Optimization of Chemical Processes Using Mathematical Programming

A thesis submitted to the
Budapest University of Technology and Economics

for the degree of
Doctor of Philosophy

by
Abdulfatah M. Emhamed

under the supervision of
Zoltán Lelkes Ph. D.
associate professor

January 2009
Acknowledgements

First of all, I would like to thank my supervisor, associate professor Zoltan Lelkes, for his guidance, and for his support during my time at the Department of Chemical and Environmental Process Engineering at the Budapest University of Technology and Economics.

I am also very grateful to associate professor Endre Rév and Dr. Tivadar Farkas for their continuous assistance and for readily helping me throughout my PhD studies. Their suggestions greatly improved the level of the research I have done.

Special thank for Barbara Czuczai for her help and patience to finish this work and I will never ever forget that for her.

Herewith I would like to thank all the colleagues at the Department of Chemical and Environmental Process Engineering at the Budapest University of Technology and Economics, for maintaining a loving atmosphere at the department.

I am forever indebted to my all family especially my father that they encouraged and supported my studies from the very beginning of my life.

Finally, I would like to thank the Higher Education Secretary for the financial support they provided during my PhD studies.
Table of Contents

1. Introduction .................................................................................................................... 6
2. Literature Review ........................................................................................................... 8
   2.1. Optimization in general ........................................................................................... 9
      2.1.1. Practical applications ..................................................................................... 10
      2.1.2. Solution strategies ......................................................................................... 10
          2.1.2.1. Solution strategies for LP ........................................................................ 10
          2.1.2.2. Solution strategies for NLP ..................................................................... 11
          2.1.2.3. Solution strategies for MILP ................................................................ 12
          2.1.2.4. Solution strategies for MINLP .............................................................. 14
      2.1.3. Commercial tools for optimization problems ................................................ 14
   2.2. Extractive Distillation ........................................................................................... 16
      2.2.1. Extractive distillation process ........................................................................ 16
      2.2.2. Strange behavior of extractive distillation ..................................................... 18
      2.2.3. Selection of the solvent .................................................................................. 18
      2.2.4. Simulation and optimization of extractive distillation ................................... 19
   2.3. Mass Exchange Networks ..................................................................................... 25
      2.3.1. Problem Statement of Mass Exchange Network ........................................... 25
      2.3.2. Mass Exchange Processes .............................................................................. 26
      2.3.3. Mass Exchangers ........................................................................................... 27
      2.3.4. Mass Exchange Networks synthesis methods ................................................ 27
          2.3.4.1. Pinch Technology ................................................................................... 27
          2.3.4.2. Mathematical Programming .................................................................... 30
              2.3.4.2.1. Superstructure .................................................................................. 31
              2.3.4.2.2. MINLP Formulation ........................................................................ 33
          2.3.4.3. The Hybrid Method ................................................................................. 34
   2.4. Desalination processes .......................................................................................... 36
      2.4.1. Desalination Technologies ............................................................................. 36
          2.4.1.1. Reverse Osmosis (RO) ............................................................................ 37
          2.4.1.2. Solar Distillation ..................................................................................... 38
      2.4.2. Desalination in Libya ..................................................................................... 38
      2.4.3. Desalination Cost ........................................................................................... 38
      2.4.4. Desalination Plant Location ........................................................................... 39
3. Scope and challenges of the thesis ............................................................................... 42
   3.1. Analysis of Extractive Distillation with Mathematical Programming .................. 42
   3.2. New hybrid method for mass exchange network optimization ............................ 43
   3.3. Optimization of desalination location problem using MILP ................................ 43
4. Analysis of Extractive Distillation with Mathematical Programming ......................... 44
   4.1. Proposed method ................................................................................................... 44
      4.1.1. Model and software ....................................................................................... 45
      4.1.2. Exploration of the feasible region .................................................................... 45
      4.1.3. Studying the effect of cost factors on the optimal configuration .................... 46
4.2. Examples and results.............................................................................................. 49
  4.2.1. Acetone/methanol mixture............................................................................. 49
    4.2.1.1. General specifications........................................................................... 49
    4.2.1.2. Exploration of the feasible region....................................................... 50
      4.2.1.2.1. Minimum stage numbers in the extractive column......................... 50
      4.2.1.2.2. Limiting flows............................................................................... 52
    4.2.1.3. The effect of cost factors on the optimal configuration......................... 54
  4.2.2. Ethanol/water mixture.................................................................................. 62
    4.2.2.1. General specifications........................................................................... 62
    4.2.2.2. Exploration of the feasible region....................................................... 62
      4.2.2.2.1. Minimum stage numbers in the extractive column......................... 62
      4.2.2.2.2. Limiting flows............................................................................... 64
    4.2.2.3. The effect of cost factors on the optimal configuration......................... 66
  4.3. Conclusion ............................................................................................................ 75

5. New hybrid method for mass exchange network optimization............................ 76
  5.1. Motivation for developing a new method......................................................... 76
  5.2. New hybrid method algorithm......................................................................... 79
  5.3. Demonstration of the new hybrid method ....................................................... 82
  5.4. Conclusion ............................................................................................................ 87

6. Optimization of desalination location problem using MILP................................. 88
  6.1. Problem statement and test problems.............................................................. 88
    6.1.1. Problem I.................................................................................................... 88
    6.1.2. Problem II.................................................................................................. 89
  6.2. Model versions and development strategy....................................................... 89
    6.2.1. The Basic model...................................................................................... 90
      6.2.1.1. Superstructure................................................................................... 90
      6.2.1.2. Model formulation............................................................................ 90
        6.2.1.2.1. Material balance constraints....................................................... 91
        6.2.1.2.2. Distances.................................................................................... 91
        6.2.1.2.3. Capacity constraints.................................................................. 92
        6.2.1.2.4. Logic expressions..................................................................... 92
        6.2.1.2.5. Transformation of logic expressions........................................ 93
        6.2.1.2.6. Assignment of feasible and infeasible regions........................... 93
          6.2.1.2.6.1. Use of triangles.................................................................... 93
          6.2.1.2.6.2. Use of polygons.................................................................. 95
        6.2.1.2.7. Objective function.................................................................... 96
        6.2.1.2.8. Additional equations.................................................................. 96
      6.2.1.3. Model variants.................................................................................... 97
    6.2.1.4. Test runs of the Basic model.................................................................. 98
      6.2.1.4.1. Results of applying the Basic model to Problem I.......................... 98
        6.2.1.4.1.1. Results with triangles............................................................ 98
        6.2.1.4.1.2. Results with the use of polygons......................................... 99
      6.2.1.4.2. Results of applying the Basic model to Problem II....................... 100
      6.2.1.5. Evaluation of the results of the Basic model...................................... 102
    6.2.2. The Improved model............................................................................... 102
      6.2.2.1. Superstructure.................................................................................. 103
1. Introduction

Optimization is a very important task in chemical industry because it can be applied to maximize the profit (or minimize the consumption of the energy or production cost, etc.). Optimization techniques have many applications in science and engineering. As the systems become more complex so become controlling and optimizing their performance increasingly important; hence optimization is positioned as an indispensable and fundamental tool. Most chemical engineering design tasks can be modeled as mixed integer programming problems.

The overall objective of the thesis is to apply mixed integer programming (MIP) to find the optimal solution of some important industrial chemical problems. Because of the importance of this topic in the field of chemical industries located in Libya, three projects were selected for this purpose, such as 1) analysis and optimization of extractive distillation process, 2) mass exchange network synthesis and 3) desalination location problems.

These problems involve mass transfer and heat transfer processes.

Application of the mathematical programming approach to problems of design, process integration, and operation consists of three major steps. The first step is developing a representation of alternatives from which the optimal solution is selected. The second is formulating a mathematical program that generally involves discrete and continuous variables for the selection of the configuration and operating levels, respectively. The third one is solving the optimization model and thus finding the optimal solution. A mathematical program consists of variables, constraints, and an objective. The constraints are mathematical statements expressing physical, technological, economical, or other kind of relations and prohibitions of variable value coexistences. A complete set of variable values is called a solution; those possible solutions which satisfy all the constraints are feasible solutions. Solutions are compared to each other and their value is quantified by an objective function (e.g., cost, profit, generated waste) which is to be minimized or maximized.

Section 2 provides a deep insight into the state-of-the-art of the selected topics (optimization in general, extractive distillation, mass exchange network synthesis and desalination processes). Section 3 outlays the aim and scope of this thesis.
Section 4 deals with studying and analysis of extractive distillation through optimization. Section 5 presents a new hybrid method for synthesizing mass exchange networks. Section 6 handles the problem of finding the optimal location of desalination plants. Finally, in the last section, the major new results are summarized.
2. Literature Review
The present thesis deals with different fields of process optimization and analysis. The three areas are 1) analysis of extractive distillation using optimization, 2) synthesis of mass exchange networks, and 3) finding the optimal location of desalination plants.

The common feature of these subfields that optimization (i.e. mathematical programming) is used as a tool for design.

Therefore before discussing the available literature and state of the art of these research areas, I provide a short fundamental survey about optimization in general: the problem formulation, practical applications, and solution approaches, in Section 2.1.

A thorough review of extractive distillation is presented in the next section: about its theory, its strange behavior, and about its modeling and design aspects.

The question of optimal design of mass exchange network and its available tools – Pinch technology and different mathematical programming approaches – are discussed in Section 2.3.

Finally the need of developing a model for optimally locating desalination plants is explained. Although there are no even similar problems arising in the literature, there are some theoretical approaches for much simpler problems, which are worth to mention: those are also discussed in Section 2.4.
2.1. Optimization in general

Optimization deals with finding the optimum (minimum or maximum) value of a function of variables while satisfying a system of constraints. A thorough survey on the field is written by Floudas (1995).

The mathematical form of a mathematical programming problem is the following:

Formulation $P$

$$\min_{\bar{x}, \bar{y}} f(\bar{x}, \bar{y}) \text{ subject to}$$

$$h_i(\bar{x}, \bar{y}) = 0 \quad i = 1, ..., M$$
$$g_j(\bar{x}, \bar{y}) \leq 0 \quad j = 1, ..., P$$

$$\bar{x} \in X \subseteq \mathbb{R}^n$$
$$\bar{y} \in Y \quad \text{integer}$$

where $\bar{x}$ is a vector of $n$ continuous variables, $\bar{y}$ is a vector of integer variables. $h_i(\bar{x}, \bar{y}) = 0$ are $M$ equality constraints, $g_j(\bar{x}, \bar{y}) \leq 0$ are $P$ inequality constraints, and $f(\bar{x}, \bar{y})$ is the so-called objective function, whose minimum (or maximum) value is to be found.

Formulation $P$ contains a number of classes of optimization problems, by appropriate consideration or elimination of its elements. If the set of integer variables is empty, and the objective function and constrains are linear, then Formulation $P$ becomes a Linear Programming (LP) problem. If the set of integer variables is empty, and there exist nonlinear terms in objective function and/or constrains, then Formulation $P$ becomes a NonLinear Programming (NLP) problem. If the set of integer variables is nonempty, the set of continuous variables is empty, and the objective function and constrains are linear, then $P$ represents an Integer Linear Programming (ILP) problem. If the set of integer variables is nonempty, the integer variables participate linearly and separately from the continuous, and the objective function and constrains are linear, then $P$ represents a Mixed-Integer Linear Programming (MILP) problem. If the set of integer variables is nonempty, and there exist nonlinear terms in objective function and/or constrains, then Formulation $P$ becomes a Mixed-Integer NonLinear Programming (MINLP).
Remark: It is important to emphasize that the above terms (LP, MILP etc.) here refer to the type of the optimization problem and not to the solution strategy.

2.1.1. Practical applications
Optimization is frequently applied in numerous areas of chemical engineering including the development of process models from experimental data, design of equipment, and the analysis of chemical processes under uncertainty and adverse conditions.
Each the discrete optimization problem can be formulated as an MI(N)LP problem.
For example, the classical traveling salesman problem (TSP) deals with finding the shortest route that a salesman has to take to visit a set of towns. Currently intensively researched area is, for example, the production scheduling, where binary variables are used to assign resources to tasks and to decide on their sequence.
MINLP models are often used in the field of process synthesis. Important area is, for example, the optimal design of distillation columns. The existence of equilibrium stages is modeled with binary variables, and the relationships for phase equilibrium, and component balances are described using nonlinear equations.
MINLP models are also used for synthesizing mass exchange networks; see a detailed literature survey on the topic in Section 2.3.

2.1.2. Solution strategies

2.1.2.1. Solution strategies for LP
The simplex algorithm, see Dantzig (1963), developed by the American mathematician George Dantzig in 1947, is the popular algorithm for numerical solution of LP problems.
It tests adjacent vertices of the feasible set (which is a polytope) in sequence so that at each new vertex the objective function improves or is unchanged. The simplex method is very efficient in practice, generally taking $2m$ to $3m$ iterations at most (where $m$ is the number of equality constraints), and converging in expected polynomial time for certain distributions of random inputs, although theoretically it requires exponential time for
general problem, i.e. the computational time necessary for the algorithm to solve the problem cannot be expressed as a polynomial function of the problem size.

In 1984, N. Karmarkar (see Karmarkar, 1984) proposed a new interior point projective method for linear programming. In contrast to the simplex algorithm, which finds the optimal solution by progresses along points on the boundary of a polyhedral set, interior point methods move through the interior of the feasible region. The new algorithm guarantees optimal solution in polynomial time even in the worst-case problem class. Since then, many interior point methods have been proposed and analyzed. An exhausting review on the history of interior point algorithms in recent decades can be found in Potra and Wright (2000).

Remark: Dantzig's simplex method should not be confused with the downhill simplex method (Spendley et al., 1962, Nelder and Mead, 1965). The latter method solves an unconstrained minimization problem in n dimensions by maintaining at each iteration n+1 points that define a simplex. At each iteration, this simplex is updated by applying certain transformations to it so that it "rolls downhill" until it finds a minimum.

Using the commonly available computational resources, even large scale LP problems can be easily solved to optimality within seconds.

2.1.2.2. Solution strategies for NLP

The algorithms for nonlinear optimization problems can be broadly classified as 1) search without use of derivatives, 2) search using first order derivatives, 3) search using second order derivatives. The simplex method and the Hooke-Jeeves method belong to the first group. The steepest descent method is an example for the second group: it performs a line search along the direction of the negative gradient of the minimized function. The well-known Newton method uses second order derivatives.
2.1.2.3. Solution strategies for MILP

The major difficulty that arises in MILP problems is due to the combinatorial nature of the domain of discrete variables. Any choice of 0 or 1 for the elements of the vector \( y \) results in an LP problem on the variables \( x \) which can be solved easily to optimality. Theoretically it is possible to follow the brute force method enumerating all possible combinations of the 0-1 values for the elements of vector \( y \). However, such an approach would result in a computation time increasing exponentially with the size of the problem, i.e. with the number of integer variables. In case of one hundred binary variables, the number of possible combinations is \( 2^{100} \), and that is the number of LP problems to be solved with enumeration. That will remain impossible for the foreseeable future; therefore this solution strategy is not an option for real scale problems.

Mixed-integer programming problems have still no known algorithms that can solve them in polynomial time.

The solution algorithms available for discrete optimization problems can be classified into two categories: 1) metaheuristics that may be able to quickly find solutions of good quality but not to prove that that is the optimal one, and 2) exact algorithms that insure to find the global optimum but require long computation time. The former group includes methods like random search, taboo search, simulated annealing, and genetic algorithm. As the scope of the present thesis is not concerned with the former one, here only exact algorithms are dealt with in detail.

The proposed algorithms for MILP can be classified into as:

- Branch and bound methods
- Cutting plane methods
- Decomposition methods
- Other

The most commonly used algorithm by commercial solvers is the first one. In the branch and bound algorithm, a binary tree is applied for representing the 0-1 combinations and valid upper and lower bounds are generated at different levels of the binary tree. The lower bound, for minimization problems, is calculated from the partially relaxed version of the original problem where the remaining binary variables, which are still not branched according to their 0-1 values, can take any values between 0 and 1. The upper bound is
the value of the momentarily best integer solution. Branches forking from nodes representing infeasible relaxed problems or, for minimization problems, having lower bounds higher than the best current solution, can be pruned off, thereby speeding up the solution process. The algorithm terminates when the lower bound becomes equal to the upper bound.

Cutting plane methods are sometimes effective alternatives of the branch and bound algorithm. In that case valid inequalities are added to the linear relaxation problem, which do not cut off any feasible integer solution from the search space but cut off the current feasible non-integer optimum. The resulted LP-relaxation problem is optimized again. Adding cuts to the relaxation problem is repeated until the current optimum is also an integer one. A useful survey on cutting planes is written by Marchand et al. (2002).

Decomposition methods decompose the original problem into two or more subproblems which are easier, and solve them independently. The two most frequently used are Bender’s decomposition and Lagrangean decomposition. The first one first creates a master problem and a subproblem, the way of the decomposition is based on the actual structure of the original problem. The algorithm iteratively first solves the master problem and the obtained values are fixed at their actual values; then it solves the subproblem at the fixed values of the variables and the obtained solution is used for creating cuts to the next masterproblem.

The Lagrangean method decomposes the set of integer variable into two subsets and creates two subproblems defined by the two subsets. Some constraints are relaxed and their infeasibilities are used as penalty in the objective function. It can be proved that for maximization problems, the sum of the objective functions of the two subproblems is upper bound for the original problem, and at suitably selected weighting factor in the objective function, the sum of the objective functions equals to the original optimum value.

Besides the above mentioned techniques, there are different approaches, integrating interdisciplinary areas of mathematics. For example, graph theoretic methods are used for optimizing MILP programs by Hegyháti et al. (2008).
2.1.2.4. Solution strategies for MINLP

One of the main two algorithms for solving MINLP problems is the Outer Approximation (OA) algorithm. It can be classified as a decomposition scheme in which the continuous optimization and the discrete optimization are performed separately. The continuous optimization is performed through NLP subproblems that arise for fixed choices of the binary variables. The discrete optimization is performed via MILP master problems which are linear approximations to the MINLP problem. An iterative bounding procedure is carried out in which, for convex problems, the master problems provide lower bounds and the NLP subproblems provide upper bounds for the optimal objective function to be minimized. One step of the bounding procedure (solution of the succeeding MILP and NLP problems) is called a major iteration cycle.

Global optimality for a given MINLP solution by the OA algorithm is guaranteed when both the objective function and the feasibility region of the MINLP model are convex. The feasibility region of the MINLP problem is convex when all the equalities consist of linear functions and the inequalities consist of convex functions (Duran and Grossmann, 1986a,b).

A different solution algorithm was invented ARKI Consulting & Development A/S, based on the branch and bound algorithm originally developed for MILP problems. Its principle is the same; the difference is that here NLP problems are to be solved at the level of the leaves of the binary tree. Finding the global optimum is not guaranteed in case of nonconvex problems.

2.1.3. Commercial tools for optimization problems

There are several available commercial softwares for solving every type of mathematical programming problems.

The most effective for solving LP and MILP problems is the advanced CPLEX (ILOG, 2000) from ILOG. It uses the well-known simplex algorithm for LP, and mainly the Branch&Bound algorithm for MILP problems, complemented with highly effective heuristics to speed up the solution process. The two altogether are able to solve even large-scale MILP problems.
A widely used solver for NLP problems is CONOPT (Drud, 1996) which uses Generalized Reduced Gradient (GRG) method for solving NLP problems.

Only few tools exist for solving MINLP problems: one is DICOPT (Grossmann et al., 2008) based on the above detailed Outer Approximation algorithm, another is SBB (GAMS, 2008) based on the nonlinear variant of the Branch&Bound concept, and there exists an advanced tool, a customizable OA-based solver from AIMMS (Bisschop and Roelofs, 2007).

Besides commercial tools, open source solvers are also available. Several examples are GNU LP Kit, COIN LP, LPSOLVE, and QSOPT for LP; and LPSOLVE, COIN Brach and Cut for MILP.

The mathematical programming problems themselves can be formulated using modeling systems such as GAMS (Brooke et al., 1992), AMPL (Fourer et al., 1992) or LINGO from the LINDO Systems Inc, and AIMMS (Bisschop and Roelofs, 2007). The latter is the most advanced tool amongst the listed. AIMMS distinguishes itself from other optimization software by a number of advanced modeling concepts, as well as a full graphical user interface both for developers and end-users.
2.2. Extractive Distillation

Distillation is the most widely used technique to separate a mixture of liquid components; the separation is based on the difference in relative volatility of the components (V. Hoof et. al. 2004). Distillation as large scale industrial application is carried out in vertical distillation columns. According to column internals, packed columns and plate columns are distinguished.

Separation of close boiling or azeotropic mixtures by distillation is not (economically) feasible. When two phases are present, most economical methods may be decanting and two-phase splitting. If it is not the case, a possibility may be pressure-swing distillation. When pressure-swing distillation is not applicable because of an insensitivity of the azeotropic composition to pressure change, two common alternatives are azeotropic and extractive distillation (Widagdo and Seider, 1996). In both cases an additional component, the so-called entrainer is introduced to the original mixture to facilitate the separation (Kossack et al., 2008).

Extractive distillation can be more energy efficient than heterogeneous azeotropic distillation (Lei et al., 2003).

2.2.1. Extractive distillation process

In the conventional extractive distillation setup shown in Figure 1, the entrainer is the least volatile component in the system and is fed into the extractive column above the process feed with the azeotrope mixture. One of the azeotrope forming components is withdrawn at the top of the extractive column, while the other one, together with the entrainer, forms the bottoms product of the extractive column. In a second column, the entrainer is separated from the other component and recycled to the first column. Which component is carried with the entrainer depends on the actual mixture. The lightest component is produced at the top of the first column in the case shown in Figure 1.
However, an alternative way called reverse extractive distillation is outlined and studied by Hunek et al. (1989). In that case, the entrainer is the most volatile component and it interacts with the lighter azeotrop-forming component, carrying it to the top of the column. The structure also changes as it is shown in Figure 2: The heavier azeotrope forming component is withdrawn at the bottom of the extractive column, while the other, together with the entrainer, forms the top product of the extractive column. In the second column, the entrainer is separated from the second feed component and recycled to the first column.
2.2.2. Strange behavior of extractive distillation
Laroche et al. (1991) showed that some extractive processes displays strange behavior compared to conventional distillation systems.
For some mixtures, separation as a function of reflux goes through a maximum. At infinite reflux, no separation is achieved. Thus sometimes determining the minimum number of trays by examining the infinite reflux case is not possible, because we often obtain better results at finite reflux.
In some cases, achieving the same specifications with a larger number of trays requires a larger reflux.

2.2.3. Selection of the solvent
Solvent selection is the key factor in extractive distillation due to the high influence of the solvent recovery system and recycles in the global process (L. Jimenez et. al 2001). Good entrainer is a component which breaks the azeotropes easily (Laroche et. al. 1991). Three types of solvents (entrainers) have been used in extractive distillation:
- Heavy solvents (i.e., solvents which have a higher boiling point than both azeotropic components).
- Intermediate solvents (i.e., solvents which have boiling point located between those of the two azeotropic constituents).
- Light solvents (i.e., solvents which have a lower boiling point than both azeotropic components).
Extractive distillation utilizes a solvent to enhance the separation among chemical species. The entrainer flowrate has a direct influence not only on minimum reflux, but also on the maximum reflux bound. Therefore the solvent must be carefully chosen (J.Gentry et. al 2003, Z. Lelkes et al. 1998, Winkle 1967, Doherty and G. A. Caldarola 1985).
- Selectively removes the desired component
- Is easily separable from the extracted components
- It has to be available and inexpensive
- It has to be thermally stable, noncorrosive, nontoxic
• It must be have a low molar latent heat since it is to be vaporized

Kossack et al. (2008) developed a systematic framework for selecting the most suitable entrainer.

2.2.4. Simulation and optimization of extractive distillation

Distillation as industrial application is used for working up large volume; and it requires high investment and operation outlays. Therefore, the significance of the analysis of the process and design of economically optimal distillation processes is of no question. Simulators are often used to study the strange behavior of extractive distillation processes. Generally HYSYS is applied for rigorous simulation and sometimes optimization of extractive distillation. Munoz et al. (2006) compared the extractive distillation to pressure swing distillation, and concluded that plants with smaller capacity more efficiently work with pressure swing distillation, while for larger plants, the extractive distillation option will be more attractive.

Hilal et al. (2001) studied the possibility of reducing the specific quantity of the extractive agent (and consequently, the energy consumption) by changing the location of the feed stage(s).

Langston et al. (2005) investigated the effects of solvent feed entry stages, solvent split stream feed and solvent condition on the separation.

Research of optimal design of distillation processes has considerably long history. Due to improving the computational capability of modern computers, rigorous mathematical modeling have already become an alternative to traditional shortcut methods. Mathematical programming is a widespread approach in process synthesis. This method has three main steps. (1) First a superstructure containing all the considered alternatives, and its graph representation, are generated. (2) Then a mathematical representation is formulated, based on the graph representation of the superstructure. Existence of structure elements are modeled with binary variables in an algebraic model. Validity of some continuous relations (e.g. material balances) depends on such existence. The relations referring to the existence and coexistence of the structure elements are complicated expressions of mixed binary and continuous variables. These expressions can be formulated using well known techniques like Big-M or Convex Hull, see
Vecchietti et al. (2003). Finally, the extreme of the objective function is looked for, the mathematical model is optimized.

Multiplicity causes serious problems in MINLP models. Multiplicity means that several solutions of the mathematical representation define the same structure. It can also be said that a structure is represented by isomorphic graphs. In this case, the objective function has the same value in several different points. It makes more difficult finding the optimum, and the search space is unnecessarily wide.

There have been a lot of interesting and effective approaches in the literature for optimizing distillation processes in the recent years. These models simultaneously determine the optimal configuration of the column and the operational parameters of the distillation process. As the number of trays is also to be determined, binary variables are needed to handle their existence. Since nonlinear equations are needed to describe the equilibrium, the problem will result in an MINLP model.

Mathematical formulations that represent rigorous distillation column configuration fall into two categories: (1) one task–one equipment (OTOE) representations (Viswanathan and Grossmann, 1993) and (2) variable task–equipment (VTE) representations (Yeomans and Grossmann, 1999).

In the OTOE representation, each stage has one task to work as an equilibrium stage. If the actual stage does not work as an equilibrium stage, than it is by-passed (Viswanathan and Grossmann, 1993). Such a superstructure is shown in Figure 3. Binary variables \( z_{j}^{\text{ref}} \) and \( z_{j}^{\text{bu}} \) represent the location of the returning streams after condensation and boiling, i.e. the stages where the reflux stream (ref) and the reboiled vapor stream (bu) are led to, beside the feed location. Stages above the reflux stream location and below the reboiled vapor stream location are considered non-existing stages; whereas all the other stages exist.
Figure 3. Superstructure of Viswanathan-Grossmann

In the VTE representation, each stage has two tasks: 1) an equilibrium stage or 2) an input-output operation through the stage with no mass transfer. Binary variables assign which task of a stage is performed in a structure (Yeomans and Grossmann, 2000). The superstructure of this approach is depicted in Figure 4.

Figure 4. Superstructure of Yeomans-Grossmann

In both methods, the MINLP model involves the total and component mass balances, the enthalpy balances, and the equations of the physical-chemical equilibrium, at each stage.
Both method have subsequently been improved and applied to complex distillation systems, see Ciric and Gu (1994), Smith and Pantelides (1995), Bauer and Stilchmair (1998).

Farkas et al. (2008) developed a new superstructure and MINLP model. In order to reduce structural redundancy already during the phase of creating the superstructure, the equilibrium stages are contained by conditional units of different sizes as it is shown in Figure 5. These conditional units contain different numbers of equilibrium stages. If the binary variable of a unit equals one then the unit is included in the structure, all the equilibrium stages contained in that unit work, and their input and output streams may take positive value. If the binary variable of a unit equals zero then the unit is not included in the structure, none of the equilibrium stages contained in that unit works, and their input and output streams must equal zero. In this case, the liquid and vapor streams flow through transmitter units instead of the units containing equilibrium stages.

As a consequence, the model uses minimal number of binary variables, thereby reducing the size of the problem and thus the computation time.
Modeling and optimizing extractive distillation is more difficult problem that in the case of conventional distillation. The number of columns, the sum allowed number of stages, the number of components all increase, thus increasing the number of variables and constraints. The entrainer choice and the entrainer flowrate comprise additional degrees of freedom.

Because of the huge number of nonlinear equations and strong non-convexities, the NLP sub-problems may be practically unsolvable for real-scale problems without good initial
values. Therefore Farkas et al. (2008) modified the original Outer Approximation algorithm (Figure 6), in order to provide good, i.e. near-feasible, initial values for each variable in each of the iterations. The initial values were calculated from trays to trays using the vector of the binary variables provided by the preceding MILP master-problem. The scheme of the modified algorithm can be seen in Figure 7.

The above model using the modified OA algorithm was subsequently applied to a four-component-extractive distillation system designed to separate a three-component (methanol-)ethanol-water azeotropic system using methanol or ethylene-glycol as extractive agent. The optimal structure (which uses ethylene-glycol as solvent) is selected successfully.

Figure 6. Original Outer Approximation algorithm

Figure 7. Modified Outer Approximation algorithm
2.3. Mass Exchange Networks

Mass exchange networks (MEN) are systems of interconnected direct-contact mass transfer units using process lean stream and/or external separating agents (MSAs) to selectively remove certain components (often pollutants) from rich process streams. The purpose of MENs is to clean the departing process streams before they are released to the sewage system or into the air. By fully exploiting the cleaning capacities of the process streams that are available on site, mass exchange networks reduce both the amount of expensive external cleaning agents and the costs of the end-of-pipe wastewater treatment. Consequently, mass exchange networks serve direct pollution prevention goals.

2.3.1. Problem Statement of Mass Exchange Network

A good compendium of mass exchange networks and their applications can be found in the book of El-Halwagi (1997). He gives the following detailed definition for the MENS task:

Given a number $N_R$ of rich streams and a number $N_S$ of MSAs (lean streams); it is desired to synthesize a cost-effective network of mass exchangers that can preferentially transfer certain species from the rich streams to the MSAs. Given also are the flowrate of each rich stream, $G_i$, its supply (inlet) composition, $s_{iy}$, and its target (outlet) composition, $t_{iy}$, where $i = 1, 2, \ldots, N_R$.

In addition, the supply and target compositions, $s_{jx}$ and $t_{jx}$, are given for each MSA where $j = 1, 2, \ldots, N_S$. The mass transfer equilibrium relations are also given for each MSA. The flowrate of each MSA is unknown and is to be determined as part of the synthesis task. The candidate MSAs (lean streams) can be classified into $N_{SP}$ process MSAs and $N_{SE}$ external MSAs (where $N_{SP} + N_{SE} = N_S$). The process MSAs already exist on the plant site and can be used for a low cost (often virtually free). The flowrate of each process MSA, $L_j$, is bounded by its availability in the plant and may not exceed a value of $L^c_j$. On the other hand, the external MSAs can be purchased from the market and their flowrates are to be determined by economic considerations. Scheme of the definition of the MENS task is shown in Figure 8.
The objective of the synthesis problem is to determine: 1) which mass exchange operations should be used, 2) which MSA should be used and what should be their flowrate, 3) how should be the rich streams matched with the MSA-s and in what sequence, and finally 4) what should be the size of the mass exchanger units. The objective function to be minimized is usually the total annual cost ($TAC$), the sum of the investment and operating costs.

### 2.3.2. Mass Exchange Processes

The realm of mass exchange includes the following operations (El-Halwagi-1997):

- **Absorption**: a liquid solvent is used to remove certain compounds from a gas mixture, by virtue of their preferential solubility.

- **Adsorption** utilizes the ability of a solid adsorbent to adsorb specific components from a gaseous or a liquid solution onto its surface.

- **Liquid-Liquid Extraction**: it employs a liquid solvent to remove certain compounds from another liquid using the preferential solubility of these solutes in the MSA.

- **Ion-Exchange**: where the cation and/or anion resins are used to replace undesirable anionic species in liquid solutions with nonhazardous ions.

- **Leaching**, which is a selective solution of specific constituents of a solid mixture when brought in contact with a liquid solvent.

- **Stripping** is the transfer of gas, dissolved in a liquid into a gas stream.
2.3.3. Mass Exchangers
The mass exchange processes can be accomplished in different kinds of equipment. The most common are plate and packed columns.

**Plate columns:** Plate columns can be classified according to mode of flow in their internal contacting devices (Perry, 1984), as (1) Cross-flows plates. (2) Counter flow plates.

**Packed Columns:** Usually the columns are filled with structured packing in the most cases. The packed column is a simple device compared with the plate column. A typical column consists of a cylindrical shell containing a support plate for the packing material and a liquid distributing device designed to provide effective irrigation of the packing.

2.3.4. Mass Exchange Networks synthesis methods
The area of chemical process synthesis has received considerable attention over the past two decades (El-Halwagi and Manousiouthakis, 1989). Process synthesis is concerned with the systematic development of process flowsheets that transform available raw material into desired products and which meet specified performance criteria of maximum profit or minimum cost, energy efficiency, maximum raw material recovery, minimum waste production and good operability.

In most chemical industries, the synthesis of the processes can be done by one of the following approaches: 1) physical and thermodynamic insight which is called now Pinch technology, and 2) mathematical programming. The combination of them can also be used, which is called hybrid method.

2.3.4.1. Pinch Technology
Pinch technology is an insight-based approach and was first developed for the design of heat exchanger networks (HENS) in the late 1970’s (Hallale 1998). Pinch technology is based on two synthesis steps: a pinch analysis step and a flowsheet synthesis step. The pinch analysis step involves specifying a minimum energy transfer driving force to locate the operating conditions where there is constrained energy transfer and determine the minimum energy requirements ahead of any detailed design. In the flowsheet synthesis step, the conditions at the location where there is constrained energy transfer (the pinch point) are used to separate the energy streams into design regions.

The pinch technology has been extended by El-Halwagi and Manousiouthakis (1989) to handle mass exchange network synthesis (MENS).
At first, the rich and lean process streams are represented in a diagram which shows their in and outlet concentrations \((y)\) versus their mass load. Since in case of MENs the driving force of the mass transfer is the difference between the actual and equilibrium concentrations, the concept of corresponding concentration scales had to be introduced (Figure 9).

![Composite curves](image1)

**Figure 9. Composite curves**

Then the rich and the lean composite curves are constructed by combining the individual rich and lean stream mass loads in concentration intervals (Figure 10). The composite curves can be shifted along the mass axis. The point where the two composite curves approach each other by the minimum composition difference \((\varepsilon)\), is called the pinch point. The vertical overlap of the
composite curves shows where mass can be recovered from the rich streams to the lean ones (Figure 11).

Thus the pinch represents a bottleneck to mass recovery (El-Halwagi, 1997). The composite curve plot shows the excess capacity of the process MSA-s and the necessary capacity for external MSAs as well. From that, the operating cost target can be determined.

Lacking the tool for capital cost targeting, the synthesis method of El-Halwagi and Manousiouthakis (1989) did not enable supertargeting (targeting for $TAC$).

Hallale and Fraser (1998) extended the pinch design method by setting up capital cost targets ahead of any design. Using the $y$-$x$ composite curve plot (rich stream concentrations versus lean stream concentrations) and the $y$-$y^*$ composite curve plot, the overall number of equilibrium stages for the network can be estimated, similarly to the McCabe-Thiele method developed for designing distillation columns.

Hallale and Fraser (2000) presented a new method for targeting the minimum capital costs of mass exchange networks. The minimum operational and capital costs were estimated as functions of the minimum composition difference ($\varepsilon$). The development of capital cost targets means that the capital/operating cost trade-off can be explored ahead of any design. For a particular value of or $\varepsilon$, both the cost of MSAs (operating cost) and the capital cost can be predicted as targets. Both the operating and the capital costs are annualized and then added to yield a target for the $TAC$ of the network. This is repeated over a range of or $\varepsilon$ values, and the resulting minimum $TAC$ is selected as target, as it is shown in Figure 12.
2.3.4.2. Mathematical Programming

The mathematical programming approaches can be classified into two categories: sequential programming approaches and simultaneous programming approaches.

The first category is practically the automated version of the pinch technology. El-Halwagi and Manousiouthakis (1990) presented an automated synthesis procedure. This procedure first used linear programming to determine the pinch points and minimum utility targets. Mixed integer linear programming was then used to synthesize all possible networks featuring the minimum number of units. The completed networks were then costed and the one featuring the lowest cost was selected. This was carried out iteratively for a range of ε values in an attempt to minimize the annualized total cost of the network.

The attribute “sequential” denotes that the synthesis is still decomposed into targeting and design steps. As a consequence, the trade-off between investment and operating costs is not taken into account rigorously; therefore, it can provide a possibly suboptimal solution only.

Simultaneous programming approaches, on the other hand, keep the integrity of the problem and consider the possibility of trade-off between investment and operating costs properly, see Papalexandri et al. (1994).

As in mass exchange network synthesis, binary variables are necessary to represent the existence of mass exchanger units, and nonlinear constraints describe mass balances, the mathematical programming is used in the form of mixed-integer nonlinear programs (MINLP).

The design of mass exchange networks using MINLP approach is performed through three main steps:

1) Development of a representation of alternatives (i.e. the superstructure) from which the optimum solution is selected.
2) Formulation of a mathematical program that generally involves discrete and continuous variables for the selection of the configuration and operating levels, respectively.

3) Solving of the optimization model from which the optimal solution is determined.

2.3.4.2.1. Superstructure
The superstructure is a representation of alternatives from among the optimum solution is selected. There are different superstructures have been proposed in the MINLP approach. The first for handling mass exchange networks was developed by Papalexandri et. al. (1994).

The MEN superstructure by Szitkai et al. (2006) keeps most of the properties of the original stagewise HEN superstructure of Yee and Grossmann (1990). It consists of several serially connected identical stages \((k)\) where streams are driven in a counter-current way. The number of the superstructure stages is arbitrary but has to be large enough to be able to embed the solution. The sum of the number of the lean and the rich streams is usually a good choice for the maximum number of stages. Any of the rich \((R)\) and lean streams \((L)\) can be matched once in each of the stages. The stream matching in the stages are formed in a way that streams are split towards the possible mass exchangers as they enter the stages. After leaving the exchangers, the streams are mixed again. The main assumption of the model is that in the superstructure only mixing of streams with equal concentrations is defined. Two rich streams, two lean streams and two stages are shown in Figure 13.
Isafiade and Fraser (2008) developed an interval based MINLP superstructure (IBMS) which harnesses the strengths of the stagewise superstructure and the pinch technology methods. In the IBMS approach, the superstructure intervals are defined by the supply and target compositions of either the rich or lean set of streams. The superstructure composition interval defining approach introduced in this paper enforces the mixing of split streams at equal compositions; hence there is no need to include mixing equations in the model. Fixing the interval boundaries helps to eliminate the complexities involved in initializations, thus the region of search for the optimum solution is reduced. Two rich streams, two lean streams and three intervals are shown in Figure 14.
2.3.4.2.2. MINLP Formulation

There are two main difficulties in MENS using MINLP: (1) the non-linearity of the model equation system, and (2) the structural multiplicity of the superstructure. The first problem was addressed by Szitkai et al. (2006). His MINLP model is formulated as to include the following constraints:

- Total mass exchanged by a stream.
- Mass exchanged by a stream in an interval.
- Mass exchanged at connection ijk.
- Big-M equations for the mass exchanged at ijk.
- Concentration monotony.
- Concentration is less than the target.
- Concentration at the beginning is the source composition.
- Calculation of composition differences.
- Relaxed binary variable is equal to binary variable.
- Constraints for the number of units.
- Objective function.
The main characteristics of the model is that in the superstructure only mixing of streams with equal concentrations is allowed. This allows streams to only be characterized by concentrations at the stage borders (concentration locations). No separate in- and outlet concentrations for the exchangers are introduced. There are no individual stream flow rate variables; flow rates of the mass exchangers can be back-calculated after obtaining the solution. Mass balances are set up only for the stages and not for the mass exchangers. This allows an almost linear formulation, which has better numerical characteristics.

Similar technique was used by Isafiade and Fraser (2008) for reducing nonlinearity: their interval-based superstructure enforces the isocomposition mixing of split streams at every interval where splits takes place by the fixed composition interval boundaries. This helps to simplify the model since there is no need to include non-linear mass balance and mixing equations in order to determine optimal split flows.

2.3.4.3. The Hybrid Method

Hybrid synthesis method for mass exchange network was first presented by Msiza and Fraser (2003), as depicted in Figure 15.

The components making up the hybrid tool are a total annual cost target of the expected flowsheet, a physically meaningful initial flowsheet, a driving force (DF) diagram and an MINLP model. The total annual cost target represents the best cost scenario for the flowsheet, an initial flowsheet is used to initialize the MINLP solver and the driving force plot assists the designer to generate alternative initial solutions to improve the generated MINLP solution. The hybrid tool produces mass exchange networks whose total annual costs are within 10% of previously reported pinch solutions and networks that are similar or better than those obtained from the MINLP approach alone.

However, this method has three main disadvantages: (1) Generating an initial flowsheet in each iteration is a very time-consuming step. (2) Generation of a new initial flowsheet is not automatic; it needs human interactivity. (3) Only the initial flowsheet is changed during the iterations, the MINLP model does not.
Figure 15. Scheme of the hybrid method of Msiza and Fraser (2003)
2.4. Desalination processes
Water is the primary basis for life on earth and without it we would not exist here today. It is an extremely plentiful compound, yet a majority of it is unusable in its current form. 99% of the total surface and ground waters above the ground are either salt laden or locked up as ice in the polar region (Younos 2005).

Water scarcity is now a common occurrence in all countries. Water is a unitary resource and its scarcity is related to many factors such as; 1) Population growth is directly affecting availability of water resources. 2) Degraded water quality and pollution of surface and groundwater sources. The application of desalting technologies over the past 50 years has changed this in many places. Villages, cities, and industries have now developed or grown in many of the arid and water-short areas of the world where sea or brackish waters are available and have been treated with desalting techniques (Buros 1990).

A desalting device essentially separates saline water into two streams: one with a low concentration of dissolved salts (the fresh water stream) and the other containing the remaining dissolved salts (the concentrate or brine stream) (Buros 1990). The device requires energy to operate and can use a number of different technologies for the separation.

2.4.1. Desalination Technologies
The desalination technologies that are used around the world can be divided into two categories, as follows:

- **Major processes**
  - **Thermal Technologies**
    - Multistage Flash Distillation (MSF)
    - Multi Effect Distillation (MED)
    - Vapor Compression Distillation (VC)
  - **Membrane Technologies**
    - Electrodialysis (ED)
    - Reverse Osmosis (RO)

- **Minor Technologies**
  - Freezing
  - Membrane Distillation
  - Solar Distillation
In this work two types of desalination processes were selected: Reverse Osmosis (RO) and Solar Distillation (SD). RO desalination plants are available in larger sizes and produce more water today than they did a few years ago because of increased productivity (an increase of 94% from 1990 to 2002). The rapid growth in the application of RO process is due to its ability to produce fresh water at lower cost. Another advantage of the RO process is that it is able to meet varying feed water concentration and varying production water quantity and quality requirement through change of system construction and operation condition (Lu et al. 2007).

Solar energy is cost effective in terms of fuel (because no fuel is required) and its price would not be affected by the supply and demand of fuels. Solar energy is also pollution free. Technology for solar desalination has been studied for a long time as a process friendly to the environment and also energy saving (Toyama and Murase 2004).

**2.4.1.1. Reverse Osmosis (RO)**

Osmosis is a natural phenomenon by which water from a low salt concentration passes into a more concentrated solution through a semi-permeable membrane. When pressure is applied to the solution with the higher salt concentration solution, the water will flow in a reverse direction through the semi-permeable membrane, leaving the salt behind.

In the RO process shown in Figure 16, the feed is pretreated first to remove any solid could be found. Then pressurized by a high-pressure pump and is made to flow across a semi-permeable membrane. The feed pressure should exceed the osmotic pressure of the salted water in order for the separation to take place. Typical pressures for seawater range from 50 to 80 bars. Water passes through the membrane, and the rejected salt emerges from the membrane molecules as a concentrated reject stream, still at high pressure. In large plants the reject brine pressure energy is recovered in a turbine.

![Figure 16. Scheme of Reverse Osmosis (RO) process](image-url)
2.4.1.2. Solar Distillation

Desalination by means of solar energy becomes more competitive, especially for remote and rural areas, where small quantities of water for human consumption are needed.

Solar distillation is the simplest desalination process and is based on the green-house effect. Glass and other transparent materials have the property of transmitting incident short wave solar radiation but do not transmit infrared radiation. Incident short wave solar radiation passes through the glass into the still where it is trapped and evaporates the water, which is condensed on the glass surface and collected as distillate, as shown in Figure 17.

![Figure 17. Scheme of the Solar Distillation (SD) process](image)

2.4.2. Desalination in Libya

Libya is a mostly arid and semiarid, sparsely populated, large North African country. Over 80% of Libya's population reside along a mild thin strip on its 1900 km long Mediterranean coast which also contains the country's most fertile lands and its major industrial projects (Abufayed and El-Ghuel 2001, Bindra and Abosh 2001).

Large increase in water demand with very little recharge has strained Libya's groundwater resources, especially along the Mediterranean coast. Both thermal and membrane desalination technologies have extensively been used. Elhassadi’s (2008) study has illustrated Libya’s need of a well defined plan to include desalination to develop a Sea Water Desalination River.

2.4.3. Desalination Cost

The cost of desalination is decreasing in recent times due to the developments in desalination technologies. However, it is still essential to select an appropriate desalination technology that
produces desalinated water at a low cost for any site under consideration (Reddy and Ghaffour, 2007).

The unit cost of desalination is a function of many factors such as plant capacity, feed water quality, desalination process and technology, energy cost, plant life, and amortization of capital investment. In general, the unit cost of desalted water is inversely proportional to the plant size, i.e. if all other factors remain the same; the cost of product water would be lower for bigger plants (Ettouney et al., 2002).

The water needs to be transported from the plant to the consumers, mainly by pipelines. Long distance pipelines attract huge investment costs, and it makes a large proportion of the total desalination cost. It is therefore necessary to optimize the existence and length of pipelines transporting the water.

Minimizing the total desalination cost therefore involves finding not only the optimal number and type, but also the location of desalination plants.

2.4.4. Desalination Plant Location

The problem of optimally locating desalination plants onto the plane in order to minimize their distance measured from fixed points (seawater intakes and cities) can be mathematically formulated as a (facility) location problem.

Facility location problems are frequently studied in the continuous and discrete optimization literature. The essential elements of such a problem are customers with specified demands and locations, and facilities that are to be located on a planar ground. The goal is to find the optimal position of the facilities such that the objective function, generally a cost function depending on the sum of the distances between the facilities and the customers, is minimized. The area on which the facilities will be sited can be restricted. The (in)feasible region can be convex or nonconvex, see Nickel (1998) and Özyurt and Realtf (1999). A useful survey related to location problems is available by ReVelle and Eiselt (2005).

The facility location problems are usually solved with heuristics. The problem is NP-hard, hence exact algorithms can solve problems of moderate size only, and they rarely occur in the literature, see Boland et al. (2006) and Wu et al. (2006).

The problem studied in this thesis later on, that is, finding the optimal location of desalination plants and assigning the connection from seawater intakes to plants and from plants to cities, has not been addressed in the literature so far.

As it is detailed later in Section 6.1, the seawater is transported by pipelines from water intakes to plants and, the desalinated water is also transported by pipelines from plants to customers. In
our problem, the minimization of the investment cost arising from the pipelines is the main issue; that constitutes the main difficulty. The pipeline is needed only if it transports water. If there is no water transportation between two particular objects, then there is no need for the pipeline between them to exist. In other words, our problem is a location–allocation problem, where logic (or binary) variables are needed to represent the existence of connections between objects. If the pipeline exists, then the investment cost—calculated from its length, which is a variable—will be considered in the objective function. If the pipeline does not exist then its investment cost is 0. No similar problems possessing all of these features are found in the literature. Some location problems existing in the literature (Capacitated Facility Location Problem or CFLP, Klose and Drexl (2005)) do not capture the substance of our problem, the fixed connections between the facilities and the customers. They are only trade-off problems between the variable and fixed cost of the production, and they do not consider the variable distances. Most of the previous works, see Özyurt and Realf (1999) and Klose and Drexl (2005), deal with problems to locate facilities in the plane so that the sum of the distances measured between the facilities and the customers is at minimum. The main difference between these and our problem is that we consider the pipelining cost in that case only when the connection between the facility and customer really exists. The objective function of the literature problems is the sum of all the distances. It follows that no binary variables to represent the existence of distinct pipelines are needed in their models. They can be solved with linear programming. Some other location problems, see Boland (2006) and Klose and Drexl (2005), are more similar to ours. These models already consider the distinct connections, but the possible locations of the facilities are discrete and known, so the distances between the facilities and the customers are also known in advance as parameters. The problem most reminiscent to ours is the theoretical multi-Weber problem where the task is to partition the full set of fixed points into subsets and finding several minimizing points (one per set) simultaneously. This problem is mainly solved by heuristics. However, Rosing (1991) showed that the objective function of the problem is very steep near the optimum. Therefore, it is essential to apply an exact solution method; otherwise, the results would be very far from the optimum. The only exact method is worked out by Rosing (1992), but his problem and model are still very different from ours and the technique applied by him is not suitable to solve our problem because of the aforementioned differences. Additionally, his solution method is based on the possibility of encircling the (customers’) locations by disjunctive convex hulls each of which includes the
(plant’s) location serving the points in the actual hull only. Since the plant capacities are finite, and there are forbidden zones, the optimal solution may, and in practice will, contain pipeline connections from plants to customers in such a way that the above convex hull assumption is not satisfied.

Since no similar problems and methods applicable to our problem were found in the literature, a new model had to be worked out to solve the problem. This new model and the obtained results are detailed in the present work. Besides mixed-integer linear programming (MILP), it is also possible to apply other techniques like constraint logic programming (CLP) and generalized disjunctive programming (GDP), see Yeomans and Grossmann (1999).
3. Scope and challenges of the thesis
The present thesis deals with different applications of mathematical programs in chemical industry, all of which are important from the point of view of Libya’s development.

3.1. Analysis of Extractive Distillation with Mathematical Programming
The first topic handles analysis and optimization of extractive distillation processes which, among others, have close relationships to the extensive oil refinery industry. Studying these distillation processes is important for deeply understanding them and successfully designing economically optimal distillation systems. Such a research is often performed using process simulators but application of optimization tools is more convenient. When the problem is to find the minimum reflux ratio or the economically optimal reflux ratio for a fixed configuration and product composition, the user of the simulator software has to make many trials to determine the optimal solution. The same results can be achieved in one step using NLP model, without making a lot of useless attempts. Similarly, the optimal structure can be calculated in one step with some iteration cycles by solving an MINLP problem.

In Section 4 an analysis method based on applying an optimization tool is presented for complex distillation processes. The analysis method is used for deeply understanding the studied processes through exploring their operational and design limits, and collecting preliminary data for optimal design and analysis from economical point of view. The feasible region of the process is explored in order to find good bounds and initial estimates of the design and operation parameters.

Using the initial data, MINLP is applied for determining the optimal structure and operational parameters in one step, in order to study the effect of cost factors on the optimal configuration and analyze the process from economical point of view. Extractive distillation (in general) is found to be a stable process in the sense that the optimal structure is rather insensitive to the cost factors.

Two different types of extractive distillation systems, namely acetone/methanol azeotropic mixture with water as heavy entrainer, and ethanol/water azeotropic mixture with methanol as light entrainer, are used as examples to demonstrate the application of the method.
3.2. New hybrid method for mass exchange network optimization

The second topic deals with the development of a new hybrid method for synthesizing mass exchange networks. These synthesis applications, although yet unknown and unapplied in Libya, will have attracted immense attention in few years in the extensively developing oil industry.

As it is pointed out in the literature review, the hybrid method of Msiza and Fraser (2003) has three main disadvantages: (1) Generating an initial flowsheet in each iteration is a very time-consuming step. (2) Generation of a new initial flowsheet is not automatic; it needs human interactivity. (3) Only the initial flowsheet is changed during the iterations, the MINLP model is not. My target was to develop such a systematic method for mass exchange network synthesis that addresses these issues and minimizes the need for human interaction.

In Section 5, a new hybrid optimization method is presented, that is fairly robust and can be accomplished in an automatic way. The main idea is using integer cuts and bounds, based on driving force plot analysis, for the lean streams to decrease the possibility of being trapped in local optima. A new initial solution is constructed if the MINLP solution is infeasible; otherwise the earlier found best solution is used. In consequence, the MINLP model is modified in the iteration steps. The iteration is stopped when the total annual cost in the solution reaches the 110% of its target.

3.3. Optimization of desalination location problem using MILP

As it is pointed out in Section 2.4 of the literature review, the need to outline a plan for supplying desert settlements with potable water through desalination has already been recognized.

In Section 6, a new mixed-integer linear programming (MILP) model for the location problem is presented in order to find the optimal co-ordinates of the desalination plants. The model takes into account the given locations and capacities of the water incomes, the demands, and the costs of plants and pipelining. Feasible and infeasible plant regions are distinguished for locating the plants.

The model has been developed in two consecutive phases. First a basic model is developed that provides a solution within short time but does not take into account the possibility of pipeline branching. Application of this model gives rise to redundant pipelines to some connections, involving extra costs. Pipeline branching is dealt with by an improved model developed in the second phase. This improved model provides realistic solution but with much longer computation time.

The results of applying the different models on motivated examples of different sizes are detailed.
4. Analysis of Extractive Distillation with Mathematical Programming

As Section 2.2 of the literature review shows, analysis of extractive distillation processes from both operational and economical points of view is, although possible but, inconvenient using process simulators. It implies a lot of necessary trial and error experiments; see e.g. the iterative method given by Chadda, Malone, and Doherty (2000).

Determining the minimum and maximum values of both the stage numbers and reflux ratio can be difficult, and can be more easily accomplished by optimization. Minimum and maximum reflux ratios for a certain number of stages can be found in one step using NLP. Also the minimum and maximum stage numbers can be found by systematic search in less time and with less trouble than with use of simulators.

Therefore in this section I present a method which can be used in an automated way for studying these processes. The method is based on the use of a previously elaborated MINLP model.

4.1. Proposed method

The main objective of this chapter is to present an optimization-based method for studying complex distillation processes from economical (financial) point of view. The optimization is carried out using Mixed-Integer Nonlinear Programming. Due to the size and complexity of the problem, this requires a good (feasible) initial configuration and tight bounds for the number of stages. Therefore, preparatory to solving the MINLP model, the first step of the analysis is exploring the feasible region. After completing the analysis, the next phase is the optimization itself, where the optimal configuration is determined at different cost-factors combinations.

The feasible region of the extractive column is determined by four parameters jointly: the three column section stage numbers (rectifying, extractive, and stripping sections), and the reflux ratio. This implies first to find the minimum number of stages in each column section with which the separation may be feasible. These minima can be used as lower bounds in subsequent optimization steps. The relationship between actual stage numbers in the column sections and minimum and maximum reflux ratios constitute the boundary of the feasible region. These data can be used for providing tight bounds and good, near-optimal, initial values for the reflux ratio in subsequent optimization steps. Additionally, they can be utilized to understand the behavior of these processes from chemical engineering point of view.
Errors in estimating possible future cost factors can affect the efficiency of optimization. If the price of the energy is estimated inaccurately, the calculated optimum may be far from the real future optimum. How the different cost-factors will affect the optimal configuration is investigated with MINLP in the second, final, step using the previously determined bounds and initial data.

As the effect of cost factors is investigated through solving mathematical programming problems, it seems to be reasonable to apply mathematical programming also in the phase of exploration of the feasible region, in the preparatory first step. Therefore, instead of using the usual process simulators, the minimal number of stages and minimal and maximal reflux ratios are also determined through optimization.

### 4.1.1. Model and software

The MINLP model developed by Farkas et al. (2008) is used for analyzing extractive distillation. The characteristics of that model have been summarized in the literature review.

The superstructures of the extractive distillation processes detailed below are built up using the concept of that model. As the number of nonlinear equations increase with the numbers of components, columns, and stages, it becomes more and more difficult to solve the model with the optimization tools. In order to facilitate the solution, a modification of the original Outer Approximation algorithm is suggested and applied by Farkas et al. (2008). By modifying the Outer Approximation algorithm in this way, i.e. inserting a step calculating initial values, it has become possible to reliably solve the model and to use the method as an analyzing tool for studying complex distillation processes. The step of calculating initial values is also performed when only NLP problems are solved, namely when the minimum reflux ratio is computed for fixed stage numbers.

### 4.1.2. Exploration of the feasible region

For exploring the feasible region, each section of the extractive column is allowed to have maximum 31 trays, and a 3-dimensional grid covering the $[0, 31] \times [0, 31] \times [0, 31]$ domain of the stage number combinations is created first. The number of the grid points is determined such a way, that it can cover the whole domain and a detailed diagram showing the relationship between the reflux ratios and number of stages can be plotted. When determining the borders of the feasible region, the section stage numbers are fixed as parameters, the binary variables are eliminated from the model, and NLP problems are solved. Instead of developing partial models
for determining the feasible region of the extractive column, the model of the full process can be applied with fixed stage numbers and fixed reflux ratio in the conventional column.

The resulting NLP problems are solved in AIMSS 3.7, using CONOPT 3.14G as NLP solver.

After finding the minimum stage number in each section, the relationship between the sections stage numbers (as independent variables) and minimum and maximum reflux ratios (as dependent variables) is explored. These five data altogether determine the feasible region. How these data depend on each other, i.e. interaction between sections, is also studied.

4.1.3. Studying the effect of cost factors on the optimal configuration

As a result of the above detailed study, good initial values and bounds are available that can be utilized to investigate the effect of changing cost factors. Here the total process is optimized, i.e. the operation parameters (reflux ratios, reboiler duties, etc) as well as the stage numbers in the columns are considered unknown, and thus none of the binary variables are fixed. This implies that an MINLP is solved because the model contains both binary and continuous variables and nonlinear equations. Unlike the exploration of the feasible region, investigation of the effect of cost factors cannot be done using process simulators like ASPEN. As the optimum structure needs to be determined at different parameter settings, this step requires the use of optimization tools, and solving MINLP-s.

Solving large MINLP problems with huge number of nonlinear (and nonconvex) constraints may require good bounds and initial values. The results of the preparatory step exploring the feasible region can provide us with tight bounds for both the continuous and the discrete variables. The 3-dimensional grid described in the previous section results in a number of NLP problems, some of which are infeasible because of the insufficient number of stages. These infeasibilities show the lower bound of the number of stages of each column section. The minimum of the reflux ratios of all the feasible runs provides the lower bound; the maximum provides the upper bound of the reflux ratio. Those tight bounds help the solver to find the solution over a reduced search space.

Guaranteeing tight bounds described in the previous section might have been carried out using engineering criteria or heuristics. Using the same optimization tool for both tasks, however, is an automated and more convenient way.

The cost function given by Luyben and Floudas (1994) for calculating the total annual cost of distillation columns is used as objective function to be minimized. The form of the cost function is the following:

\[
TotalCost = btax \cdot C_s \cdot Q_R + btax \cdot C_v \cdot Q_c + UF \cdot \frac{f(N,D)}{bpay}
\]  

(4.1)
where $C_s$ is unit price of steam, $Q_R$ is reboiler duty, $C_w$ is unit price of cooling water, $Q_C$ is condenser duty, $UF$ is update factor of the price of column installation, $bpay$ is payback period, $N$ is number of stages, and $D$ is column diameter. The four cost factors ($C_s$, $C_w$, $UF$, $bpay$) are chosen to investigate their effect. A steady increase of the objective function with increasing $C_s$, $C_w$, $UF$ and decreasing $bpay$ is predictable from its mathematical form. Our goal here is to explore which factors have significant effect on the total annual cost and the optimal structure, and how the optimal structure changes with these cost factors.

An orthonormal matrix containing the set values of the cost factors is designed for this aim, based on the composite design used for multiple regression in the statistical literature, see Box, Hunter and Hunter (2005).

Two basic, reasonably estimated, levels (two different values) are set to each cost factor. The process is optimized at each possible combination of these values; this involves $2^4$ experiments. In order to determine the nonlinear parts in the regression polynomial, an additional experiment is made at the center point of the design and at excessive points (star-points) outside the core of the design, as is visualized in Figure 18.

The orthonormal matrix is presented in Table 1. The columns of the table represent the four cost factors; the “$-1$” and “$+1$” signs denote the lower and upper levels of the factors; “$0$” marks the centre point; “$-\sqrt{2}$” and “$+\sqrt{2}$” are used for the star-points.
Table 1. The orthonormal matrix

<table>
<thead>
<tr>
<th>Test</th>
<th>bpay</th>
<th>C_s</th>
<th>C_w</th>
<th>UF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>+1</td>
</tr>
<tr>
<td>3</td>
<td>-1</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
</tr>
<tr>
<td>4</td>
<td>-1</td>
<td>-1</td>
<td>+1</td>
<td>+1</td>
</tr>
<tr>
<td>5</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>6</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
</tr>
<tr>
<td>7</td>
<td>-1</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
</tr>
<tr>
<td>8</td>
<td>-1</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
</tr>
<tr>
<td>9</td>
<td>+1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>10</td>
<td>+1</td>
<td>-1</td>
<td>-1</td>
<td>+1</td>
</tr>
<tr>
<td>11</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
</tr>
<tr>
<td>12</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>+1</td>
</tr>
<tr>
<td>13</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>14</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
</tr>
<tr>
<td>15</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
</tr>
<tr>
<td>16</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
</tr>
<tr>
<td>17</td>
<td>( \sqrt{2} )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>( +\sqrt{2} )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>-( \sqrt{2} )</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>( +\sqrt{2} )</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>0</td>
<td>-( \sqrt{2} )</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>0</td>
<td>( +\sqrt{2} )</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-( \sqrt{2} )</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>( +\sqrt{2} )</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The cost factor values at the two levels and in the center point are shown in Table 2.

Table 2. Values of factors at levels

<table>
<thead>
<tr>
<th></th>
<th>(-\sqrt{2})</th>
<th>- center</th>
<th>+</th>
<th>+(\sqrt{2})</th>
</tr>
</thead>
<tbody>
<tr>
<td>bpay</td>
<td>2.7874</td>
<td>4</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>( C_s )</td>
<td>1.6645-10(^{-6})</td>
<td>1.7667-10(^{-6})</td>
<td>2.01335-10(^{-6})</td>
<td>2.26-10(^{-6})</td>
</tr>
<tr>
<td>( C_w )</td>
<td>4.2867-10(^{-8})</td>
<td>4.3923-10(^{-8})</td>
<td>4.6472-10(^{-8})</td>
<td>4.902-10(^{-8})</td>
</tr>
<tr>
<td>UF</td>
<td>1.4715</td>
<td>1.5657</td>
<td>1.7934</td>
<td>2.021</td>
</tr>
</tbody>
</table>

Table 1 assigns 25 MINLP problems. These are solved in AIMMS 3.7 with modified Outer Approximation algorithm. CPLEX 10.0 is applied as MILP solver; CONOPT 3.14A is used for NLP subproblems. TotalCost calculated according to Equation (4.1) is used as objective function. The value of TotalCost, the sum of the stage numbers (\( \sum N \)), and the total reboiler duty (\( \sum Q_b \)), are the recorded measurement values. After determining the three measurement values at
each design point, multivariate nonlinear regression is applied to calculate the parameters of the following polynomial:

\[ \hat{Y}_i = b'_0 + \sum_j b_j x_{ji} + \sum_{j<k} b_{jk} x_{ji} x_{ki} + \sum_j b_{jj} (x_{ji}^2 - x_{ji}^2) \]  \hspace{1cm} (4.3)

where \( \hat{Y}_i \) is the regressed actual objective function \((Total Cost, \Sigma Q_r, \Sigma N)\). \( b'_0 \) is the intersection; \( b_j \) is the coefficient of \( j^{th} \) factor; \( b_{jk} \) is the coefficient of the interaction of \( j^{th} \) and \( k^{th} \) factors; \( b_{jj} \) is the coefficient of the squared \( j^{th} \) factor; \( x_{ji} \) is the value of \( j^{th} \) factor at \( i^{th} \) experiment.

Once the coefficients are computed, their effect can be analyzed. We are interested in the question which factors have significant effect, and what kind of shift in the optimal configuration is expected.

### 4.2. Examples and results

The two systems, used for studying the method, are (1) separation of acetone from methanol using water as heavy solvent, and (2) separation of ethanol from water using methanol as light solvent. In both cases, the product purity requirements are highly above the azeotropic composition, so that using extractive distillation is necessary.

#### 4.2.1. Acetone/methanol mixture

##### 4.2.1.1. General specifications

The process scheme is shown in Figure 19. 100 kmol/h acetone/methanol mixture containing 50 mol% acetone is fed to the extractive column. Water is fed to the column above the main feed. Pure acetone is produced at the top. Methanol/water mixture leaves the column at the bottom and is separated in the solvent recovery column. The purity requirement is 98 mol% for each product.

The vapor phase has been considered ideal, and Wilson equations are used for modeling the activity coefficients in the liquid phase, see the applied Wilson-coefficients in Section 11.1. The column diameters are calculated from the vapor flowrate using the \( F_f \)-procedure, see Kister (1992), i.e. the diameters are considered as free variables, both in the preliminary exploration of the feasible region and in the analysis of the effect of the cost factors. The maximum allowed stage numbers is set to 31 in each section of the extractive column.

When exploring the feasible region in the preparatory step, the water feed flow rate is set to 100 kmol/h. In the conventional column used for separating the methanol from water, the stage
numbers are set to 7 in both sections and the reflux ratio is fixed at 5. The results of exploring the feasible region are detailed in the next section.

4.2.1.2. Exploration of the feasible region

4.2.1.2.1. Minimum stage numbers in the extractive column

The minimum stage number means the number of stages minimally required to reach the specified product purity. This value in each section depends on the actual stage numbers of the other sections. The minima are determined by systematically fixing the stage numbers in two column sections and finding in the actual section the fewest stages with which the separation is still feasible. In Table 3 to Table 5, only a few configurations are shown but they represent the behavior of the minimum stage numbers in function of the stage numbers of the other sections. $N_R$, $N_E$, and $N_S$, are stage numbers of the rectifying section, extractive section, and stripping section, respectively. These numbers do not count the feed stages themselves.

As it is shown in Table 3, the increase of the stage number of stripping section has no significant effect on the minimum stage number of the rectifying section; the extractive section has a considerable effect.

Figure 19. Process scheme of acetone-methanol separation
Table 3. Minimum number of stages in the rectifying section at different stage numbers in the other two sections with acetone/methanol mixture

<table>
<thead>
<tr>
<th>N_S</th>
<th>N_E</th>
<th>N_R,min</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>31</td>
<td>10</td>
</tr>
<tr>
<td>31</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>31</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
<td>25</td>
</tr>
</tbody>
</table>

As is shown in Table 4, the range of feasible stage numbers of the extractive section is considerably smaller than that of the other sections. The specified level of separation requires many stages in this section. The situation is similar to the previous one. A change in the length of the rectifying section has a greater effect than that of stripping section on the minimum stage number of the extractive section.

Table 4. Minimum number of stages in the extractive section at different stage numbers in the other two sections with acetone/methanol mixture

<table>
<thead>
<tr>
<th>N_S</th>
<th>N_R</th>
<th>N_E,min</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>31</td>
<td>15</td>
</tr>
<tr>
<td>31</td>
<td>11</td>
<td>24</td>
</tr>
<tr>
<td>6</td>
<td>31</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>25</td>
</tr>
</tbody>
</table>

The minimum number of stages is the smallest in the stripping section, compared to the others, as it is presented in Table 5. The interactions have smaller effect than in the case of rectifying and extractive section. In other words, the minimum stage number does not change strongly with the number of stages in the other sections.
Table 5. Minimum number of stages in the stripping section at different stage numbers in the other two sections with acetone/methanol mixture

<table>
<thead>
<tr>
<th>$N_E$</th>
<th>$N_R$</th>
<th>$N_{S,min}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>31</td>
<td>2</td>
</tr>
<tr>
<td>31</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>16</td>
<td>31</td>
<td>6</td>
</tr>
<tr>
<td>16</td>
<td>21</td>
<td>9</td>
</tr>
</tbody>
</table>

As a summary, a large number of stages are needed in both the rectifying and extractive sections; and only few stages are necessary in the stripping section. There is considerable interaction between the rectifying and the extractive sections; when one of them contains fewer trays, the other must include more. The stripping section has only a slight effect on the others. The minimum stage numbers are 10, 15, and 2 in the rectifying, extractive, and stripping sections, respectively.

4.2.1.2.2. Limiting flows
Minimum and maximum reflux ratios are computed with NLP at fixed feasible stage number combinations determined in the previous section. Results are collected and shown in Figure 20 to Figure 22. The envelope of the plots constitutes a borderline of the feasible region. Each figure plots the minimum and maximum reflux ratios against varied stage numbers in a particular column section at two characteristic extreme stage number combinations of the other two sections. The envelope of the plots constitutes a borderline of the feasible region. The figures have similar shapes. The gap between the upper and the lower limits becomes smaller toward smaller stage numbers, and the increase of the gap in increasing stage number is smaller at higher stage numbers.
Figure 20. Relationship between minimum and maximum reflux ratios and number of stages in the rectifying section with acetone/methanol mixture

Figure 21. Relationship between minimum and maximum reflux ratios and number of stages in the extractive section with acetone/methanol mixture
Figure 22. Relationship between minimum and maximum reflux ratios and number of trays in the stripping section with acetone/methanol mixture

The width of the feasible reflux ratio interval strongly depends on the number of stages in the rectifying and the extractive sections, but has an almost constant plateau at higher stage numbers of the stripping section.

4.2.1.3. The effect of cost factors on the optimal configuration

For studying the effect of the cost factors, the water feed flow rate is enabled to vary between 0 and 100 kmol/h, the stage numbers in each column section of the extractive column may vary between its minimum (obtained from the results presented in Section 4.2.1.2.1) and 31, and between 0 and 15 in the solvent recovery column. The reflux ratio is also considered variable, between its smallest minimum and highest maximum ($R_{min}=2$ and $R_{max}=10$, respectively), presented in Section 4.2.1.2.2, as bounds in the extractive column. The lower and upper bounds of the reflux ratio in the solvent recovery column are 0.5 and 20, respectively.

For obtaining a good initial estimate, first the total annual cost is calculated to the results presented in Section 4.2.1.2.2. Since the minimum reflux ratio belongs to high stage numbers, three cost curves are calculated and plotted in Figure 23. In each curve the stage number in one column section is varied whereas each other column section contains 31 stages. An optimal stage number belonging to minimum cost is read from each curve: 18, 20, and 7 stages in the rectifying, extractive, and stripping sections, respectively. These are merely local optima belonging to fixed (maximum allowed) stage numbers in other sections. When the stage numbers
are decreasing in the other sections, the optimum number of stages increases in the actual section currently studied. Therefore, the actual optimal stage numbers, when all of them are variable, may be higher or lower because of the expected interactions. Combination of the locally optimal stage numbers determined independently of each other may even be infeasible. This is just the case in this example, too. Thus, the initial stage number configuration is set to 31, 15, and 7 stages in the rectifying, extractive, and stripping sections, respectively, because this configuration is the closest feasible one to the simple but infeasible combination of the locally optimal stage numbers. This initial configuration is applied in each of the twenty-five runs. The initial reflux ratio in the extractive column must be inside the feasible region; $R=5$ is selected arbitrarily.

Figure 23. TotalCost as function of stage numbers section by section

The total annual cost values resulted by solving the 25 MINLP problems, total number of stages in both columns ($\Sigma N$), and total reboiler duty ($\Sigma Q_r$), at each point of the orthonormal matrix are collected in Table 6. Figure 24 shows the optimal structure belonging to the 25th MINLP problem.
Table 6. Run results of the experimental design for acetone/methanol system

<table>
<thead>
<tr>
<th>Test</th>
<th>bpa</th>
<th>C_s</th>
<th>C_w</th>
<th>UF</th>
<th>TotalCost</th>
<th>ΣQR</th>
<th>ΣN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>249.823</td>
<td>12.04</td>
<td>63</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>281.8288</td>
<td>12.044</td>
<td>63</td>
</tr>
<tr>
<td>3</td>
<td>-1</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
<td>212.0692</td>
<td>9.613</td>
<td>70</td>
</tr>
<tr>
<td>4</td>
<td>-1</td>
<td>-1</td>
<td>+1</td>
<td>+1</td>
<td>277.4572</td>
<td>10.793</td>
<td>69</td>
</tr>
<tr>
<td>5</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
<td>-1</td>
<td>253.6883</td>
<td>12.036</td>
<td>63</td>
</tr>
<tr>
<td>6</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>285.3449</td>
<td>11.066</td>
<td>64</td>
</tr>
<tr>
<td>7</td>
<td>-1</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
<td>238.3446</td>
<td>10.933</td>
<td>64</td>
</tr>
<tr>
<td>8</td>
<td>-1</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>295.5067</td>
<td>10.723</td>
<td>70</td>
</tr>
<tr>
<td>9</td>
<td>+1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>129.9131</td>
<td>10.665</td>
<td>71</td>
</tr>
<tr>
<td>10</td>
<td>+1</td>
<td>-1</td>
<td>-1</td>
<td>+1</td>
<td>155.1512</td>
<td>10.672</td>
<td>79</td>
</tr>
<tr>
<td>11</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
<td>130.0188</td>
<td>10.721</td>
<td>70</td>
</tr>
<tr>
<td>12</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>+1</td>
<td>152.6864</td>
<td>11.596</td>
<td>64</td>
</tr>
<tr>
<td>13</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
<td>-1</td>
<td>151.5152</td>
<td>10.663</td>
<td>79</td>
</tr>
<tr>
<td>14</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>172.2935</td>
<td>10.786</td>
<td>77</td>
</tr>
<tr>
<td>15</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
<td>148.3437</td>
<td>10.886</td>
<td>69</td>
</tr>
<tr>
<td>16</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>166.6504</td>
<td>10.79</td>
<td>69</td>
</tr>
<tr>
<td>17</td>
<td>+√2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>346.767</td>
<td>10.668</td>
<td>71</td>
</tr>
<tr>
<td>18</td>
<td>+√2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>147.346</td>
<td>10.201</td>
<td>63</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>-√2</td>
<td>0</td>
<td>0</td>
<td>162.036</td>
<td>9.202</td>
<td>81</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>+√2</td>
<td>0</td>
<td>0</td>
<td>171.702</td>
<td>8.835</td>
<td>74</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>0</td>
<td>-√2</td>
<td>0</td>
<td>175.067</td>
<td>10.845</td>
<td>67</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>0</td>
<td>+√2</td>
<td>0</td>
<td>188.346</td>
<td>10.791</td>
<td>78</td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-√2</td>
<td>165.273</td>
<td>11.596</td>
<td>64</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+√2</td>
<td>200.174</td>
<td>10.791</td>
<td>69</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>187.691</td>
<td>12.032</td>
<td>63</td>
</tr>
</tbody>
</table>

Figure 24. Optimal structure belonging to the 25th MINLP problem

Coefficients of Equation (4.3) are computed to each measured pointer functions (TotalCost, ΣN and ΣQR) with multivariate nonlinear regression, and are collected in Table 7 - Table 12.
Table 7. Coefficients of the cost factors to TotalCost

<table>
<thead>
<tr>
<th>$b'_0 = 201.802$</th>
<th>$bpay$</th>
<th>$C_s$</th>
<th>$C_w$</th>
<th>$UF$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_j$</td>
<td>-58.476</td>
<td>6.821</td>
<td>-1.983</td>
<td>16.128</td>
</tr>
<tr>
<td>$b_{jj}$</td>
<td>31.603</td>
<td>-8.491</td>
<td>-1.069</td>
<td>-0.564</td>
</tr>
</tbody>
</table>

Table 8. Coefficients of interactions to TotalCost

<table>
<thead>
<tr>
<th>$bpay$</th>
<th>$C_s$</th>
<th>$C_w$</th>
<th>$UF$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$bpay$</td>
<td>-</td>
<td>1.208</td>
<td>2.258</td>
</tr>
<tr>
<td>$C_s$</td>
<td>-</td>
<td>-</td>
<td>1.906</td>
</tr>
<tr>
<td>$C_w$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$UF$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

As it could be predicted, the small price of cooling water has a coefficient smaller by one order of magnitude than the others to TotalCost. The payback period has a strong negative effect: The total annual cost decreases with increasing payback period because the installation costs are broken down to more years. The price of the steam and the cost of column installation have a positive effect.

Table 9. Coefficients of the cost factors to $\sum N$

<table>
<thead>
<tr>
<th>$b'_0 = 69.36$</th>
<th>$bpay$</th>
<th>$C_s$</th>
<th>$C_w$</th>
<th>$UF$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_j$</td>
<td>2.0343</td>
<td>-0.195</td>
<td>0.0778</td>
<td>0.654</td>
</tr>
<tr>
<td>$b_{jj}$</td>
<td>-1.9</td>
<td>3.35</td>
<td>0.85</td>
<td>-2.15</td>
</tr>
</tbody>
</table>

Table 10. Coefficients of interactions to $\sum N$

<table>
<thead>
<tr>
<th>$bpay$</th>
<th>$C_s$</th>
<th>$C_w$</th>
<th>$UF$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$bpay$</td>
<td>-</td>
<td>0.875</td>
<td>-3.375</td>
</tr>
<tr>
<td>$C_s$</td>
<td>-</td>
<td>-</td>
<td>-0.5</td>
</tr>
<tr>
<td>$C_w$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$UF$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 11. Coefficients of the cost factors to $\Sigma Q_R$

<table>
<thead>
<tr>
<th></th>
<th>$b'_0$</th>
<th>$b_{ij}$</th>
<th>$b_{jj}$</th>
<th>$b_{ijj}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10.91</td>
<td>-0.0272</td>
<td>-0.039</td>
<td>-0.205</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.0113</td>
<td>-0.0113</td>
<td>-0.0113</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.403</td>
<td>-0.762</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.326</td>
</tr>
</tbody>
</table>

Table 12. Coefficients of interactions to $\Sigma Q_R$

<table>
<thead>
<tr>
<th></th>
<th>$b_{pay}$</th>
<th>$b_{Cs}$</th>
<th>$b_{Cw}$</th>
<th>$b_{UF}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_{pay}$</td>
<td>-</td>
<td>-0.049</td>
<td>0.396</td>
<td>0.0566</td>
</tr>
<tr>
<td>$Cs$</td>
<td>-</td>
<td>-</td>
<td>0.0924</td>
<td>-0.2012</td>
</tr>
<tr>
<td>$Cw$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.162</td>
</tr>
<tr>
<td>$UF$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 25. $TotalCost$ as function of $b_{pay}$ and $Cs$

The regressed functions of $TotalCost$ are shown in Figure 25 - Figure 27. $b_{pay}$ has the largest (negative) effect; its coefficient is $-58.5$. $UF$ and $Cs$ have smaller (positive) effects; their coefficients are 16.13 and 6.82, respectively. These results are not surprising since $b_{pay}$ varies in a wider interval than the other two parameters. $Cw$ has no significant effect; its coefficient is smaller than the others by an order of magnitude. This is due to the fact that the price of the cooling water is negligible compared to the others.
Figure 26. TotalCost as function of bpay and UF

Figure 27. TotalCost as function of Cs and UF
The regressed functions of $\Sigma N$ and $\Sigma Q_R$ are visualized in Figure 28 - Figure 33. Note that these values vary in a very narrow interval. The relative widths of these intervals are smaller than 5%, i.e. the cost-factors have no considerable effect on the number of stages and reboiler duty.

Figure 34 - Figure 42 show the effects of cost factors on the stage numbers in the three column sections. The extractive and stripping section lengths are insensitive to the cost factors, whereas the rectifying section has some sensitivity, although the latter may still be due to random variation because the difference of 2-3 in the stage numbers may hardly be significant as it causes only 2 to 4 % difference in the total annual cost, imperceptible by MINLP. This insensitivity is due to the fact that the minimum reflux ratio does not depend on the stage numbers in the stripping and extractive section, and, therefore, the total annual cost increases more steeply with increasing stage numbers in these sections than in the rectifying section.
In summary, the optimal structure is widely independent of the cost factors. It follows that (1) the investor has no reason to be afraid of investing in extractive distillation plants, and (2) an error in the estimation of the future prices will not affect the future optimality of the designed structure.
4.2.2. Ethanol/water mixture

4.2.2.1. General specifications
The process scheme is shown in Figure 43. 100 kmol/h ethanol/water mixture containing 80 mol% ethanol is fed to the extractive column. Methanol is fed to the column below the main feed. Pure water is produced at the bottom. Methanol/ethanol mixture leaves the column at the top and is separated in the solvent recovery column. The purity requirement is 98 mol% for each product. The vapor phase has been considered ideal, and Wilson equations are used for modeling the activity coefficients in the liquid phase, see the applied Wilson-coefficients in Section 11.1. The column diameters are calculated from the vapor flowrate using the \( F_f \)-procedure, see Kister (1992), i.e. the diameters are considered as free variables, both in the preliminary exploration of the feasible region and in the analysis of the effect of the cost factors.

When exploring the feasible region in the preparatory step, the methanol feed flow rate is fixed at 300 kmol/h. In the extractive column, the maximum allowed stage numbers is set to 31 in each section. In the conventional column used for separating the methanol from ethanol, the stage numbers are set to 15 in both sections and the reflux ratio is fixed at 5. The results of exploring the feasible region are detailed in the next section.

![Figure 43. Ethanol-water separation applying methanol solvent](image)

4.2.2.2. Exploration of the feasible region

4.2.2.2.1. Minimum stage numbers in the extractive column
The minimum number of stages in the rectifying section strongly depends on the other two stage numbers, as is clear from Table 13, and can be approximated as a function of the total number of stages in the other two sections. According to Table 14, the minimum number of stages in the
extractive section also strongly depends on the other two stage numbers. The same applies to Table 15 and stage numbers of the stripping section.

Table 13. Minimum number of stages in the rectifying section at different stage numbers in the other two sections with ethanol/water mixture

<table>
<thead>
<tr>
<th>$N_S$</th>
<th>$N_E$</th>
<th>$N_{R,\text{min}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>31</td>
<td>9</td>
</tr>
<tr>
<td>31</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>0</td>
<td>31</td>
<td>11</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>25</td>
</tr>
</tbody>
</table>

Table 14. Minimum number of stages in the extractive section at different stage numbers in the other two sections with ethanol/water mixture

<table>
<thead>
<tr>
<th>$N_S$</th>
<th>$N_R$</th>
<th>$N_{E,\text{min}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>31</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>11</td>
<td>19</td>
</tr>
</tbody>
</table>

Table 15. Minimum number of stages in the stripping section at different stage numbers in the other two sections with ethanol/water mixture

<table>
<thead>
<tr>
<th>$N_E$</th>
<th>$N_R$</th>
<th>$N_{S,\text{min}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>31</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>11</td>
<td>13</td>
</tr>
</tbody>
</table>

It can be concluded that there is strong interaction between sections. The minimum number of stages of a certain section depends on the stage numbers in the other sections. The most stages are needed in the rectifying section.
4.2.2.2. Limiting flows

Minimum and maximum reflux ratios are computed with NLP at fixed feasible stage number combinations determined in the previous section. The results are collected and shown in Figure 44 to Figure 46.

It is worth noting that while there is an interaction between sections having effect on the number of stages, there is no effect on the minimum reflux ratio in Figure 44.

![Figure 44. Relationship between the minimum and maximum reflux ratios and number of stages in the rectifying section with ethanol/water mixture](image)

The situation in the extractive section (Figure 45) is different from the rectifying one. Neither the minimum, nor the maximum reflux ratio does significantly depend on the number of extractive stages when the other two sections have maximum stage numbers, but there is a considerable tightening when the stage numbers approach the minimum. This applies to the stripping section, as well (Figure 46).
Figure 45. Relationship between the minimum and maximum reflux ratios and number of stages in the extractive section with ethanol/water mixture

Figure 46. Relationship between minimum and maximum reflux ratios and number of stages in the stripping section with ethanol/water mixture

The most stages are needed in the rectifying section. The minimum reflux ratio does not decrease with increasing stage numbers in the extractive and stripping sections if they are above their minimums and, therefore, there is no need and it is not economical to use more stages than the minimum. However, there is a trade-off between the reflux ratio and the number of stages in the rectifying section. Here the optimal configuration is far above the minimum number of stages.
Consequently, the optimal stage numbers in the stripping and extractive sections are the minimum values whereas it is much higher in the rectifying section.

4.2.2.3. The effect of cost factors on the optimal configuration
For studying the effect of the cost factors, the methanol feed flow rate is enabled to vary between 0 and 300 kmol/h, the stage numbers in each column section of the extractive column may vary between its minimum (obtained from the preparatory step’s results presented in Section 4.2.2.2.1) and 31, and between 0 and 15 in the solvent recovery column. The reflux ratio is also considered variable, between its smallest minimum and highest maximum ($R=1$ and $R=10$, respectively), presented in Section 4.2.2.2, as bounds in the extractive column. The lower and upper bounds of the reflux ratio in the solvent recovery column are 0.5 and 20, respectively.

For obtaining a good initial estimate, first the total annual cost is calculated to the results presented in Section 4.2.2.2. Since the minimum reflux ratio belongs to high stage numbers, three cost curves are calculated and plotted in Figure 47. In each curve the stage number in one column section is varied whereas each other column section contains 31 stages.

An optimal stage number belonging to minimum cost is read from each curve: 21, 0, and 0 stages in the rectifying, extractive, and stripping sections, respectively. These are merely local optima belonging to fixed (maximum allowed) stage numbers in other sections. When the stage numbers are decreasing in the other sections, the optimum number of stages increases in the actual section currently studied. Therefore, the actual optimal stage numbers, when all of them are variable, may be higher or lower because of the expected interactions. Combination of the locally optimal stage numbers determined independently of each other may even be infeasible. This is just the case in this example, too. Thus, the initial stage number configuration is set to 15, 3, and 3 stages in the rectifying, extractive, and stripping sections, respectively, because this configuration is the closest feasible one to the simple but infeasible combination of the locally optimal stage numbers. This initial configuration is applied in each of the twenty-five run. The initial reflux ratio in the extractive column must be inside the feasible region; $R=5$ is selected arbitrarily.
The resulted total annual cost, total number of stages in both columns ($\sum N$), and total reboiler duty ($\sum Q_R$), at each point of the orthonormal matrix are collected in Table 16. Figure 48 shows the optimal structure belonging to the 25th MINLP problem.
Table 16. Run results of the experimental design for ethanol/water system

<table>
<thead>
<tr>
<th>Test</th>
<th>bpa</th>
<th>y</th>
<th>C</th>
<th>w UF</th>
<th>TotalCost</th>
<th>ΣQR</th>
<th>ΣN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>836.757</td>
<td>74.533</td>
<td>60</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>-1</td>
<td>+1</td>
<td>+1</td>
<td>954.395</td>
<td>77.368</td>
<td>56</td>
</tr>
<tr>
<td>3</td>
<td>-1</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
<td>892.274</td>
<td>83.592</td>
<td>56</td>
</tr>
<tr>
<td>4</td>
<td>-1</td>
<td>-1</td>
<td>+1</td>
<td>+1</td>
<td>953.452</td>
<td>74.534</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
<td>-1</td>
<td>921.158</td>
<td>69.87</td>
<td>63</td>
</tr>
<tr>
<td>6</td>
<td>-1</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>1072.075</td>
<td>74.533</td>
<td>60</td>
</tr>
<tr>
<td>7</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>931.871</td>
<td>71.534</td>
<td>61</td>
</tr>
<tr>
<td>8</td>
<td>-1</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>1080.1</td>
<td>77.367</td>
<td>56</td>
</tr>
<tr>
<td>9</td>
<td>+1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>610.887</td>
<td>77.367</td>
<td>56</td>
</tr>
<tr>
<td>10</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>+1</td>
<td>655.759</td>
<td>77.367</td>
<td>56</td>
</tr>
<tr>
<td>11</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
<td>599.894</td>
<td>74.529</td>
<td>60</td>
</tr>
<tr>
<td>12</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>+1</td>
<td>646.08</td>
<td>74.53</td>
<td>60</td>
</tr>
<tr>
<td>13</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
<td>-1</td>
<td>735.305</td>
<td>77.367</td>
<td>56</td>
</tr>
<tr>
<td>14</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>764.6671</td>
<td>74.52938</td>
<td>60</td>
</tr>
<tr>
<td>15</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
<td>695.818</td>
<td>71.532</td>
<td>61</td>
</tr>
<tr>
<td>16</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>765.905</td>
<td>74.5297</td>
<td>60</td>
</tr>
<tr>
<td>17</td>
<td>+√2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1126.072</td>
<td>74.196</td>
<td>57</td>
</tr>
<tr>
<td>18</td>
<td>+√2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>640.662</td>
<td>71.532</td>
<td>61</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>-√2</td>
<td>0</td>
<td>0</td>
<td>656.491</td>
<td>71.532</td>
<td>61</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>+√2</td>
<td>0</td>
<td>0</td>
<td>844.953</td>
<td>74.529</td>
<td>60</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>0</td>
<td>+√2</td>
<td>0</td>
<td>759.3304</td>
<td>74.5302</td>
<td>60</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>0</td>
<td>+√2</td>
<td>0</td>
<td>731.206</td>
<td>69.042</td>
<td>66</td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-√2</td>
<td>666.047</td>
<td>67.962</td>
<td>63</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+√2</td>
<td>806.825</td>
<td>74.531</td>
<td>60</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>737.834</td>
<td>71.533</td>
<td>61</td>
</tr>
</tbody>
</table>

Figure 48. Optimal structure belonging to the 25th MINLP problem

Coefficients of Equation (4.3) are computed to each measured pointer functions (TotalCost, ΣN and ΣQR) with multivariate nonlinear regression, and are collected in Table 17 - Table 22.
Table 17. Coefficients of the cost factors to $TotalCost$

<table>
<thead>
<tr>
<th></th>
<th>$b'_0 = 803.43$</th>
<th>$b_{ij}$</th>
<th>$b_{jj}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_{ij}$</td>
<td>-142.7</td>
<td>54.196</td>
<td>-1.269</td>
</tr>
<tr>
<td>$b_{jj}$</td>
<td>72.651</td>
<td>6.3289</td>
<td>3.602</td>
</tr>
</tbody>
</table>

Table 18. Coefficients of interactions to $TotalCost$

<table>
<thead>
<tr>
<th></th>
<th>$bpay$</th>
<th>$C_s$</th>
<th>$C_w$</th>
<th>$UF$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$bpay$</td>
<td>-</td>
<td>5.0468</td>
<td>-8.265</td>
<td>-17.97</td>
</tr>
<tr>
<td>$C_s$</td>
<td>-</td>
<td>-</td>
<td>-3.338</td>
<td>8.0451</td>
</tr>
<tr>
<td>$C_w$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-1.069</td>
</tr>
<tr>
<td>$UF$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

As it could be predicted, the price of cooling water has a coefficient smaller by one order of magnitude than the others to $TotalCost$. The payback period has a negative effect: The total annual cost decreases with increasing payback period because the installation costs are broken down to more years. The price of the steam and the cost of column installation have a positive effect.

The regressed $TotalCost$ function is visualized in Figure 49 - Figure 51. $C_w$ and $C_s$ are fixed at their zero levels in Figure 49; $bpay$ and $UF$ are independent variables. $bpay$ has the more significant effect, and there is no considerable interaction between the two factors.
$C_w$ and $UF$ factors are fixed at their zero levels in Figure 50; $bpay$ and $C_s$ are independent variables. $bpay$ has the more significant effect, and there is no considerable interaction between the two factors.
$C_w$ and $bpay$ factors are fixed at their zero levels in Figure 51; $UF$ and $C_s$ are independent variables. The steam price has larger effect than the cost of the column installation, probably because the heating costs usually reach 60-70% of the total annual costs of distillation processes. There is no significant interaction of the two factors.

![Figure 51. TotalCost as function of UF and Cs](image)

Table 19. Coefficients of the cost factors to $\sum N$

<table>
<thead>
<tr>
<th></th>
<th>$bpay$</th>
<th>$C_s$</th>
<th>$C_w$</th>
<th>$UF$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_0'$</td>
<td>59.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b_j$</td>
<td>0.1328</td>
<td>0.5793</td>
<td>0.7743</td>
<td>-0.4621</td>
</tr>
<tr>
<td>$b_{jj}$</td>
<td>-1.875</td>
<td>-1.125</td>
<td>0.125</td>
<td>-0.625</td>
</tr>
</tbody>
</table>

Table 20. Coefficients of interactions to $\sum N$

<table>
<thead>
<tr>
<th></th>
<th>$bpay$</th>
<th>$C_s$</th>
<th>$C_w$</th>
<th>$UF$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$bpay$</td>
<td>-</td>
<td>-0.1875</td>
<td>1.1875</td>
<td>0.6875</td>
</tr>
<tr>
<td>$C_s$</td>
<td>-</td>
<td>-</td>
<td>-0.5625</td>
<td>-0.3125</td>
</tr>
<tr>
<td>$C_w$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.0625</td>
</tr>
<tr>
<td>$UF$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 21. Coefficients of the cost factors to $\Sigma Q_R$

<table>
<thead>
<tr>
<th></th>
<th>$b'_0 = 74.176$</th>
<th>$b_{pay}$</th>
<th>$C_s$</th>
<th>$C_w$</th>
<th>$UF$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_j$</td>
<td>-0.2674</td>
<td>-0.916</td>
<td>-0.4274</td>
<td>0.6904</td>
<td></td>
</tr>
<tr>
<td>$b_{jj}$</td>
<td>1.6263</td>
<td>1.7096</td>
<td>1.0874</td>
<td>0.8026</td>
<td></td>
</tr>
</tbody>
</table>

Table 22. Coefficients of interactions to $\Sigma Q_R$

<table>
<thead>
<tr>
<th></th>
<th>$b_{pay}$</th>
<th>$C_s$</th>
<th>$C_w$</th>
<th>$UF$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_{pay}$</td>
<td>-</td>
<td>0.6805</td>
<td>-1.3896</td>
<td>-0.257</td>
</tr>
<tr>
<td>$C_s$</td>
<td>-</td>
<td>-</td>
<td>-0.1179</td>
<td>1.0549</td>
</tr>
<tr>
<td>$C_w$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.3054</td>
</tr>
<tr>
<td>$UF$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Regarding the effect of cost factors on the stage numbers and reboiler duty, notice that the cost factors have much smaller effect on them than on the TotalCost. As discussed previously, there is no trade-off between the reflux ratio and the number of stages. It follows that the minimum of total costs of stripping and extractive sections is, independently of cost factors, situated at a point in or near the border of feasibility. Some trade-off can be experienced in the case of rectifying section but a large contribution to the objective function is constant and independent of the number of stages in rectifying section and, therefore, this effect can be neglected. The regressed function for the total number of stages is shown in Figure 52 - Figure 57.
It is clear from the figures that the optimal configuration is independent of the cost factors, at least within a considerable interval. The total number of stages is always between 58 and 61, a 5% variation, practically imperceptible by MINLP. The optimal number of stages in rectifying section varies between 21 and 23, and between 4 and 7 in the extractive section, still negligible variations compared to the size of the system.

One could experience some dependence between cost-factors and the optimal configuration if the change in the cost factors would be larger, for example 100 to 200 %. However, applying reasonable values at different levels of the factors, the optimal configuration does not depend on them.

This phenomenon has significant consequences. The economical risk is relatively small when investing into extractive distillation plants because the optimal configuration remains optimal or near-optimal even in changing economic environment. Although it seems to be valid generally for extractive distillation, its general validity is worth to analyze further.
Figure 58 to Figure 66 show the effect of the cost factors on the optimal number of stages in different sections.

<table>
<thead>
<tr>
<th>Rectifying section</th>
<th>$b_{pay}$-$UF$</th>
<th>$C_s$-$UF$</th>
<th>$b_{pay}$-$C_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 58. $N_R$ as function of $b_{pay}$ and $UF$</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>Figure 59. $N_R$ as function of $C_s$ and $UF$</td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
<tr>
<td>Figure 60. $N_R$ as function of $b_{pay}$ and $C_s$</td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Extractive section</th>
<th>$b_{pay}$-$UF$</th>
<th>$C_s$-$UF$</th>
<th>$b_{pay}$-$C_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 61. $N_E$ as function of $b_{pay}$ and $UF$</td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
<tr>
<td>Figure 62. $N_E$ as function of $C_s$ and $UF$</td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
</tr>
<tr>
<td>Figure 63. $N_E$ as function of $b_{pay}$ and $C_s$</td>
<td><img src="image16.png" alt="Image" /></td>
<td><img src="image17.png" alt="Image" /></td>
<td><img src="image18.png" alt="Image" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stripping section</th>
<th>$b_{pay}$-$UF$</th>
<th>$C_s$-$UF$</th>
<th>$b_{pay}$-$C_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 64. $N_S$ as function of $b_{pay}$ and $UF$</td>
<td><img src="image19.png" alt="Image" /></td>
<td><img src="image20.png" alt="Image" /></td>
<td><img src="image21.png" alt="Image" /></td>
</tr>
<tr>
<td>Figure 65. $N_S$ as function of $C_s$ and $UF$</td>
<td><img src="image22.png" alt="Image" /></td>
<td><img src="image23.png" alt="Image" /></td>
<td><img src="image24.png" alt="Image" /></td>
</tr>
<tr>
<td>Figure 66. $N_S$ as function of $b_{pay}$ and $C_s$</td>
<td><img src="image25.png" alt="Image" /></td>
<td><img src="image26.png" alt="Image" /></td>
<td><img src="image27.png" alt="Image" /></td>
</tr>
</tbody>
</table>
4.3. Conclusion

A method is proposed to accurately analyze complex distillation processes with (MI)NLP model as optimization tool. Application of this tool is demonstrated on two extractive distillation processes: one with heavy entrainer, and another with light entrainer. First the feasible region is explored in order to find the lower bounds of stage numbers in the column sections. Then the minimum and maximum reflux ratios are calculated at different stage number combinations.

Due to nonconvexity of the model, multiple local optima in the parameters space may exist. It may be especially important in cases not studied here. In case of heterogeneous systems, it is worthwhile to restart the searching process from several different initial points in order to find the global optimum.

The minimum and maximum reflux ratio values are used to generate near optimal initial data for designing optimal configuration and operation parameters in one step with MINLP. Using the initial configuration determined in the previous step, the effect of the cost factors on the optimal configuration is studied.

The optimal structures are found widely independent from the weights of different cost parts in the objective function. It follows that some inaccuracy in estimating the cost-factors will not cause serious problem in determining the economically optimal configuration. In turn, small changes in the economic environment will not alter the optimal configuration.

The methodology presented here is quite general and can be applied to other complex distillation processes, as well.
5. New hybrid method for mass exchange network optimization

5.1. Motivation for developing a new method

There are two main difficulties in MENS using MINLP: (1) the non-linearity of the model equation system, and (2) the structural multiplicity of the superstructure.

(1) Even using a fairly linear model based on pinch insight superstructure (Szitkai et al., 2006), some of the equations in the model (e.g. the calculation of logarithmic mean concentration differences) remain nonlinear. Therefore, finding the global optimum is not guaranteed, and the found optimum strongly depends on the initial point.

(2) In each hitherto known superstructures formulated for MENS, there is great structural multiplicity. Structural multiplicity means that several solutions of the mathematical model represent the same structure. As a consequence, the search space is unnecessarily large, and the solution time is considerably increased.

The above mentioned difficulties in MENS using MINLP are demonstrated by solving example 3.2 of Hallale (1998), using the fairly linear MINLP model of Szitkai et al. (2006). Sulphur dioxide (SO₂) is to be removed from four gaseous process streams, by absorption into water which is an external MSA. Stream and equilibrium data are given in Table 35 of the Appendix. All exchangers are packed columns, according to the specifications. Equipment and capital cost data for the columns are given in Table 36 of the Appendix.

The column diameters (D) and cross-sectional areas (S) are previously calculated, based on the velocity of the gaseous streams (here the rich streams). These data are given in advance (before the design step), and are collected in Table 23.

<table>
<thead>
<tr>
<th></th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
</tr>
</thead>
<tbody>
<tr>
<td>D (m)</td>
<td>0.66</td>
<td>0.72</td>
<td>0.59</td>
<td>0.51</td>
</tr>
<tr>
<td>S (m²)</td>
<td>0.342</td>
<td>0.407</td>
<td>0.273</td>
<td>0.204</td>
</tr>
</tbody>
</table>

The calculated total annual cost target is \( TAC_{target} = 642\ 000 \text{ $/yr} \) by Hallale (1998).

The initial structure is generated in a way that the only external lean stream is split into the same number of branches as the number of rich streams, and each branch removes all the sulphur-
dioxide from the rich stream it is matched with in one step. Finally, the four branches are mixed. The flowrates of the lean stream branches are calculated with the assumption of the MINLP model of Szitkai et al. (2006), i.e. only streams with the same concentration can be mixed. The initial structure and the initial driving force plot containing the composite operating line are shown in Figure 67 and Figure 68, respectively. In Figure 68 the composite operating line is dashed, and the operating lines of the particular units in the initial structure are shown by solid lines numbered according to the units.

![Figure 67. The initial structure of example 3.2 of Hallale (1998)](image)

![Figure 68. Initial driving force for example 3.2 of Hallale (1998)](image)
The problem has been solved three times, consecutively. The same initial structure was used each time, but the earlier solutions were excluded by integer cut. The general formula of the integer cuts is taken from Floudas (1995):

$$\sum_{i \in B'} z_i' - \sum_{i \in NB'} z_i' \leq |B'| - 1$$  \hspace{1cm} f \in F \hspace{1cm} (5.1)$$

The objective function values are shown in Table IV; the optimal structures found are shown in Figure 69 - Figure 71.

The TAC results are very close to each other, and all were less than 1.1*TAC$_{target}$. From these results one can conclude that there are a lot of local optima, and the global optimum is not guaranteed.

<table>
<thead>
<tr>
<th>Runs</th>
<th>TAC $/year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>653 770</td>
</tr>
<tr>
<td>2</td>
<td>654 643</td>
</tr>
<tr>
<td>3</td>
<td>660 644</td>
</tr>
</tbody>
</table>

Table 24. Values of the objective function obtained for example 3.2 of Hallale (1998)

Figure 69. Optimal structure in the 1st run
5.2. New hybrid method algorithm

The main idea of the new hybrid method is using integer cuts and bounds to the lean streams in order to exclude the local optima found earlier. The MINLP model developed by Sztikai et al. (2006) is applied because its superstructure is based on pinch insight and, therefore, this model is more suitable for a hybrid method than that of Papalexandri et al. (1994).

MSA regeneration is not taken into account in the algorithm developed for mass exchange synthesis problems. The method can be extended to consider MSA regeneration by modifying the MINLP model.

The algorithm of the new hybrid method is shown in Figure 72.
The steps of the new hybrid method are set out as follows:

Step 1: First, the total annual cost target $TAC_{\text{target}}$ is determined with the supertargeting method of Hallale and Fraser (1998). 110% of this value is used as the stop criterion in the iterative process. Driving force plot (DFP) to the minimum external lean stream flow rates is drawn, and an initial flowsheet is determined. Only one of the external lean streams is used in the initial flowsheet. The pollutant from each rich stream is removed in one stage, where the selected external lean stream is splitted into the same number of branches as the number of the rich streams. The flowrates of the branches are calculated with the assumption used in the MINLP model of Szitkai et al. (2006), i.e. only streams with the same concentration may be mixed.

Step 2: The flowsheet is optimized using the MINLP technique of Szitkai et al. (2006). If the solution is feasible and the $TAC$ is smaller then $1.1*TAC_{\text{target}}$ then this solution is accepted, and the iteration stops. If the solution is feasible but the $TAC$ does not reach $1.1*TAC_{\text{target}}$, then the MINLP model is extended in Step 3. If the solution is infeasible, then a new initial structure is generated in Step 4.

Step 3: If the solution was feasible in the previous iteration but the $TAC$ was not smaller than $1.1*TAC_{\text{target}}$, then the MINLP model is modified by performing the following actions:
(a) The former binary solution is excluded from the MINLP representation by an integer cut (equation 5.1).
(b) The best solution found earlier is used as initial flowsheet.
(c) A constraint that \( TAC \) has to be smaller than the best earlier found \( TAC \) is inserted.

(d) A bound is given to a lean stream flowrate. In order to determine whether an upper or a lower bound is necessary, the driving force plot is drawn for the current solution. If the average driving force of the units is higher than the composite operating line which has been drawn in Step 1 then an upper bound, otherwise a lower bound is used.

Choice of the bounded lean stream is based on the connection(s) which is (are) above or below the composite operating line, according to whether an upper or a lower bound is necessary. If those connections whose lines are far from the composite operating line belong to the same lean stream, then that stream has to be bounded. Otherwise, an efficiency factor of the referred lean streams has to be calculated, and the lean stream with the highest efficiency factor has to be bounded.

The efficiency factor for upper bounding can be calculated as

\[
E_{j}^{up} = \frac{A_{j}}{ME} \quad j \in J
\]

and for lower bounding as

\[
E_{j}^{lo} = \frac{(B_{j} - A_{j})}{ME} \quad j \in J
\]

where \( A_{j} \) is current mass exchange by lean stream \( j \), calculated with equation (5.4); \( B_{j} \) is the maximum mass exchange that can be done by lean stream \( j \), calculated with equation (5.5); and \( ME \) is the total mass exchange by the whole system, calculated with equation (5.6).

\[
A_{j} = L_{j}(x_{j}^{out} - x_{j}^{t}) \quad j \in J
\]

\[
B_{j} = L_{j}^{up}(x_{j}^{t} - x_{j}^{r}) \quad j \in J
\]

\[
ME = \sum_{i} \left[ G_{i}(y_{i}^{r} - y_{i}^{t}) \right]
\]

Once the lean stream to be bounded has been chosen, then such an upper/lower bound is taken that is in the range 105%/95% of the actual flowrate of the stream.

Step 4: In the case of an infeasible solution, a new feasible initial structure has to be generated. The MINLP model is modified by introducing relaxed binary variables which will be forced by equation (5.7) to take binary values. By equation (5.7), the integer-infeasible path MINLP (IIP-MINLP) model formulation of Sorsak and Kravanja (2002) is introduced. The equation states that the relaxed version of the binary variable \( r_{i,j,k} \) must be equal to the real binary variable \( z_{i,j,k} \) within a positive \( p_{i,j,k} \) and a negative \( n_{i,j,k} \) tolerance. Tolerances can take non-negative values only.

\[
r_{i,j,k} = z_{i,j,k} + p_{i,j,k} - n_{i,j,k} \quad i \in I, j \in J, k \in K
\]
The objective function is modified at the same time equation (5.8), and the target is to find the minimum of the sum of the tolerances:

\[
OBJ = \sum_{i} \sum_{j} \sum_{k} \left( p_{i,j,k} + n_{i,j,k} \right)
\]  

(5.8)

In the case of a solution with zero objective value, a feasible solution is found. If in this solution the TAC is smaller than 1.1*TAC\text{target} then this solution is accepted, and the iteration stops. If the TAC is not smaller than 1.1*TAC\text{target}, then the feasible solution found can be used as an initial structure in Step 3.

In case of solution of Step 4 with objective value greater than zero, feasible solution cannot be found. In this case, feasibility of the problem is considered. If at least one feasible solution, which does not fulfill the 1.1*TAC\text{target} stop criterion, was earlier found then the best solution hitherto found is reported. That is, a broader stop criterion is accepted. If no solution is found at all, the problem is infeasible.

### 5.3. Demonstration of the new hybrid method

The new hybrid method has been demonstrated on example 4.1 of Hallale (1998). The target is removal of ammonia from five gaseous streams which are composed mainly of air. Three water-based streams are available for service. Two of them, S1 and S2, are process MSAs and one, S3, is an external MSA. Stream data are given in Table 25. Carbon steel columns packed with 25.4 mm Raschig rings are used as mass exchangers, with the lumped coefficient \( K_W = 0.02 \text{ kg NH}_3/\text{s/kg exchanger mass} \), and shell cost \( $618M^{0.66} \), where \( M \) is exchanger mass (kg).
Table 25. Stream data for the example

<table>
<thead>
<tr>
<th>Rich streams</th>
<th>$G$ (kg/s)</th>
<th>$y^i$ (mass fraction)</th>
<th>$y^j$ (mass fraction)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>2</td>
<td>0.005</td>
<td>0.0010</td>
</tr>
<tr>
<td>R2</td>
<td>4</td>
<td>0.005</td>
<td>0.0025</td>
</tr>
<tr>
<td>R3</td>
<td>3.5</td>
<td>0.011</td>
<td>0.0025</td>
</tr>
<tr>
<td>R4</td>
<td>1.5</td>
<td>0.010</td>
<td>0.0050</td>
</tr>
<tr>
<td>R5</td>
<td>0.5</td>
<td>0.008</td>
<td>0.0025</td>
</tr>
<tr>
<td>MSAs</td>
<td>$L^{up}$ (kg/s)</td>
<td>$x^i$ (mass fraction)</td>
<td>$x^j$ (mass fraction)</td>
</tr>
<tr>
<td>S1</td>
<td>1.8</td>
<td>0.0017</td>
<td>0.0071</td>
</tr>
<tr>
<td>S2</td>
<td>1</td>
<td>0.0025</td>
<td>0.0085</td>
</tr>
<tr>
<td>S3</td>
<td>$\infty$</td>
<td>0</td>
<td>0.0170</td>
</tr>
</tbody>
</table>

First, in Step 1, the target total annual cost of the problem is determined with the supertargeting method of Hallale and Fraser (1998). $TAC_{target}$ of this example is $788,405$ $\$/yr at 0.0007 minimum approach composition. The initial flowsheet is then constructed based on the assumptions mentioned above that one (in this case the only S3) external lean stream takes all the pollutant from all the rich streams in one stage. The generated initial structure is shown in Figure 73.
The driving force plot for the minimum external lean stream flowrate is also determined, and depicted in Figure 74. In this figure, the composite operating line is dashed, and the operating lines of the particular units in the initial structure are shown by solid lines. The driving forces of units 3, 4, and 5 are too high compared to the composite operating line, indicating that the initial flowsheet is far from the optimum.
In the next step, the MINLP model was optimized using GAMS (Brooke et al., 1998) and DICOPT++ solver on a Sun Sparc Station. The stop criterion in a run was the worsening of the objective in the NLP subproblem. The optimal flowsheet of the first iteration is shown in Figure 75.

The optimum was found at objective value $TAC = 928342 \$/yr. This solution is not better than $1.1 \times TAC_{target}$; therefore, the MINLP model is extended in Step 3.

The found binary solution is excluded using integer cut. The same found solution is used as initial flowsheet. The earlier found $TAC$ value is given as an upper constraint to the $TAC$. The driving force plot for the found solution is shown in Figure 76.
Figure 76. Driving force plot based on the MINLP solution of the 1st iteration

The average driving force is situated above the composite operating line; therefore, an upper bound has to be applied. The least efficient unit is unit 5; therefore, its lean stream S1 has to be bounded.

Calculation of the efficiency factor for all the lean streams it is not necessary in either case. The actual value of the flowrate of lean stream S3 is 2.971 kg/s. The new upper bound, which is less then this value with 5%, is 2.822 kg/s.

The modified MINLP representation was then solved. The optimal solution of the problem has been found in 4 iteration steps. The consecutive improvements in $TAC$ are shown in Table 26. The $TAC$ of the optimal structure is 808,986 $/yr. The optimal structure is shown in Figure 77.

<table>
<thead>
<tr>
<th>No of iteration:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$TAC$ ($/year$)</td>
<td>928,342</td>
<td>894,872</td>
<td>868,571</td>
<td>808,986</td>
</tr>
</tbody>
</table>
5.4. Conclusion
A new hybrid optimization method for mass exchange network synthesis has been developed. The new method uses integer cuts and bounds calculated on the basis of driving force plot analysis. A new initial solution is constructed only if the MINLP solution is infeasible; otherwise the earlier found best solution is used. The method has been demonstrated on a middle scale MENS problem involving five rich streams and two process lean streams, and one external lean stream. The optimal solution has been found in four iteration steps; the value of the objective was improved in each step. The new method is fairly robust and can be accomplished in an automatic way.
6. Optimization of desalination location problem using MILP

6.1. Problem statement and test problems

The desalination plant location problem presented here involves a given number of cities each of which contains at least one storage tank to receive the imported potable water. The potable water will be produced in different kinds of desalination plants from seawater which is delivered to the plants by pipelines from seawater intakes located at the sea coast. The locations of the seawater intakes and the potable water storage tanks are known. The demands for water at each storage location (at each city) are also known. The storage tanks, the desalination plants, and the seawater intakes are considered punctually, i.e. their extension on the plane is not considered. The problem also includes forbidden zones such as cities, sea surface, and some other places at which desalination plants are not allowed to exist. An example of the location problem plane is depicted in Figure 78. The total cost of the problem consists of the fixed costs of the plants and pipelines, and the variable costs of the production and transportation. Our task is to find the optimal number and coordinates of the plants, such that the total cost is at minimum.

F: feasible region, IF: infeasible region

Figure 78. The location problem plane

Two example problems have been worked out for testing the capability of the developed models.

6.1.1. Problem I

A 20 km by 10 km rectangle contains 5 storage tanks, and a part of the sea. The complementary area can be subdivided into at least 12 triangles. The locations of two seawater intakes and five fresh water storage tanks are given. At most 8 SD plants with capacity of 10 m$^3$/day, 3 RO plants...
with capacity of 250 m$^3$/day, and at most 3 RO plants with capacity of 500 m$^3$/day may be used. Detailed data of the Problem I are shown in Table 37 to Table 40 of Appendix.

### 6.1.2. Problem II
A 50 km by 30 km rectangle contains 6 cities and a part of the sea. The complementary area can be subdivided into at least 23 triangles. The locations of two seawater intakes and six fresh water storage tanks are given. At most 8 SD plants with capacity of 10 m$^3$/day, 3 RO plants with capacity of 250 m$^3$/day, and at most 3 RO plants with capacity of 500 m$^3$/day may be used. Detailed data of the Problem II are shown in Table 52 to Table 55 of Appendix.

### 6.2. Model versions and development strategy
Mixed-integer linear programming (MILP) models have been developed to determine the optimal location of the desalination plants, the amount of seawater transported from the water intakes to the plants, and the amount of potable water transported from the plants to the storage tanks.

The model is developed in three consecutive steps:
1. Superstructure development
2. GDP model development
3. MP model development

Developing a mathematical programming (MP) model directly based on the superstructure is a rather demanding task. The model developer should take into account several viewpoints including feasibility of the structures and the models, and relations between the imagined structures and their algebraic models. Existence of structure elements are modeled with binary variables in an algebraic model. Validity of some continuous relations (e.g. material balances) depends on such existence. The relations referring to the existence and coexistence of the structure elements are complicated expressions of mixed binary and continuous variables.

Generalized Disjunctive Programming (GDP) models, see Yeomans and Grossmann (1999), are much more similar to the human approach than an MP model because GDP applies logic variables for describing existence and coexistence of conditional entities in the superstructure. Therefore, instead of directly developing an MP model from the superstructure, it is much easier to develop a GDP model first. The logic relations of the GDP model are then transformed to binary and mixed binary relations with well known techniques like Big-M or Convex Hull, see
Vecchietti et al. (2003). In this way, a basic MP model is obtained that can be improved in subsequent development steps.

The structure of this section is the following: first the superstructure is presented, and then the MILP model is detailed. For the sake of intelligibility, the logic expressions and relations are explained first in their GDP form, and then their transformations to algebraic form are detailed. Where it was possible, convex hull was applied since it results in better relaxation than the Big-M technique. In case of some expressions, however, the Big-M technique was the only solution. Finally, the additional constraints which are used merely for improving the performance and accelerating the solution are presented in their algebraic form only.

6.2.1. The Basic model

6.2.1.1. Superstructure

The superstructure includes three kinds \((p)\) of desalination plants \((u_p)\), seawater intakes \((w)\) and fresh water storage tanks \((s)\). The superstructure is based on R-graph representation introduced by Farkas et al. (2005), and is depicted in Figure 79. The units and the pipelines are all conditional elements of the superstructure.

6.2.1.2. Model formulation

The model contains material balances, logic expressions on the existence of the plants and the pipelines, and constraints restricting the plant locations to the feasible region. In order to decrease the computation time, additional constraints are also introduced, see Section 6.2.1.3.
6.2.1.2.1. Material balance constraints
The material balances between the water intakes and the plants, and also between the plants and the storage tanks, must be satisfied. The following indices are used in this part of the model: \( w \) for seawater intakes, \( s \) for potable water storage tanks, \( p \) for the type of desalination plants (three types are used in our case), \( u_p \) for the units of type \( p \) of desalination plant, and \( k \) for the \( x\)-\( y \) coordinates. The material balance constraints describe the requirement of satisfying the conservation of material between the objects; between water intakes and plants (equations (6.1) to (6.2)), between plants and storage tanks (equations (6.4) to (6.5), and the equation describing the efficiency of the plants (equation (6.3)). The equations are given below, as follow.

\[
\sum_p \sum_{u_p} w_{iw_{p,u_p},w} p \in PT, u_p \in U_p (6.1)
\]

\[
\sum_{w} \sum_{u_p} w_{iw_{p,u_p},w} w \in W (6.2)
\]

\[
w_{p,u_p} = E_p \cdot w_{p,u_p} p \in PT, u_p \in U_p (6.3)
\]

\[
wp_{p,u_p} = \sum_s w_{p,u_p,s} p \in PT, u_p \in U_p (6.4)
\]

\[
ws_s = \sum_p \sum_{u_p} w_{p,u_p,s} s \in S (6.5)
\]

6.2.1.2.2. Distances
The transportation costs are proportional to the length of each pipeline; these in turn depend on the distances between sea water intakes and plants, and between plants and storage tanks. Distances are defined according to the Manhattan metric, see Peer et al. (2006). Manhattan distance is the expression \( D_{m}^{12} = |x_2 - x_1| + |y_2 - y_1| \) used for approximating the nonlinear Euclidian distance \( D_{eu}^{12} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \) of two points \((x_1, y_1)\) and \((x_2, y_2)\). Although the absolute-value function is not linear, it can be modeled using two linear non-equalities \( D_{m}^{12} \geq x_2 - x_1 \) and \( D_{m}^{12} \geq x_1 - x_2 \) if \( D_{m}^{eu} \) is minimized in the objective function. Of course, as the Manhattan-distance is only an estimation of the Euclidean distance, the optimal solution can be affected by the position of the Cartesian system.

Equations computing the distance between seawater intake and any plant are:

\[
D_{w_{p,u_p,k}} \geq (p_{w,k} - pp_{p,u_p,k}) p \in PT, u_p \in U_p, w \in W, k \in K (6.6)
\]

\[
D_{w_{p,u_p,k}} \geq (pp_{p,u_p,k} - p_{w,k}) p \in PT, u_p \in U_p, w \in W, k \in K (6.7)
\]
Equations computing the distance between the plants and storage tanks are:

\[ D_{s,p,u,r,k} \geq (pp_{s,p,u,r,k} - ps_{s,k}) \quad p \in PT, u_p \in U_p, s \in S, k \in K \] (6.8)

\[ D_{s_{s,k}} \geq (ps_{s,k} - pp_{p,u,k}) \quad p \in PT, u_p \in U_p, s \in S, k \in K \] (6.9)

### 6.2.1.2.3. Capacity constraints

The amount of produced potable water is constrained by the capacity of the plant:

\[ wp_{p,u} \leq CAP_p \quad p \in PT, u_p \in U_p \] (6.10)

### 6.2.1.2.4. Logic expressions

Logic expression on the existence of the connection between seawater intake \( w \) and plant \( u_p \) of type \( p \). The amount of water transported through a non-existing pipeline must be zero. The fixed cost of an existing pipeline is calculated from its length:

\[
\begin{align*}
Z_{wp_{w,p,u}} & \\
0 \leq \text{woi}_{w,p,u} & \leq CAPWP_{w,p} \\
\text{fcostwp}_{w,p,u,k} & = fcp \cdot Dw_{w,p,u,k}
\end{align*}
\]

\[
\begin{align*}
-w_{wp_{w,p,u}} & \\
\text{woi}_{w,p,u} & = 0 \\
\text{fcostwp}_{w,p,u,k} & = 0
\end{align*}
\]

\[ w \in W, p \in PT, u_p \in U_p, k \in K \] (6.11)

Logic expression on the existence of plants \( u_p \) of type \( p \). The amount of water produced by a non-existing plant must be zero. The fixed cost of an existing plant is equal to its fixed cost parameter:

\[
\begin{align*}
Z_{p_{p,u}} & \\
0 \leq wp_{p,u} & \leq wp_{p,u_p} \\
\text{fcostp}_{p,u,k} & = \text{fcup}_p
\end{align*}
\]

\[
\begin{align*}
-Z_{p_{p,u}} & \\
wp_{p,u} & = 0 \\
\text{fcostp}_{p,u,k} & = 0
\end{align*}
\]

\[ p \in PT, u_p \in U_p \] (6.12)

Logic expression on the existence of the connection between plant \( u_p \) of type \( p \) and storage tank \( s \). The amount of water transported by a non-existing pipeline must be zero. The fixed cost of an existing pipeline is calculated from its length:

\[
\begin{align*}
Z_{ps_{p,u_r}} & \\
0 \leq \text{wps}_{p,u_r} & \leq CAPPS_{p,s} \\
\text{fcostps}_{p,u_r,k} & = fcp \cdot Ds_{p,u_r,k}
\end{align*}
\]

\[
\begin{align*}
-Z_{ps_{p,u_r}} & \\
\text{wps}_{p,u_r} & = 0 \\
\text{fcostps}_{p,u_r,k} & = 0
\end{align*}
\]

\[ p \in PT, u_p \in U_p, s \in S, k \in K \] (6.13)
6.2.1.2.5. Transformation of logic expressions

The above logic expressions are transformed to algebraic expressions using the Big-M technique.

For example, equation (6.13) is reformulated as a system of equations below:

\[ \text{wps}_{p,u,s} \leq \text{CAPPS}_{p,s} \cdot \text{yps}_{p,u,s}, \quad p \in PT, u_p \in U_p, s \in S \]  \hspace{1cm} (6.14)

\[ f_{\text{costp}}_{p,u,s,k} - f_{\text{cp}} \cdot D_{p,u,s,k} \leq U_{\text{D}}_{s,k} \cdot (1 - \text{yps}_{p,u,s}) \]

\[ p \in PT, u_p \in U_p, s \in S, k \in K \]  \hspace{1cm} (6.15)

\[ -U_{\text{D}}_{s,k} \cdot (1 - \text{yps}_{p,u,s}) \leq f_{\text{costp}}_{p,u,s,k} - f_{\text{cp}} \cdot D_{p,u,s,k} \]

\[ p \in PT, u_p \in U_p, s \in S, k \in K \]  \hspace{1cm} (6.16)

Reformulation of the other logic expressions is performed with the same technique.

6.2.1.2.6. Assignment of feasible and infeasible regions

The plane of the location problem is divided into two regions. Feasible region (F) is the area where the desalination plants may be located, and infeasible region (IF) or forbidden zone is the area where plants should not be located (e.g. sea and cities, see the problem statement), as depicted in Figure 78. The plants must be located in the feasible region. Two distinct methods are separately used to assign the feasible region.

6.2.1.2.6.1. Use of triangles

The whole feasible region is subdivided into disjunctive triangles denoted by index \( j \). This subdivision is made in advance to any optimization. The number of triangles should be as small as possible because the number of binary variables used for modeling the problem strongly depends on it. How to find a subdivision with a minimum or at least a small number of triangles is not discussed here. The subdivision actually applied is considered as a constituting part of the problem statement.
Such a triangle is shown in Figure 80. The $\Pi_1$, $\Pi_2$ and $\Pi_3$ directed edges of the triangle in Figure 80 are used as the basis vectors of an oblique frame of axes. In general, any point $P$ inside the rectangular shown in Figure 80 can be represented by the following expression:

$$P_k = \Pi_{1k} + \lambda_{21} \cdot (\Pi_{2k} - \Pi_{1k}) + \lambda_{31} \cdot (\Pi_{3k} - \Pi_{1k}) \quad (k \in K)$$  \hspace{1cm} (6.17)

where $\lambda_{21}$ and $\lambda_{31}$ are coordinates of the oblique system. If $\lambda_{21}$ and $\lambda_{31}$ are non-negative and their sum is less than or equal to 1 then $P$ is situated in the triangle constituted by $\Pi_1$, $\Pi_2$ and $\Pi_3$. Why it is so is shortly explained in Section 11.3 of the Appendix.

We introduce a new logic variable $Z_{pj \_up \_j}$ which represents the information whether a plant $u_p$ of type $p$ is located inside triangle $j$ (i.e. $Z_{pj \_up \_j}$ is true) or outside it (i.e. $Z_{pj \_up \_j}$ is false). The following mixed logic expression is applied in the model for obtaining correct $\lambda_{21}$ and $\lambda_{31}$ values. If plant $u_p$ of type $p$ is located inside triangle $j$ then $\lambda_{21 \_up \_j}$ and $\lambda_{31 \_up \_j}$ belonging to them are non-negative and their sum is less than or equal to 1:

$$\begin{align*}
&\left[ Z_{pj \_up \_j} \\
&0 \leq \lambda_{21 \_up \_j} \\
&0 \leq \lambda_{31 \_up \_j} \\
&\lambda_{21 \_up \_j} + \lambda_{31 \_up \_j} \leq 1
\end{align*} \quad \land \quad \begin{bmatrix} -Z_{pj \_up \_j} \\
L_{21 \_j} \leq \lambda_{21 \_up \_j} \leq U_{21 \_j} \\
L_{31 \_j} \leq \lambda_{31 \_up \_j} \leq U_{31 \_j}
\end{bmatrix} \quad p \in PT, u_p \in U_p, j \in J \quad (6.18)
$$

Equation (6.18) is reformulated in the following way to represent the plant locations with binary variables instead of logic ones:

$$pp_{p \_up \_k} = P_{1 \_j \_k} + \lambda_{21 \_up \_j} \cdot (P_{2 \_j \_k} - P_{1 \_j \_k}) + \lambda_{31 \_up \_j} \cdot (P_{3 \_j \_k} - P_{1 \_j \_k})$$  \hspace{1cm} (6.19)

$$L_{21 \_j} \cdot (1 - y_{pj \_up \_j}) \leq \lambda_{21 \_up \_j} \quad p \in PT, u_p \in U_p, j \in J \quad (6.20)$$

$$L_{31 \_j} \cdot (1 - y_{pj \_up \_j}) \leq \lambda_{31 \_up \_j} \quad p \in PT, u_p \in U_p, j \in J \quad (6.21)$$
\[ \lambda_{21\ p, u_j} + \lambda_{31\ p, u_j} - 1 \leq (U\ 21\ j + U\ 31\ j) \cdot (1 - y_{pj\ p, u_j}) \]
\[ p \in PT, u_p \in U_p, j \in J \]

Due to the above equations, the lambdas are non-negative (equations (6.20) to (6.21)), and their sum is less than or equal to 1 (equation (6.22)) if the binary variable \( y_{pj} \) of the existence of the plant \( u_p \) of type \( p \) inside the triangle \( j \) equals to 1. Equation (6.23) expresses the requirement that plant \( u_p \) of type \( p \) must be contained by exactly one of the triangles.

\[ \sum_j y_{pj\ p, u_j} = 1 \quad p \in PT, u_p \in U_p \]  

(6.23)

### 6.2.1.2.6.2. Use of polygons

Any convex polygon can be drawn as an intersection of triangles. A particular point is contained by the polygon if and only if it is contained by each triangle enclosing the polygon. A sample of such a polygon is shown in Figure 81.

![Figure 81. Polygon assignment](image)

In this version, polygons (indexed by \( i \)) are used to assign the feasible region. Each non-triangle polygon can be assigned by two or more triangles. They must share the whole area of the polygon. The set of triangles (indexed by \( j \)) is sorted to subsets (each indexed by \( j_i \)) covering polygon \( i \). The triangles belonging to different subsets may overlap. The binary variable \( y_{pi\ p, u_j} \) expresses the existence of plant \( u_p \) of type \( p \) within polygon \( i \).

Equation (6.19) remains unchanged in this model. Equations (6.20) to (6.23) are substituted with equations (6.24) to (6.27), as follows:

\[ L_{21\ j_i} \cdot (1 - y_{pi\ p, u_j}) \leq \lambda_{21\ p, u_j, j_i} \quad i \in I, p \in PT, u_p \in U_p, j_i \in J_i \]  

(6.24)

\[ L_{31\ j_i} \cdot (1 - y_{pi\ p, u_j}) \leq \lambda_{31\ p, u_j, j_i} \quad i \in I, p \in PT, u_p \in U_p, j_i \in J_i \]  

(6.25)

\[ \lambda_{21\ p, u_j, j_i} + \lambda_{31\ p, u_j, j_i} - 1 \leq (U\ 21\ j_i + U\ 31\ j_i) \cdot (1 - y_{pi\ p, u_j}) \]
\[ i \in I, p \in PT, u_p \in U_p, j_i \in J_i \]  

(6.26)
\[ \sum_{i} y_{p,u,i} = 1 \quad p \in PT, u_p \in U_p \]  

(6.27)

6.2.1.2.7. Objective function

The objective function of the desalination location model is the total cost. The total cost is the sum of the capital costs of the existing plants, the variable production costs of the plants, the variable transportation costs to and from these plants, and the capital costs of the pipelines:

\[
\text{obj} = \sum_{p} \sum_{u_p} f_{\text{cost}}_{p,u_p} + \sum_{p} \sum_{u_p} \left( v_{c_p, u_p} \cdot w_{p,u_p} \right) + \sum_{w} \sum_{p} \sum_{u_p} \left( v_{c_p} \cdot w_{o_i, u_p} \right) + \\
+ \sum_{s} \sum_{p} \sum_{u_p} \left( v_{c_p} \cdot w_{s,u_p} \right) + \sum_{w} \sum_{p} \sum_{u_p} \sum_{k} f_{\text{cost}}_{w, p, u_p, k} + \sum_{s} \sum_{p} \sum_{u_p} \sum_{k} f_{\text{cost}}_{s, p, u_p, r, k}
\]

(6.28)

6.2.1.2.8. Additional equations

These equations are used merely for accelerating the solution procedure; they are not part of the Basic MP but the results of the model-improving phase; therefore, they are presented in their algebraic form only. In order to decrease the solution time, relaxation of the Big-M equations must be as tight as possible. For example, an upper bound to the distance parameter \( UDS_{s,k} \) between a particular storage tank and any plant can be calculated as the maximum of the distances between the storage tank and the borders of the plane.

In order to enhance the lower bound of the total cost (and thereby to make faster the optimization process) another constraint, using the minimum distance between the storage and the nearest water intake (\( D_{\text{min} s,k} \)), is employed for calculating the minimum of the pipeline costs. If a particular connection between a storage tank and a plant exists then the sum of the pipelining costs calculated from the length of the pipelines between the storage tank and the plant, and between the plant and all the water intakes, must be larger than the pipelining cost calculated from the minimum distance between the storage tank and the nearest water intake:

\[
f_{c_p} \cdot D_{\text{min} s,k} \cdot y_{s, u_p, s} \leq f_{\text{cost}}_{s, p, u_p, s, k} + \sum_{w} f_{\text{cost}}_{w, p, u_p, k}
\]

(6.29)

Calculation of the minimum cost for water intakes is performed with the same method.

Additional equations with binary variables for describing co-existence, or exclusion of co-existence, of some structure elements are also applied. With these equations, infeasible solutions can be excluded without examining the node in the Branch and Bound algorithm.
If a particular plant exists then at least one pipeline should be directed to this plant from one of the water intakes, and at least one pipeline should connect this plant to one of the storage tanks. These constraints are described by equations (6.30) and (6.31).

\[ \sum w y_{wp} \leq \sum p y_{wpu}, \quad p \in PT, u_p \in U_p \tag{6.30} \]
\[ \sum s y_{ps} \leq \sum p y_{ps}, \quad p \in PT, u_p \in U_p \tag{6.31} \]

If a particular pipeline from a particular water intake to a particular plant exists then that plant must also exist (equation (6.32)). Similarly, if a particular pipeline from a particular plant to a particular storage tank exists then that plant must also exist (equation (6.33)).

\[ y_{wp} \leq y_{pu}, \quad w \in W, p \in PT, u_p \in U_p \tag{6.32} \]
\[ y_{ps} \leq y_{pu}, \quad p \in PT, u_p \in U_p, s \in S \tag{6.33} \]

6.2.1.3. Model variants

The general model hitherto explained is formulated in such a way that it may have multiple solutions with the same objective function value. Therefore, additional constraints are inserted to the model in order to decrease the number of equivalent solutions, without restricting the generality of the model. In order to decrease the computation time, six different variants are created.

**Variant 0:** No additional constraint is used; this is the basic variant.

**Variant 1:** A particular plant has a higher \( x \) coordinate than its predecessor in the increasing order of its index:

\[ p_{p,u_p,1,x} \leq p_{p,u_p,x}, \quad p \in PT, u_p \in U_p \tag{6.34} \]

**Variant 2:** The plants are ordered along the \( y \) axis:

\[ p_{p,u_p,1,y} \leq p_{p,u_p,y}, \quad p \in PT, u_p \in U_p \tag{6.35} \]

**Variant 3:** A particular plant cannot exist if its predecessor does not exist:

\[ y_{p,u_p} \geq y_{p,u_p}, \quad p \in PT, u_p \in U_p \tag{6.36} \]

**Variant 4:** Combination of the first and the third variants.

**Variant 5:** Combination of the second and the third variants.
6.2.1.4. Test runs of the Basic model
GAMS 20.0. modeling system was used for all of the problems described in this work. CPLEX version 7.0.0 was applied as the MILP solver. The runs were performed on a PC with Intel 2.39 GHz and 752 MB RAM.

6.2.1.4.1. Results of applying the Basic model to Problem I

6.2.1.4.1.1. Results with triangles
The complementary area is subdivided into 12 triangles. The model statistics are detailed in Table 27.

<table>
<thead>
<tr>
<th>Number of constraints</th>
<th>Number of continuous variables</th>
<th>Number of binary variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 076</td>
<td>1 320</td>
<td>280</td>
</tr>
</tbody>
</table>

The variants of additional constraints are applied on the above example problem. The optimum was found $1 036 557/day with all the six variants, but the computation time was different, as shown in Table 28.

<table>
<thead>
<tr>
<th>Variant</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU Time [sec]</td>
<td>774.77</td>
<td>221.02</td>
<td>512.77</td>
<td>278.08</td>
<td>103.39</td>
<td>98.42</td>
</tr>
</tbody>
</table>

The results in Table 28 show that the shortest computation time is found with Variant 5. The optimal plant allocation and the structure of water distribution found with Variant 5 are shown in Figure 82 and Figure 83, respectively. One plant is situated between S1 and S4 storages, one is mapped to S2 storage only, and one is located in the same point as S5 storage. There are no U1 or U3 plants in use because their capacities are smaller than of the U2 plants. It is not economical to use plants with small capacities if there are no consumers with small demands, or if the consumers are close to each other.
6.2.1.4.1.2. Results with the use of polygons

In order to reduce the computation time, the feasible region is assigned with polygons instead of triangles. Use of 10 polygons is sufficient to cover the feasible region. Variant 5 is selected to solve because it has been the best one with triangles. The model statistics and results are detailed in Table 29. The same optimal solution is found ($1 036 557/day); the number of the binary variables is smaller by 10%; the computational time is decreased by 30%.

Table 29. Model statistics and results with Variant 5 of the Basic Model for Problem I, using polygons

<table>
<thead>
<tr>
<th>Number of constraints</th>
<th>Number of continuous variables</th>
<th>Number of binary variables</th>
<th>Solution [$/day]</th>
<th>CPU time [sec]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 566</td>
<td>1 488</td>
<td>252</td>
<td>1 036 557.463</td>
<td>65</td>
</tr>
</tbody>
</table>
The optimal plant allocation and the structure of the distribution found with Variant 5 are shown in Figure 84 and Figure 85, respectively. Not surprisingly, the structure found is similar to that found with triangles.

![Figure 84. Optimal allocation of the plants found with Variant 5, using polygons](image)

![Figure 85. Optimal water distribution found with Variant 5, using polygons](image)

### 6.2.1.4.2. Results of applying the Basic model to Problem II

Variant 5 of the Basic model is applied for 17 polygons instead of 23 triangles. The model statistics and results are detailed Table 30.
Table 30. Model statistics and results with Variant 5 of the Basic Model for Problem II, using polygons

<table>
<thead>
<tr>
<th>Number of constraints</th>
<th>Number of cont. variables</th>
<th>Number of bin. variables</th>
<th>Solution [$/day]</th>
<th>CPU time [sec]</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 659</td>
<td>1 992</td>
<td>364</td>
<td>3 019 669.4613</td>
<td>886.41</td>
</tr>
</tbody>
</table>

The optimum is found ($3 019 669/day) in 886.41 seconds. The optimal plant allocation and the structure of the distribution are shown in Figure 86 and Figure 87, respectively. The U1(2), U2(1) and U2(2) plants have all distinct pipelines. Obviously, it would be more economical if there was a splitting point in the center of the area enclosed by S1-S2-S3-S4 storage tanks. In this way only one pipeline would be needed between the W(2) water intake and the central splitting point. This problem will be addressed subsequently.

![Figure 86. Optimal allocation of the plants, using polygons](image-url)
Figure 87. Optimal water distribution, using polygons

6.2.1.5. Evaluation of the results of the Basic model
The Basic model is able to find an optimal solution for the defined problems in reasonable time. The effect of using additional equations to cut off the redundant solutions considerably accelerates the solutions process. Variant 0 without any additional equation decreasing the redundancy was the slowest one. The best result can be reached with Variant 5 where the plants are ordered along the y-axis, and a particular plant does not exist if its predecessor does not exist. It is therefore worthwhile to use this variant henceforward, although the difference of computational time between the Variants 4 and 5 is minimal.
The use of the polygons instead of triangles provides us with the same optimal solution. The saving in time was more than 30 \% in the case of Problem I; therefore that was used in the case of Problem II. When solving Problem II, the number of the binary variables increased by 44\%, and the solution time increased to its 1300 \%. This is a common phenomenon because in general the computational time changes exponentially with the number of binary variables.

6.2.2. The Improved model
As it can be seen in Figure 82, the pipelines between the first water intake and the first and second u\_2 plants partially coincide. In the Basic model, distinct pipelines are designed to and from different plants even if their trace lines coincide. This gives rise to redundant costs compared to the real life solution of designing common pipeline to the common section, and then branching it.
One of the consequences of neglecting the branching option is that plant types with low capacity (i.e. the solar distillation plants) will almost never be found economical. If the supplement of a
storage tank needs the usage of more solar distillation plants then each plant to storage tank connection requires a particular pipeline even if they are located at the same point of the plane. This is due to the inability of the pipeline to branch.

In order to deal with branching, so-called transportation units are introduced in the superstructure. The transportation units cannot produce potable water from seawater; the water flows unchanged through them. These new type units can have several inlet and outlet pipelines; this option provides us with the possibility of pipeline branching and/or merging.

In order to avoid mixing the potable water with seawater, two different types of transportation units are introduced. One is used for branching and merging seawater, the other one for branching and merging potable water. The transportation units are located either between the water intakes and the plants (denoted with ut1), or between the plants and the storage tanks (denoted with ut2).

### 6.2.2.1. Superstructure

The superstructure includes three kinds ($p$) of desalination plants ($u_p$), seawater intakes ($w$), fresh water storage tanks ($s$), first transportation units ($ut1$) and second transportation units ($ut2$). The superstructure is depicted in Figure 88.

![Figure 88. The improved superstructure of location problem](image-url)
6.2.2.2. Model formulation

6.2.2.2.1. Material balance constraints

The material balance constraints describe the requirement of conserving the water between the water intakes and first transport units (equations (6.37) to (6.38)), between the first transport units and plants (equations (6.39) to (6.40)), between the plants and second transport units (equations (6.42) to (6.43)), and between the second transport units and storage tanks (equations (6.44) to (6.45)). Equation (6.41) describes the efficiency of the plants. The equations are given below, as follow.

\[ w_{o_w} = \sum_{w} w_{wut_{1_{w,ut}}} \quad w \in W \quad (6.37) \]

\[ w_{ut_{1_{ut}}} = \sum_{w} w_{wut_{1_{w,ut}}} \quad ut1 \in UT1 \quad (6.38) \]

\[ w_{ut_{1_{ut}}} = \sum_{p} \sum_{u_p} w_{ut_{1_{p,u_p,ut}}} \quad ut1 \in UT1 \quad (6.39) \]

\[ w_{i_{p,u_p}} = \sum_{u_{11}} w_{p_{11}} \quad p \in PT, u_p \in U_p \quad (6.40) \]

\[ w_{p_{11}} = E_{w} \cdot w_{i_{p,u_p}} \quad p \in PT, u_p \in U_p \quad (6.41) \]

\[ w_{p_{11}} = \sum_{u_{12}} w_{p_{12}} \quad p \in PT, u_p \in U_p \quad (6.42) \]

\[ w_{ut_{1_{ut}}} = \sum_{p} \sum_{u_p} w_{ut_{1_{p,u_p,ut}}} \quad ut2 \in UT2 \quad (6.43) \]

\[ w_{ut_{1_{ut}}} = \sum_{s} w_{ut_{2_{s,ut}}} \quad ut2 \in UT2 \quad (6.44) \]

\[ w_{s} = \sum_{s} w_{ut_{2_{s,ut}}} \quad s \in S \quad (6.45) \]

6.2.2.2.2. Distances

Distances between seawater intake and first transportation units, between first transportation units and plants, between plants and second transportation units, and between second transportation units and storage tanks have to be calculated in the model in order to determine the transportation cost. Distances are defined according to Manhattan metric, see Peer et al. (2006).

Equations computing the distance between seawater intake and any first transportation unit are:

\[ D_{w_{1_{w,ut1,k}}} \geq (p_{w_{1_{w,ut1,k}}}) \quad ut1 \in UT1, w \in W, k \in K \quad (6.46) \]

\[ D_{w_{1_{w,ut1,k}}} \geq (p_{1_{w,ut1,k}} - p_{w_{1_{w,ut1,k}}}) \quad ut1 \in UT1, w \in W, k \in K \quad (6.47) \]
Equations computing the distance between the first transportation unit and any plant are:

\[ \text{Dut}_1 p_{u \in U, \, p \in P} \geq (\text{put}_1 - pp_{u \in U, \, p \in P}) \quad p \in PT, \, u \in U, \, p \in U, \, ut1 \in UT1, k \in K \]  

(6.48)

\[ \text{Dut}_2 p_{u \in U, \, p \in P} \geq (pp_{u \in U, \, p \in P} - \text{put}_1) \quad p \in PT, \, u \in U, \, p \in U, \, ut1 \in UT1, k \in K \]  

(6.49)

Equations computing the distance between plant and any second transportation unit are:

\[ \text{Dput}_2 p_{u \in U, \, p \in P} \geq (\text{put}_2 - pp_{u \in U, \, p \in P}) \quad p \in PT, \, u \in U, \, p \in U, \, ut2 \in UT2, k \in K \]  

(6.50)

\[ \text{Dput}_2 p_{u \in U, \, p \in P} \geq (pp_{u \in U, \, p \in P} - \text{put}_2) \quad p \in PT, \, u \in U, \, p \in U, \, ut2 \in UT2, k \in K \]  

(6.51)

Equations computing the distance between the second transportation unit and storage tanks are:

\[ \text{Dut}_2 s_{s \in S, \, k \in K} \geq (\text{put}_2 - ps_{s \in S, \, k \in K}) \quad ut2 \in UT2, s \in S, k \in K \]  

(6.52)

\[ \text{Dut}_2 s_{s \in S, \, k \in K} \geq (ps_{s \in S, \, k \in K} - \text{put}_2) \quad ut2 \in UT2, s \in S, k \in K \]  

(6.53)

6.2.2.2.3. Logic expressions

Logic expression on the existence of the connection between sea water intake \( w \) and first transportation unit \( ut1 \). The amount of water transported through a non-existing pipeline must be zero. The fixed cost of an existing pipeline is calculated from its length:

\[ \begin{align*}
 Zw_{ut1} & \leq \text{CAPWUT}_{ut1} \\
 f_{costut1} &= f_{cp} \cdot D_{ut1} \\
 w & \in W, ut1 \in UT1, k \in K 
\end{align*} \]

(6.54)

Logic expression on the existence of the connection between first transportation unit \( ut1 \) and plant \( up \) of type \( p \). The amount of water transported through a non-existing pipeline must be zero. The fixed cost of an existing pipeline is calculated from its length:

\[ \begin{align*}
 Z_{ut1} p_{u \in U, \, p \in P} & \leq \text{CAPUT}_{ut1, p} \\
 f_{costut1} p_{u \in U, \, p \in P} &= f_{cp} \cdot \text{Dut}_1 p_{u \in U, \, p \in P} \\
 w_{ut1} p_{u \in U, \, p \in P} &= 0 \\
 f_{costut1} p_{u \in U, \, p \in P} &= 0 \\
 ut1 & \in UT1, p \in PT, u \in U, k \in K 
\end{align*} \]

(6.55)

Logic expression on the existence of plants \( up \) of type \( p \). This is same as used in the Basic model.

Logic expression on the existence of the connection between plant \( up \) of type \( p \) and second transportation unit \( ut2 \). The amount of water transported by a non-existing pipeline must be zero. The fixed cost of an existing pipeline is calculated from its length:
Logic expression on the existence of the connection between second transportation unit $ut_2$ and storage tanks $s$. The amount of water transported by a non-existing pipeline must be zero. The fixed cost of an existing pipeline is calculated from its length:

$$Z_{put_2} = \begin{cases} \leq w_{put_2} \cdot C_{APPUT} & \text{if } w_{put_2} \neq 0 \\ \geq C_{APPUT} & \text{if } w_{put_2} = 0 \\ \text{if } \text{some other condition}\end{cases}$$

$$f_{cost_{put_2}} = f_{cap} \cdot D_{put_2}$$

$$p \in P_T, u_p \in U_p, ut_2 \in UT_2, k \in K$$  \hspace{1cm} (6.56)

Transformation to algebraic equations. Equation (6.57) is reformulated to algebraic equations using the same method as used for transforming equations (6.11) to (6.13):

$$w_{ut_2} \cdot C_{CAPUT} \leq y_{ut_2} \cdot C_{CAPUT}$$  \hspace{1cm} (6.58)

$$f_{cost_{ut_2}} - f_{cap} \cdot D_{ut_2} \leq U_{DS} \cdot (1 - y_{ut_2})$$  \hspace{1cm} (6.59)

Reformulation of the other logic expressions is performed with the same technique.

6.2.2.2.4. Assignment of the feasible region

The equations of the assignment of feasible region are the same as in the Basic model.

6.2.2.2.5. Objective function

The objective function is the total cost. The total cost is the sum of the capital costs of the existing plants, the variable production costs of the plants, the variable transportation costs to and from these plants, and the capital costs of the pipelines:
\[ \text{obj} = \sum_{p} \sum_{u_p} f\text{cost}_{p,u_p} + \sum_{p} \sum_{u_p} (vcpup_p \cdot w\text{p}_{p,u_p}) + \sum_{w} \sum_{u_{wl1}} (vcp \cdot w\text{ut1}_{w,u_{wl1}}) + \]
\[ + \sum_{w} \sum_{u_{wl1}} \sum_{p} (vcp \cdot w\text{ut1}_{w1,u_{wl1},p}) + \sum_{w} \sum_{u_{wl2}} \sum_{p} (vcp \cdot w\text{ut2}_{p,u_{wl2}}) + \sum_{w} \sum_{u_{wl2}} \sum_{s} (vcp \cdot w\text{ut2}_{s,u_{wl2},s}) + \]
\[ + \sum_{w} \sum_{u_{wl1}} \sum_{k} f\text{cost}_{wut1}_{w,u_{wl1},k} + \sum_{w} \sum_{u_{wl2}} \sum_{p} \sum_{u_p} f\text{cost}_{p,u_{wl2},k} + \sum_{w} \sum_{u_{wl2}} \sum_{p} \sum_{u_p} f\text{cost}_{p,u_{wl2},k} + \]
\[ + \sum_{s} \sum_{u_{wl2}} \sum_{k} f\text{cost}_{2s}_{u_{wl2},s,k} \]
\[(6.61)\]

### 6.2.2.2.6. Additional equations

A relation similar to equation (6.29) is used in the Improved model, but not the same one because it cannot enhance the solution process here. Namely, if a particular connection from a storage tank to a plant exists, then the pipeline cost of this connection and the sum of the costs of the connections to the same plant from the water intakes must be at least as high in the Basic model as the cost of a pipeline to the storage tank from the nearest water intake.

This constraint is changed in the Improved model because of the presence of the transportation units. If a particular connection to a storage tank from a transportation unit \( ut2 \) exists, then the pipeline cost of this connection and the sum of the costs of the connections to this transportation unit \( ut2 \) from any plant, plus the sum of the costs of the connections to any plant from any transportation unit \( ut1 \), plus the sum of the costs of the connections to any transportation unit \( ut1 \) from any water intake, must be at least as high as the cost of a pipeline to the storage tank from the nearest water intake. This constraint is expressed by equation (6.62).

\[ f\text{cp} \cdot D\text{mins}_{s,k} \cdot y\text{ut2}_{s,u_{wl2},s} \leq f\text{cost}_{2s}_{u_{wl2},s,k} + \sum_{p} \sum_{u_p} f\text{cost}_{p,u_{wl2},k} + \sum_{w} \sum_{u_{wl1}} f\text{cost}_{wut1}_{w,u_{wl1},k} \]
\[ \text{ut2} \in UT2, s \in S, k \in K \]
\[(6.62)\]

The following relations are analogous to equations (6.30) to (6.33):

\[ y\text{p}_{p,u_p} \leq \sum_{w} y\text{ut1}_{p,u_{wl1},p,u_p} \]
\[ p \in PT, u \in U_p \]
\[(6.63)\]

\[ y\text{p}_{p,u_p} \leq \sum_{w} y\text{put2}_{p,u_{wl2},p,u_{wl2}} \]
\[ p \in PT, u \in U_p \]
\[(6.64)\]

\[ y\text{ut1}_{p,u_{wl1},p,u_p} \leq y\text{p}_{p,u_p} \]
\[ ut1 \in UT1, p \in PT, u \in U_p \]
\[(6.65)\]

\[ y\text{put2}_{p,u_{wl2},p,u_{wl2}} \leq y\text{p}_{p,u_p} \]
\[ p \in PT, u \in U_p, ut2 \in UT2 \]
\[(6.66)\]
6.2.2.2.7. Tightening the relaxation

The above model in its hitherto presented form cannot be used within reasonable time. The relaxation made with the Big-M equations is therefore tightened by applying reasonable estimation for the maximum distance Big-M parameters. Such estimations can be made by using the solution obtained with the Basic model.

In the original form, the upper bound of the distance is equal to the extent of the plane in a direction. This is decreased by creating a particular feasible region for each transport unit and plant. The plane is divided to three approximately equal size sub-regions along the x-axis. One third of the plants and transportation units are located in the left hand side of the plane, another third in the middle, and the last third in the right hand side of the plane.

The transportation units located in the left hand side region are not connected to the storage tanks located in the right hand side region of the plane, according to the estimation used for bounding. It is rather improbable to have an optimum solution with such a connection because of the high transportation costs due to the long pipelines. Therefore, the connections between units being far from each other are forbidden. The Big-M equations, i.e. equations (6.59) and (6.60), are applied to the allowed connections only:

\[
\begin{align*}
& f_{cost} 2s_{2s, 2s, k} - f_{cp} \cdot D_{ut} 2s_{2s, 2s, k} \leq f_{cp} \cdot U_{DUT} 2S_{ut2s, 2s, k} - (1 - y_{ut} s_{ut, s}) \\
& (ut2, s) \in \{(ut2, s) | ut2 \in UT2, s \in S, y_{ut} s_{up}(ut2, s) = 1\}, k \in K \\
& - f_{cp} \cdot U_{DUT} 2S_{ut2s, 2s, k} - (1 - y_{ut} s_{ut, s}) \leq f_{cost} 2s_{2s, 2s, k} - f_{cp} \cdot D_{ut} 2s_{2s, 2s, k} \\
& (ut2, s) \in \{(ut2, s) | ut2 \in UT2, s \in S, y_{ut} s_{up}(ut2, s) = 1\}, k \in K
\end{align*}
\]

Consequently, the Big-M parameters will not be the maximum distances (i.e. the width or height in the plane), but the maximum allowed distances only.

This approach basically is a heuristic. In order to avoid cutting off the optimum, it has to be used in an iterative form. First, the Big-M parameters of the distances in equations (6.67)-(6.68) can be estimated using the results obtained with the Basic model. Then after increasing the maximum allowed distances and the number of the allowed connections, the resulted change in the best value of the objective function can be checked. The algorithm is terminated when there is no further improvement.

Application of restrictions (6.33) to (6.35) for the locations and connections of the plants as in the Basic model can cut off the optimum or even all the feasible solutions in case of the Improved model. Therefore these constraints are not applied in the Improved model.

The actual feasible regions and forbidden connections used in the test examples are detailed in Table 41 to Table 51 and Table 56 to Table 66 of the Appendix.
6.2.2.3. Test runs of the Improved model
The problems were solved in the same computing environment used for the Basic model.

6.2.2.3.1. Results of applying the Improved model to Problem I
The model statistics and results are detailed in Table 31. It can be seen that the computational time is increased by more than an order of magnitude, although the number of binary variables increased with 8% only. This is due to the deteriorating relaxation.

Table 31. Model statistics and results with Improved Model for Problem I, using polygons

<table>
<thead>
<tr>
<th>Number of constraints</th>
<th>Number of cont. variables</th>
<th>Number of bin. variables</th>
<th>Solution [$/day]</th>
<th>CPU time [sec]</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 826</td>
<td>2 373</td>
<td>272</td>
<td>926 338.8335</td>
<td>3 917.63</td>
</tr>
</tbody>
</table>

The optimal solution found is $926 338/day, 10.63% less than the solution obtained with the Basic model; the computation time was 3 617.63 seconds. The optimal plant allocation and the structure of the distribution are shown in Figure 89 and Figure 90, respectively. Here the S2 storage tank obtains its water from W(2) water intake. In point (35; 20) there is a splitting centre. There is only one pipeline between W(2) and this splitting point. This could not be possible with the Basic model. Similarly, there is another splitting centre in (20; 25) where the water comes to S4 storage tank from S1. This could not be possible with the Basic model, either. The effect of these two splitting centers provides the decreasing in the total cost. New phenomena is that one U3 solar distillation plant is used as well. This, again, was not possible with the Basic model because this plant would need a distinct pipeline. Here, the plant is situated in the splitting point. In this way, only one pipeline is needed between the W(1) water intake and the splitting point; and only one is needed between splitting point and S3 storage tank.
6.2.2.2. **Results of applying the Improved model to Problem II**

The model statistics and the results are detailed in Table 32. It can be seen that the computational time increased by its 700 % although the number of the binary variables increased with 1 % only. This is due to the deteriorating relaxation, again.

<table>
<thead>
<tr>
<th>Number of constraints</th>
<th>Number of cont. variables</th>
<th>Number of bin. variables</th>
<th>Solution [$/day]</th>
<th>CPU time [sec]</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 384</td>
<td>2 815</td>
<td>368</td>
<td>2 546 144.5489</td>
<td>7 051 75</td>
</tr>
</tbody>
</table>

The optimal solution found is $2 546 144/day, which is 13.23% less than the previous solution; the computation time is 7 051.75 seconds.
The plant allocation and the structure of the distribution are shown in Figure 91 and Figure 92, respectively. There is a huge water distribution hub in point (100; 75). Between W(2) and this splitting point, there is only one pipeline. This could not be possible with the Basic model.

Figure 91. Optimal allocation of the plants using polygons with transportation units

Figure 92. Optimal water distribution using polygons with transportation units

6.2.2.4. Conclusion with the Improved model
Using well-estimated Big-M values, the Improved model provides a solution which takes into account the possibility of pipeline branching. Consequently, the solution obtained using the Improved model designs network for the studied examples with 10 to 15% less total cost than the Basic model.
6.3. Conclusions

Successfully applicable MILP models have been developed for determining optimal desalination plant positions in a given feasible subset of the plane.

The Basic model is able to find a near optimal solution with the deficiency of allocating redundant multiple pipelines. Equivalent solutions can be excluded by applying additional constraints. All the six model variants obtained in this way result in identical solution. Variant 5 is the fastest one; this variant applies a special ordering of locations (equation (6.35)) and plant existence (equation (6.36)).

The feasible region can be allocated with using triangles or polygons; use of polygons decreases the computation time.

The deficiency of the Basic model can be removed by introducing transportation units in the superstructure. The Improved model is capable of finding optimal solution with branching pipelines.
7. Major new results

I. I have proven that optimization can effectively be used for analyzing complex distillation processes. Analysis of extractive distillation processes from both operational and economical points of view is, although possible but, inconvenient using process simulators. It implies a lot of necessary trial and error experiments; see e.g. the iterative method given by Chadda, Malone, and Doherty (2000). Determining the minimum and maximum values of both the stage numbers and reflux ratio can be difficult, and can be more easily accomplished by optimization. Minimum and maximum reflux ratios for a certain number of stages can be found in one step using NLP. Also the minimum and maximum stage numbers can be found by systematic search in less time and with less trouble than with use of simulators.

II. I have proven that the optimal configuration of extractive distillation plants is widely independent of the weight of the cost factors. It follows that some inaccuracy in estimating the cost-factors or small changes in the economic environment will not alter the optimal configuration. The economical risk is relatively small when investing into extractive distillation plants because the optimal configuration remains optimal or near-optimal even in changing economic environment.

III. I have developed a new hybrid method for synthesizing mass exchange network systems (MENS). The new method is able to eliminate the disadvantages of earlier methods. The new insight based method uses integer cuts and bounds calculated on the basis of driving force plot analysis. A new initial solution is constructed only if the MINLP solution is infeasible; otherwise the earlier found best solution is used. The method has been demonstrated on a middle scale MENS problem involving five rich streams and two process lean streams, and one external lean stream. The optimal solution has been found in four iteration steps; the value of the objective was improved in each step. The new method is fairly robust and can be accomplished in an automatic way.

IV. I have defined a new and practically important class of location-allocation problems, and developed a new MILP model for solving it. The task is to find the optimal type, number and coordinates of desalination plants, and the pipeline connections between seawater intakes, plants and cities, such that the potable water
demand of a set of cities is satisfied, while the total cost is at minimum. The total cost of the problem consists of the fixed costs of the plants and pipelines, and the variable costs of the production and transportation. The new MILP model takes into account the given locations and capacities of the water incomes, the potable water demands, and the costs of plants and pipelining. Feasible and infeasible plant regions are distinguished for locating the plants. The model has been developed in two consecutive phases. First a basic model is developed that provides a solution within short time but does not take into account the possibility of pipeline branching. Application of this model gives rise to redundant pipelines to some connections, involving extra costs. Pipeline branching is dealt with by an improved model developed in the second phase. This improved model provides realistic solution but with much longer computation time. The results of applying the different models on motivated examples of different sizes are detailed.

V. I have proven that the solar distillation desalination technology can be an economical alternative of the reverse osmosis plants. Due to the smaller capacity, the specific cost of the potable water is higher in the case of solar distillation plants. However, they can still be economical if the potable water demand arisen by small villages is small enough that it can be satisfied by just a few solar distillation plants.
8. Publications

8.1. Articles
Abdulfatah M. Emhamed, Zoltán Lelkes, Endre Rév, Tivadar Farkas, Zsolt Fonyó and Duncan
M. Fraser: “New hybrid method for mass exchange network optimization”, Chemical
Abdulfatah M. Emhamed, Barbara Czuczai, László Horváth, Endre Rév and Zoltán Lelkes:
2367-2383.
Abdulfatah M. Emhamed, Barbara Czuczai, Endre Rév and Zoltán Lelkes: “Analysis of
Extractive Distillation with Mathematical Programming”. Accepted for publication by Industrial
& Engineering Chemistry Research
Abdulfatah M. Emhamed, Barbara Czuczai, László Horváth, Endre Rév and Zoltán Lelkes: “An
Improved Desalination Location Model Using Mixed-Integer Linear Programming”, Accepted
by Alacademia Journal for Basic and Applied Sciences/ Libya

8.2. Conferences

8.2.1. Presentations
Abdulfatah M. Emhamed, Barbara Czuczai, László Horváth, Endre Rév and Zoltán Lelkes:
“Desalination location model using mixed integer linear programming” CHISA-17, Prague, 27-
31 August 2006.
Abdulfatah M. Emhamed, Zoltán Lelkes, Endre Rév, Tivadar Farkas, Zsolt Fonyó and Duncan

8.2.2. Posters

115
9. References

9.1. Optimization in general


ILOG (2000) CPLEX 7.0 user’s manual. ILOG CPLEX Division, Incline Village, NV., USA


9.2. Extractive Distillation


### 9.3. Mass Exchange Network Synthesis


**9.4. Desalination Location**


10. Notations

10.1. Extractive Distillation

\( b' \) intersection
$b_j$ coefficient of $j^{th}$ factor

$b_{jj}$ coefficient of the squared $j^{th}$ factor

$b_{jk}$ coefficient of the interaction of $j^{th}$ and $k^{th}$ factors

$hpay$ payback period [yr]

$btax$ tax factor [-]

$C_s$ unit price of steam [$/kJ$]

$C_w$ unit price of cooling water [$/kJ$]

$D$ column diameter [m]

$N$ number of stages [-]

$Q_C$ condenser duty [kJ/h]

$QR$ reboiler duty [kJ/h]

$UF$ update factor of the price of column installation [-]

$x_{ji}$ value of $j^{th}$ factor at $i^{th}$ experiment

$x_{ki}$ value of $k^{th}$ factor at $i^{th}$ experiment

$x_j$ value of $j^{th}$ factor

$\hat{Y}_i$ regressed actual objective function ($TotalCost$, sum of the reboiler duties $QR$, or sum of stage numbers $N$).

### 10.2. Mass Exchange Network Synthesis

DFP  Driving Force Plot

MEN  Mass Exchange Network

MENS  Mass Exchange Network Synthesis

MINLP  Mixed Integer NonLinear Programming

MSA  Mass Separating Agent

NLP  NonLinear Programming

**Sets**

$B$ set of existing units

$F$ set of feasible solutions

$I$ set of lean streams

$J$ set of rich streams
K set of stages in the superstructure

NB set of non-existing unit

**Superscripts**

- \( f \) feasible solution found earlier
- \( lo \) lower
- \( out \) output concentration
- \( s \) source
- \( t \) target
- \( up \) upper

**Subscripts**

- \( best \) earlier found best
- \( i \) rich stream
- \( j \) lean stream
- \( k \) stage in the superstructure
- \( t \) total
- \( target \) target

**Parameters and variables:**

- \( a \) specific surface area [\( m^2/m^3 \)]
- \( A \) actual mass exchange of stream [kg/s]
- \( b \) intercept of the equilibrium line [-]
- \( B \) maximal mass exchange of stream [kg/s]
- \( D \) column diameter [m]
- \( G \) rich stream flowrate [kg/s or kmol/h]
- \( H \) height of the column [m]
- \( K_y \) overall rich phase mass transfer coefficient [kg pollutant/s/m²]
- \( K_w \) mass lumped coefficient [kg pollutant/s/kg]
- \( L \) lean stream flowrate [kg/s or kmol/h]
- \( m \) slope of the equilibrium line [-]
- \( M \) mass of the exchanger [kg]
**10.3. Desalination location**

**Sets:**

- $I$: set of polygons
- $J$: set of triangles
- $J_i \subseteq J$: set of triangles belonging to polygon $i$
- $K$: set of coordinates
- $PT$: set of plant types
- $S$: set of storage tanks
- $U_p$: set of plants of type $p$
- $UT1$: set of first transportation units
- $UT2$: set of second transportation units
- $W$: set of water intakes

**Indices:**

- $i$: index for polygons
- $j$: index for triangles
- $j_i$: index for triangles belonging to polygon $i$
- $k$: index for coordinates
- $p$: index for plant types
- $s$: index for storage tanks
- $u_p$: index for plants of type $p$
$ut1$  
index for first transportation units

$ut2$  
index for second transportation units

$w$  
index for water intakes

**Parameters:**

$CAP_p$  
capacity of plants of type $p$ [m$^3$/day]

$CAPPS_{p,s}$  
capacity of the pipeline between plant $u_p$ of type $p$ and storage tank $s$ [m$^3$/day]

$CAPPUT_{p,ut2}$  
capacity of the pipeline between plant $u_p$ of type $p$ and transport unit $ut2$ [m$^3$/day]

$CAPUTP_{ut1,p}$  
capacity of the pipeline between transport unit $ut1$ and plant $u_p$ of type $p$ [m$^3$/day]

$CAPUTS_{ut2,s}$  
capacity of the pipeline between transport unit $ut2$ and storage tank $s$ [m$^3$/day]

$CAPWP_{w,p}$  
capacity of the pipeline between water intake $w$ and plant $u_p$ of type $p$ [m$^3$/day]

$CAPWUT_{w,ut1}$  
capacity of the pipeline between water intake $w$ and transportation unit $ut1$ [m$^3$/day]

$Dmins_{s,k}$  
minimal distance measured from storage tank $s$ to the nearest water intake

$E_p$  
efficiency of plants of type $p$ [$/day]

$fcp$  
fixed cost factor of the pipeline [$/day]

$fcup_p$  
fixed cost factor of plants of type $p$ [$/day]

$L21_{j,k}$  
lower bound of lambda of edge $P1P2$ of triangle $j$

$L31_{j,k}$  
lower bound of lambda of edge $P1P3$ of triangle $j$

$NW$  
number of water intakes

$NP$  
number of plant types

$NP1$  
number of plants of type ‘1’

$NP2$  
number of plants of type ‘2’

$NPP_p$  
number of plants of type ‘NP’

$NUT1$  
number of first transport units

$NUT2$  
number of second transport units

$NS$  
number of storage tanks

$P1_{j,k}$  
coordinate $k$ of point 1 of triangle $j$

$P2_{j,k}$  
coordinate $k$ of point 2 of triangle $j$
$P3_{j,k}$ coordinate $k$ of point 3 of triangle $j$

$ps_{s,k}$ coordinate $k$ of storage tank $s$

$pw_{w,k}$ coordinate $k$ of water intake $w$

$Π1_k$ coordinate $k$ of point $Π1$ used in the explanation of the idea of triangles

$Π2_k$ coordinate $k$ of point $Π2$ used in the explanation of the idea of triangles

$Π3_k$ coordinate $k$ of point $Π3$ used in the explanation of the idea of triangles

$U21_j$ upper bound of lambda of edge $Π1Π2$ of triangle $j$

$U31_j$ upper bound of lambda of edge $Π1Π3$ of triangle $j$

$UDS_{s,k}$ maximal distance measured from storage tank $s$

$UDUT1P_{ut1,p,u,p,k}$ maximal allowed distance measured from transport unit $ut1$ to plant $u_p$ of type $p$

$UDPUT2_{p,u,p,ut2,k}$ maximal allowed distance measured from plant $u_p$ of type $p$ to transport unit $ut2$

$UDUT2S_{ut2,s,k}$ maximal allowed distance measured from transport unit $ut2$ to storage tank $s$

$UDWUT1_{w,ut1,k}$ maximal allowed distance measured from water intake $w$ to transport unit $ut1$

$vcp$ variable cost factor of water transported by pipeline [\$/m^3]

$vcpup_p$ variable cost factor of plants of type $p$ [\$/m^3]

$ws_s$ potable water transported to storage tank $s$ [m³/day]

$xplo_{p,u_p}$ lower bound for the x-coordinate of plant $u_p$ of type $p$

$xpup_{p,u_p}$ upper bound for the x-coordinate of plant $u_p$ of type $p$

$xut1lo_{ut1}$ lower bound for the x-coordinate of transport unit $ut1$

$xut1up_{ut1}$ upper bound for the x-coordinate of transport unit $ut1$

$xut2lo_{ut2}$ lower bound for the x-coordinate of transport unit $ut2$

$xut2up_{ut2}$ upper bound for the x-coordinate of transport unit $ut2$

$yplo_{p,u_p}$ lower bound for the y-coordinate of plant $u_p$ of type $p$

$ypup_{p,u_p}$ upper bound for the y-coordinate of plant $u_p$ of type $p$

$yut1lo_{ut1}$ lower bound for the y-coordinate of transport unit $ut1$

$yut1up_{ut1}$ upper bound for the y-coordinate of transport unit $ut1$

$yut2lo_{ut2}$ lower bound for the y-coordinate of transport unit $ut2$

$yut2up_{ut2}$ upper bound for the y-coordinate of transport unit $ut2$
**Variables:**

- $D_{put2,p,u_p,ut2,k}$: distance in coordinate $k$ between plant $u_p$ of type $p$ and transportation unit $ut2$
- $D_{s,p,u_p,s,k}$: distance in coordinate $k$ between plant $u_p$ of type $p$ and storage tank $s$
- $D_{ut1p,ut1,p,u_p,k}$: distance in coordinate $k$ between transportation unit $ut1$ and plant $u_p$ of type $p$
- $D_{ut2s,ut2,s,k}$: distance in coordinate $k$ between transportation unit $ut2$ and storage tank $s$
- $D_{w,w,p,u_p,k}$: distance in coordinate $k$ between seawater intake $w$ and plant $u_p$ of type $p$
- $D_{wut1,w,ut1,k}$: distance in coordinate $k$ between seawater intake $w$ and transportation unit $ut1$
- $f_{costp,p,u_p}$: fixed cost of plant $u_p$ of type $p$ [$$/day]
- $f_{costps,p,u_p,s,k}$: fixed cost of pipeline between plant $u_p$ of type $p$ and storage tank $s$ [$$/day]
- $f_{costp2,ut2,p,u_p}$: fixed cost of pipeline between plant $u_p$ of type $p$ and transport unit $ut2$ [$$/day]
- $f_{costut1p,ut1,p,u_p}$: fixed cost of pipeline between transport unit $ut1$ and plant $u_p$ of type $p$ [$$/day]
- $f_{costut2s,ut2,s,k}$: fixed cost of pipeline between transport unit $ut2$ and storage tank $s$ [$$/day]
- $f_{costwp,w,p,u_p}$: fixed cost of pipeline between water intake $w$ and plant $u_p$ of type $p$ [$$/day]
- $f_{costwut1,w,ut1}$: fixed cost of pipeline between water intake $w$ and transport unit $ut1$ [$$/day]
- $\lambda_{21,p,u_p,j}$: lambda of plant $u_p$ of type $p$ on edge $\overline{P1P2}$ belonging to triangle $j$
- $\lambda_{31,p,u_p,j}$: lambda of plant $u_p$ of type $p$ on edge $\overline{P1P3}$ belonging to triangle $j$
- $\Lambda_{21}$: lambda of $P$ on edge $\overline{P1P2}$ used in the explanation of the idea of triangles
- $\Lambda_{31}$: lambda of $P$ on edge $\overline{P1P3}$ used in the explanation of the idea of triangles
- $obj$: objective function [$$/day]
coordinate $k$ of a general point used in the explanation of the idea of triangles

coordinate $k$ of plant $u_p$ of type $p$

coordinate $k$ of transport unit $ut1$

coordinate $k$ of transport unit $ut2$

seawater produced by water intake $w$ \[ m^3/\text{day} \]

seawater transported to plant $u_p$ of type $p$ \[ m^3/\text{day} \]

seawater transported from seawater intake $w$ to plant $u_p$ of type $p$ \[ m^3/\text{day} \]

potable water produced by plant $u_p$ of type $p$ \[ m^3/\text{day} \]

potable water transported from plant $u_p$ of type $p$ to storage tank $s$ \[ m^3/\text{day} \]

potable water transported from plant $u_p$ of type $p$ to transportation unit $ut2$ \[ m^3/\text{day} \]

seawater transported by transportation unit $ut1$ \[ m^3/\text{day} \]

seawater transported from transportation unit $ut1$ to plant $u_p$ of type $p$ \[ m^3/\text{day} \]

potable water transported by transportation unit $ut2$ \[ m^3/\text{day} \]

potable water transported from transportation unit $ut2$ to storage tank $s$ \[ m^3/\text{day} \]

seawater transported from seawater intake $w$ to transportation unit $ut1$ \[ m^3/\text{day} \]

Logic variables:

true, if there is a pipeline between water intake $w$ and plant $u_p$ of type $p$

true, if plant $u_p$ of type $p$ exists

true, if plant $u_p$ of type $p$ exists within triangle $j$

true, if there is a pipeline between plant $u_p$ of type $p$ and storage tank $s$

true, if there is a pipeline between plant $u_p$ of type $p$ and transportation unit $ut2$

true, if there is a pipeline between transportation unit $ut1$ and plant $u_p$ of type $p$
true, if there is a pipeline between transportation unit \( ut2 \) and storage tank \( s \)

true, if there is a pipeline between water intake \( w \) and transportation unit \( ut1 \)

**Binary variables:**

\[ y_{p,u} \] 1, if plant \( u_p \) of type \( p \) exists; 0 otherwise

\[ y_{pi, p, i} \] 1, if plant \( u_p \) of type \( p \) exists within polygon \( i \); 0 otherwise

\[ y_{pj, p, j} \] 1, if plant \( u_p \) of type \( p \) exists within triangle \( j \); 0 otherwise

\[ y_{ps, p, s} \] 1, if there is a pipeline between plant \( u_p \) of type \( p \) and storage tank \( s \); 0 otherwise

\[ y_{put2, p, u} \] 1, if there is a pipeline between plant \( u_p \) of type \( p \) and transportation unit \( ut2 \); 0 otherwise

\[ y_{ut1, p, u} \] 1, if there is a pipeline between transportation unit \( ut1 \) and plant \( u_p \) of type \( p \); 0 otherwise

\[ y_{ut2, s} \] 1, if there is a pipeline between transportation unit \( ut2 \) and storage tank \( s \); 0 otherwise

\[ y_{wp, w, u} \] 1, if there is a pipeline between water intake \( w \) and plant \( u_p \) of type \( p \); 0 otherwise

\[ y_{wut1, w, ut1} \] 1, if there is a pipeline between water intake \( w \) and transportation unit \( ut1 \); 0 otherwise

**11. Appendix**
11.1. Wilson coefficients used in the examples

Table 33. Wilson coefficients for acetone/methanol mixture

<table>
<thead>
<tr>
<th></th>
<th>acetone</th>
<th>methanol</th>
<th>water</th>
</tr>
</thead>
<tbody>
<tr>
<td>acetone</td>
<td>0</td>
<td>-157.981</td>
<td>393.27</td>
</tr>
<tr>
<td>methanol</td>
<td>592.638</td>
<td>0</td>
<td>-52.6055</td>
</tr>
<tr>
<td>water</td>
<td>1430.000</td>
<td>620.633</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 34. Wilson coefficients for ethanol/water mixture

<table>
<thead>
<tr>
<th></th>
<th>methanol</th>
<th>ethanol</th>
<th>water</th>
</tr>
</thead>
<tbody>
<tr>
<td>methanol</td>
<td>0</td>
<td>1082.600</td>
<td>306.450</td>
</tr>
<tr>
<td>ethanol</td>
<td>-628.570</td>
<td>0</td>
<td>308.270</td>
</tr>
<tr>
<td>water</td>
<td>436.590</td>
<td>949.750</td>
<td>0</td>
</tr>
</tbody>
</table>

11.2. Data of example 3.2

Table 35. Stream data for demonstration example 1

<table>
<thead>
<tr>
<th>Rich streams</th>
<th>G (kmol/hr)</th>
<th>y^s (kmol/kmol inert)</th>
<th>y^t (kmol/kmol inert)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>50</td>
<td>0.01</td>
<td>0.004</td>
</tr>
<tr>
<td>R2</td>
<td>60</td>
<td>0.01</td>
<td>0.005</td>
</tr>
<tr>
<td>R3</td>
<td>40</td>
<td>0.02</td>
<td>0.005</td>
</tr>
<tr>
<td>R4</td>
<td>30</td>
<td>0.02</td>
<td>0.015</td>
</tr>
<tr>
<td>MSAs</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S1</th>
<th>L^(top) (kmol/hr)</th>
<th>x^s (kmol/kmol inert)</th>
<th>x^t (kmol/kmol inert)</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>∞</td>
<td>0.0002</td>
<td>-</td>
<td>26.1</td>
<td>-0.00326</td>
</tr>
</tbody>
</table>

Table 36. Equipment data for demonstration example 1

<table>
<thead>
<tr>
<th>Column capital cost (installed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shell: $12800 \cdot H_r^{0.95} \cdot D^{0.6}$</td>
</tr>
<tr>
<td>Packing (3.8mm Intalox saddles): $1130 \cdot m^3 = 887 \cdot D^2 \cdot H$</td>
</tr>
<tr>
<td>$K_{r,a}$: 0.02 kmol SO2/m3/s</td>
</tr>
<tr>
<td>Inactive height: 3 m</td>
</tr>
</tbody>
</table>
11.3. Visualized explanation of the use of triangles

It is well known from linear algebra that in Figure 93 the following relations hold:

1.)
In point $\Pi_2$, $A21 = 1$ and $A31 = 0$.
In point $\Pi_3$, $A21 = 0$ and $A31 = 1$.
In point $\Pi_4$, $A21 = 0$ and $0 \leq A31 \leq 1$.
In point $\Pi_5$, $0 \leq A21 \leq 1$ and $A31 = 0$.

![Figure 93. Coordinates of specific points](image)

2.) Any point $\Pi_6$ in the plane can be obtained by adding an appropriate $A21 \cdot \overrightarrow{\Pi_1\Pi_2}$ vector to an appropriate point $\Pi_4$ lying somewhere in the $\overrightarrow{\Pi_1\Pi_3}$ edge. The straight line in which the $\overrightarrow{\Pi_1\Pi_3}$ edge lies splits the plane into two parts. If $A21$ is nonnegative then point $\Pi_6$ lies on that side of plane in which point $\Pi_2$ lies, see Figure 94:

![Figure 94. Representation of point $\Pi_6$ inside the triangle with coordinate $A12$.](image)

3.) Similarly, point $\Pi_6$ in the plane can be obtained by adding an appropriate $A31 \cdot \overrightarrow{\Pi_1\Pi_3}$ vector to an appropriate point $\Pi_5$ lying somewhere in the $\overrightarrow{\Pi_1\Pi_2}$ edge. The straight line in which the
edge lies splits the plane into two parts. If $A3I$ is nonnegative then point $Π6$ lies on that side of plane in which point $Π3$ lies, see Figure 95:

![Figure 95. Representation of point $Π6$ inside the triangle with coordinate $A13$.](image)

4.) The point lies accurately on the $Π2Π3$ edge if the sum of the lambda-s is equal to 1. This can be shown as follows.
Let $A2I$ be equal to $1 - A3I$. By equivalent transformations, one can obtain the following series of expressions:

\[ P_k = Π1_k + A2I \cdot (Π2_k - Π1_k) + A3I \cdot (Π3_k - Π1_k) = \]
\[ = Π1_k + (1 - A3I) \cdot (Π2_k - Π1_k) + A3I \cdot (Π3_k - Π1_k) = \]
\[ = Π1_k + Π2_k - Π1_k - A3I \cdot Π2_k + A3I \cdot Π1_k + A3I \cdot Π3_k - A3I \cdot Π1_k = \]
\[ = Π2_k - A3I \cdot Π2_k + A3I \cdot Π3_k = Π2_k + A3I \cdot (Π3_k - Π2_k) \]

As a result, $P_k = Π2_k + A3I \cdot (Π3_k - Π2_k)$, i.e. $P_k$ lies on the $Π2Π3$ edge.

5.) Any point $Π6$ in the plane can be obtained by adding an appropriate $L2I \cdot Π1Π3$ vector to an appropriate point $Π7$ lying somewhere in the $Π2Π3$ edge. The straight line in which the $Π2Π3$ edge lies splits the plane into two parts. If $L2I$ is nonnegative then point $Π6$ lies on that side of plane in which point $Π1$ lies, see Figure 96:
Any point $\Pi_6$ in the plane can be obtained from an appropriate point $\Pi_7$ point (lying somewhere on the $\Pi_2\Pi_3$ edge) by an appropriate $L_{21} \cdot \Pi_3\Pi_1$ vector. The $\Pi_2\Pi_3$ edge splits the plane into two parts. If the sum of the lambdas belonging to $\Pi_6$ point is less than 1, then $L_{21}$ must be negative. If $L_{21}$ is negative, then we are approaching the $\Pi_1$ point, i.e. the $\Pi_6$ point is lying on that side of plane in which the $\Pi_1$ point lies.

A summary is shown in Figure 97. If $A_{21}$ is nonnegative then the point lies in the horizontal hatched area. If $A_{31}$ is nonnegative then the point lies in the vertical hatched area. If their sum is less than 1 then the point lies in the bricked-hatched area.
### 11.4. Data of Problem I

Table 37. Number, capacity and efficiencies of plants of different types

<table>
<thead>
<tr>
<th>$p$</th>
<th>$NPP_p$</th>
<th>$CAP_p$</th>
<th>$E_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>250</td>
<td>0.9</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>500</td>
<td>0.85</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>10</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Table 38. Coordinates of the triangles:

<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>$\Delta$</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x</td>
<td>y</td>
<td>x</td>
<td>y</td>
<td>x</td>
<td>y</td>
<td>x</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>40</td>
<td>0</td>
<td>5</td>
<td>31.27389</td>
<td>40</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>13.40305</td>
<td>20</td>
<td>0</td>
<td>5</td>
<td>25</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>13.40305</td>
<td>20</td>
<td>25</td>
<td>8</td>
<td>42.647</td>
<td>20</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>33</td>
<td>34</td>
<td>25</td>
<td>20</td>
<td>40</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>33</td>
<td>34</td>
<td>40</td>
<td>20</td>
<td>40</td>
<td>34.8936</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>31.27389</td>
<td>40</td>
<td>25</td>
<td>32.9787</td>
<td>80</td>
<td>40</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 39. Coordinates of the water intakes

<table>
<thead>
<tr>
<th>Sea water intake</th>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>6.636364</td>
</tr>
<tr>
<td>2</td>
<td>35</td>
<td>7.545455</td>
</tr>
</tbody>
</table>

Table 40. The demands and coordinates of storage tanks:

<table>
<thead>
<tr>
<th>Sea water intake</th>
<th>Demand m3/day</th>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>250</td>
<td>10</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>300</td>
<td>45</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>350</td>
<td>70</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>80</td>
<td>50</td>
<td>15</td>
</tr>
</tbody>
</table>

### 11.4.1. Actual feasible regions and forbidden connections for Problem I
Table 41. Lower and upper bounds for the x- and y-coordinates of transport unit *ut1*

<table>
<thead>
<tr>
<th></th>
<th>xut1lo</th>
<th>xut1up</th>
<th>yut1lo</th>
<th>yut1up</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>27</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>27</td>
<td>45</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>55</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>70</td>
<td>10</td>
<td>20</td>
</tr>
</tbody>
</table>

Table 42. Lower and upper bounds for the x- and y-coordinates of plant *u_p* of type *p*

<table>
<thead>
<tr>
<th></th>
<th>u_p</th>
<th>xplo</th>
<th>xpup</th>
<th>yplo</th>
<th>ypup</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>10</td>
<td>65</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>65</td>
<td>10</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>65</td>
<td>10</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>10</td>
<td>65</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>65</td>
<td>10</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>10</td>
<td>65</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>65</td>
<td>10</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>65</td>
<td>10</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>65</td>
<td>10</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>65</td>
<td>10</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>65</td>
<td>10</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>65</td>
<td>10</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

Table 43. Lower bounds for the x-coordinates of transport unit *ut2*

<table>
<thead>
<tr>
<th></th>
<th>xut2lo</th>
<th>xut2up</th>
<th>yut2lo</th>
<th>yut2up</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>27</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>27</td>
<td>45</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>55</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>70</td>
<td>10</td>
<td>30</td>
</tr>
</tbody>
</table>
Table 44. Forbidden connections from water intake $w$ to transport unit $ut1$

<table>
<thead>
<tr>
<th>$w$</th>
<th>$ut1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

The upper bound of the binary variable $y_{wut1_{w,ut1}}$ is equal to 0.

Table 45. Forbidden connections from transport unit $ut1$ to plant $u_p$ of type $p$

<table>
<thead>
<tr>
<th>$ut1$</th>
<th>$p$</th>
<th>$u_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

The upper bound of the binary variable $y_{ut1p_{ut1,p,u_p}}$ is equal to 0.

Table 46. Forbidden connections from plant $u_p$ of type $p$ to transport unit $ut2$

<table>
<thead>
<tr>
<th>$p$</th>
<th>$u_p$</th>
<th>$ut2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

The upper bound of the binary variable $y_{put2p_{p,u_p,ut2}}$ is equal to 0.
Table 47. Forbidden connections from transport unit $ut2$ to storage tank $s$

<table>
<thead>
<tr>
<th>$ut2$</th>
<th>$s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>

The upper bound of the binary variable $y_{ut2s_{ut2,s}}$ is equal to 0.

Table 48. Upper bounds for the distance between water intakes and transportation units

<table>
<thead>
<tr>
<th>Upper bound</th>
<th>$k$</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$UDWUT1_{w,ut1,k}$</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>$UDWUT1_{w,ut1,k}$</td>
<td>2</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 49. Upper bounds for the distance between transportation units and plants

<table>
<thead>
<tr>
<th>Upper bound</th>
<th>$k$</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$UDUT1P_{ut1,p,up,k}$</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>$UDUT1P_{ut1,p,up,k}$</td>
<td>2</td>
<td>25</td>
</tr>
</tbody>
</table>

Table 50. Upper bounds for the distance between plants and transportation units

<table>
<thead>
<tr>
<th>Upper bound</th>
<th>$k$</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$UDPUT2_{p,up,ut2,k}$</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>$UDPUT2_{p,up,ut2,k}$</td>
<td>2</td>
<td>25</td>
</tr>
</tbody>
</table>

Table 51. Upper bounds for the distance between transportation units and storage tanks

<table>
<thead>
<tr>
<th>Upper bound</th>
<th>$k$</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$UDUT2S_{ut2,s,k}$</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>$UDUT2S_{ut2,s,k}$</td>
<td>2</td>
<td>20</td>
</tr>
</tbody>
</table>
### 11.5. Data of Problem II

Table 52. Number, capacity and efficiencies of plants of different types

<table>
<thead>
<tr>
<th>p</th>
<th>$NPP_p$</th>
<th>$CAP_p$</th>
<th>$E_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>250</td>
<td>0.9</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>500</td>
<td>0.85</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>10</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Table 53. Triangle coordinates:

<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>$\Delta$</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x</td>
<td>y</td>
<td>x</td>
<td>y</td>
<td>x</td>
<td>y</td>
<td>x</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>150</td>
<td>0</td>
<td>10</td>
<td>35</td>
<td>125</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>35</td>
<td>125</td>
<td>0</td>
<td>10</td>
<td>30</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>33.952</td>
<td>103</td>
<td>30</td>
<td>20</td>
<td>57</td>
<td>103</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>38.132</td>
<td>45</td>
<td>70</td>
<td>25</td>
<td>30</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>38.132</td>
<td>45</td>
<td>70</td>
<td>25</td>
<td>70</td>
<td>45</td>
<td>17</td>
</tr>
<tr>
<td>6</td>
<td>57</td>
<td>103</td>
<td>49.51</td>
<td>80</td>
<td>70</td>
<td>80</td>
<td>18</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>150</td>
<td>35</td>
<td>125</td>
<td>250</td>
<td>150</td>
<td>19</td>
</tr>
<tr>
<td>8</td>
<td>35</td>
<td>125</td>
<td>100</td>
<td>104.35</td>
<td>103.8</td>
<td>133</td>
<td>20</td>
</tr>
<tr>
<td>9</td>
<td>46.658</td>
<td>121.296</td>
<td>70</td>
<td>80</td>
<td>120</td>
<td>98</td>
<td>21</td>
</tr>
<tr>
<td>10</td>
<td>70</td>
<td>80</td>
<td>120</td>
<td>25</td>
<td>100</td>
<td>90.8</td>
<td>22</td>
</tr>
<tr>
<td>11</td>
<td>70</td>
<td>80</td>
<td>70</td>
<td>25</td>
<td>120</td>
<td>25</td>
<td>23</td>
</tr>
<tr>
<td>12</td>
<td>110</td>
<td>57.9</td>
<td>120</td>
<td>25</td>
<td>138.57</td>
<td>70</td>
<td></td>
</tr>
</tbody>
</table>

Table 54. Coordinates of the water intakes

<table>
<thead>
<tr>
<th>Sea water intake</th>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>25</td>
</tr>
</tbody>
</table>
Table 55. The demands and coordinates of storage tanks:

<table>
<thead>
<tr>
<th>Sea water intake</th>
<th>Demand m³/day</th>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>250</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>2</td>
<td>300</td>
<td>40</td>
<td>110</td>
</tr>
<tr>
<td>3</td>
<td>350</td>
<td>120</td>
<td>80</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>120</td>
<td>110</td>
</tr>
<tr>
<td>5</td>
<td>80</td>
<td>210</td>
<td>60</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>200</td>
<td>100</td>
</tr>
</tbody>
</table>

11.5.1. Actual feasible regions and forbidden connections for Problem II

Table 56. Lower and upper bounds for the x- and y-coordinates of transport unit ut1

<table>
<thead>
<tr>
<th>ut1</th>
<th>xut1lo</th>
<th>xut1up</th>
<th>yut1lo</th>
<th>yut1up</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>120</td>
<td>30</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>130</td>
<td>200</td>
<td>30</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 57. Lower and upper bounds for the x- and y-coordinates of plant $u_p$ of type $p$

<table>
<thead>
<tr>
<th>$p$</th>
<th>$u_p$</th>
<th>$x^{plo}$</th>
<th>$x^{pup}$</th>
<th>$y^{plo}$</th>
<th>$y^{pup}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>40</td>
<td>210</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>210</td>
<td>20</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>210</td>
<td>20</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>40</td>
<td>210</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>210</td>
<td>20</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>210</td>
<td>20</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>40</td>
<td>210</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>210</td>
<td>20</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>210</td>
<td>20</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>210</td>
<td>20</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>210</td>
<td>20</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>210</td>
<td>20</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>40</td>
<td>210</td>
<td>20</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>40</td>
<td>210</td>
<td>20</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

Table 58. Lower bounds for the x-coordinates of transport unit $ut2$

<table>
<thead>
<tr>
<th>$ut2$</th>
<th>$x^{ut2lo}$</th>
<th>$x^{ut2up}$</th>
<th>$y^{ut2lo}$</th>
<th>$y^{ut2up}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>110</td>
<td>30</td>
<td>90</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>110</td>
<td>30</td>
<td>90</td>
</tr>
<tr>
<td>3</td>
<td>130</td>
<td>200</td>
<td>30</td>
<td>90</td>
</tr>
<tr>
<td>4</td>
<td>130</td>
<td>200</td>
<td>30</td>
<td>90</td>
</tr>
</tbody>
</table>

Table 59. Forbidden connections from water intake $w$ to transport unit $ut1$

<table>
<thead>
<tr>
<th>$w$</th>
<th>$ut1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 60. Forbidden connections from transport unit $ut1$ to plant $u_p$ of type $p$

<table>
<thead>
<tr>
<th>$ut1$</th>
<th>$p$</th>
<th>$u_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 61. Forbidden connections from plant $u_p$ of type $p$ to transport unit $ut2$

<table>
<thead>
<tr>
<th>$p$</th>
<th>$u_p$</th>
<th>$ut2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 62. Forbidden connections from transport unit $ut2$ to storage tank $s$

<table>
<thead>
<tr>
<th>$ut2$</th>
<th>$s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 63. Upper bounds for the distance between water intakes and transportation units

<table>
<thead>
<tr>
<th>Upper bound</th>
<th>$k$</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$UDWUT1_{w,ut1,k}$</td>
<td>1</td>
<td>70</td>
</tr>
<tr>
<td>$UDWUT1_{w,ut1,k}$</td>
<td>2</td>
<td>70</td>
</tr>
</tbody>
</table>
Table 64. Upper bounds for the distance between transportation units and plants

<table>
<thead>
<tr>
<th>Upper bound</th>
<th>$p$</th>
<th>$u_p$</th>
<th>$k$</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$UDUT1P_{ut1,p,up,k}$</td>
<td>1,2</td>
<td>1</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>$UDUT1P_{ut1,p,up,k}$</td>
<td>1,2</td>
<td>1</td>
<td>2</td>
<td>30</td>
</tr>
<tr>
<td>$UDUT1P_{ut1,p,up,k}$</td>
<td>1,2</td>
<td>2</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>$UDUT1P_{ut1,p,up,k}$</td>
<td>1,2</td>
<td>2</td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td>$UDUT1P_{ut1,p,up,k}$</td>
<td>1,2</td>
<td>3</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>$UDUT1P_{ut1,p,up,k}$</td>
<td>1,2</td>
<td>3</td>
<td>2</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 65. Upper bounds for the distance between plants and transportation units

<table>
<thead>
<tr>
<th>Upper bound</th>
<th>$p$</th>
<th>$u_p$</th>
<th>$k$</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$UDPUT2_{p,up,ut2,k}$</td>
<td>1,2</td>
<td>1</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>$UDPUT2_{p,up,ut2,k}$</td>
<td>1,2</td>
<td>1</td>
<td>2</td>
<td>30</td>
</tr>
<tr>
<td>$UDPUT2_{p,up,ut2,k}$</td>
<td>1,2</td>
<td>2</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>$UDPUT2_{p,up,ut2,k}$</td>
<td>1,2</td>
<td>2</td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td>$UDPUT2_{p,up,ut2,k}$</td>
<td>1,2</td>
<td>3</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>$UDPUT2_{p,up,ut2,k}$</td>
<td>1,2</td>
<td>3</td>
<td>2</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 66. Upper bounds for the distance between transportation units and storage tanks

<table>
<thead>
<tr>
<th>Upper bound</th>
<th>$k$</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$UDUT2S_{ut2,s,k}$</td>
<td>1</td>
<td>80</td>
</tr>
<tr>
<td>$UDUT2S_{ut2,s,k}$</td>
<td>2</td>
<td>50</td>
</tr>
</tbody>
</table>
12. Declaration
I declare that no portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

All sources I used, every part I quoted word for word or redrafted, are listed in the Reference section.


………………………………………...
Abdulfatah M. Emhamed