürger, 1999).

The properties of the suspension are described by $f(\varphi)$ and $\sigma_e(\varphi)$. This requires that a suitable model be available for the settling velocity and effective solid stress, which can be obtained from a batch settling test and rheology test, respectively. The Richardson-Zaki model for the sedimentation velocity can be used to build a simple, but self-consistent, model for the ideal thickener. Models for effective solid stress can be represented by empirical equations that describe some compressible sediment. Burger *et al* (1999) have used the following equation:

$$\sigma_e(\varphi) = \begin{cases} \alpha \cdot e^{\beta\cdot\varphi}., & if\ \varphi \geq \varphi_C \\ 0, & if\ \varphi < \varphi_C \end{cases} \tag{4.11}$$

where φ_C is the critical concentration and α, β are empirical variables.

The conditions given by (4.3) and (4.7) correspond to an initial distribution of a given concentration; the conditions of (4.4) and (4.8) are a prescribed value of concentration in *z=L;* the conditions in (4.5) and (4.9) are the reduction of the flux density in the bottom of the vessel to a convective part $q(t).\varphi(0,t)$ and the condition (4.8) is a flux condition at *z=L* (Bustos et al, 1999).

IV.3. FINITE DIFFERENCE

The resolution of differential equations is a mathematical problem consisting of evaluation of any function that satisfies a differential equation. This can be achieved by transforming the differential equation or by applying a specific method of resolution (Toro, 1997). As an example of the transforming method, a numerical method like finite differences is used, which permits solving of partial differential equations by finding an approximation of the equation in finite spaces (Toro, 1997). Thus, the partial differential equation is transformed in an algebraic equation which can be solved using numerical methods. This process is termed *discretization* (Toro, 1997).

In order to achieve discretization, a specific point in space and time is chosen to develop the PDE (Toro, 1997). Due to this, a non-linear (in most cases) system of algebraic equations is obtained.

As an example, a linear advection equation is used for method description:

$$\frac{\partial q}{\partial t} + u\frac{\partial q}{\partial x} = 0 \qquad (4.12)$$

where *u* is a constant.

The special dimension, *x*, can be represented in a *net* or *point mesh* (Toro, 1997). It is easier to consider that the point distance is constant, which means that thee points are equally spaced with a distance of $h = \Delta x$ between adjacent points (Toro, 1997). For example, the interval $0 \le x \le X$ can be divided in i points equally spaced.

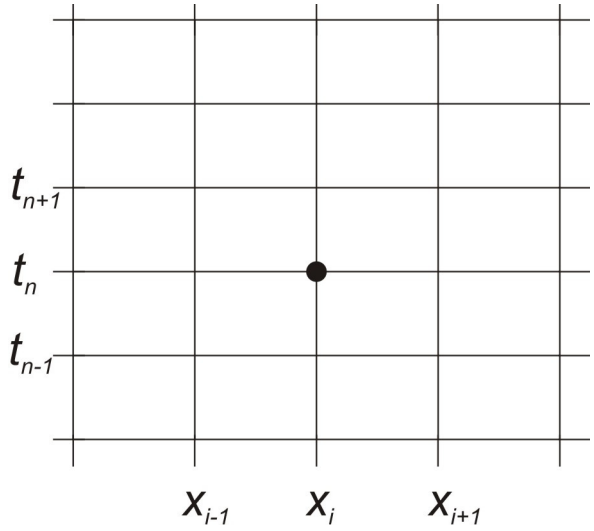


Figure 4.3: Discretization nomenclature in space and time (Toro, 1997)

The other independent variable, time *t*, also has to be discretized, and its step can be represented as $k = \Delta t$. In some cases, the time step is not uniform.

The values of the dependant variable, Q, in a spatial point with subscript *i*, and in a time step with superscript *n*, give a numerical solution for the partial differential equation at $x = x_i$ and at $t = t_n$ as (Toro, 1997):

$$Q_i^n \approx q(x_i, t_n) \tag{4.13}$$

and in the next time step, the solution takes the form:

$$Q_i^{n+1} \approx q(x_i, t_n + 1) \tag{4.14}$$

Using the finite difference method, the derivatives in the partial differential equation are discretized to obtain a system of algebraic relations among the values (approximations) in each point of the mesh.

In the example of the linear advection equation (Equation (4.12)) described above, the numerical solution Q_i^n in each point of the mesh i at time step *n* is used to advance to the next time step at $t = t_{n+1}$.

The solution in the point x_i at $t = t_{n+1}$ can be obtained developing a Taylor series around of Q_i^n (Toro, 1997):

$$Q_i^{n+1} = Q_i^n + \left(t_{n+1} - t_n\right) \cdot \left.\frac{\partial Q}{\partial t}\right|_n + O\left(\Delta t^2\right) \tag{4.15}$$

where $\Delta t = t_{n+1} - t_n$. Hence, the temporal derivative can be defined as a *forward difference* plus a *truncation error* (Toro, 1997), since the Taylor series is truncated after the first derivative:

$$\frac{\partial Q}{\partial t} = \frac{Q_i^{n+1} - Q_i^n}{\Delta t} + O(\Delta t) \tag{4.16}$$

In the case of the spatial derivative, there are several possibilities (Toro, 1997):

- Forward difference: the first derivative in the Taylor series is considered for Q_{i+1}^n around Q_i^n.

$$\frac{\partial Q}{\partial x} = \frac{Q_{i+1}^n - Q_i^n}{\Delta x} + O(\Delta x) \tag{4.17}$$

- Backward difference: first derivative in the Taylor series is considered for Q_{i-1}^n around Q_i^n.

$$\frac{\partial Q}{\partial x} = \frac{Q_i^n - Q_{i-1}^n}{\Delta x} + O(\Delta x) \tag{4.18}$$

- Central difference: the second derivative in the Taylor series is considered for Q_{i+1}^n and Q_{i-1}^n around Q_i^n, and then subtracted them:

$$\frac{\partial Q}{\partial x} = \frac{Q_{i+1}^n - Q_{i-1}^n}{2\Delta x} + O\left(\Delta x^2\right) \tag{4.19}$$

The terms $O(\Delta t)$, $O(\Delta x)$ and $O\left(\Delta x^2\right)$, represent discards terms during the development of the series and are known as *truncation errors* (Toro, 1997), i.e., the error introduced by the series truncation after first derivative (or second). The meaning of these terms is that the finite difference forms are accurate only at Δt,

Δx order. It can be seen that central differences form has a truncation error of Δx^2 order, while forward and backward difference forms have an order of Δx. This means that central differences are more accurate while $\Delta x \to 0$ (Toro, 1997).

The advection equation (4.12) now can be constructed using three different types of finite difference, based on the way of approximation of the spatial derivative. These forms of expression are known as *explicit* (Toro, 1997), because they use the actual values to calculate the values in the next time step.

Using forward difference, the advection equation acquires the form (Toro, 1997):

$$\frac{Q_i^{n+1} - Q_i^n}{\Delta t} + u \cdot \left(\frac{Q_{i+1}^n - Q_i^n}{\Delta x} \right) = 0 \qquad (4.20)$$

which can be rewritten as:

$$Q_i^{n+1} = Q_i^n - \frac{u\Delta t}{2\Delta x}\left(Q_{i+1}^n - Q_{i-1}^n\right) \qquad (4.21)$$

This last form of the equation is known as *forward time centred space* (FTCS) (Toro, 1997).

In theory, the forms of centred difference give more precise results due to the order of *O* $(\Delta t, \Delta x^2)$, while other schemes are O $(\Delta t, \Delta x)$ (Toro, 1997). Scheme stability and convergence are required for the correct solution. *Stable* means that the errors in the approximated solution decrease with time, i.e., errors do not increase in the next time steps (Toro, 1997).

IV.3.1. UPWIND DIFFERENCE

In order to represent a partial differential equation using upwind difference, the central difference approximation to spatial derivative $\partial Q / \partial x$ is replaced by a first-order one-side approximation (Toro, 1997). The upwind scheme to be used can be chosen among the following relations (Toro, 1997):

$$\frac{\partial Q}{\partial x} = \frac{Q_i^n - Q_{i-1}^n}{\Delta x} \tag{4.22}$$

$$\frac{\partial Q}{\partial x} = \frac{Q_{i+1}^n - Q_i^n}{\Delta x} \tag{4.23}$$

The correct choice of either (4.22) or (4.23) will depend of the *sign of the wave of propagation speed* u of the differential Equation (4.12) (Toro, 1997). In the case of a positive u, the scheme takes the form of:

$$Q_i^{n+1} = Q_i^n - c\left(Q_i^n - Q_{i-1}^n\right) \tag{4.24}$$

where *c* is:

$$c = u \cdot \frac{\Delta t}{\Delta x} \tag{4.25}$$

which is known as the *Courant number*, or the Courant-Friedrichs-Lewy number (CFL) (Toro, 1997). It can be regarded as the ratio of two speeds, namely the wave propagation speed u and the grid speed $\Delta t / \Delta x$ (Toro, 1997).

This scheme is stable under Von-Neumann criteria which can be denoted by the stability condition (Toro, 1997):

$$0 \leq c \leq 1 \tag{4.26}$$

This scheme in accordance with Equation (4.24) is called *first-order upwind* method (Toro, 1997), because the spatial differencing is performed using mesh point on the side from which information (wind) flows (Toro, 1997).

Chapter V

NEURAL NETWORKS

V.1. INTRODUCTION

Computational techniques and their applications vary from the computer industry to company production chains. To this end, multiple techniques related to artificial intelligence have been developed. The best known are fuzzy logic, genetic algorithms and neural networks.

In the 1950s, there were great expectations for research related to artificial intelligence and, above all, for research related to artificial neural networks (ANN). The studies, *Principles of neurodynamic and the perceptron: A probabilistic model for information storage and organization in the brain*, developed by Rosenblat (1958), gave a new perspective on the subject. However, these theories were challenged by Minsky and Papert (1969) in their work: *Perceptrons*, who showed that such networks cannot classify input patterns that are not linearly separable.

Twenty years later, the 1970s saw a renewed interest in these theories and their applications. Neural networks are currently employed in various fields (Deboeck, 2000):

- Financial and economical models
- Market profiling and clients
- Medical applications
- Knowledge management and “data discovery”
- Optimization and identification of industrial processes and quality control
- Scientific research.

The use of neural networks for system identification may provide relatively simple models from complex models. Neural networks can handle multiple parameters, such as a number of hidden layers, the number of neurons in each layer, number of inputs, outputs and some learning parameters. One of biggest difficulties, however, is the determination of the topological network configuration, that provides the best

approximation of the real system, since these criteria are not predetermined. This may be solved by trial and error (Haykin, 1999).

V.2. BIOLOGICAL NEURONS

The brain consists of a large number (approximately 10^{11}) of highly interconnected elements (approximately 10^4 connections per element), *called neurons*. These neurons have three main components, the *dendrites*, the *cell body* or *soma*, and the *axon* (Haykin, 1999). The dendrites, the receptors of the network, are the nerve fibres that carry electrical signals from the body of the cell. The cell body sums those input signals. The axon is a long fibre that carries the signal from the soma to other neurons. The contact point between an axon of a cell and a dendrite of another cell is called a synapse; the length of the synapse is determined by the complexity of the chemical process that stabilizes the function of the neural network (Haykin, 1999). A scheme of a biological neural network and the interconnection between its component is shown in the Figure (5.1):

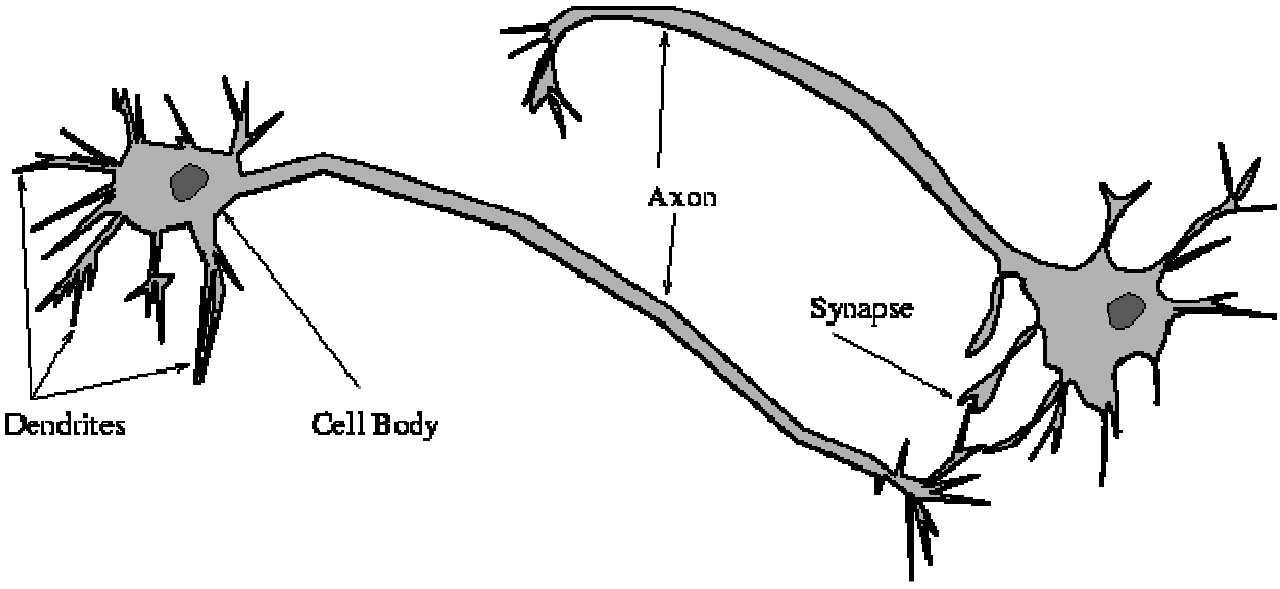


Figure 5.1: Biological neural network (Schmidt, 2000)

Some of the neural structures are determined at birth, another part is developed through learning, a process in which neural connections are made and others are completely lost (Schmidt, 2000). Neurological development is critical during the first years of life; for instance, it is shown that if a kitten is prevented from using one eye for a short period of time, it will never develop normal vision in that eye (Schmidt, 2000).

The neural structures change throughout life; these changes include the strengthening or weakening of synaptic connections. For example, new memories are believed to form by changing the intensity between synapses. Thus, the process of remembering a new friend's face is due to alteration of several synapses.

Based on the first studies on the neural basis of *mnemic systems* (related to memory), the storage of associative memory, both implicit and explicit, was believed to require a very complex neural circuit (Schmidt, 2000). Among those who began to disagree with this approach was Donald O. Hebb, a professor at Milner University. Hebb suggested that associative learning could be produced by a single cellular mechanism and indicated that partnerships could be formed by a coincident neuronal activity: "When an axon of a cell A excites cell B and participates in its activation, there is a developmental process or metabolic change in one or both cells, so that the effectiveness of A, as excitatory cell of B, intensifies (Haykin, 1999). According to the Hebbian learning rule, the fact that pre-synaptic neuron activity (providing the input pulse) matches with post-synaptic (receiving the pulse) is very important for strengthening of the connections between them; this mechanism is called *pre-post associative* (Schmidt, 2000).

All neurons convey information in a similar way: it travels along axons in brief electrical impulses, called action potentials (Schmidt, 2000). The action potentials, which reach a maximum amplitude of 100 mV and lasting $1\,\mu s$, are the result of displacement through the cell membrane of sodium ions endowed with positive charge, which pass from the extracellular fluid to the intracellular cytoplasm, as the extracellular concentration of sodium far exceeds the intracellular concentration (Schmidt, 2000).

The resting membrane maintains an electrical potential gradient of –70mV; the negative sign appears because the intracellular cytoplasm is negatively charged with respect to the outside (Schmidt, 2000). The sodium ions do not easily cross the resting membrane, and physical or chemical stimuli that reduce the potential gradient, or depolarize the membrane, increase its permeability to sodium. The flow of this ion increases the outward membrane depolarization, thus further increasing the permeability to sodium (Schmidt, 2000).

Although axons may seem similar to insulated conductor wire, electrical impulses are not propagated in the same way. An axon would not be very valuable as an electrical wire because its strength along the axis is too large and the membrane resistance is too low (Haykin, 1999); the injected positive charge to the axon during the action potential is dissipated one or two millimetres below. In order for the signals to travel several centimetres, it is often necessary to repeatedly regenerate the action potential along the way (Schmidt, 2000). The need to strengthen the electric is current limited to approximately 100 meters per second, the maximum travel speed of the impulses. This speed is less than a millionth of the speed of an electrical signal through a copper wire (Schmidt, 2000).

When an action potential reaches an axon terminal, transmitters housed in tiny vesicles are released into a gap slit of about 20 nanometres wide that separates the pre-synaptic of the post-synaptic membrane (Schmidt, 2000). During the climax of the action potential, calcium ions penetrate into the nerve terminal, this ion flow is the determining signal of the synchronized exocytosis; that is, the coordinated release of neurotransmitter molecules (Schmidt, 2000). As they are released, the neurotransmitters bind to postsynaptic receptors, invoking a change in membrane permeability (Schmidt, 2000).

When the load displacement at the membrane approaches the threshold for generating action potentials, an excitatory effect is produced and when the membrane is stabilized in the vicinity, an inhibitory effect is produced (Schmidt, 2000). Each synapse produces only a small effect; to determine the intensity

(frequency of action potentials) of each neuron's response is to seamlessly integrate synaptic signals until about 1000, adding in the soma or cell body (Schmidt, 2000).

In some neurons, impulses are initiated in the junction between the axon and soma, and then transmitted along the axon to other nerve cells. When the axon is close to its target cells, it is divided into many branches that form synapses with the soma or axons of other cells (Schmidt, 2000). Synapses can be excitatory or inhibitory according to the neurotransmitter released. Each neuron receives from 10000 to 100000 synapses and its axon makes a similar number of synapses (Schmidt, 2000).

V.3. ARTIFICIAL NEURAL NETWORKS (ANN)

According to Freeman and Skapura (1993), a neural network is a system of interconnected parallel processors in the form of a directed graph. Schematically, each processing element (neuron) of the network is represented as a node. These connections provide a hierarchical structure which emulates the physiology of the brain, seeking new processing models to solve specific problems in the real world. What is important in developing the technique of ANN is its ability to learn, recognize and apply relationships between objects and patterns of objects in the real world. In this regard, an ANN is a tool that may be used to solve difficult problems. The ability to solve difficult problems is feasible because of the principles of neural networks. The five most important, listed by Hilera and Martinez (1995), are as follow:

- *Adaptative learning:*

  This is perhaps the most important feature of neural networks; they can be trained using a series of illustrative examples. Thus, it is not necessary to model *a priori*, or to establish probabilistic functions. An artificial neural network can be considered as adaptive because it can be modified constantly in order to adapt to new working conditions.
- *Self-organization:*

  While learning is a process that modifies the internal information of the artificial neural network, self-organization is the modification of the entire network in order to carry out a specific goal.

Self-organization means generalization, so that a network can respond to data or situations not experienced before, but which can be inferred on the basis of their training. This feature is very useful when the input information is unclear or is incomplete.

- *Fault tolerance:*

  In traditional computing, the loss of a small piece of information can often lead to deactivation of the system. Artificial neural networks possess a high capacity for fault tolerance. This means that networks can recognize patterns of information with noise or distortion or that is incomplete, but also can keep working even if part of the network (with some degradation) is destroyed. The explanation for this phenomenon is that while a traditional computer stores information in unique spaces, localized and addressable, neural networks do it in a distributed way and with a high degree of redundancy.

- *Operating in real time:*

  Artificial neural networks, of all existing methods, are the best suited for pattern recognition in real time, because they work in parallel to update all instances simultaneously. It is important to point out that this feature is seen only when networks are implemented with hardware specifically designed for parallel processing.

- *Easy integration with existing technologies:*

  It is relative easy to obtain specialized chips for neural networks that improve their ability in certain tasks. This facilitates the modular integration into existing systems.

In elaborating on the principles of artificial neural network and continuously observing the term *neuron*, it is not surprising that the analogy with the brain is analysed. This may be because the artificial neural networks are based on biological inspiration. The human has about 10^{10} massively interconnected neurons; the neuron is a specialized cell that can propagate an electrochemical signal. As mentioned before, neurons have a branching

structure input (dendrites) and branching structure of outputs (axons). The axons of one cell are connected to the dendrites of another; by way of the synapse, neuron is activated and an electrochemical signal is relayed through the axon. This signal transfers from the synapses to other neurons, which in turn can excite. The neurons are excited only if the total signal received at the cell body through the dendrites exceeds a certain level (threshold of excitation).

Artificial neural networks try to mimic this principle of brain functioning.

V.3.1. NEURAL NETWORKS STRUCTURE

The three concepts for emulating nervous systems are: *parallel computing*, *distributed memory*, and *adaptability to the environment* (Haykin, 1999).

Parallel processing is essential in this type of task, in order to perform many calculations in a time interval as short as possible.

Another important concept that appears in the brain is the *distributed memory* (Haykin, 1999). While the information in a computer occupies well defined memory positions, in neuronal systems, this is distributed by network synapses. Therefore, if a synapse is damage, just a very small part of the information is lost (Haykin, 1999).

The last key concept is *adaptability* (Haykin, 1999). The artificial neural networks readily adapt to the environment by modifying their synapses, and learn from experience, being able to generalize concepts from particular cases (Haykin, 1999).

From the three previous properties, we conclude that in pursuit of an artificial neural system, a similar hierarchical structure can be established. The key starting point is the artificial neuron, which is organized in layers; several layers will form a neural network, and finally, a neural network (or set of these), along with input and output interfaces, plus the requirements of additional conventional modules, constitute the

overall system process. The hierarchical structure can be seen in Figure (5.2) (Haykin, 1999):

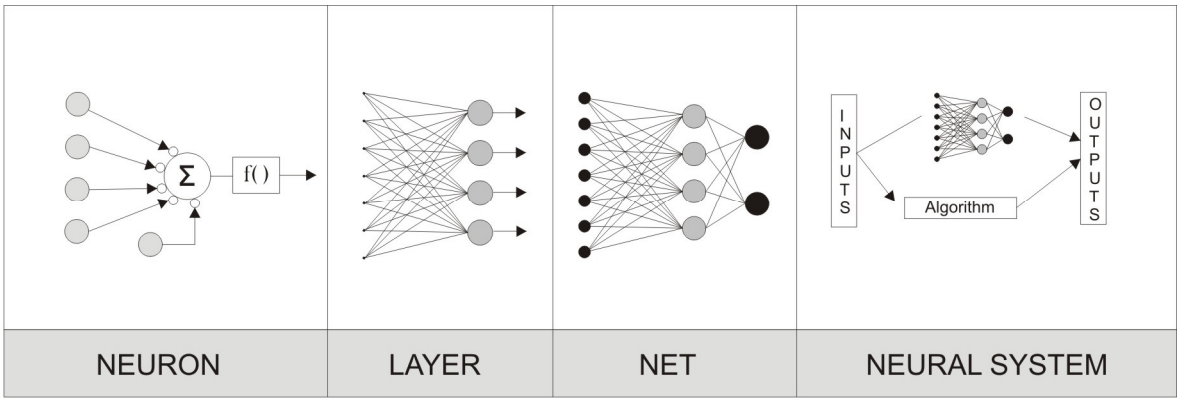


Figure 5.2: Hierarchical structure of an artificial neural system (Haykin, 1999)

V.3.2. ARTIFICIAL NEURON MODEL

A neuron is an elementary processor that, from an input vector from outside or from other neurons, provides a single response or output (Figure (5.3)).

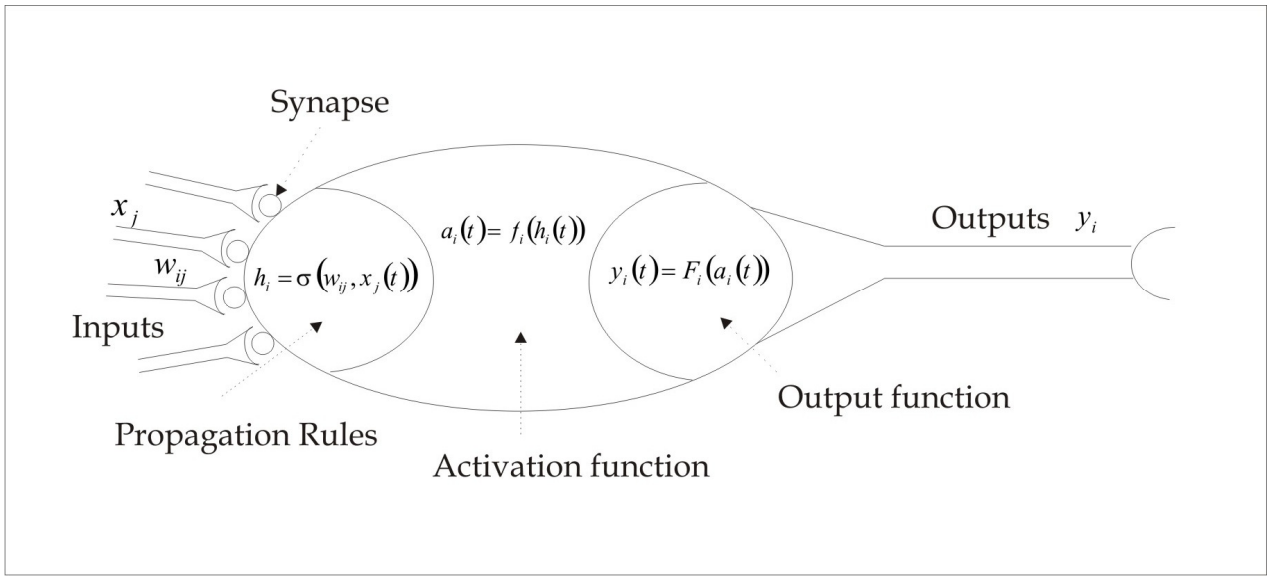


Figure 5.3: Single model of an artificial neuron (Beale, 1990)

The constitutive elements of a single neuron are (Engelbrecht, 2008):

- *Inputs:* $x_j(t)$

  The inputs and output variables can be binary (digital) or continuous (analog) depending on the type of application.

- *Weights:* w_{ij}

  Represent the intensity of interaction between each pre-synaptic neuron *j* and post-synaptic neuron *i*.

- *Propagation Rules:* $\sigma(w_{ij}, x_j(t))$

  Provide the value of the post-synaptic potential (PSP), $h_i(t)$ of the neuron i as a function of their weights and input:

$$h_i = \sigma(w_{ij}, x_j(t)) \tag{5.1}$$

  The most common function is linear, and is based on a weighted sum of the inputs and synaptic weights:

$$h_i = \sum_j w_{ij} \cdot x_j \tag{5.2}$$

  The synaptic weight w_{ij} is defined in this case as the intensity of interaction between the pre-synaptic neuron j and post-synaptic neuron i. Given a positive input, if the weight is positive, it tends to excite the postsynaptic neuron; if the weight is negative, it will tend to inhibit it. Thus, excitatory synapses (positive weight) and inhibitory synapses (negative weight) can be found.

- *Activation or transfer functions:* $f_i(a_i(t-1), h_i(t))$

  Provides the status of current activation, $a_i(t)$ of the neuron i as a function of its previous state, $a_i(t-1)$ and its current postsynaptic potential:

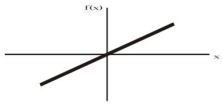
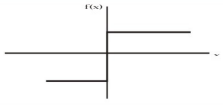
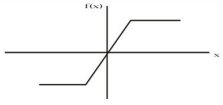
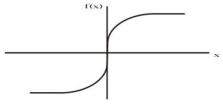
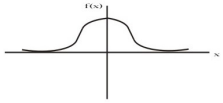
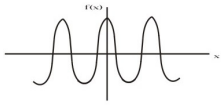
$$a_i(t) = f_i(a_i(t-1), h_i(t)) \tag{5.3}$$

  In many models of artificial neural networks (ANN), the current state of the neuron does not depend on its previous state but only the current state:

$$a_i(t) = f_i(h_i(t)) \tag{5.4}$$

The activation function $f(\cdot)$ is usually considered deterministic and, in most of the models, is monotone increasing and continuous. The form $y = f(x)$ of the most used activation functions in the ANN is shown in the Table (5.1), which represents the postsynaptic potential and the activation status.

Table 5.1: Activation functions (Haykin, 1999)

	FUNCTION	RANGE	GRAPH
Identity	$y = x$	$(-\infty, \infty)$	
Step	$y = sign(x)$ $y = H(x)$	$(-1,1)$ $(0,1)$	
Linear in Traces	$y = \begin{cases} -1 & if\ x < -l \\ x & if\ +l \le x \le -l \\ +1 & if\ x > +l \end{cases}$	$(-1,1)$	
Sigmoid	$y = \frac{1}{1+e^{-x}}$ $y = tgh(x)$	$(0,1)$ $(-1,1)$	
Gaussian	$y = A \cdot e^{-B \cdot x^2}$	$(0,1)$	
Sin	$y = A \cdot \sin(\omega \cdot x + \varphi)$	$(-1,1)$	

- *Outputs:* $F_i(a_i(t))$

  Provides current output, $y_i(t)$ of neuron i in terms of its current activation state, $a_i(t)$. Very often the output function is simply the identity $F(x) = x$, so that the activation state of the neuron is considered as the actual output:

$$y_i(t) = F_i(a_i(t)) = a_i(t) \tag{5.5}$$

Therefore, the operation of the neuron i can be expressed as follows:

$$y_i(t) = F_i\left(f_i\left(a_i(t-1), \sigma_i\left(w_{ij}, x_j(t)\right)\right)\right) \tag{5.6}$$

Figure 5.4 shows the interconnection between various neurons:

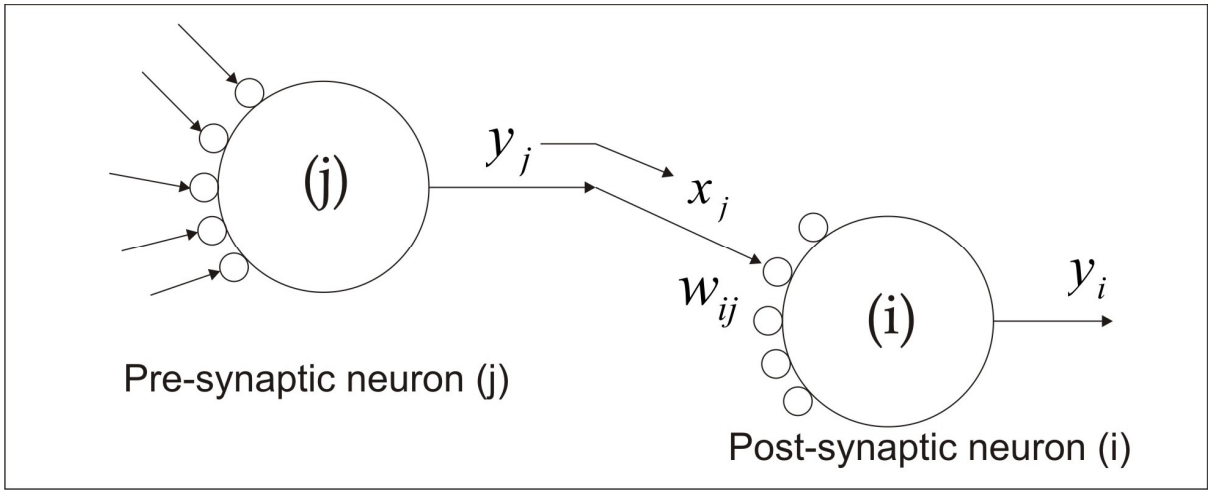


Figure 5.4: Interconnection between a pre-synaptic and a postsynaptic neuron (Haykin, 1999)

V.3.2.1. STANDARD MODEL OF AN ARTIFICIAL NEURON

Whereas the rule of propagation is the weighted sum and the output function is the identity function, the standard neuron is (Engelbrecht, 2008):

- *Inputs*: $x_j(t)$
- *Synaptic weights*: w_{ij} - related to the inputs
- *Propagation rules*: $h_i(t) = \sigma\left(w_{ij}, x_j(t)\right)$ - the most common is $h_i(t) = \sum w_{ij} \cdot x_j$
- *Activation function*: $y_i(t) = f_i(h_i(t))$ – this simultaneously represents the output of the neuron and its activation state.

Figure 5.5 shows a graphical representation of these elements:

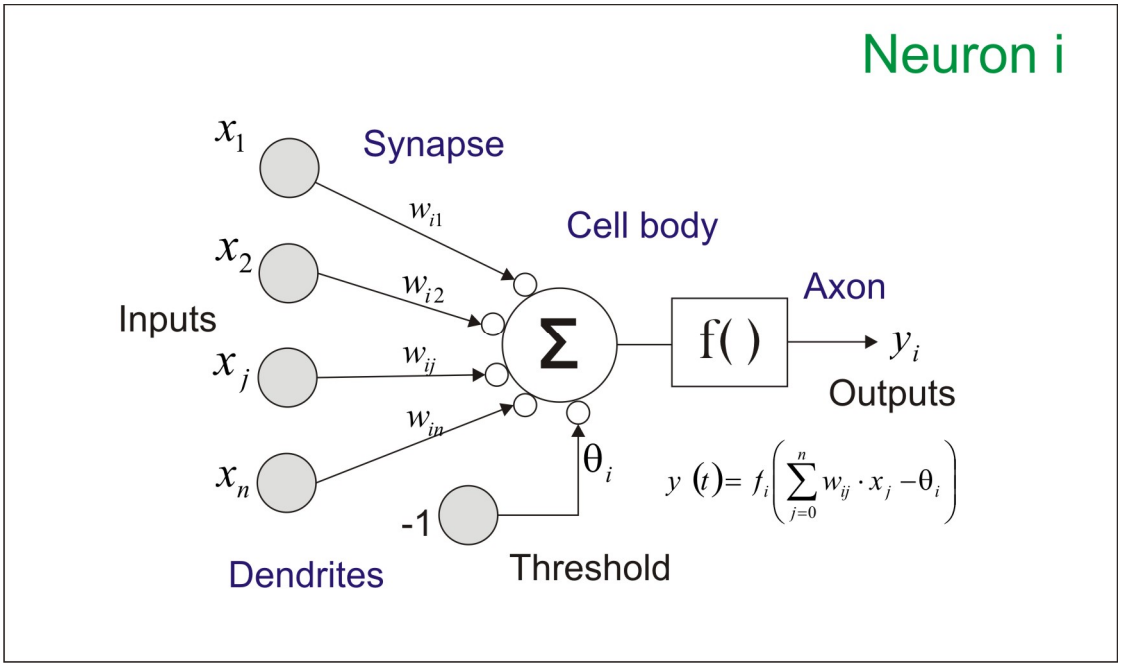


Figure 5.5: Standard neuron (Haykin, 1999)

A parameter that is often added to the set of weights of the neuron is the threshold θ_i. This is subtracted from the postsynaptic potential, so the argument of the activation function is (Haykin, 1999):

$$\sum_j w_{ij} \cdot x_j - \theta_i \tag{5.7}$$

Correspondingly, if the indices i and j at 0 and setting $w_{i0} = \theta_i$ and $x_0 = -1$ (constant), the behaviour of the neuron can be described by (Haykin, 1999):

$$y\ (t) = f_i\left(\sum_{j=0}^{n} w_{ij} \cdot x_j\right) \tag{5.8}$$

V.3.3. NEURAL NETWORK ARCHITECTURE

The topology or structure in which different constituent neurons of the neural network are associated is called an *architecture* (Haykin, 1999). In an ANN, the nodes are connected by synapses; the synaptic structure determines the behaviour of the network. Synaptic connections are directional; i.e., information can only flow in one direction (from pre-synaptic neuron to postsynaptic neuron) (Haykin, 1999).

In general, neurons are grouped into structural units called *layers* (Haykin, 1999). Within a layer, neurons can be grouped to form *neuronal groups*. Within a *layer* or

group, neurons are often the same type. The set of one or more layers is called a *neural network* (Haykin, 1999). Normally, all neurons receive signals from one layer and send signals to the next layer.

There are three types of layers (Engelbrecht, 2008) (see Figure (5.6)):

- *Input layer:*

  The input layer is composed of neurons that receive data or signals from the environment.
- *Hidden layer:*

  This layer has no direct connection with the environment. The function of the hidden neurons is to intervene between the external input and the network output.
- *Output layer:*

  The neurons in this layer provide the response of the neural network.

Typically, the inputs in each layer of the network are the output signals of the preceding layer only.

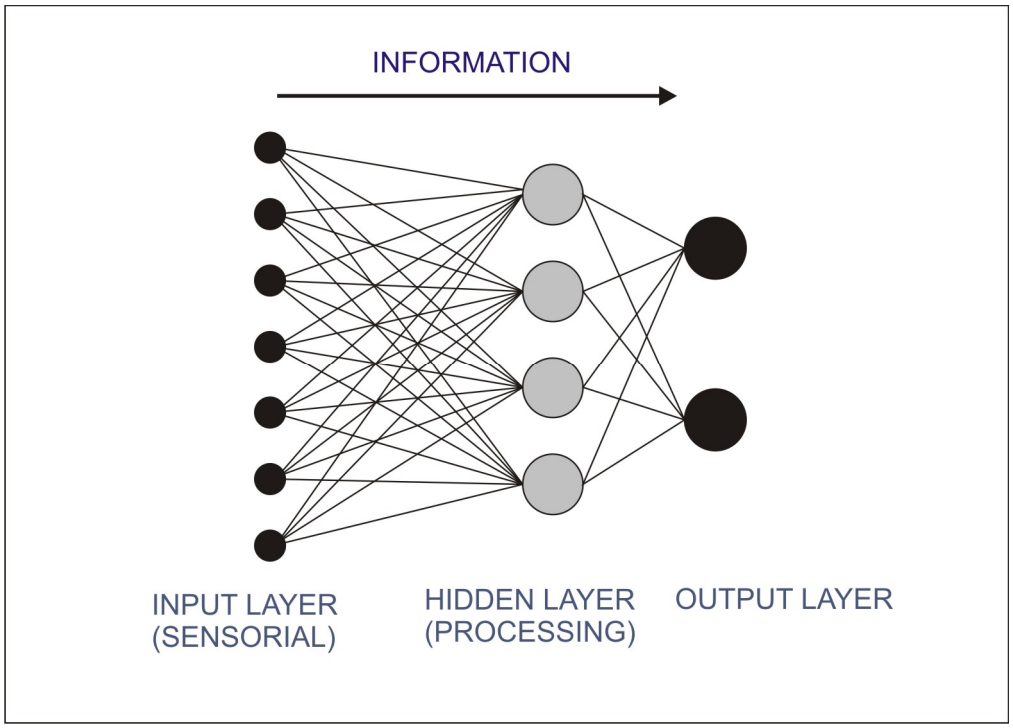


Figure 5.6: Layers of a network (Haykin, 1990)

The connections between neurons can be excitatory or inhibitory, depending on the sign of synaptic weight associated with the connection. If the synaptic weight is negative, an inhibitory connection is obtained; if on contrary the sign is positive, an excitatory connection is found. This distinction is not frequently used, as the weight and its size will be determined at each instant by the training algorithm (Haykin, 1999).

The connections may also be classified into interlayer connections which correspond to the connections between neurons in different layers, and connections between neurons in a single layer (Haykin, 1999).

V.3.3.1. DEFINITION OF NEURAL NETWORK

In order to find a definition of neural network, a mathematical concept of a graph will be used. Through this term, a neural network can be defined as follows:

A *neural network* is a directed graph with the following properties (Freeman & Skapura, 1993):

- For each node i, a variable x_i is associated.
- Each connection (i, j) of nodes i and j have a weight associated $w_{ij} \in \Re$.
- Each node i has associated a threshold θ_i.
- For each node i a function $f_i\left(x_j, w_{ij}, \theta_i\right)$ is defined, which depends on the weights of their connections, the threshold and the states of j nodes connected to it. This feature provides the new state of the node.

V.3.3.2. TYPES OF NEURAL NETWORK

As previously specified, the topology or architecture of a neural network consists of the organization and arrangement of neurons in the net, forming layers or clusters of neurons more or less distant from the entrance and exit of the network. In this sense, one of the basic parameters of the network is the number of layers (Haykin, 1999).

V.3.3.2.1. SINGLE LAYER NETWORKS

In single layer networks, there are connections between elements or neurons of the same layer. Single layer networks are used generally in tasks related to what is

known as self-association (regenerate input information that is incomplete or distorted).

This network corresponds to the simplest neural network (Figure (5.7)) since it has a layer of input neurons that project them to a layer of output neurons which perform various calculations.

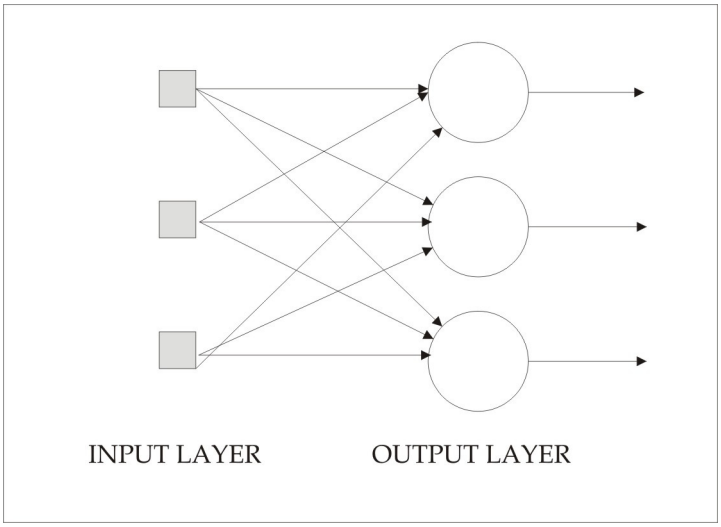


Figure 5.7: Single layer Network (Haykin, 1999)

V.3.3.2.2. MULTILAYER NETWORKS

Multilayer networks are those that have a set of neurons grouped in several (2, 3, etc.) levels or layers (see Figure (5.8)). In these cases, one way to distinguish the layer where a neuron belongs would be to focus on the origin and destination of the signals (Haykin, 1999). Usually, all neurons of a layer receive input signals from other previous layers (which are closer to the entrance of the network), and send output signals to the next layer (which is closer to the output of the network). These connections are called forward or *feedforward* connections (Haykin, 1999).

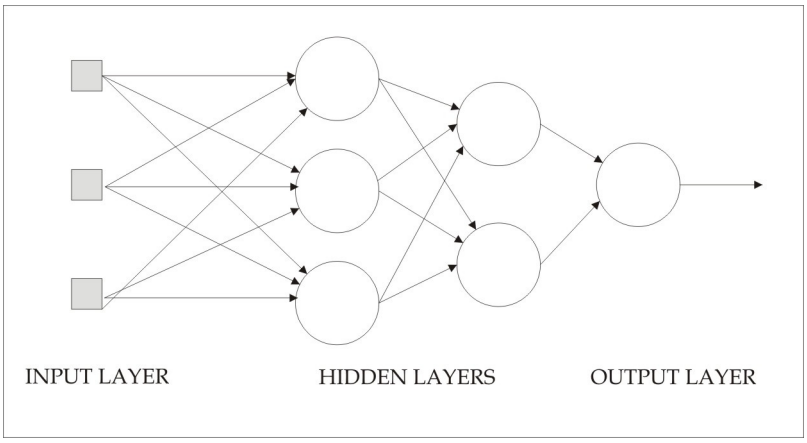


Figure 5.8: Multilayer Network (Haykin, 1999)

However, in many of these networks, it is also possible to connect the output of neurons in layers after the entry of previous layers; these connections are called back or *feedback* connections (Haykin, 1999).

These two possibilities allow two types of networks with multiple layers to be distinguished: forward network connections or *feedforward* networks, and networks that have connections to both forward and backward or *feedforward/feedback* networks (Haykin, 1999).

V.3.3.2.3. RECURRENT NEURAL NETWORKS

The connectivity between nodes in a neural network is related to how the outputs of neurons are channelled to become inputs of other neurons. The output signal of a node can be an input to another processing element (as shown in Figure (5.9)), or even be an entry of itself (recurrent connection) (Haykin, 1999).

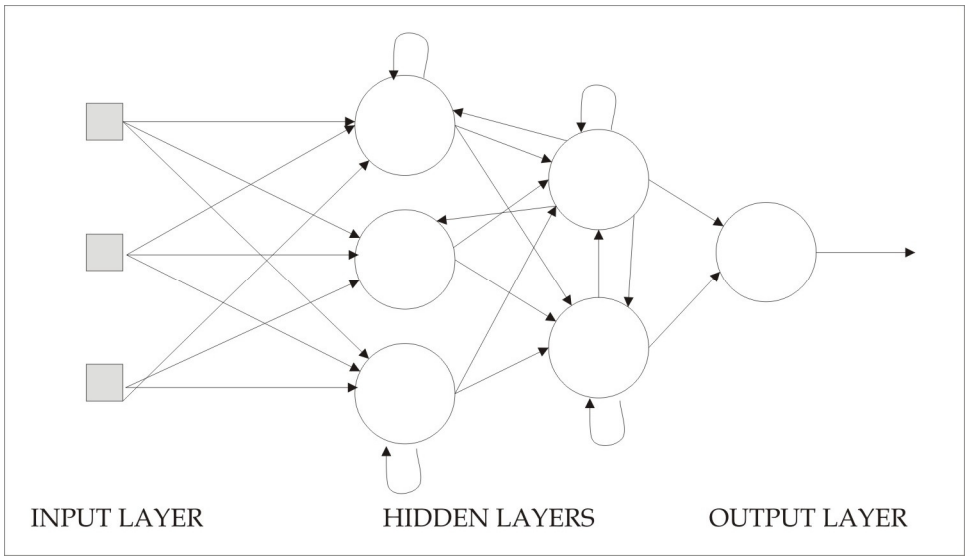


Figure 5.9: Recurrent neural network (Haykin, 1999)

V.3.4. LEARNING MECHANISM

The input data are processed through the neural network in order to achieve an output. A neural network is also able to draw generalizations from a specific set of historical examples of these types of decision problems (Haykin, 1999). A neural network must learn to calculate the correct output for each constellation (array or vector) of input from the set of examples. This learning process is called the *training process* or *upgrading* (Engelbrecht, 2008).

If the network topology and the different functions of each neuron (input, output, and activation) do not change during learning while the weights on each of the connections do, the learning of the network is called *weight adaptation* (Engelbrecht, 2008).

In other words, learning is the process by which a neural network modifies its weights in response to input information (Haykin, 1999). The changes that occur during this process are reduced to the destruction, modification and creation of connections between neurons. In biological systems, there is a continual destruction and creation of connections between neurons (Beale, 2000). In models of artificial neural networks, creating new connections implies that the weight of this connection varies to a nonzero value. Similarly, a connection is destroyed when their weight becomes zero (Beale, 2000).

During the learning process, the weights of the network connections are modified; therefore, this process can be considered complete (the network has learned) when the weight values remain stable ($d\omega_{ij}/dt = 0$) (Haykin, 1999).

An important aspect in the learning of neural networks is understanding how to modify the weight value, i.e., to know the criteria used to change the assigned value to the connections when the network has to learn new information (Engelbrecht, 2008).

There are two major learning methods that can be distinguished (Engelbrecht, 2008):

- Supervised learning
- Unsupervised learning

Another criterion that can be used to differentiate the learning rules is to consider whether the network can learn during its normal operations or if learning involves disconnecting from the network, i.e., disabling until the process ends. In the first case, this would be an *on-line* learning, while the second is what is known as *off line* (Haykin, 1999).

When learning is off line, it can be distinguished as either a *learning* or *training phase* and or an *operation* or *functioning phase* (Haykin, 1999), where a set of training data and a set of test data will be used in the corresponding phase. In addition, the weights of the connections remain fixed after ending the training stage of the network (Haykin, 1999). Precisely because of their static nature, these systems do not have functioning stability problems (Engelbrecht, 2008).

A generalization of the formula or rule to decide changes in the weights is as follows (Haykin, 1999):

$$\omega_{ij}(t+1) = \omega_{ij}(t) + \Delta\omega_{ij}(t) \tag{5.9}$$

where *t* is the learning stage.

V.3.4.1. SUPERVISED LEARNING

Supervised learning is characterized as a learning process that is done through a controlled training by an external agent (supervisor, teacher) that determines the desired response from a given input (Haykin, 1999). The supervisor controls the output of the networks, so in case of mismatch with the desired output, a modification of the weights will be made in order to approximate the results (Haykin, 1999).

This type of learning can be done using one of the three following methods (Haykin, 1999; Beale, 2000):

- Learning by error correction
- Learning by reinforcement
- Stochastic learning
-

V.3.4.1.1. LEARNING BY ERROR CORRECTION

This method consists of adjusting the connection weights of the network based on the difference between the desired values and those obtained at the output of the network, i.e., as a function of the output error (Haykin, 1999).

An example of this type of algorithm is the *Perceptron learning rule* used in the training of the network of the same name, which was developed by Rosenblatt (1958). According with this rule for each neuron in the output layer, an output deviation error, δ, is calculated. This error is then used to change the weights for the connection of the previous neuron. The change in weights through the Perceptron learning rule is performed according to the following rule (Rosenblatt, 1958):

$$\Delta w_{ij} = \sigma \cdot (out)_j \cdot (a_{qi} - (out)_i) \tag{5.10}$$

where a_{qi} is the desired output (target) of the output neuron N_i, $\delta_i = (a_{qi} - (out)_i)$ is the deviation of the neuron N_i and σ is the learning (Rosenblatt, 1958).

The output of the neuron N_j (*out* $_j$) is used because this value influences the overall input as well as the activation and the output of the neuron N_i. This can be represented as "*the domino effect*". See Figure (5.10):

Figure 5.10: Influence of the output neuron N_j over neuron N_i (Beale, 2000)

Another well-know algorithm that belongs to this classification is the *Delta learning rule* or *rule of minimum square error* (LMS error: Least Mean Square Error), which uses the deviation of the output but takes into consideration all the previous neurons of the output neuron (Haykin, 1999). This allows quantifying the global error at any time during the network training process, which is important because the more error information the faster the learning. The error δ is then divided between previous neurons (Haykin, 1999).

Finally, the *backpropagation learning rule*, also known as multilayer LMS rule, should be mentioned. This is a generalization of the Delta learning rule and is the first learning rule that allows changes in the weights of the neurons of the hidden layer (Haykin, 1999). This learning algorithm will be explained later on.

V.3.4.1.2. LEARNING BY REINFORCEMENT

This learning process is slower than learning by error correction. This process is based on lack of example of the desired behaviour, i.e., there is no indication of the desired output for a given input during training (Haykin, 1999).

In this rule, the role of the supervisor is to indicate through a reinforcement signal if the output from the network follows that desired (success = +1, failure =-1), and according to this, the weights are adjusted based on a mechanism of probabilities (Engelbrecht, 2008). This type of learning is a case where the supervisor role is more like a critic (who judges about network response) than a teacher (who indicates the response that should be generated), as in the case of supervision by error correction (Haykin, 1999).

V.3.4.1.3. STOCHASTIC LEARNING

Stochastic learning essentially consists of random changes in the values of the connection weights of the network and it assesses the effect on the desired target and probability distributions (Beale, 2000).

In stochastic learning, an analogy is often made in terms of thermodynamics, associating the neural network with a solid with energy (Beale, 2000). In the case of the network, the energy would represent the degree of stability, so that the minimum energy state corresponds to a situation in which the weights of the connections meet the desired target (Beale, 2000).

As before, learning is affected through random changes in the values of the weights and then determining the network energy (usually the energy function is of Liapunov type). If energy is lower after change, i.e., if the behaviour of the network approaches the desired change, this is accepted (Beale, 2000). Conversely, if the energy is not lower, the change is accepted as a function of a given and preset probability distribution (Haykin, 1999).

V.3.4.2. UNSUPERVISED LEARNING

Networks with unsupervised learning (also known as self-supervising) require no external influence to adjust the weights of the connections between its neurons (Haykin, 1999). The network does not receive any information from the environment to indicate whether the output generated in response to a given input is correct or not (Haykin, 1999).

These networks have to find the features, regularities, correlations or categories that can be established among input data (Haykin, 1999). There are several possibilities as to the interpretation of the output of these networks, which depend on their structure and learning algorithm used (Haykin, 1999).

In some cases, the output represents the degree of familiarity or similarity between the information that is presented at the input and information that have been shown before (in the past) (Haykin, 1999). In another case, a clustering could be performed

or categories established, indicating to which category the input information belongs. In this case, the network itself must find the appropriate categories from the correlations between information provided (Haykin, 1999).

In general, two types of unsupervised learning algorithms are often considered, which give rise to the following (Haykin, 1999):

- Hebbian learning
- Competitive and comparative learning

V.3.4.2.1. HEBBIAN LEARNING

This learning rule is the basis for many others, which aims to measure the familiarity or to extract features from input data. The foundation is a fairly simple assumption: if two neurons N_i and N_j take the same state simultaneously (both active or inactive), the weight of the connection between them increases (Beale, 2000).

The inputs and outputs allowed to the neuron are: $\{-1,1\}$ or $\{0,1\}$ (binary neurons). This can be explained by the Hebb's learning rule, which was originated from the classical biological neuron, that can only have two states: active or inactive (Beale, 2000).

V.3.4.2.2. COMPETITIVE AND COMPARATIVE LEARNING

In competitive learning, neurons "fight" each other to represent a class or standard input (Engelbrecht, 2008). The selected neuron is one whose incident weights are more similar to the entry pattern. The learning consists in strengthening the winning connections and weakening the others, so that the weights of the winning unit resemble the input pattern (Engelbrecht, 2008).

The reconstruction of an input pattern from a winning neuron takes the weight of that neuron because these values have the most resemblance (Engelbrecht, 2008).

V.3.5. THE PERCEPTRON

One of the most significant features of neural networks is their ability to learn from some source of information interacting with their environment. In 1958, psychologist Frank Ronsenblant developed a simple model of a neuron based on the model of

McCulloch and Pitts and a learning rule based on error correction. This model was called the *Perceptron*. One feature of paricular interest in this model was its ability to learn to recognize patterns. The Perceptron consists of input sensors that receive the input patterns for recognizing or classification and an output neuron which classifies input patterns into two classes according to the output, 1 (activated) or 0 (off) (Haykin, 1999). However, this model had many limitations; for example, it is not capable of learning the XOR logic function. Some years later, the error back-propagation learning rule was proposed for the Multilayer Perceptron, which demonstrated that it can be considered as a universal approximate (Engelbrecht, 2008).

Figure (5.12) illustrates a monolayer or simply perceptron with a single neuron, which has a limited use in pattern classification.

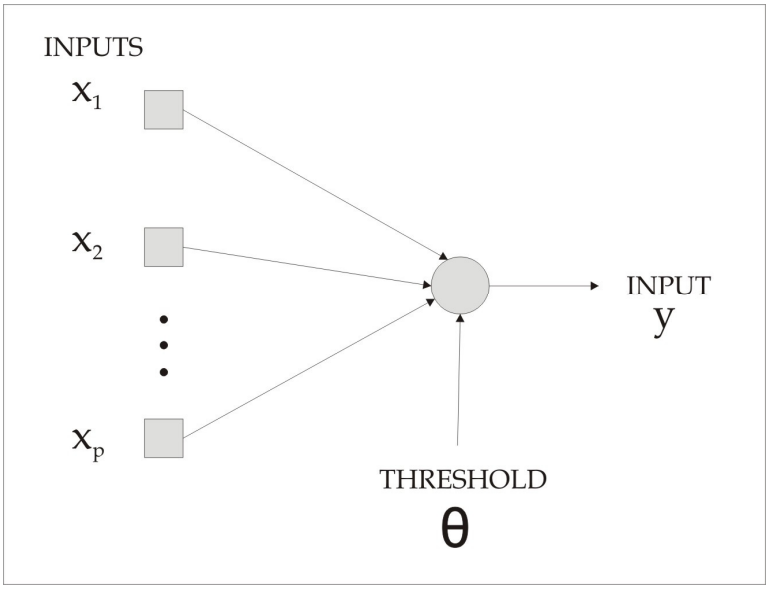


Figure 5.11: Single neuron (Ronsenblant, 1958)

A single perceptron has many limitations; the most important is that it cannot distinguish classes that are not linearly separable. If a group of these are organized in layers, this limitation is exceeded (Haykin, 1999). Each layer consists of multiple neurons whose input comes from the units of the previous layer and whose output goes to the units of the posterior layer, without any contact with any other layers, or between neurons in a single layer. The first layer, called the input layer, receives the

input signal but does not perform any type of processing. The result of the network is given by the state of neurons in the last layer, called the output layer. The intermediate layers are called hidden (Engelbrecht, 2008).

Learning is done by submitting to the network at each step the value of the input and desired output. Following the rule of Rosenblat, minimizing the least square error (LSE) between the output obtained and the desired output is obtained by the generalized delta rule or backpropagation of error algorithm; i.e., once the output and further comparison with the desired output is made, an error in each neuron of the last layer is obtained, and thus their synaptic weights are changed (Haykin, 1999). However, that error is transmitted to the layers above, and corrects their weights for all the neurons in these layers. Although this algorithm can be applied to networks with any number of layers, a single hidden layer (Homik, 1989) is sufficient to approximate, with a given precision, any function with a finite number of discontinuities, using nonlinear activation functions for neurons of the hidden layer. This network can also be achieved in a finite number of learning steps (Homik, 1989).

In general, multilayer neural networks are used with error backpropagation learning to simulate an unknown function from pairs of input and output signals obtained from learning examples (Haykin, 1999). In these cases, the network is expected to be able to generalize properly, so that even with input signals which have not been previously presented as patterns of learning, the proper output is obtained (Engelbrecht, 2008).

Each hidden or output neuron of a multilayer perceptron is thus designed to do two calculations: the signal propagation forward (*forward pass*) calculation of the gradient vector required for the retro-propagation (*backward pass)* (Haykin, 1999).

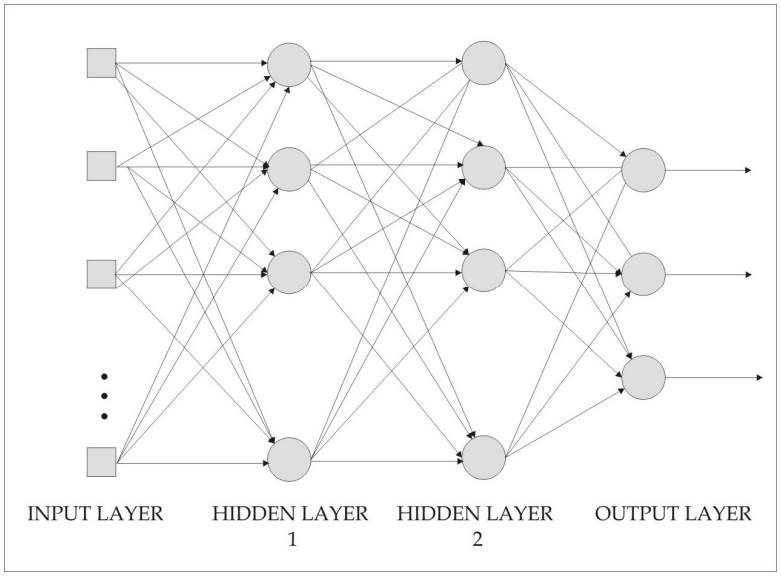


Figure 5.12: Multilayer perceptron with two hidden layers (Haykin, 1999)

The typical structure of a multilayer perceptron can be seen in Figure (5.12), where two types of signals can be identified (Figure (5.13)):

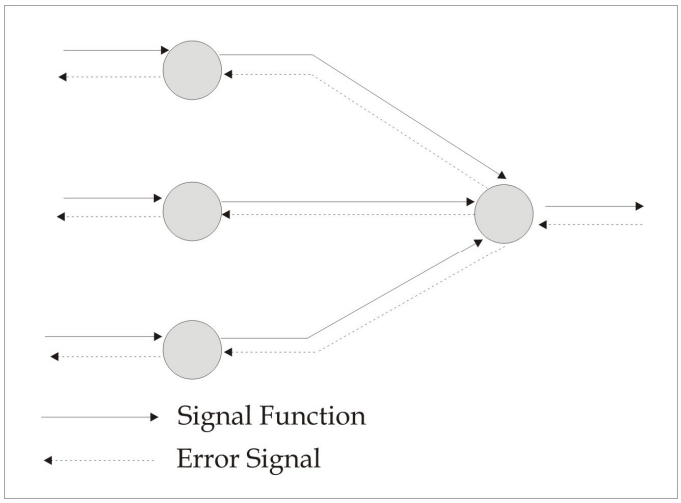


Figure 5.13: Signals propagation in a multilayer perceptron (Haykin,1999)

- *Function signal*: is the signal that is propagated from the input (left) to the output layer (right).
- *Error signals*: are those generated by the output neurons and are back-propagated in the form of adjustment of the synaptic connections to the input in order to adjust the obtained output as closely to the desired output.

V.3.6. BACKPROPAGATION ALGORITHM

The main idea of this rule or algorithm is to calculate the errors for the units of the hidden layers from the error of the units in the output layer being propagated, layer by layer, toward the entrance (Haykin, 1999).

The error signal at the output neuron *j* at iteration *n* (presentation of the n^{th} training pattern) is defined as (Haykin, 1999):

$$e_j(n) = d_j(n) - y_j(n) \tag{5.11}$$

where $e_j(n)$ is the output error of the neuron *j*, $d_j(n)$ is the desired response of the node *j*, $y_j(n)$ is output signal of the neuron *j* which is the output node.

The sum of the square errors of the network can be defined as (Haykin, 1999):

$$E(n) = \frac{1}{2}\sum_{j \in C} e_j^2(n) \tag{5.12}$$

where set *C* includes all nodes in the output layer of the network.

Figure (5.14) shows the neuron *j* fed by a set of signals from one layer of neurons on the left.

The number of patterns or examples is known as the Training set and is denoted by *N*. Thus, the mean square error (E_{av}) is obtained by summing $E(n)$ for all iterations and normalizing for *N* (Haykin, 1999):

$$E_{av} = \frac{1}{N}\sum_{n=1}^{N} E(n) \tag{5.13}$$

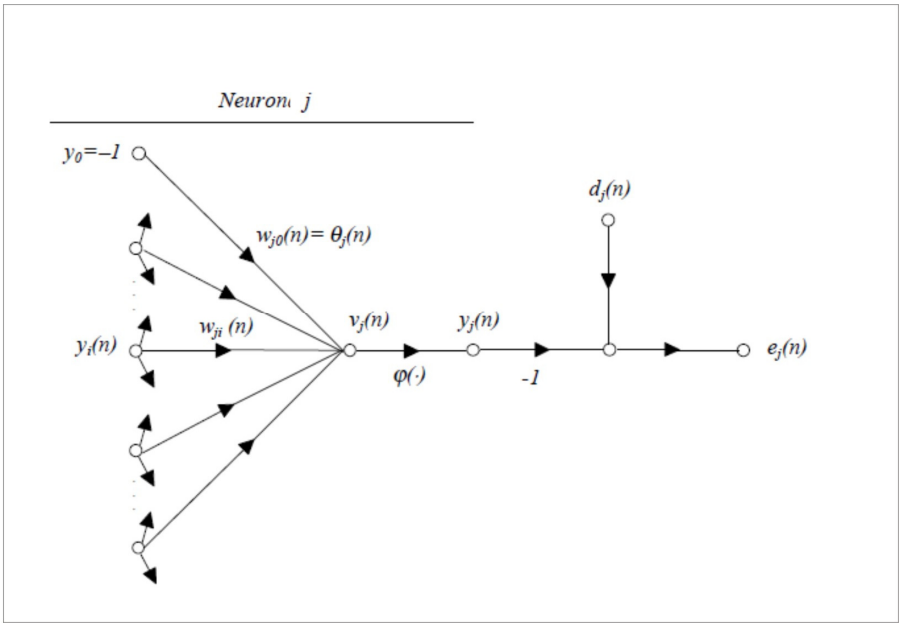


Figure 5.14: Signals in an output neuron (Haykin, 1999)

For a given training set, E_{av} is a measure of learning of RNA called *cost function* (Haykin, 1999). The objective of this learning process is to minimize this error function by adjusting the free parameters (weights, adaptation constant, etc) (Haykin, 1999).

V.4. NEURAL NETWORKS IN THE PROCESS INDUSTRIES

In recent times, historical databases with real time capture and large storage capacity have been set up to capture data for industrial processes. In refineries and petrochemical industries, these types of databases collect process data in real time using distributed control systems (DCS), with a typical sampling rate of 1 minute (Fogelman-Soulie & Gallinari, 1998).

In general, these databases are available online for many years and also include the results of routine laboratory analysis. In a typical refinery, it is normal to collect thousands of variables (tags), including process variables and set points, valve outputs and controller modes (e.g., automatic, manual, local, remote, etc.) (Fausset, 1994). These real-time histories produce very large databases from which very rich information, although not always easy to find, can be drawn regarding the process

(Fausset, 1994). Due to special compression techniques, these databases have tremendous storage capacity.

As example: if 40,000 tags from DCS are sampled every minute (a modest number of an oil refinery average size), 57.6 million data points are collected every day, totalling about 21 billion points of data annually (Fausset, 1994). It is not unusual to have on-line data collected over several years, easily accessible using standard PC tools such as an *Excel* spreadsheet.

The term "data mining" (also known as Knowledge Discovery in Databases - KDD) has been defined as “the nontrivial extraction of implicit information, previously unknown and potentially useful” from data (Frawley et al., 1999). It uses machine learning, as well as statistical and visualization techniques, to discover and present information in a form that is easily understandable by humans. The metaphor of the mine is truly powerful: there are rich veins of material (i.e., useful data from the point of view of economy, security and operation of the process) hidden in a vast amount of raw stored data, much of which can be considered as scum (Fausset, 1994). The effort to extract the precious from the dross material is based on multiple mathematical and IT skills that help in the task. One such methodology is the Neural Network (NN) (Fausset, 1994).

However, for the NNs to successfully exploit the rich veins of material found in large databases, they not only need to be fed with the appropriate data input and output, but also must complete the following 2 important requirements (Fausset, 1994):

- The NN should have a structure capable of adequately representing the problem.
- The NN should be trained with robust and reliable algorithm that is able to converge to an acceptable solution.

Not surprisingly, many of the items available in this area are rare in industrial applications of NNs. Many times, much effort can be spent fighting with the best NN topology, while losing many CPU hours trying to get a fit and reasonable convergence (Fausset, 1994).

Neural networks allow the engineer to create process models using historical data for the same process. The models predict how the process will respond to changes of inputs and different working conditions. The optimum operating conditions, subject to restrictions, may also be determined with the NNs correctly stated. The models identified using NNs can be used in studies of off-line processes or be installed on-line to provide early detection of process problems and to determine the set-point for continuous optimization of the process, to maximize profits (Fogelman-Soulie & Gallinari, 1998).

The NNs bring to life the historical data, revealing the most important factors affecting the quality and performance of products. This knowledge can often detect improvements without any capital investment. The general areas of potential use of NNs are (Fogelman-Soulie & Gallinari, 1998):

- *Quality Control, Inferential Sensors and Model Reduction:* In the actual globalized economy, the management of quality in real time is a critical application, but the quality tests are rarely available without delay and are usually expensive. The neural network-based models provide "virtual" measures in real time, allowing fast control actions to maintain quality in the desired goal. Models can be obtained not only from plant data and laboratory data but also from runs generated with rigorous simulation models (developed, for example, in HYSYS). This last procedure is known as "model reduction".

- *Process Optimization*: The value of model-based optimization is well established but, in general, the analytical models of a process can be very difficult to obtain. When using neural networks in conjunction with their ability to optimize online and in real time, it may be possible to obtain the best economic potential of a process.

- *Predictive Maintenance and Security:* Models based on neural networks can be used to monitor the performance of machinery and equipment. Errors in operational models or sensors can be detected with these shifts, allowing engineers to correct problems before they result in major incidents. The availability of the plant and equipment can thus be improved. Continuous

monitoring of the content of emissions (CEM Continuous Emissions Monitoring) of NO_x, CO_2, SO_2 in the exhaust gases of furnaces and boilers is a typical application in this area.

- *Sensor Validation:* The gradual drift and / or abrupt failure of sensor signals are the major source of unplanned plant stops in production and off-specification products. With models based on neural networks it is possible to follow the sensor values and generate alarms when the measurements from physical sensors disagree with the values inferred for them. The inferred value also can be used as a baseline where the instrument is recalibrated or repaired.

- *Prediction and Estimation:* The future can be predicted within the accuracy given by models based on behaviours. Neural networks can learn the best models, continuously adapted to the use of the latest measured data. Engineers can use these predictions to estimate the demand for short-term markets, predict future states of the process or even weather conditions that affect emissions and impact the vicinity of the plant.

SECTION C:
MODELLING AND SIMULATION RESULTS

Chapter VI

MODEL BASED SIMULATION

VI.1. INTRODUCTION

The Kynch theory provides a good starting point for understanding the sedimentation phenomenon based on the propagation of concentration waves in a suspension (Kynch, 1952). However, Kynch settling behaviour occurs only in zones where no consolidation effects are involved. This theory was discussed in detail in section III.2.2.

Ideal Kynch settling is sometimes referred to as free settling behaviour. However, it does not describe the entire thickening process because it is not applicable when the sediment is under compression at the bottom of the thickener. In this lower region, a different model is required to describe the behaviour of the sediment. In particular, the settling velocity of the flocs in the sediment depends not only on the local concentration of the solids (Kynch, 1952), but on the gradient of the concentration as well. Thus, the dynamic differential equation that describes the behaviour of the entire process in a thickener is obtained from differential mass and momentum balances. The result is a hyperbolic conservation law which degenerates in a non-linear partial differential equation that can be solved numerically using proper initial and boundary conditions as well as appropriate parameters of the constitutive function related to the slurry used for the study.

The aim of this work is to present a simulation of the sedimentation process based on the phenomenological theory of sedimentation of flocculated suspensions (Concha et al. 1996), in order to predict the behaviour of the washer used in aluminium industries.

VI.2. RED MUD CONSTITUTIVE EQUATIONS:

The specific material properties of the suspension are described by $f(\varphi)$ and $\sigma_e(\varphi)$, which can be stated as Kynch solid flux function and effective solid stress, respectively.

In the case of the Kynch solid flux function, a function of the Richardson-Zaki type can be used to describe the settling behaviour. On the other hand, for the effective solid stress function, an empirical equation (Bürger *et al*, 1999) will be used, as well as the Herschel-Bulkley model. The methods to construct these models will be explained below.

VI.2.1. SETTLING BEHAVIOUR:

The behaviour of settling particles in a slurry can be conveniently studied in small batch experiments. In these tests, settling velocity can be found as a function of a concentration.

Settling of noncolloidal suspension has been investigated by Kynch (1952), who obtained a relation for sedimentation velocity, as follows:

$$u_C = u_0 \cdot h(\varphi) \tag{6.1}$$

where u_0 is the settling velocity of a single particle which can be determined by Stokes' law, φ is the volumetric fraction, and $h(\varphi)$ is a settling function which is defined as a concentration-dependent function.

In the case of low concentration regimes the concentration-dependent function is readily resolved, and has a reasonable agreement between theory and experiments (Davis *et al* 1988). On the other hand, for higher concentration regimes, several experimental investigations have found a relative agreement with the empirical relation of Richardson-Zaki (1954), which can be written as:

$$h(\varphi) = (1-\varphi)^n \tag{6.2}$$

where n is an index that depends on the Reynolds number ($n \approx 5.5$ at low Re) which can be found experimentally.

The relationship described above is acceptable in most applications, even if it does not completely match low concentration data.

The sedimentation velocity u_C in the case of settling of suspensions can be expressed as a function of volumetric concentration φ, while the volumetric rate of sedimentation per unit area, or flux, can be represented by the product of the sedimentation velocity and volumetric concentration.

According to this, the flux can be written as follows:

$$f(\varphi) = u_C \cdot \varphi \tag{6.3}$$

Using (6.1) and (6.2) in (6.3), a new relationship for sedimentation flux is obtained:

$$f(\varphi) = \varphi \cdot u_0 \cdot (1-\varphi)^n \tag{6.4}$$

Thus, for the sedimentation of a suspension of uniform particles in a liquid, the relation between observed sedimentation velocity u_C and fractional volumetric concentration φ is given by the Richardson-Zaki equation which takes the form:

$$u_C = u_0 \cdot (1-\varphi)^n \tag{6.5}$$

In order to determine the settling properties of bauxite residue and to fit the results according to Equation (6.5), a series of batch settling tests were made at the Rio Tinto Alcan Alumina refinery.

VI.2.1.1. EXPERIMENTAL: BATCH SETTLING TEST

The equipment used for the batch settling test was a series of settling cylinders of 1 litre of volume and a stopwatch. In these vessels, different concentrations of red mud from the washer's underflow were allowed to settle in order to find out the settling characteristics of the alumina residue. The layout of a settling vessel is shown in Figure (6.1).

The list of materials used for the settling test was as follows:

- Graduated settling cylinders (1L).
- Underflow suspension (washers 20, 50, 70)
- Overflow liquor for dilution (washers 30, 60 and water for washer 70)

- Flocculant (the dosage was varied according to the washer used for testing)
- Stopwatch

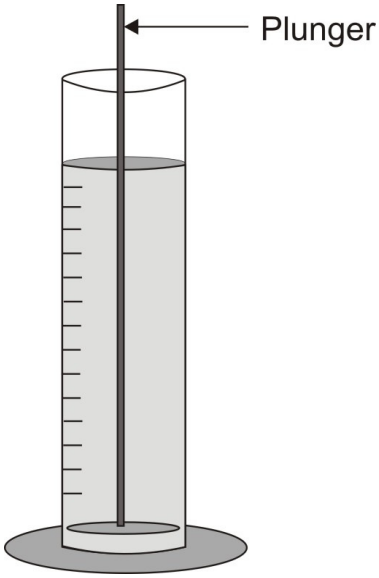


Figure 6.1: Experimental settling vessel

The samples for the characterization of the residue were collected from three different washers, -first operational washer (20), last washer (70) and an intermediate washer (50). Together with the samples taken from the underflow of the washer mentioned, enough overflow liquor was collected in order to enable dilution, which was helpful to obtain different suspensions of variable concentrations. The dilutions used for experiments were 50 g/L, 70 g/L and 90 g/L.

Considering the countercurrent character of the operation, the overflow liquor used to dilute the washer's underflow came from the next washer, i.e., the liquor used to dilute the underflow of the washer 50 came from the overflow of washer 60.

The flocculant used in the experiments varied depending on the characteristic of the underflow and the dosage could be calculated based on the concentration of the slurry.

As indicated above, once the washer underflow and overflow are collected, dilution has to be made in order to obtain different suspension concentration. To obtain the volume of liquor needed for dilution, a simple calculation is made according to the following relation:

$$C_1 \cdot V_1 = C_2 \cdot V_2 \tag{6.6}$$

where C_1 and V_1 are the initial concentration and volume respectively (original slurry concentration), and C_2, V_2 are the target concentration and volume.

Several cylinders of different slurry concentration were prepared in order to obtain representative and repetitive data for this study. Thus, for each washer, 12 samples were prepared in order to have enough cylinders to study different effects. For example, three columns of 50 g/L were allowed to settle using different flocculant dosages to observe the effects on sedimentation velocity.

The procedure carried out for batch settling test was as follows:

Note: all settling experiments were carried out under a controlled temperature environment using a water bath. The temperature was the same used in the process vessel.

a. Prepare underflow thickener suspensions with known initial concentration.
b. Add homogenised suspension sample to settling cylinders.
c. Record the initial height.
d. Homogenise, once again, the suspension using a flocculation test plunger.
e. Add flocculant to each cylinder (one at a time) under the same conditions.
f. Use a flocculation test plunger to gently homogenise the suspension. It is very important to be consistent, to allow a valid comparison of results.
g. Allow the contents of the vessel to settle. During this process, record the height of the interface regularly until a constant height is obtained.

As mentioned above, this procedure was made for each settling column and for each sample collected from the washers 20 (the first washer in operation was 20, washer 10 was *on duty* by the time of the experiments), 50 and 70.

As an example of the results obtained, the graph below shows the suspension settling rate (height of interface versus time) for washer 70 using a concentration of 50 grams per litre and with different flocculant dosages.

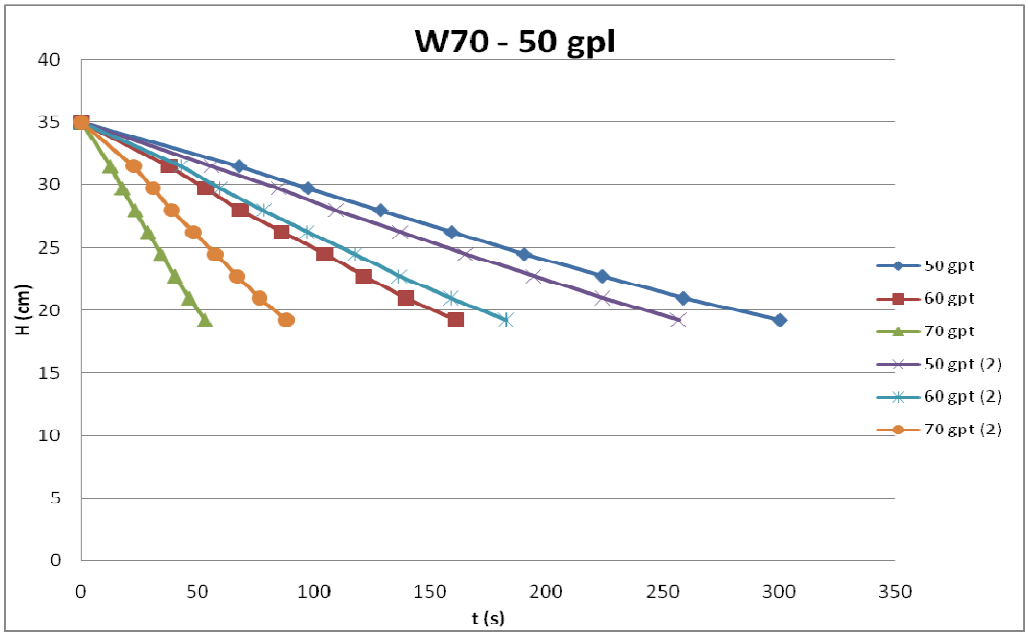


Figure 6.2: Settling rates at different flocculant dosage

In this case, settling tests were made in duplicate in order to improve data accuracies. The settling rates (indicated by the slope of the graph for each concentration) can be seen to increase when more flocculent is added.

Once the different data of height of interface against time are obtained, it necessary to fit those data according to the Richardson-Zaki equation (Equation (6.5)).

In order to do so, Equation (6.5) can be rewritten in logarithmic terms:

$$\log u_C = \log u_0 + n \cdot \log(1-\varphi) \tag{6.7}$$

According to this equation and for a given suspension, a plot of $\log u_C$ versus $\log(1-\varphi)$ can be made in order to find the Richardson-Zaki index n. In most cases, the graph of the Equation (6.7) is a straight line with slope n and intercept $\log u_O$.

In order to illustrate this procedure, washer 20 settling data is used below to find the index n, with the help of the Kynch theory of sedimentation. The specifications of the underflow are shown in the table below:

Table 6.1: Washer 20 specification for batch settling test

Washer No.	**20**		
Samples Origin	U/F Clarifier 30	**Density (g/mL)**	1.5
Solids Concentration (g/L)	50		
Dilution Origin	O/F Washer 30	**Density (g/mL)**	1.1
Flocculant	Nalco 85252 RRA		

In this case, the slurry for testing came from clarifier 30 because washer 10 was on duty.

First, a typical sedimentation curve (Figure (6.3)) showing the height of the interface as a function of time, is plotted using the settling data:

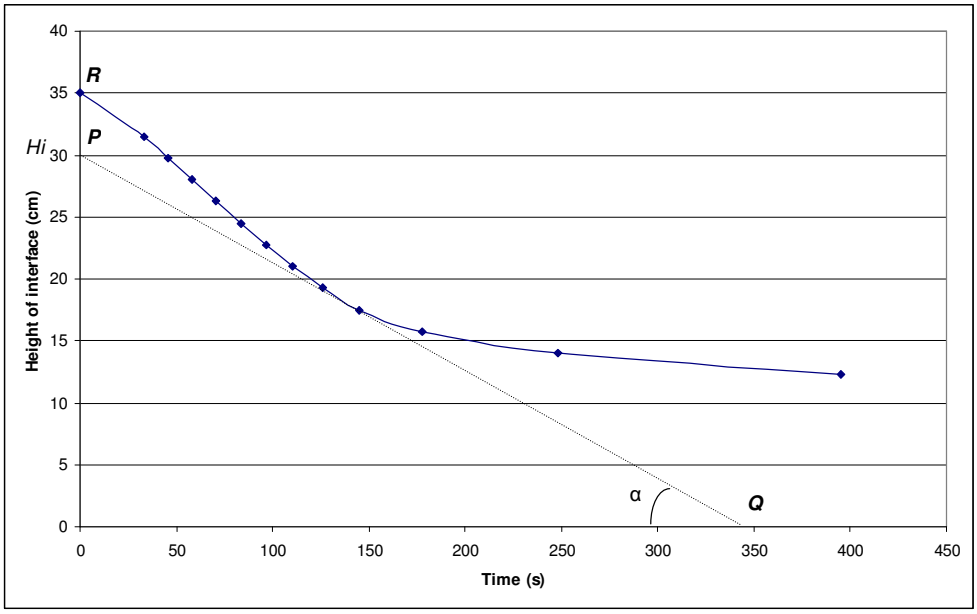


Figure 6.3: Height of the interface as a function of time for washer 20

The initial height is given by OR (35 cm) and the initial constant slope of the curve gives the sedimentation velocity $(u_c)_o$ for a concentration of 50 g/L (C_o). For some other heights, such as OP (30 cm), the slope of the tangent gives the sedimentation velocity u_c for a given concentration (Kynch, 1952):

$$C = C_O \cdot \frac{OR}{OP} \tag{6.8}$$

Thus, for each height, the corresponding concentration may be calculated and the slope of the tangent measured to give the sedimentation velocity. These values are tabulated in Table (6.2):

Table 6.2: Settling velocities

L (ml)	H(cm)	t (s)	C (g/L)	u_C (cm/s)
1000	35	0	50.00	0.1250
900	31.5	33.26	55.56	0.0980
850	29.75	45.73	58.82	0.0840
800	28	58.45	62.50	0.0710
750	26.25	70.86	66.67	0.0610
700	24.5	83.57	71.43	0.0500
650	22.75	96.73	76.92	0.0400
600	21	110.51	83.33	0.0280
550	19.25	126.02	90.91	0.0208
500	17.5	144.73	100.00	0.0150
450	15.75	177.39	111.11	0.0080
400	14	248.29	125.00	0.0030

From Figure (6.3), the critical concentration can be found graphically by drawing tangents in sectors of nearly constant velocity in Figure (6.3). Then, an angle bisection is constructed; the bisector intersects the settling curve at Z_c = 17.8 cm. These results can be seen in Figure (6.4).

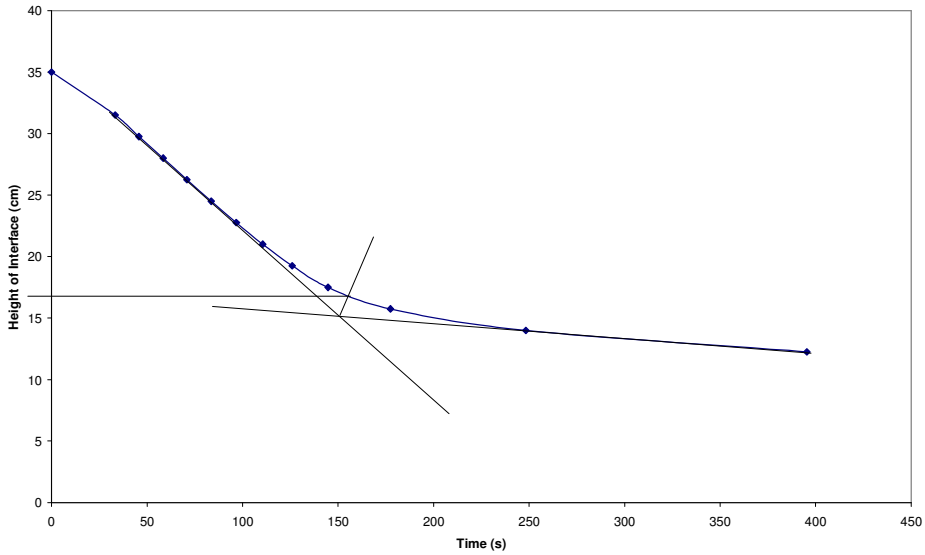


Figure 6.4: Critical point graphic determination

From Table (6.2), a plot (Figure (6.5)) of $\log u_C$ versus $\log(1-\varphi)$, an experimental value for Richardson-Zaki index n can be found.

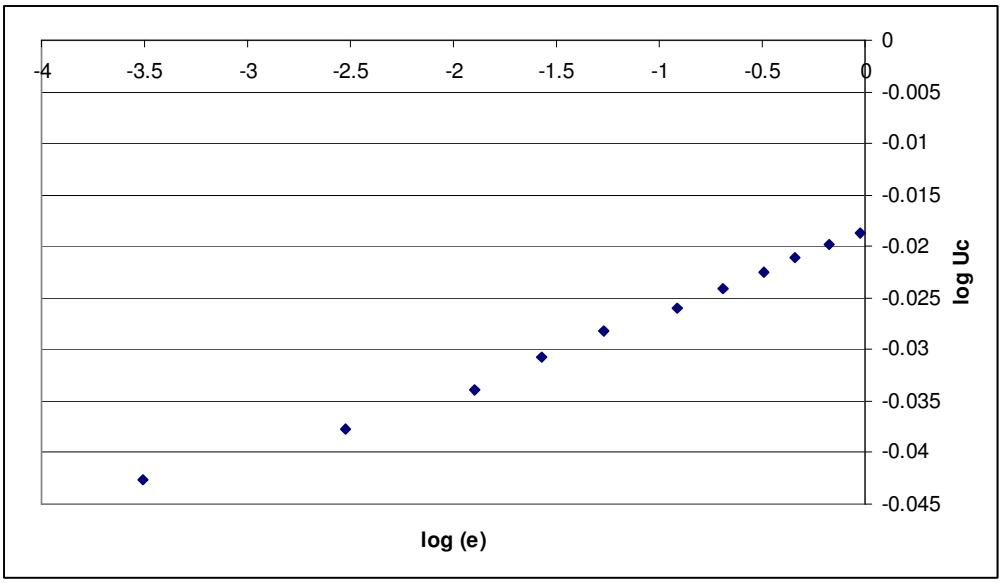


Figure 6.5: Voidage (e) versus $\log u_C$

According to the graph the value for the index (slope of the line) is 0.0065 and the intercept is -0.0173, which corresponds to sedimentation velocity for a single particle (u_0). Therefore, applying the inverse logarithm $u_0 = 0.96\ cm/s$; this can be replaced in the Equation (6.4) to obtain the sediment flux function.

The same procedure is used for each settling test obtained in the experiments.

VI.2.1.2. RHEOLOGY TESTS:

The rheological behaviour of concentrated mud residues is affected strongly not only by solids loading, particle morphology and particle size distribution but also by particle interactions. These interactions are influenced by chemical factors such as the presence of flocculant agents.

In order to determine the rheology of red mud slurry, samples of the washer underflows for 20, 50 and 70 were taken and different concentrations were prepared using caustic liquor for dilution for washer 20 and 50, while in the case of washer 70, water was used.

The samples, once collected, were stored in screw top containers and the rheological measurements were carried out with a Haake VT 550 rheometer.

The measured shear stress and shear rate data for the residue samples can be analysed according to the Herschel-Bulkley model (Green, 1997), which combines the Power law model with a yield stress variable. The model equation is expressed in the following form:

$$\tau = \tau_y + K \cdot \gamma^n \tag{6.9}$$

where τ is the shear rate (Pa), τ_y is the yield stress , γ is the shear rate (sec^{-1}), K is the power law constant (consistency index), and *n* is a constant (flow behaviour index).

In addition, fundamental models for the relationship between effective solid stress in the compressible sediment as a function of concentration have not yet developed but the Burger *et al* (1999) empirical equation has been found to describe some typical compressible sediments:

$$\tau(\varphi) = \begin{cases} \alpha \cdot e^{\beta \cdot \varphi} & for\ \varphi \geq \varphi_C \\ 0 & for\ \varphi < \varphi_C \end{cases} \tag{6.10}$$

where φ_C is the critical volume concentration, α and β are constants.

The rheological measurement of shear stress as a function of shear rate were conducted using the vane technique (Nguyen, 1983; Nguyen and Boger, 1983, 1985) which allows direct and accurate determination of the yield stress for red mud.

This technique involves a vane attached to a torsion measuring head, which is slowly rotated and immersed into a fluid material, where the torque as a function of time is measured.

The vane is four (or more depending on the tested material) thin blades set up at equal angles around a small cylindrical shaft. Figure (6.5) illustrates a four blade vane and a torsion blade.

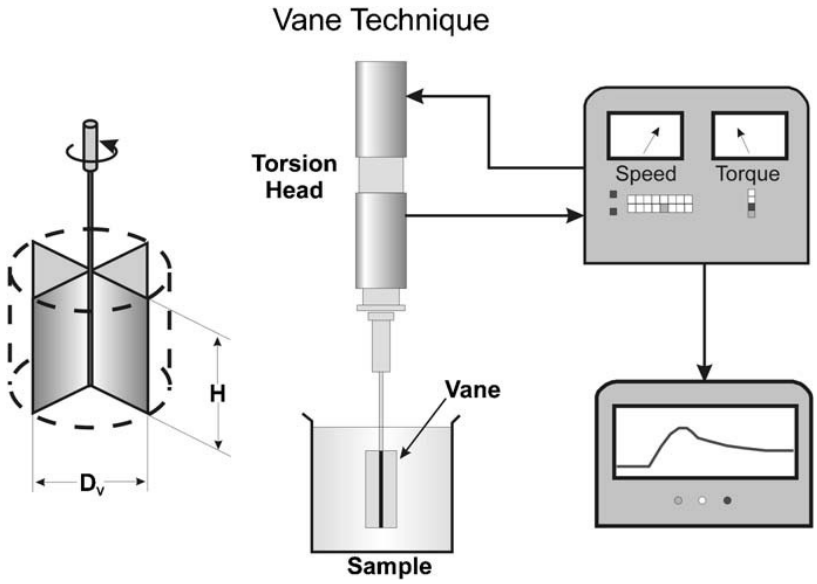


Figure 6.6: Vane technique for suspension rheology characterization (Boger, 1983)

The maximum torque can be related to the yield stress and the vane dimensions according to the Equation (6.11) (Boger, 1983):

$$T_m = \frac{\pi}{2} D_v \left(\frac{H}{D_v} + \frac{1}{3} \right) \cdot \tau_y \qquad (6.11)$$

As mentioned above, once samples from washers 20, 50 and 70 were collected, a methodology for rheology test was implemented which can be described as follows:

Note: all rheology experiments were carried out in a controlled temperature environment using a water bath. Thus, for washer 20, the controlled temperature was 50^{o}C, which corresponds to the process temperature inside the process vessel.

- ***a.*** Collect the samples from washers 20, 50 and 70. Prepare slurry of different concentrations (200, 300, 400 and 450 g/L).
- ***b.*** Store the suspensions in screw top containers and keep them in a controlled temperature environment (water bath).
- ***c.*** Select the desired vane and connect it to the rheometer.

d. Equilibrate the slurry of known concentration at the test temperature and homogenise it in order to eliminate the solids concentration distribution which can occur due to particle settling.

e. Use a cup to place the test sample (approx. 600 ml) and put it on a stand under the rheometer.

f. Raise the stand slowly so the vane penetrates the surface of the suspension.

g. Start the motor rotating at a constant rate and watch the build up of torque, which will rise until maximum reading is reached, followed by stabilisation to a slightly lower value over time.

h. These measurements were made in triplicate to assure accuracy.

Figure (6.7) shows typical torque as a function of time for washer 20 with a concentration of 300 grams per litre:

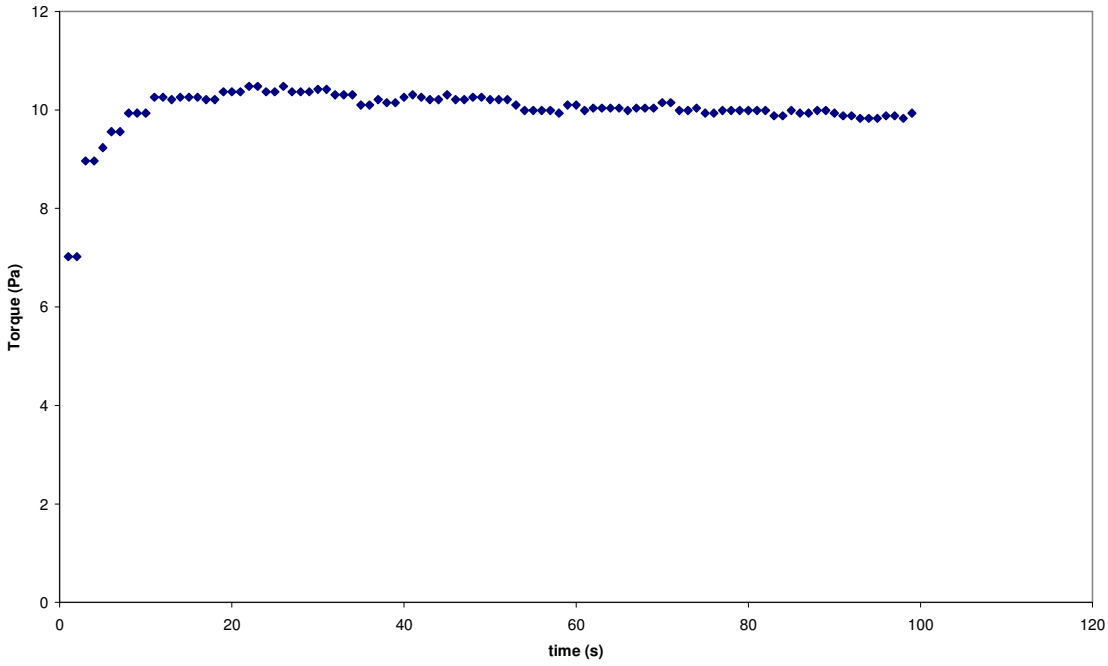


Figure 6.7: Torque versus time for washer 20 at 300 g/L

As shown in Figure (6.7), the maximum torque for a concentration of 300 g/L in washer 20 was approximately 10.5 Pa.

Once maximum torque is found, the suspension yield stress can be obtained using the vane dimension and the Equation (6.11).

Thus, in this case the yield stress is:

$$\tau_y = \frac{T_m}{\frac{\pi}{2} D_v \left(\frac{H}{D_v} + \frac{1}{3} \right)} = 2.4\, Pa$$

According to this, a relation between suspension yield stress as a function of concentration can be made in order to obtain Equation (6.10).

Table (6.3) below shows the results for red mud from washer 20 at different concentrations:

Table 6.3: Yield stress for washer 20 at different concentrations

Conc. (g/L)	Conc. (v/v)	Yield Stress(Pa)
200	0.0606	0.87
300	0.0909	2.4
400	0.1212	6.23

Considering the results obtained in Table (6.3), a fitting curve correlating the suspension concentrations with yield stress (Equation (6.10)) can be obtained by plotting yield stress (τ_y) versus concentration, v/v (φ) (Fig 6.7):

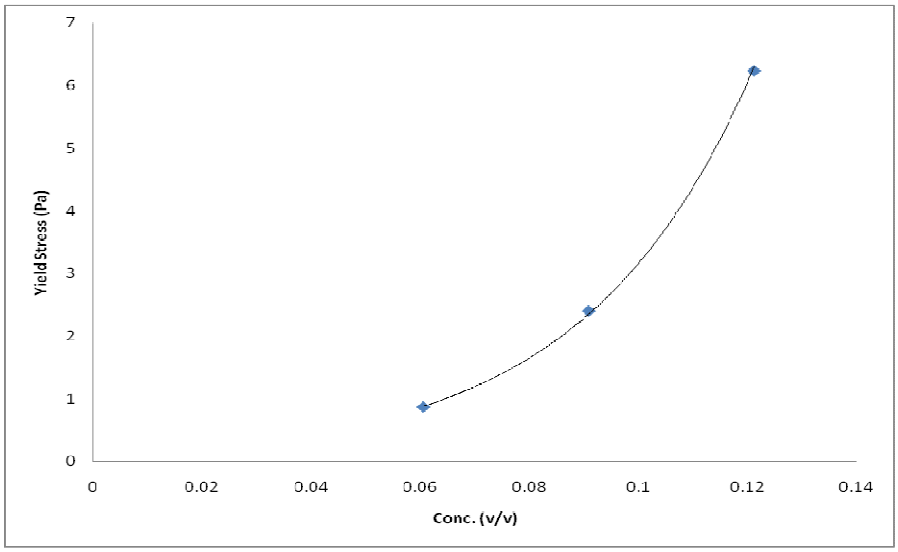


Figure 6.8: fitting graph Equation (6.10)

This graph allows the parameters of Equation (6.10) to be obtained, in which the slope of the line is β and the intercept is related with $\log\alpha$ .Equation (6.10) then takes the form:

$$\tau_y = 0.12 \cdot e^{32.55\cdot\varphi} \qquad (6.12)$$

The procedure described above is repeated for each washer in this study and the remaining results are tabulated in Table (6.4):

Table 6.4: Settling and rheology parameters

	Experimental Bürger parameters (eq. 6.10)		Settling parameters (eq. 6.5)	
Washer	α (Pa)	β	*n*	u_o *(cm/s)*
20	0.12	32.55	0.0065	0.96
50	0.0054	50.797	0.0062	0.93
70	0.0882	29.158	0.0024	0.82

VI.3. MATHEMATICAL MODEL SOLUTION:

For the solution of the convection-diffusion law (4.2), a finite difference method of the upwind type is used for the discretization of the PDE:

$$\frac{\varphi_j^{n+1}-\varphi_j^n}{\Delta t}+q(n\Delta t)\frac{\varphi_{j+1}^- -\varphi_{j-1}^+}{\Delta t}+\frac{f^{EO}(\varphi_j^n,\varphi_{j+1}^n)-f^{EO}(\varphi_{j-1}^n,u_j^n)}{\Delta z}=\frac{A(\varphi_{j+1}^n)-2A(\varphi_j^n)+A(\varphi_{j-1}^n)}{(\Delta z)^2} \qquad (6.13)$$

The so-called Engquist-Osher numerical flux (1981):

$$f^{EO}(\varphi_j^n,\varphi_{j+1}^n)=f^+(\varphi_j^n)+f^-(\varphi_{j+1}^n) \qquad (6.14)$$

is determined by:

$$f^+(\varphi)=f(0)+\int_0^{\varphi}\max(f'(s),0)ds \quad ; \quad f^-(\varphi)=\int_0^{\varphi}\min(f'(s),0)ds \qquad (6.15)$$

The upwind explicit discretization is stable under the following stability condition (CFL: Courant-Friedrichs-Lewy condition):

$$CFL := \max|f'(\varphi)|\frac{\Delta t}{\Delta z}+2\max|a(\varphi)|\frac{\Delta t}{(\Delta z)^2}\le 1 \qquad (6.16)$$

This solving method is of first order and can present problems in smooth regions. As an alternative, an upgrade of the Engquist-Osher scheme using the generalized MUSCL extrapolation (*Monotonic Scheme for Conservation Laws*) to extend the accuracy to second order is used (Toro, 1997; Evje & Karlsen, 2000). To this end, a linear function $\varphi^n(z)$ is introduced and defined by:

$$\varphi^n(z) = \varphi_j^n + s_j^n\left(z - z_j\right) \quad , z \in \left]z_{j-1/2}, z_{j+1/2}\right[\tag{6.17}$$

Where s_j^n is a slope at u^n. At $s_j^n = 1$ the reconstruction is linear and the error is $O(\Delta z)^2$. At $s_j^n = 0$ the reconstruction is in constant parts and the error is $O(\Delta z)$. Slope limiters have to be used in order to force monotony of the reconstruction. In this case, a $\theta_L - limiter$ is used (Sweby, 1984; Toro, 1997).

$$s_j^n = MM\left\{\theta_L \frac{\varphi_j^n - \varphi_{j-1}^n}{\Delta z}, \frac{\varphi_{j+1}^n - \varphi_{j-1}^n}{2\Delta z}, \theta_L \frac{\varphi_{j+1}^n - \varphi_j^n}{\Delta z}\right\}, \theta_L \in [0,2] \tag{6.18}$$

Where MM is a *Min-Mod* function defined by:

$$MM(a,b,c) = \begin{cases} \min(a,b,c) & if\ a,b,c > 0 \\ \max(a,b,c) & if\ a,b,c < 0 \\ 0 & Otherwise \end{cases} \tag{6.19}$$

The values to the borders are now extrapolated:

$$\varphi_j^L = \varphi_j^n - \frac{\Delta z}{2} s_j^n \qquad ; \qquad \varphi_j^R = \varphi_j^n + \frac{\Delta z}{2} s_j^n \tag{6.20}$$

The second order upwind takes the form:

$$\frac{\varphi_j^{n+1} - \varphi_j^n}{\Delta t} + q(n\Delta t)\frac{\varphi_{j+1}^L - \varphi_j^R}{\Delta z} + \frac{f^{EO}(\varphi_j^R, \varphi_{j+1}^L) - f^{EO}(\varphi_{j-1}^R, \varphi_j^L)}{\Delta z} = \frac{A(\varphi_{j-1}^n) - 2A(\varphi_j^n) + A(\varphi_{j+1}^n)}{(\Delta z)^2} \tag{6.21}$$

The border condition (4.8) is expressed as:

$$\frac{\varphi_{No}^{n+1}-\varphi_{No}^{n}}{\Delta t}+\frac{\Psi(n\Delta t)-q(n\Delta t)\cdot\varphi_{No}^{n}}{\Delta z}-\frac{f^{EO}(\varphi_{No-1}^{n},u_{jNo}^{n})}{\Delta z}=\frac{A(\varphi_{No-1}^{n})-A(\varphi_{No}^{n})}{(\Delta z)^{2}} \quad (6.22)$$

And the condition (4.9) can be discretized:

$$\frac{\varphi_{0}^{n+1}-\varphi_{0}^{n}}{\Delta t}+q(n\Delta t)\frac{\varphi_{1}^{n}-\varphi_{0}^{n}}{\Delta z}+\frac{f^{EO}(\varphi_{0}^{n},u_{1}^{n})}{\Delta z}=\frac{A(\varphi_{1}^{n})-A(\varphi_{0}^{n})}{(\Delta z)^{2}} \quad (6.23)$$

The method described above can be summarized using the following algorithm:

i. Initial conditions:

φ_0 : Initial concentration

φ_C : Critical concentration

ρ_S, ρ_f : Solid and liquor density

ii. Parameter for functions f_{bk} (Equation (6.4)) and τ (Equation (6.10))

iii. Domain length (height of the vessel L)

iv. Number of mesh points in order to determine Δz

v. Time step from CFL condition (Equation (6.16))

$$\Delta t=\frac{CFL\cdot\Delta z^{2}}{\Delta z\cdot\max\left|f'(\varphi)\right|+2\max\left|a(\varphi)\right|}$$

vi. Discretization of the boundary condition using Equation (6.22) and (6.23)

vii. Application of upwind scheme for interior cells using Equation (6.21) with slope limiters (6.18).

Figure (6.9) shows the concentration profile obtained using the previous algorithm:

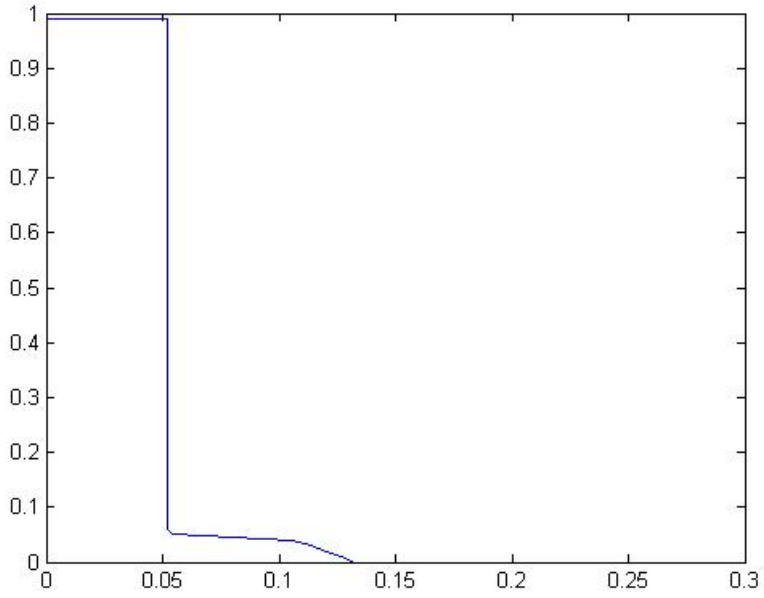


Figure 6.9: height of interface as a function of concentration (W20)

VI.3.1. STEADY STATE SOLUTION:

A particularly important solution of Equation (4.2) is the steady state condition, which describes the behaviour of operating industrial thickeners. At steady state, Equation (4.2) can be expressed as:

$$\frac{d}{dz}\left(q\cdot\varphi+f_{bk}(\varphi)\right)=\frac{d}{dz}\left(-\frac{f_{bk}(\varphi)\cdot\sigma_e'(\varphi)}{\Delta\rho\cdot g\cdot\varphi}\cdot\frac{d\varphi}{dz}\right) \tag{6.24}$$

and it can be integrated to obtain the solution:

$$q\cdot\varphi+f_{bk}(\varphi)=-\frac{f_{bk}(\varphi)\cdot\sigma_e'(\varphi)}{\Delta\rho\cdot g\cdot\varphi}\cdot\frac{d\varphi}{dz}+K \tag{6.25}$$

The constant *K* can be evaluated using Equation (9) at the bottom of the vessel, where the total flux must be referred to the convective part $q\cdot\varphi_D$. The concentration profile in the consolidation zone acquires the following form:

$$\frac{d\varphi}{dz}=-\frac{\Delta\rho\cdot g\cdot\varphi}{\sigma_e'(\varphi)\cdot f_{bk}(\varphi)}\left(q\cdot\varphi+f_{bk}(\varphi)-q\cdot\varphi_D\right)\quad,\quad z>0 \tag{6.26}$$

According to hydrodynamics stability, for sedimentation to take place, $du/dz \leq 0$ throughout the compression zone where $u_C \leq u \leq u_D$ and the inequality:

$$q \cdot \varphi + f_{bk}(\varphi) \leq f_f \,, \tag{6.27}$$

must be attained (Bustos *et al.* 1999). In the previous condition f_f is the feed flux and, in a continuous thickener, this value can be expressed as:

$$f_f = \frac{Q_f \cdot \varphi_f}{A} \tag{6.28}$$

Several steady-state stages can be obtained by properly manipulating q and f_f in Equation (6.27).

An alternative form of the Equation (6.26) is:

$$\frac{dz}{d\varphi} = -\frac{\sigma_e'(\varphi) \cdot f_{bk}(\varphi)}{\Delta\rho \cdot g \cdot \varphi \cdot \left(q \cdot \varphi + f_{bk}(\varphi) - q \cdot \varphi_D\right)} \tag{6.29}$$

Equation (6.27) is evaluated until critical concentration is reached, at which the height of the consolidation zone is encountered.
In order to test the model obtained above (Equation (6.29)), red mud samples were collected from different washers (20, 50 and 70). The red mud was a typical underflow sample from the washers.

The suspension (underflow washer 20) characteristic equations identified by Equations (6.4) and (6.12), whose parameters were obtained according to the method stated in sections VI.2.1.1 and VI.2.1.2, are then introduced into the model.

This model can be solved using the following algorithm developed in Matlab:

i. Determine input variables and parameter of Equations (6.4) and (6.12): density of solid and liquid, washer dimensions, underflow concentration, gel point or critical concentration, which can be found either by bisection or by approximation using:

$$\varphi_C \approx \frac{\varphi_0 \cdot z_0}{z_f} \tag{6.30}$$

where φ_0, z_0, z_f, are the initial concentration, initial and final height of the suspension (batch test).

ii. Find the concentration in the hindered settling zone using Equation (6.27), locating the root of the equation.

iii. In order to solve the differential equation, a Runge Kutta of fourth order is applied, selecting the step size of the independent variable φ.

iv. The function is evaluated at every step from the bottom concentration until critical concentration is reached.

The results showing the concentration profile inside the thickener are illustrated in Figure (6.10):

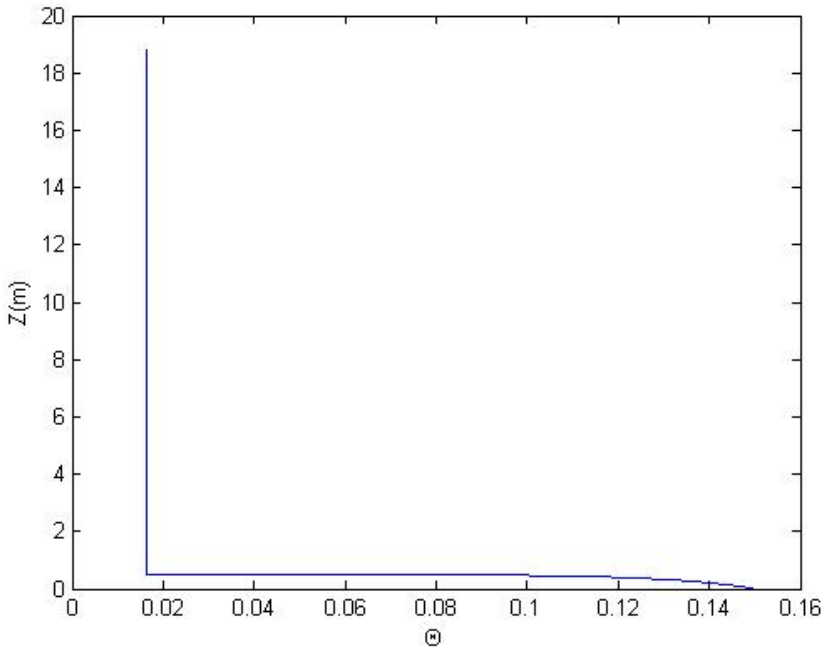


Figure 6.10: height of interface as a function of concentration (W20) – Steady State

The height of the heavy mud level for different underflow concentrations can be observed in Table (6.5) and Figure (6.11):

Table 6.5: Mud level at different underflow concentrations

Densities		Settling parameters		Rheology parameters		Critical Conc.	Feed flux	U/F Conc	Height
Liq (kg/m3)	Sol (kg/m3)	uo (cm/s)	C	α (Pa)	β	(v/v)	(t/h)	(v/v)	(m)
1100	3000	0.96	0.0065	0.12	32.55	0.1	51.25	0.171	1.5
							112.83	0.174	1.815
							49.87	0.177	2.130
							85.92	0.182	2.835
							82.08	0.183	2.020
							97.35	0.187	3.750
							95.92	0.187	3.750
							58.79	0.191	4.430
							11.43	0.191	4.686
							79.16	0.197	6.560

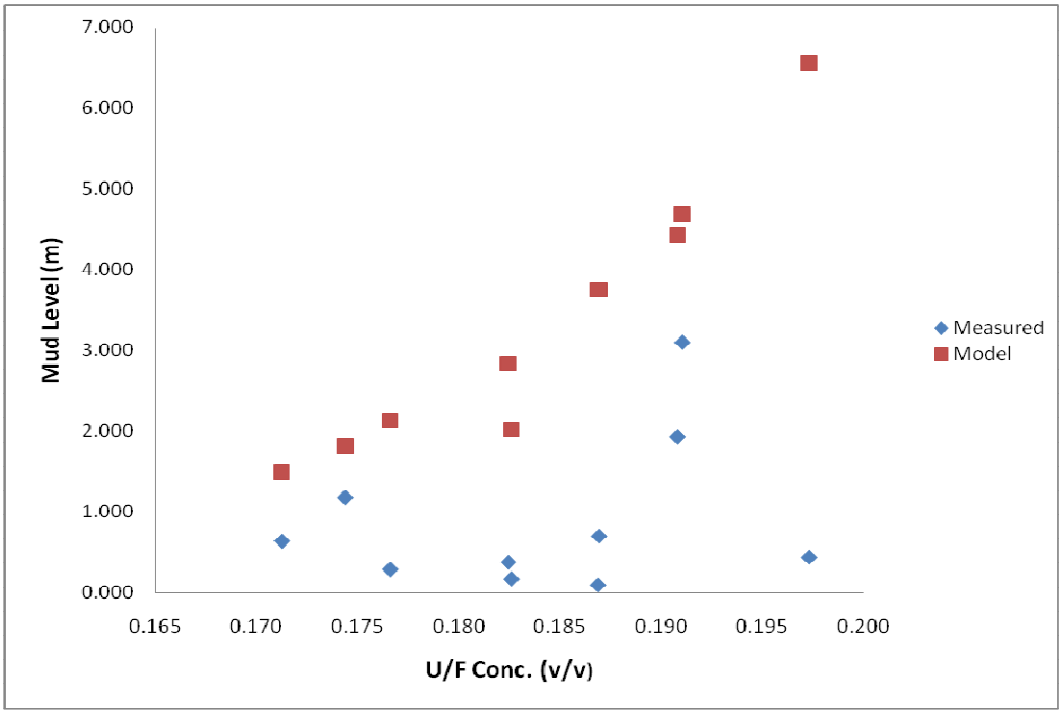


Figure 6.11: Mud level data comparison between measured data and extended Kynch data

The results predicted by the simulation can be seen to be a poor fit to the measured data. For instance, in the case of an underflow concentration of $\varphi_D = 0.177$, the measured value of the mud level is 0.29 m, whereas the predicted value is 2.13 m.

After considering the difference between measured and predicted data, other methods for measurement of the suspension properties were used. These are based on the dewaterability theory (Landman *et al*, 1988) and further treatment (Usher, 2002). These models for compressive yield stress and hindered settling function have the following form:

$$P_y(\varphi) = \left\{1 - \left(\frac{\varphi_C}{\varphi}\right)^a e^{b\cdot\varphi^c + d}\right\} \tag{6.31}$$

$$R(\varphi) = A \cdot \varphi^B + C \tag{6.32}$$

where *a, b, c, d* and *A, B, C* are parameters which can be found experimentally.
The dewaterability theory describes the sediment height in terms of shape of the bottom section of the thickener. This can be described as a differential equation between height of the heavy mud level and concentration as independent variable:

$$\frac{dz}{d\varphi} = \frac{P_y'(\varphi)}{\{R(\varphi)/(1-\varphi)^2\}\{q/\alpha(z)\}\{1-\varphi/\varphi_D\} - \Delta\rho g\varphi} \tag{6.31}$$

Again, this equation is solved from the bottom ($\varphi = \varphi_D$) until critical concentration is reached.

The solution to differential Equation (6.31) is showed in Figure (6.12), where a concentration profile is plotted:

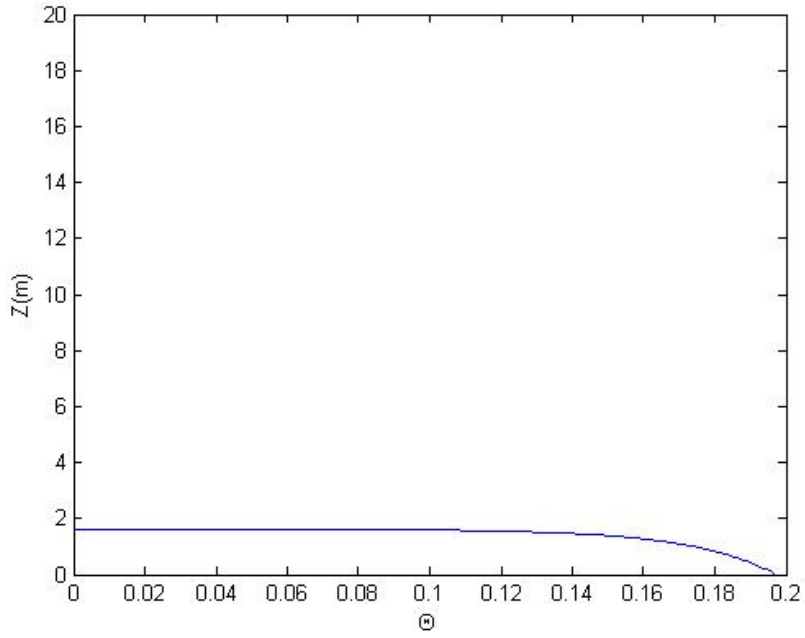


Figure 6.12: Concentration profile – dewaterability model (Landman, 1988)

Once more, model results are plotted against measured values to observe the discrepancies between both sets of the data (Figure (6.13)):

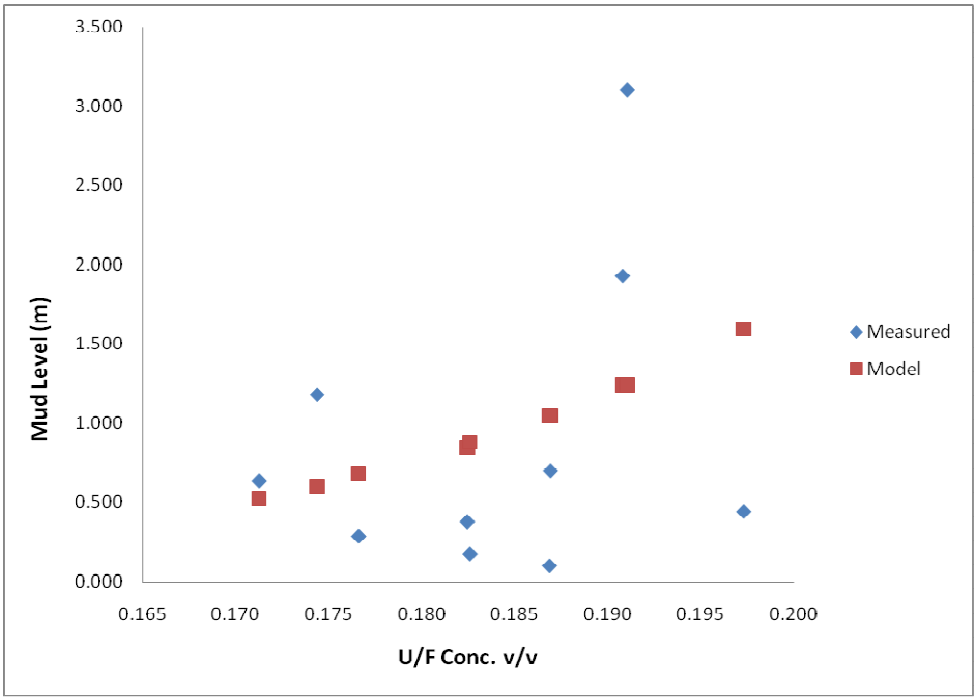


Figure 6.13: Comparison between mud level plant values and model data (Eq. 6.31)

Simulation outputs from the model predict values closer to the measured data, indicating the influence of rheology parameters. However, the fit is still not very good. Thus, these parameters need to be determined using a different method, such as filtration. There may be several reasons for the discrepancy between measured and predicted data. One of these reasons is that the effect of rakes was not considered in the mathematical model (raking action in the bottom of the vessel generates stress). The rakes move the suspension through the discharge point in the bottom of the vessel, but also the effect of the rakes help to create a more compact bed in the compression zone, obtaining a value of mud level lower than simulation result. This effect may be overcome introducing into the model a relationship between the torque of the rakes and slurry yield stress.

On the other hand, the range of concentrations used to evaluate the suspension properties was low, so it did not cover the entire range of concentration that may be present in the vessels.

Chapter VII

NEURAL NETWORK IDENTIFICATION

VII.1. INTRODUCTION

Modelling and simulation have become powerful tools for design, analysis and optimization systems and industrial processes. The availability of personal computers that are continuously becoming more powerful, lower in cost and easy to use, accompanied by software or application programs and programming languages with high flexibility, has allowed the mass use of different techniques of simulation and process control.

Unlike the method discussed in the previous chapter, which describes the representation of a system based on a mathematical model, this part of the work will propose the utilization of computational tools based on historical data in order to obtain an instrument for the prediction of control variables. In order to achieve this, historical data were collected from Rio Tinto's database and a multilayer perceptron was trained to learn the relationship between input and the target variables.

VII.2. DEVELOPMENT OF THE NEURAL MODEL

For the neural network design, numeric values of the variables involved in the washing process from Rio Tinto's data base were considered. The data set available was part of the first ten months of 2009, randomly divided into two groups: 80% for training and 20% for validation and testing.

The development of the neural network can be defined by the following steps:

- Variable selection
- Normalization
- Architecture
- Training and testing of the network

VII.2.1. VARIABLE OR DATA SELECTION

Working continuous data come from the washer train and belong to washers 10 and 60. Figure (7.1) illustrates the vessel (washer 10) and the process variables taken into account for the design of the neural network.

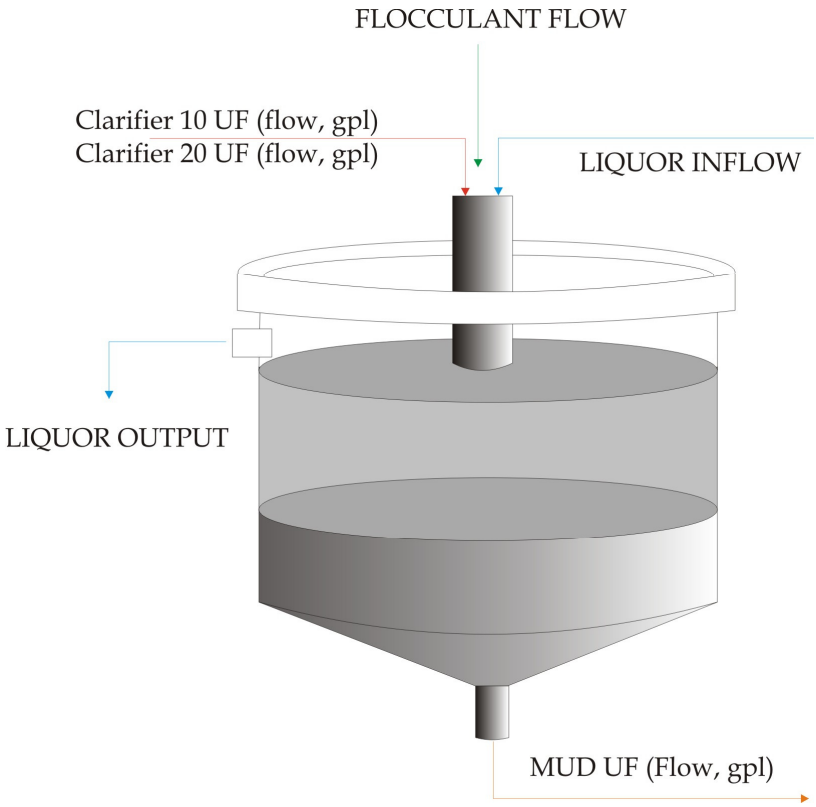


Figure 7.1: Washer vessel with operating variables

The variables can be arranged in a table in order to show the relationships between the input variables and the target. The number of rows depends on the number of relations between variables presented. Table (7.1) presents input and output variables for washer 10.

Table 7.1: Variables relation

	NETWORK INPUT VARIABLES										NET. OUTPUT
Input No.	Clar. 10 UF m3/h	Clar. 10 UF gpl	Clar. 20 UF m3/h	Clar. 20 UF gpl	Liq. In m3/h	Filt. Liq m3/h	Flocc. m3/h	Liq. Out m3/h	Was UF m3/h	Was UF gpl	Mud lev. m
1											
…											
n											

Values belonging to the database (29180 readings) correspond to data collected by the sensors every 15 minutes. Because of the large number of data available, this number was reduced to 139, which corresponds to daily average values. In addition, only values corresponding to a normal operation of the vessel were accepted (vessel in DUTY was rejected).

Figures (7.2) and (7.3) show, in detail, the behaviour of some of the variables that were used as an input pattern for network training.

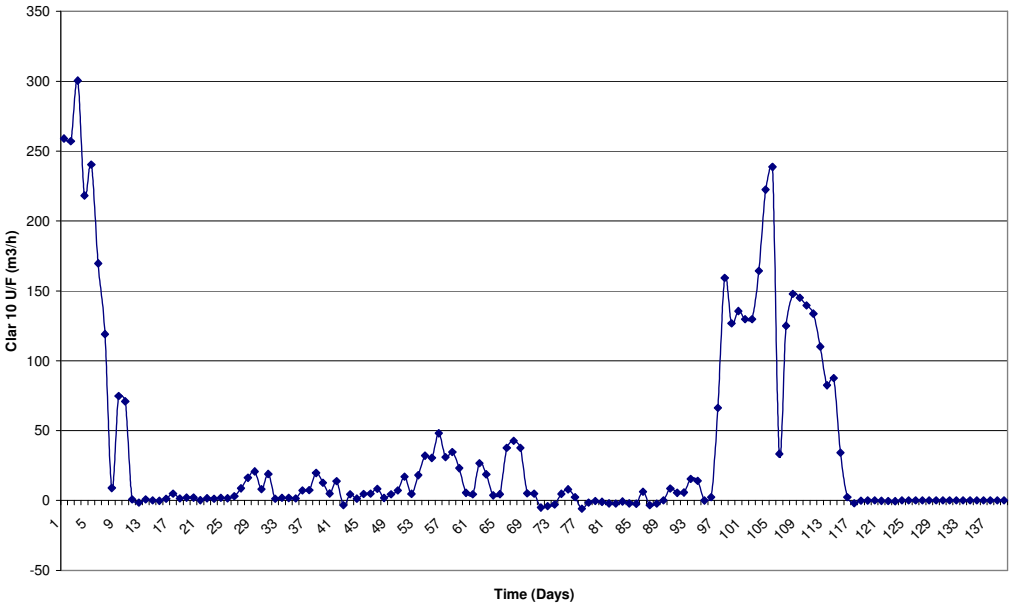


Figure 7.2: Clarifier 10 U/F – Rio Tinto Yarwun (2009)

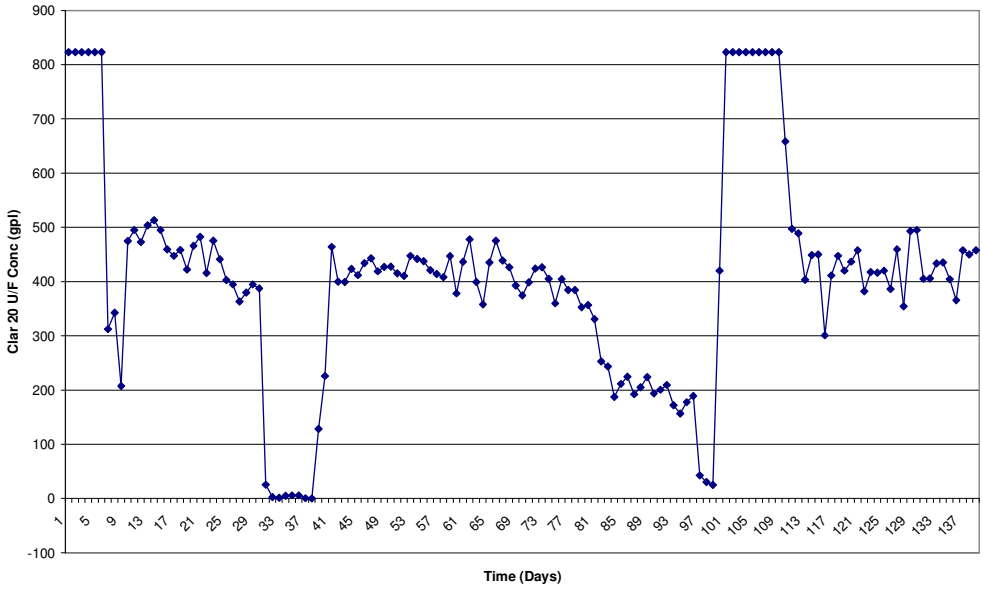


Figure 7.3: Clarifier 20 U/F conc. – Rio Tinto Yarwun (2009)

VII.2.2. PREPARATION OF THE DATA

Once process variables were identified, preparation of the data was performed in order to train the network. These variables are those measured daily by sensors in the plant, which will be used as input and output patterns for both training and testing of the network.

VII.2.2.1. DATA NORMALIZATION

It is useful to normalize the input data before submission to the neural network. This normalization is used to reduce the range of the set of values, which become appropriate for the transfer or activation function that will be used. In general, the data are normalized between 0 and 1 (or -1 and 1) to prevent the saturation effect of the transfer function.

This process will help the network to learn a greater number of relationships between inputs and outputs, avoiding uncertainties that would be induced when input values, due to their size, cannot be processed. The real value of the output variables will be obtained by inverting the normalization process applied to each of them.

In the standardization process, the data are transformed in order to obtain a zero mean and standard deviation equal to 1 according to the equation:

$$x' = \frac{x - \bar{x}}{\sigma_P} \tag{7.1}$$

Where x' are the standardized data, x are the original values, $\bar{x}$ is the mean of the values and σ_P the standard deviation.

VII.2.3. ARCHITECTURE

As specified above, the architecture of the network is related to the organization and arrangement of neurons forming layers. Thus, the basic parameters of the network are: number of layers, number of neurons per layer, learning algorithms and activation function.

The number of input and output layers of the network is dictated by the number of inputs and outputs considered. However, the number of neurons in the hidden layer is a key point in the design strategy of the network. If a network is designed with a small number of neurons, it will not reproduce the dynamics of the system and therefore will not provide an accurate forecast. On the other hand, excessive numbers of neurons can provide an appropriate behaviour but with high computational cost. Knowledge from non-learned patterns will also not be generalizable.

Considering the number of variables of the washer system (see VII.2.1 – washer 10) the following consideration was opted:

- Feed forward backpropagation neural network
- One input layer formed by 10 neurons
- One hidden layer formed by 9 neurons with activation function sigmoid.
- One output layer formed by one neuron with activation function *Tansig*.

The network architecture considered, which contains the consideration above stated, is shown in Figure (7.4).

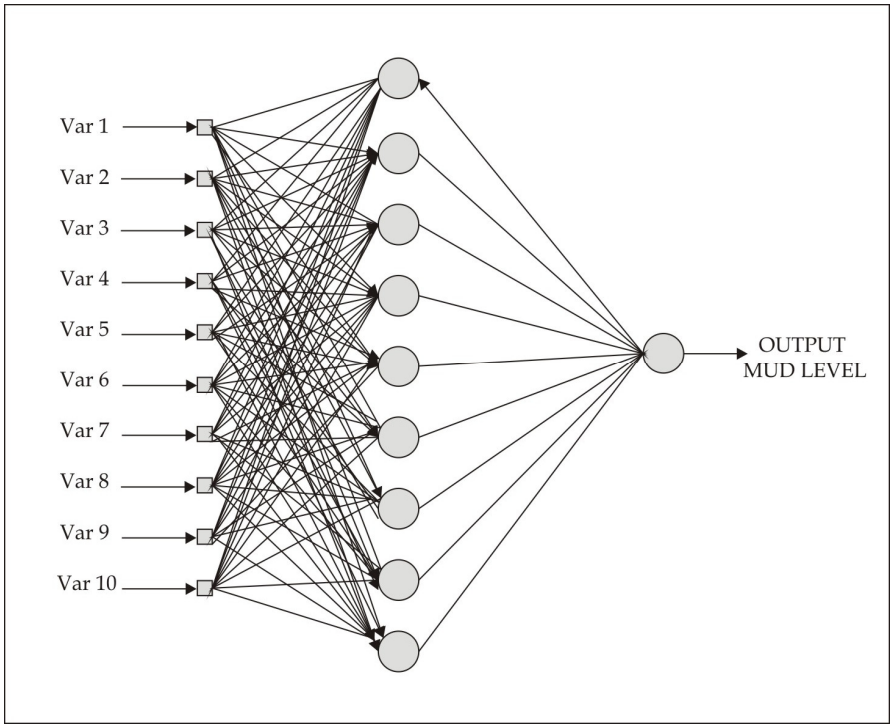


Figure 7.4: Neural network model

The selection of an ideal activation function, as well as, layers in the hidden layers depends of the nature of the problem and the researcher discretion, that in most of occasions include trial and error. In literature, standard criteria for the selection of the "correct" activation function cannot be found, and its choice would depend of the network response in comparison with the desired output. In most of cases, sigmoid functions are used because is a monotonous crescent function which exhibits a good balance between a linear and non-linear behaviour. On the other hand, problems that require more than one hidden layer are rarely found and there is not theoretical reason for that matter. In fact, using more than one hidden layer can affect the convergence to local minima, introducing more complexity which rarely improves the network results.

VII.2.4. TRAINING AND TESTING

The training algorithms selected are the most commonly used. These algorithms can be summarized by the following functions:

- *Quasi-Newton:*

  The threshold and weights are updated according to the following algorithm:

  $$x_{k+1} = x_k - H_k^{-1} \cdot \nabla_k \tag{7.2}$$

  where H is the Hessian matrix of the performance index at the current values of the weights and biases, and ∇_k is the gradient.

- *Levenberg-Marquardt:*

  This algorithm is one of the fastest for training of the neural network, but it may require more memory than do other algorithms.

  $$x_{k+1} = x_k - \left(J^T J + \mu \cdot I\right)^{-1} \cdot J^T e \tag{7.3}$$

  where J is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, ***e*** is a vector of network errors and μ is a scalar which can be zero (Newton method) or larger (gradient descent).

- *Gradient descent with momentum backpropagation:*
  This is a descent steps algorithm, which updates weights and gains in the negative direction of the gradient error function.

  Momentum allows a network to respond not only to the local gradient, but also to recent trends in the error surface.

The training algorithm is used to find weights that minimize a global measure of the error. This can be called performance function, which can be described as a Mean Square Error (MSE):

$$E = \frac{1}{2}\sum_{i=n+1}^{N}(y_i - a_i)^2 \tag{7.4}$$

According to the training algorithm stated above, the Levenberg-Marquardt was selected due to the speed in obtaining an acceptance MSE value, which ensures a maximum accuracy during the learning process.

All the design process, training and testing of the network were developed using Matlab, which possesses a toolbox for that purpose.

Starting from a random set of initial weights, and applying the Levenberg-Marquardt training algorithm using the standard input/output values (see Table (7.1)) of the variables of the final database (139 values per variable), a convergence process is needed, which will be achieved when the error reaches a value of 10^{-10}.

For washer 10 data, the error proposed was found after 2422 iterations, as shown in the plot below:

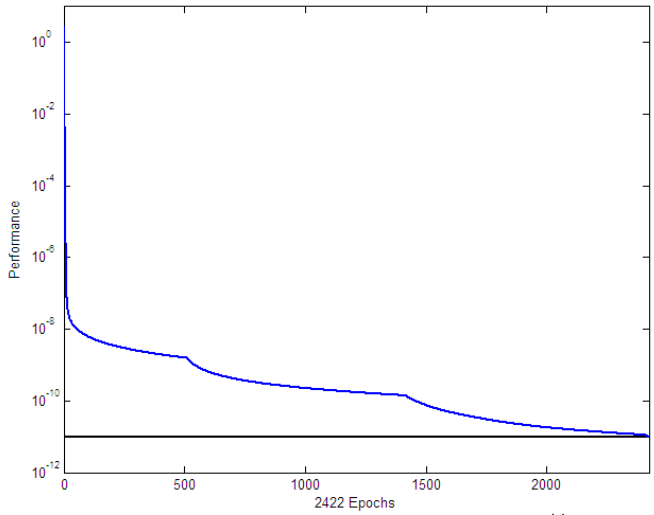


Figure 7.5: Error convergence. $E=10^{-11}$

After arriving at a satisfactory network design, this was tested with 20% of the data reserved, allowing assessment of the ability of the neural model in describing the phenomenon. The results are shown in Figure (7.6).

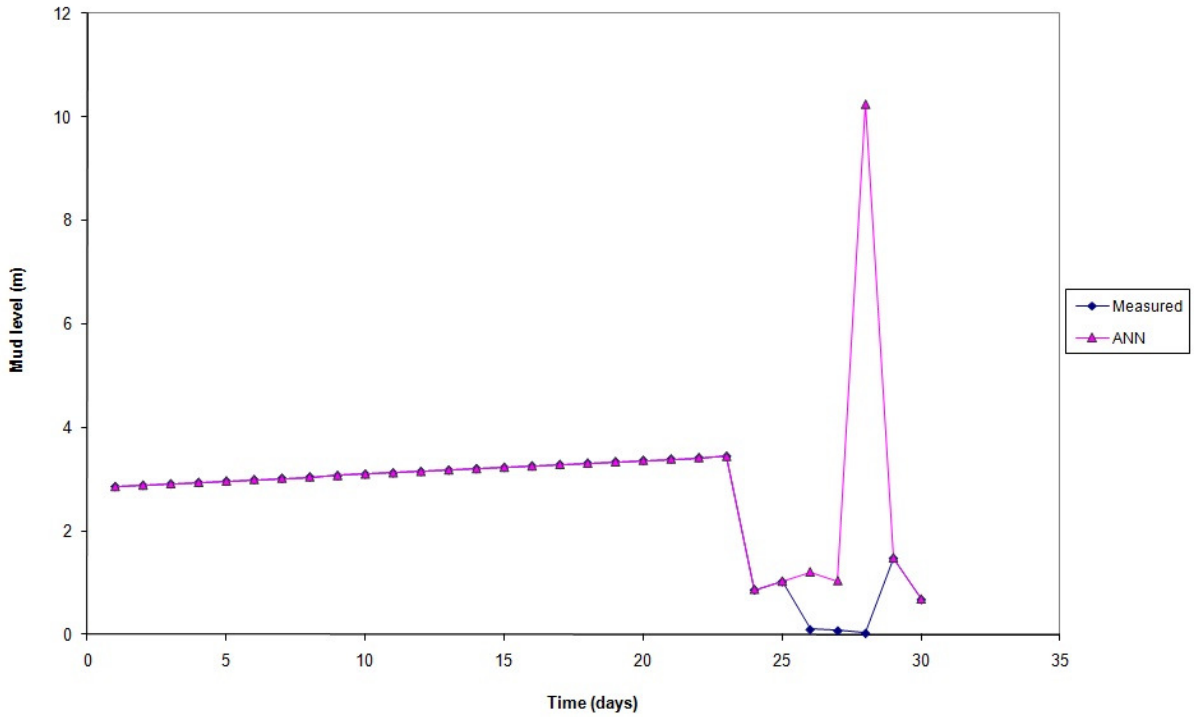


Figure 7.6: Mud level values for washer 10 (measured data - neural network result)

The Figure above shows a comparison between measured data and those obtained by the neural network in a 30 day period. In the first 25 days, the network response produced a good agreement with the measured data from washer 10. However, this trend is different for the next four days, having a significant difference at day 28. This

can be explained by the fact that there were insufficient training data in this specific range (0.02 m – 1.4 m approximately), so the network did not have enough information to achieve a better generalization. A good generalization occurs when training data are sufficiently representative of the domain to which they belong.

Similarly, a neural network was employed for identification of washer 60. Three months of data were collected.

The data, again, were separated into two categories, training data (60% of total) and testing data (40%). In this case, different topologies were tested in order to find the one most appropriate for the data under consideration.
First, a network with one hidden layer of 10 neurons and *Tansig* transfer functions for both hidden and output layer was considered. According to this configuration, a performance graph is shown in Figure (7.7):

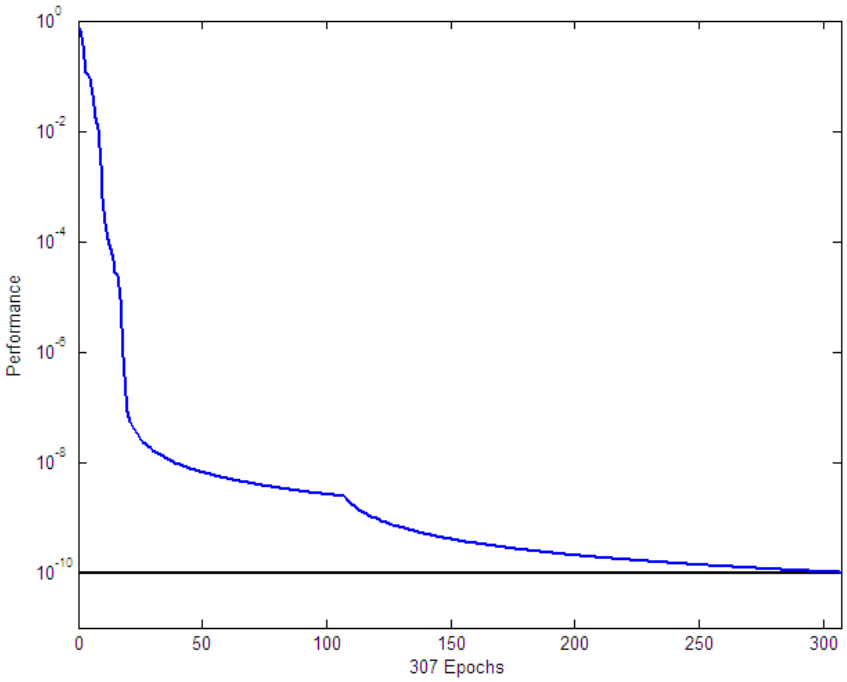


Figure 7.7: Washer 60 error performance

Figure (7.7) shows that the target error was reached after about 300 iterations, and whereas the regression graph (Figure (7.8)) shows a poor fit for the response. In order to increase the degree of fit, different approaches were used.

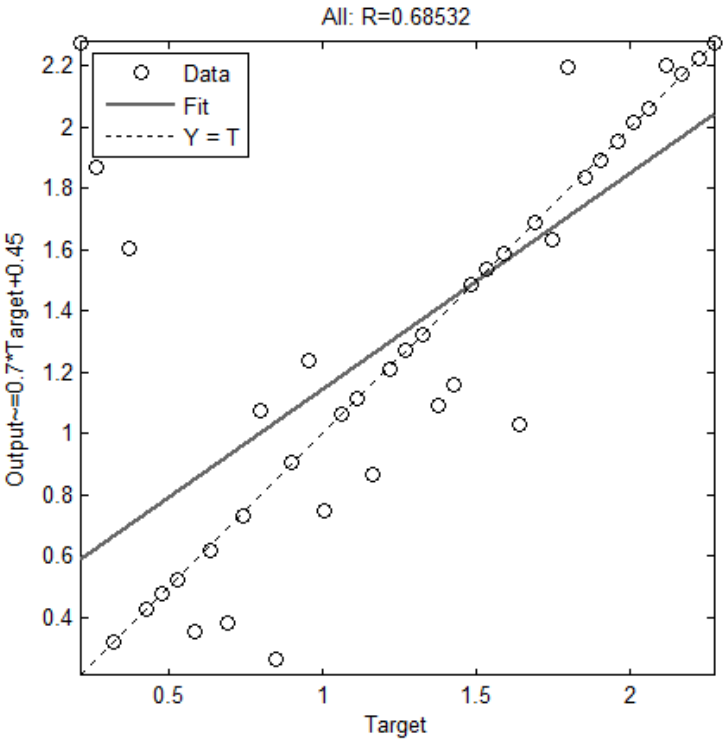


Figure 7.8: Washer 60 regression graph

In the new approach, the connection weights were reset, and the network was trained again. The corresponding regression is presented in Figure (7.8):

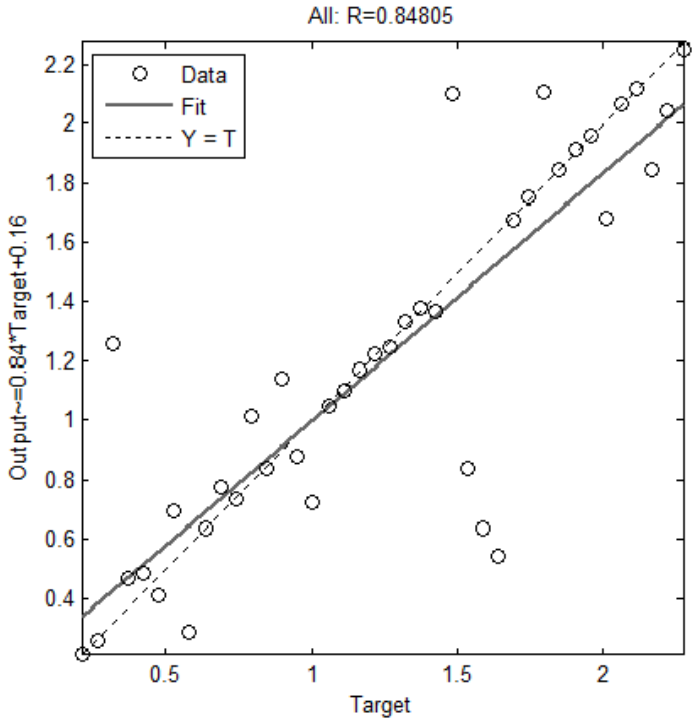


Figure 7.9: Washer 60 regression graph (2nd run)

Figure (7.9) shows that the correlation was increased to 0.84. It was of interest to see if this could be improved, so the number of hidden neurons was increased from 10 to 12 to watch the network response.

After several weight initializations, the new network reached the best regression value achievable, according to Figure (7.10):

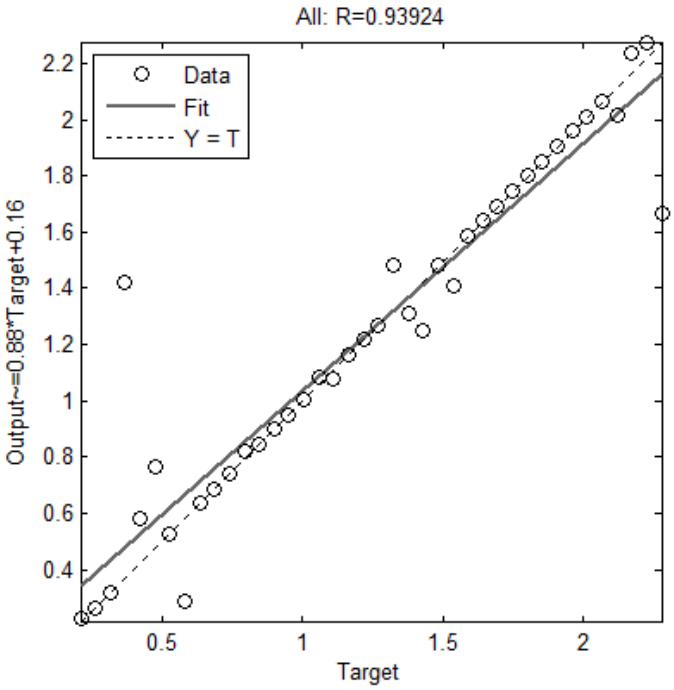


Figure 7.10: Washer 60 regression graph (3rd run)

Once this point had been reached, the network was tested with 40% of the remaining data, as shown in Figure (7.11):

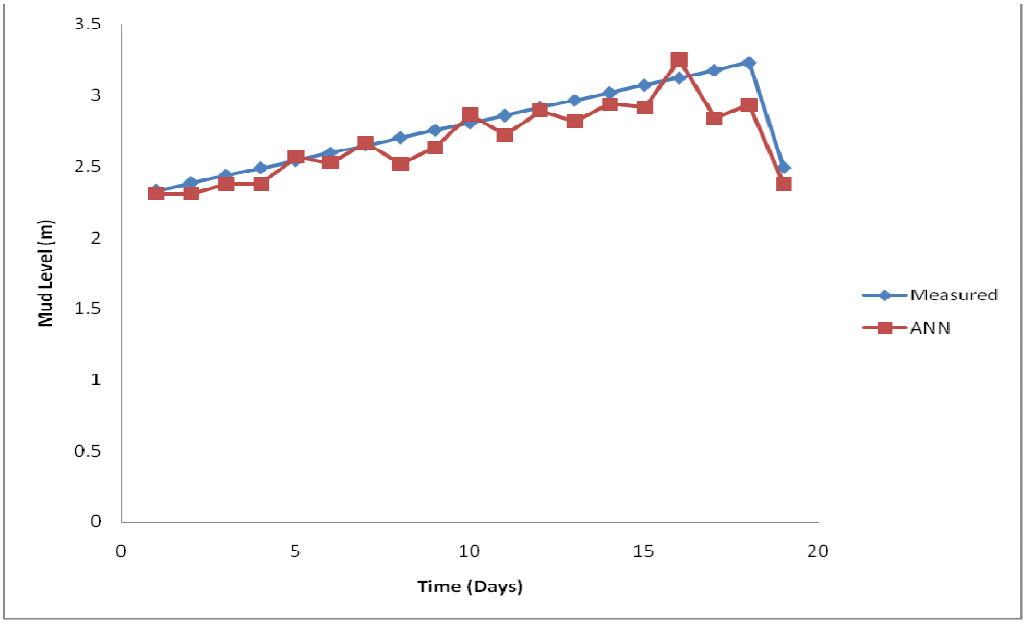


Figure 7.11: Mud level values for W60 (reference data - neural network result)

Figure (7.11) shows that the predicted data fit the measured data quite well, which indicates that the number of neurons in the hidden layer has increased the accuracy of the network. However, the neurons should not exceed a number that permits over-learning (Haykin, 1999). Although over-learning achieves a high rate of success, this is reached at the expense of ability to generalize, since the network extracts too much information and focuses on peculiarities from the training set, forgetting relevant information for a more general case.

SECTION D:
SUMMARY OF RESULTS, CONCLUSIONS AND FUTURE WORK

Chapter VIII

OVERALL OUTCOMES

VIII.1. INTRODUCTION

This work was focused on understanding the sedimentation phenomenon that occurs in industrial equipment such as thickener/washers. Simulation of a continuous thickener was made in order to obtain the concentration profile and height of the heavy mud level, which is an important controlled variable of the process. Mathematical modelling and neural networks were chosen as simulation methods.

The mathematical models used in this work were tested using actual measured data from Rio Tinto. These models require obtaining parameters for solids flux density function (or hindered settling) according to the Richardson-Zaki model (Richardson *et al* 1997) and effective solid stress indentified by the Bürger empirical Equation (Bürger *et al*, 1999). Experimental work was conducted in order to obtain these parameters.

The use of computational techniques as an alternative for process identification was adopted using historical data from Rio Tinto's data base. Two neural models were developed for washers 10 and 60.

This chapter presents the major outcomes obtained in this work, as well as general conclusions based on results and observations. Finally, recommendations for further work will be presented.

VIII.2. SUMMARY OF RESULTS

The following are the major findings of this research:

- During settling test (Figure (6.2)) the settling rate was verified as strongly dependant on the concentration of flocculant. Thus, a higher concentration of flocculant resulted in a faster settling rate.

- The Richardson-Zaki index found experimentally differs from theoretically predicted values (assuming spherical particles). This index (Richardson & Zaki,

1997) is said to take values greater than 1, which was not reflected experimentally. The initial concentration considered in a settling test should be as close as possible to the actual initial concentration of the process.

- The critical concentration obtained ($\varphi_C \approx 0.1$ at $Z_C = 17.8\,cm$) by the angle bisection method (Figure (6.4)) showed a good agreement with literature values.

- The relationship of yield stress as a function of volume concentration for red mud showed a good fit according to the empirical Bürger power law (Equation (6.10)).

- Dynamic simulation of a washer vessel based on Kynch theory of sedimentation and its extension to flocculated suspension allowed obtaining of the concentration profile inside the vessel.

- The steady state model of a continuous thickener predicted both the concentration profile and the height of the mud level. The model was very sensitive to the measured of solid stress function of the suspension. On the other hand, the model was weakly sensitive to the settling function.

- Simulation results, according to Figure (6.10), were significantly different from the measured data; further analysis based on dewaterability theory (Landman *et al*, 1988) illustrated in Figure (6.12) showed better results, but there were still differences with measured data.

- The concentration ranges used for rheology characteristic determination were low. Thus, extrapolation of the data was necessary in order to cover the entire range of concentration present in the washer operation.

- As an alternative to mathematical models, neural networks were developed for 2 different washers (washer 10 and 60). The architecture that reported the best results for washer 10 was formed by 10 neurons in the input layer, 9 neurons in the hidden layer and one neuron in the output layer. On the other hand, the

architecture chosen for washer 60 consisted of 10 neurons in the input layer, 12 neurons in the hidden layer and one neuron in the output layer.

- The Levenberg-Marquardt algorithm was chosen as a training algorithm due to its speed in obtaining the best mean square error.

- The neural identification for washer 10 (Figure (7.6)) showed a very good fit for about 85% of the data considered for network testing. The remaining 15% did not show the same behaviour. On the other hand, results for washer 60 (Figure (7.11)) showed a good agreement with measured data for a 19 days period. The differences presented between predicted and measured data in both cases can be explained as being due to network generalization errors.

VIII.3. CONCLUSIONS

In this thesis, the identification of the concentration profile and the height of the heavy mud level using two simulation methods (mathematical based simulation and neural networks) are predicted. The outcomes of this work lead to the following conclusions:

The disagreement between measured and predicted data, according to model based simulation, is likely due to errors in the determination of suspension properties. The range of concentrations used to determine both solid stress and hindered settling functions, was not sufficient. Thus, in order to cover the entire range of concentration inside the thickener, extrapolations of the data are needed at lower concentrations.

The mathematical model used to determine the heavy mud level and the concentration profile does not consider the effect of rakes. This could be one of the main reasons for poor agreement of data. Raking is used to remove sediments in the bottom of a settling vessel, leading the slurry towards a discharge point, but raking also influences the underflow concentration. The action of raking improves the sediment permeability by liberating bound water, thus creating a more compact structure on the bottom. The raking shear effect could increase underflow concentration.

The neural network approach shows better results in comparison with mathematical models. However, some considerations are needed in order to use this method for process prediction.

The quality, as well as the quantity, of the data has to be considered when a neural network is developed for process identification. Although the range of data was quite large for both washers, in the case of washer 60, these values are low in a certain range of operation. The input data for network training should contain sufficient information pertaining to the domain.

VIII.4. FUTURE WORK

As a consequence of the results found in this work, some recommendations for further work can be made in order to improve the mathematical model used, as well as data mining for process identification.

New experiments on settling and rheology should be conducted, increasing the range of concentrations. This would improve the accuracy of the hindered settling and effective solid stress functions. In addition, another method, such as pressure filtration, should be used to obtain a relationship for the compression zone (compressive yield stress). The effect of raking should be considered in a new model to observe its influence over the concentration profile and heavy mud level.

Once a better understanding of the washer vessel dynamics is achieved, a simulation of the whole washer train can be made. This would further increase the robustness of the model.

If a new set of data with a larger range was collected, the neural network learning process could be improved. New network architectures could be tested. These might improve the behaviour of the network and achieve a better fit to the measured data.

REFERENCES

Adorjan, L.A 1976, 'Determination of Thickener Dimensions from Sediment Compression and Permeability Tests', *Inst. Min. Met*., pp. 157-163.

Bacri J-C., Hoyos, M., Rakotomalala, N. Salin, N., Bourlion, M., Daccord, G., Lenormand, R., & Soucemarianadin, A 1991 'Ultrasonic diagnostic in porous media and suspensions', *J. Phys. France*, vol III, pp.1455

Beale, R 1990, 'Neural computing', *IOP Publishing*, London.

Bell, F.G., Donnelly, L.J 2006, *Mining and its impact on the environment*, Taylor & Francis, New York.

Botin, J.A 2009, *Sustainable management of mining operations*, Society for mining, Metallurgy and Exploration.

Bustos, M.C., Concha, F., Bürger, R., Tory, E.M 1999, 'Sedimentation and Thickening', *Kluwer Academic Publishing*, Dordrecht, The Netherlands.

Bürger, R., Evje, S., Karlsen, K.H., Lie, K 2000, 'Numerical methods for the simulation of the settling of flocculated suspensions', *Chem Eng J*, vol 80, pp 91-104.

Coe. H.S., and Clevenger G.H 1916. 'Methods for determining the Capacity of Slime Settling Tanks', *AIME*, pp. 356-385

Cybenko, G 1989, 'Approximation by superpositions of a sigmoidal functions', *Mathematics of Control, Signal and Systems* Nº 2, pp.303-314.

Davis R. H & Birdsell K. H 1988, 'Hindered settling of semidilute monodisperse and polydisperse suspensions', *AIChE J*., vol 34, pp. 123

Deboeck, G., Kohonen, T 2000, 'Visual explorations in finance with self-organizing maps', *Springer Finance*.

Engelbrecht, A.P 2008, *Computational intelligence an introduction*, 2nd edn, John Wiley & Sons, Chichester.

Engquist, B., Osher, S 1981, 'One-sided difference approximations for nonlinear conservation laws', *Mathematics of Computation*, vol. 36, pp. 321-351.

Evje, S., Karlsen, K.H 2000, 'Monotone difference approximation of BV solutions to degenerate convection-diffusion equations', *SIAM J. Num. Analysis*, vol. 37, no. 6, pp. 1838-1860.

Fogelman-Soulie, F., Gallinari 1998, *Industrial applications of neural networks*, World Scientific, Singapore.

Fitch E. B. and Stevenson D.G 1977, *Gravity Separation Equipment in Solid Separation Equipment Scale-Up*, D.B Purchas, Uplands Press Ltd., Croydon England.

Fausset, L 1994, *Fundamentals of Neural Networks. Architecture, Algorithms and applications*, Prentice-Hall, New Jersey.

Foust, A.S 1960. *Principles of unit operations*, Wiley, New York.

Frawley W., Piatetsky-Shapiro G., Matheus, C 1992, 'Knowledge Discovery in Databases: An Overview', *AI Magazine*, pp. 213-228.

Freeman, J.A, Skapura, D. M 1993, *Redes Neuronales. Algoritmos, aplicaciones y técnicas de programación*, Addison-Wesley/Diaz De Santos

Funahashi, K.I 1989, 'On the approximate realization of continuous mappings by neural networks', *Neural Networks*, Nº 2, pp. 183-192.

Gupta, A 2006, *Mineral Processing design and operation: An introduction*, Elsevier, Amsterdam.

Green, M.D 1997, 'Characterisation of suspensions in settling and compression', Ph.D thesis, Department of Chemical Engineering, The University of Melbourne, Melbourne.

Hartman, E.J., Keeler, J.D. & Kowalsky, J.M 1989, 'Layered Neural Networks with Gaussian hidden units as universal approximations', *Neural Computation*, Nº 2, pp. 210-215.

Hangos, K.M., Cameron, I.T 2001, *Process modelling and model analysis*, Academic Press, San Diego.

Hanselman, D.C., Littlefield, B 2005, *Mastering Matlab 7*, Pearson Prentice Hall, London.

Haykin, S.S 1999, *Neural networks: a comprehensive foundation*, 2 edn, Prentice Hall, Upper Saddle River, NJ.

Higham, D.J., Higham, N.J 2005, *Matlab guide*, 2nd edn, Society for industrial and applied mathematics, Philadelphia.

Hornik, K., Stinchcombe, M & White, H 1989, 'Multilayer feedforward networks are universal approximations', *Neural Networks*, Nº 2, pp. 359-366.

Hudson, L. K 1987, 'Alumina Production', *Production of Aluminium and Alumina*. A. R. Burkin (ed), J. Wiley, Chichester, West Sussex, pp. 11-46.

International Aluminium Institute, 2002, website www.world-aluminium.org

King, R.P 2002, *Introduction to practical fluid flow*, *Elsevier Science*, Boston.

Kynch G. J 1952, 'Theory of sedimentation', *Trans. Faraday Sot*, vol 48, pp 166

Landman, K.A., White, L.R., Buscall, R 1988, 'The continuous flow gravity thickener: steady state behaviour', *AIChE J.*, vol 34 (2), pp. 239–252.

Lottermoser, B.G, 2007, *Mine wastes: characterization, treatment and environmental impacts*, 2nd edn, Springer, London.

McCabe, W.L, Smith, J.C 2005, *Unit operations of chemical engineering*, 7th edn, McGraw-Hill, Boston.

Menzies, N. W., Fulton, I. M., & Morrell, W. J 2004, 'Seawater Neutralization of Alkaline Bauxite Residue and Implications for Revegetation', *J. ENVIRON. QUAL*, vol 33, pp. 1877-1884

Minsky, M., Papert, S 1969, 'Perceptrons'. Cambridge University, Massachusetts pp. 226.

Paradis, R.D 1992, 'Disposal of red mud using wet stacking technology', *Proc International Bauxite Tailings Workshop*, Perth, Western Australia.

Parekh, B.K, Goldberger, W.M 1976, 'An assessment of technology for possible utilization of Bayer process muds', *U.S Environmental Protection Agency*.

Perry, R.H, Green, D.W, Maloney, J.O 1997, *Chemical's engineers' handbook*, McGraw Hill, New York.

Pryor, E.J 1965, *Mineral processing*, 3rd edn, Elsevier, Amsterdam.

Richardson, J.F., Coulson, J.M., Harker, J.H. & Backhurst, J.R 2002, *Chemical Engineering Volume 2: Particle Technology and Separation Process*, 5th edn, Heinemann-Butterworth, Oxford.

Richardson, J.F., Zaki, W.N 1997, 'Sedimentation and Fluidisation: Part I', *Trans IChemE*, vol. 75

Rio Tinto Yarwun 2008, Alumina refinery, website www.riotinto.com

Rosenblatt, F 1958, 'The perceptron: A probabilistic model for information storage and organization in the brain', *Psychol. Rev.*, vol 65(3) pp. 386-408.

Sweby, P.K 1984, 'High resolution schemes using flux limiters for hyperbolic conservation laws', *SIAM Journal on Numeric Analysis*, vol. 21, no. 5, pp. 995-1011

Toro, E.F 1997, *Riemann solvers and numerical methods for fluid dynamics: a practical introduction*, Springer, New York.

Talmage, W.P., Fitch, E.B., 1955, 'Determining Thickener Unit Areas', *Ind. Eng. Chem.*, vol 47, pp. 38-48.

Whittaker, C.W., Armiger, W.H., Chichilo, P.P., Hoffman, W.M., 1955. 'Brown mud from the alumina industry as a soil liming material'. Soil Science Society Proceeding, pp. 288-292.

Williams, R.E 1975, *Waste production and disposal in mining, milling, and metallurgical industries*, Miller Freeman Publications, San Francisco.

Wills, B.A. 1997, *Mineral processing technology: an introduction to the practical aspects of ore treatment and mineral recovery*, 6th edn, Butterworth-Heinemann, Boston.

Yoshioka, N 1957., 'Continuous Thickening of Homogeneous Flocculated Slurries', *Soc. Chem. Eng. Japan*, vol 2, pp. 66-75.

APPENDIX A:
SETTLING TEST DATA

APPENDIX A.1: Settling test at different flocculant dosages (Washer 20)

Settling test for washer 20 at different flocculant dosages (initial solids concentration: 50g/L).

Table A.1: flocculant dosage: 50 gpt

MSlurry (g)	187.8	
MLiquor (g)	962.7	
FLOCCULANT DOSAGE (g/T) 50		
SOLIDS CONCENTRATION (g/L) 50		
L (ml)	H(cm)	t (s)
1000	35	0
900	31.5	33.26
850	29.75	45.73
800	28	58.45
750	26.25	70.86
700	24.5	83.57
650	22.75	96.73
600	21	110.51
550	19.25	126.02
500	17.5	144.73
450	15.75	177.39
400	14	248.29
350	12.25	395.32

Table A.2: flocculant dosage: 60 gpt

MSlurry (g)	187.3	
MLiquor (g)	964.1	
FLOCCULANT DOSAGE (g/T) 60		
SOLIDS CONCENTRATION (g/L) 50		
L (ml)	H(cm)	t (s)
1000	35	0
950	33.25	30.5
900	31.5	45
850	29.75	56.89
800	28	68.76
750	26.25	80.7
700	24.5	94.9
650	22.75	107.87
600	21	122.12
550	19.25	140
500	17.5	172.65
450	15.75	242.3
400	14	380.65

Table A.3: flocculant dosage: 70 gpt

MSlurry (g)	187
MLiquor (g)	962.3
FLOCCULANT DOSAGE (g/T)	70
SOLIDS CONCENTRATION (g/L)	50

L (ml)	H(cm)	t (s)
1000	35	0
900	31.5	30.23
850	29.75	42.23
800	28	54.26
750	26.25	66.13
700	24.5	77.7
650	22.75	91.16
600	21	104.07
550	19.25	117.7
500	17.5	131.8
450	15.75	169.7
400	14	236.64
350	12.25	376.57

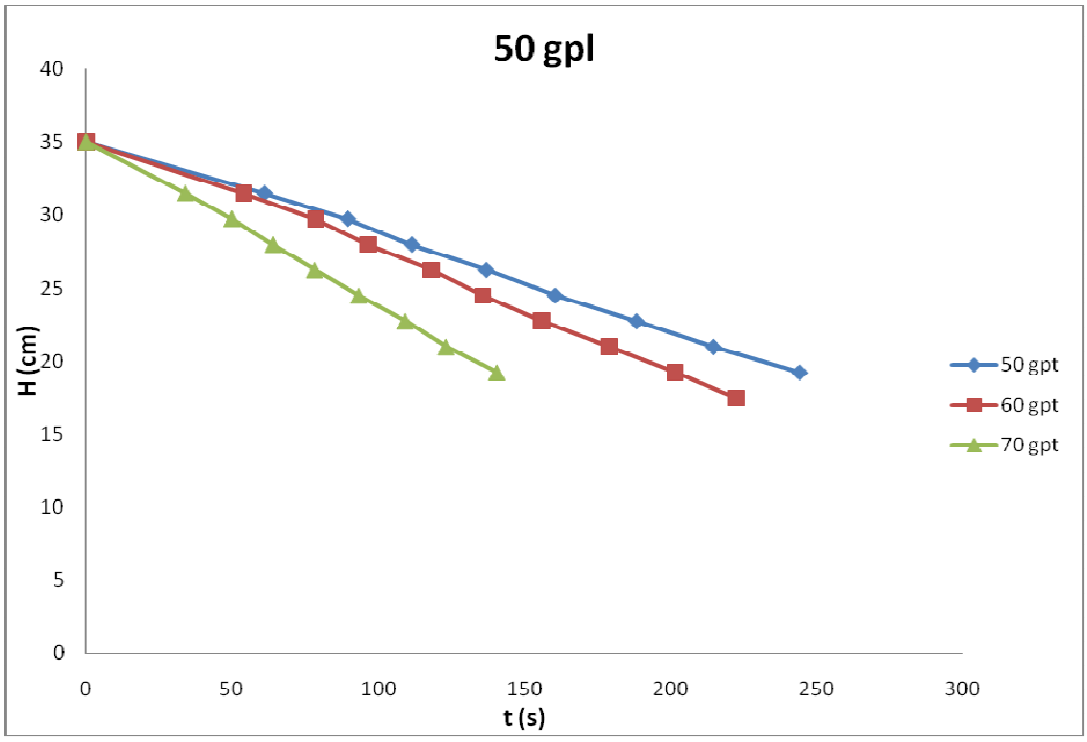


Figure A.1: initial settling rate at different flocculant dosages

APPENDIX A.2: Settling test at constant flocculant dosages (Washer 20)

Settling test for washer 20 at constant flocculant dosage (70 gpt).

Table A.4: solids concentration 70 g/L

MSlurry (g)	262
MLiquor (g)	907.2
FLOCCULANT DOSAGE (g/T)	70
SOLIDS CONCENTRATION (g/L)	70

L (ml)	H(cm)	t (s)
1000	35	0
900	31.5	43.16
850	29.75	65.07
800	28	86.92
750	26.25	109.2
700	24.5	129.42
650	22.75	153.64
600	21	175.64
550	19.25	217.86
500	17.5	300.76
450	15.75	455.89
400	14	744.86

Table A.5: Solids concentration 90 g/L

MSlurry (g)	337.1
MLiquor (g)	856.3
FLOCCULANT DOSAGE (g/T)	70
SOLIDS CONCENTRATION (g/L)	90

L (ml)	H(cm)	t (s)
1000	35	0
900	31.5	157.76
850	29.75	212.64
800	28	255.7
750	26.25	308.86
700	24.5	357
650	22.75	422.64
600	21	566.8
550	19.25	901.02

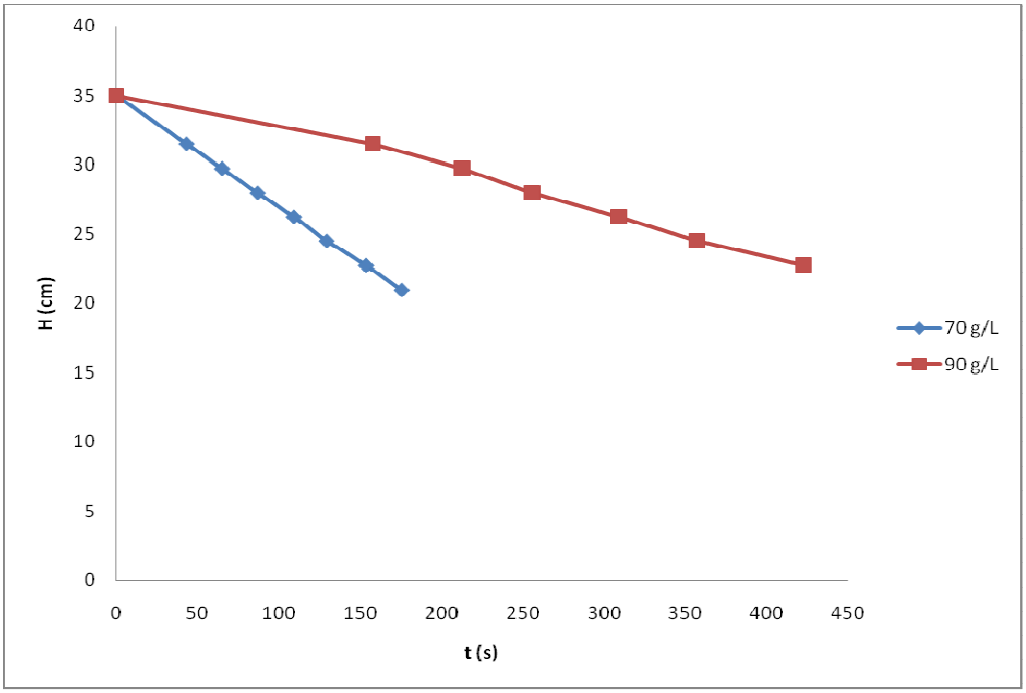


Figure A.2: Initial settling rate washer 20 (flocculant dosage 70 gpt)

APPENDIX A.3: Settling test at different flocculant dosages (Washer 50)

Settling test for washer 50 at different flocculant dosages (initial solids concentration: 50g/L).

Table A.6: Settling test at 10 gpt

MSlurry (g)	140	
MLiquor (g)	990	
FLOCCULANT DOSAGE (g/T)	10	
SOLIDS CONCENTRATION (g/L)	50	
L (ml)	H(cm)	t (s)
1000	35	0
900	31.5	46.16
850	29.75	56.42
800	28	67
750	26.25	78.32
700	24.5	90.54
650	22.75	103.23
600	21	115.92
550	19.25	130.82
500	17.5	146.82
450	15.75	176.95
400	14	235.89
350	12.25	
300	10.5	
250	8.75	

Table A.7: Settling test at 20 gpt

MSlurry (g)	141	
MLiquor (g)	990	
FLOCCULANT DOSAGE (g/T)	20	
SOLIDS CONCENTRATION (g/L)	50	
L (ml)	H(cm)	t (s)
1000	35	0
900	31.5	12.3
850	29.75	15.61
800	28	19.58
750	26.25	23.8
700	24.5	27.83
650	22.75	32.45
600	21	38.02
550	19.25	43.48
500	17.5	49.11
450	15.75	56.04
400	14	71.95
350	12.25	107.04
300	10.5	180.04
250	8.75	

Table A.8: Settling test at 30 gpt

MSlurry (g)	140.7
MLiquor (g)	990.6
FLOCCULANT DOSAGE (g/T)	30
SOLIDS CONCENTRATION (g/L)	50

L (ml)	H(cm)	t (s)
1000	35	0
900	31.5	8.13
850	29.75	10.36
800	28	12.13
750	26.25	13.95
700	24.5	15.95
650	22.75	18.48
600	21	20.7
550	19.25	23.54
500	17.5	26.64
450	15.75	29.73
400	14	35.92
350	12.25	53.16
300	10.5	87.57
250	8.75	249.67

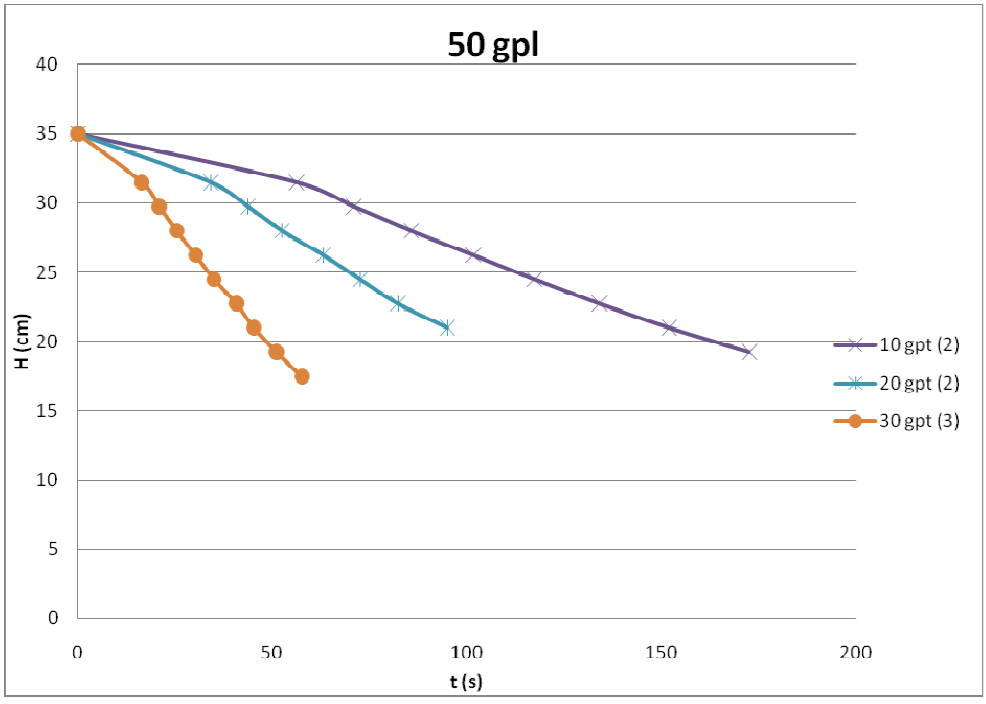


Figure A.3: initial settling rate at different flocculant dosages

APPENDIX A.4: Settling test at constant flocculant dosages (Washer 50)

Settling test for washer 50 at constant flocculant dosage (20 gpt).

Table A.9: Settling test at 20 gpt

MSlurry (g)	196.7
MLiquor (g)	946
FLOCCULANT DOSAGE (g/T)	20
SOLIDS CONCENTRATION (g/L)	70

L (ml)	H(cm)	t (s)
1000	35	0
900	31.5	20.07
850	29.75	29.95
800	28	37.2
750	26.25	45.23
700	24.5	53.7
650	22.75	62.73
600	21	72.7
550	19.25	85.2
500	17.5	108.67
450	15.75	152.42
400	14	233.8

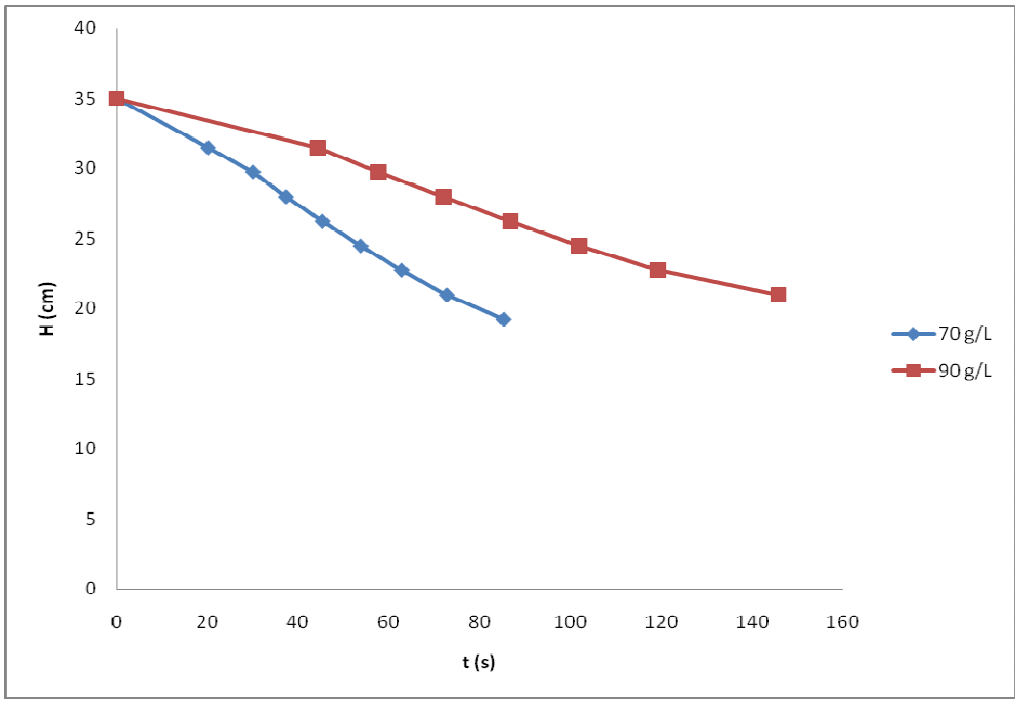


Figure A.4: Settling curve at 20 g/T

APPENDIX A.5: Settling test at different flocculant dosages (Washer 70)

Settling test for washer 70 at different flocculant dosages (initial solids concentration: 50g/L).

Table A.10: Settling test at 50 gpt

MSlurry (g)	127.9	
MLiquor (g)	910	
FLOCCULANT DOSAGE (g/T) 50		
SOLIDS CONCENTRATION (g/L) 50		
L (ml)	H(cm)	t (s)
1000	35	0
900	31.5	56.02
850	29.75	84.8
800	28	109.64
750	26.25	137.2
700	24.5	165
650	22.75	194.57
600	21	224.2
550	19.25	256.89
500	17.5	297.89
450	15.75	375.1
400	14	531.1
300	10.5	960

Table A.11: Settling test at 60 gpt

MSlurry (g)	127.7	
MLiquor (g)	910	
FLOCCULANT DOSAGE (g/T) 60		
SOLIDS CONCENTRATION (g/L) 50		
L (ml)	H(cm)	t (s)
1000	35	0
900	31.5	11.67
850	29.75	17.3
800	28	22.48
750	26.25	28.36
700	24.5	33.77
650	22.75	40.51
600	21	46.74
550	19.25	53.7
500	17.5	62.7
450	15.75	80.3
400	14	114.51
350	12.25	193.04

Table A.12: Settling test at 50 gpt

MSlurry (g)	127.4
MLiquor (g)	910
FLOCCULANT DOSAGE (g/T)	30
SOLIDS CONCENTRATION (g/L)	50

L (ml)	H(cm)	t (s)
1000	35	0
900	31.5	22.36
850	29.75	30.8
800	28	38.64
750	26.25	48.04
700	24.5	57.32
650	22.75	66.82
600	21	76.51
550	19.25	88.02
500	17.5	101.67
450	15.75	128.73
400	14	176.1
350	12.25	254.23

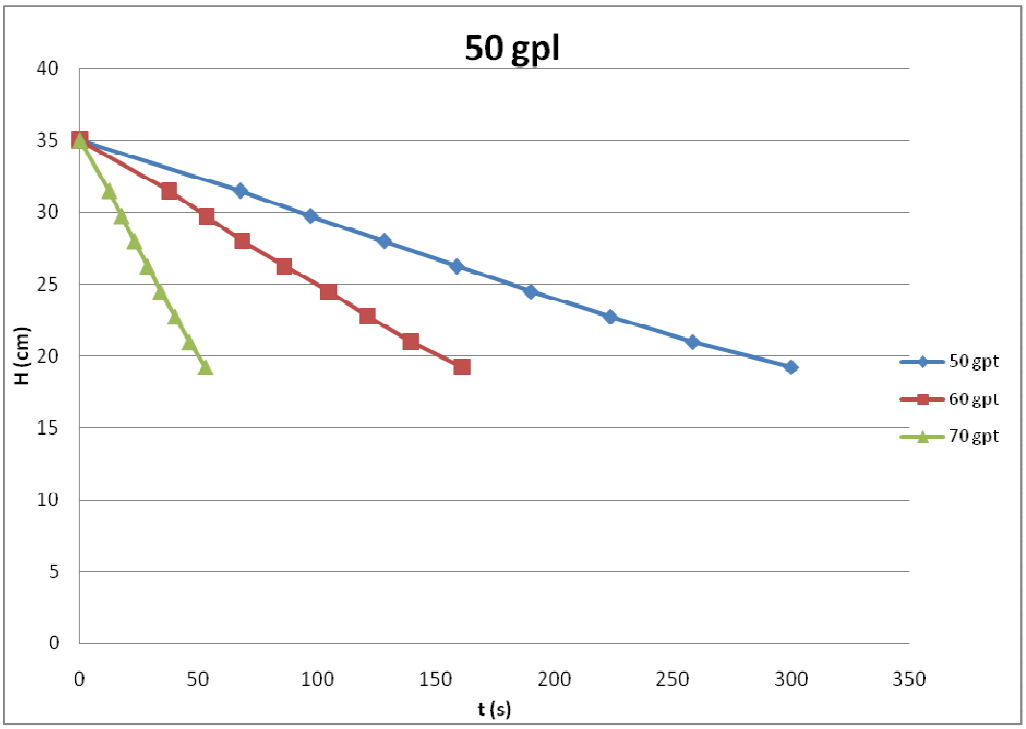


Figure A.5: initial settling rate at different flocculant dosages

APPENDIX A.6: Settling test at constant flocculant dosages (Washer 70)

Settling test for washer 70 at constant flocculant dosage (70 gpt).

Table A.13: Settling test at 70 gpt

MSlurry (g)	178.8
MLiquor (g)	872.4

FLOCCULANT DOSAGE (g/T)		
70		
SOLIDS CONCENTRATION (g/L)		
70		
L (ml)	H(cm)	t (s)
1000	35	0
900	31.5	78.7
850	29.75	107.16
800	28	136.86
700	24.5	197.7
650	22.75	230.6
600	21	269.82
550	19.25	341.51
500	17.5	458.32

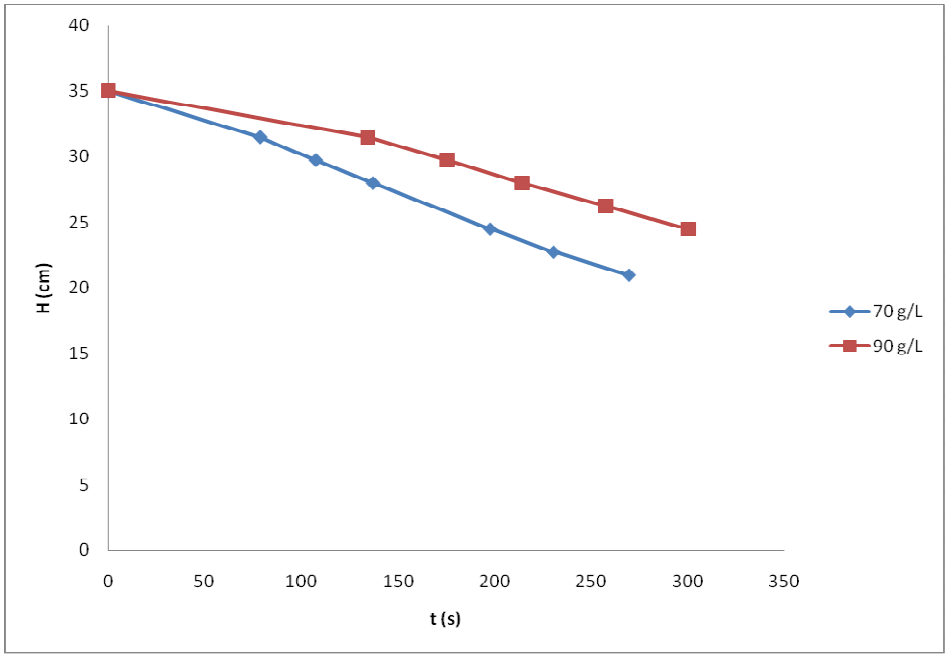


Figure A.6: initial settling rate at 70 g/L

APPENDIX B:
RHEOLOGY TEST DATA

APPENDIX B.1: Torque vs time (Washer 20)

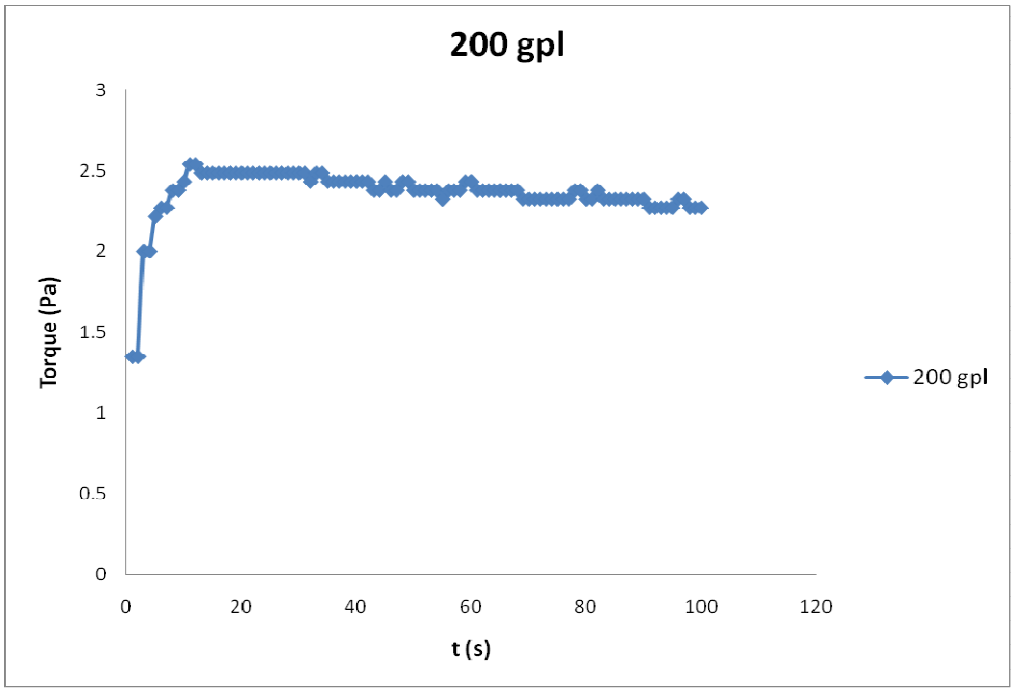


Figure B.1: Torque vs time at 200 g/L

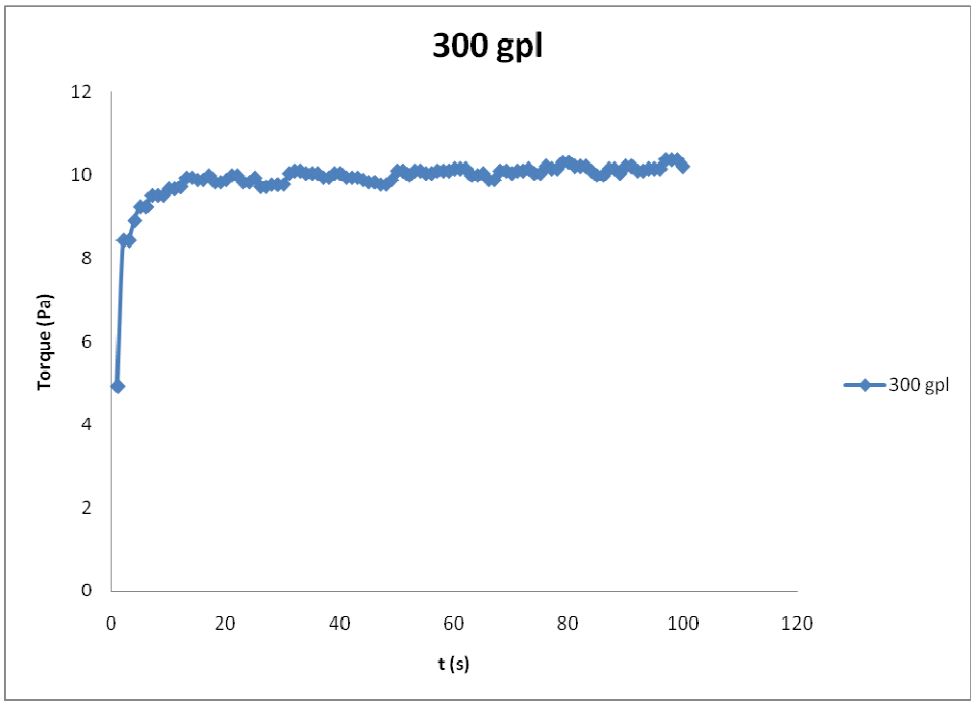


Figure B.2: Torque vs time at 300 g/L

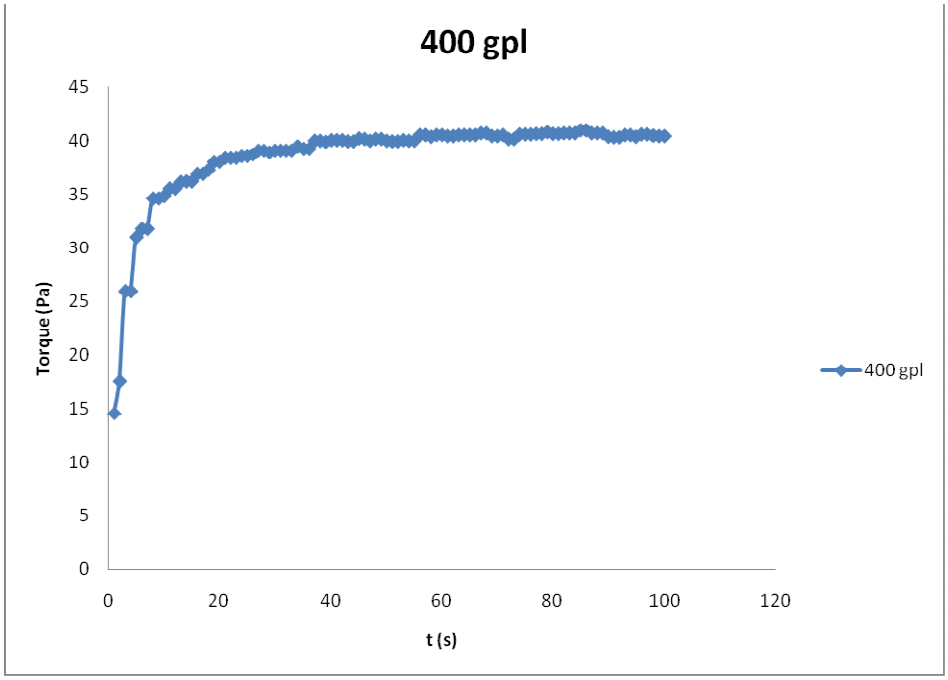


Figure B.3: Torque vs time at 400 g/L

APPENDIX B.2: Torque vs time (Washer 50)

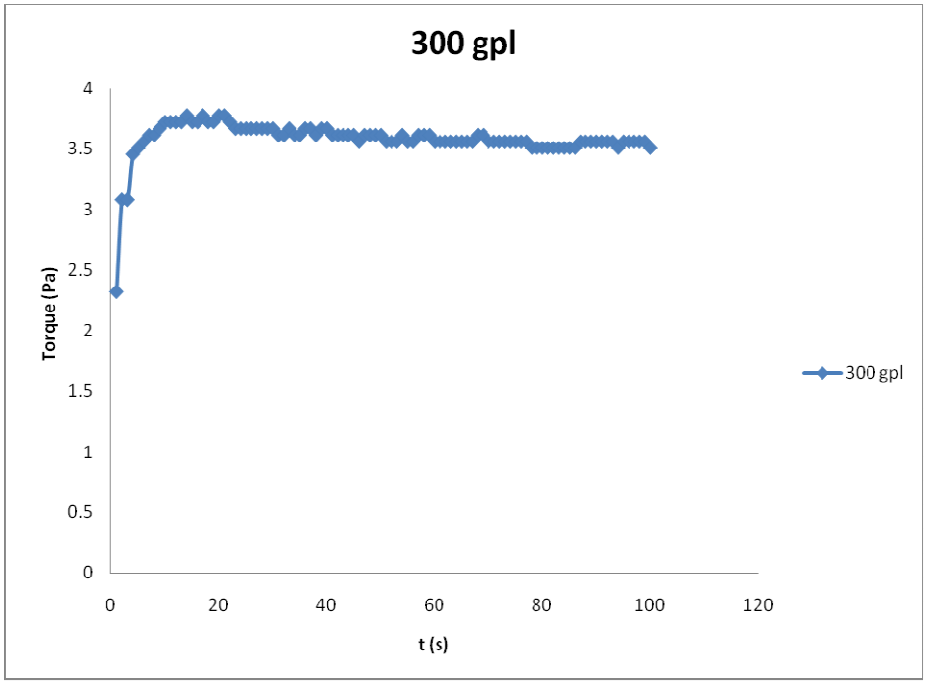


Figure B.4: Torque vs time at 300 g/L

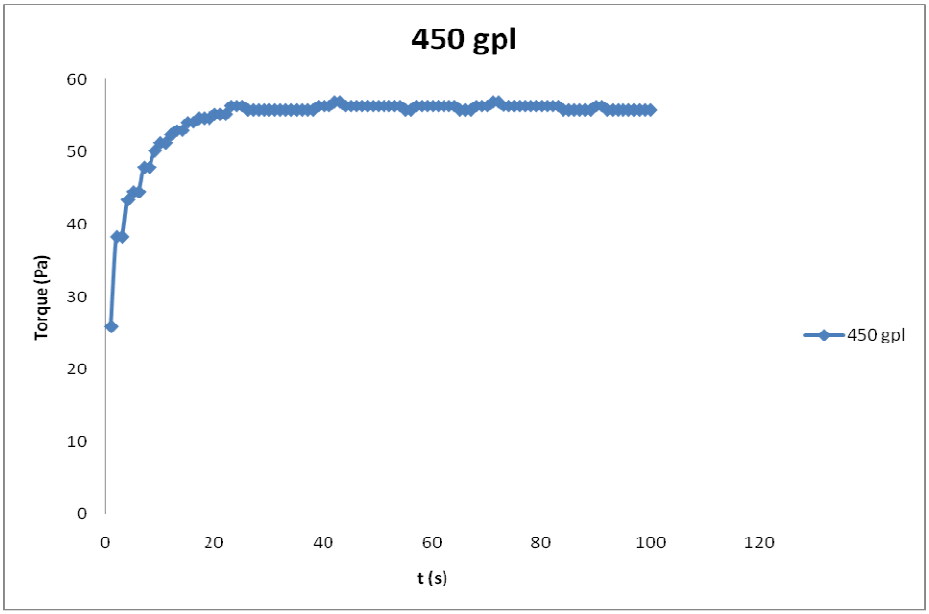


Figure B.5: Torque vs time at 450 g/L

APPENDIX B.3: Torque vs time (Washer 70)

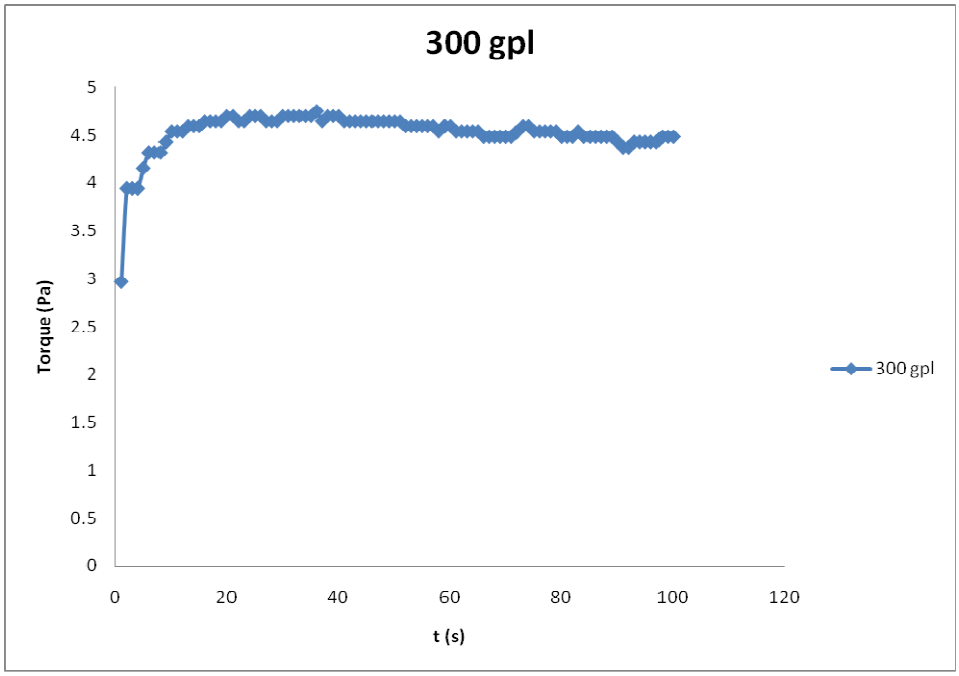


Figure B.6: Torque vs time at 300 g/L

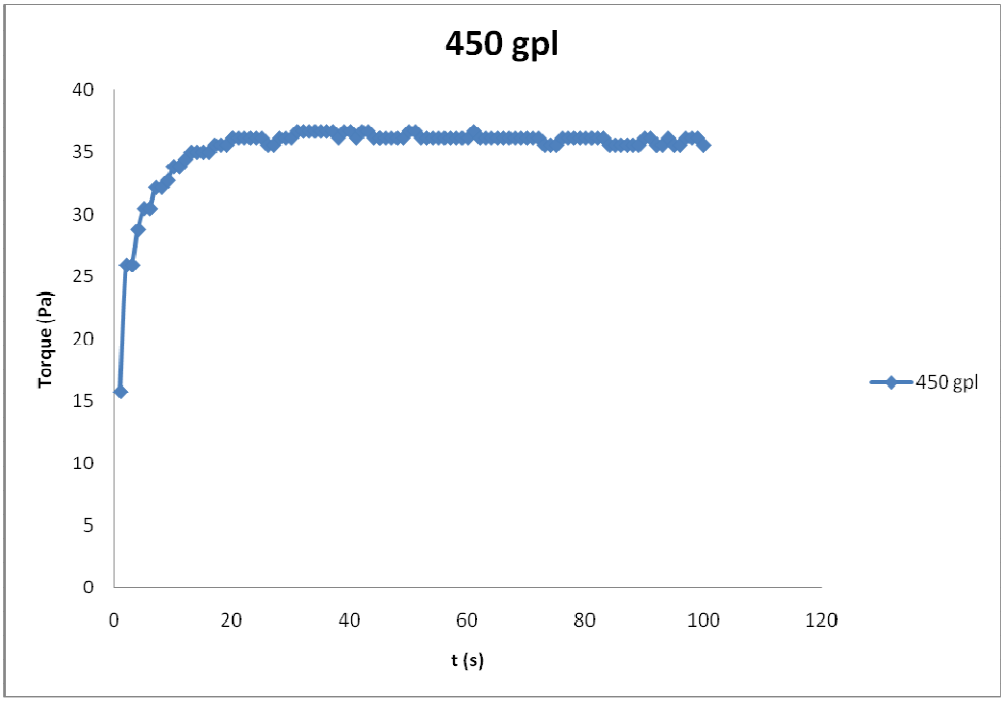


Figure B.7: Torque vs time at 450 g/L

APPENDIX B.4: Herschel-Bulkley parameters:

Table B.1: Herschel-Bulkley parameters

Washer No.	Concentration (gpl)	τy (Pa)	k	n
20	200	2.5	1.3	0.21
	300	10.97	7.5	0.12
	400	47	39.41	0.04
50	200	3.58	1.83	0.21
	300	4.2	2.42	0.18
70	200	1.66	0.94	0.19
	300	5.3	3.7	0.12
	450	33.08	27.37	0.05