SIMULTANEOUS MINIMISATION OF WATER AND ENERGY WITHIN A WATER AND MEMBRANE NETWORK SUPERSTRUCTURE

Esther Buabeng-Baidoo

Master of Science in Engineering by research:

“A dissertation submitted to the Faculty of Engineering and Built Environment, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Engineering.’’

June, 2016

Declaration

I declare that this dissertation is my own unaided work. It is being submitted for the Degree of Master of Science in Chemical Engineering to the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination to any other University.

....

(Signature of Candidate)

.... day of .... year..

In loving memory of my dearest brother

Cephas Buabeng-Baidoo

Acknowledgements

First, I would like to thank almighty God, without whom none of this would have been possible. I would like to thank my supervisor, Prof Thokozani Majozi for his support, guidance and patience over the past two and a half years.

I would like to thank my family and friends for all the support and love they showed me during the good and bad times. This has been a long Journey but you never gave up on me. A special thanks to my colleagues especially, Nielsen Mafukidze, Dematria Nezungai and Musah Abass. Thank you for the many good laughs, your insight, thoughts and love.

I would also like to acknowledge the National Research Foundation (NRF) for funding this work under the NRF/DST Chair in Sustainable Process Engineering at the University of the Witwatersrand, Johannesburg, South Africa.

Synopsis

The scarcity of water and strict environmental regulations have made sustainable engineering a prime concern in the process and manufacturing industries. Water minimisation involves the reduction of freshwater use and effluent discharge in chemical plants. This is achieved through water reuse, water recycle and water regeneration. Optimisation of the water network (WN) superstructure considers all possible interconnections between water sources, water sinks and regenerator units (membrane systems). In most published works, membrane systems have been represented using the "black-box" approach, which uses a simplified linear model to represent the membrane systems. This approach does not give an accurate representation of the energy consumption and associated costs of the membrane systems.

The work presented in this dissertation therefore looks at the incorporation of a detailed reverse osmosis network (RON) superstructure within a water network superstructure in order to simultaneously minimise water, energy, operating and capital costs. The WN consists of water sources, water sinks and reverse osmosis (RO) units for the partial treatment of the contaminated water. An overall mixed-integer nonlinear programming (MINLP) framework is developed, that simultaneously evaluates both water recycle/reuse and regeneration reuse/recycle opportunities. The solution obtained from optimisation provides the optimal connections between various units in the network arrangement, size and number of RO units, booster pumps as well as energy recovery turbines. The work looks at four cases in order to highlight the importance of including a detailed regeneration network within the water network instead of the traditional "black-box'' model. The importance of using a variable removal ratio in the model is also highlighted by applying the work to a literature case study, which leads to a 28% reduction in freshwater consumption and 80% reduction in wastewater generation.

Contents

List of figures

List of tables

CHAPTER

1

INTRODUCTION

1.1 Background

Strict environmental regulations and social pressures have created the need for water and energy minimisation in the process industries (Bandyopadhyay & Cormos, 2008). Water minimisation involves the reduction of freshwater use and effluent discharge in chemical plants. This is achieved through water reuse, water recycle and water regeneration. Water reuse involves the use of wastewater in operations other than the process where it was originally produced. Water recycle, however, allows the effluent to be used in any process including the process in which it was produced. In water regeneration-reuse/recycle, the effluent is partially treated before it is recycled or reused in other processes. Partial treatment can be achieved by using water purification units often classified as membrane and non-membrane processes, e.g. reverse osmosis (RO) membranes and steam stripping respectively (Khor et al., 2011).

The purification of water through membrane systems is an energy intensive process. The minimisation of energy within the water networks is also needed for sustainable development. Energy usage within the water network is largely associated with the regeneration units (membrane units). In most published works, however, membrane systems have been represented using the “black-box” approach, which uses a simplified linear model to represent the membrane systems (Tan et al., 2009; Alva-Argáez et al., 1998; Khor et al., 2012). The performance of the regenerators in most cases was also represented by a fixed removal ratio (RR), which is the fraction of mass load into the regenerator that exits in the retentate stream (Khor et al., 2011).

RO membranes are highly favoured amongst other separation units due to their relatively low energy consumption, ease of operation, high product recovery and quality (El-Halwagi, 1992). In past studies of RO membranes, more attention has been given to incorporating it in a water network superstructure in order to minimise the amount of water usage. Some papers have also focused on minimising the energy used by the RO systems by using energy recovery turbines (El-Halwagi, 1992). Very little focus has, however, been given to synthesising the RO network and incorporating it in a WN superstructure. This approach minimises freshwater, energy and also synthesises the optimal number of RO units, booster pumps, energy-recovery turbines, operating conditions and to allow for parallel and series connections. Most of the work on RO systems has failed to achieve these objectives simultaneously.

There are two major approaches adopted in addressing water network synthesis, namely, insight based techniques and mathematical model-based optimisation methods. Insight-based techniques involve the water pinch analysis, which is a graphical method based on the concept of a limiting water profile. This method was first proposed by Wang and Smith (1994a). Hallale (2002) then proposed a graphical method that was based on non-mass transfer operations with single contaminants. Recently, the water pinch method has been extended to only include algebraic methods such as the water cascade analysis (Ng et al., 2007; Manan et al., 2004). The water pinch method proves unsuccessful for complex problems involving multiple contaminants (Faria & Bagajewicz , 2009) and various topological constraints (Khor et al., 2012). The computational burden of this method is, however, lower than that experienced by mathematical model based optimisation methods.

The mathematical optimisation approach employs a superstructure, which identifies an optimal configuration for the process from a number of alternatives. This idea was first proposed in the work of Takama et al. (1980). They proposed a nonlinear model that incorporates both water using and wastewater treating units for multiple contaminant systems. Significant developments in the area have been achieved including the work of Galan and Grossmann (1998), Karuppiah and Grossmann (2006) and Tan et al. (2009) who explored different techniques for modeling regenerators and developing strategies for the complex mixed-integer nonlinear programming (MINLP) problems. However, mathematical optimisation is computationally expensive.

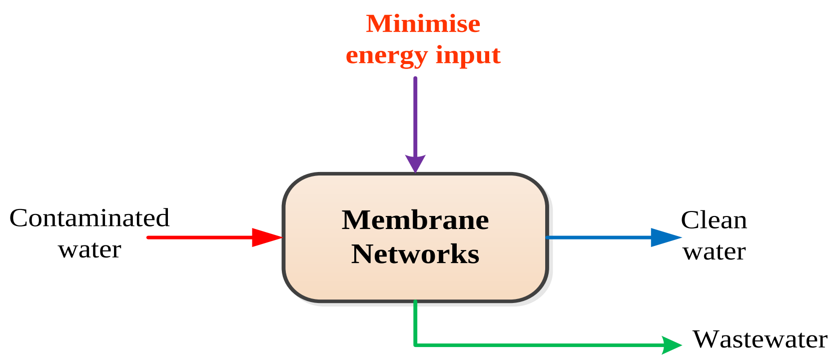
1.2 Motivation

The first motivation behind this work is that most works on WN synthesis does not consider regeneration reuse/recycling. The incorporation of a regenerator is proven to reduce the amount of freshwater usage and wastewater generation in the process industries. The second motivation behind this work is that, most work on WN synthesis that does incorporate regeneration units, describes the performance of the regenerators by means of the "black-box'' approach. This approach does not give an accurate representation of the energy consumption and associated costs of the membrane systems. The treatment units cannot be clearly identified with this method and no design considerations are indicated. This, therefore, means that a more rigorous and detailed design and synthesis model of the regeneration units is needed (Khor et al., 2014). This will allow the incorporation of parallel and series configuration of the regenerators, as this is not taken into account with the "black-box" approach.

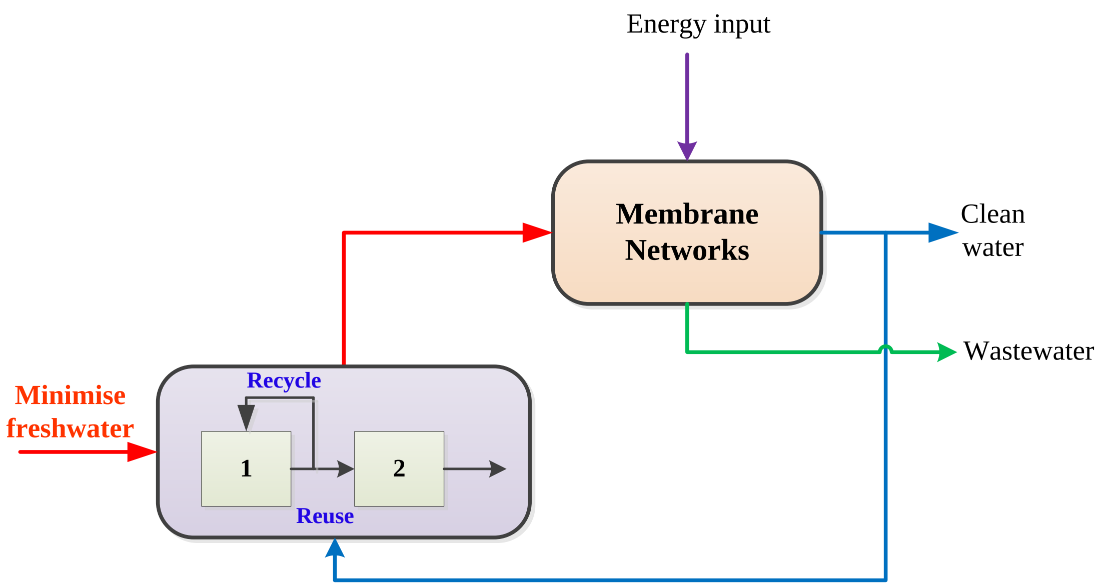
The final motivation for this work is that most work on water networks has not focused on minimising both water and energy simultaneously within WN superstructure. In past studies of RO membranes, more attention has been given to incorporating it in a WN in order to minimise the amount of water usage. Some papers have also focused on minimising the energy used by the RO systems by using energy recovery turbines (El-Halwagi, 1992). Very little focus has, however, been given to synthesising the RO network and incorporating it in a WN that minimises freshwater, energy and also synthesise the optimal number of RO units, booster pumps and energy-recovery turbines at optimal operating conditions. Most of the works on RO systems have not addressed these objectives simultaneously.

Figure 1.1 shows a schematic representation of the motivations behind this work with regards to energy and water minimisation. In Figure 1.1(a) the idea was to minimise the amount of energy used by the membrane networks and this was achieved in the work of

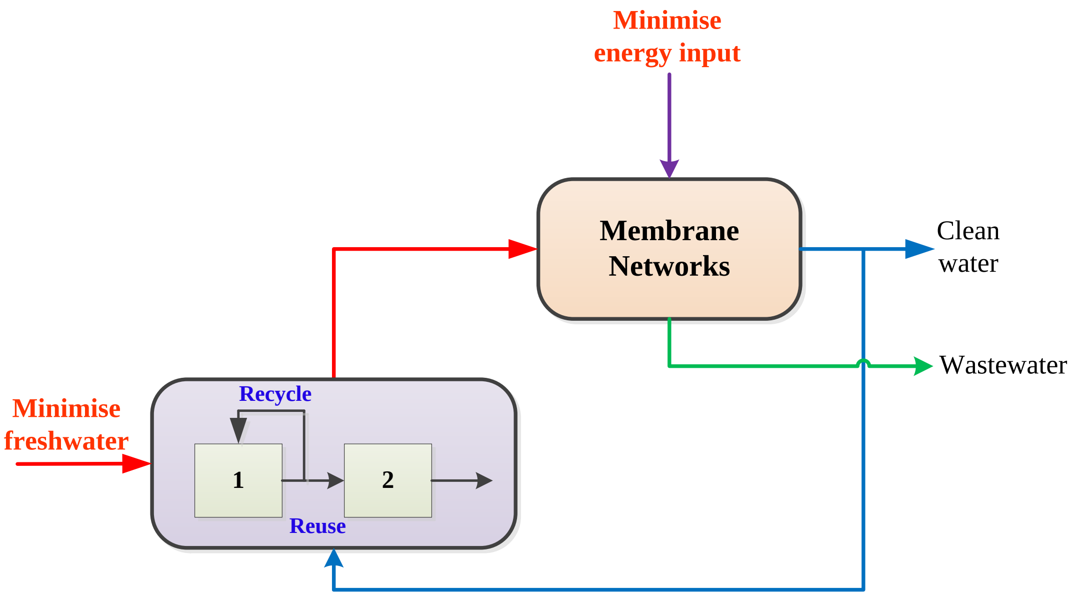
Tsiakis and Papageorgiou (2005). Figure 1.1(b) shows the scenario where freshwater minimisation was the main objective of the optimisation, as the minimisation of energy was not considered. Figure 1.1(c) shows the scenario where the objective of the problem is to simultaneously minimise energy and water.



(a)



(b)



(c)

Figure 1.1: Illustration of the motivation behind the work.

1.3 Objectives

The objectives of the research can be summed up as follows:

(i) To develop a mathematical model to synthesise a detailed network of RO membranes to purify industrial wastewater and also minimise the energy used by the regenerators.

(ii) To develop a mathematical model for WN superstructure that treats wastewater with multiple contaminants for further recovery to minimise fresh water consumption.

(iii) To combine the reverse osmosis network (RON superstructure and the water network superstructure (WNS) in order to simultaneously minimise energy and water use.

(iv) To explore the idea of using a variable RR to describe the performance of the RO membranes.

(v) Validate the model with a literature study in order to show the practicality of the model.

1.4 Problem Statement

The problem addressed in this work can be stated as follows:

Given:

(i) A set of water sources, I, $i \in I$, with known flowrates and known contaminant concentration, M, $m \in M$.

(ii) A set of water sinks, J, $j \in J$, with known flowrates and known maximum allowable contaminant concentration.

(iii) A network of RO regenerators, Q, $q \in Q$, with known liquid recovery and design parameters.

(iv) A freshwater source, *FW*, with known contaminant concentration and variable flowrate.

(v) A wastewater sink, *WW*, with known maximum allowable contaminant concentration and variable flowrate.

Determine:

(i) The minimal freshwater intake, wastewater generation, the energy consumed in the RON and the total annualised cost (TAC).

(ii) The optimal configuration of the water network.

(iii) The optimal number of RO units, pumps and energy recovery turbines.

(iv) The optimal operation and design conditions of the RON such as feed pressure, number of hollow fibre modules per regenerator, stream distributions, separation levels etc.

1.5 Dissertation Structure

Chapter 2 gives a comprehensive survey of the literature connected to this work. Literature review is given on the synthesis of RO membranes as well as the different types of membranes used for the purification of wastewater. Literature is also given on the different techniques used in solving WN problems such as insight-based methods and mathematical model based optimisation methods. The review also looks at the work that has used the “black-box’’ approach and the work that has considered a detailed design of regenerators within the WN. The mathematical model is then developed in Chapter 3. Chapter 4 shows the results obtained when the model is applied to a petroleum refinery case study and Chapter 5 gives recommendations and considerations for future work drawn from the study. Conclusions are presented in Chapter 6. References follow each chapter.

1.6 References

Alva-Argáez, A., Kokossis , A. & Smith , R., 1998. Wastewater minimisation of industrial systems using an integrated approach. Computers and Chemical Engineering, 22, pp.741-44.

Bandyopadhyay , S. & Cormos, C., 2008. Water management in process industries incorporating regeneration and recycle through a single treatment unit. Industrial and Engineering Chemistry Research , 47, pp.1111-19.

El-Halwagi, M., 1992. Synthesis of reverse osmosis networks for waste reduction. American Institute of Chemical Engineers, pp.1185-98.

Faria, D. & Bagajewicz , M., 2009. On the appropriate modelling of process plant water systems. American Institute of Chemical Enginering Journal, 56, pp.668-89.

Galan , B. & Grossmann, I., 1998. Optimal design of distributed wastewater treatment networks. Industrial and Engineering Chemistry Research, 37(10), pp.4036-48.

Hallale , N., 2002. A new graphical targeting method for water minimisation. Advances in Environmental Research, 52, pp.377-90.

Karuppiah, R. & Grossmann, I., 2006. Karuppiah, R.; Grossmann, I. Global optimisation for the synthesis of integrated water systems in chemical processes. Computers and Chemical Engineering 2006, 30 (4), 650-673

Khor , C., Chachuat , B. & Shah, N., 2014. Optimisation of water network synthesis for single-site and continous processes: milstones, challenges, and future directions. Industrial & Engineering Chemistry Research , 53, pp.10257-75.

Khor, C., Chachuat , B. & Shah, N., 2012. A superstructure optimisation approach for water network synthesis with membrane separation-based regenerators. Computers and Chemical Engineering , 42, pp.48-63.

Khor, CS, Foo, DCY., El-Halwagi, MM., Tan, RR., Shah, N., 2011. A superstructure optimisation approach for membrane separation-based water regeneration networks synthesis with detailed nonlinear mechanistic reverse osmosis model. Industrial and Engineering Chemistry Research , 50, pp.13444-56.

Manan , Z., Tan , Y. & Foo, D., 2004. Targeting the minimum water flow rate using water cascade analysis technique. AIChE Journal , 50(12), pp.3169-3183.

Ng, D., Foo , D. & Tan , R., 2007. Targeting for total water networks.1. waste stream identification. Industrial and Engineering Chemistry Research , 46, pp.9107-13.

Takama, N., Kuriyama , T., Shiroko, K. & Umeda , T., 1980. Optimal water allocation in a petroleum refinery. Computers and Chemical Engineering, 4, pp.251-58.

Tan, R., Ng, D., Foo, D. & Aviso, K., 2009. A superstructure model for the sybthesis of single-contaminant water network with partitioning regenerators. Process Safety and Environmental Protection, 87(3), pp.197-205.

Tsiakis, P. & Papageorgiou, L., 2005. Optimal design of an electrodialysis brackish water desalination plant. Desalination , 173(2), pp.173-86.

Wang , Y. & Smith, R., 1994a. Wastewater minimisation. Chemical Engineering Science, 49(7), pp.981-1006.

CHAPTER

2

LITERATURE REVIEW

2.1 Introduction

As the work in this dissertation focuses on the incorporation of a RON superstructure within a WN, background is given on the synthesis of a RON and the WN. Different membrane technologies are also discussed with their respective advantages and disadvantages in order to highlight the importance of using RO membranes for the minimisation of water and energy in the process industries.

The incorporation of the RON superstructure within the WN superstructure leads to an overall mixed-integer nonlinear programming (MINLP) framework. The overall mathematical model consists of binary variables that are used to account for the existence of units and streams. An MINLP model is, however, difficult to solve due to bilinear terms (which create nonconvex functions) in the mass balance equations and the concave cost terms in the objective function (Ahmetović & Grossmann , 2010). This section therefore looks at the different approaches that have been used over the years to solve nonlinear problems (NLP) and MINLP problems with regards to WN synthesis problems.

Finally, the chapter includes at a detailed discussion of the synthesis of a RON. The works that have looked at a “black-box’’ representation of the regenerators and those that consider a detailed synthesis of the regenerators have been discussed.

2.2 Process Integration

Process integration is defined as “a holistic approach to process design, retrofitting and operation of existing plants which emphasises the unity of the process and considers the interactions between different unit operations from the outset rather than optimising them separately” (El-Halwagi, 1997). The main advantage of this method is that, it looks at the system as a whole unlike analytical approaches that attempt to optimise or improve a process unit by looking at each unit separately.

These techniques are often used at the beginning of a project in order to screen all the possible options to optimise the design and/or operations of the plant. The objective of process integration is therefore to optimise the use of resources, energy and equipment and to produce sustainable methods, which in turn can have a significant effect on the efficiency and revenue of the plant. Process integration methods are, therefore, used in conjunction with mathematical optimisation methods.

2.3 Mathematical Optimisation

2.3.1 Optimisation Theory

According to Snyman (2005) mathematical optimisation is defined as “the science of determining the best solutions to mathematically defined problems, which may be models of physical reality or of manufacturing and management systems”. This is needed in engineering in order to not only minimise or maximise cost, but also to develop designs that enhance sustainable developments.

Optimisation in engineering is concerned with the selection of the best solution (global optimum) or one of the best solutions (local optimum) among an entire set by an efficient quantitative method. Every optimisation problem consists of at least one objective function, and equality and inequality constraints (Edgar & Himmelblau , 1988). Different mathematical solvers are used to obtain the optimal solutions. The problem that needs to be solved has to be written in a mathematical form in order for the solvers to obtain the solutions. The following is the typical mathematical form of an optimisation problem (Song, 1999):

Objective: minimise $f(x)$

Subject to $h(x) = 0$

$g(x) \leq 0$

The aim is therefore to minimise an objective function $f(x)$ that is subject to equality $h(x)$ and inequality constraints $g(x)$. The mathematical model obtained from the RON and WN can be a linear model, nonlinear (NLP) model, mixed integer nonlinear programming (MINLP) model or a mixed integer linear programming (MILP) model. A feasible solution in an optimisation problem is when a set of variables satisfies the constraints of the problem. A feasible region of an optimisation problem represents all the possible feasible solutions to the problem (Edgar & Himmelblau , 1988). An optimal solution is a set of feasible solutions that give the best solution to the objective function (Edgar & Himmelblau , 1988).

Different papers use different methods and computer programming solvers to solve the mathematical models in order to obtain an optimal solution. The optimal solution can be a local minimum, local maximum, global minimum or a global maximum solution (best solution). The solution is, however, dependant on whether a model is convex or concave. This will help determine if a locally optimal solution is also a globally optimal solution. A function is convex if a line drawn arbitrarily between two points on a convex curve, has all its values above the curve. This therefore means that the points on the curve must be less than or equal to the points on the straight line. This observation is best depicted in Figure 2.1(a). A concave function is when all the points on the curve are greater than or equal to the points on the straight line. This observation is depicted in Figure 2.1(b) (Edgar & Himmelblau , 1988).

The function can also be classified as strictly convex or concave. A strictly convex function has the greater or equal to sign replaced by just a greater than sign while a strictly concave function has its less than or equal to sign replaced by just a less than sign. This therefore means that strictly convex or concave function provides a single optimum solution. A nonconvex function may, however, have multiple optimum solutions (local optima).

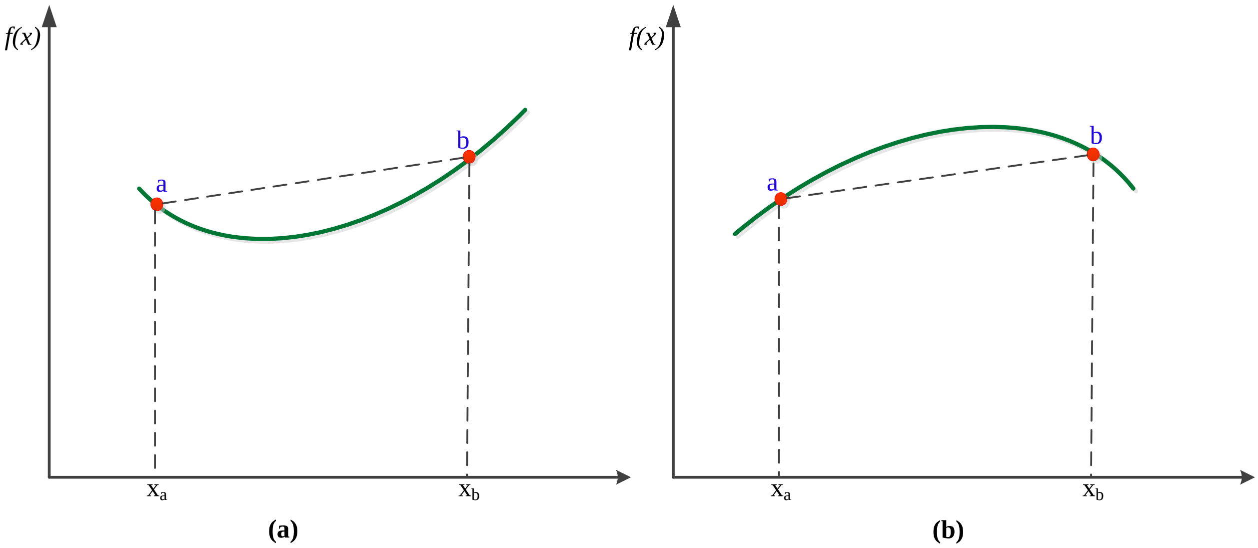


Figure 2.1: Comparison of a) Convex function and b) Concave function.

2.3.2 Convexification methods

A global optimum, which is the best solution, can be guaranteed if a function is convex (Lundell & Westerlund, 2012). This therefore means that MINLP models have to be convexified in order for a globally optimal solution to be obtained. There is currently no method available that can guarantee global optimality for nonconvex MINLP problems (Pörn et al., 1999) .

Relaxation methods are a modelling strategy used to approximate difficult problems by means of nearby problems that are easy to solve. A solution obtained from the relaxed problem is used to provide information about the original problem. Linear programming relaxations replace the 0-1 variables by variables belonging to the interval between 0 and 1. This relaxation results in a linear program

The convexification methods that will be discussed in this review are McCormick (1976) over and under estimators, Glover (1975) transformations, reformulation-linearisation techniques (RLT), transformations for other nonlinear terms and piecewise linearisation methods.

a) McCormick (1976) *over and under estimators*

Linear programming methods are used to drive the solution process of NLP and MINLP algorithms. The success of these algorithms therefore depends on the strength and tightness of the linear programming algorithms (Sherali & Adams, 1999).

Linearisation of nonlinear terms within NLP and MINLP models lead to a convex model which can then be solved to obtain a globally optimal solution. Linearisation methods have been developed for bilinear terms. Bilinear terms are a product of two continuous variables or of a product a binary variable and a continuous variable (Zamora & Grossmann, 1998). A product of two continuous variables within a model gives rise to an NLP model. The linearisation of these terms can help accelerate the convergence of the model.

McCormick (1976) introduced a general method for linearising the concave/convex envelopes of these functions that involves a set of LP relaxation models which use linear convex underestimators and linear concave overestimators for a tight upper bound on the global optimum with regards to bilinear terms. It was assumed that convex and concave envelopes can be provided for any function of a single variable. The convex envelope was defined as the highest convex function, which everywhere underestimates the function and the concave envelope was defined as the lowest concave function, which everywhere overestimates the function. This method was, however, limited to NLP problems that are factorable. This method can therefore be used to handle nonconvexities in the concentration balance of the WN. A solution is obtained by solving the linearised model.

The method proposed by McCormick (1976) can be explained as follows: A bilinear term, which arises from a product of two continuous variables, is defined as

$$\begin{aligned} x &\in R, \\ y &\in R \end{aligned} \tag{2.1}$$

The following substitution shown in constraint (2.2) can therefore be made for the product of the two continuous variables.

$$w = xy \tag{2.2}$$

The two continuous variables each have a lower and upper bound and this is shown in constraints (2.3) and (2.4).

$$x^L \le x \le x^U \tag{2.3}$$

$$y^L \le y \le y^U \tag{2.4}$$

The following constraints therefore arise as a result

$$x - x^L \ge 0 \text{ and } y - y^L \ge 0 \tag{2.5}$$

$$x^U - x \ge 0 \text{ and } y^U - y \ge 0 \tag{2.6}$$

Taking the product of the constraints in constraint (2.5), one gets

$$xy - x^L y - y^L x + x^L y^L \ge 0 \tag{2.7}$$

which inherently is positive as the product of two positive terms must also be positive. Rearranging the terms in constraint (2.7) gives rise to constraint (2.8).

$$w \ge x^L y + y^L x - x^L y^L \tag{2.8}$$

Three more constraints can also be derived in the same way using different combinations in constraints (2.5) and (2.6). These constraints are shown in constraints (2.9) to (2.11).

$$w \geq x^U y + y^U x - x^U y^U \tag{2.9}$$

$$w \geq x^L y + y^U x - x^L y^U \tag{2.10}$$

$$w \geq x^U y + y^L x - x^U y^L \tag{2.11}$$

The bilinear term can therefore be replaced by *w*, which has only linear terms. The upper and lower bounds of *x* and *y* are shown in constraints (2.3) and (2.4) respectively.

Constraints (2.8) to (2.17) represent the McCormick (1976) overestimators and underestimator and this is shown in Figure 2.2. This method is, however, not an exact linearisation technique, but does lead to the creation of a convex solution space as all bilinear terms are replaced by linear constraints. The method also allows the resulting system to be solved easily and does not require an initial starting point. It, however, leads to an increase in the number of constraints and can also be cumbersome when applied to a large scale NLP or MINLP problem.

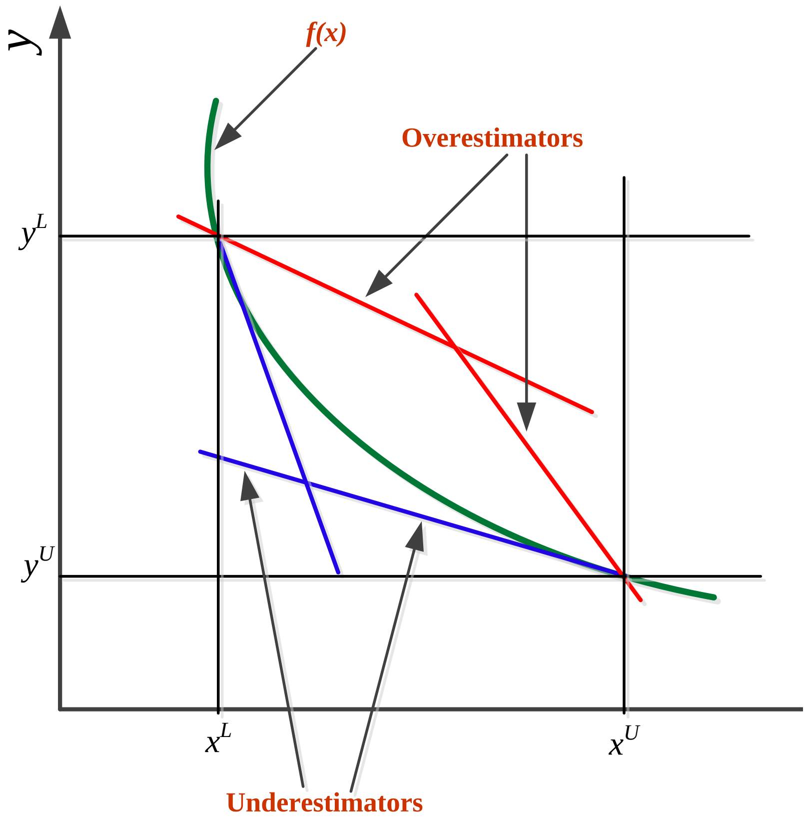


Figure 2.2: Graphical representation of the McCormick (1976) overestimators and underestimators.

b) Glover (1975) *Transformations*

Glover (1975) introduced a novel method to linearise the nonlinear terms in NLP and MINLP problems which are due to the product of a discrete variable and a continuous variable. The method proposed by Glover (1975) can be explained as follows:

Let x be a continuous variable and y a binary variable. Both x and y are defined in constraint (2.12).

$$x \in R, \quad y \in [0,1] \tag{2.12}$$

The product of x and y can be replaced with a new continuous variable, H and this is shown in constraint (2.13).

$$H = xy \tag{2.13}$$

From constraint (2.13), it can be seen that, H can assume a value of 0 when y is equal to 0 and 1 when y is equal to 1. The lower and upper bounds of x, if known, are expressed in constraint (2.14).

$$x^L \leq x \leq x^U \tag{2.14}$$

Constraint (2.14) is then multiplied with y in order to replace the nonlinear term xy with H and this is shown in constraint (2.15). Constraint (2.16) is then derived from the understanding of binary variables and upper and lower bounds of x.

$$x^L y \leq H \leq x^U y \tag{2.15}$$

$$x - x^U(1-y) \leq H \leq x + x^L(1-y) \tag{2.16}$$

It can be seen that constraints (2.15) and (2.16) are linear in terms of x and y as it is assumed that the upper and lower bounds of x are known. These constraints are then used to replace constraint (2.13) in order to eliminate the nonlinear term. This method is an exact transformation technique and a globally optimal solution can therefore be guaranteed provided that the rest of the formulation is also linear.

c) Transformations

Other nonlinear terms within an NLP and an MINLP problem can be linearised by means of transformations, which reformulate an MINLP problem to a convex MINLP problem. This is achieved by transforming the original nonconvex problem into a convex problem, which is then solved using an MINLP solver (Pörn et al., 1999). The nonlinear terms include exponential terms, positive and negative power terms and mixed power and continuous terms, which are grouped together as signomial terms (Lundell & Westerlund, 2012). Different transformations are therefore developed for these terms individually, namely (Lundell & Westerlund, 2012):

(i) Positive power transformations (PPT) for positive power terms.

(ii) Exponential transformation (ET) for exponential terms

(iii) Mixed power and exponential transformations (MPET) for a mixed power and exponential term.

(iv) Power transformations (PT) for negative power terms, which are applied term-wise and include the αBB-reformulation (for nonconvex twice-differentiable functions).

(v) Inverse transformations for positive terms.

The number of discrete and continuous variables that are needed in the reformulation can, therefore, vary depending on the combination of transformations that are chosen. This in turn leads to reformulated problems with different styles (Lundell & Westerlund, 2012). Different transformation methods have been developed over the years to handle nonconvex MINLP problems.

Pörn et al. (1999) looked at a large number of general convexification techniques which were applicable to large class of MINLP problems. The extended cutting plane method was used by convexifying all the inequality constraints and by making sure that all the equality constraints and objective function were linear. Their method showed how posynomials and binomials could be convexified within the discrete optimisation and was a general method, which could incorporate continuous variables. Pörn et al. (2008) then applied the ET, IT and PT for NLP and MINLP problems where the nonlinear transformation constraints were discretised in order to obtain a piecewise linear transformation.

Lundell and Westerlund (2012) introduced a set of transformations for convexifying nonconvex twice-differentiable problems in an extended variable space. An MILP was solved in order to obtain the transformation. The solution that was obtained for the MILP problem rendered a minimal set of convex transforms for the nonconvex MINLP problem. Their method included the αBB convex reformulation technique, which made it possible to obtain a set of transformations for any MINLP problem containing nonconvex twice-differentiable functions.

d) Reformulation-Linearisation Methods (RLT)

RLT methods generate tight linear programming relaxations in order to design heuristic procedures for discrete and continuous nonconvex programming problems (Sherali & Liberti, 2007). The method consists of two basic steps known as reformulation and linearisation. Given a mixed 0-1 linear program and n binary variables, additional constraints are created by multiplying the constraints by product factors of binary variables x and their complements $(1-x)$ in the reformulation step. The linearisation step replaces the continuous variables for each product of variables by means of McCormick (1976) over and under estimators or any linearisation method. This results in a hierarchy of linearisation, which is dependent on the form of the product factors, employed. RLT generates an explicit algebraic characterisation of the convex hull which is available at the highest level, level-n. This method can be applied to discrete optimisation problems where the bound-factors are replaced by suitable Lagrange interpolating polynomials (Sherali & Liberti, 2007).

Quesada and Grossmann (1995) used the RLT for the linearisation of bilinear terms. In their method, the bilinear terms are eliminated by creating a convex solution space. This was achieved by substituting the bilinear term with four constraints that contained the upper and lower bounds of each continuous variable within the bilinear term. This technique, however, was not an exact linearisation technique, even though a convex solution space was created from the method. Like the method proposed by McCormick (1976), overestimating and underestimating envelops are created around the nonlinearities. The result of the LP model was then used as a starting point for the original NLP model. Quesada and Grossmann (1995) then showed that, if the solution of the LP and NLP match, then the solution is a globally optimal solution. If the solutions, however, did not match, then the locally optimal solution found was therefore not a globally optimal solution. Figure 2.3 shows a general algorithm procedure for the RLT.

Meyer and Floudas (2006) presented a global optimisation algorithm based on the piecewise RLT based on the approach by Sherali and Alameddine (1992). The method was applied to a complex generalised pooling problem. Binary variables were used to indicate the existence of treatment units, which were described by a removal ratio. This therefore rendered the problem an MINLP with nonconvex bilinearities. In the reformulation stage, nonlinear constraints were formed by multiplying groups of valid constraints from the original formulation. In the linearisation stage, every product was substituted for a new variable and new constraints were added by multiplying inequality constraints on the bounds in order to generates lower bounds and produces an MILP model. The method was therefore able to reduce the gap between the lower and upper bounds by augmenting the lower bounding problem by using 0-1 variables to partition the continuous space. The method, when applied to a complex industrial case study (multiple contaminants), was able to generate tight lower bounds. This method was, however, computationally expensive due to the increase in the number of 0-1 variables that were used.

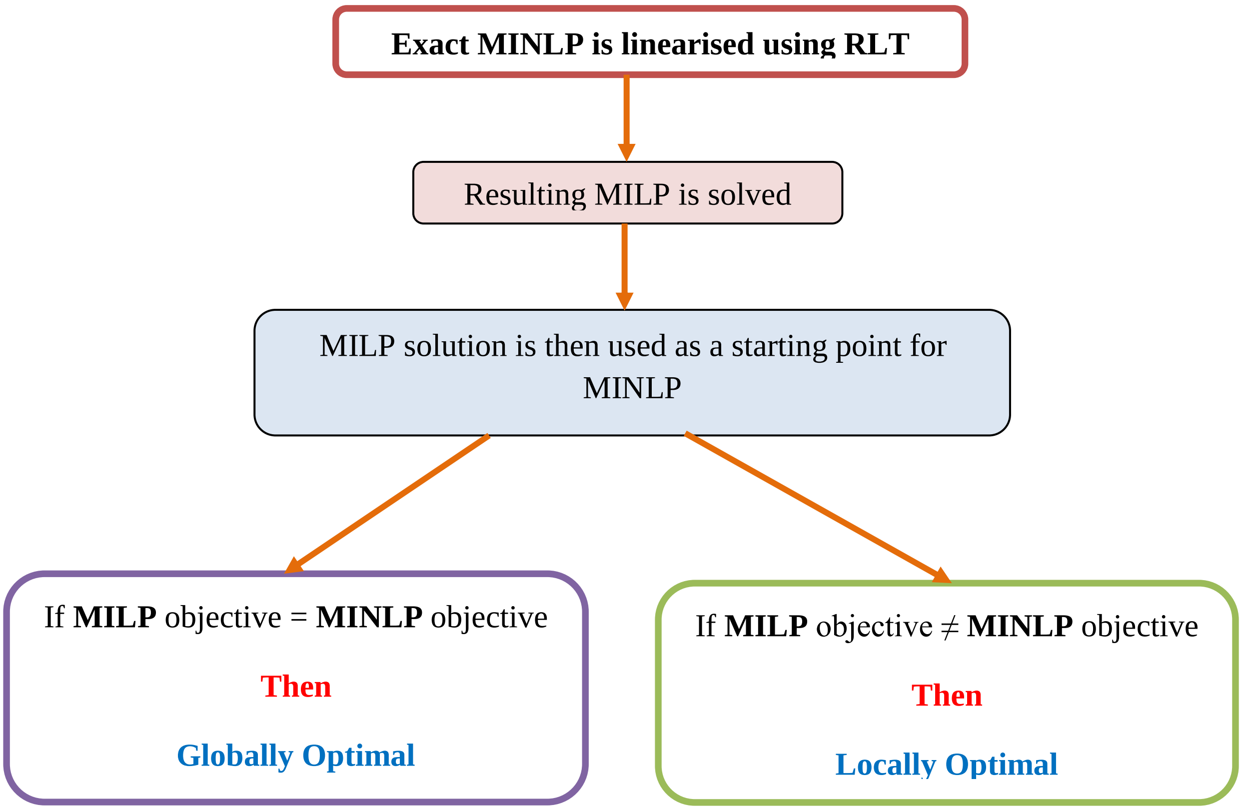


Figure 2.3: Solution algorithm for RLT by Quesada and Grossmann (1995).

e) Piecewise linear approximation

This method solves MINLP problems by approximating all the nonlinearities as piecewise linear functions as can be seen in Figure 2.4 for the approximation of *f(x)* by *L(f(x))*. The benefit of this method is that, the piecewise linear functions can be modelled by linear constraints in mixed integer variables which in turn opens the possibility of applying MILP solvers to the approximated MINLP (D'Ambrosio et al., 2015). This is achieved by partitioning the domain of a univariate (function, polynomial or an expression of only one variable) function into several intervals. The function can then be approximated by means of a line segment that connects the end points of the intervals known as breakpoints. The accuracy of the approximation is therefore dependent on the number of breakpoints. This method can also be applied to multivariate functions by portioning the domain (instead of intervals for univariate functions) of the function into several simplices and then approximating over each simplex with an affine function (function is the composition of a linear function with a translation function).

The globally optimal solution obtained for the MILP is, however, not necessarily a global or local optimal solution for the MINLP as the method only approximates the original problem (D'Ambrosio et al., 2015). Figure 2.5 shows the general framework for incorporating piecewise-affine relaxations into a GO algorithm within a spatial branch and bound framework (Khor et al., 2014).

Karuppiah and Grossmann (2006) introduced a new deterministic spatial branch and contract algorithm in order to obtain a global optimum solution for the minimisation of freshwater for the design of integrated water systems which combines both water using and water treating operations within a superstructure. The model was first formulated as an NLP problem and then modelled as a general disjunctive program (which is an MINLP) in order to allow for the selection of different technologies. A general disjunctive programming problem uses logic-based methods to represent discrete and continuous decisions. Piecewise linear under- and over-estimators were used to approximate the nonconvex terms by means of McCormick convex and concave envelops. This resulted in an MILP problem whose solution was used as a tight lower bound for every node within the spatial branch and bound tree. The lower bounds were then compared to the upper bounds (obtained by solving the nonconvex problem) within a branch and bound enumeration. The lower bound tightening cuts they proposed was only applicable for fixed load formulation and was also computationally expensive.

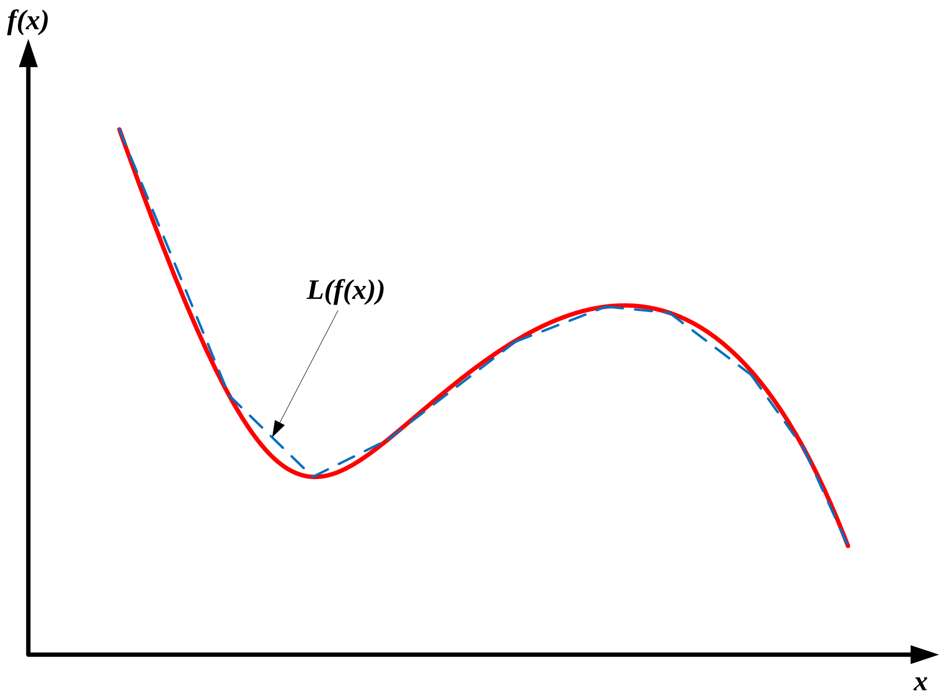


Figure 2.4: Piecewise linearisation of *f(x).*

This method has been used by many authors in order to obtain a globally optimal solution for the WN problem (Faria & Bagajewicz , 2011; Gounaris et al., 2009; Misener & Floudas , 2013).

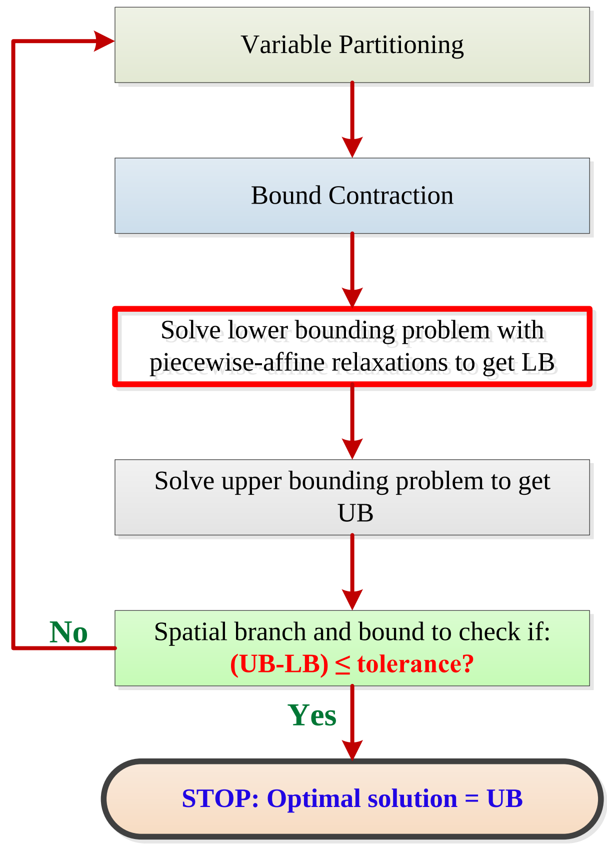


Figure 2.5: General framework for incorporating piecewise-affine relaxations into GO algorithm (Khor et al., 2014).

2.4 Wastewater Minimisation

Wastewater is generated in industry by processes and their utilities. Reducing the amount of wastewater, however, affect both effluent treatment and freshwater costs. Wastewater minimisation involves the reduction in freshwater consumption and wastewater generation. This is achieved through water reuse, water recycle and water regeneration (Wang & Smith, 1994). Figure 2.6 illustrates the different methods used in the minimisation of wastewater.

(i) *Water Reuse:* Water reuse involves the use of wastewater in other operations except the process where it was originally used. These operations do not need freshwater. This process leads to a reduction in the effluent volume, but the contaminant mass load is often unchanged. This principle is illustrated in Figure 2.6(a).

(ii) *Water recycling:* Water recycle, however, allows the effluent to be used in any process including the process in which it was produced. This principle is illustrated in Figure 2.6(b).

(iii) *Regeneration reuse:* During regeneration reuse, wastewater is regenerated by partial treatment to remove the contaminants. Water is therefore regenerated to be used in other operations. The regenerated water does not, however, go back to the operation it was originally used for. The benefit of this process is the volume of the freshwater used, the wastewater generated and that the contaminant mass load in the wastewater will decrease. Partial treatment can be achieved by using water purification units often classified as membrane and non-membrane processes, e.g. RO membranes and steam stripping respectively (Khor et al., 2011). This principle is illustrated in Figure 2.6(c).

(iv) *Regeneration recycling:* In regeneration recycling, water is regenerated and can be used in any process. The regenerated water can, therefore, be recycled to processes in which it had been used previously. This therefore means that the freshwater volume required, effluent volume and the contaminant mass load in the wastewater will most probably all be decreased by more than that achieved with regeneration reuse (Wang & Smith, 1994a). This principle is illustrated in Figure 2.6(d).

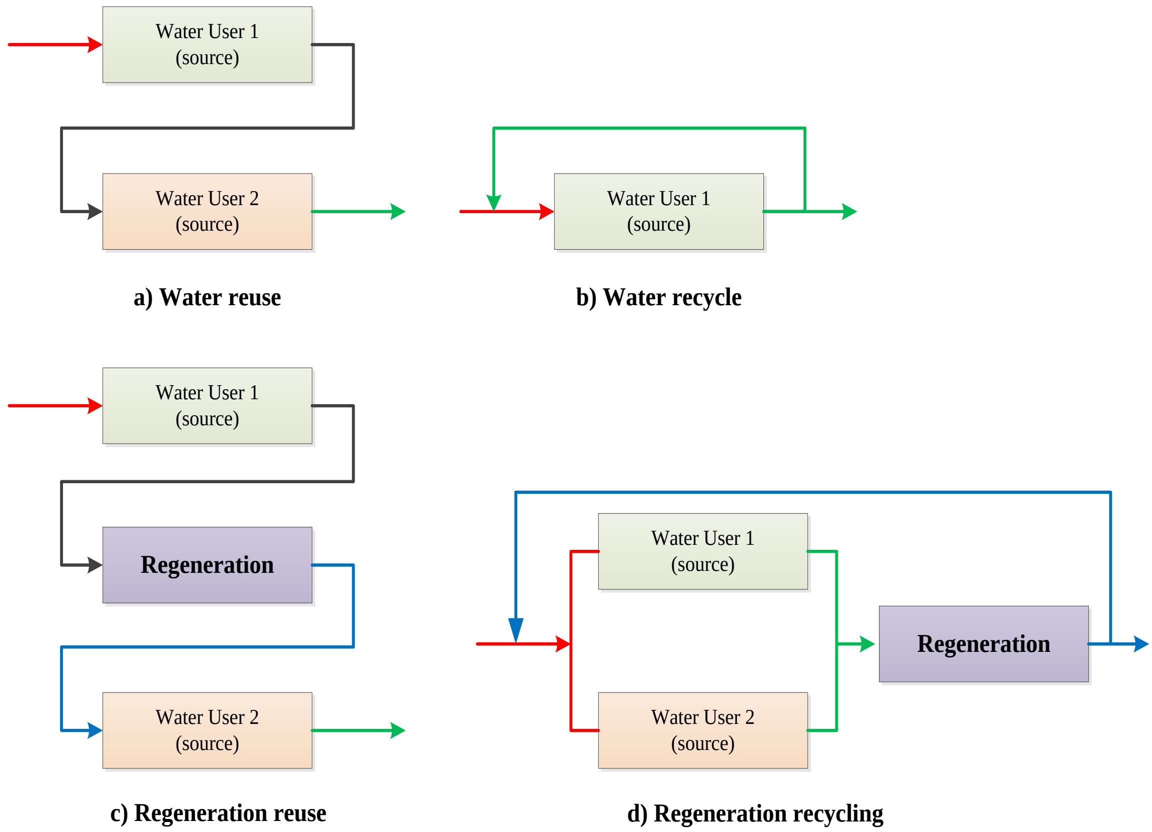


Figure 2.6: Schematic representation of different ways of achieving wastewater minimisation (Khor et al., 2014).

2.5 Water Networks

The idea of reuse, regeneration reuse and regeneration recycling is incorporated in a water network superstructure. The water network (WN) is a collection of processes that require or produce water called water-using processes and operations that clean wastewater known as regenerators (treatment units). Most works involve the use of WN for continuous operation mode (Jeżowski, 2010). The aim of WN synthesis is to synthesise a network which integrates water using operations and/or water treatment operations by optimising an objective function, which is based on economics and/or environmental sustainability while obeying certain discharge limits to the environment (Khor et al., 2014).

In order to define and solve a WN design problem, the following minimum information is needed: concentration of contaminants, mass loads of contaminants transferred to the water, or the flowrate of the water streams (Jeżowski, 2010). This therefore means that either mass load or flowrate data are required in order to solve a WN problem.

2.5.1 Basic elements of WN

A WN also consists of freshwater sources and wastewater as well as wastewater disposal sites. It also encompasses mixers and splitters for the distribution of streams within the network.

Water using processes are classified into mass transfer operations and non-mass transfer units. Mass transfer operations are also known as quality controlled or fixed load processes and involve the mass load of contaminants that have to be carried by a medium such as water. Examples of mass transfer operations are absorption, liquid-liquid extraction and fractional distillation (Treybal , 1981). Non-mass transfer operations are also known as quantity controlled or fixed flowrate operations. These operations are further classified into sources and sinks (Jeżowski, 2010).

Water sources are processes that supply streams with a fixed flowrate and contaminant concentration that enable direct reuse/recycle or regeneration–reuse/recycle. The freshwater source has an unknown flowrate. One of the objectives of solving WN optimisation problems is to minimise the amount of freshwater needed. The other aim of the optimisation is also to determine the optimal split ratios of the water source flowrates at a particular contaminant concentration for regeneration and for use by the water sinks (Khor et al., 2012).

Water sinks are water-using units or operations that use water from the sources or the regenerators. The water sinks have a fixed known flowrate and a maximum allowable contaminant concentration limit. The optimisation therefore aims at finding the optimal mixing ratios of the source and regeneration streams that are needed for reuse/recycle in the sink operations. The water sinks also consist of a wastewater stream that consists of streams from the regenerator or from the sources to be discharged to the environment (Khor et al., 2012).

Fixed load processes can be transformed into fixed flowrate operations in a case where a single contaminant problem is considered. This is achieved by dividing the fixed load process into a pair of sink-source and then setting their concentrations (inlet and outlet) to a maximum (Poplewski et al., 2010). This method is needed to reach the minimum freshwater intake for single contaminant cases according to the necessary conditions of Savelski and Bagajewicz (2000) and is valid if freshwater minimisation is the only objective of the problem. Savelski and Bagajewicz (2003) also showed that a “key” contaminant is a necessary condition for freshwater minimisation for multiple contaminant problems. The key component is however unknown for multiple contaminant problems and as such, fixed load processes cannot be transformed to fixed flowrate operations for multiple contaminants.

It should also be noted that a WN problem (non-mass transfer processes) for a case where there are no regenerators and the optimisation criterion is only the minimisation of freshwater flow, the problem becomes a linear problem (Jeżowski, 2010).

2.5.2 Superstructures

WN optimisation methods are often applied to a superstructure. A superstructure encompasses all the feasible structures of a particular network (Jeżowski, 2010). It is used to identify the optimal configuration for the process from a number of alternatives. Superstructures therefore generate multiple alternative solutions, which are then used by the designer to make a well informed decision (Alnouri & Linke, 2012). The WN superstructure consists of water sources, water sinks and water regeneration units. An example of the WN superstructure with a regeneration unit is shown in Figure 2.7 (Khor et al., 2011).

The treatment units can be centralised or distributed. In a centralised treatment system, wastewater from different operations is mixed and then treated in one centralised treatment facility (Wang & Smith, 1994b). Distributed effluent treatment systems lead to lower capital and operating costs (compared to centralised treatment) as streams are treated separately or partially mixed which reduces the flowrate to be treated. The capital cost reduction is proportional to the flowrate of wastewater in most treatment operations. The operational cost, however, increases with decreasing concentration for a specific mass of contaminants.

This therefore means that streams should be segregated and only combined where appropriate (Galan & Grossmann, 1998).

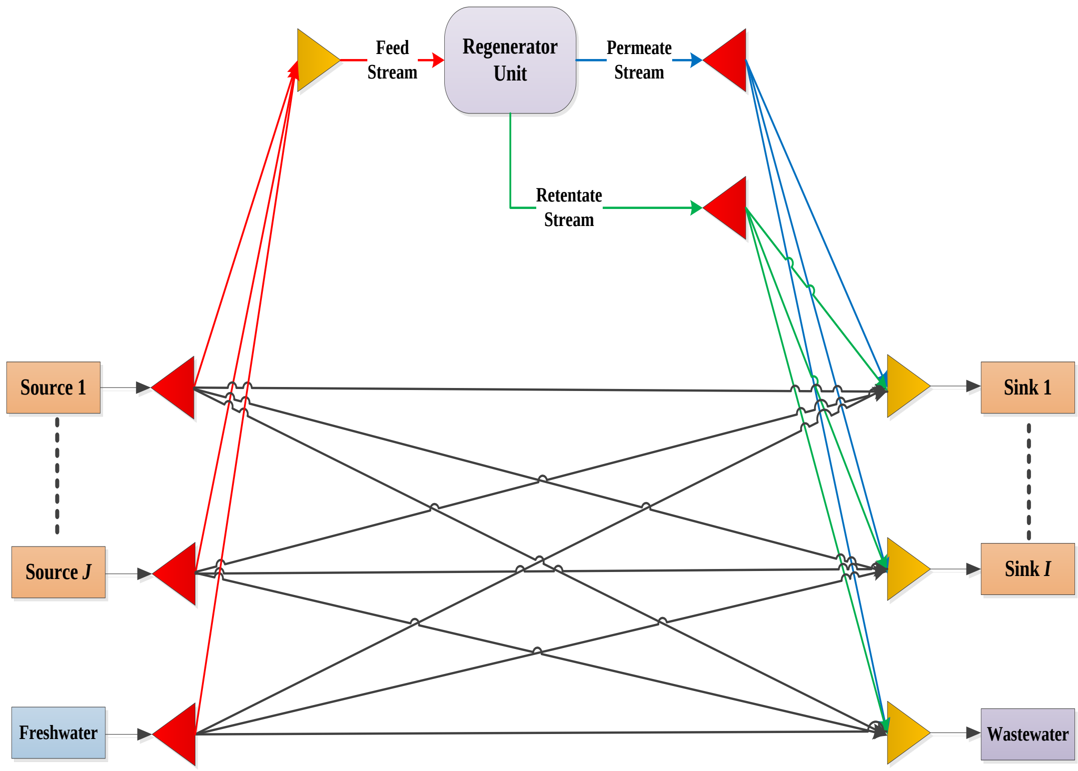


Figure 2.7: General representation of a WN superstructure (Khor et al., 2011).

2.5.3 Optimisation of WN

There are two major approaches adopted in addressing water network synthesis namely, insight based techniques and mathematical model-based optimisation methods. These two methods will be discussed in detail in the preceding sections. Insight based techniques will be addressed in Section 2.6 and mathematical model-based optimisation methods will be discussed in depth in Section 2.7. A combination of these two methods can be used to solve water network problems (Jeżowski, 2010).

2.6 Insight based techniques

2.6.1 Basic concept

Insight based techniques involve water pinch analysis, which is a graphical method based on the concept of a limiting water profile which is the most contaminated water that can be fed into a particular operation (Wang & Smith, 1994a). During water pinch analysis, the amount of freshwater that is needed has to be determined in order for specific targets to be set.

El-Halwagi and Manousiouthakis (1989) were the first to use this method for a mass exchange between a set of rich and lean streams. Their work defined a minimum allowable concentration difference, which was applied throughout the mass exchange network and was applied to a single key component. El-Halwagi and Manousiouthakis (1989) then extended their method to include regeneration. The model used a variety of mass transfer agents for the lean streams and with the aid of mathematical optimisation which, sought to design a mass exchange network and then to minimise the annual cost of the system by allocating the right mass exchange agents. This method was, however, complicated.

Wang and Smith (1994a) developed a limiting water composite curve for minimising the generation of wastewater when water is the only lean stream for single and multiple contaminants. The method was based on the grounds that, all the water using operations require clean water and can handle a maximum level of contamination. In their work, targets were first set, which included regeneration reuse and recycling. The method begins by developing an understanding of how the water using processes behave in an overall sense. A limiting composite curve of contaminant concentration versus the mass load is then constructed which defines concentration intervals by means of the inlet and outlet concentrations of the processes and this is illustrated in Figure 2.8(a). The operations are then combined within the concentration intervals to form the limiting composite curve (or grand composite curve), which is shown in Figure 2.8(b). This curve represents how the total system will behave if it was a single water using process.

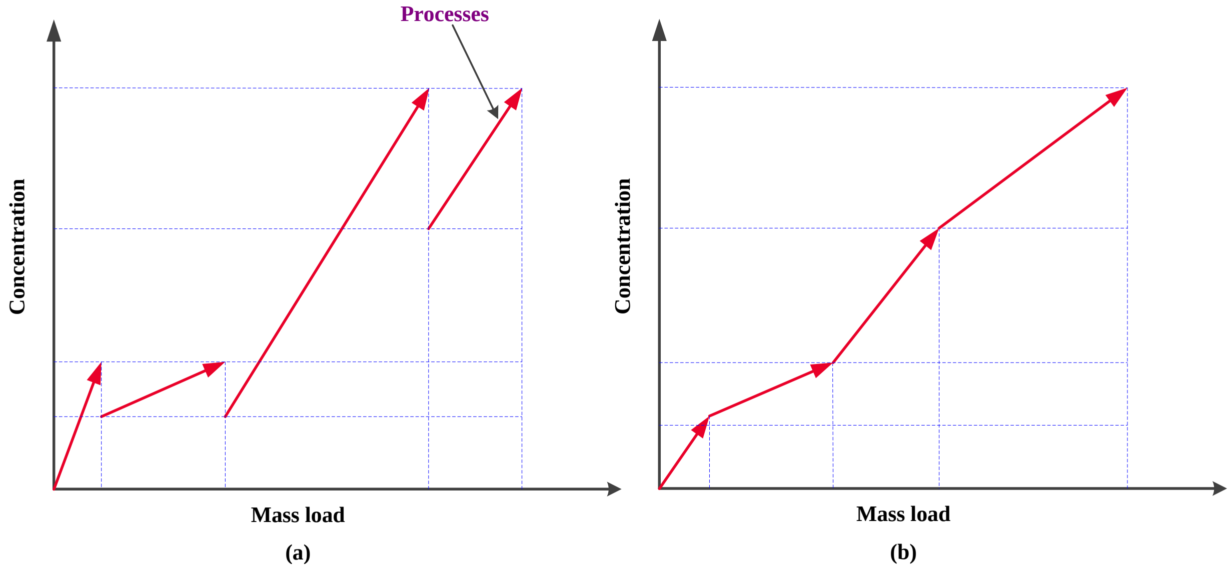


Figure 2.8: (a) Limiting water profile (b) Limiting composite curve (Wang & Smith, 1994a).

The limiting water line (or water supply line) is determined by drawing a line that is just below the grand composite curve as shown in Figure 2.9. The inlet and outlet of the limiting water line is set to zero. This therefore meant that by maximising the outlet concentrations of the water supply line, one is also able to minimise the amount of freshwater consumption and wastewater generation. The gradient of this line is inversely proportional to the flowrate. The line is used to define a boundary between feasible and infeasible concentrations. Lines below the limiting water profile signify feasible water streams and the lines above it lie in the infeasible region. It creates a pinch point where it just touches the grand composite curve (Wang & Smith, 1994a) and this can be seen in Figure 2.9. The pinch point represents the minimum feasible flowrate of wastewater. The relationship between the concentration (C), mass load (m) and flowrate (F) is shown in constraint (2.17).

$$F = \frac{\Delta m}{\Delta C} \tag{2.17}$$

It can be seen from constraint (2.17) that once the concentration and mass loads are known, the flowrate can be determined and vice versa. The driving force for water pinch is, therefore, the change in contaminant concentration. The flowrate corresponding to the maximum inlet and outlet concentrations is known as the limiting flowrate (Doyle & Smith, 1997). The minimum flowrate is therefore defined, as the flowrate required if the operation is supplied with pure water.

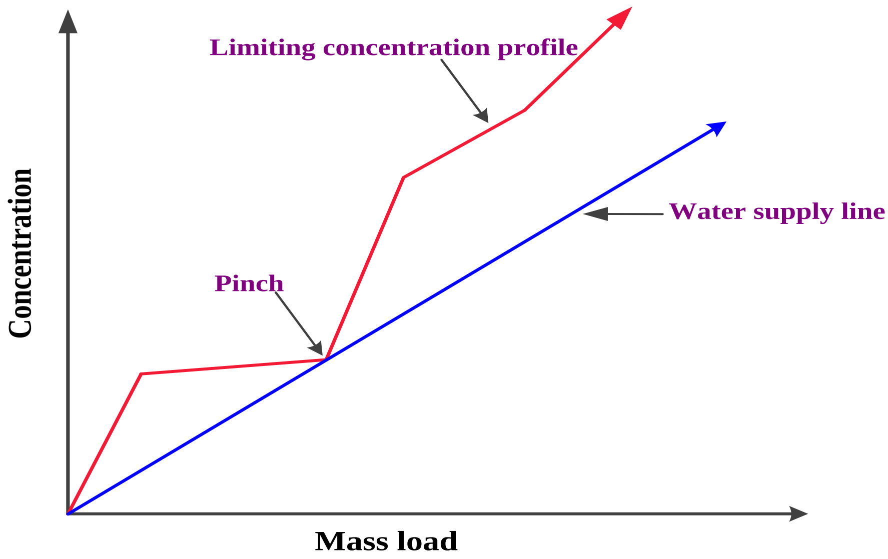


Figure 2.9: Grand composite curve with the water supply line for targeting minimum water flowrate (Wang & Smith, 1994a).

Once the minimum amount of freshwater is known for the entire design and for each process, the design that meets the target can therefore be determined. The amount of freshwater required below and above the pinch was determined by constraint (2.17). Each operation is then drawn with respect to the concentration intervals. The final design is drawn with reuse and recycling of streams. In the design, the amount of freshwater used was equal to the amount of wastewater generated from the system (Wang & Smith, 1994a).

The advantage of the method was that no knowledge of the equipment performance or mechanics of the mass transfer was needed as it only required a limiting maximum concentration for each process.

Wang and Smith (1994b) then extended the method for single contaminants to multiple contaminants. Their method initially sets targets for each contaminant in isolation. The highest flowrate that was obtained for all the contaminants for a treatment process was then taken as the target for the treatment process. Network designs were obtained for each contaminant in isolation and a final network was then obtained by merging all the subnetworks together. This therefore meant that each contaminant was taken into account for targeting and design. This method, however, becomes unsuccessful when applied to large and complex problems. There are also a number of drawbacks in the methodology presented by Wang and Smith (1994b):

(i) In some cases, the method fails to give the best targets as the pinch position could move to different positions after regeneration (Kuo & Smith, 1998b).
(ii) It was difficult to apply the method to cases involving multiple contaminants.
(iii) Stream splitting was also allowed in the operation. This was, however, impractical as some operations might require more water over and above the amount predicted by the targets that were set (Kuo & Smith, 1998b).

2.6.2 Extension of the water pinch method

The work proposed by Wang and Smith (1994a) was modified and extended by many authors in order to improve its application for the minimisation of wastewater for both single and multiple contaminants for mass transfer operations.

Wang and Smith (1994b) extended their methodology to distributed effluent treatment systems. In distributed treatment systems, streams are segregated for treatment and are only combined if appropriate. The method they proposed was, however, a general approach for both centralised and distributed systems. Targets are first set for the effluent flowrates through the treatment processes in order to determine the minimum treatment costs in the case of single contaminants. Design rules are then used for the design of the final networks. These rules were, however, based on the location of the pinch for the particular system. In their design, streams starting above the pinch for the treatment system were treated fully. Streams starting at the pinch were partially treated and streams starting below the pinch completely bypassed the treatment units. The performance of the treatment processes was defined by either an achievable outlet concentration or a RR. The methodology could also predict the number of treatment units needed. The method was extended to multiple contaminants by an extension of that used for single contaminants. Subnetworks were generated for each contaminant and the final network is achieved by a combination of the different subnetworks. This methodology, however, had several drawbacks.

(i) A detailed design was not used to represent the treatment processes.

(ii) In the development of the multiple contaminant model, a single treatment unit was assumed for each targeting stage and was therefore not known beforehand. The performance of the treatment processes was also fixed.

(iii) The prediction of the lowest possible flowrate was not always possible.

(iv) Important features of the design for multiple treatment processes for both single and multiple contaminants were also not taken into consideration during the design process.

Kuo and Smith (1997) pointed out the above draw backs of the method proposed by Wang and Smith (1994b) by introducing a modified method for the design of the distributed effluent treatment systems and extended the concepts to retrofit cases. The method was able to choose the appropriate type and number of treatment operations. Their methodology also allowed the effluent streams to reach their consent limits for discharge at a minimum cost. This was achieved by setting targets for minimum flowrates in the case of single contaminants where the optimum solution could be achieved. In the case of multiple contaminants, the treatment network was achieved by means of a repeated use of targets and design. The methodology was based on a "composite effluent curve" instead of the water supply line.

Kuo and Smith (1998a) then looked at the interactions between water use and effluent treatment systems in the process industries for the design of minimum water use. Their method used a conceptual and graphical approach based on the "composite effluent curve". They introduced a new method for the design of the water networks, which does not only achieve the target for minimum water consumption, but also leads to the lowest effluent treatment cost. This new approach involved the construction of the "water mains" which helps satisfy the requirements of each operation. Water mains (imaginary) were used to store the minimum freshwater below and above the pinch. Figure 2.10 shows the water mains used for the design of the WN. From Figure 2.10, the intermediate water main acts as a source and a sink. From this design, the final WN is then drawn. Kuo and Smith (1998b) then expanded on this method to include regeneration reuse and recycling.

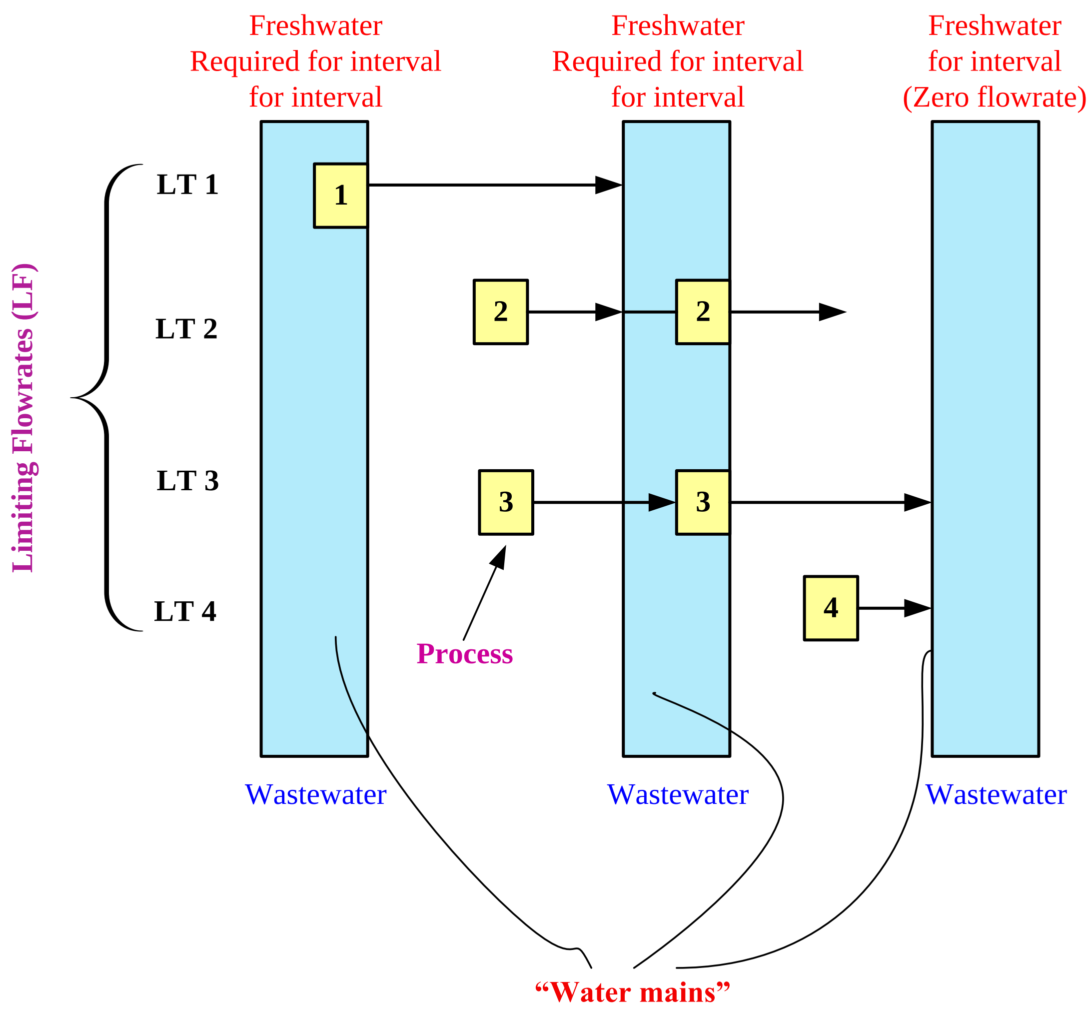


Figure 2.10: Design grid (Kuo & Smith, 1998a).

2.6.3 Water pinch for non-mass transfer processes

The methodology proposed by Wang and Smith (1994a; 1994b) and Kuo and Smith (1997; 1998a; 1998b), however, treats the water using processes as mass transfer operations. According to Dhole et al. (1996), most process units (reactors, boilers, cooling towers etc.) cannot be modelled as mass transfer operations. This is because these operations are based on flowrate of water rather than the amount of contaminants. Also, mass transfer models cannot be easily adapted to situations where several aqueous streams enter and leave a unit at different concentrations and changes in water flowrate are also not easily accounted for within the model formulation (Dhole et al., 1996).

Dhole et al. (1996) therefore proposed a new graphical technique to overcome the limitations experienced by mass transfer models. In their method, each operation was considered to have an inlet and outlet stream (different concentrations and flowrates). The input streams were all plotted together to form a demand composite curve and the water sources formed the source composite curve. Their graphical technique represented concentration versus flowrate and not the original concentration versus mass load used by previous studies and this is shown in Figure 2.11. The two composite curves were then plotted together on the same axes and were then shifted together until they just touched. The point at which the two curves touched was identified as the water pinch and the potential for water reuse was therefore identified. They therefore concluded that freshwater should be added below the pinch and that sources above the pinch should not be discharged as wastewater in order for targets to be met.

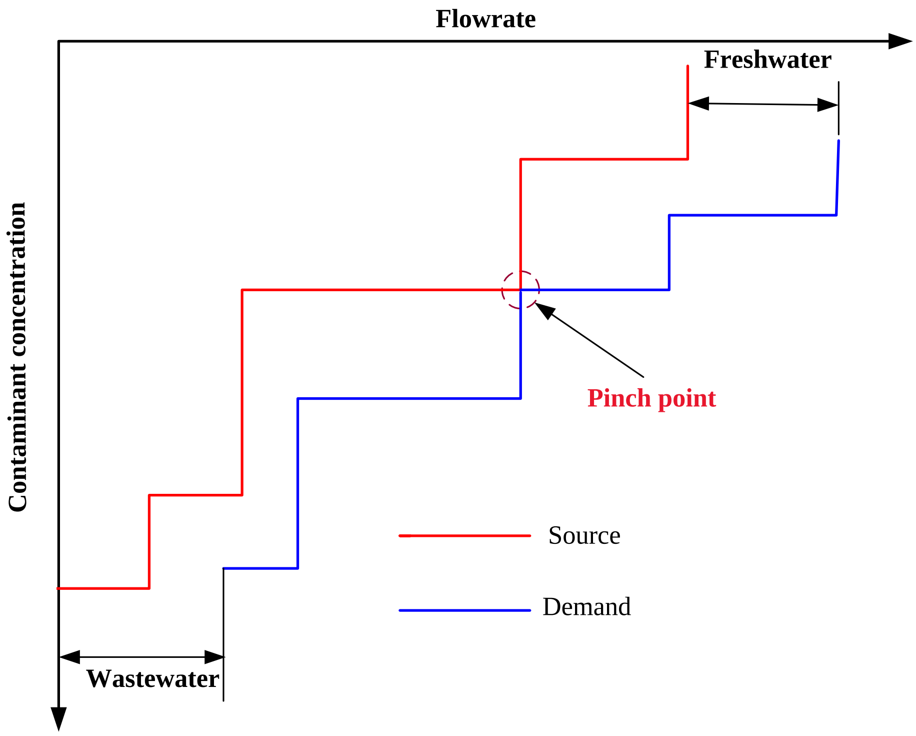


Figure 2.11: Water composite curve representation of Dhole et al. (1996).

The method proposed by Dhole et al. (1996) , however, did not provide a systematic method for the elimination of pinch points by mixing. This is because the mixing of water sources could easily change the shape of the source composite curve and as a result change the targets. The targets given by Dhole et al. (1996) can, therefore, not be considered as true targets due to these limitations (Hallale , 2002).

Hallale (2002) then proposed a graphical method that was based on non-mass transfer operations with single contaminants in order to overcome the drawbacks of the methodology proposed by Dhole et al. (1996). This was firstly achieved by plotting a new demand and supply curve, which had purity of water on the vertical axis rather than the contaminant concentration and this, is shown in Figure 2.12(a). The initial value of freshwater flowrate was now assumed as can be seen from Figure 2.12(a). The assumed value for freshwater was then tested (too high or too low). This was achieved by knowing that sufficient pure water is needed at all points within the network. A new diagram known as the Water Surplus diagram was therefore constructed to account for all possible missing arrangements. This diagram was plotted by knowing from Figure 2.12(a) that there are regions that lie below (deficit of pure water) and above (surplus of pure water) the demand composite. The regions are illustrated in Figure 2.12(a).

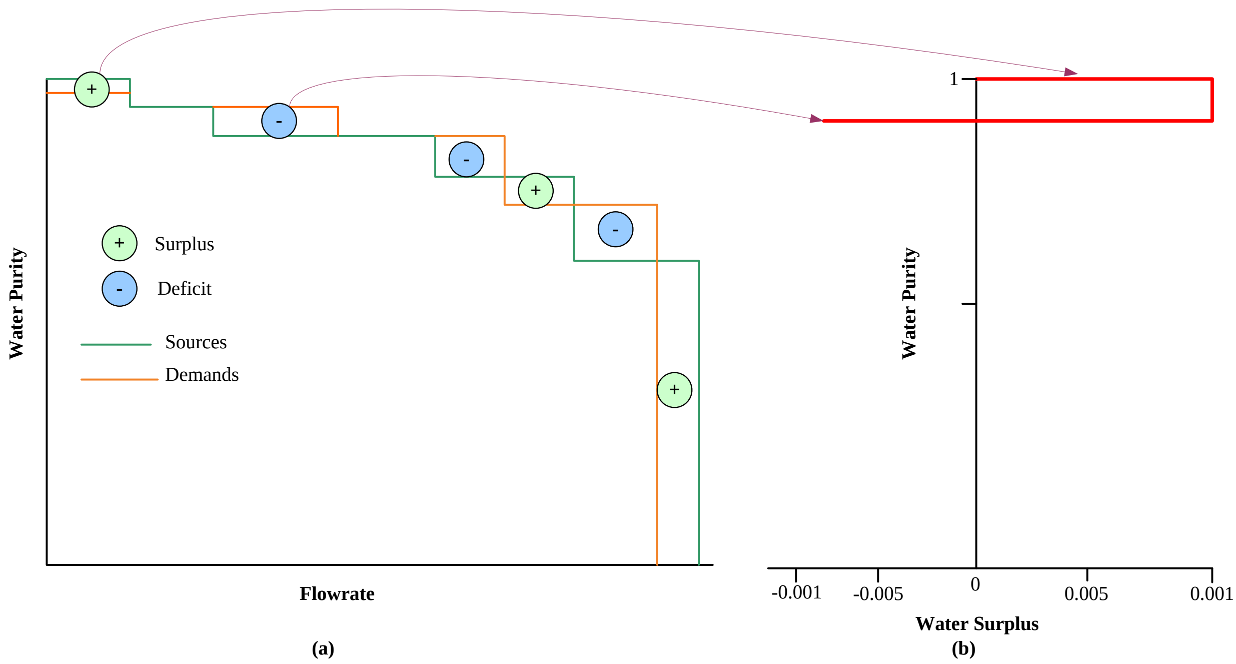


Figure 2.12: An illustration of the surpluses and deficits plotted to form the water surplus diagram (Hallale , 2002) .

The surplus or deficit of pure water for each region was then determined by calculating the area enclosed by each rectangle (Hallale , 2002). These values were plotted against the water purity to form the water surplus diagram shown in Figure 2.12(b). From Figure 2.12 (b) it can be seen that the cumulative surplus is plotted and if the region of deficit is greater than the previous value, the graph moves in the negative water surplus direction. The complete water surplus diagram is illustrated in Figure 2.13(a).

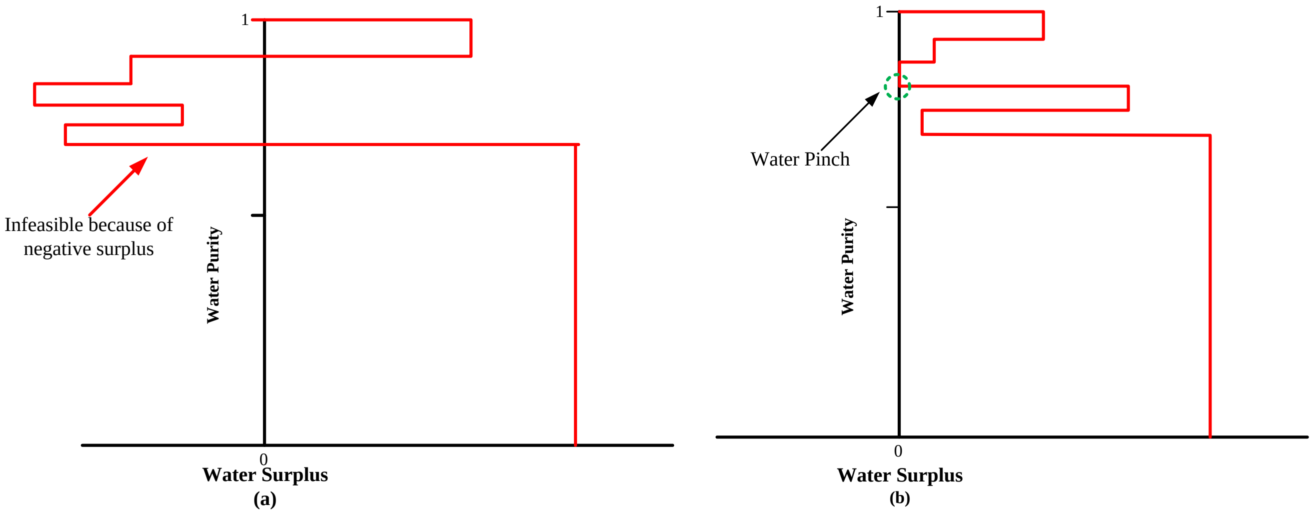


Figure 2.13: (a) The complete surplus diagram. (b) The freshwater flowrate is increased until surplus diagram becomes pinched.

Figure 2.13(a) shows that the assumed value for the freshwater flowrate causes part of the plot to lie in the negative region, which is an indication of insufficient water purity at all points within the network. This therefore means that more freshwater is needed until no part of the surplus diagram lies in the negative region. The minimum water flowrate is therefore the flowrate that causes the surplus diagram just to touch the vertical axis and this is illustrated in Figure 2.13(b). The method also gives flowrate targets that ensure that all demands with regards to flowrate and purity are satisfied. This method helps determine the true pinch points and reuse targets as it incorporates all the mixing opportunities.

The method by Hallale (2002) was able to deal with a wider range of water using operations but was, however, limited to single contaminant cases. Hallale (2002) also built a mathematical model to determine the minimum freshwater and wastewater flowrates. According to Hallale (2002), insight based methods have an advantage over mathematical models as they:

(i) provide increased insight into the problem
(ii) give clear guidelines about the process modifications that are beneficial to the designer.

(iii) have a lower computational burden.

The above analysis, however, shows that the water pinch method proves unsuccessful for complex problems involving multiple contaminants (Faria & Bagajewicz , 2009) and various topological constraints (Khor et al., 2012). The methodology cannot allow for constraints other than concentration and flowrates. Constraints for forbidden matches, safety and distances between processes cannot be specified by the water pinch method (Doyle & Smith, 1997). Insight-based techniques therefore offer good insights into the water network problem with a low computational burden but require a significant amount of problem simplification (Khor et al., 2014).

2.6.4 Recent works on water pinch methods

Recent studies have extended water pinch analysis to algebraic methods primarily water cascade analysis (WCA) (Ng et al., 2007; Manan et al., 2004). The WCA is a numerical technique that is used to establish the minimum water and wastewater targets in a maximum water recovery network. This method establishes the minimum targets by looking at the possibility of using available water sources within the process in order to satisfy the water demand (Manan et al., 2004). The method is advantageous to other previous methods as it determines the exact utility targets and pinch locations without any tedious iterative steps. It can also be applied not only to non-mass transfer operations, but to a wide range of water using operations (Manan et al., 2004).

2.7 Mathematical model optimisation methods

Mathematical optimisation is capable of handling WN problems in their full complexity by considering a wider range of constraints in the objective function, multiple contaminants, representative cost functions as well as various topological constraints (Khor et al., 2014) can be included in the model formulation. Optimal water allocation and treatment is therefore moving towards the use of mathematical techniques also due to the tedious nature of insight based techniques (Karuppiah & Grossmann, 2006) .

The mathematical optimisation approaches employ a superstructure, which identifies an optimal configuration for the process from a number of alternatives. This idea was first proposed in the work of Takama et al. (1980). In their work, they proposed a nonlinear model that incorporates both water using and wastewater treating units for multiple contaminant systems. The objective of the work was the minimisation of freshwater consumption and wastewater generation for a refinery problem. This was achieved by the removal of uneconomical and irrelevant connections from the superstructure. The model was transformed into problems without inequalities by means of penalty functions and was solved using a method they proposed. Their method was, however, restricted and included a centralised treatment system (Gunaratnam et al., 2005) and was highly complex to apply. The method was also applied to small-scale problems. The solution they obtained was feasible, but was far from the optimum solution.

WN problems therefore result in NLP and MINLP models. Binary variables are needed to account for the existence of units, streams, piping interconnections and for topological constraints. This therefore results in an MINLP model, which is difficult to solve when a global optimum is desired (Grossmann & Biegler , 2004). The complexity arises due to bilinear terms (which create nonconvex functions) in the mass balance equations and the concave cost terms in the objective (Ahmetović & Grossmann , 2010), which result in nonconvexities within the model. The complexities are also due to the existence of integer variables, nonlinearities and nonconvexities within the model (Ahmetović & Grossmann , 2010).

Nonconvex models give rise to many suboptimal solutions and lead to certain complications that cause the failure of most local optimisation models (Zamora & Grossmann, 1998). In the absence of convexity, NLP methods fail to locate the global optimum solution (Ryoo & Sahinidis, 1996). This difficulty can, however, be handled in a number of ways (Jeżowski, 2010) through direct linearisation, generating a “good” starting point, using sequential solution procedures and by means of global (deterministic) optimisation methods.

2.7.1 Direct Linearisation

This method involves the linearisation of the nonlinear terms in the mathematical model. This is achieved by the selection of linear conditions for optimality. In the context of WN, linearity constraints exist for non-mass transfer processes as well as processes with or without regeneration, which are defined by fixed outlet concentrations (Jeżowski, 2010). Relaxation methods proposed by McCormick (1976) and Glover (1975) can be used to linearise an MINLP problem. Different methods for linearising NLP and MINLP models have been proposed over the years.

Bagajewicz and Savelski (2001) showed that a WN with mass transfer processes and single contaminants can easily be linearised when freshwater minimisation is the only objective of the optimisation. They proposed an iterative method, which involved LP formulation for the optimal solution of the single contaminant problem and an MILP for the design of the different possible network alternatives. Their method was based on the previously developed necessary conditions of optimality. Partial regeneration of wastewater was also considered in the formulation. In the case where no regeneration was considered, a sequential two-step procedure was proposed in which the LP (freshwater minimisation) solution was made the starting point of the MILP, which minimises the number of interconnections. The bilinearities were eliminated in this case by setting the outlet concentrations to their maximum values. In the case where regeneration was considered, an additional step which involved the MILP solution being the starting point of another LP with the objective of determining the minimum amount of water through the regenerator. The optimality conditions for water regeneration without recycle were also determined. This method, however, uses the fixed load method and was limited to single contaminants.

Savelski and Bagajewicz (2003) then extended the work by Bagajewicz and Savelski (2001) for multiple contaminants through the selection of a key component. Their work was the first to provide proof for optimality conditions for multiple contaminants. They proved that at least one contaminant reaches its maximum allowable concentration at the outlet of the freshwater-using process and that concentration monotonicity only holds certain key contaminants. The first condition was that, at every outlet of a partial water provider, the outlet concentration of a key component should not be lower than the concentration of the same component from the precursors. The second condition states that, the outlet concentration of a key component of a partial provider head process must be equal to its maximum concentration and the third condition was that, the outlet concentration of at least one component of an intermediate process reaches its maximum value. Regeneration of streams was, however, not considered in their work and the model was based on a fixed load model. Freshwater minimisation was the only objective of the work.

The methods provided by Bagajewicz and Savelski (2001) and Savelski and Bagajewicz (2003) provide an exact linearisation method as the method is applied to LP and MILP problems. Exact linearisation is, however, not possible for nonconvex MINLP models.

2.7.2 Generation of a "good" starting point

This method determines a global optimum or "good" optimal solutions. This is achieved by using problem linearisation to provide a good starting point for the nonconvex MINLP problem. The initial point can be obtained by stochastic optimisation or through problem linearisation. The most common practice for mass transfer water using operations is to remove the bilinear term by fixing outlet concentrations in all operations to their maximum values (Jeżowski, 2010). The initial guesses adopted for solving NLP and MINLP models have a significant impact on the convergence process and must therefore be chosen with reliable methods (Zamora & Grossmann, 1998).

Li and Chang (2007) proposed an efficient initialisation strategy to solve NLP and MINLP models for WN synthesis problems with multiple contaminants by generating near feasible guesses. The model was based on a superstructure and the initialisation strategy was based on knowing the mass load of contaminants in every water-using unit, the rate of water loss in each unit and the upper bounds of the corresponding inlet and outlet concentrations (Li & Chang, 2007). The computational time for solving the NLP and MINLP models was reduced as a result. The NLP model was, however suited for small-scale problems while the MINLP model could be used to optimise larger water using systems by including structural constraints for the simplification of the network configuration. Their method, however, did not guarantee global optimality.

Teles et al. (2008) proposed a initialisation procedure that replaces the NLP with a succession of LP models that are then solved for all operation sequences. The LP model was first relaxed and used as a starting point for the NLP model. The paper therefore looked at four initialisation methods for the NLP model which were proposed and tested. The first method looks at a single starting point by linearising the NLP by looking at the maximal concentrations or by removing connections among the fixed load operations. The other method looks at using multiple starting points. Each point is, however, related to a predefined sequence of fixed load operations and the LP model is also generated by the two methods used in the single starting point scenario. The best solutions were obtained in the case were multiple starting points where used with the maximal concentration linearisation method. This method, however, was computationally expensive. The procedure they proposed does not guarantee global optimality but provides a large probability of finding the globally optimal solution. The model does not also consider regeneration.

Galan and Grossmann (1998) looked at the optimum design of a distributed wastewater network where multiple contaminants were taken into account. They proposed an NLP and MINLP model for the superstructure that was presented by Wang and Smith(1994b). Their paper was the first to address the synthesis regeneration networks within the WN. The paper presents three formulations. The first formulation looks at an NLP model for the distributed wastewater treatment network synthesis with nonlinear bilinearities in a mixer unit. The second formulation looks at an MINLP model that employs 0-1 variables for the selection of different treatment technologies. The treatment units in this case were described by a constant removal ratio. The final formulation looks at an NLP model for membrane based treatment technologies by using short-cut design equations instead of a fixed removal ratio.

They proposed a search procedure that is based on a relaxed linear model. The LP relaxation was based on the method proposed by Quesada and Grossmann (1995). The solution from the LP model was then used as a lower bound as well as a starting point for the NLP model. Different objective functions were used in the LP model to provide different starting points for the NLP model (the best objective function was then selected). This therefore led to different locally optimal solutions. The best solution was then chosen as the upper bound for the globally optimal solution. The nonconvex exponential terms in the objective function was linearised by using linear underestimators proposed by Zamora and Grossmann (1998). The procedure was able to find near global or global optimum solutions. This method was, however, computationally demanding even though it was very effective.

NLP and MINLP models can therefore be solved with less computational time once a “good initial starting point” is provided. This therefore aids in the convergence process of the model. This method, however, does not also guarantee a globally optimal solution and also minimises the chance of a nonlinear solution becoming a local solution, which is far from the globally optimal solution (Doyle & Smith, 1997).

2.7.3 Sequential solution procedures

This involves the use of iterative methods. With regards to WNs, the concentration intervals are divided into smaller intervals until convergence is achieved. The work by Takama et al. (1980) was the first to use sequential optimisation procedures for solving WN problems.

Doyle and Smith (1997) then presented the first model for a sequential superstructure optimisation approach for WN synthesis, which was based on an iterative procedure. The superstructure used considered direct reuse and recycle. The solution procedure they proposed, involves a sequential procedure that uses a linear programming approximation (LP) as an initial guess to solve an NLP. The model considered multiple contaminants and water regeneration was not considered. The Linearisation was based on assuming a fixed maximum outlet concentration and the water using processes were then modelled by assuming a fixed mass load for the NLP. The LP problem is solved first and used as a starting point for the NLP problem. Convergence was, however, achieved by the introduction of additional constraints on the maximum wastewater flows and forbidden stream matches. Feasibility was also achieved by relaxing the concentration balance as an inequality. The method they proposed, however, does not guarantee a globally optimal solution, but does reduce the difficulties that are associated with NLP problems.

Gunaratnam et al. (2005) used the sequential superstructure optimisation approach to generate a WN which considers both water-using operations and water-treating systems. Their procedure was developed in three steps. In the first step, the material balance equations are relaxed by setting the outlet concentration at a maximum and introducing slack variables in order to create an MILP. In the second step, the flowrate solutions are then used as the starting point for solving the LP relaxation. This generates new concentration values that can be used in the MILP in the next step. The objective of the LP problem is to minimise the summation of the slack variables. In the last step, convergence is achieved when the sum of the slack variables becomes small and this then becomes the solution for the MINLP. The LP and MILP models are therefore solved iteratively until convergence and then used as a starting point for the MINLP model. Network complexity was also reduced through the specification of the minimum permissible flowrate, maximum number of streams allowed at a mixing point and piping costs. Binary variables are also used to enforce/eliminate certain substructures from consideration.

This method is, however, computationally demanding and does not necessarily guarantee a global optimum solution. Regeneration recycling was also eliminated in order to avoid concentration build. The number of water-treating operations was fixed and was modelled using the removal ratio. This therefore means that a detailed design was not used to describe the treatment systems. The cost of effluent treatment was also assumed to be proportional to the effluent flowrate. The methodology proposed by Gunaratnam et al. (2005) is shown in Figure 2.14.

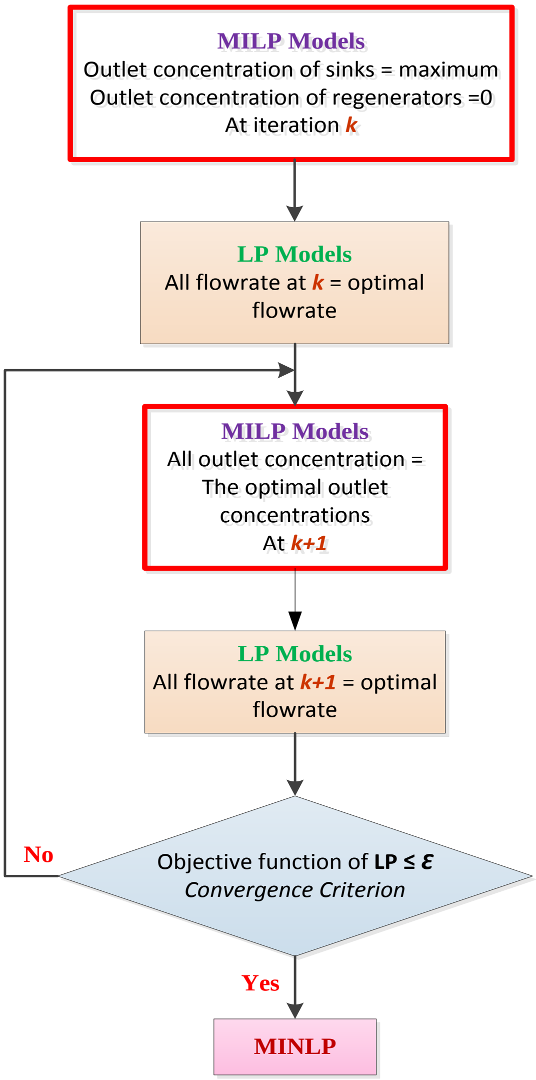


Figure 2.14: Solution strategy for sequential solution procedures (Gunaratnam et al., 2005)

The sequential superstructure optimisation approach also helps reduce the computational burden experienced in solving MINLP models. This method is, however, not straight forward as it includes different iterative procedures as can be seen from Figure 2.14. The method also does not guarantee a globally optimal solution.

2.7.4 Global (deterministic) Optimisation (GO)

GO methods can be described as either stochastic or deterministic (Grossmann & Biegler , 2004).

Stochastic Methods

The term stochastic refers to systems that are based on the theory of probability. Randomness is present in these models and unique variables are not used to describe variables as they are described by probability distributions. These methods often rely on physical analogies for the generation of trial points, which in turn mimic the approach to an equilibrium condition. They are easy to implement, but require that the problem is modelled in terms of recursive moves. This is, however, not easy with continuous variables. Stochastic methods are not rigorous and also have difficulty in handling complex constrained problems. Examples of stochastic methods include simulated annealing and genetic algorithms (Grossmann & Biegler , 2004). Global optimality with a probability approaching one can be achieved as the running time of this method can go to infinity (Ryoo & Sahinidis, 1996). This method cannot, however, guarantee convergence to a global optimum in a finite number of iterations (Zamora & Grossmann, 1998).

Deterministic Methods for NLP and MINLP

a) Basic concept

In a deterministic model, every variable is determined by parameters in the model and by the previous values of the variables. It performs the same way for a given set of initial conditions. GP algorithms are deterministic and converge to a global optimum value. Deterministic global optimisation techniques are designed to converge to a global optimum solution or to prove that a particular point (solution) does indeed exist. This is, however, achieved by making certain specific assumptions and is also restricted to specific classes of problems (Zamora & Grossmann, 1998). GO algorithms use subsolvers to solve LP and NLP subproblems. This method includes Lipschitzian methods, branch and bound (BB) methods, cutting planes methods, difference of convex and convex methods, outer-approximation methods, primal-dual methods and reformulation-linearisation methods (Grossmann & Biegler , 2004) and piecewise affine linearisation methods. Deterministic methods take advantage of the mathematical structure of the problem and often guarantee finite convergence within a particular pre-specified accuracy (Ryoo & Sahinidis, 1996).

Most GO methods work by using convex envelopes or underestimators to formulate the lower-bounding convex MINLP problems. These techniques are then combined with GO techniques for continuous variables, which are usually spatial branch and bound methods. Spatial branch and bound methods divide the feasible region of continuous variables and then compare each lower and upper bound in order to unravel each subregion. The subregion that contains the optimal solution is then found by eliminating the subregions that do not contain the optimal solution (for NLP nonconvex problems). An example of this approach is the method proposed by Quesada and Grossmann(1995).

Zamora and Grossmann (1998) were the first to propose a global (deterministic) optimisation algorithm for addressing nonconvexities in MINLP for distributed wastewater network synthesis problems (Khor et al., 2014). They were also the first to apply global optimisation to a superstructure model. The MINLP model consisted of nonconvex bilinear, linear fractional, and concave univariate objective function terms. A branch and bound based algorithm with bound contraction was proposed which led to the elimination of a large portion of the search space and a reduction in the number of nodes within the search tree.

The deterministic global optimisation methods that will be discussed are the branch and bound method, branch and reduce methods, cutting planes methods and outer approximation methods.

b) Branch and Bound (BB)

Land and Doig (1960) were the first to propose a BB method for discrete programming.BB methods develop lower (LB) and upper (UB) bounds of the optimal value of the objective function over subregions within the particular search space. Branching refers to the successive subdivision of the feasible domain while bounding refers to the computation of the lower and upper bounds for the global optimum. The branch is then checked against upper and lower estimated bounds on the optimal solution and then discarded if it cannot produce a better solution than the best solution found by the model. The main feature of this method is its ability to delete inferior subsets of the original subspace during the iteration process. During the iteration process, the subregions whose lower bounds are no longer better than the current upper bound is then deleted from the search (Ryoo & Sahinidis, 1996). Ryoo and Sahinidis (1996) then used the methodto handle nonconvex MINLP problems. The procedure for the branch and bound tree is illustrated in Figure 2.15.

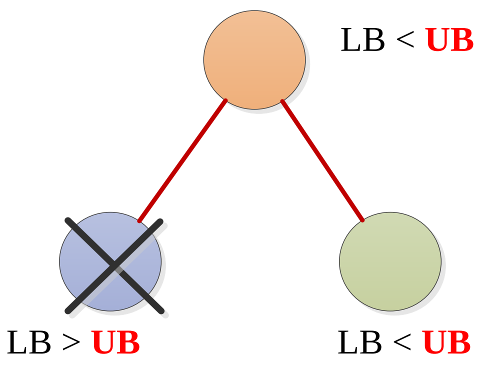
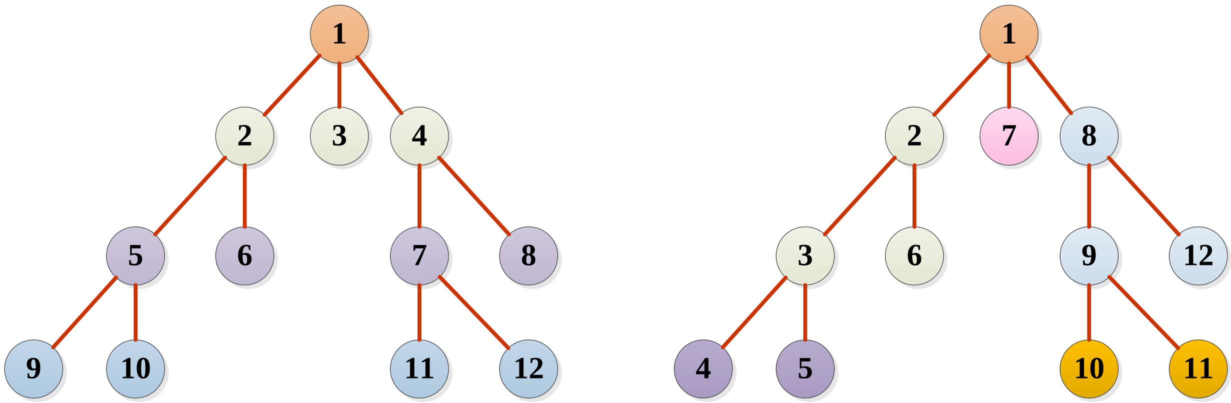


Figure 2.15: Branch and bound tree.

The branching can either take place through the depth first-approach or the breadth-first approach. The depth-first approach performs branching on the most recently created node within the tree. If no nodes are expanded, the method then backtracks to a node whose successor has not been examined. The breadth-first approach, however, selects a node with the best value at each level and then expands on all its successor nodes. The two different methodologies are illustrated in Figure 2.16. The breadth-first approach requires examination of fewer nodes and backtracking is also not required. The depth-first approach, however, requires less storage and can find the optimal solution early in the enumeration procedure. The branch and bound tree, however, uses a breadth-first enumeration (Ryoo & Sahinidis, 1996).



a) Breadth-first b) Depth-first

Figure 2.16: Schematic representation of the breadth and depth-first approach.

c) Branch and reduce (BR)

Reduction techniques are used to pre-process a global optimisation problem before the global optimisation algorithm is applied. Range reduction is used in BB algorithms to improve the performance of the bounding procedure at every node for a particular search tree (Sahinidis , 2000). The resulting algorithm is known as a branch and reduce (BR) algorithm. The reduction test is therefore applied to every subproblem of the search tree in pre-processing and post-processing steps in order to contract the space and to reduce the relaxation gap. The relaxation tests are often based on duality (Sahinidis , 2000). Certain subregions are therefore excluded by employing optimality and feasibility criteria and also refine other subregions dynamically (Ryoo & Sahinidis, 1996). The method therefore branches on the continuous and discrete variables (Grossmann & Biegler , 2004). This concept is implemented in the global optimisation solver known as BARON (Branch and Reduce Optimisation Navigator) (Tawarmalani & Sahinidis , 2005).

BARON therefore integrates the BB with a wide variety of range reduction tests (Sahinidis , 2000). Heuristic techniques are also implemented in BARON for the approximate solution of optimisation problems that yield solution bounds for the variable. This is known as feasibility based tightening. Convergence is accelerated by the incorporation of a number of compound branching schemes. Additional constraints are therefore required in order to achieve global optimality. These constraints may therefore speed up the solution time and also increase the probability of obtaining a solution. The solver also requires finite lower and upper bounds on the problem variables in order for BARON not to infer the bounds from the problem constraints. BARON does not require a starting point in solving NLP and MINLP problems (GAMS, 2013). Subsolvers for LP, MIP and NLP are incorporated in the solver.

Most GO methods incorporate the BB and BR method (Grossmann & Biegler , 2004). This includes the work of Zamora and Grossmann (1998), Karuppiah and Grossmann (2006) and Misener and Floudas (2010). Figure 2.17 describes the solution algorithm scheme used for the implementation of BARON.

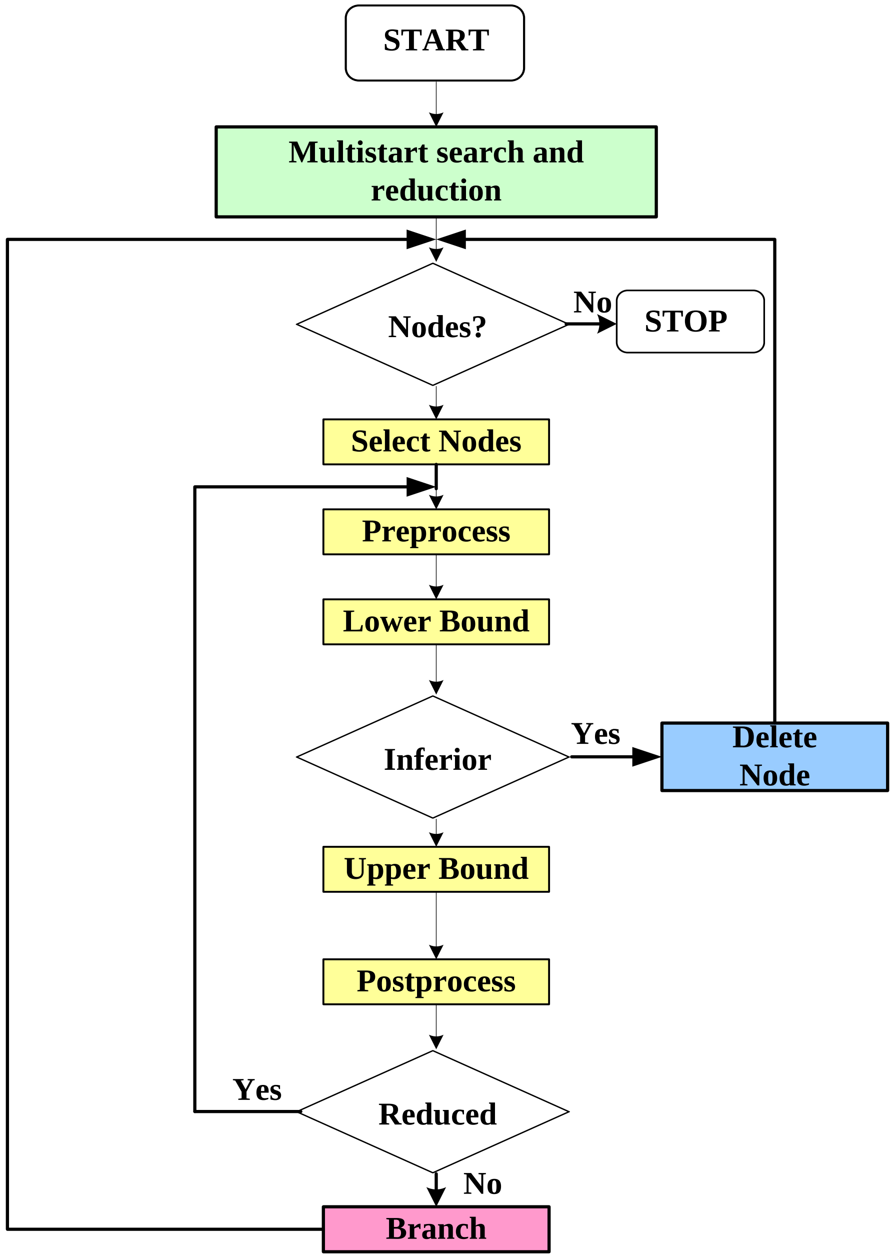


Figure 2.17: BARON algorithm.

d) Cutting planes Method

These methods iteratively refine a feasible set (or objective function) by using linear inequalities known as “cuts”. Optimisation problems are therefore solved through a series of relaxations whose feasible sets are progressively tightened through the addition of valid cuts (valid inequalities). These methods are used to find integer solutions to MILP problems to solve convex optimisation problems. CP methods are used in solving NLP problems by approximating a feasible region of a nonlinear (convex) model by means of a finite set of closed half spaces, which are then solved by a sequence of approximating linear programs. An example of a cutting planes method is known as outer approximation (OA).

The OA method was first proposed by Duran and Grossmann (1986). This method applies convex (or concave) functions and convex sets. The OA method approximates a function by means of a polyhedral, which contains the set. The function is therefore approximated by piecewise-linear functions. The procedure is an iterative method that generates an upper and lower bound on an MINLP solution. The disadvantage of this method is that, a large number of approximations may be required for an adequate approximation to be obtained. OA is implemented within the MINLP optimisation solver known as DICOPT (discrete and continuous optimiser) (Lee & Leyffer, 2012).

2.8 Membrane Technology

Membrane technology has gained a growing level of application in the process industry (Galan & Grossmann, 1998). This is because membrane technology is less energy intensive than the traditional separating processes such as distillation. Membrane systems also have a low capital and utility cost (El-Halwagi, 1997). They are thin film-like structures that separate two fluids and act as selective barriers to retain pollutants in a contaminated stream in order to allow water (solvent) to permeate into a purified stream (Saif et al., 2008b). Membrane systems are therefore impermeable to certain particles when exposed to a specific driving force such as pressure. The feed stream is split into two product streams namely permeate and retentate. The permeate stream has a low contaminant concentration and the retentate has a high contaminant concentration level. A schematic representation of a simple membrane separation process is shown in Figure 2.18.

There are many different types of membranes used in the process industry for the treatment of wastewater and seawater. Membranes are selected based on the types of material that passes through their pore, the type of wastewater that needs to be treated and the driving force for the separation process. The focus of this research will, however, be on RO membranes due to their distinct characteristics. The different types that will be discussed briefly in this review are namely:

(i) Microfiltration membranes
(ii) Ultrafiltration membranes
(iii) Nanofiltration membranes
(iv) Reverse osmosis membranes
(v) Forward osmosis
(vi) Membrane distillation
(vii) Electrodialysis

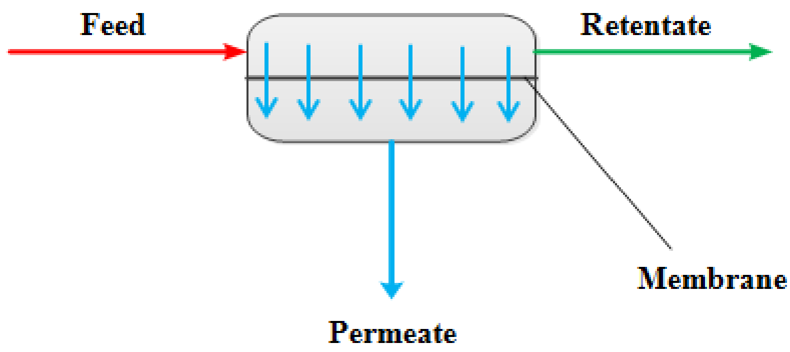


Figure 2.18: A simple schematic representation of a membrane separation.

(i) *Microfiltration (MF)*

MF is a separation process that allows a solution to flow perpendicular to a porous membrane. The pore of MF ranges from 0.1µm to 10µm (Baker , 2012). It is a low-pressure separation process with pressures of 0.2bar to 5bar (Perry & Green , 2007). This therefore means that any particle that exceeds the pore size is retained on the membrane and as such the solution then filters out of the membrane. MF is used to remove pathogens such as sediments, algae and protozoa within the wastewater. They are therefore used in the pharmaceutical industry, clarification of juices/wine/beer, oil/water separation, water treatment, dairy processing etc. (Baker , 2012). MF membranes are often used as pretreatment for UF, RO and NF membranes.

(ii) *Ultrafiltration (UF)*

UF is a membrane separation process that involves the use of a pressure gradient to separate solvents from solutes through a semipermeable membrane. UF is similar to MF with a smaller pore size of 1nm to 100nm. The membranes are characterised by the molecular weight cut-off (MWCO) of the membrane, which refers to the lowest molecular weight solute in which 90% of the solute is retained by the membrane. UF membranes are used to remove particulates, macromolecules, bacteria, colloids, dispersed fluids and suspended solids from the contaminated solution (KOCH, 2013).

(iii) *Nanofiltration (NF)*

NF is a high pressure process which is similar to RO but is however used to remove only divalent and large ions. NF membranes have a low rejection to monovalent ions and are therefore used mainly for de-salting of a process stream. In water treatment NF membranes are used to remove pesticides and for colour reduction (KOCH, 2013). It uses nanometer sized cylindrical through-pores which, penetrate the membrane at an angle of 90°C. NF membranes have a pore size that ranges from 1nm to 10nm. NF is, however, the least used method in industry as the pore size has to be in nanometers and incurs high maintenance costs (Baker & Martin, 2007).

(iv) *Reverse Osmosis (RO)*

RO membraneshave the smallest pore size which ranges from 0.0001μm to 0.001μm. RO membranes separate a water stream into a lean stream of low contaminant concentration known as the permeate and a highly contaminated stream known as the retentate stream. The process is achieved by applying an external pressure to the feed solution in order to reverse the osmotic phenomenon. As a result of this process, retentate streams exit the membrane at a high pressure. RO membranes are used to remove different types of molecules and ions (Saif et al., 2008a). RO membrane systems are often used for seawater and brackish water desalination (Maskan et al., 2000).

(v) *Forward Osmosis (FO)*

FO membranes are similar to RO membranes, but the driving force for the separation is an osmotic pressure gradient. More energy is, however, required for RO than FO. FO is used in desalination and wastewater treatment. FO membranes are often used as pretreatment for RO membranes (Lee, 1981).

(vi) *Membrane distillation (MD)*

MD is a thermally driven separation system and separation is brought about by a phase change. The driving force is due to a partial vapour pressure, which is driven by a temperature difference. The membrane is hydrophobic and displays a barrier for the liquid phase, which in turn allows the vapour phase to pass through the pores of the membrane. This technology is applied in seawater desalination, water treatment and water purification (Winter et al., 2011).

(vii) *Electrodialysis (ED)*

ED is a process, which is based on the electromigration of ions across cation and anion exchange permselective membranes by means of direct electric current (Tsiakis & Papageorgiou, 2005). The ED membrane allows the movement of positive and negative ions through its pores. ED is used for the desalination of high salinity water, wastewater minimisation etc.

The focus of this review will, however, be on reverse osmosis membranes due to their low energy consumption (compared to multistage flash distillation), high quality and product recovery. RO units are also easy to operate and have a modular plant design. They are also attractive as they are able to meet varying feed water concentrations and varying production water qualities (Lu et al., 2012). The RO system is also moderate in energy consumption when compared to thermal separation systems (Marcovecchio et al., 2005) and other separation systems. Cost of maintenance is also significantly lower (compared to thermal separation processes) for RO units (Voros et al., 1997). These advantages therefore make the RO system more attractive than other conventional separation processes (Saif et al., 2012).

2.9 Reverse Osmosis Membranes (RO)

2.9.1 Basic concepts

RO is a pressure driven process, where solute is retained on the pressurised side known as the retentate side and the solvent is allowed to pass through to the less pressurised side known as the permeate. RO membranes are able to retain molecules and ions due to their small pore size, which is less than 0.5 nm (Saif et al., 2012).

Solutions with different solute concentrations create a chemical potential difference when separated by a semi-permeable membrane (Saif et al., 2008a). Chemical potential difference in a mixture is defined as the slope of free energy of the system with respect to a change in the number of moles of just that species. The chemical potential difference allows the carrier solvent to be transported from a low concentration side to a high concentration side. This phenomenon is known as osmotic flow. Osmotic flow causes an increase in pressure on the retentate side. The system will then reach equilibrium when the pressure difference across the membrane balances the chemical potential across the membrane. An external pressure, which is larger than the osmotic pressure is applied to the solution in order to reverse the osmotic phenomenon. The external pressure allows the solvent to pass through the membrane while the solute remains in the retentate stream (El-Halwagi, 1997).

An energy recovery turbine is used to harness pressure energy from the retentate stream as it leaves the RO membrane at a high pressure (El-Halwagi, 1992). The principle behind RO is shown in Figure 2.19, where π is the osmotic pressure and ΔP is the pressure difference across the membrane. The presence of the osmotic pressures of the solutions limits the expansion of the RO membrane, as the value must not exceed the applied pressure (Evangelista , 1989).

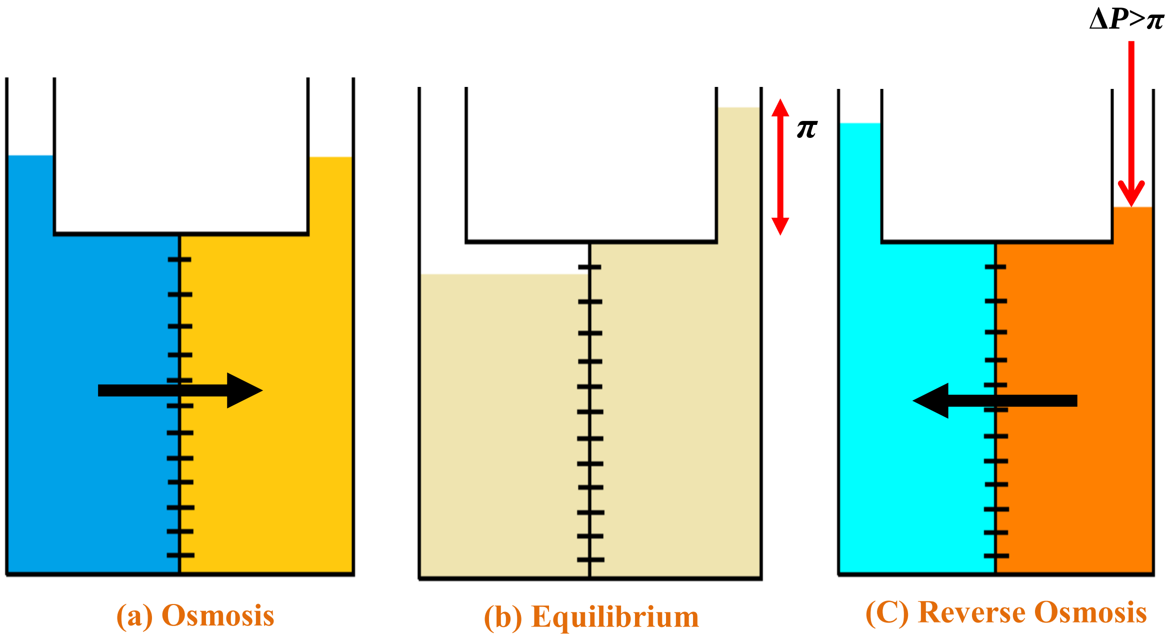


Figure 2.19: Principle of RO (El-Halwagi, 1997).

The performance of RO membranes (membranes in general) is, however, affected by fouling (Sassi & Mujtaba, 2011). Fouling can be in the form of suspended solids such as silica, iron oxides, inorganic compounds, organic compounds and biological compounds. The mass transfer on the high pressure side of the RO membrane (retentate), causes fouling (Evangelista, 1985). Fouling affects membrane performance as it deteriorates membrane permeability. Fouling also results in decreased product quality and increased feed pressure in order to maintain the freshwater demand (Sassi & Mujtaba, 2011). It also increases the energy consumption and because chemicals are needed to remove the foulants, this results in an increase in the total treatment cost. This therefore means that the membranes have to undergo regular maintenance (Zhu et al., 1997). The performance of the RO membranes is, however, recovered by being chemically or mechanically regenerated (Zhu et al., 1997).

The performance of RO units is also affected by concentration polarisation. Concentration polarisation is the accumulation of solute on the membrane surface. This therefore means that the solute concentration at the membrane wall becomes greater than that of the bulk feed solution (Kaushik, 2008). This affects the solvent and solute recovery as they are dependent on the wall concentration, which in turn is a function of the solvent and solute fluxes (Evangelista, 1985).

In order to minimise capital cost, the membrane module must provide a large area per unit volume. This creates a more efficient separation system. RO units consist of four module configurations: hollow fibre, plate and frame, spiral and tubular wound (Evangelista, 1985). The choice of a module configuration therefore depends on ease and cost of module manufacture, energy efficiency, fouling tendency, required recovery and the capital cost of auxiliary equipment (Maskan et al., 2000). Hollow-fibre reverse osmosis and spiral wound modules are commonly used in industrial processes as they offer a large surface area to volume ratio, self-supporting strength of fibres and negligible concentration polarisation (El-Halwagi, 1997).

2.9.2 Hollow-fibre reverse osmosis modules (HFRO)

HFRO modules consist of a large number of membrane tubes, which are placed in a module shell. The fibres have a small diameter of approximately $1x10^{-5}$m (Marriott, 2001). This therefore increases the packing density of the configuration (larger than spiral-wound configuration). HFRO modules are commonly used in industrial processes as they offer a large surface area to volume ratio, self-supporting strength of fibres and negligible concentration polarisation (El-Halwagi, 1997). The concentration polarisation is negligible because the permeability is about ten times less than that of flat sheet membranes (Evangelista, 1985).

HFRO modules will therefore be used in the design of the RO membrane. The feed stream is introduced outside the hollow-fibres and the material permeates into the interior to form the permeate stream (Marriott, 2001).

Figure 2.20 illustrates the main features of a hollow-fibre module configuration. The fibres are grouped together in a bundle with one exposed to the atmosphere while the other end is sealed. The open ends are potted into an epoxy sealing head after which the permeate is collected. The feed solution flows around the outer side of the fibres towards the perimeter of the shell, the permeate solution penetrates through the fibre wall into the bore by means of reverse osmosis (El-Halwagi, 1997). The permeate stream is then collected at the open ends of the fibres while the retentate stream is collected at the porous wall of the shell (El-Halwagi, 1997).

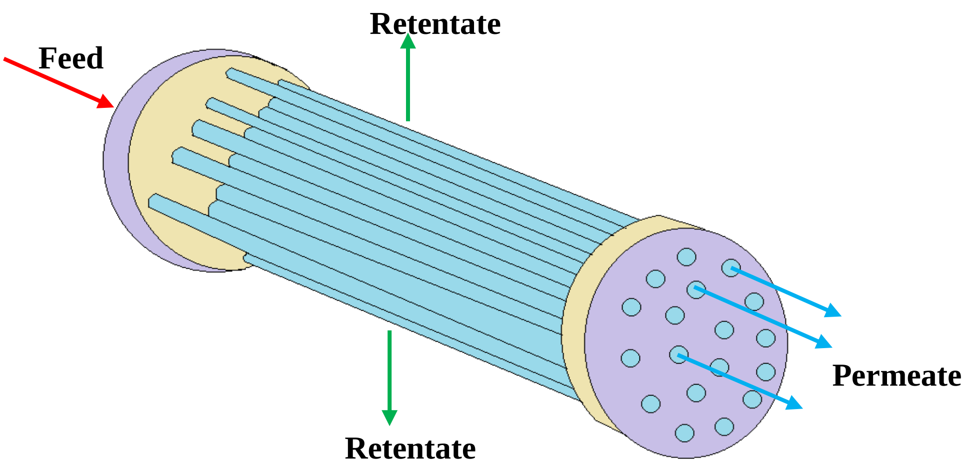


Figure 2.20: HFRO membrane.

2.9.3 Modelling of HF reverse osmosis units

Two aspects are considered when modelling the RO module such as the membrane transport equations and the hydrodynamic modelling of the RO module. The membrane transport equations describe the process taking place at the membrane surface. The hydrodynamic model describes the macroscopic transport of the different species along with the momentum and energy associated with them (El-Halwagi, 1997).

In the hydrodynamic models, two approaches have been adopted to describe the pressure variation through the shell side. The first approach assumes a constant pressure on the shell side while the second approach treats the fibre bundle as a porous medium where flow is described by Darcy's equation with an arbitrary constant. The pressure drop inside the model along the membrane creates a pressure difference across the fibre and as a result, the permeation rate may change considerably along the fibre length. This therefore means that an axial component of the shell side pressure arises in addition to the radial component. The model must therefore capture both radial and axial flows within the HF reverse osmosis module.

Two transport equations are used to predict the flux of water and solute. The water flux through the membrane is described in constraint (2.18).

$$N_{water} = A\left(\Delta P - \frac{\pi_f}{C_f} C_s\right)\gamma \qquad (2.18)$$

Where N_{water} is the water flux, ΔP is the pressure difference across the membrane, π_f is the osmotic N_{water} feed pressure, C_f is the solute concentration in the feed, C_s is the average solute concentration in the shell side and γ is a dimensionless constant described by constraint (2.19).

$$\gamma = \frac{\eta}{1 + \dfrac{16 A \mu r_o L L_s \eta}{1.0133x10^5 r_i^4}} \qquad (2.19)$$

In constraint (2.19) is also a dimensionless constant and is described in constraint (2.20). The description of θ is shown in constraint (2.21).

$$\eta = \frac{tanh\,\theta}{\theta} \tag{2.20}$$

$$\theta = \left(\frac{16A\mu r_o}{1.0133x10^5 r_i^2} \right)^{\frac{1}{2}} \frac{L}{r_i} \tag{2.21}$$

The solute flux is shown in constraint (2.22).

$$N_{solute} = \left(\frac{D_{2M}}{K\delta} \right) C_s \tag{2.22}$$

where $\left(\frac{D_{2M}}{K\delta} \right)$ is the salt flux constant.

2.9.4 Synthesis of RO Membrane Networks

Research towards the optimum design and synthesis of RO networks has increased considerably (Marcovecchio et al., 2005). Designing a cost effective RO unit depends on the determination of the optimal operational and structural schemes (Voros et al., 1997). The optimum design includes the generation of the optimum number of RO units, booster pumps, energy recovery turbines, optimum stream distributions, operating conditions and separation levels of the streams (El-Halwagi, 1992). This therefore allows a detailed synthesis and design of the RON system. The RO system is typically installed in order to meet the environmental, technical and economic requirements needed for the separation process (Maskan et al., 2000). The optimisation of a RON has been studied extensively (Khor et al., 2011). A WN has, however, not been included in most studies of the RON synthesis.

El-Halwagi (1992) was the first to introduce the idea of using a sequence of reverse osmosis networks for wastewater minimisation. The paper led to the development of a superstructure system that considered all the possible processing unit configurations (membranes, pumps and energy recovery turbines) and full stream connectivity. El-Halwagi (1992)investigated the optimum synthesis of RO networks by using the state space approach. In the state space approach the RO networks are split into four distribution boxes: a pressurisation/depressurisation stream-distribution box (PDSDB), pressurisation/depressurisation matching box (PDMB), a RO stream-distribution box (ROSDB) and a RO matching box (ROMB). The purpose of the distribution boxes was to allow all possible combinations of stream mixing, splitting, recycle and bypass. This therefore allows all possible network configurations. The mathematical model was therefore formulated as an MINLP. The objective of the optimisation problem was to minimise the total annualised cost (TAC), which consisted of the annual installation cost of the RO module (including annualised installed cost, membrane replacement, labour and maintenance), fixed cost of the pumps, turbine installation and cost of electrical power.

The model was applied to a seawater desalination problem and a pulp-bleaching plant. The solution to the optimisation problem provided the optimum arrangement, size and type of the RO units, energy recovery turbines and booster pumps, optimum stream-distributions and operating conditions. Figure 2.21 shows the state space representation of the RO network proposed by El-Halwagi (1992). The methodology proposed by El-Halwagi (1992) does not, however, guarantee global optimality and did not include a WN (Khor et al., 2011). El-Halwagi (1993) then combined the RON with other separation processes. The maximum allowable inlet concentration into the regenerator was also not specified or included in the model formulation.

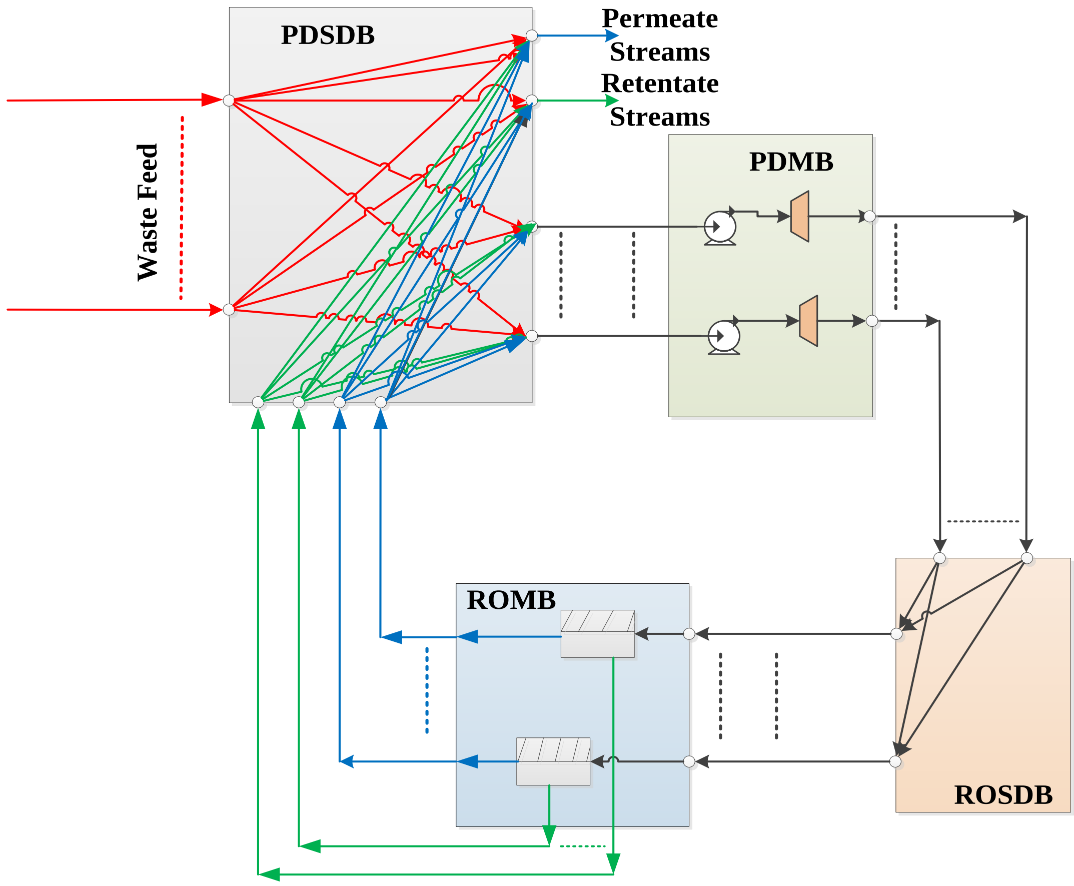


Figure 2.21: State space representation of the RO network (El-Halwagi, 1992).

Zhu et al. (1997) then presented a technique for the optimal design and scheduling of a flexible RO network. In the model formulation the decline in permeate flux was described as an exponential decay due to fouling. This therefore meant that the model was influenced by variable feed conditions and system performance. The mathematical model was formulated as an MINLP problem whose objective was to minimise the TAC, which included thermodynamic, technical, fouling and flexibility constraints. Different schedules for the membrane regeneration were determined. The overall minimum TAC generated from the schedules was then chosen as the best configuration. The effect of fouling on the membrane performance was taken into account by Zhu et al. (1997) but was not taken into consideration in the model proposed by El-Halwagi (1992). This is important as it affects the TAC. The model by Zhu et al. (1997) did not, however, guarantee global optimality.

Maskan et al. (2000) optimised a two-staged RO network. The problem was formulated as a constrained multivariable nonlinear optimisation problem. The objective of the problem was to maximise the annual profit that was obtained from the permeate, capital cost from process units, the operation costs associated with the maintenance and energy consumption. A nonlinear correlation between the concentration and the osmotic pressure was used to estimate the osmotic pressure better instead using of the common linear relationship. A linear correlation was used to calculate the osmotic pressure in the model presented by El-Halwagi (1992) and Zhu et al. (1997). The model also accounted for the pressure losses (due to the friction and flow in the module) and concentration polarisation in order to calculate the osmotic pressure accurately. The objectives of the study were to determine the optimum dimensions for the RO module and the optimum layout of the treatment network. The objective function also included the sale of permeate obtained by external and internal customers. The model also used decision variables to distinguish between module types like tubular, spiral wound and hollow fibre modules. This was also not considered in previous papers. The model was applied to the desalination of brackish water and seawater and the analysis showed that the optimum network designs were the ones that produced the maximum permeate.

Saif et al. (2008a) proposed a superstructure which was a modification of the superstructure presented by El-Halwagi (1992). In their superstructure, several alternatives among the utility units (e.g. pumps and turbines) that have common RO design concepts where reduced. The superstructure also provides other network alternatives, which enabled meaningful connectivity between the RO units and also simplified the modelling of the RO network. The superstructure consisted of only the PDSDB and unit operation boxes (pumps, turbines, RO membrane) that contained different units, which treat the various feed streams. The mathematical programming model was formulated as a nonconvex MINLP for water desalination and wastewater treatment from a pulp and paper industry. A mixed integer linear problem (MILP) was then derived from the original nonlinear problem by means of convex relaxation of the nonconvex MINLP model. The MILP was solved iteratively in order to supply different initial guesses for the nonconvex MINLP model. The method was effective in finding several local optimum solutions as the convergence difficulty experienced by most MINLP local search methods was overcome. Global optimality was, however, not guaranteed. The model was able to minimise treatment costs and led to minimisation of wastewater.

Saif et al. (2008b) then extended the superstructure by Saif et al. (2008a) by applying an efficient branch-and-bound algorithm in order to obtain global optimality for the RON. This was achieved by solving the MILP model at every node in the branch-and-bound tree in order to verify the global optimality of the RO network. Additional constraints where derived in order to tighten the mathematical programming structure and the RO design. This helped tighten the bounds of the variables and to accelerate the convergence of the algorithm. The tightening constraints were provided in order to prevent a stream discharge from an RO stage being recycled back to the same stage due to the pressure drop at every stage; to prevent mixing of streams from reject to turbine and streams from turbine to the RO and to prevent the mixing of low and high-pressure streams mixing at the same pump nodes. The optimal treatment cost obtained by Saif et al. (2008b) was 14.8% lower than that obtained by El-Halwagi (1992). The mathematical model of Saif et al. (2008b) also guarantees global optimality. Most of the earlier authors of RO synthesis proposed models that do not guarantee global optimality (Saif et al., 2008b).

Sassi and Mujtaba (2011) optimised an RO network using the MINLP approach and also incorporated fouling effects. The effect of fouling was described by an exponential function, which also represented the decline in water permeability using a spiral wound membrane element. The objective of the problem was to minimise the TAC in order to find the optimal design and configuration of the RO system. The model was solved using outer approximation algorithm within the gPROMS software. The results showed that, the optimal solution was sensitive to the fouling distribution between stages. The overall fouling however remained constant. The fouling effects in their formulation was therefore not assumed be equal as in previous optimisation of RO networks.

Lu et al. (2012) presented a systematic methodology for the optimal design of RO desalination systems with multiple feed streams (seawater, brackish water and regenerated water) and multiple product streams of different quality. The problem was formulated as an MINLP whose objective was to minimise the TAC in order to determine the optimal system structure, operation conditions and stream distributions when subjected to constraints of the multiple feed and multiple product system. The superstructure was a modified version of the state-space approach presented by El-Halwagi(1992). An example of the modification was that a turbine is used only for the final retentate stream produced. Stream split ratios and a logical expression of stream mixing was included in the mathematical formulation. This made the mathematical model easier to handle as it reduced the number of binary variables and solving space. The model was also formulated to select different types of spiral-wound membrane modules for each stage. The minimum desirable product flowrate with its corresponding maximum concentration was given as a parameter in the case study. The model was able to produce multiple permeate streams of different quality and the optimal design of the RO network was obtained.

Saif et al. (2012) considered the minimisation of wastewater and freshwater in the pulp and paper industry by using RO membranes. The RON was synthesised in order to regenerate streams with reduced salt concentration at a minimum cost. The model by Saif et al. (2008a) was used for the RON. Their work optimised the RO but did not optimise the WN. This therefore meant that the RON was optimised separately then incorporated within the ready existing water network. The minimisation freshwater intake and wastewater generation was therefore not included in the objective function. The maximum allowable inlet concentration into the regenerator was also not specified or included in the model formulation.

2.10 Optimisation of the WN

The idea of a WN was first proposed by Takama et al. (1980). Different papers were written after that based on the idea of using mathematical programming to optimise a WN. The difference in ideas, however, arises with the modelling of the regenerators and the methods used to solve the MINLP model. Some papers have looked at a detailed model of the regenerators while others have represented the regenerator with a "black-box'' (without any detail). A detailed design of the regenerator can help reduce the amount of energy used for the treatment of wastewater. There have, however, been few works that consider a detailed nonlinear regeneration model for the synthesis of water networks (Khor et al., 2014). The cost of designing the regenerator can also be optimised, as it will be included in the objective function.

The use of regenerators in a WN leads to a reduction in both freshwater usage and wastewater generation. A few works have, however, looked at a detailed model of the regenerators but these are often limited to only a single treatment technology of fixed type (Khor et al., 2011). It is also ideal for the number and type of treatment units not to be fixed, but rather chosen among others through the optimisation process.

2.10.1 WN optimisation with a "black-box" Regenerator

Models for regenerators have, however, been described in most works by means of a fixed outlet concentration and a fixed removal ratio for contaminants (Jeżowski, 2010). This is known as the "black box" approach (Alva-Argáez et al., 1998; Khor et al., 2012; Tan et al., 2009), which uses a simplified linear model with constant removal ratios (RR) to represent the membrane systems (Tan, et al., 2009). The RR is defined as the fraction of mass load in the regeneration unit from the feed stream that exits in the retentate stream (Khor et al., 2011). In networks where multiple regenerators are considered, the RR and allowable contaminant concentration are varied. This approach allowed the simplification of complex networks that consisted of multiple water sources, sinks and regenerators (Khor et al., 2012).

The fixed liquid phase recovery (α) factor is also used to represent the performance a regenerator unit. α is the fixed fraction of the inlet stream into a regenerator that exits in the permeate stream (Khor et al., 2011). This is achieved by expressing the objective function in terms of the total inlet flow of streams into the regenerator (Galan & Grossmann, 1998). In some works, the cost of regeneration is neglected and only the cost of freshwater, wastewater and the capital cost of the network are taken into account. This therefore means that the actual cost of the regenerator is not considered in the model formulation. This approach does not give an accurate representation of the energy consumption and associated costs of the membrane systems.

Tan et al. (2009) developed a WN with partitioning based regenerators for total WN synthesis. The partitioning regenerators split the wastewater into regenerated lean streams (permeate) and low quality reject streams (retentate). Membrane separation processes such, as reverse osmosis and ultrafiltration are examples of partitioning regenerators. A fixed RR and α were used to represent the function of the regenerators. A centralised single portioning regenerator was used with a source-sink superstructure. The formulation resulted in a nonconvex NLP, which was solved using a branch and bound method via Lingo. The water streams that are linked to the regenerator were already identified. The problem was, however, restricted to single contaminated streams. The model did lead to the minimisation of fresh water consumption. The minimisation of energy was, however, not considered in the model formulation, as the objective was to minimise freshwater. The model was not formulated to select multiple regenerators as a single portioning regenerator was used. A detailed representation of the regenerator was also not considered in the model formulation.

Chew et al. (2008) focused on interplant water integration (IPWI) by looking at the geographical location of the water-using processes. They combined different water networks together instead of the usual single water network that has been used by most researchers. In their work they looked at two IPWI schemes, direct and indirect integration, which are then solved by mathematical optimisation techniques. In the direct integration, water from the different networks is integrated directly via pipes. The indirect integration utilises a centralised utility hub to integrate the water from the different networks together. The centralised hub was then used to distribute water to the different plants and was also modelled in a different case study as a water regeneration unit. The removal ratio of the contaminants from the regenerator was used to define the lean and concentrate streams. The scenario with the regenerator led to lower freshwater and wastewater flowrates in the overall water networks. A detailed representation of the regenerator was, however, not considered in the mathematical modelling.

Khor et al. (2012) proposed a WN that consisted of partitioning regenerators (RO and ultrafiltration) and non-membrane regenerators. Their superstructure consisted of sources, sinks and regenerators. The sinks, however, consisted of an end-of-pipe effluent treatment system (ETS). The water sources also consisted of multiple freshwater streams. A linear model with fixed removal ratios and liquid-phase recovery factors was developed for the membrane regenerators. The permeator and rejector streams from the regenerators were treated as tasks instead of states. This meant that,the permeator and rejector streams were treated as units which can accept water streams from the sources. Khor et al. (2012) also incorporated linear logical constraints by using 0-1 variables in order to tighten the model formulation and to enhance solution convergence. The model was formulated as an MINLP. A global optimum WN was developed with 27% savings in freshwater use. The mathematical model, however, did not include a detailed representation of the regenerators. This therefore means the solution is not a true reflection as a “black-box’’ was used to describe the regenerators. The consumption of energy by the multiple regenerators was also not taken into account in the objective function.

Similar approaches of using the “black-box” method have been used in most published work with regards to WN synthesis. This includes the work of Galan and Grossmann(1998), Karuppiah and Grossmann (2006), Koppol et al. (2004), Meyer and Floudas (2006) and Faria and Bagajewicz (2011)

The ‘black-box’ representation does not present a good description of the regenerator as removal ratios and liquid recoveries are used to represent the treatment by the regenerators. The minimisation of energy used by the regenerators in all the papers discussed is also not considered in the model formulation. The optimum solution is therefore not a good representation of the cost and a detailed design of the regenerator is therefore not obtained at the optimum solution. The specific regeneration technology is also not considered in the model.

2.10.2 WN optimisation with a detailed Regenerator

A detailed model of the regenerator makes the optimum cost of the WN more realistic as its design is also included in the optimisation model. The type of regenerator for the treatment of the wastewater can also be specified in the model instead of using a “black-box’’ representation. Faria and Bagajewicz (2009) showed the importance of modelling the regenerators in the WN. In their discussion, they showed that every WN needs a detailed model of the regenerator. The use of a fixed RR to represent treatment units also limits their application in industrial processes (Yang et al., 2014). This therefore means that a more rigorous representation of the regeneration unit is needed (Khor et al., 2014).

Khor et al. (2011) presented a detailed model of a regenerator, which was incorporated into a water network superstructure. The superstructure they proposed consisted of continuous variables for the contaminants and flowrates and binary variables for the piping interconnections. The superstructure consisted of the nonlinear detailed RON model, water sources and water sinks. An MINLP mathematical model was therefore proposed. The model enabled direct water reuse/recycle, regeneration reuse or regeneration recycling. The model was also formulated to incorporate multiple contaminants. They proposed equations to represent the RO membrane, but a superstructure of the RON model was not incorporated into the model formulation.

The work of Khor et al. (2011) assumed a single regenerator with a fixed design, which implies that the number of regenerators needed, number of pumps, number of energy recovery turbines were specified a priori. This limited the flexibility of the model, which could result in a suboptimal solution. As such, the model was not programmed to select if a series connection or a parallel connection between the regenerators was the optimal choice. The model, however, did lead to a 58% saving in freshwater use and a reduction in the capital cost of the regeneration unit with a payback period of 2.1 years when applied to a petroleum refinery. The maximum allowable inlet concentration into the regenerator was also not specified or included in the model formulation.

Yang et al. (2014) addressed the problem of using a RR to represent the performance of treatment units. The objective of the work was to consider the trade-off between the removal efficiency of a unit and treatment cost and their impact on the WN. The work combines various technologies in order to remove total dissolved solids, total suspended solids and organics. Unit specific short-cut models were used to describe each treatment system instead of a fixed RR. Uncertain parameters are used to account for the change of condition for a particular process during the course of the operation. The model also looks at the best available technology to remove a specific contaminant. This was achieved by using disjunction in the GDP. This is, however, computationally expensive to solve to global optimality. The treatment units that were considered are RO, UF, ion exchange, sedimentation, activated sludge and trickling bed. A spiral-bound RO module was used instead of the hollow fibre RO module. The model was applied to a metal finishing and petroleum refining industrial case study. The work of Yang et al. (2014) used only one RO unit and also did not consider the detail synthesis of the RO unit.

2.11 References

Ahmetović, E. & Grossmann , I., 2010. General superstructure and global optimisation for the design of integrated process water networks.

Alnouri, S. & Linke, P., 2012. A systematic approach to optimal membrane network synthesis for seawater desalination. Journal of membrane science, 417, pp.96-112.

Alva-Argáez, A., Kokossis , A. & Smith , R., 1998. Wastewater minimisation of industrial systems using an integrated approach. Computers and Chemical Engineering, 22, pp.741-44.

Bagajewicz, M. & Savelski, M., 2001. On the use of linear models for the design of water utilization systems in process plants with a single contaminant. Trans. Institute of Chemical Engineers, 79, pp.600-10.

Baker , R., 2012. Microfiltration in Membrane Technology and Applications. California : John Wiley amd Sons Ltd.

Baker , L. & Martin, 2007. Nanotechnology in Biology and Medicine: methods, devices and applications. Nanomedicine; Nanotechnology, Biology and Medicine , 9, pp.1-24.

Chew, I. et al., 2008. Synthesis of direct and indirect interplant water network. Industrial and Engineering Chemistry Research, 47, pp.9485-96.

D'Ambrosio, C., Lodi , A., Wiese, S. & Bragalli, C., 2015. Mathematical programming techniques inwater network optimisation. European Journal of Operational Research , 243, pp.774-88.

Dhole , V., Ramchandi, N., Tainsh , R. & Wasilewski, M., 1996. Make your process water pay for itself. Chemical Engineering , 103(1), pp.100-03.

Doyle , S. & Smith, R., 1997. Targeting water resue with multiple contaminants. Institution of Chemical Engineers, 75, pp.181-89.

Duran, M. & Grossmann, I., 1986. An outer approcimation algorithm for a class of mixed integer nonlinear programs. Mathematical Programming, 36, pp.307-39.

Edgar , T. & Himmelblau , D., 1988. Optimisation of Chemical Processes. New York : McGraw-Hill.

El-Halwagi, M., 1992. Synthesis of reverse osmosis networks for waste reduction. American Institute of Chemical Engineers, pp.1185-98.

El-Halwagi, M., 1993. Optimal design of membrane-hybrid systems for waste reduction. Separation Science and Technology, 28, pp.283-307.

El-Halwagi, M., 1997. Pollution Prevention through Process Integration. California: Academic Press.

El-Halwagi, M. & Manousiouthakis, V., 1989. Synthesis of mass-exchange networks. AIChE J, 3, pp.1233-44.

Evangelista , F., 1989. Design and preformance of two-stage reverse osmosis plants. Chemical Engineering and Processing , 25, pp.119-25.

Evangelista, F., 1985. A short-cut method for the design of reverse-osmosis desalination plants. Industrial & Engineering Chemistry Process Design and Development, 24, p.211.

Faria , D. & Bagajewicz , M., 2011. Global Optimisation of water management problems using linear relaxations and bound contraction methods. Industrial and engineering chemistry research , 50(7), pp.3738-53.

Faria, D. & Bagajewicz , M., 2009. On the appropriate modelling of process plant water systems. American Institute of Chemical Enginering Journal , 56, pp.668-89.

Galan , B. & Grossmann, I., 1998. Optimal design of distributed wastewater treatment networks. Industrial and Engineering Chemistry Research, 37(10), pp.4036-48.

GAMS, 2013. GAMS: The Sover Manuals. Washington DC: GAMS Development Corporation.

Glover , F., 1975. Improved linear integer programming formulations of nonlinear inteher pronlems. Management Science , 22(4), pp.455-60.

Gounaris , C., Misener , R. & Floudas , C., 2009. Computational comparison of piecewise-linear relaxations for poooling problems. Industrial and Engineering Chemistry Research, 48(12), pp.5742-66.

Grossmann, I. & Biegler , L., 2004. Future perspective on optimisation. Computers and Chemical Engineering , 28, pp.1193-218.

Gunaratnam, M. et al., 2005. Automated design of total water systems. Industrial and Engineering Chemistry Research, 44, pp.588-99.

Hallale , N., 2002. A new graphical targeting method for water minimisation. Advances in Environmental Research, 52, pp.377-90.

Jeżowski, J., 2010. Review of water network design methods with literature annotations. Industrial and Engineering Chemistry Research, 49, pp.4475-516.

Karuppiah, R. & Grossmann, I., 2006. Global optimisation for the synthesis of integrated water systems in chemical processes. Computers and Chemical Engineering, 30(4), pp.650-73.

Kaushik, N., 2008. Membrane separation processes. New Delhi: Prentice-Hall.

Khor , C., Chachuat , B. & Shah, N., 2014. Optimisation of water network synthesis for single-site and continous processes: milstones, challenges, and future directions. Industrial & Engineering Chemistry Research , 53, pp.10257-75.

Khor, C., Chachuat , B. & Shah, N., 2012. A superstructure optimisation approach for water network synthesis with membrane separation-based regenerators. Computers and Chemical Engineering, 42, pp.48-63.

Khor, CS, Foo, DCY., El-Halwagi, MM., Tan, RR., Shah, N., 2011. A superstructure optimisation approach for membrane separation-based water regeneration networks synthesis with detailed nonlinear mechanistic reverse osmosis model. Industrial and Engineering Chemistry Research , 50, pp.13444-13456.

KOCH Membrabe Systems , 2013. KOCH Membrane Systems. [Online] Available at: HYPERLINK "http://www.kochmembrane.com/PDFs/Other-Documents/Membrane-Filtration-Technology-Koch-Membrane-Sys.aspx "[Accessed Monday July 2015].

Koppol, A., Bagajewicz , M., Dericks, B. & Savelski, M., 2004. On zero water discharge solutions in the process industry. Advances in Environmental Research, 8(2), pp.151-71.

Kuo, W. & Smith, R., 1997. Effluent treatment system design. Chemical Engineering Science , 52, pp.4273-90.

Kuo, W. & Smith, R., 1998a. Designing for the interactions between water-use and effluent treatment. Trans IChemE, 76((Part A)), pp.87-301.

Kuo, W.C.J. & Smith, R., 1998b. Design of water-using systems involving regeneration. Transactions of the Indtitute of Chemical Engineers, 76(Part B), pp.94-114.

Land, A. & Doig, A., 1960. An automatic method of solving discrete programming problems. Econometrica , 28(3), pp.497-520.

Lee, K., 1981. Membranes for power-generation by pressure-retarded osmosis. Journal of Membrane Science, 8(2), pp.141-71.

Lee, J. & Leyffer, S., 2012. Mixed integer nonlinear programming. New York : Springer.

Li, B. & Chang, C., 2007. A simple and efficient initialisation strategy for the optimising water-using network designs. Industrial and Engineering Chemistry Research, 26, pp.8781-86.

Lu, Y., Liao, A. & Hu, Y., 2012. The design of reverse osmosis systems with multiple-feed and multiple-product. Desalination, 307, pp. 42-50.

Lundell, A. & Westerlund, T., 2012. Finding an optimised set of transformations for convexifying nonconvex MINLP problems. In Karimi, I. & Srinivasa, R., eds. 11th International Symposium on Process Systems Engineering. Singapore. Elsevier B.V.

Manan , Z., Tan , Y. & Foo, D., 2004. Targeting the minimum water flow rate using water cascade analysis technique. AIChE Journal , 50(12), pp.3169-3183.

Marcovecchio, M., Aguirre, P. & Scenna, J., 2005. Global optimal design of reverse osmosis networks for seawater desalination: modeling and algorithm. Desalination, 184(1), pp.259-64.

Marriott, J., 2001. Detailed modelling and optimal design of membrane separation systems. PhD thesis. London: University of London University of London.

Maskan, F., Wiley, D., Johnston, L. & Clements , D., 2000. Optimal design of reverse osmosis module networks. American Institute of Chemical Engineers Journal, 46, pp.946-54.

McCormick, G., 1976. Computability of global solutions to factorable nonconvex ptogramms-part I- convex underestimating problems. Mathematical Programming, 10, pp.146-75.

Meyer , C. & Floudas , C., 2006. Global optimisation of a combinatorially complex generalised pooling problem. American Institute of Chemical Engineers , 52(3), pp.1027-37.

Misener , R. & Floudas , C., 2013. GloMIQC: Global mixed integer quadratic optimiser. Journal of Global Optimisation, 57, pp.3-50.

Misener, R. & Floudas, C., 2010. Global Optimisation of large-scale generalised pooling problems: Quadratically Constrained MINLP Models. Industrial and Engineering Chemistry Research, 49(11), pp.5424-38.

Ng, D., Foo , D. & Tan , R., 2007. Targeting for total water networks.1. waste stream identification. Industrial and Engineering Chemistry Research , 46, pp.9107-13.

Perry, R. & Green , D., 2007. Perry's Chemical Engineers' Handbook. New York: McGraw-Hill Professional.

Poplewski, G., Walczyk, K. & Jeżowski, J., 2010. Optimisation-based method for calculating water networks with user specified characteristics. Chemical Engineering Research and Design , 88, pp.109-20.

Pörn, R., Björk, K. & Westerlund, T., 2008. Global solution of optimisation problems with signomial parts. Discrete Optimisation , 5, pp.108-20.

Porn, R., Harjunkoski, I. & Westerlund, T., 1999. ConvexiÞcation of di¤erent classes of non-convex MINLP problems. Computers and Chemical Engineering, 23, pp.439-48.

Quesada , I. & Grossmann, I., 1995. Global optimisation of bilinear process networks with multi component flows. Computers and Engineering , 19(21), pp.1219-42.

Ryoo, H. & Sahinidis, N., 1996. A branch and reduce approach to global optimisation. Journal of global optimisation, 8, pp.107-38.

Sahinidis , N., 2000. Branch and reduce optimisation navigator. 4th ed. Illinois : University of Illinois at Urbana-Champaign.

Saif , Y., Elkamel , A. & Pritzker , M., 2008b. Global optimisation of reverse osmosis network for wastewater treatment and minimisation. Industrial and Engineering Chemistry Research , 47(1), pp.3060-70.

Saif, Y., Almansoori, A. & Elkamel , A., 2012. Wastewater minimisation in pulp and paper industries through energy-efficient membrane processes. Istanbul, 2012. Industrial Engineering and Operations Management.

Saif, Y., Elkamel , A. & Pritzker, M., 2008a. Optimal design of reverse osmosis networks for wastewater treatment. Chemical Enineering and Processing , pp.2163-74.

Sassi, K. & Mujtaba, I., 2011. Optimisation of design and operation of reverse osmosis based desalination process using MINLP approach incorporating fouling effects. 21st European Symposium on Computer-Aided Process Engineering, 21, pp.206-10.

Savelski, M. & Bagajewicz , M., 2000. On the optimality conditions of water utilization systems in process plants with single contaminants. Chemical Engineering Science , 55(21), pp.5035-48.

Savelski, M. & Bagajewicz, M., 2003. On the necessary conditions of optimality of water utilization systems in process plants with multiple contaminants. Chemical Engineering Science, 58(23-24), pp.5349-62.

Sherali, H. & Adams, W., 1999. A reformulation-linearisation technique for solving discrete and continous nonconvex problems. AH Dirdrecht : Kluwer Academic publishers.

Sherali, H. & Alameddine , A., 1992. A new reformulation linearisation technique for bilinear programming problems. Journal of Global Optimisation , 2, pp.379-410.

Sherali, H. & Liberti, L., 2007. Reformulation-Linearisation Methods for Global Optimisation. Virginia: Department of Industrial and Systems Engineering.

Snyman , J., 2005. Practical Mathematical Optimisation: An Introduction to Basic Optimisation Theory and Classical and New Gradient-Based Algorithms. New York : Springer Science & Business Media.

Savelski, M. & Bagajewicz, M., 2003. Savelski, MJ and Bagajewicz, MJ (2003) "On the necessary conditions of optimality of water utilization systems in process plants with multiple contaminants", Chemical Engineering Science, 58(23-24), 5349-5362. Chemical Engineering Science, 58(23-24), pp.5349-62.

Takama, N., Kuriyama , T., Shiroko, K. & Umeda , T., 1980. Optimal water allocation in a petroleum refinery. Computers and Chemical Engineering, 4, pp.251-58.

Tan, R., Ng, D., Foo, D. & Aviso, K., 2009. A superstructure model for the sybthesis of single-contaminant water network with partitioning regenerators. Process Safety and Environmental Protection, 87(3), pp.197-205.

Tawarmalani , M. & Sahinidis , N., 2005. A polyhedral branch-and-cut approach to global optimizatiom. Mathematical Programming , 103(2), pp.225-49.

Teles, J., Castro, P. & Novais, A., 2008. LP-based strategies for the optimal design of industrial water networks with multiple contaminants. Chemical Engineering Science, 63, pp.376-94.

Treybal , R., 1981. Mass-transfer operations. 3rd ed. Singapore : McGraw-hill.

Tsiakis, P. & Papageorgiou, L., 2005. Optimal design of an electrodialysis brackish water desalination plant. Desalination , 173(2), pp.173-86.

Voros , N., Maroulis , Z. & Marinos-Kouris , D., 1997. Short-cut structral design of reverse osmosis desalination plants. Journal of Membrane Science , 127, pp.47-68.

Wang , Y. & Smith, R., 1994a. Wastewater minimisation. Chemical Engineering Science, 49(7), pp.981-1006.

Wang, Y. & Smith, R., 1994b. Design of distributed effluent treatment systems. Chemical Engineering Science , 49(18), pp.3127-45.

Winter, D., Koschikowski, J. & Wieghaus, M., 2011. Desalination using membrane distillation: Experimental studies on full scale spiral wound modules. Journal of Membran Scince, 375(1-2), pp.104-12.

Yang , L., Salcedo-Diaz , R. & Grossmann, I., 2014. Water network optimisation with wastewater regeneration models. Industrial and Engineering Chemistry Research , 53, pp.17680-95.

Zamora, J. & Grossmann, I., 1998. Continuous optimisation of structured process systems models. Computers and Chemical Engineering, 22(12), pp.1749-70.

Zhu, M., El-Halwagi, M. & Al-Alhmad, M., 1997. Optimal design and scheduling of flexible reverse osmosis networks. Journal of Membrane Science, 129, pp.161- 174.

CHAPTER

3

MATHEMATICAL MODEL

3.1 Introduction

This chapter gives the development of the mathematical model for the incorporation of a RON superstructure within a WN superstructure for the simultaneous minimisation of water and energy. The overall MINLP model is based on the superstructure represented in Figure 3.1.

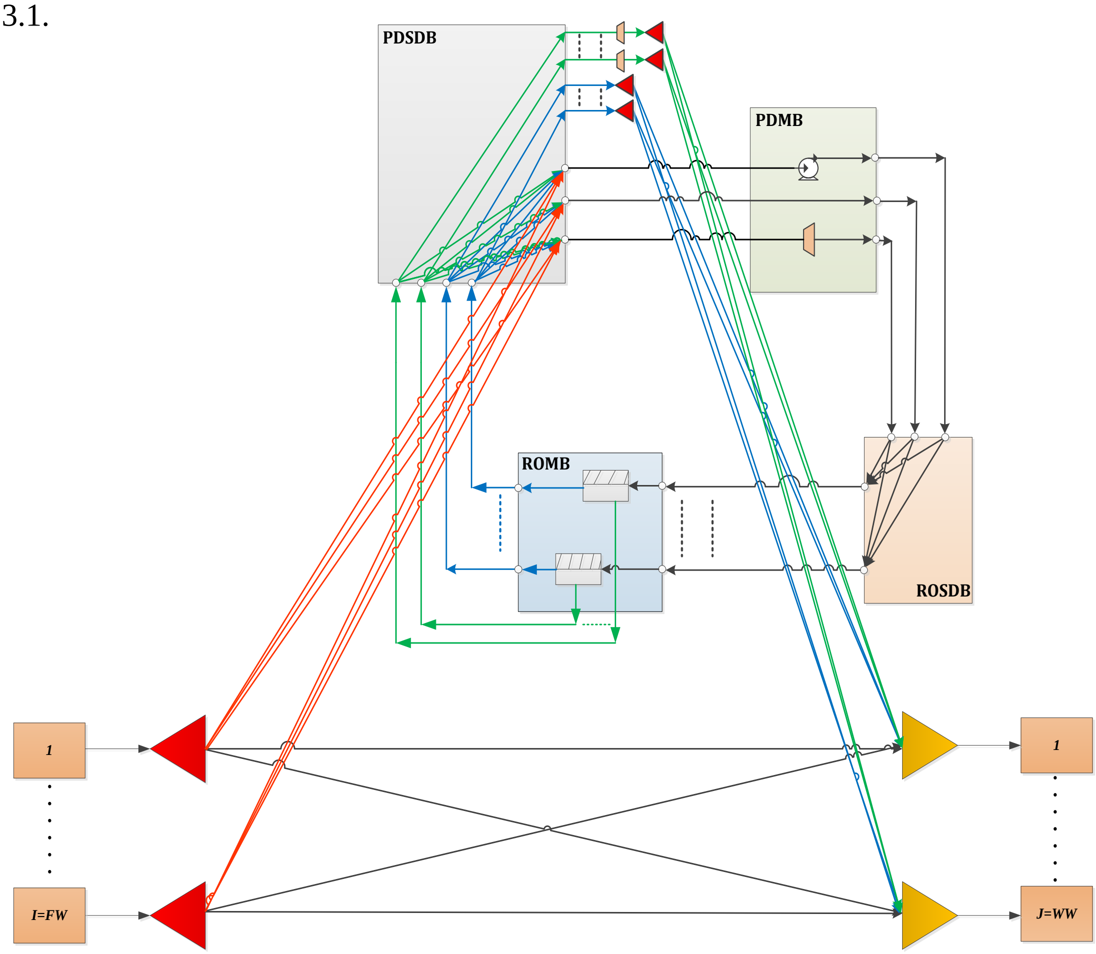


Figure 3.1: Superstructure representation of the RON superstructure within the WN superstructure

3.2 Mathematical Model

The RON superstructure proposed by El-Halwagi (1992) must however be modified in order to incorporate it within the WN. This is achieved by modifying the pressurisation/depressurisation stream-distribution box (PDSDB) and the pressurisation/ depressurisation matching box (PDMB) section of the superstructure. The properties of the updated pressurisation/depressurisation stream-distribution box PDSBD and pressurisation/depressurisation matching box PDMB are detailed below:

(i) Water sources are fed directly to node *n* for regeneration and are not mixed with retentate or permeate streams. This was incorporated to ensure that each reteantate and permeate stream leaves its respective regenerator, without further contamination.

(ii) Permeate and retentate streams are not allowed to mix in order for each stream to be fed directly from regenerator to the sinks. It is also assumed that each permeate stream will leave the regenerator at atmospheric pressure. Retentate streams, however, leave the RO at high pressures and are therefore passed through an energy recovery turbine for reduction in pressure to atmospheric pressure before distribution to the sinks.

(iii) Different retentate streams or permeate streams are also not allowed to mix in order to feed each stream directly to the water sinks. Mixing of the streams within the water sinks is decided by water quality requirements of the sink.

(iv) Each retentate stream or permeate stream can, therefore, go directly to a retentate node or can be recycled back to node *n* for further cleaning by the regenerators.

(v) A stream that does not require a pressure change can be fed directly to the ROSDB where it is then fed to the ROMB.

(vi) Inlet streams to box PDMB can either go to a pump or to an energy recovery turbine. The illustration of this idea is modified in order to clearly explain the original idea proposed by El-Halwagi (1992).

These modifications are illustrated in Figure 3.2(a) and 3.2(b). Figure 3.2(a) shows the original PDSDB proposed by El-Halwagi (1992) and Figure 3.2(b) shows the modified PDSDB and PDMB which will be incorporated with the WN.

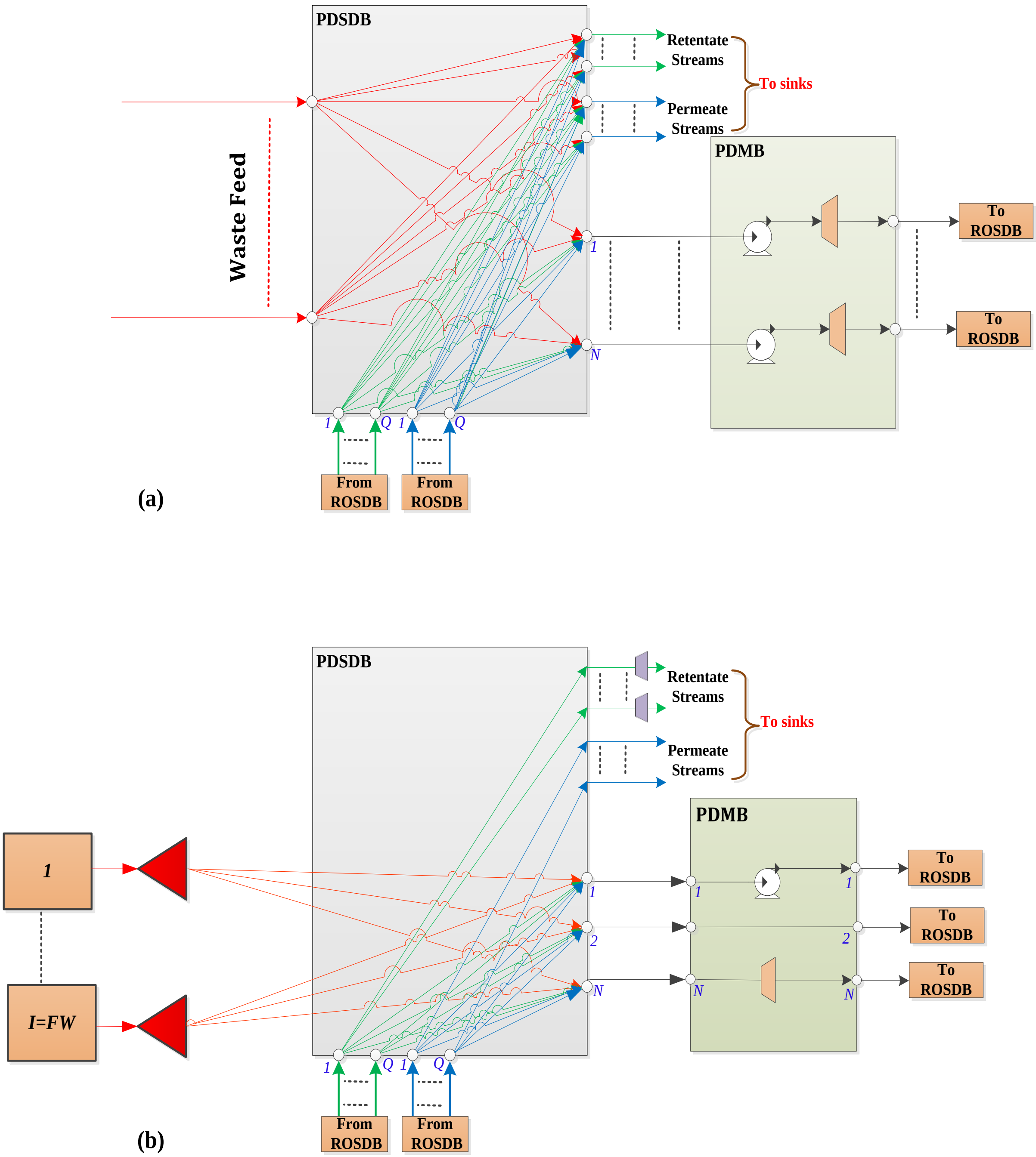


Figure 3.2: a) Original PDSDB and PDMB by El-Halwagi (1992). b) New modification to PDSDB and PDMB.

3.2.1 Water balances for the sources

Figure 3.3 shows a schematic representation of the water sources. From the diagram it can be seen that a water source can be fed to the PDSDB, wastewater sink or to the water sinks. The flowrate balance is shown in constraint (3.1).

$$F_i = \sum_{j=1}^{J} F_{i,j}^{s} + \sum_{n=1}^{N} F_{i,n}^{d} \qquad \forall i \in I \qquad (3.1)$$

It should also be noted that the freshwater source is included in the model as the last source within the model formulation. It can also be sent to the regenerators for further cleaning as its contaminant concentration is not zero.

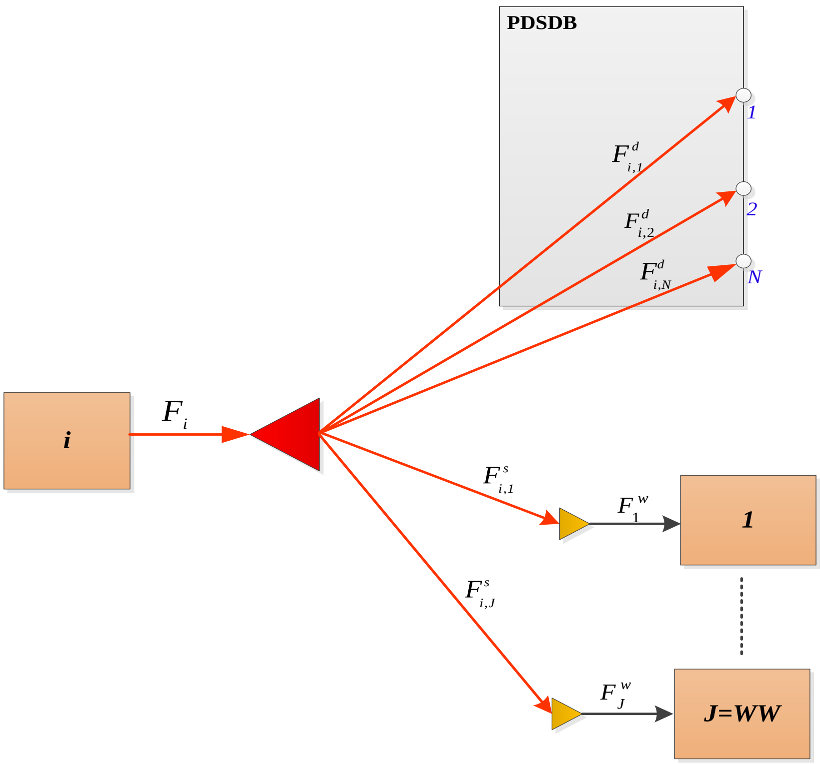


Figure 3.3: Schematic representation of the water sources.

3.2.2 Water balances for the sinks

Figure 3.4 shows a schematic representation of the water sinks. From the diagram it can be seen that the water sinks receive water from the water sources, permeate and retenetate of the regeneration units as well as the freshwater source. This flowrate balance is shown in constraint (3.2).

$$F_j^w = \sum_{i=1}^{I} F_{i,j}^s + \sum_{q=1}^{Q} F_{q,j}^r + \sum_{q=1}^{Q} F_{q,j}^p \qquad \forall j \in J \qquad (3.2)$$

Each sink can however handle a certain concentration limit. Constraint (3.3) implies that, the load to each sink must not exceed the maximum allowable load to that particular sink.

$$C_{j,m}^U \geq \frac{\sum_{i=1}^{I} F_{i,j}^s C_{i,m}^s + \sum_{q=1}^{Q} F_{q,j}^r C_{q,m}^r + \sum_{q=1}^{Q} F_{q,j}^p C_{q,m}^p}{F_j^w} \qquad \begin{array}{l} \forall j \in J \\ \forall m \in M \end{array} \qquad (3.3)$$

It should be noted that the wastewater sink is considered as the last sink. The maximum allowable load to this sink is also given in order to comply with the standard effluent discharge limits imposed by environmental regulations.

In order to forbid the mixing of permeate and retentate streams from one regenerator in the same sink, constraint (3.4) is added to the model as follows:

$$y_{q,j}^p + y_{q,j}^r \leq 1 \qquad \begin{array}{l} \forall q \in Q \\ \forall j \in J \end{array} \qquad (3.4)$$

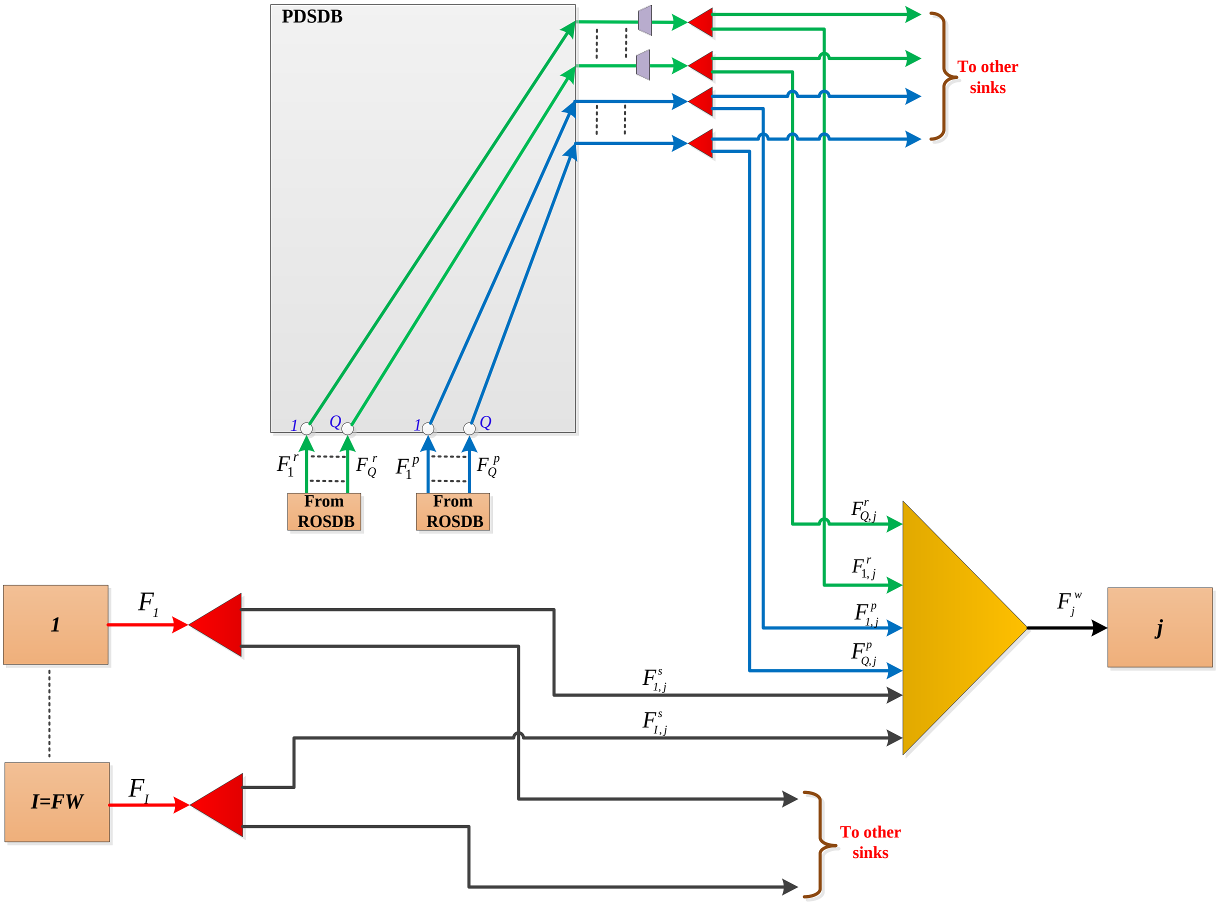


Figure 3.4: Schematic representation of the water sinks.

3.2.3 Regeneration unit (RON superstructure)

Figure 3.1 shows the schematic representation of the updated RON superstructure within the WN superstructure. Figure 3.5 therefore shows the interaction of the PDSBD with the sources and sinks of the WN.

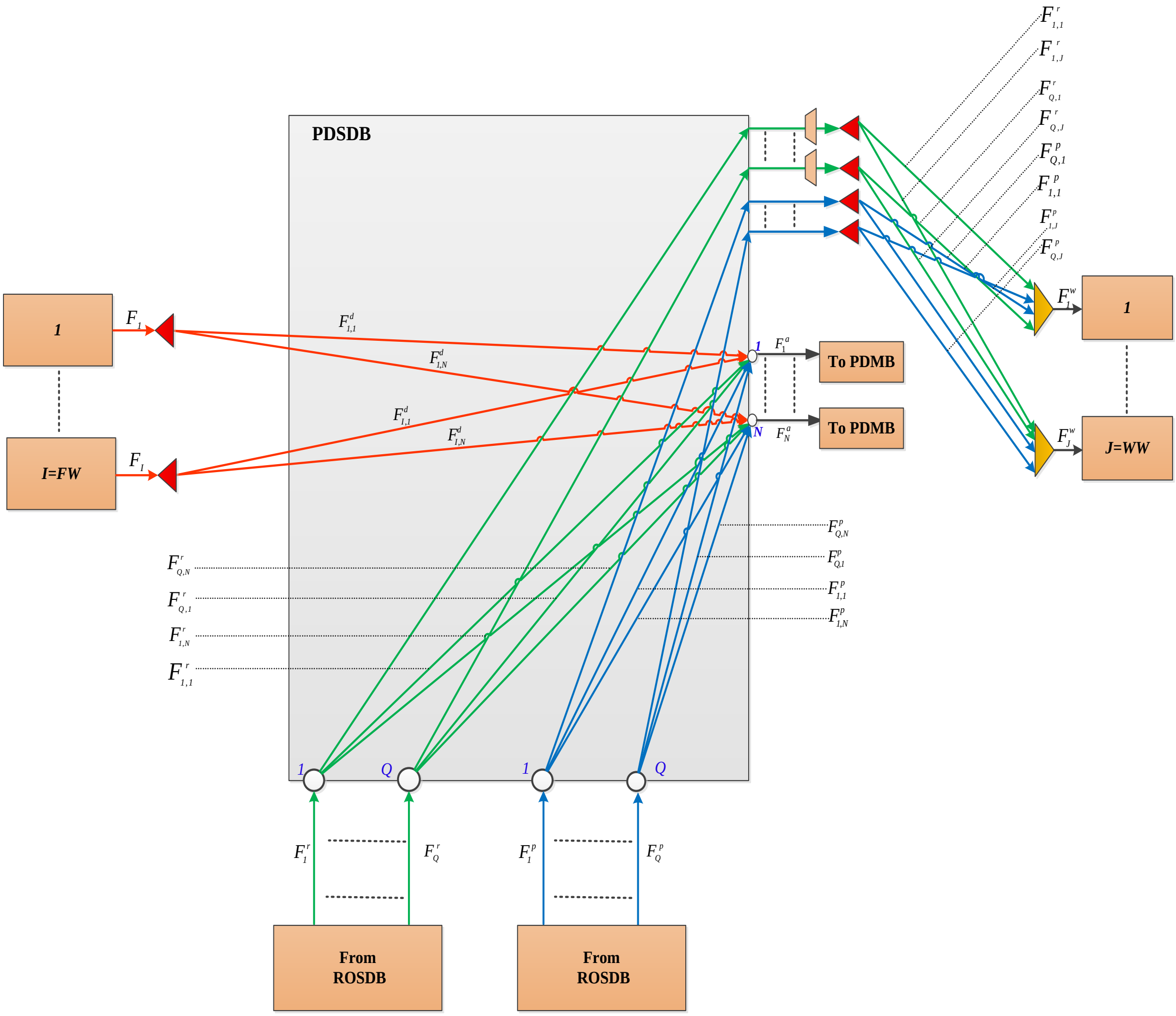


Figure 3.5: Schematic representation of the modified PDSDB.

Performance equations

The performances of the RO regenerators are represented by means of the liquid recovery α_q and removal ratio $RR_{q,m}$. The liquid recovery is the amount of the feed flowrate into the regenerator that exits in the permeate stream. The removal ratio $RR_{q,m}$ refers to the fraction of the inlet mass load that exits in the retentate stream of the regenerators (Khor et al., 2011). Constraints (3.5) and (3.6) represent the α_q and $RR_{q,m}$ respectively.

$$\alpha_q = \frac{F_q^p}{F_q^f} \qquad \forall q \in Q \qquad (3.5)$$

$$RR_{q,m} = \frac{C_{q,m}^r F_q^r}{C_{q,m}^f F_q^f} \qquad \forall q \in Q \quad \forall m \in M \qquad (3.6)$$

The recommended operating flowrate for RO modules is given in constraint (3.7) and is determined by the manufacturers. Constraint (3.8) gives the upper bound for the feed pressure into the RO membranes.

$$F^L \le \frac{F_q^f}{N_q^s} \le F^U \qquad \forall q \in Q \qquad (3.7)$$

$$P_q^f \le P_{\max} \qquad \forall q \in Q \qquad (3.8)$$

RON superstructure equations

a) Constraints for PDSDB

Constraint (3.9) shows the flowrate balance for the outlet junction of the PDSDB as can be seen in Figure 3.5. The node *n* represents a mixing junction at the outlet of the PDSBD.

$$F_n^a = \sum_{i=1}^{I} F_{i,n}^d + \sum_{q=1}^{Q} F_{q,n}^p + \sum_{q=1}^{Q} F_{q,n}^r \qquad \forall n \in N \qquad (3.9)$$

Constraint (3.10) shows the corresponding concentration balance for the outlet junction of the PDSDB.

$$F_n^a C_{n,m}^a = \sum_{i=1}^{I} F_{i,n}^d C_{i,m}^d + \sum_{q=1}^{Q} F_{q,n}^p C_{q,m}^p + \sum_{q=1}^{Q} F_{q,n}^r C_{q,m}^r \qquad \forall n \in N \quad \forall m \in M \qquad (3.10)$$

The balance for the flowrate and concentration of the permeate stream entering the PDSDB to the sinks is shown in constraint (3.11) and (3.12) respectively.

$$F_q^p = \sum_{n=1}^{N} F_{q,n}^p + \sum_{j=1}^{J} F_{q,j}^p \qquad \forall q \in Q \qquad (3.11)$$

$$F_q^p C_{q,m}^p = \sum_{n=1}^{N} F_{q,n}^p C_{q,m}^p + \sum_{j=1}^{J} F_{q,j}^p C_{q,m}^p \qquad \forall q \in Q \quad \forall m \in M \qquad (3.12)$$

The balance for the flowrate and concentration of the retentate stream entering the PDSDB to the sinks is shown in constraint (3.13) and (3.14) respectively.

$$F_q^r = \sum_{n=1}^{N} F_{q,n}^r + \sum_{j=1}^{J} F_{q,j}^r \qquad \forall q \in Q \qquad (3.13)$$

$$F_q^r C_{q,m}^r = \sum_{n=1}^{N} F_{q,n}^r C_{q,m}^r + \sum_{j=1}^{J} F_{q,j}^r C_{q,m}^r \qquad \forall q \in Q \quad \forall m \in M \qquad (3.14)$$

Since the permeate and retentate streams from the regenerator are at different pressures, constraints have to be given in order to ensure that streams are at the same pressures before they mix. This is shown in constraint (3.15), (3.16) and (3.19) for the feed, permeate and retentate streams. Constraint (3.18) shows the isobaric mixing of streams within the ROSDB.

$$(P_n^a - P_i^w) F_{i,n}^d = 0 \qquad \forall n \in N \quad \forall i \in I \qquad (3.15)$$

$$(P_n^a - P_q^p) F_{q,n}^p = 0 \qquad \forall n \in N \quad \forall q \in Q \qquad (3.16)$$

$$(P_n^a - P_q^r) F_{q,n}^r = 0 \qquad \forall n \in N \quad \forall q \in Q \qquad (3.17)$$

$$(P_n^a - P_n^o) F_{n,q}^a = 0 \qquad \forall n \in N \quad \forall q \in Q \qquad (3.18)$$

b) Constraints for PDMB and ROSDB

In the PDMB, the turbine is used to reduce the pressure of a stream while the pump is used to increase the pressure. Constraints (3.19) and (3.20) represent the principles of an energy recovery turbine and a pump respectively. Figure 3.6 shows the schematic representation of the PDMB and RODB.

$$\left(P_n^i - P_n^a\right) \geq 0 \qquad \forall n \in N \qquad (3.19)$$

$$\left(P_n^i - P_n^o\right) \geq 0 \qquad \forall n \in N \qquad (3.20)$$

The flowrate balance for the inlet of the ROSDB is given in constraint (3.21).

$$F_n^a = \sum_{q=1}^{Q} F_{n,q}^a \qquad \forall n \in N \qquad (3.21)$$

The outlet flowrate and concentration balance for the ROSDB is given in constraints (3.22) and (3.23) respectively.

$$F_q^f = \sum_{n=1}^{N} F_{n,q}^a \qquad \forall q \in Q \qquad (3.22)$$

$$F_q^f C_{q,m}^f = \sum_{q=1}^{Q} F_{n,q}^a C_{n,m}^a \qquad \forall q \in Q \quad \forall m \in M \qquad (3.23)$$

The maximum inlet concentration limit to the regenerators must also be specified since not all of the waste streams can be fed to the RO membrane and this is shown in constraint (3.24).

$$C_{q,m}^U \geq \frac{\sum_{q=1}^{Q} F_{n,q}^a C_{n,m}^a}{F_q^f} \qquad \forall q \in Q \quad \forall m \in M \qquad (3.24)$$

c) Binary variables for the existence of units

Constraint (3.25) shows that a booster pump exits in the RON if the P_n^i is larger than the pressure of the stream entering the PDMB and this forces the binary variable b_n to become one. A similar concept is used to represent the existence of an energy recovery turbine and is given in constraint (3.26) It is however illogical to pressurise and depressurise a stream simultaneously. Constraint (3.27) is therefore needed to prevent a turbine and pump from appearing in series.

$$P^L b_n \le P_n^i - P_n^a \le P^U b_n \qquad \forall n \in N \qquad (3.25)$$

$$P^L t_n \le P_n^i - P_n^o \le P^U t_n \qquad \forall n \in N \qquad (3.26)$$

$$b_n + t_n \le 1 \qquad \forall n \in N \qquad (3.27)$$

Constraint (3.28) indicates the existence of RO unit which is defined by the flowrate of the permeate stream from the regenerator *q*.

$$Fl^L r_q \le F_q^p \le Fl^U r_q \qquad \forall q \in Q \qquad (3.28)$$

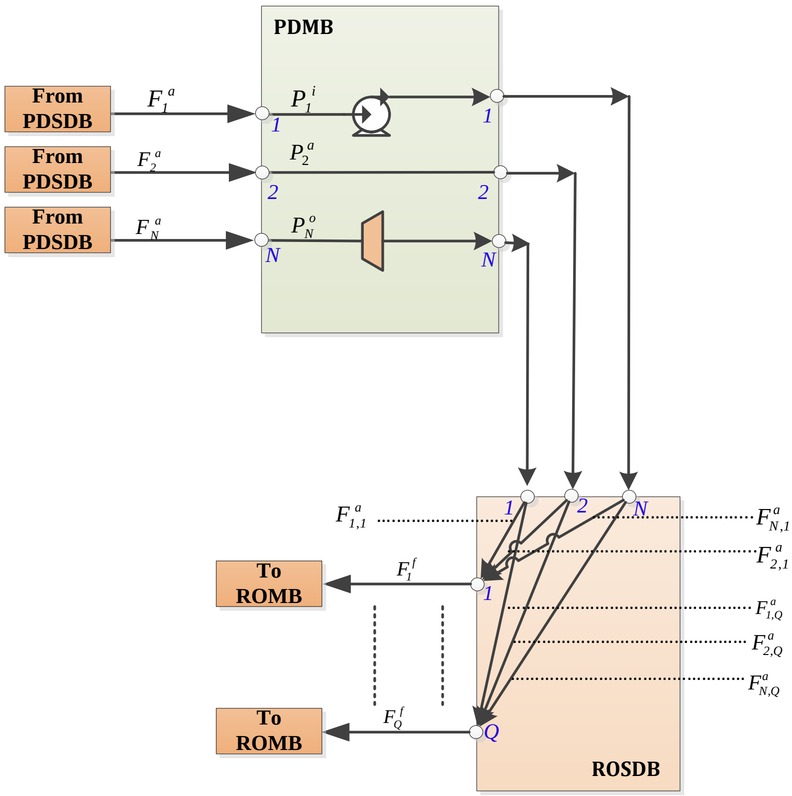


Figure 3.6: Schematic representation of the PDMB and ROSDB.

The characteristic of the RO membrane needs to be described in order to relate flowrate to pressure. The pressure drop across the membrane is calculated as the difference in pressure between the feed side and permeate side pressure and is shown in constraint (3.29) (El-Halwagi, 1997). The pressure on the retentate side is calculated as the pressure difference between the feed and the shell side pressure drop per module, ΔP_q^m . This is shown in constraint (3.30). The pressure drop across the membrane ΔP_q in terms of shell side pressure drop per module is given in constraint (3.31) (Khor et al., 2011). The equation was simplified by assuming a linear-shell side concentration and pressure profiles (El-Halwagi, 1997). The schematic representation of the ROMB is given in Figure 3.7.

$$\Delta P_q = \frac{P_q^f + P_q^r}{2} - P_q^p \qquad \forall q \in Q \qquad (3.29)$$

$$P_q^r = P_q^f - \Delta P_q^m \qquad \forall q \in Q \qquad (3.30)$$

$$\Delta P_q = P_q^f - \left(\frac{\Delta P_q^m}{2} + P_q^p \right) \qquad \forall q \in Q \qquad (3.31)$$

The osmotic pressure, $\Delta\pi_q$, is defined as a function of the contaminant concentration on the feed side (Saif et al., 2008a) and is shown in constraint (3.32). The osmotic pressure on the permeate side is however neglected due to its low contaminant concentration.

$$\Delta\pi_q = OS \sum_{m=1}^{M} C_{q,m}^{av} \qquad \forall q \in Q \qquad (3.32)$$

The permeate flowrate per module is given in constraint (3.33).

$$\frac{F_q^p}{N_q^s} = AS_m (\Delta P_q - \Delta\pi_q) \qquad \forall q \in Q \qquad (3.33)$$

The average concentration $C_{q,m}^{av}$ on the feed side is given by constraint (3.34).

$$C_{q,m}^{av} = \frac{C_{q,m}^f + C_{q,m}^r}{2} \qquad \forall q \in Q \quad \forall m \in M \qquad (3.34)$$

The concentration of contaminants on the feed side must also be described in terms of the pressure drop and the osmotic pressure. This is described in constraint (3.35).

$$C^{p}_{q,m} = \frac{k_m C^{av}_{q,m}}{A(\Delta P_q - \Delta\pi_q)\gamma} \qquad \forall q \in Q \quad \forall m \in M \qquad (3.35)$$

A mass and concentration balance around the regenerator is also needed and is described in constrain (3.36) and (3.37) respectively.

$$F^{f}_{q} = F^{p}_{q} + F^{r}_{q} \qquad \forall q \in Q \qquad (3.36)$$

$$F^{f}_{q} C^{f}_{q,m} = F^{p}_{q} C^{p}_{q,m} + F^{r}_{q} C^{r}_{q,m} \qquad \forall q \in Q \quad \forall m \in M \qquad (3.37)$$

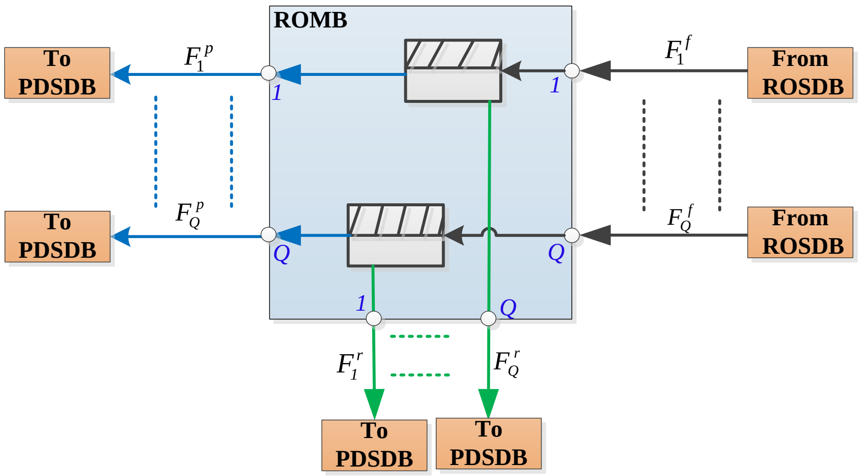


Figure 3.7: Schematic representation of the ROMB.

3.3 Big-M constraints

In order to determine the existence of piping interconnections, logical constraints and discrete variables will be adopted. This is adopted in order to reject small flowrates that are unnecessary for the design of the plant. This formulation makes use of the big-*M* parameters adopted by Khor et al. (2011). In the big-*M* parameters, *M* is a valid upper/lower bound denoted by *U* and *L* respectively. The piping interconnections of flowrates below the lower bound are then eliminated from the final design. Constraints (3.38) to (3.41) represent the big-*M* parameters for the piping interconnections between the different units. The upper and lower bounds are chosen based on the information given by the water sources.

$$M_{i,j}^{L} y_{i,j} \leq F_{i,j}^{s} \leq M_{i,j}^{U} y_{i,j} \tag{3.38}$$

$$M_{q,j}^{L} y_{q,j}^{p} \leq F_{q,j}^{p} \leq M_{q,j}^{U} y_{q,j}^{p} \tag{3.39}$$

$$M_{q,j}^{L} y_{q,j}^{r} \leq F_{q,j}^{r} \leq M_{q,j}^{U} y_{q,j}^{r} \tag{3.40}$$

$$M_{i,n}^{L} y_{i,n}^{d} \leq F_{i,n}^{d} \leq M_{i,n}^{U} y_{i,n}^{d} \tag{3.41}$$

3.4 Objective function

The objective function of the combined RON superstructure and WN superstructure is used to minimise the overall cost of the regeneration network on an annualised basis which consists of:

(i) TAC of the RON
(ii) cost of freshwater (*FW*)
(iii) treatment cost of wastewater (*WW*)
(iv) capital and operation costs of the piping interconnection

The total annualised cost of the RON consists of the capital cost of RO modules, pump, and energy recovery turbines, operating cost of pumps and turbines as well as pretreatment of chemicals. The operating revenue of the energy recovery turbine is also considered in the determination of the TAC and is shown in constraint (3.42).

$$TAC(q,n) = C^{pump}\left(\sum_{n=1}^{N}(P_n^i - P_n^a)F_n^a\right)^{0.65} + C^{tur}\left(\sum_{n=1}^{N}(P_n^i - P_n^o)F_n^a\right)^{0.43} \qquad \begin{array}{l}\forall q \in Q \\ \forall n \in N\end{array} \qquad (3.42)$$
$$+ C^{tur}\left(\sum_{j=1}^{J}(P_q^r - P_j^r)F_{q,j}^r\right)^{0.43} + C^{elec}AOT\left(\frac{\sum_{n=1}^{N}(P_n^i - P_n^a)F_n^a}{\eta^{pump}}\right)$$
$$- C^{elec}AOT\left(\sum_{j=1}^{J}(P_q^r - P_j^r)F_{q,j}^r\right)\eta^{tur} - C^{elec}AOT\left(\sum_{j=1}^{J}(P_n^i - P_n^o)F_n^a\right)\eta^{tur}$$
$$+ C^{\mathrm{mod}}\sum_{q=1}^{Q}N_q^s + C^{chem}AOT\sum_{i=1}^{I}F_{i,n}^d$$

The piping cost of components will be formulated by assuming a linear fixed-charge model. In their formulation, the cost particular of a pipe is incurred if the flowrate through the pipe falls below the threshold value. This is achieved by using 0-1 variables. Constraint (3.43) represents the objective function of the total regeneration network. It is also assumed that all the pipes share the same properties of p_c and q_c and a 1-norm distance D. The cost of piping also includes the approximate length and the material of construction.

$$\min\left(\begin{array}{l} \sum_{q=1}^{Q}\sum_{n=1}^{N}TAC_{q,n} + AOTC^{water}FW + AOTC^{waste}WW \\ + AA\left(\sum_{i=1}^{I}\sum_{j=1}^{J}D_{i,j}\left(\frac{p_c F_{i,j}^s}{3600v} + q_c y_{i,j}\right)\right) \\ + AA\left(\sum_{q=1}^{q}\sum_{j=1}^{J}D_{q,j}^p\left(\frac{p_c F_{q,j}^p}{3600v} + q_c y_{q,j}^p\right)\right) \\ + AA\left(\sum_{q=1}^{q}\sum_{j=1}^{J}D_{q,j}^r\left(\frac{p_c F_{q,j}^r}{3600v} + q_c y_{q,j}^r\right)\right) \\ + AA\left(\sum_{i=1}^{I}\sum_{n=1}^{N}D_{i,n}^d\left(\frac{p_c F_{i,n}^d}{3600v} + q_c y_{i,n}^d)\right)\right) \end{array}\right) \qquad (3.43)$$

where $AA = \left(\dfrac{m(1+m)^n}{(1+m)^n - 1^n} \right)$ is the annualisation factor.

The overall model results in a nonconvex MINLP due to the bilinear terms as well as the power function in the constraints.

3.5 Nomenclature

3.3.1 Sets

I={i|i=water source}
J={j|j=water sink}
M={m|m=contaminants}
Q={q|q=regeneration units}

3.3.2 Parameters

α_q	liquid recovery
$RR_{q,m}$	removal ratio
F^U	maximum flowrate per hollow fiber module
F^L	minimum flowrate per hollow fiber module
ΔP_q^m	shell side pressure drop per module
M^U	upper bound of big-M constant for interconnections between streams
M^L	lower bound of big-M constant for interconnections between streams
AOT	annual operating time
p_c	parameter for carbon steel piping based on CEPCI value of 318.3
q_c	parameter for carbon steel piping based on CEPCI value of 318.
v	velocity
A	water permeability coefficient
P_{max}	maximum allowable pressure for the regenerators
k_m	solute permeability constant
L	fiber length
L_s	seal length
r_o	outside radius of fiber
r_i	inner radius of fiber
S_m	membranearea per module
P^U	an arbitrary big value for pressure
P^L	an arbitrary small value for pressure
Υ	a dimensionless constant
η^{pump}	pump efficiency
η^{tur}	turbine efficiency
OS	proportionality constant between the osmotic pressure

	and average salt mass fraction on the feed side
$C^U_{j,m}$	maximum allowable contaminant concentration *m* in sink *j*
$C^U_{q,m}$	maximum allowable contaminant concentration *m* into a regenerator *q*
$D^a_{i,j}$	manhattan distance between water source *i* and sink *j*
$D^p_{q,j}$	manhattan distance between regenerator *q* and sink *j*
$D^r_{q,j}$	manhattan distance between regenerator *q* and sink *j*
$D^d_{i,n}$	manhattan distance between source *i* and node *n*
$C_{i,m}$	mass fraction of contaminant *m* within water source *i*
C^{chem}	cost parameter for chemicals
C^{elec}	cost of electricity
C^{mod}	cost per module of HFRO membrane
C^{pump}	cost coefficient for pump
C^{tur}	cost coefficient for turbine
C^{waste}	freshwater cost
C^{water}	waste water cost
μ	water viscosity
P^p_q	pressure of a permeate stream from regenerator *q*
P^w_i	pressure of source *i*
P^r_j	pressure of the retentate stream in sink *j*
Fl^L	lower bound on flowrate
Fl^U	upper bound on flowrate
P^L	lower bound on pressure
P^U	upper bound on pressure

3.3.3 Continuous Variables

$F^s_{i,j}$	allocated flowrate between source *i* and sink *j*
$F^d_{i,n}$	allocated flowrate between source *i* and node *n*
F_i	flowrate of sources *i*
$F^p_{q,j}$	flowrate of the permeate stream from regenerators *q* to sink *j*
$F^r_{q,j}$	flowrate of the retentate stream from regenerators *q* to sink *j*
$F^a_{n,q}$	flowrate of streams from node *n* to regenerator *q*
F^f_q	flowrate leaving the outlet junction of ROSDB
F^p_q	flowrate of permeate stream leaving the regenerator *q*
$F^p_{q,n}$	flowrate of permeate stream regenerator *q* to node *n*
F^r_q	flowrate of retentate stream leaving the regenerator *q*
$F^r_{q,n}$	flowrate of retentate stream from regenerator *q* to node n
F^a_n	flowrate of streams from node

n

F_j^w flowrate of sink j

$C_{n,m}^a$ concentration of contaminant m in stream leaving node n

$C_{q,m}^f$ concentration of contaminant m in the feed to the regenerator q

$C_{q,m}^p$ concentration of contaminant m in permeate stream leaving regenerator q

$C_{q,m}^r$ concentration of contaminant m in retentate stream leaving regenerator q

$C_{q,m}^{av}$ average concentration of contaminant m in the high-pressure side of regenerator q

P_n^a pressure of streams leaving node n

P_n^i pressure of an inlet stream to an energy recovery turbine from node n

P_n^o pressure of an outlet stream from an energy recovery turbine from node n

ΔP_q pressure drop over regenerator q

P_q^f feed pressure into regenerator q

P_q^r pressure of a retentate stream from regenerator q

P_q^p pressure of a permeate stream from regenerator q

$\Delta\pi_q$ osmotic pressure on the retentate side of regenerator q

FW freshwater flowrate

WW wastewater flowrate

3.3.4 Binary Variables

$$b_n = \begin{cases} 1 \leftarrow \text{if a pump exits} \\ 0 \leftarrow \text{otherwise} \end{cases}$$

$$t_n = \begin{cases} 1 \leftarrow \text{if a turbine exits} \\ 0 \leftarrow \text{otherwise} \end{cases}$$

$$r_q = \begin{cases} 1 \leftarrow \text{if regenerator } q \text{ exits} \\ 0 \leftarrow \text{otherwise} \end{cases}$$

$$y_{q,j}^p = \begin{cases} 1 \leftarrow \text{if piping exits between the permeate streams and sink } j \\ 0 \leftarrow \text{otherwise} \end{cases}$$

$$y_{q,j}^r = \begin{cases} 1 \leftarrow \text{if piping exits between the retentate streams and sink } j \\ 0 \leftarrow \text{otherwise} \end{cases}$$

$$y_{i,j} = \begin{cases} 1 \leftarrow \text{if piping exits between source } i \text{ and sink } j \\ 0 \leftarrow \text{otherwise} \end{cases}$$

$$y_{i,n}^d = \begin{cases} 1 \leftarrow \text{if piping exits between source } i \text{ and node } n \\ 0 \leftarrow \text{Otherwise} \end{cases}$$

3.3.4 Integer Variables

N_q^s the number of hollow fiber modules of regenerator *q*

3.6 References

El-Halwagi, M., 1992. Synthesis of reverse osmosis networks for waste reduction. American Institute of Chemical Engineers, pp.1185-98.

El-Halwagi, M., 1997. Pollution Prevention through Process Integration. San Diego: Academic Press.

Khor, CS, Foo, DCY., El-Halwagi, MM., Tan, RR., Shah, N., 2011. A superstructure optimisation approach for membrane separation-based water regeneration networks synthesis with detailed nonlinear mechanistic reverse osmosis model. Industrial and Engineering Chemistry Research , 50, pp.13444-13456.

Saif, Y., Elkamel , A. & Pritzker, M., 2008a. Optimal design of reverse osmosis networks for wastewater treatment. Chemical Enineering and Processing , pp.2163-74.

CHAPTER

4

RESULTS AND DISCUSSION

4.1 Introduction

The results obtained for applying the model to a petroleum refinery case study are presented. The model was applied to four cases in order to highlight the importance of incorporating a detailed RON superstructure within a WN superstructure. The schematic diagram for all four cases is also presented in order to show the complete design of the WN. A comparison of the best case and the “black-box” approach is also discussed in order to highlight the importance of using a detailed model for the regenerators.

4.2 Illustrative Example

The above model was applied to a literature based refinery case study (Khor et al., 2011). The model was implemented in GAMS 24.2 using the general purpose global optimisation solver BARON which obtains a solution by using a branch-and-reduce algorithm. The network consists of 4 sources and 4 sinks. The limiting water data for the sources and sinks is given in

Table **4.1**.

Table 4.2 shows the Manhattan distances between different units. The distances between the regenerators and the sinks for both permeate and retentate streams are the same.

Table **4.3** presents the process and economic data for the detailed RON. The economic data and the model parameters are given in

Table 4.1: Limiting data for water network.

Sources, *i*					**Sinks, *j***				
i	**Unit**	**Flowrate (kg/s)**	**Contaminant Concentration (kg/m^3)**		***J***	**Unit**	**Flowrate (kg/s)**	**Max Contaminant Concentration (kg/m^3)**	
			TDS	COD				TDS	COD
1	Amine Sweeting	7.3	3.5	3.5	1	Caustic Treating	0.83	2.5	2.5
2	Distillation	10.65	4	4	2	Menox-I Sweeting	40	2	2
3	Hydrotreating	3.5	1	3	3	Desalting	5.56	2.5	2.5
4	Freshwater	___	2	1	*4*	Wastewater	___	25	25

Table 4.2: Manhattan Distance for the case study.

	Sinks				**Regenerator unit**	
Sources	**1**	**2**	**3**	**4**	**1**	**2**
1	50	50	50	60	50	50
2	60	50	60	70	40	40
3	50	50	50	60	65	50
4	60	50	60	70	100	50
Regenerator unit						
1	80	70	60	70		
2	60	10	40	20		

Table 4.3: Process and economic data for the detailed RON.

Parameter	Value
Pure water permeability, A	$5.50\text{x}10^{-13}$ m/(s.Pa)
Shell side pressure drop per module per regenerator , P_m	$4.05\text{x}10^{4}$ Pa
Solute permeability coefficient, k_m	$1.82\text{x}10^{-8}$ m/s
Fibre length, L	0.75 m
Seal length, L_s	0.075 m
Outside radius of fiber, r_o	$42\text{x}10^{-6}$ m
Inner radius of fiber,r_i	$21\text{x}10^{-6}$ m
Membrane area, S_m	180 m
Water viscosity, μ	0.001 kg/(m.s)
Dimensionless constant, Υ	0.69
Permeate pressure per regenerator, $P_p(q)$	101325 Pa
Pump efficiency, η_{pump}	0.7
Turbine efficiency, $\eta_{turbine}$	0.7
Liquid recovery for all regenerators, $\alpha(q)$	0.7
Osmotic constant, OS	$4.14\text{x}10^{-7}$ Pa
Cost parameter for chemicals, $C_{chemical}$	0.11$/kg
Cost of electricity, C_{elec}	0.06 $/(kW.h)
Cost coefficient for pump, C_{pump}	6.5 $/(yearW0.65)
Cost coefficient for pump, C_{tur}	18.4 $/(yearW0.43)

Cost per module of HFRO membrane, C_{mod}	2300 \$/(year.module)
Maximum flowrate per hollow fiber module, F^U	0.27 kg/s
Minimum flowrate per hollow fiber module, F^L	0.21 kg/s

Table 4.4: Economic data and the model parameters for WN.

Parameter	**Value**
Annual operating time, *AOT*	8760 h
Unit cost of freshwater, C_{water}	1 \$/kg
Unit cost of wastewater, C_{waste}	1 \$/kg
Interest rate per year, *m*	5%
Number of years, *n*	5 year
Parameter *p* for carbon steel piping	7200
Parameter *q* for carbon steel piping	250
Velocity, *v*	1 m/s

4.3 Scenarios Considered

Four scenarios will be compared in order to highlight the importance of incorporating a detailed RON superstructure within the water network.

(i) Firstly, the case in which no regeneration is considered within the water network is modeled in order to provide a basis (base case) for comparison (Case 1).
(ii) In the second case, a single regenerator is incorporated within the WN with fixed removal ratio (Case 2).
(iii) The third case looks at multiple regenerators within the WN with fixed removal ratio (Case 3).

(iv) In case 4 multiple regenerators with variable removal ratio are considered.

4.4 Results and Discussion

The results obtained from the optimisation are given Table 4.5 for case 1 to 3. In case 2 and 3, the regenerators had a fixed removal ratio of 0.95. In the first scenario, the water network with no regeneration had a higher total cost due to the high consumption of freshwater which can be seen in Table 4.5. The network is shown in Figure 4.1. The second scenario where a single regenerator was used led to a 15.26% reduction in freshwater usage and a 43.36% in reduction in wastewater generation in comparison with the base case. The overall cost of network was minimised by 17.6% due to the incorporation of the RO regenerator. The use of the energy recovery turbines in the RON led to a reduction in the regeneration cost of the network.

Figure 4.2 shows the complete water network and RON obtained for case 2. This diagram includes the distribution boxes as shown in Figure 3.1. Figure 4.2 can be translated into a simplified schematic diagram showing only the relevant physical units, i.e. RO membranes, pumps, turbines, mixes and splitters. Figure 4.3 shows the water network for case 2. In Figure 4.3 it can be seen that, one pump and turbine are needed for the regeneration as well as 20 HFRO modules. For simplicity in cases 3 and 4 only the simplified water network is presented.

Case 3 led to a 24.82% reduction in freshwater consumption and 70.82% reduction in wastewater generation in comparison with case 1. The total cost of the network was also reduced by 22.35%. The low cost of the water network is due to the low freshwater consumption and wastewater generation. The introduction of a second regenerator, Case 3, leads to further reduction in the total cost. This is due to the lower consumption in freshwater and wastewater generation. Figure 4.4 shows the water network for case 3. In Figure 4.4 it can be seen that, two pumps and turbines are needed for the regeneration as well as 37 HFRO modules per regenerator. A parallel configuration was chosen by the model.

Table 4.5: Summary of results for case 1 to 3

	No regeneration	**Single regenerator**	**Two regenerators**
		Fixed RR	Fixed RR
	(Case1)	(Case 2)	(Case 3)
Freshwater flowrate (kg/s)	38.40	32.54	28.87
Wastewater flowrate (kg/s)	13.40	7.59	3.91
Cost of regeneration (million $/year)	—	0.068	0.23
Total cost (million $/year)	1.70	1.40	1.32
CPU time (h)	0	0.13	6

Table 4.6 shows the comparison between case 3 and 4. The removal ratio chosen by the model in case 4 was 0.97 for all contaminants. Case 4 led to 3.12% reduction in freshwater and 30.43% reduction in wastewater generation in comparison with case 3. A 15.91% reduction in the total network cost was also achieved. The large decrease in the total cost of the network in case 4 can be attributed to the high removal ratio which was selected by the model rather than the value that was initially predicted. In comparison with the case where no regeneration was considered, case 4 leads to a 28% reduction in freshwater consumption and 80% reduction in wastewater generation. The

modeling of case 4 is however computationally expensive as can be seen in. The best case used 15 HFRO modules per regenerator. The model selected two regenerators, two pumps and two energy recovery turbines as can be seen in. It can also be seen that a parallel configuration of the network was chosen by the model. Flowrates obtained for the different streams are indicated on Figure 4.1 to Figure 4.5.

Table 4.6: Summary of results for case 3 and 4.

	Multiple regenerators	
	Fixed RR (Case 3)	Variable RR (Case 4)
Freshwater flowrate (kg/s)	28.87	27.68
Wastewater flowrate (kg/s)	3.91	2.72
Cost of regeneration (million \$/year)	0.23	0.096
Total cost (million \$/year)	1.32	1.11
Network configuration	Parallel	Parallel
Number of HFRO modules	37 for each regenerator	15 for each regenerator
CPU time (h)	6	54

The high computational time for solving the model in case 3 was due to the complexity of the problem as well as the large number of 0-1 variables. The model solves quicker when tighter bounds are imposed on the feed and retentate pressure. The use of the energy recovery turbines in the RON led to a reduction in the regeneration cost of the network, and as a result, a reduction in energy usage by the system was achieved. The statistics of the model for all the four cases is shown in. It can be seen from the

Table 4.7 that the amount of discrete variables increased from case 1 to case 4. This was due to the introduction of binary variables for indication of piping interconnection between the sources and RON as well as between the sinks and the RON. This increase is also as a result of the integer variables that are used to determine the number of HFRO modules within each RO membrane.

Table 4.7: Model statistics for case 1 to 4.

	No regeneration	**Single regenerator**	**Multiple regenerators**	
		Fixed RR	Fixed RR	Variable RR
	(Case1)	(Case 2)	(Case 3)	(Case 4)
Number of equations	60	168	282	282
Number of continuous variables	46	134	208	212
Number of discrete variables	16	32	48	48
Optimality gap	0.1	0.1	0.1	0.1

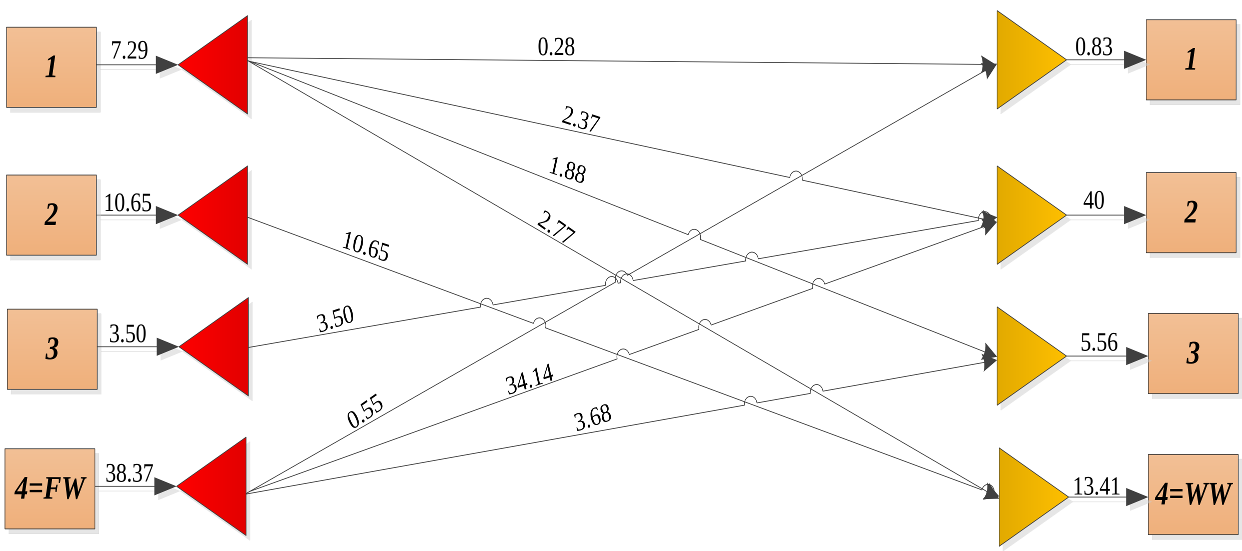


Figure 4.1: Network obtained for case 1 (No regeneration).

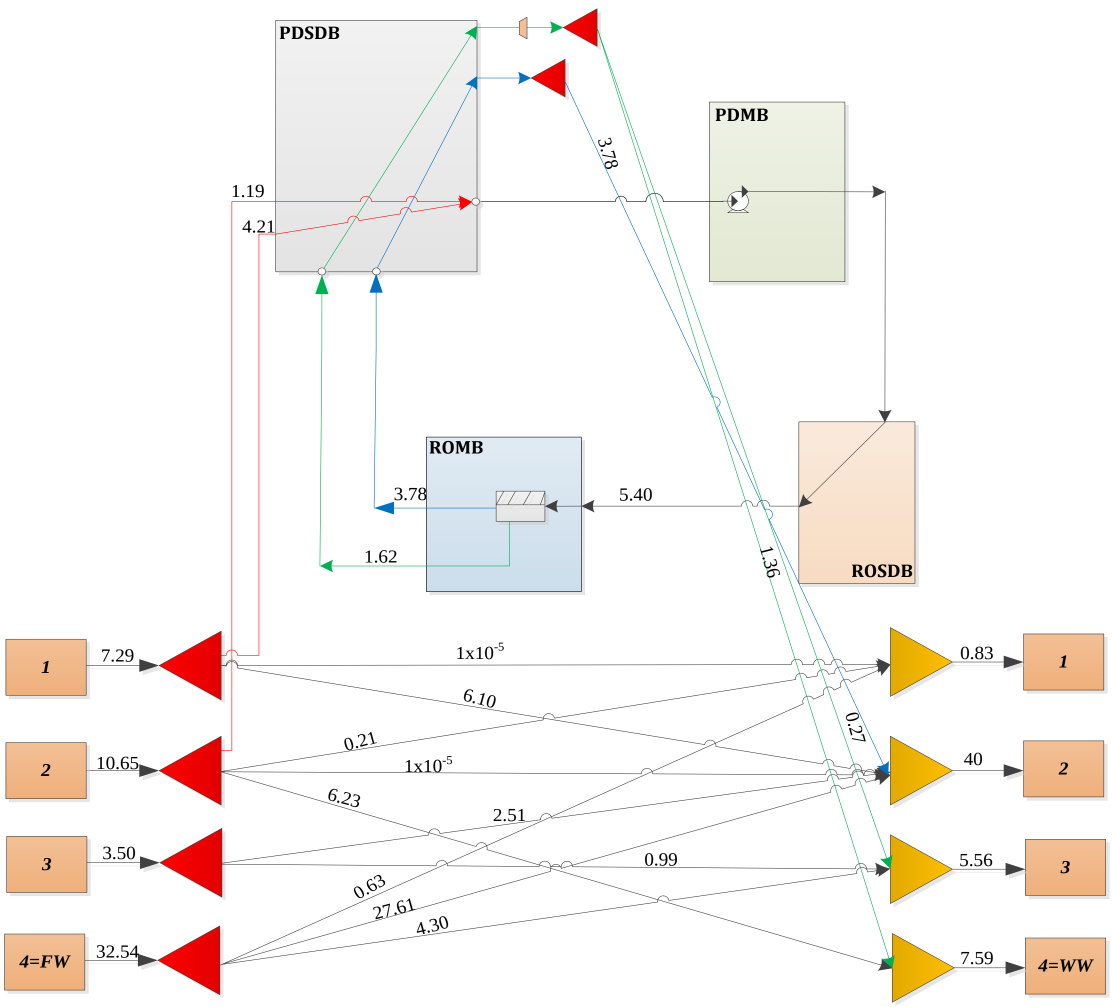


Figure 4.2: Network for case 2 based on the distribution boxes.

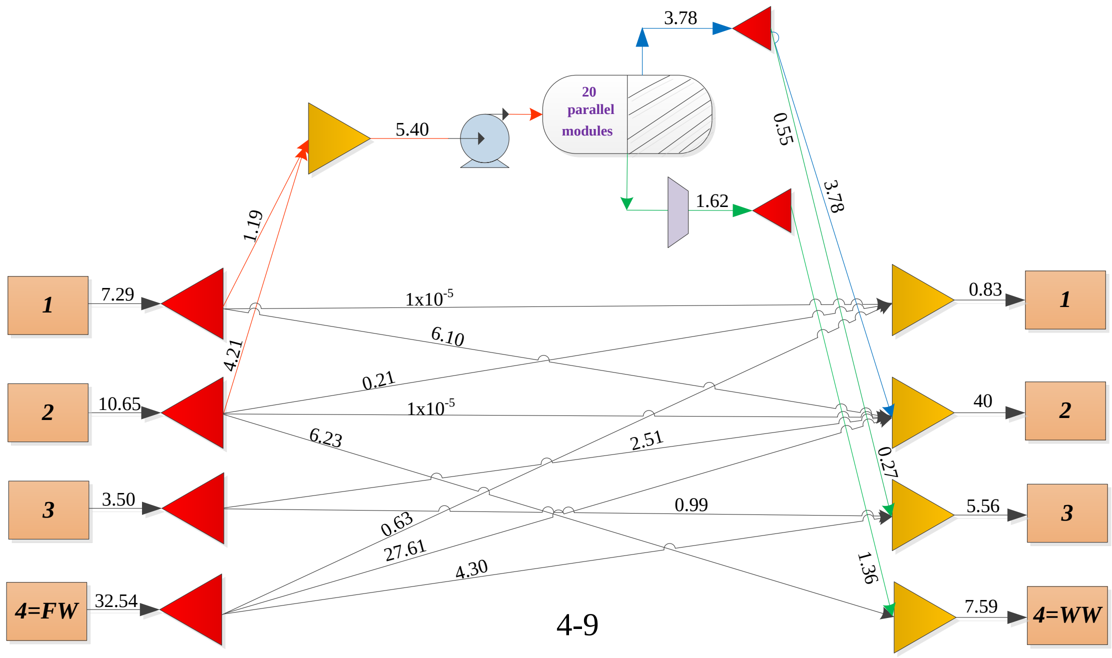


Figure 4.3: Network obtained for case 2 (Single regenerator with fixed removal ratio)).

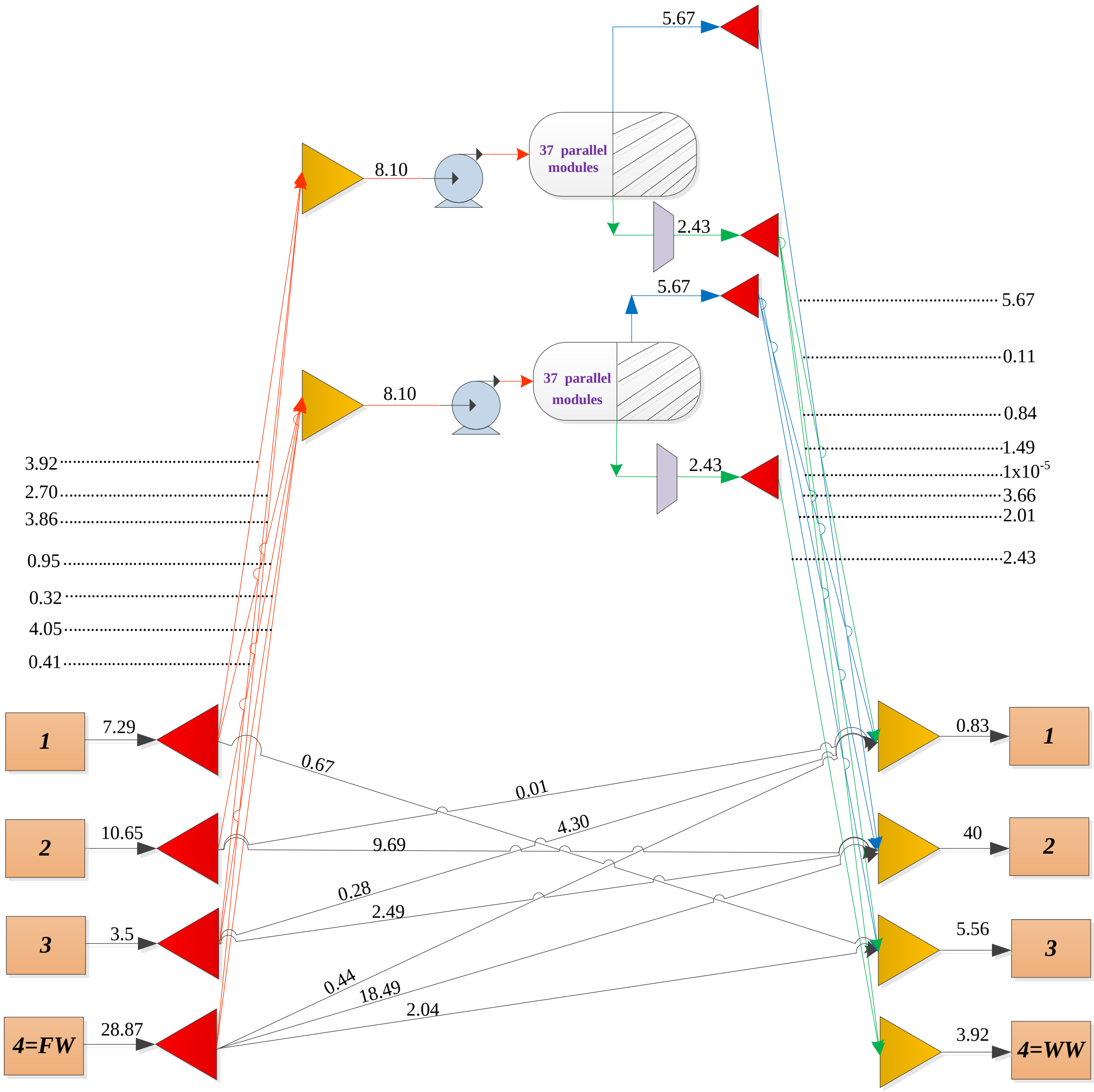


Figure 4.4: Network obtained for case 3 (Multiple regenerators with fixed removal ratio).

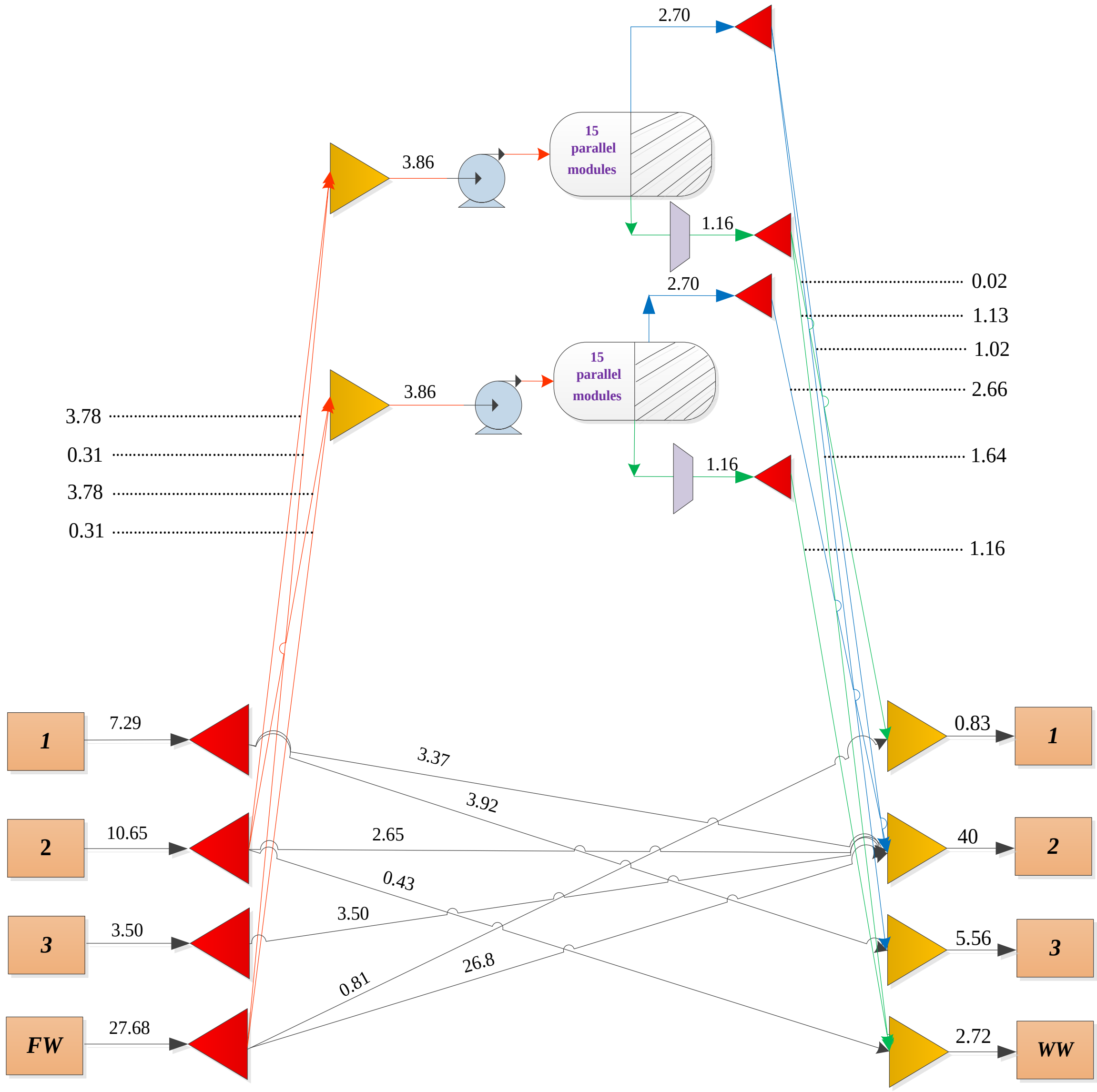


Figure 4.5: Network obtained for case 4 (multiple regenerators with variable removal ratio).

4.5 Black-box” and Detailed model

The importance of the proposed model can also be demonstrated by comparing the results obtained in case 4, which was the best case, to a “black box” model. The TAC of the “black-box” regenerators was adopted from the work Tan et al. (2009) which is dependent on the inlet flowrate into the regenerators. Table 4.8 shows the results for the apparent and true “black-box” models and case 4 (detailed model) which consisted of two regenerators. The true “black-box” model only considers the actual cost of regeneration. This was estimated using a detailed standalone regeneration model in order highlight the short comings of the “black-box” model in terms of regeneration cost.

Table 4.8: Summary of results for the “black-box” approach and case 4.

	Apparent “black-box” model	**True “black-box” model**	**Multiple regenerators** (Case 4)
Freshwater flowrate (kg/s)	28.15	28.15	27.68
Wastewater flowrate (kg/s)	3.20	3.20	2.72
Cost of regeneration (million $/year)	0.034	0.13	0.096
Total cost (million $/year)	1.06	1.15	1.11

From Table 4.8 it can be seen that, the “black-box” model led to a higher freshwater consumption and wastewater generation than in case 4. The regeneration cost estimated by the apparent “black-box” model was 73.85% less than that estimated by the true “black-box” model. The total cost of the true “black-box” model was 3.48% higher than that obtained by the by case 4. It can therefore be seen from the above results that, the "black-box" model does not give a true cost representation of the RO regenerators and as such a detailed model of the regenerators is needed to accurately determine the cost of regeneration.

4.6 References

Khor, CS, Foo, DCY., El-Halwagi, MM., Tan, RR., Shah, N., 2011. A superstructure optimisation approach for membrane separation-based water regeneration networks synthesis with detailed nonlinear mechanistic reverse osmosis model. Industrial and Engineering Chemistry Research , 50, pp.13444-13456.

Tan, R., Ng, D., Foo, D. & Aviso, K., 2009. A superstructure model for the sybthesis of single-contaminant water network with partitioning regenerators. Process Safety and Environmental Protection, 87(3), pp.197-205.

CHAPTER

5

RECOMMENDATIONS AND CONSIDERATIONS

5.1 Introduction

This chapter shows the methods that can be used to improve the solution procedure used in solving the model. Further considerations for a detailed design of the RO membranes are also discussed. It looks at the variables and equations that are critical to obtaining a solution to model with less computational time. The chapter also gives the relevant recommendation in order to highlight the shortcomings of the proposed model.

5.2 Detail design of RO membranes

In the modeling of the RO membranes, the membrane area, length of fiber, outer and inner radius of the HFRO modules were fixed. This parameters can however be used in the model as variables in order to obtain a complete design of the RO membranes obtained by the model as this values will be at the optimal solution. This therefore means that, the fixed values used in solving the model, are not necessary the optimal values.

In order to illustrate the importance of considering these parameters as variables in the model, the membrane area was made a variable (case 5) for the case 2 in Chapter 4. Table 5.1 shows the comparison between the case 2 where the membrane area was a parameter and case 5 where it is a variable.

Table 5.1: Summary of results for case 2 and 5.

	Single regenerator	**Single regenerator**
	Fixed S_m (Case 2)	Variable S_m (Case 5)
Membrane area (m^2)	180	316
Number of HFRO modules	20	32
Freshwater flowrate (kg/s)	32.54	28.0
Wastewater flowrate (kg/s)	7.59	3.04
Cost of regeneration (million $/year)	0.068	0.10
Total cost (million $/year)	1.40	1.12
CPU time (h)	0.13	0.083

It can be seen from Table 5.1 that, the membrane area chosen by the model in case 5 was higher than that in case 2. There was however a reduction in the freshwater usage (13.95%) and wastewater generation (60%) in comparison with case 2. The total cost of regeneration was also lower for case 5 as it was reduced by 20% compared to case 2 even though the cost of regeneration was higher in case 5 due to the increase in the number HFRO modules. It can therefore be concluded from the following results that, it is better for the model to choose the design variables for the RO membrane. It is therefore recommended that, the membrane area, length of fiber, outer and inner radius of the HFRO modules should be made variables in future work.

5.3 Method used to reduce computational time

From Chapter 4 it was seen that the computational time for case 4 with multiple regenerators was high due to the increase in the complexity of the problem as the number of discrete variables increased. The following methods were considered in aiding the convergence of the model:

(i) The relaxed model (RMINLP) was first solved using BARON. The solution obtained from the RMINLP model was then used as a starting point for the MINLP model and was then solved using DICOPT. This method was used to provide a better starting point that can aid in the convergence of the MINLP model. The method however failed to solve the problem. This could be attributed to the large number of discrete variables that can be seen from case 2 to case 4 of Table 4.7.

(ii) An adaptive numerical optimisation procedure proposed by Arora (2012) also used to help increase the convergence of the MINLP problem. In this method, the RMINLP is solved and the variables that are close to their discrete or integer value are then assigned that value. The variables are then held fixed and the optimisation problem was solved again. This procedure was then continued until all the variables were assigned their discrete or integer values. This method also failed and was tedious as the model consisted of a large number of discrete variables (Arora, 2012).

5.4 Recommended methods for Convergence

The optimisation problem was solved successfully with BARON even though the problem was computationally expensive. The following methods can be used to accelerate the converge rate of this model:

(i) The RLT method proposed by Quesada and Grossmann (1995) can be used to decrease the computational time of the model. This method can aid in the convergence of the problem as a linear model is first solved to provide a starting point and a lower bound for the nonlinear model. The procedure can aid in finding a near global or global optimum solutions.

(ii) The whole model can also be solved be linearising all the bilinear terms in the model. This can be achieved by using McCormick’s over and under estimators for the product of two continuous terms (not an exact method) and Glover transformations for the product of a continuous variable and an integer variable (exact method). Transformation methods can be used for other nonlinear terms within the model.

(iii) Piecewise linearisation can also be used in aiding the convergence of the proposed model (not an exact method).

5.5 Preprocessing of Variables

The idea behind preprocessing is that by reducing problem size one is able to also reduce data storage requirements and computational time (Hare et al., 2010). This involves the reduction of variables and constraints. The convergence of the MINLP problem can also be improved by knowing before optimisation, the variables that are critical for obtaining an optimal solution and quick solution convergence. This can be achieved through a thorough inspection of the mathematical model. The methods used in preprocessing include:

(i) Standard linear reductions

(ii) Knowing more information about problem to further inspect variables

(iii) Removing redundant constraints

(iv) Removing linear dependencies

(v) Eliminating fixed variables

(vi) Substituting out free variables with their complementary equations from the model.

(vii) Checking for consistency of bounds

Preprocessing steps are used by solvers like BARON. The preprocessing step used by BARON can be summarised as follows (Tawarmalani & Sahinidis, 2002):

a) The solver starts by looking at the user supplied starting points

b) It then looks at the user supply LPs that maximise/minimise the problem by using OA.

The methods i, ii and vii for the preprocessing of optimisation problems can be applied to the proposed model to reduce computational time and to aid in the convergence of the model. The following method can be used in the preprocessing stage:

(i) Knowing more information about problem to further inspect variables can be applied to the model during the preprocessing stage. The results indicated that, the feed position of streams and outlets of the system where the critical variables that need to be optimised (Lu et al., 2012). This therefore means that these variables are one of the many variables that are critical to the optimisation process. The mass load of permeate and retentate streams from the outlet of the regenerators therefore need to be linearised in order to aid in the convergence process of the MINLP problem.

(ii) Bounds were provided for the number of HFRO modules, feed pressure, retentate and permeate pressure, upper and lower bounds as this was needed for the model to be solved. This therefore means that, a bilinear term consisting of any of these variables can be linearised in other to aid the convergence process of the problem. These variables can, therefore be classified as critical variables.

(iii) Concentrations of the permeate and retentate streams are critical as this are dependent on the osmotic pressure and indicate the performance of the RO membrane. In the preprocessing stage, a product of these variables with any other variable should therefore be linearised.

(iv) The flowrate of the permeate stream is also associated with the osmotic pressure and should therefore be considered a critical variable for the optimisation process.

5.6 Recommendations for future work

The recommendations for the proposed model can be summed up as follows:

(i) Additional design variables of the RO membrane should be incorporated in order to obtain a more detailed design of the units. It is therefore recommended that the design of the RO unit should be fully chosen through the optimisation process. Parameters like the membrane area, inner and outer radius and length of fibers must be made variables in order to obtain the optimal design of the RO units. The importance of this is demonstrated in Table 5.1.

(ii) A more detailed model should be used to account for fouling and concentration polarisation within the model as their influence affects the performance of the RO membranes.

(iii) Other membrane technologies such as ultrafiltration can be included in the model to increase the flexibility of the model.

(iv) Convex relaxation methods must be used within the model in order to accelerate the convergence process of the model.

(v) The proposed model can be applied to large-scale petroleum case studies with multiple contaminants and this is therefore recommended.

5.7 References

Arora, J., 2012. Intoduction to optimum design. London: Academic Press Publications.

Hare, W., Liu, M. & Terlaky, T., 2010. Efficient preprossing for VLSI optimisation problems. Computational Optimisation and Applications, pp.475-493 (45).

Lu, Y., Liao, A. & Hu, Y., 2012. The design of reverse osmosis systems with multiple-feed and multiple-product. Desalination, 307, pp. 42-50.

Quesada , I. & Grossmann, I., 1995. Global optimisation of bilinear process networks with multi component flows. Computers and Engineering , 19(21), pp.1219-42.

Tawarmalani, M. & Sahinidis, N., 2002. Convexification and global optimisation in continuous and mixed-intege nonlinear programming. AA Dordrecht: Kluwer Academic Publishers.

CHAPTER

6

CONCLUSIONS

This work has addressed the synthesis of an optimal water regeneration network that incorporates the detailed synthesis of a RON. The model formulation resulted in an MINLP problem. The use of water and energy were optimised simultaneously. Both fixed and variable RR has been considered. Streams with multiple contaminants have also been considered in the model formulation.

The proposed model was applied to a literature case study with 4 sources and 4 sinks with multiple contaminants. It was then solved using GAMS/BARON in order to highlight its practicality. The results show that the use of multiple regenerators in the water network, can lead to a reduction in the total cost of the network due to the significant reduction in freshwater consumption and wastewater generation. It can also be concluded that, there is a significant benefit in allowing the removal ratio in the model to be a variable as this has significant impact on the cost and structure of the network.

Large computational times were however incurred due to the complex nature and structure of the model and relaxation methods must therefore be used together with the MINLP solver. It is also noteworthy that the proposed model was limited to one membrane technology. Multiple membrane technologies such as ultrafiltration can however be incorporated in the membrane network and thus offering a scope for future work. This is needed to increase the flexibility of the proposed model.