Delft University of Technology
Faculty of Technology, Policy and Management
Energy and Industry section

Non-deterministic networked infrastructure design of multiple sources and multiple sinks

An agent-based implementation of Ant Colony Optimization

Author: Nguyen Quoc Viet

Date: 30th June 2015

MSc Engineering and Policy Analysis - Master Thesis

Graduation committees:

Chairman:	Prof.dr.ir. P.M. Herder
First supervisor:	dr.ir. P.W. Heijnen
Second supervisor:	dr. A.F. Correljé

Acknowledgements

This thesis would not have been possible without the guidance and the help from the graduation committees and friends at TPM.

Foremost, I would like to express my gratitude to Petra Heijnen for her precise guidance from the problem definition to the model development. It was a rewarding and exciting time to work with her on the same problem and to improve the solution literally day by day through the comparison between my agent-based model and her geometric graph model. I believe it is definitely one of the most effective ways to carry out a research. I would also like to thank her for her detailed comments on earlier drafts of this thesis.

Furthermore, I am deeply grateful to Paulien Herder and Aad Correljé for their valuable comments which have substantially improved the value of this research.

Last, I have to thank Yesh and Sizhe for their helpful suggestions on network flexibility.

Delft, 30th June 2015

Nguyen Quoc Viet.

Abstract

Networked infrastructures such as gas and water pipelines, roads, railroads or power grids provide essential utilities and services to society. Common characteristics of such infrastructures include high capital costs, generally long lifetimes and irreversibility once the construction of such networks have finished. In the design of these networks, the planners face a multitude of challenges ranging from traditional factors such as technical complexities, space-constrained areas to emerging factors such as complex multi-actor contexts and climate change. A chronic challenge is a multi-actor context in which supply capacities from the supplier side, demand capacities from the consumer side and information about the actual commitment of the network participants, who are about to be connected to the network, can remain uncertain for a long time. The uncertainty is especially high when the network involves multiple suppliers and multiple consumers. While deterministic network design ignores these uncertainties, non-deterministic network design takes them into consideration.

The goal of this research is to develop an approach of designing a multiple source (supplier) - multiple sink (consumer) network layout that minimizes the initial investment costs while remaining flexible in its response to future changes of network participants. To this end, firstly an agent-based deterministic modeling method of Ant Colony Optimization was developed, which proved to be feasible in finding cost-minimized network layouts of multiple sources and multiple sinks. Next, the method was extended to a non-deterministic method by embedding flexibility options in order to deal with the uncertainty on network participants. The modeling methods of Ant Colony Optimization were found to be intuitive, extensible and customizable. Based on the modeling outputs, the final design approach can be a supportive decision-making tool in the network planning stage. Future work needs to incorporate more practical criteria required by decision makers into the modeling to bridge the gap between scientific knowledge and decision making process in reality.

Keywords: networked infrastructures, uncertainty, agent-based, Ant Colony Optimization, network flexibility, real options, multiple sources, multiple sinks

Contents

Chapter 1. Introduction

This chapter opens the topic about networked infrastructures in section 1.1 and discusses the challenges in the design of these infrastructures in section 1.2. Next, the network categorization and the target network of this research, multiple source - multiple sink networks, are introduced in section 1.3 and section 1.4 respectively. Also in section 1.4, the specific challenges in the design of multiple source - multiple sink networks are highlighted. The next section presents a literature review on what have been studied on network design in order to identify the knowledge gaps that this research is to fill up. Finally the chapter is ended by the research goal, scope and the thesis outline.

1.1 Networked infrastructures

Networked infrastructures such as gas and water pipelines, roads, railroads or power grids provide essential utilities and services to society. Viewed as large-scale social-technical system (Bijker, et al., 1987), their effects on society are profound. Spatially, since these infrastructure networks are normally fixed, heavy and space-consuming, they have had huge effects on the environment, the spatial organization of the society and the built environment at both macro and micro scales which involve many actors of the society (Feitelson & Salomon, 2000). Economically and temporally, their common characteristics are capital intensive and long-lived, e.g. 40 - 50 years for gas pipeline networks (Ajah & Herder, 2005). Moreover, it is often irreversible once the construction of such a network has been completed (Zhao & Tsend, 2003). For these reasons, the designs are of the paramount importance to the success of these networked infrastructures.

1.2 The challenges in the design of networked infrastructures

The design of infrastructural networks pervades many application contexts due to its significant influences on the full hierarchy of strategic, tactical and operational level decision-making in multistage infrastructure systems (Ukkusuri, et al., 2007). The manifestation of the network design varies in some extent across different infrastructure sectors. For instance, in transportation, network design mainly focuses on the selection of the arcs in the network (Ukkusuri, et al., 2007) while in gas infrastructures, it is necessary to concern at the same time the supply, demand, the pipeline routes and the station locations (Herder & Pulles, 2011).

Social, technological and environmental changes have brought many difficulties to the network planners. Within a limited time of planning, they face a multitude of challenges ranging from traditional factors such as technical difficulty and cost of the projects, organizational changes, space-constrained areas, regional cost differentiation to emerging factors such as complex multi-actor contexts and climate change (Heijnen, et al., 2013). These challenges are more visible in the initial planning[1] and routing stage than anywhere else. An exact choice of the route will greatly reduce technical difficulty and the initial cost, improve the return on investments and mitigate possible social and environmental impacts. One can imagine the big difference between laying a short pipeline through a residential area and a long meandering one through rivers or grasslands. The focus of this study is on the initial planning and routing issues in infrastructural network design and is referred as the network design hereinafter.

---

[1] Initial planning here is the overview of who will possibly participate in the network and their roles.

Among the aforementioned challenges, a multi-actor context is a chronic challenge which makes the initial planning and routing infrastructures an extremely difficult task. The main involved actors in the beginning are the network planner, the suppliers and the consumers[2]. The network planner is the business initiator who works with both suppliers and consumers to plan the network. In reality, at the planning stage, information related to the major actors is often unknown. Namely, supply capacities from the supplier side, demand capacities from the consumer side, and information about the actual commitment of relevant parties, who are about to be connected to the network, can remain uncertain for a long time. To illustrate, the current CO_2 network in the Netherlands from industrial plants to greenhouses was initiated by the company OCAP[3] in 2001. As the network planner, OCAP signed 30-year contracts with two suppliers from the Rotterdam industrial zone: Shell Pernis refinery plant and Abengoa bio-ethanol plant. The network supplies 45 kilotonnes of pure CO_2 per year to its consumers, who are (more than 500) greenhouses in different horticulture areas in the country such as Westland, B-driehoek and Zuidplaspolder. At the start of planning the pipeline network, a team of 16 people from OCAP personally visited every greenhouse in these areas for 6 months to investigate the demand, and also examined the CO_2 treatment systems at the emitter sites to learn about the supply. "We only could calculate the rough demand and supply before coming back to the farmers to convince them to join the network. Some postponed their decisions at the time but eventually became our consumers years later" (Veenstra & Limbeek, 2014).

In regard to those uncertainties in network planning, the author classified network design into two approaches as follows.

- *Deterministic network design.*
  The presence of uncertainties is ignored. The design goal is to locate an optimal network that satisfies a fixed set of inputs including the number and the location of connectors and the supply/demand capacities they require from the network. The common criterion is the total building cost of the network.

- *Non-deterministic network design.*
  The uncertainties are recognized and taken into account. As a result of these uncertainties, the need for an in-built flexibility into the infrastructures becomes highly imperative (Ajah & Herder, 2005). The common design goal is a flexible network that performs well when the uncertainty is high. It enables the network planners to gain from upside opportunities when the future turns out to be favorable and to minimize downside risks when the future turns out to be unfavorable (Ajah & Herder, 2005). Flexibility from real options perspective is a practical approach in infrastructure network design to deal with the uncertainties.

Nevertheless, the uncertainty degree in the design in multi-actor contexts can widely vary depending on different types of networks with different number of actors involved. The next section introduces the network categorization based on the number of major actors participating in the network.

---

[2] Consumers here refers to either big-scale consumers or a group of small-scale consumers in an area

[3] OCAP comes from Organic Carbondioxide for Assimilation of Plants, available at http://www.ocap.nl/

1.3 Network categorization

In network design, Feitelson & Salomon (2000) identified three important dimensions: node, link and temporality. Nodes represent the physical infrastructure facilities of the suppliers and the consumers, while links are the physical connections between the nodes and temporality refers to the use of infrastructures and the coordination among suppliers and consumers. When the number of nodes becomes higher, the design becomes an increasingly complicated task due to the fact that it requires more links in the network and more efforts in arranging coordination among the actors.

Placed upon the number of nodes in the network, networked infrastructures can be categorized into three groups: (1) one source - one sink networks, (2) one source - multiple sinks / multiple sources - one sink networks and (3) multiple source - multiple sink networks. Hereinafter, a source represents a supplier and a sink represents a consumer. For the sake of simplicity, examples given in this study will be mainly drawn from CO_2 pipeline networks.

- *Type-1 network:* one source - one sink or *1 : 1* networks. They are simple networks such as one big CO_2 emission source connected to one big CO_2 injection site. The uncertainty comes solely from demand and supply.
- *Type-2 network:* one source - multiple sinks or *1 : n* networks and multiple sources - one sink or *n : 1* networks. These two types of networks share the same conceptually geometric topology. For example, one central CO_2 supplier is connected to multiple greenhouses. Another example is multiple small CO_2 sources are connected to one big sink of a beverage factory. The uncertainty comes from both demand/supply and additionally network participants.
- *Type-3 network:* multiple sources - multiple sinks or *n : n* networks. For instance, CO_2 emission sources (*sources*) and geologic injection reservoirs or greenhouses (*sinks*) are spatially dispersed; an efficient pipe network connecting all of them will substantially reduce the transportation cost (Wiley, et al., 2013). Although there is a currently gradual uptake of carbon capture and storage (CCS), it is possible that *not all* the CO_2 sources will implement CCS at the same time. Therefore whether or not they will join the network and when they will join the network are not known in the planning stage. Moreover, along with the CO_2 market's expansion, there will be probably more CO_2 consumers joining the network in the future. In other words, there is a high uncertainty not only about the demand and supply but about the network participants as well. The development of this type of networks and the specific challenges in the design of these networks are elaborated in details in the next section.

1.4 Multiple source - multiple sink networks

Recent development in networked infrastructures shows that there appears to be a growing number of multiple source - multiple sink networks (Veltin & Belfroid, 2013). Below are the discussions of three representative examples of such networks.

- *The connection of off-shore small fields to Dutch gas grid*
  The Groningen gas field is the largest on-shore field in the Europe. With the aim to conserve the reserve of the Groningen field for as long as possible, Dutch government has pursued its *small fields* policy from 1974. The policy is to bring as many small fields, mostly off-shore, as possible into production so that the Groningen field is able to play the strategic *swing supplier*

role to the Dutch gas system. Gasunie (a leading European gas infrastructure company) and NAM (Nederlandse Aardolie Maatschappij, the exploration and production company) have great expectations of obtaining necessary licenses for the exploration and production of these discovered fields. Several small fields such as Grijpskerk and Waddenzee were successfully added to the Dutch gas supply network. However, because the reserves of new fields in the Dutch sector on the North Sea (figure 1) are getting smaller and smaller, the exploitation costs and the profits are critically dependent on the proximity of the fields to the transport infrastructure (pipelines) (Correljé, et al., 2003). Due to the long and complex licensing procedures, the planning of new gas pipelines to connect the new fields to the multiple entry points on the mainland is under a high uncertainty. While the physical proximity plays a critical role, the non-deterministic design of this multiple source - multiple sink network must provide a flexibility to deal with the uncertain connecting locations, timings and/or required capacities.

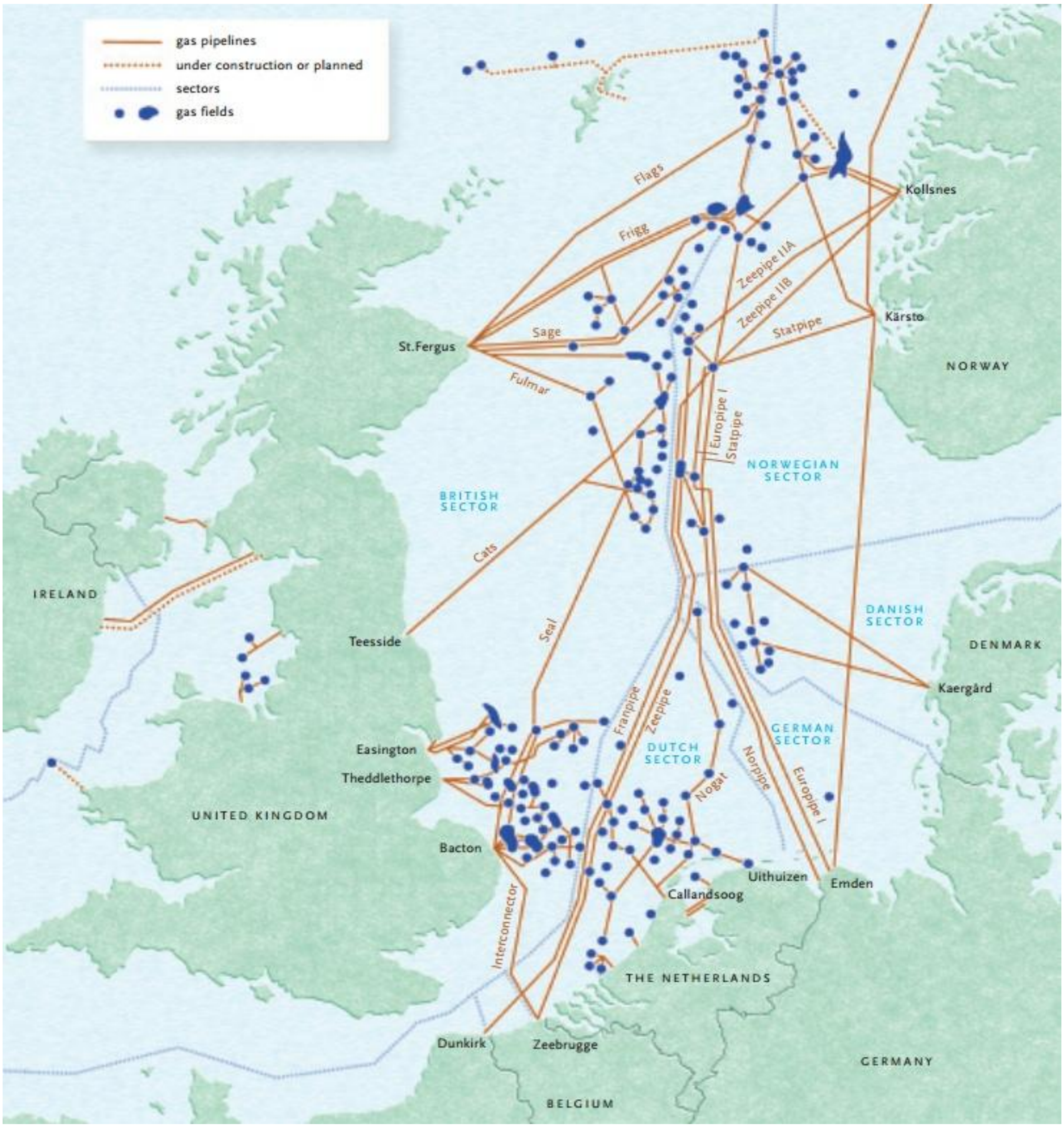


Figure 1. Gas production and transport infrastructure on the Continental Shelf (Correljé, et al., 2003, p. 101)

- *The penetration of renewable energy sources.*
  The overall future demand for energy is less uncertain than the demand for specific energy carriers, particularly for renewable energy (Herder, et al., 2011). By 2025, the electricity demand in Europe is forecasted to be 15% higher than it was in 2010. Besides, the CO_2 emissions decrease due to the higher CO_2 price and CO_2 emissions eliminating policies by the European Commission lead to the growing penetration of hydro, wind and solar energy as main renewable sources into the electricity networks (Martinez-Anido, 2013). The integration of these new sources and the widespread deployment of distributed energy resources require the existing grid networks undergo fundamental changes on the electricity generation and transmission. Thus there is a need to re-design and upgrade the multiple source - multiple sink networks in order to prepare for these changes. Since not all the sources enter the networks at the same time and their supply capacities are not certainly known, the re-designs need to be flexible to reduce the cost of future connections.

- *CO2 network expansion in the Netherlands.*
  From 2004 to 2014, a research program on carbon capture and storage (CCS) by the Dutch government called CATO[4] explored how CCS chains can be built on the local scale. It demonstrated that the costs will be substantially reduced when different companies (sources) in an industrial park that emit CO_2 join forces to design an optimal pipe network for CO_2 transportation to sharing CO_2 storage sites or to CO_2 separation units (sinks) (Ramirez, et al., 2014). Thereby it is an example of connecting multiple sources to multiple sinks. Veenstra & Limbeek from the company OCAP also stated that the present CO_2 transportation pipelines need to be extended because of the demand increase from new consumers. In order to meet this demand increase, the current two unique sources, Shell and Abengoa, will have to increase the supply capacities in near future and OCAP has to expand its business with new potential high-quality CO_2 emission sources (2014).

The growth of these multiple source - multiple sink networks leads to a great demand for effective design approaches, which should be comprehensive and intuitive, to be the guidance tools in decision-making process between the multiple involved actors on planning such networks. Yet the task is very *challenging* as the networks consist of multiple sources and multiple sinks.

The fundamental *difference* regarding the design of multiple source networks and one source networks (including type-1 and type-2 networks) is the starting point of the design (Correljé, 2015). The design of one source networks has to map the unique source to multiple sinks whereas theoretically in the design of multiple source networks, there are numerous choices of how to connect a sink to a source or to a group of sources. This makes the routing issue more complicated due to the much larger routing solutions space. In details, the numerous options to connect a sink to the network bring a multitude of possible network topologies. In addition to the dimension of connecting options, the *order* of connecting a sink or a source to the network can lead to entirely different layouts. It is clear that different network layouts present different initial building costs and likely different degrees of network flexibility. Selecting among those layouts one which is the best combination between cost and flexibility is the main challenge for

---

[4] CATO is an abbreviation for **C**O_2 **A**fvang, **T**ransport en **O**pslag (CO_2 capture, transport and storage), available at http://www.co2-cato.org/

the design. In most of the cases, defining the whole space of the possible network layouts is already a next-to-impossible task, especially when the number of sources and sinks are high.

The uncertainty is much higher in the multiple source - multiple sink networks compared to one source networks. The fact is that the uncertainty not only rises from the sink side but also from the source side. The penetration of new sources to a network requires big modifications of the network configuration. Besides, connecting the sinks in the network to an uncertain source is not an elaborated choice. For instance, if the supply capacity of the source widely fluctuates, the links between that source and the sinks are considered as uncertain links on the capacity as well.

Nonetheless, the *opportunity* still exists for the design. Because a sink can be connected to any source(s), the network planners can strategically set up the priority to connect the sinks to the network. For instance, the rule is set as at first connecting the highest demand sink to the biggest capacity source regardless of their distance. Another example is to first connect a pair of a sink and a source that has the longest distance among all the pairs of a sink and a source. These established priorities lead to various directions in the solutions space of the network layouts (Heijnen, 2014).

Accepted the need and challenges of designing multiple source - multiple sink networks, this study is devoted to developing a design approach which is able to solve the routing challenges and intuitive to support the decision making process. To this end, first of all an extensive *literature review* on deterministic and non-deterministic network design was executed to identify the available research methodologies, relevant findings and knowledge gaps in the field.

1.5 Literature review on the network design

Available literature in network design is categorized into three groups according to their target infrastructure networks.

1.5.1 The network design of type-1 networks

A single path network connecting a static source and a static sink is the simplest scenario (Lin, et al., 2013) and relatively easy to model (Wiley, et al., 2013). It is understandable that there seems no scientific papers wholly dedicated to studying the deterministic network design of this type of network. The non-deterministic network design on the uncertainty of demand and supply received more attention. The economic benefit of oversized design and parallel design in pipeline networks was studied qualitatively and quantitatively by Morbee et al. (2011) and by Wiley et al. (2013) respectively for example. However, due to the simplicity of the network, the routing problem was acceptably ignored in these two papers.

1.5.2 The network design of type-2 networks

In the deterministic network design of these 1 : n and n : 1 networks, most of the studies share the following aspects and requirements:

- The nodes have fixed locations and fixed supply/demand capacities
- The building costs mainly depend on the length and the capacities of the links
- The networks are to be built in space-constrained areas or areas with limitations such as rivers, mountains.

There are two fundamental approaches, namely Top-Down and Bottom-up. *Top-Down* approach relies on a global optimization algorithm that regularly requires complete information about the system to locate an optimal network. Several examples of Top-Down approach that have a high number of citations are the one by Kabirian & Hemmati (2007) for natural gas transmission networks, the one by Marcoulaki et al. (2012) for pipeline systems and the one by Thomas & Weng (2006) for cost flow-dependent networks. Graph theory is applied the most for Top-Down approach (Heijnen, et al., 2014). *Bottom-Up* approach uses distributed agents/entities with their local information to achieve a good enough network. The dominant algorithm is Ant Colony Optimization. One good example is the paper by Maier et al. (2003), which uses Ant Colony Optimization to design a water distribution system within limited areas by existing buildings. There are several attempts to combine these two approaches to solve network design problems such as the research by Liu & Wang (2012) and by Hu et al. (2006). Although these studies proposed the integrated approach, they do not address explicitly if the integrated approach generates better solutions or if it does efficiently utilize the differences or similarities of the Top-Down and Bottom-Up approaches.

Regardless the approach, those studies in some extent succeeded in fulfilling the requirements mentioned earlier. Heijnen et al. (2014) designed two very different algorithms for Top-Down and Bottom-Up approaches using graph theory and agent-based modeling respectively on a same set of generic deterministic problems to compare the disadvantages and the advantages of these two approaches. The results indicate that in the one hand, the Top-Down approach is able to guarantee a solution and is computationally efficient. In the other hand, the Bottom-Up approach can be easily extended and is very intuitive regarding the problem description. Consequently, Bottom-Up is a promising approach when there are more than one source in the network. Besides the aforementioned paper, to the author's knowledge there are no other studies directly comparing the performance and solution quality of the available algorithms for each approach.

About the non-deterministic network design, many scientific papers are available when it comes to the *uncertainty about demand and supply* in type-2 networks. The general aim of these papers is to discover the optimal sizes of the paths in the network with taking into account the extra capacity and/or the fluctuation of demand in future. Representative examples are the papers by Escudero et al. (1999) and Goel et al. (2006). The common limitation of the frameworks in these papers is that they do not explain how the network routing problem is solved considering the presence of the uncertainty but only focus on identifying the optimal capacity. Capturing the routing issue and the uncertainty on the capacities, Heijnen et al. (2011) employed Gilbert networks to provide a robust network topology that can minimize the network planner's regret.

The networks in this category involve only one source, thus the *uncertainty of network participants* mainly comes from the anticipating connections from future potential sinks. Despite the fact that it is intuitive to think that the new connection requests by future potential participants will significantly change the network topology, the literature search shows that there is a very limited number of literature working on this problem. Two important examples of this category are found and to be discussed in details below.

- In the research by Heijnen et al. (2013), a group of sinks with different participating probabilities are to be connected to a single source. Due to this uncertainty, the decision-making process may result in a stalemate because all involved parties are afraid of doing wrong investments. Rather than an optimal solution, a robust one that stands well in many unexpected situations is desired. The proposed approach, which is a combination of graph theory and concepts of exploratory modeling, allows the network planners to maximize the worth of the networks under the uncertainty of participants. The approach was tested on two energy networks in the Netherlands and specific high expected worth networks were identified. However it does not guarantee that the participants with high probabilities will be connected in the network which has the maximum worth. This may bring an unwanted disappointment from these participants and harm the long-term relationships. An extension of the approach is necessary to insure a set of mostly fixed participants to be connected in the network that maximizes the network's worth.

- Melese et al. (2014) proposed a different approach to solve the same problem as the one in the paper mention right above. It is a simulation based on a combination of Monte Carlo simulation and graph theory. The concept of real options and network flexibility are embedded in the approach. The results show that designing with architectural flexibility can significantly improve the value of the network compared to the deterministic network design approach.

1.5.3 The network design of type-3 network

On the *deterministic network design* of type-3 networks of multiple source and multiple sinks, all found ones are studies on CO_2 transportation networks. Weihs et al. (2011) aimed for the design of a deterministic CO_2 network of multiple emission sources and multiple injection sites with steady-state optimization; however, the objective is not the optimal network topology but the cost per tonne of CO_2 avoided solely based on the optimal technical options of pipelines and compression strategies. In another paper, a Top-Down approach is applied by Middleton & Bielicki (2009) for a scalable infrastructure model for carbon capture and storage. It offers a comprehensive deterministic routing approach which takes many technical and social criteria into consideration.

The only example on the *non-deterministic network design* of multiple source - multiple sink networks is the research by Kazmierczak et al. (2009). In this study, the uncertainties of CO_2 supply from sources as well as storage capacity of sinks are reflected by the different cost ranges for pipeline network. The proposed algorithm is concluded to be not limited to CO_2 networks only but also applicable to other type of commodities in pipeline networks.

Up to present, the uncertainty of network participants in the design of multiple source - multiple sink networks has not been studied by any research.

1.5.4 Literature review summary

Table 1. Summary of the literature review

Type of networks		Type-1 of *1 : 1*	Type-2 of *1 : n* or *n : 1*	Type-3 of *n : n*
Typical network problems		There are only two participants in the network so major problems relates to the uncertainty of demand and supply.	Routing problems in space-constrained areas, no-go areas or areas with obstacles such as rivers, buildings. Existing uncertainties: demand and supply, network participants	Available papers deal with the same problems in the type-2 networks.
Representative examples	Deterministic network design	(Lin, et al., 2013)	(Heijnen, et al., 2014) (Marcoulaki, et al., 2012) (Liu & Wang, 2012) (Kabirian & Hemmati, 2007) (Thomas & Weng, 2006) (Hu, et al., 2006) (Maier, et al., 2003)	(Weihs, et al., 2011) (Middleton & Bielicki, 2009)
	Non-deterministic network design	(Wiley, et al., 2013), (Morbee, et al., 2011)	Uncertainty of demand/supply: (Heijnen, et al., 2011) (Goel, et al., 2006) (Escudero, et al., 1999) Uncertainty of network participants: (Melese, et al., 2014) (Heijnen, et al., 2013)	Uncertainty of demand/supply: (Kazmierczak, et al., 2009) Uncertainty of network participants: not available
Methodologies and theories		Optimal technical choice based on scenario analysis	Top-Down approach and Bottom-Up approach. Dominant theories: Graph theory and Ant Colony Optimization	Top-Down approach using Graph theory
Major knowledge gaps		Sufficiently studied	(a) Direct comparisons of different approaches (b) Embedding uncertainties in routing algorithm	(c) Deterministic design framework for generic networked infrastructures (d) Uncertainty of demand/supply and network participants

Table 1 briefly summarizes the results of the literature review including the major knowledge gaps. In the multi-actor context, two major groups of actors directly involved in the network planning stage are suppliers and consumers. First, the greater the number of the connectors is, the more challenging the task of network design is (Feitelson & Salomon, 2000) and the higher the variety of network topologies is (Heijnen, 2014). Second, as discussed in section 1.4, reality shows multiple source - multiple sink networks are in a growing stage especially in energy infrastructures (Veltin & Belfroid, 2013); further research is demanded to meet the growing pace of this type of networks. Last, an efficient and effective design approach that is able to deal with the challenges in multiple source - multiple sink network is highly expectable to be applicable to the other two types of infrastructure networks. For these reasons, this research is entirely dedicated to the design of *multiple source - multiple sink networks*.

Regarding the non-deterministic network design of type-3 networks, the focuses were made only on the uncertainty of demand and supply. Firstly, it is highly possible that the available solutions for the uncertainty of demand and supply in the type-2 networks are also applicable to the type-3 networks (Heijnen, 2014). Secondly, due to the fundamental difference on the network design explained in section 1.4 between type-2 and type-3 networks, the available approaches on *1 : n* networks may not be applicable to *n : n* networks when it comes to the uncertainty of network participants. Lastly, there are no studies on the uncertainty of network participants in the design of *n : n* networks. For these reasons, within the limited available time, this research will focus only on the *uncertainty of network participants* in the design of multiple source - multiple sink networks. Based on these choices, the research goal, scope and the thesis outline will be addressed in the next section.

1.6 Research goal, scope and thesis outline

1.6.1 Research goal

The best value of a networked infrastructure is achieved when the facilities are available at the right places when needed. Likewise, the central design problem in three examples of *n : n* networks introduced in section 1.4 is the difficulty in knowing what to build, where to build it and at what time due to the uncertainty related to the participants. The guide for this problem, as indicated in the literature review, is not available at this moment. Providing a sound and implementable solution for the problem to fill up the knowledge gaps and support the network planners is the final objective of this research. The research context is described as follows.

The target network of the study is a generic networked infrastructure that involves a set of multiple suppliers and multiple consumers that are to be connected by the network. Among them, there is a group of *certain* (fixed) participants with known locations, supply and demand capacities during the project life. The challenges come from another group consisting of *uncertain* participants who have not yet committed or need to be convinced to partake in the network development. In the start of the project, the network planner wants to connect the certain participants according to their demand and supply capacity to kick-off the business. However, looking from the non-deterministic network design perspective, the network planner also wants to prepare for undesirable futures and increase the value of this large-scale, long-lasting and economically irreversible project by accepting to invest more capital in the initial investment for built-in flexibilities in the design. With the aim to support the network planner, the research goal is to

In practice, there are numerous reasons to pick a deterministic (rigid) design over a non-deterministic (flexible) design, e.g. the problem of weak dissemination of developed non-deterministic approaches and the constraints on time, money and the quantity of personnel who can understand the essence of the approaches (Herder, et al., 2011). Furthermore, while the decision-making on a deterministic design is either go or no-go, a non-deterministic design requires constant monitoring and review of the flexibility options in the design. It would be of the best convenience to show both types of designs to the decision makers. As a result, the proposed approach in this study should include both the deterministic and non-deterministic designs to comprehensively equip the network planners so that they can stay convincing in any circumstance in the decision-making process. The non-deterministic design should explicitly depict the extra steps to take from the deterministic design to demonstrate the trade-off between the initial investment cost and the built-in flexibilities. In order to reach this, this research first addresses the following sub-questions in their orders.

Methodology

The literature review identified available methodologies and theories to solve the challenges of the network design.

1. What is the most appropriate methodology for the design problem in multiple source - multiple sink networks?

The deterministic design

The network planner first wants to start the business with all the certain participants, then

2. How to develop a deterministic design method of multiple source - multiples sink networks to connect these participants?

The non-deterministic design

The future is inevitably uncertain. If the network planner does not take a range of possible future scenarios into account from the start, the designs would be misled.

3. What are the practical scenarios regarding the uncertainty of network participants in multiple source - multiple sink networks?

With the identified uncertainties in sub-question 3,

4. What are the appropriate flexibility options in the network that can handle and/or exploit those uncertainties?

The next question is

5. How to incorporate the selected flexibility options into the deterministic design method to develop a non-deterministic design method to cope with the uncertainties?

Incorporating the flexibility options into the design is costly. There is always a trade-off between the investment costs, the flexibility degree and other important factors.

6. How to evaluate the design outputs from the non-deterministic network design method?

1.6.2 Research scope

Working on a generic infrastructural network, the research is placed between the strategic level and tactical level decision-making on the planning. The criteria and requirements for the operational level of the networks do not fall into the scope of this study. In other words, the target is the architectural design and architectural flexibility, not including the operational design and operational flexibility. Nevertheless, the operational flexibility is achieved only when the architectural flexibility is available (Melese, et al., 2014). The output from this design approach will be the input for further modifications in order to obtain the design on the operational level. For example, obviously gas pipe networks and water pipe networks are entirely differently operated, yet from the strategic and tactical level, they share a similar design framework.

1.6.3 Thesis outline

The next chapter will discuss the research methodology to give a direct answer to the sub-question 1 and introduce how to answer the rest of the sub-questions. Chapter 3 deals with the sub-question 2 about the development of the deterministic design model. Next, in chapter 4 the answers for sub-question 3 and sub-question 4 about the network participant uncertainty and flexibility options, respectively are presented. Chapter 5 describes the non-deterministic design method. The output of the model is evaluated in the end of chapter 5. Finally, chapter 6 presents the final network design approach, conclusions and discussions on future work. It ends with reflections on the limitation of the research.

Chapter 2. Research methodology

This chapter provides the answer for the first sub-question about the methodology and the brief explanations for the other sub-questions.

- *What is the most appropriate methodology for the design problem in multiple source - multiple sink networks?*
  The literature review on the network design reveals that there are two dominant theories, which are Geometric Graph theory and Ant Colony Optimization (ACO) used in Top-Down approaches and Bottom-Up approaches respectively. The convincing comparative study on the network design for type-2 one source - multiple sink networks by Heijnen et al. (2014) concluded that these two algorithms are hardly different in terms of the quality of their solutions. Yet both algorithms generally have discrepant advantages and disadvantages on which the practitioners can rely to make decisions on choosing the appropriate one. Geometric Graph theory can guarantee a solution and usually efficient in computing time, but it poses the difficulty in encoding the problem. Although the model of the ACO algorithms is slightly worse in the computational performance, it has significant advantage in the ease of intuitive problem encoding and future extensibility. By extensibility, it means adding additional nodes (sources and/or sinks) and links to the modeling world. A challenge of practical implementation of the available models is the requirement of mathematical skills to use the models comfortably and knowledgeably that practitioners, even many academics do not have (Lander & Pinches, 1998). Upon this remark, the ACO appears to be more appropriate to this study than the Geometric Graph theory. Besides that, the ACO is commonly used for various routing problems in engineering domains, and therefore it is applicable for finding the good network layout of multiple sources - multiple sinks. Based on these reasons, this research adopted the ACO algorithm to solve the problem in the network design of multiple sources and multiple sinks.

- *How to develop a deterministic design method of multiple source - multiples sink networks?*
  In the comparative study, Heijnen et al. (2014) successfully built two models to solve the deterministic design problem for *1 : n* networks, namely the 1 : n Geometric Graph model (1:n GG model) and the 1 : n Ant Colony Optimization model (1:n ACO model). Adopting the ACO algorithm, this research extended the *1:n ACO model* to a deterministic design model for *n : n* networks, which is called as *deterministic n:n ACO model* as displayed in figure 2 below. The deterministic n:n ACO model in this study generates deterministic network layouts connecting multiple sources to multiple sinks. The extension required major changes in the original ACO algorithm to add more sources into the modeling. The implementation of multiple sources required the model's agents to behave differently from their behaviors in the 1:n ACO model. Details are available in chapter 3.

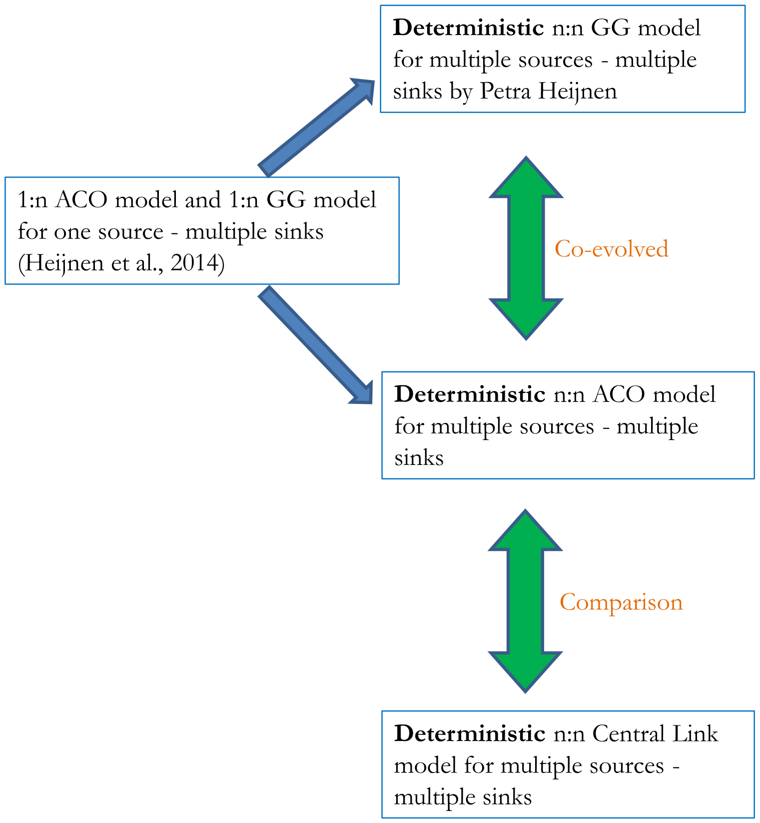


Figure 2. Deterministic model development

Nevertheless, the performance of the deterministic n:n ACO model needs to be validated. As shown in figure 2, the deterministic n:n ACO model was co-evolved with a Geometric Graph (GG) deterministic model developed by Petra Heijnen for multiple source - multiple sink networks. The *deterministic n:n GG model* is an extension of the 1:n GG model for one source - multiple sinks networks by Heijnen et al. (2014). The extended GG model provides a good theoretical benchmark to validate and evaluate the deterministic ACO model because: first, Heijnen et al. (2014) concluded that the GG method's results are slightly better than the ACO method's results and second, at present there are no available optimal deterministic models of multiple source - multiple sink networks for output's comparison.

Given the fact that these two models are theoretically developed in the lack of proven optimal deterministic models, the deterministic n:n ACO model needs to be validated by a practical benchmark as well, as shown in figure 2. Interviews with experienced network planners (Veenstra & Limbeek, 2014) and literature (Berghout, 2014) reveal that in reality the Central Link approach is one of the most common practices in infrastructure network design. The approach will be discussed in details in Chapter 3. While the deterministic n:n GG model makes a good theoretical benchmark for the comparison, the Central Link approach makes a good practical one. Therefore in this research, besides the agent-based *deterministic n:n ACO model*, another model that simulates the Central Link approach - *deterministic n:n Central Link model* - was developed for model validation and comparison purpose.

- *What are the practical scenarios regarding the uncertainty of network participants in multiple source - multiple sink networks?* and *what are the appropriate flexibility options in the network that can handle and/or exploit those uncertainties?*

  The third sub-question is partly answered by the literature review not only in the field of infrastructure network design but also in the field of economy of infrastructures. To sufficiently answer this sub-question, knowledge of real-world practices is necessary. To achieve this, interviews with experts in the field were executed. They are researchers in the field at Delft University of Technology and pipelines network planners at OCAP. Details can be found in chapter 4.

  Sub-question 4 relates to the theory of real options and flexibility in engineering design. A literature review in this field was carried out. The representative sources are (de Neufville & Scholtes, 2011), (Ajah & Herder, 2005), (Gil, 2007) and (Herder, et al., 2011).

- *How to incorporate the selected flexibility options into the deterministic design method to develop a non-deterministic design method under the uncertainty?*

  The deterministic n:n ACO model needs to be extended to a *non-deterministic n:n ACO model* which embeds the flexibility options to deal with the uncertainty of network participants. The extension is mainly placed on new behaviors of the model agents. Details are available in chapter 5.

  For the validation of the non-deterministic n:n ACO model, the literature review indicates that there is no theoretical non-deterministic network design model available. The expert interviews reveal that with the Central Link approach, the network planners always take an estimated extra capacity in the design of the central link to prepare for future demand increases or new connections (Veenstra & Limbeek, 2014). Since the Central Link approach is actually an implementation of flexibility in network design, the deterministic n:n Central Link model is also further extended to the *non-deterministic n:n Central Link model* to serve as a practical base for the model validation as shown in figure 3 below.

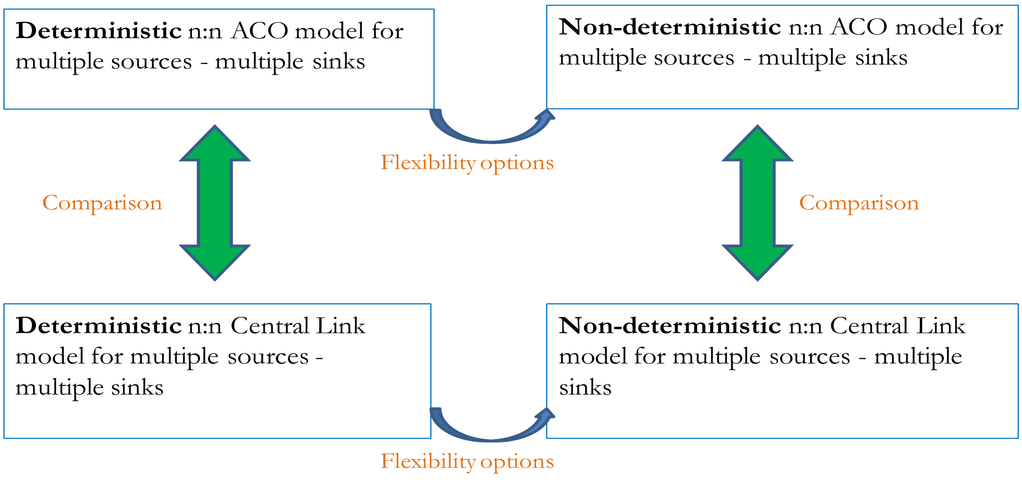


Figure 3. Non-deterministic model development

- *How to evaluate the design outputs from the non-deterministic network design model?*
  As indicated in the research goal, the evaluation is mainly on the initial investment cost. However, additional factors such as multi-actor decision making perspective should be addressed. Details are available in chapter 5.

Summary of chapter 2

The model development flow starts with the deterministic n:n ACO model and ends with the non-deterministic n:n ACO model. The incorporated flexibility options are the distinguishing points of these two models. Table 2 below summarizes all the models in this research.

Table 2. Models summary

Model	Explanation	Role
1:n ACO	Deterministic network design model for *1 : n* networks	The base to develop the deterministic n:n ACO model
1:n GG		The base to develop the deterministic n:n GG model
Deterministic n:n ACO	Deterministic network design model for *n : n* networks	The base to develop the non-deterministic n:n ACO model
Deterministic n:n GG		Theoretical benchmark to validate the deterministic n:n ACO model
Deterministic n:n Central-Link		Practical benchmark to validate the deterministic n:n ACO model
Non-deterministic n:n ACO	Non-deterministic network design model for *n : n* networks	Provide final designs to the network planners
Non-deterministic n:n Central-Link		Practical benchmark to validate the non-deterministic n:n ACO model

Chapter 3. Development of the deterministic n:n ACO model

This chapter is dedicated to describing the deterministic n:n ACO model for multiple source - multiple sink networks. In section 3.1, the original 1:n ACO model of *1 : n* networks is introduced. Section 3.2 presents the conceptual design model for *n : n* networks where the differences between the original 1:n ACO model and the deterministic n:n ACO model are made clear. Section 3.3 discusses important parameters of the deterministic n:n ACO model. The model is validated by the deterministic n:n GG model by comparing their performances on 100 randomly generated examples in section 3.4. Next, the Central Link approach and the deterministic n:n Central-Link model are introduced in section 3.5. The chapter ends by section 3.6 with the model validation by the deterministic n:n Central-Link model.

3.1 The original 1:n ACO model for *1 : n* networks

Heijnen et al. (2014) developed the original 1:n ACO model to solve the routing problem in the network design of one source - multiple sink networks. In the study, the challenges include the limitations on the possible routing of network links due to *no-go* regions which are existing buildings, obstacles such as rivers or mountains or zoning rules. The model is an agent-based implementation of an Ant Colony Optimization in Netlogo. Below is the summary of the model conceptualization which forms the base for the extension in the next sections.

The problem is formalized as finding a network that minimizes the total building cost and satisfies the following criteria

- The network connects all sinks to the source
- All the network edges must go through the *allowed* regions
- The capacities of the edges satisfy the demands of the sinks assumed that the source can always satisfy the total demand of the sinks.

The total building cost C(N) of a network N is the sum of the building costs of all the edges:

$$C(N) = \sum_{e \in E} l_e \, {q_e}^{\beta}$$

where E is the set of all the edges in the network N, l_e is the length of an edge e and ${q_e}^{\beta}$ is the cost per unit length of building the edge e with a commodity flow of q_e capacity.
β is the cost exponent for the capacity with $0 \leq \beta \leq 1$. The lower this capacity cost exponent is, the more cost-effective it is to partly *merge* the edges going to two different sinks into one bigger edge. For instance, when $\beta = 0$, the capacity of a pipeline has no influence on its building cost; when $\beta = 0.5$, the cost of building one pipeline of length l with capacity 2 is only $l\sqrt{2}$ while the cost of building two pipelines of length l with capacity 1 is $2l$; when $\beta = 1$, building two pipelines of length l with capacity 1 is just as expensive as building one pipeline of length l with capacity 2. The capacity cost exponent for gas pipelines varies due to digging costs and material costs for example.

In Netlogo, the modeling environment consists of discretized patches. The only source and the multiple sinks are situated on the specific patches according to their coordinates. The network is formed by multiple paths, which are connected sets of patches that link all the sinks to the source. An example of a simple network of 1 source and 3 sinks is shown in figure 4.

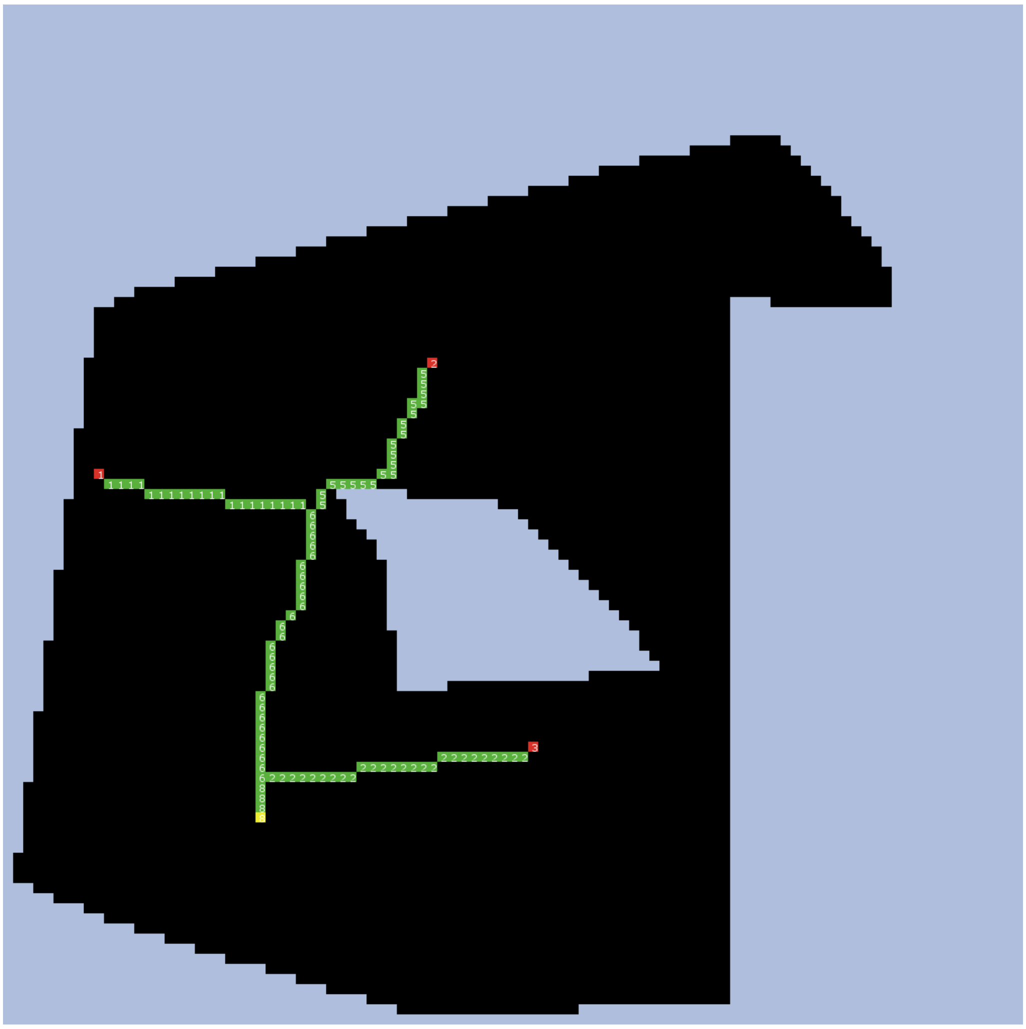


Figure 4. The yellow node is the source, the red nodes are the sinks marked as 1,2,3. The black area is the *allowed* area. The grey area is the *no-go* area. The green patches are the network paths with the patch's capacity in white number.

A brief description of the 1:n ACO model is as follows.

- *The agents*

  The agents in this agent-based model are ants looking for food sources, which are represented by the sinks in the model, and returning to their nest, which is represented by the source in the model, after reaching the food sources. The paths found by the ants that connect the food sources and the nest have different costs. Only the cheapest discovered paths are added to the network.

- *How is the model set up?*

  The unique source (nest) and the sinks (food sources) are enumerated (1, 2, ...) and situated on the entirely empty modeling environment according to their coordinates.

- *How do the ants look for food?*

  The ants semi-randomly walk until they reach a food source of which the demand is not yet fully satisfied. After reaching the food source, the ants collect the information regarding the food source number, its demand before starting to return to the nest.

- *What attract the ants on their returning to the nest?*
  There is only one nest. The ants are attracted by the unique *nest scent* spread all over the modeling world and the *pheromone* that existing network patches emit. The nest-scent is equal to the full *nest-scent power* of 200 at the nest. This is the basis parameter to which the other parameters such as *pheromone* values are adjusted. Details on the parameterizing the 1:n ACO model are referred to (Heijnen, et al., 2014). The nest scent decreases as the distance from the patches to the nest increases by the following formulation:

$$\text{nest-scent of at patch}_i = \textit{nest-scent-power} - \text{distance-from-patch}_i\text{-to-nest}$$

  The pheromone of the network is spread similarly. When the ants reach the existing network paths, they will only need to follow the paths until reaching the nest because all the paths lead to the unique nest.

- *How is a new path found by the ants added to the network?*
  After a specified number of ants have returned from the same food source without finding a cheaper alternative path, the current cheapest path will be built as a network path. When the path is built, the network capacity of the patches in the path will be increased accordingly by the demand from the food source (see the figure 4 above). The new network patches will start emit the network pheromone.

- *When does the model stop?*
  When all the demands of the sinks are fully satisfied.

Summary on the original 1:n ACO model

The original model for one source - multiple sink networks provides a sound solution for laying out networks taking into account the *no-go* regions. Another research by van Tol (2015) acceptably solves the problem of networks roll-out in cost differentiated regions. Since the solutions for these challenge are available, in this research, for simplicity,all the regions in the deterministic n:n ACO model are treated equally as *allowed* region with the same cost. The conceptualization of the deterministic n:n ACO model is in the next section.

3.2 Conceptual model for multiple source - multiple sink networks

As an extension of the original ACO model for *1 : n* networks, the deterministic n:n ACO model for *n : n* networks accepts the inputs, which are the coordinates of the fixed sources and sinks and their supply or demand capacities. It generates a network layout connecting all the sinks to the sources. The design problem is formulated similarly to the one for *1 : n* networks introduced in the previous section using the same building cost function: *To find a network that minimizes the total building costs and satisfies the following constraints.*

- All the fixed sinks must be connected to the fixed sources.
- There are no differentiated areas in the modeling world.
- Due to the availability of the total supply capacity of the sources, the demand of a sink could be either fully or partly satisfied.

The deterministic n:n ACO model keeps the established settings and parameters that directly or indirectly influence the behavior of the ants from the original model. In order to add more sources to the modeling world, the following changes were made.

- *What attract the ants on their returning to the nests?*
  There are multiple nests. Once a nest's supply capacity is entirely occupied by the sink(s), it will stop spreading its nest scent to the modeling world. The ants are attracted only by the strongest nest scent among the available nest scents from all the nests with available capacities. On each step of the ants, the strongest nest scent can belong to different nests according to the coordinates where the ants are. The nest scents are spread in the same way as in the original 1:n ACO model.

  $$\text{nest-scent of nest}_i \text{ at patch}_j = \textit{nest-scent-power} - \text{distance-from-patch}_j\text{-to-nest}_i$$

  Because there are multiple nests, different network paths may lead to different nests. For this reason, the network pheromone is not adopted. In other words, the network paths do not emit the pheromone.
  When the ants have reached the existing network paths, they will get on the paths under a defined condition; they will either keep following the paths until reaching one of the nests or get off from the paths and following the strongest nest scent under another defined condition regarding the strength of the nest scent. Details are provided in the next section on why and how the ants get on and get off the existing network paths.

- *How is a new path found by the ants added to the network?*
  After a specified number of ants have returned from the same food source without finding a cheaper alternative path, the current cheapest path will be built as a network path. When the path is built, the network capacity of the patches in the path will be increased accordingly by either *the demand* of the food source or the *highest available capacity* of the nest.

- *When does the model stop?*
  When all the demands of the sinks are fully satisfied or when the total supply capacity of the nests becomes zero.

Details of the internal states and properties of the model agents are provided in Appendix B.

3.3 Deterministic n:n ACO model parameterizing and cleaning

A detailed list of parameters with explanations about their roles in the model and their corresponding real-world network characteristics are available in Appendix B. In the implementation of adding multiple nests, two new parameters were created: *small-margin-get-on-network* and *small-margin-get-off-network*. The range of these two parameters are proportionally adjusted by hand based on the value of *nest-scent-power*, which is 200. The value of 200 is the result from the parameterizing the 1:n ACO model. It serves as the basis parameter to which the other parameters' values are related.

At each patch in the modeling world, there are multiple nest scents emitted by the different nests. Once the available capacity of a nest is zero, its nest scent will disappear from the modeling world. On their way returning to one of the nests, the ants follow the strongest nest scent among the available nest scents. In the discretized patch world, each patch has 8 neighbors consisting of 4 direct neighbors (white patches) and 4 diagonal neighbors (orange patches) as shown in figure 5. The distance between a patch to its direct neighbors is 1 and the distance between a patch to its diagonal neighbors is $\sqrt{2}$. At the moment, an ant on the green patch can only perceive the nest scents from 3 patches in its heading direction (3 patches on the vertically left of the green patch because the ant is heading to the left). As a result, the maximum difference between the strongest nest scent at the green patch and the strongest nest scents at one of its neighbor is $\sqrt{2}$. Upon this fact, the ratio of the strongest nest scent between the green patch and its neighbors is relatively estimated. The ratio helps to determine the ranges of the two parameters as follows.

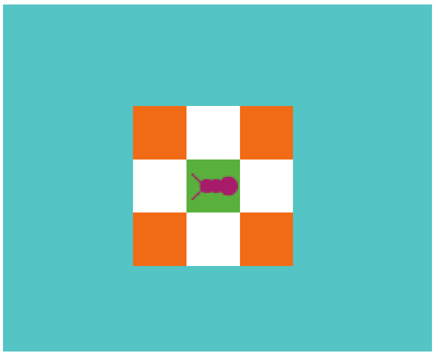


Figure 5. Patch neighborhood

3.3.1 Small-margin-get-on-network (s_1)

After reaching a food source, the ants start to return to a nest which still has available capacity. Induced from the cost function, it is cheaper to merge two paths into a bigger capacity path when the capacity cost exponent β is smaller than 1. Yet because there are multiple nests, it is not always cheaper to do so. When there are existing network patches in the neighborhood, the ants must make decision if they should get on one of the network patches following the decision-making procedure described below.

- Step 1: the neighbor patches where the ants has not yet stepped on are divided into two groups: network neighbor group and non-network neighbor group.
- Step 2: let *best-non-network* be the non-network patch with the strongest nest scent in the non-network neighbor group. Its strongest nest scent is denoted as $\text{best}_{\text{non-network}}$
- Step 3: let *best-network* be the network patch with the strongest nest scent in the network neighbor group. Its strongest nest scent is denoted as $\text{best}_{\text{network}}$
- Step 4: If

  $$\text{best}_{\text{non-network}} > (1 + s_1) * \text{best}_{\text{network}}$$

  then the ants will not get on the network but move to the *best-non-network* patch, otherwise they will get on the *best-network* patch. In other words, the $\text{best}_{\text{non-network}}$ needs to be greater than the $\text{best}_{\text{network}}$ by some extent in order to let the ants ignore the existing network patches.

The high value of s_1 encourages the ants to get on the existing network when they face one. The value of this parameter is estimated by hand as in figure 6 by comparing the nest scent ratio between a patch and its neighbors. The values range approximately from 0.0075 to 0.0175.

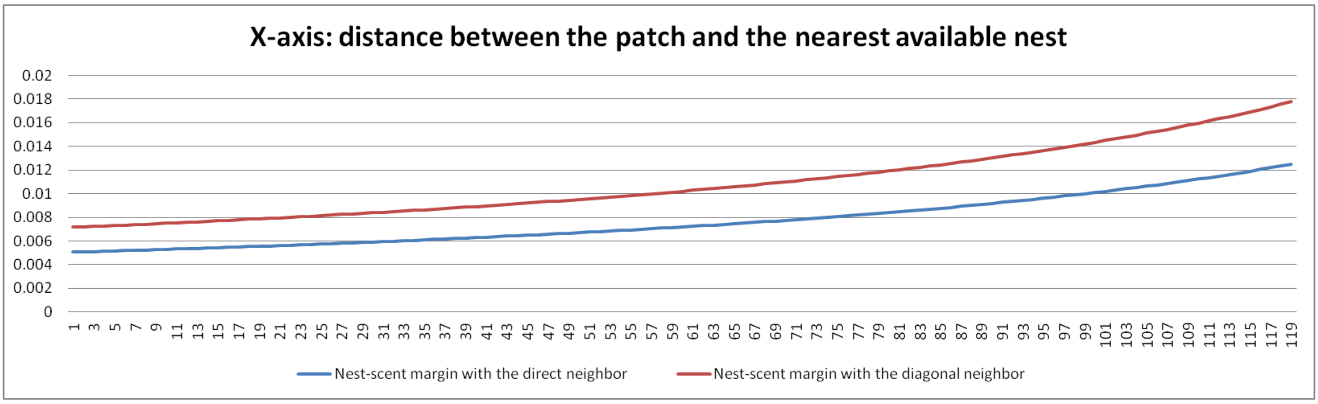


Figure 6. Small margin parameter estimation

In attempt to locate the best values of s_1 , an experiments on 100 random examples was executed. The result indicated that the lower s_1 is, the lower the possibility to find the cheapest network due to the fact the ants ignore the network patches. In the second experiment on the same set of 100 random examples, the ants were required to get on the existing network patches when they face them. The result of the second experiment is hardly different compared to the result of the first experiment, yet the computing time was much reduced. As a result, the rule is set as *when the ants find network patches on their returning to the nests, they will always get on the network patch with the strongest nest scent.*

3.3.2 Small-margin-get-off-network (s_2)

Because there are multiple nests in the model world, the existing network paths may lead the ants to a nest which will not help to minimize the cost. With the introduction of the parameter *small-margin-get-on-network*, it allows the ants to merge the small paths into one big path and with the introduction of the parameter *small-margin-get-off-network* it enables the ants to split the path into two separate paths when necessary.

After getting on the existing network, the ants must make decision on each step if they should stay on the existing network or get off from it. The decision-making procedure is as follows.

- Step 1: the neighbor patches where the ants has not yet stepped on are divided into two groups: network neighbor group and non-network neighbor group.
- Step 2: let *best-non-network* be the non-network patch with the strongest nest scent in the non-network neighbor group. Its strongest nest scent is denoted as $\text{best}_{\text{non-network}}$
- Step 3: let *best-network* be the network patch with the strongest nest scent in the network neighbor group. Its strongest nest scent is denoted as $\text{best}_{\text{network}}$
- Step 4: If

$$\text{best}_{\text{non-network}} > (1 + s_2) * \text{best}_{\text{network}}$$

then the ants will get off from the network and move to the *best-non-network* patch, otherwise they will keep follow on the network path. In other words, the $\text{best}_{\text{non-network}}$ needs to be greater than the $\text{best}_{\text{network}}$ by some extent in order to let the ants get off from the path.

Similarly, the parameter is estimated to range from 0.0075 to 0.0175. The high value of s_2 keeps the ants on the existing network. By setting the parameter high, the network planner favors the merging of the paths. The experiments on the 100 random examples show that there is no such a statistically significant value of the *small-margin-get-off-network* that gives the higher chance to find the cheapest networks as displayed in the figure below.

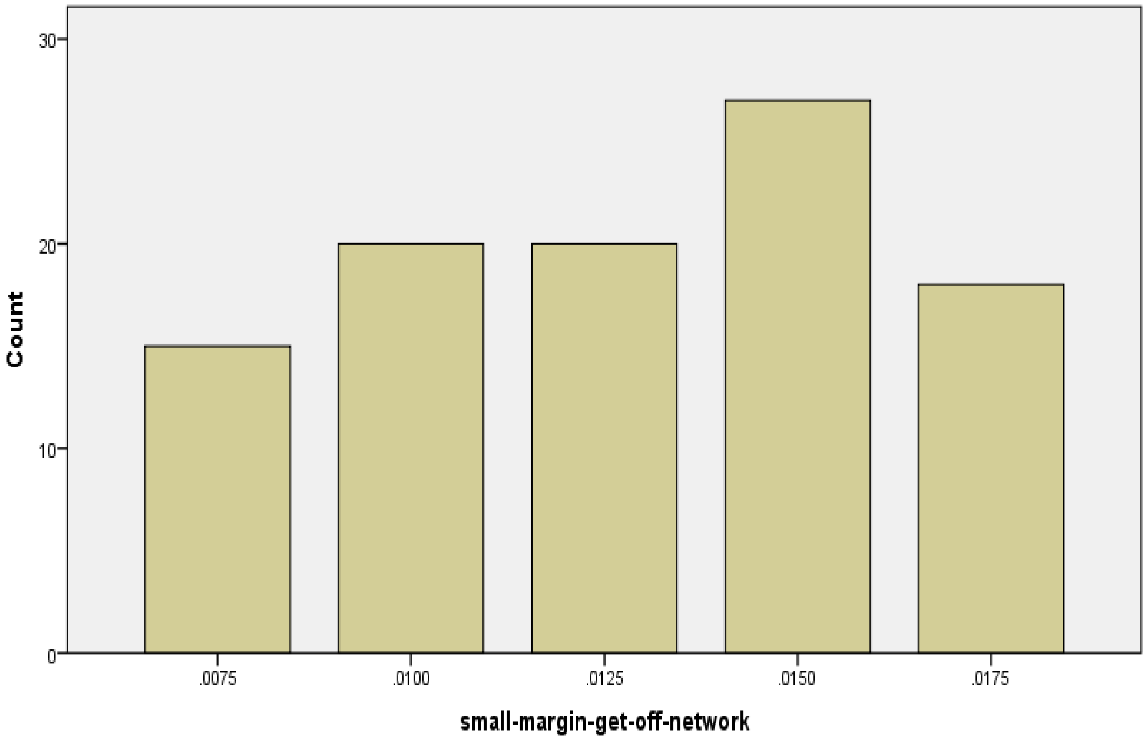


Figure 7. The frequency (out of 100 examples) that the best networks are found with each parameter value

3.3.3 Path trimming

Due to the fact that the ants at the same time are under the attractions by multiple nest scents, they walk in a multi-directional fashion corresponding to the strength of the nest scents. For this reason, the paths found by the ants can consist of many useless patches or even loops. The paths are trimmed to remove the useless patches and the loops for cost improvement.

The trimming is a practical solution for the computing time issue of the model. As introduced in the conceptual model, after *a specified number of ants* have returned from a food source without finding a cheaper alternative path, the current cheapest path will be fixed as a network path. The specified number of ants is practically set at 50 in the model. Increasing the number would lead to a higher number of different paths found by the ants and consequently improve the result, but significantly slow the simulation down. Most of the time the cheapest path is only a few patches shorter/different compared to the second and third cheapest paths before it. The trimming speeds up the simulation by removing the useless patches and loops in the paths. Without the trimming, the model needs more ants (more time) to trim the paths itself. Details are available in the Appendix A.

3.4 Model validation on comparison with the deterministic n:n GG model

As aforementioned in the research methodology, the deterministic n:n ACO model is compared to the deterministic n:n GG model developed by Petra Heijnen. The original 1:n GG model for *1 : n* networks slightly outperforms the original 1:n ACO models in terms of finding cheaper networks. For this reason the comparison between the deterministic n:n ACO model and the deterministic n:n GG model helps to theoretically reveal how good the outputs of the deterministic n:n ACO model are on the metric of total building cost.

3.4.1 Generated set of examples

In order to compare both methods, 100 random examples were generated in a 2D-plane of size 100×100 units of length. For each example, the parameters as listed in table 3 are randomly chosen from a uniform distribution within the given range. For simplicity, the total supply capacity of the sources is set equal to the total demand from the sinks.

Table 3. Input parameters for comparison on 100 examples

Parameter	**Range**
Horizontal axis total area	[0,100]
Vertical axis total area	[0,100]
Total number of nodes	[7,15]
Total number of sources	[2,4]
Maximal demand of consumers	[1,10]
Capacity cost exponent β	[0,0.9]

Following the comparison method established in the previous study for one source - multiple sink network layouts by Heijnen et al. (2014), the deterministic n:n ACO model and the deterministic n:n GG model are mainly compared on the best-cost layouts for each of 100 examples.

3.4.2 Results

For each example, the deterministic n:n GG model executed 10 runs while the deterministic n:n ACO model executed 15 runs. 55 out of the 100 examples, both methods gave approximately equal costs within the margin of 5%[5]; 43% of the examples, the GG method outperformed the ACO method (see figure 8).

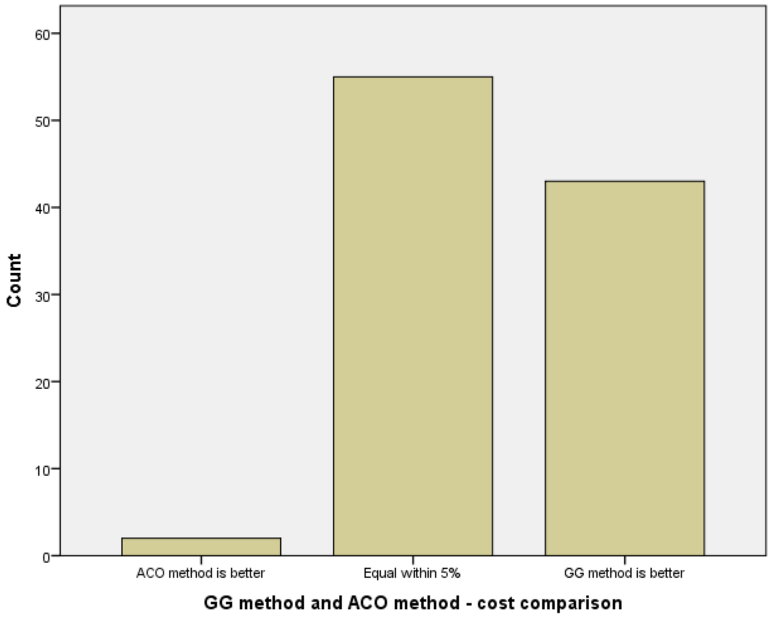


Figure 8. The deterministic GG method and the deterministic ACO method - cost comparison on 100 examples

---

[5] There are many examples in which the best costs are slightly different but the network layouts are hardly different. 5% is a pragmatic choice to fairly compare the two methods.

On average, the costs from the runs in one example by the deterministic n:n GG model show an average deviation from the best results of only 5.5% with a standard deviation of 7.9%. Around 50% of the examples, the best-cost networks are found within 7 first runs and re-found a couple of times. This leads to a conclusion that the deterministic GG method generates very *deterministic* network layouts. Different speaking, the output network typologies are mostly repeated in the runs.

With the deterministic n:n ACO model, generally the best-cost networks are found only 1 time in all 15 runs for each example. The average deviation from the best results is 6.1% with a standard deviation of 5.8%. In addition, the deterministic ACO method generated a higher number of network topologies in comparison to the deterministic GG method. It is concluded that the more runs are executed, the likelihood to find better networks increases. 15 runs per examples is merely a pragmatic choice due to the fact that the deterministic n:n ACO model is worse than the deterministic n:n GG model in terms of computing time.

For summary, the deterministic GG method's results are slightly better, yet the deterministic ACO method is *rich* in generating diverse network layouts which are promising inputs for the non-deterministic design.

3.5 Central Link approach and the deterministic n:n Central-Link model

This section introduces the Central Link approach - the real-world common practice in the network design. First, the reasons why the Central Link approach is applied in reality are presented. Next, the development of the deterministic n:n Central-Link model is explained. In the next section, the deterministic n:n ACO model is validated by comparing with the deterministic n:n Central-Link model.

3.5.1 The Central Link approach

In the design of one source networks, a common approach practiced by the network planners is to first build a very large capacity central link from the source to one of the sink; the rest of the sinks are then connected directly to the central link. The first source to be connected is usually the source with the highest demand.

For multiple source network, it is no longer an easy decision on the planning of the central link. The reason is that the starting point of the central link can be at any source and it can end at any sink. Depending on the geographical locations of the sources and sinks, the central link between the biggest capacity source and the highest demand sink are mostly, but not always the most cost-effective choice. Sometimes the central link is built between the biggest capacity source and the second or third highest demand sink.

Taking the current network of OCAP, the flows from two CO_2 sources of Shell refinery in Botlek and the Abengo's bio-ethanol factory are first merged in the middle of these two sources. A long 26-inch backbone pipeline was built starting from the merging point in Rotterdam to the north of the Netherlands. Greenhouses areas were straightly connected to this backbone pipeline (Veenstra & Limbeek, 2014). The backbone pipeline enables OCAP to implement a top-down controlling mechanism with master control system at OCAP and slave control system at each greenhouse whenever there is either a fall in the supply or an increase in the demand. In summer, by setting up at the master control system, OCAP is able to allocate specific amounts of CO_2 flown to each greenhouse. In other words, the central backbone pipeline has a great impact on the operational level.

A study on designing cost-optimal CCS configurations for an industrial cluster at Botlek area by Berghout (2014) indicated that the centralised configuration in which the CO_2 from all 16 plants at the site is jointly captured, purified and compressed to a trunk CO_2 pipeline was proven to be cost-effective and particularly interesting for the smaller emitters because of economies of scale.

3.5.2 The deterministic n:n Central-Link model

Whereas the *deterministic n:n ACO model* starts connecting the nodes from scratch, the *deterministic n:n Central Link model* starts with first building the central link. An experiment on the starting point (source) and the ending point (sink) of the central link was executed. The options for the source-sink pair were (1) the biggest capacity source and the highest demand sink, (2) a pair of a source and a sink that has the longest distance among all the source-sink pairs and (3) randomly chosen. The result revealed that the first choice worked the best in finding cheap networks. As discussed in the last sub-section, real-world central links are also often built between the biggest capacity source and the highest demand sink. As a result, the rule is set that the central link in the model is built between the biggest capacity source and the highest demand sink. The modeling steps are as follows.

- Step 1: In this step, the ants only look for the highest demand sink and return to the biggest capacity source. A path between this pair of source and sink is built after the specified number of ants have returned from the sink without finding a cheaper path.
- Step 2: After the central link is built, it starts to emit the so-called *central-link-scent* which is spread all over the modeling world in the same way the nest scents are spread. This *central-link-scent* will attract the ants in the step 3 to connect the rest of the nodes to the central link.
- Step 3: From this step, the ants start looking for the rest of the sinks. However, it is set that after reaching the sinks, the ants must first return to the nearest patches in the central link following the strength of the *central-link-scent.* Once the ants reach the central link, they will start to behave as they do in the deterministic ACO model to find the way returning to the nests.

Figure 9 below demonstrates an example of typically different layouts found by the deterministic ACO method and the Central Link method for the same set of inputs.

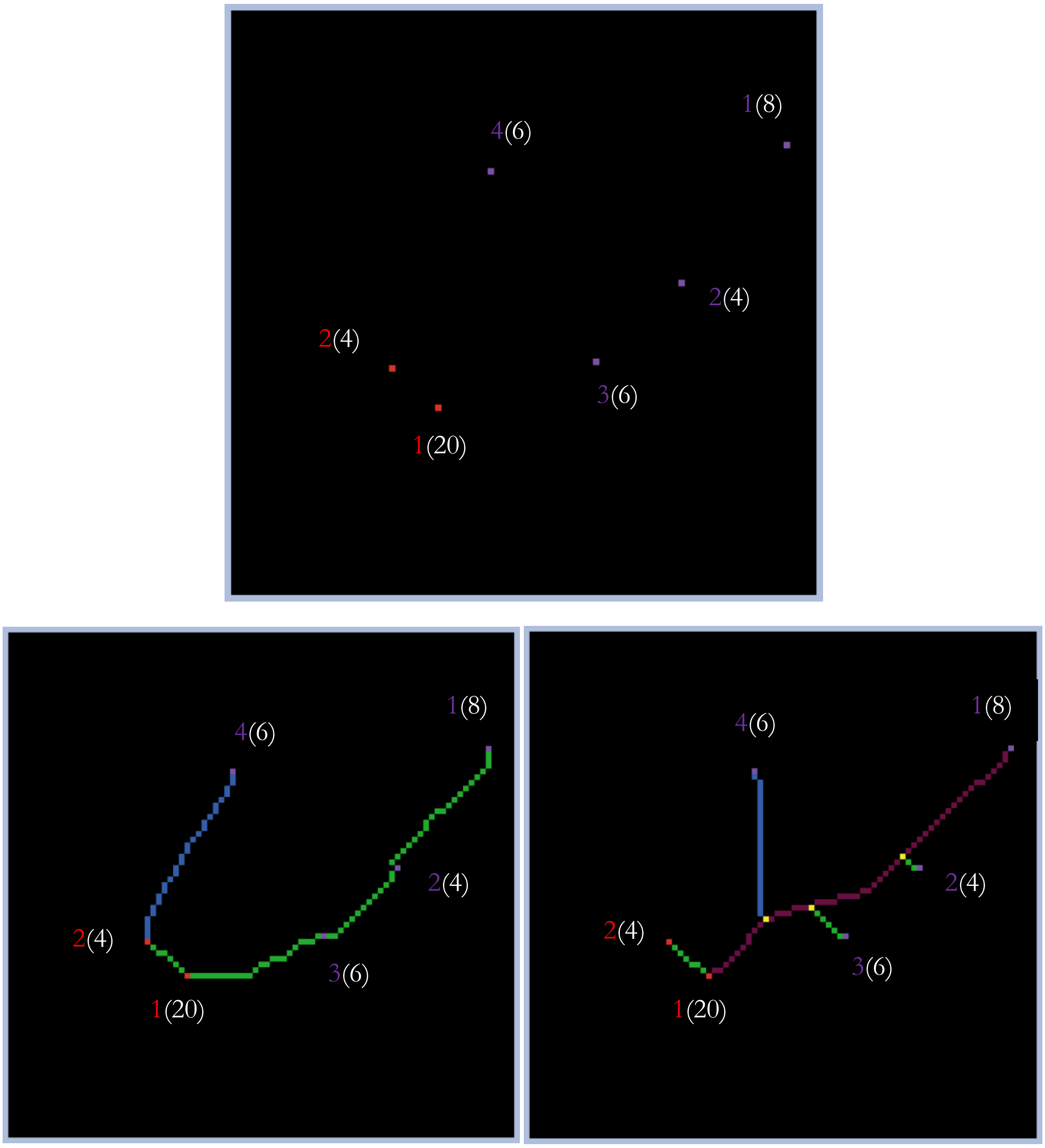


Figure 9. The upper image is the initial setup. The lower on the left is the output of the deterministic n:n ACO model and the lower on the right is the output of the deterministic n:n Central-Link model.

In the upper image, there are two sources marked with red label numbers and their supply capacities in white; there are four sinks marked with violet label numbers and their demand capacities in white. In the two lower images, the network paths are green if it is fed by only one source and blue if it is fed by more than one source; the central link is purple. The sink 4 with demand 6 is fed by two sources: capacity of 4 from the nest 2 and an additional capacity of 2 from the nest 1. The yellow patches are where the ants reached the central link The network layout generated by the deterministic ACO model costs 357 while the one generated the Central Link model costs 364. Whereas the cost gap is less than 2% of the cheaper layout, one can imagine that the network flexibility degrees of these two layouts are possibly very different. The comparison on the network flexibility of these two approaches will be addressed in chapter 5.

3.6 Model validation on comparison with the deterministic n:n Central-Link model

The purpose of comparing the deterministic n:n ACO model and the deterministic n:n Central Link model is to practically assess the design quality of the deterministic n:n ACO model. The comparison was executed on the metric of total building cost.

3.6.1 Generated set of examples

The method of comparison in this case is different from the one used to compare the deterministic n:n ACO model and the deterministic n:n GG model. Since the previous comparison is generally focused on the methodology of bottom-up and top-down approaches, the 100 examples are randomly generated and not categorized in any category regarding the capacity cost exponent, locations and the capacity/demand of the nodes. However in this comparison, the Central Link model's outputs depend greatly on those factors due to the geographical layout, the length and the capacity of the central link. In order to fairly compare two models, 36 examples are semi-randomly generated by hand in different categories based on the factors in the cost function. For simplicity, the total supply capacity is still assumed to be equal to the total demand.

- *Capacity cost exponent* β: Low (0) - Medium (0.5) - High (1). High capacity cost exponent increases the building cost substantially. The central link attracts the ants on their returning to nest, so in most of the cases the capacity of the central link will be as high as the total demand.
- *Locations of the sources*: One area - Two areas - Random. Examples are shown in figure 10 below. The number of nodes are fixed at 4 sources and 9 sinks. The sinks are assumed to be always randomly situated. The total length of the network has a direct impact on the building cost. It mostly depends on the locations of the nodes.

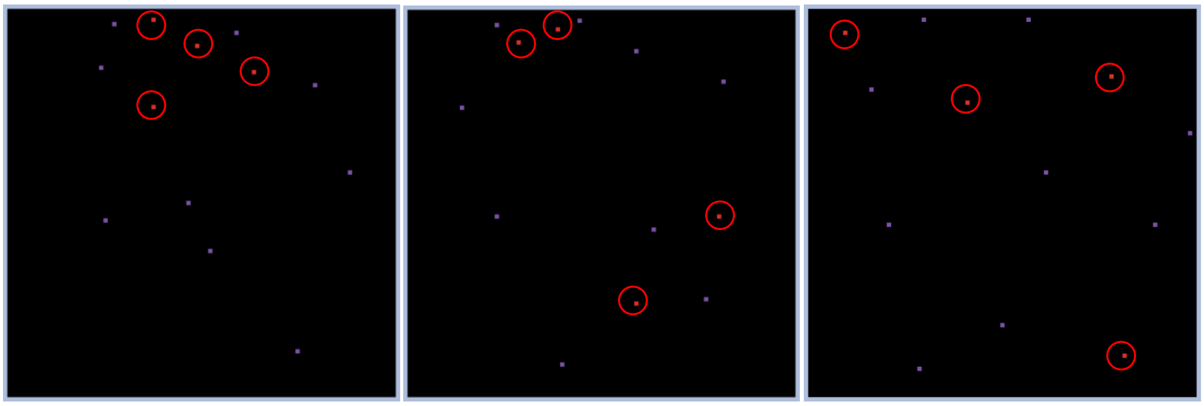


Figure 10. Sources locations from the left: One area - Two areas - Random. The sources are red nodes in the red circles. The violet nodes are the sinks.

- *Capacity variance of the sources*:
    - Low: the capacities of the sources are mostly evenly distributed (the left image in figure 11)
    - High: there is a source with a very big capacity and the rest have similarly small capacities (the right image in figure 11)

  The capacity variance of the sources can considerably affect the total cost of the network when the central link approach is applied. In details, the central link approach is conjectured to likely give bad results when the capacities of the sources are similar due to there is no clearly such the biggest capacity source.

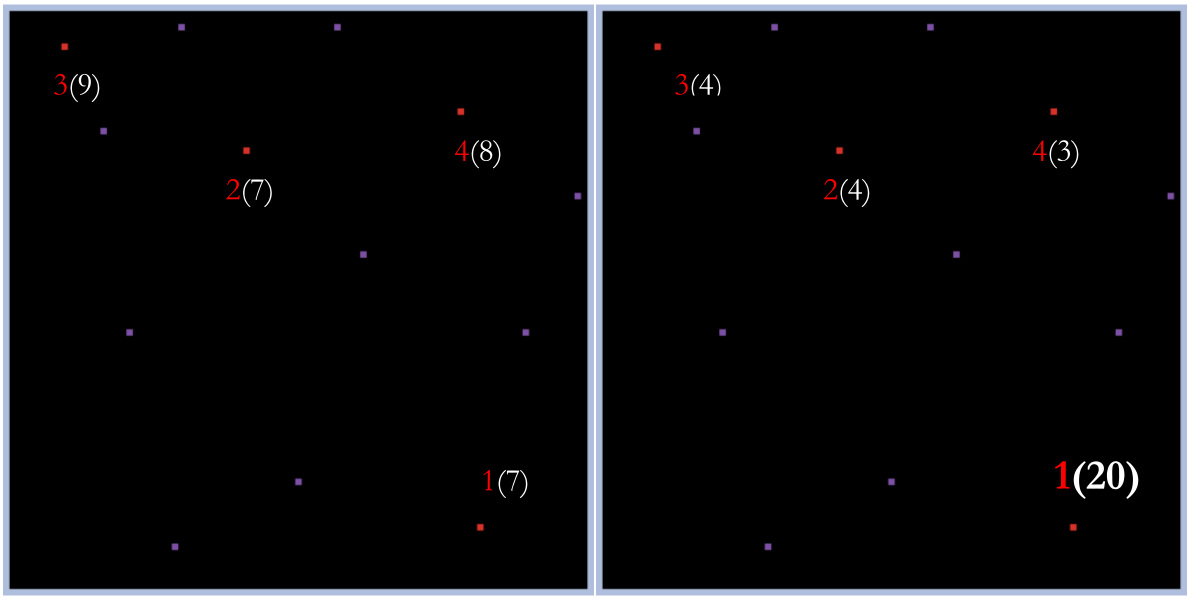


Figure 11. Capacity variance: Low (the left image) and High (the right image). The sources are marked by red label numbers and supply capacities in white. The total supply capacity in both cases are 31. In the left image, the sources have almost the same capacities while the sink 1 in the right image has a very big capacity.

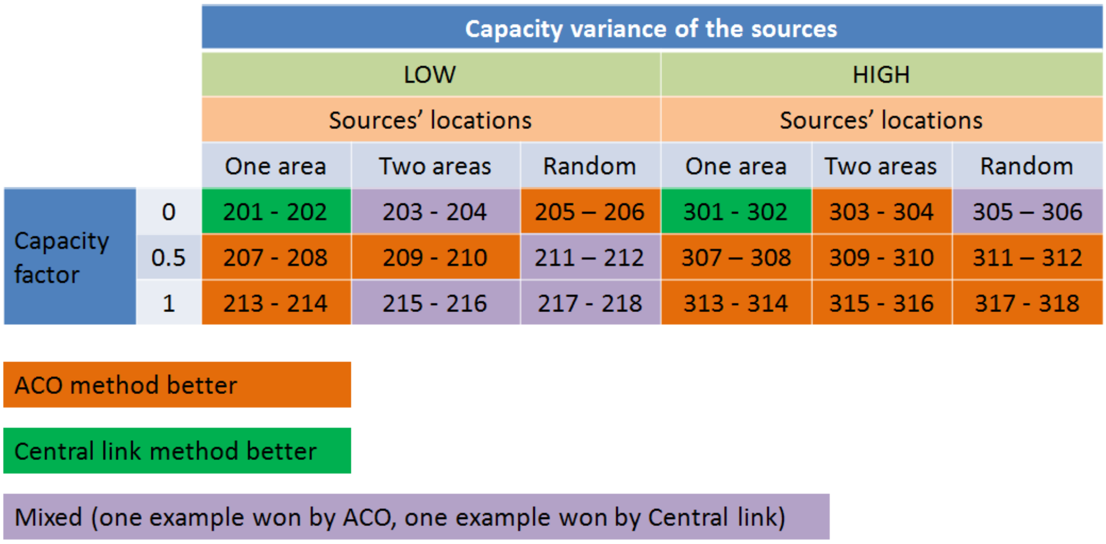
Two examples were generated for each category as shown in the figure below.

		Capacity variance of the sources					
		LOW			HIGH		
		Sources' locations			Sources' locations		
		One area	Two areas	Random	One area	Two areas	Random
Capacity cost exponent	0	201 - 202	203 - 204	205 – 206	301 - 302	303 - 304	305 – 306
	0.5	207 - 208	209 - 210	211 – 212	307 – 308	309 - 310	311 – 312
	1	213 - 214	215 - 216	217 - 218	313 - 314	315 - 316	317 - 318

Figure 12. 36 categorized examples

3.6.2 Results

20 runs are executed for each example for both models. The comparison is based on the best-cost output layouts from these 20 runs. The frequency to find the cheapest networks are counted for each method. Overall, 27 out of 36 examples, the ACO method gave better cost networks.



		Capacity variance of the sources					
		LOW			HIGH		
		Sources' locations			Sources' locations		
		One area	Two areas	Random	One area	Two areas	Random
Capacity factor	0	🟩 201 - 202	🟪 203 - 204	🟧 205 – 206	🟩 301 - 302	🟧 303 - 304	🟪 305 – 306
	0.5	🟧 207 - 208	🟧 209 - 210	🟪 211 – 212	🟧 307 – 308	🟧 309 - 310	🟧 311 – 312
	1	🟧 213 - 214	🟪 215 - 216	🟪 217 - 218	🟧 313 - 314	🟧 315 - 316	🟧 317 - 318

🟧 ACO method better

🟩 Central link method better

🟪 Mixed (one example won by ACO, one example won by Central link)

Figure 13. Results on the comparison between the ACO method and the Central Link method

For each category:

- When the capacity cost exponent is 0 (low), the capacity does not have influence on the cost. The results also indicated that when the cost exponent is 0, the ACO method and the Central Link method share 50%-50% of the times they found the best-cost networks. When the capacity cost exponent is 0.5 (medium) or 1 (high), the ACO method outperformed the Central Link method. The results can be easily understood because the central link's capacity is very big which leads to the high costs.

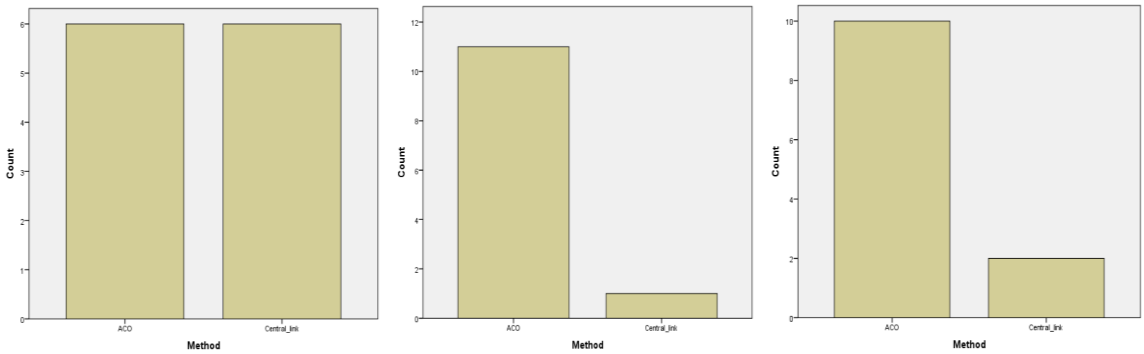


Figure 14. Capacity cost exponent from the left: Low - Medium - High

- In both cases of the capacity variance, the ACO method gave better results.

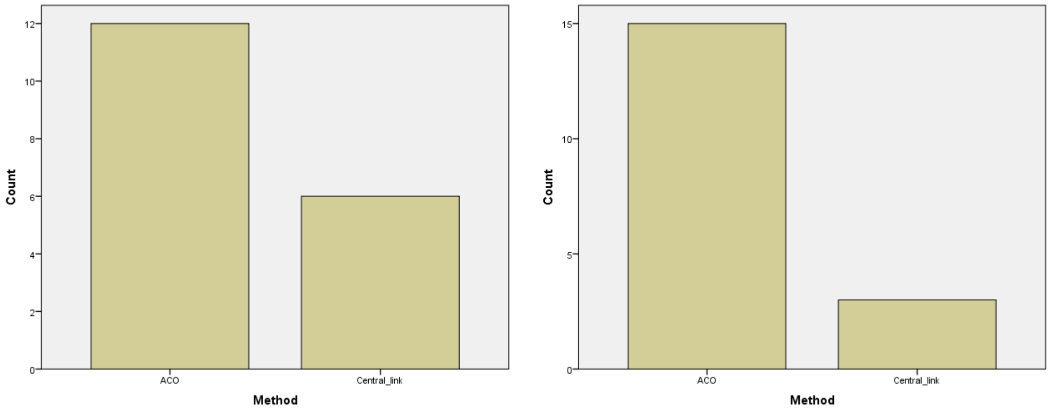


Figure 15. Capacity variance from the left: Low - High

- The location of the sources will enhance the Central Link approach if all the sources are in one area. When the sources are widely dispersed, they require longer connections from them to the central link which means higher costs. The results also align with this reasoning. The Central Link method is preferred when the sources are in one area.

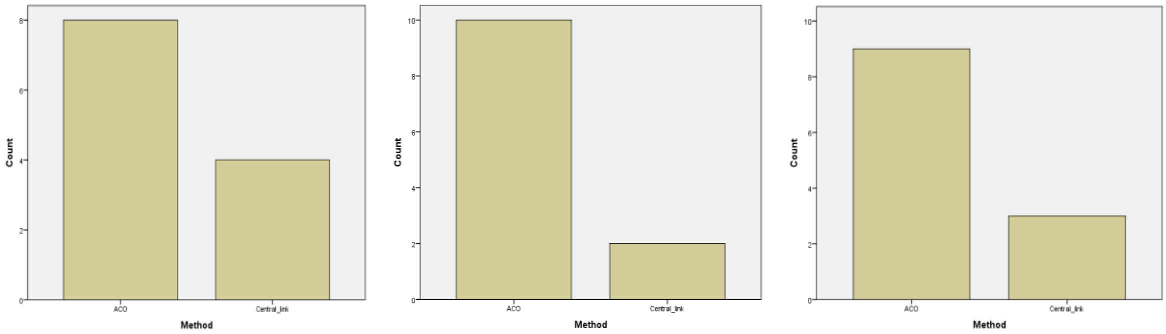


Figure 16. The locations of the sources from the left: One area - Two areas - Random

For summary, the ACO method works better in most of the cases, yet in some special categories such as when the sources are close to each other in an area and the sinks are not too dispersing, the Central Link method can be preferable.

3.6.3 The improved deterministic n:n ACO model

The previous results lead to an improvement for the deterministic n:n ACO model. Because there are cases that the Central Link approach works well, the deterministic n:n ACO model will be improved by additionally embedding the Central Link approach into the current deterministic n:n ACO model.

- *How does the improved deterministic n:n ACO model work?*
  Before the improvement, the deterministic n:n ACO model built the network from scratch. After the improvement, the model user can choose how to set up the starting point of the model by switching between two options:

(1) First option is to start as it was before the improvement: all the sinks (food sources) are treated equally. The ants look for all the sinks at the same time.

(2) The second option is a hybrid approach between the ACO approach and Central Link approach. The idea is to first build a network path between a randomly selected source and a randomly selected sink. In other words, at first the ants only target the randomly selected sink. After that sink is connected to that source, the ants can start looking for the rest of sinks as they do in the original deterministic n:n ACO model. However this hybrid approach differs from the Central Link approach as the first built path is not a central link itself, so there is no central-link scent emitted. The first built path serves merely as an existing network path in the beginning.

Multiple runs with multiple randomly selected pairs of source-sink increases the chance to find the better cost networks. Since two approaches enlarges the coverage of solutions space, the improved deterministic n:n ACO model is expected to be able to discover cheaper networks.

- *How could the improved deterministic n:n ACO model make improvements?*
  The limitation of the original deterministic n:n ACO model is that the distance between a sink and a source is prioritized. A sink seems likely to be first connected to its near capacity-available sources, then connected to further sources. For this reason, there are possible cost-minimized network layouts being ruled out from the output set. In the example shown by figure 17, the sink 1, which is the highest demand sink, is the last one to be connected to the sources because it is far from both the sources. The best-cost network by the original deterministic n:n ACO model costs 244.

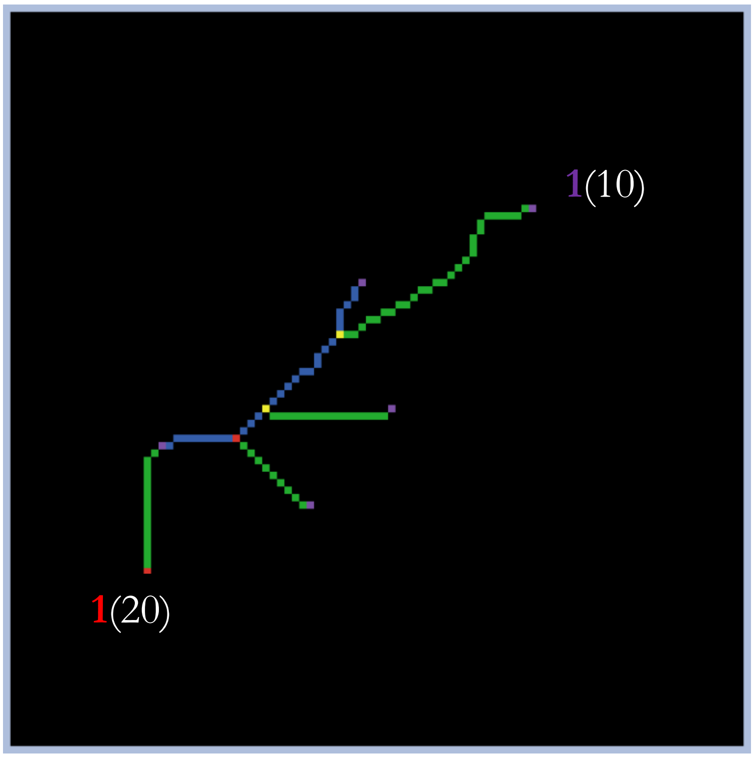


Figure 17. Best-cost output by the original deterministic n:n ACO model

With the improved deterministic n:n ACO model, by randomly connecting a sink to a source, there are numerous interesting network layouts with better costs generated within enough number of simulation runs. On the same example, the improve deterministic n:n ACO model generated a cheaper network (figure 18) in which the sink 1 is first connected to the source 1. This network only costs 230 (6% less).

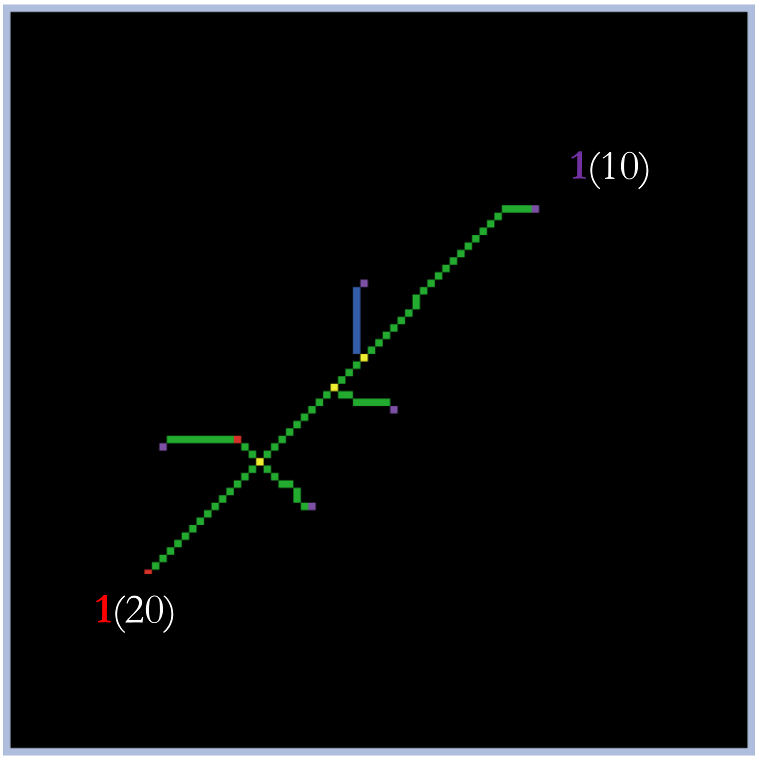


Figure 18. Best-cost output by the improved deterministic n:n ACO model

- *Results on the comparison between the original deterministic n:n ACO model and the improved deterministic n:n ACO model*
  An experiment was carried out on the same set of 36 examples with 20 runs per each model. 10 out of 36 examples the improved deterministic n:n ACO model found cheaper networks with entirely different topologies than the original deterministic n:n ACO model did. Most of the cases, the costs were reduced due to the *random* but *right* choices of the sink to connect to the right source at the first place. Hereinafter, the improved deterministic n:n ACO model is denoted as the deterministic n:n ACO model.

Summary of chapter 3

The deterministic n:n ACO model is developed based on the 1:n ACO model. Major changes that directly influence on the behaviors of the ants were made to implement the multiple source settings. The validation by the theoretical approach of the deterministic n:n GG model and the practical approach of the deterministic n:n Central-Link model confirmed the capability to find various cost-minimized networks of the deterministic n:n ACO model. This paved the way for proceeding the extension from the deterministic n:n ACO model to the non-deterministic n:n ACO model in chapter 5. Before that, the next chapter addresses the future scenarios of the uncertainty on network participants and the available flexibility options.

Chapter 4. Future scenarios and flexibility options

In the engineering design for flexibility process, recognizing the major uncertainties that the network is likely to encounter is a prerequisite step before identifying the specific parts of the network that provide the flexibility (Ajah & Herder, 2005; de Neufville & Scholtes, 2011). The first section of this chapter is dedicated to discussing the uncertainties of network participants. Taking these uncertainties into consideration, the second section highlights the real options theory, introduces the optimization problem of the non-deterministic design in *n : n* networks and at the end presents the flexibility options to deal with those uncertainties.

4.1 Future scenarios

By looking at a wide range of possible future scenarios, the network planners learn about the situations that the network has to face in the future. If the anticipation of possible scenarios is correct, the design has a high chance of being successful in dealing with upside opportunities and downside risks.

4.1.1 Scenarios on the uncertainty of network participants in the *n : n* networks

The lifetime of infrastructures can range from about 5-10 years to about 30 - 50 years. During the time, the uncertainty of network participants can theoretically stem from both sides: sources and sinks. It will be either new requests of joining the network from the new sources and/or new sinks or the closedown of the existing sources and/or the existing sinks. Since this research focuses on the initial construction cost at the upper level of the operational level, it is not unrealistic to leave out the shutdowns of several nodes in the network. These shutdowns seem to have insignificant impacts on the value of the network. For instance, the removal of a greenhouse will not cause significant changes to the CO_2 pipeline network. The utmost attention is given to the potential new sources and the potential new sinks.

Practically, the major uncertainty comes from the new potential sinks/consumers. New sources can theoretically occur as well but are more unlikely to happen in short-term of less than 5 years. Looking at the example of OCAP's CO_2 network: besides the current two CO_2 sources, there are hardly other plants at the Rotterdam industrial area that are able to provide 99.9% pure CO_2. Moreover, the CCS implementation represents substantial costs while the CO_2 price is currently far from the break-even point (Veenstra & Limbeek, 2014).

A successful networked infrastructure must continue to fulfill the actual needs from the consumers in the right place and at the right time. Placing the *potential new sinks* at the central point of the study, the uncertainty varies across the following dimensions.

- Location: where are the potential sinks? In reality, this dimension brings about the least uncertainty. Practically, the network planners can somehow spot the potential areas.
- Timing: When will the business be started with the potential sinks?
- Demand: How much capacity is required by the potential sinks?

4.1.2 The upside opportunities and downside risks

A flexible network design limits possible losses from downside risks and increases the possible gains from the upside opportunities (de Neufville & Scholtes, 2011). In this study the upside opportunities and the downside risks in regard to the uncertainty of network participants are identified as follows.

- *Upside opportunities*
  When there are requests for the new connections and the network's layout and capacity are ready for the modifications with a relative ease. Accepting the lack of flexibility in the existing CO_2 network, Veenstra & Limbeek said "at present, there are greenhouse farmers from Aalsmeer came to OCAP and suggested that they will pay partly the building cost of a new big pipeline from the backbone pipeline to their greenhouses area. It is a very tight business case because the investment cost by us for a new big pipeline will be huge. If the area were closer to the backbone pipeline then the conversation would be much easier" (2014).

- *Downside risks*
  The major risk from the network participants relate to the shutdown of one participant. This type of risk directly affect the network's operational strategies, which is beyond the research scope.

4.2 Flexibility options

The previous section identified the uncertainty sources. This section determines the flexibility options provided by specific parts of the network that could handle and exploit those uncertainties, which are the uncertainties about the location, timing and demand of the potential new sinks. First, the real option approach is highlighted. Next, the optimization problem of the non-deterministic design and the flexibility options in *n : n* networks are addressed. The Central Link approach, which is a practical implementation of flexibility, is discussed at the end of this section.

4.2.1 Real options

As widely known in engineering systems design, a real option is the right, but not the obligation to take an action such as expanding, deferring or abandoning at a specified price and for a predetermined period of time (Gil, 2007). Among various types of option strategies summarized by Ajah & Herder (2005), the following are the most relevant to this study.

- *Option to expand*
  This seems to be the most useful architectural option. The network planners desire to select among several design alternatives one which is embedded with built-in flexibility to easily expand the network in future to connect the potential sinks. The investment in this type of option can be either investing in extra capacity for some parts of the network or investing in building a flexible layout or both.

- *Option to defer*
  This option conventionally links to one of the most natural inclination "wait and see". The network can be built partly in this period while the rest can be accomplished following either a predetermined or flexible schedule. The better information obtained in the planning stage, the more the planners choose not to defer. The implementation of this type of option in infrastructure network design is expectably beneficial as the infrastructure sector is characterized with high innovation rate (Herder, et al., 2011).

4.2.2 The optimization problem in the non-deterministic network design

In the conceptual model for the deterministic ACO design method, the inputs for the optimization problem are only from the fixed network participants (physical locations, supply and demand capacities). The optimization goal of the deterministic design problem is to minimize the total building cost under the defined constraints. In non-deterministic design, besides the inputs from the fixed participants, the predicted information about the uncertain participants is also taken into account, namely their locations and estimated supply and/or demand capacities. To cope with the uncertainties, the network planner wants to build flexibility options into the deterministic design. The extra investment cost for these flexibility options leads to an increase of the initial capital expenditure and the network planner must demonstrate that they are worth their cost at the decision-making table.

Under the major uncertainties caused by the *potential new sinks* identified in the previous section, the optimization problem to be solved for the non-deterministic design is formulized as: *to find a network that minimizes the total building cost and satisfy the following constraints.*

- The constraints of the deterministic design: (1) all the fixed sinks must be connected to the fixed sources, (2) there are no differentiated areas in the modeling world and (3) due to the availability of the total supply capacity of the sources, the demand of a fixed sink could be either fully or partly satisfied.
- The total building cost includes the building cost in present to connect the fixed nodes, the cost of implementing the flexibility options in present and the cost of connecting the potential sinks in future.

There is a trade-off between the flexibility and the initial investment cost. As a result, it is possible that there would be no flexibility options in the non-deterministic design if the project budget cannot afford them. The next sub-section identifies the possible flexibility options in multiple source - multiple sink networks.

4.2.3 Flexibility options in *n : n* networks

Combining the option to expand and the option to defer results in the phasing design.

- *Phasing design*
  The investment is ramified into several phases. As a consequence, the initial cost is substantially reduced, yet the economies of scale is foreseen (de Neufville & Scholtes, 2011). In the first phase, the fixed sources and fixed sinks are considered. The connections of the uncertain sinks are deferred and planned for future expansion in the second phase. Yet the design in the first phase has to prepare to those. In this research, the first phase is designed by the *non-deterministic* ACO design method, which is an extension of the deterministic ACO design method by embedding the possible flexibility options.

The flexibility of a network in this study refers to the ease with which network nodes (potential new sinks) can be added to the existing network. In other words, it is the ease to add an additional link connecting the nodes to the network. The ease in this research is measured on the construction cost of the connection which is based on the distance (locations) and the required capacity. Generally, the possible parts of a network that can provide the flexibility are the capacities and the physical layout of the paths. Two flexibility options are identified as follows.

- *Dynamic layout based flexibility*
  Melese et al. (2014) hypothetically studied the value of designing networked energy infrastructures with architectural flexibility. The results show that the architectural flexibility can significantly improve the gains from upside opportunities. With regard to the potential new connections requested by the potential sinks, a layout which in some extent reduces the distances between the network and the potential sinks while incurs acceptable costs is expected to bring about significant benefits in future expansion.

- *Extra-capacity based flexibility*
  Investing in extra capacity is not only meant for future connections but also an option to deal with anticipated demand increase by the sinks. The challenge for this study is the question about which parts of the network need to be built with extra-capacity.

4.2.4 The Central Link approach as an implementation of flexibility

The Central Link approach is a real-world practice of implementing flexibility. First of all, the central link always has a big capacity including an extent of spare capacity. This provides an ease to let the potential consumers connect directly to the network without influencing the demand-fulfillment to the existing consumers. Moreover, the central link in some special cases largely covers the consumer area. With the fishbone-like layout, it is easy to construct new links without interrupting the operation of the existing links. Whereas there is no optimal non-deterministic design model for multiple source - multiple sink network available, the *non-deterministic n:n Central Link* model is acceptable for the model validation and comparison.

Summary of chapter 4

In the design of multiple source - multiple sink networks, the uncertainty on network participants can theoretically come from both the sources and the sinks. This study focuses on the potential new sinks who want to join the network in future. As a combination of two real options (to expand and to defer), the phasing design is suggested to the network planner. Based on the formulation of the optimization problem in non-deterministic design, two flexibility options are identified as the dynamic layout based flexibility and the extra-capacity based flexibility. The next chapter provides the answers of how to embed the two flexibility options into the deterministic design method to achieve the non-deterministic design.

Chapter 5. Development of the non-deterministic n:n ACO model

This chapter introduces the non-deterministic n:n ACO design method. Following the phasing design, the input of the non-deterministic n:n ACO method are both the fixed network nodes and the potential network nodes. The output are the network designs that minimizes the total building cost and satisfies the constraints in the optimization problem for non-deterministic design defined in the previous chapter.

Section 5.1 is to explain how the flexibility options are incorporated in the non-deterministic ACO design method. The next section validates the non-deterministic ACO method based on the comparison with the non-deterministic Central-Link method. The chapter ends with section 5.3 where the evaluation of the design outputs are discussed.

5.1 The implementation of the flexibility options in the non-deterministic n:n ACO model

The model development process is shown in the figure below. The non-deterministic n:n ACO model consists of 2 sub-models which implement the two flexibility options. The dynamic layout ACO sub-model generates flexible network layouts which provide the proximity of the networks to the potential sinks while the extra-capacity ACO sub-model locates the parts of the network that need to have extra capacity for future connections.

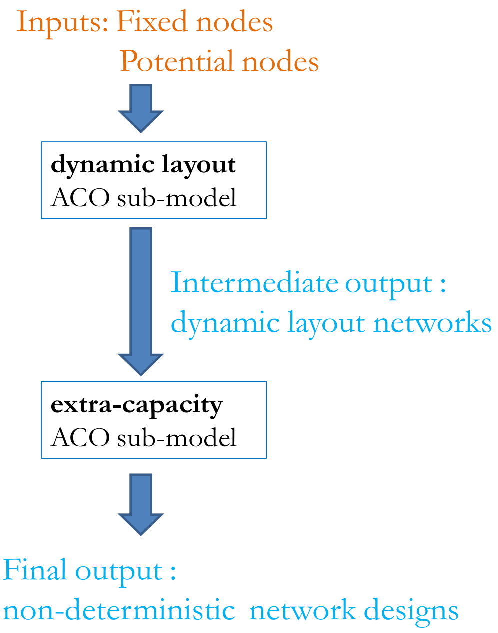


Figure 19. The non-deterministic n:n ACO model

5.1.1 The dynamic layout ACO sub-model

The dynamic layout ACO sub-model accepts the inputs including the fixed nodes and the potential sinks. A main difference distinguishing the dynamic layout ACO sub-model and the deterministic n:n ACO model is that the potential sinks are treated equally to the fixed sinks in the first stage in assumption that the total demand can be fulfilled by the total supply. The steps are explained on one simple example.

1. The potential sinks are seen as the same as the fixed nodes. Their estimated locations and potential demands are also inputted.

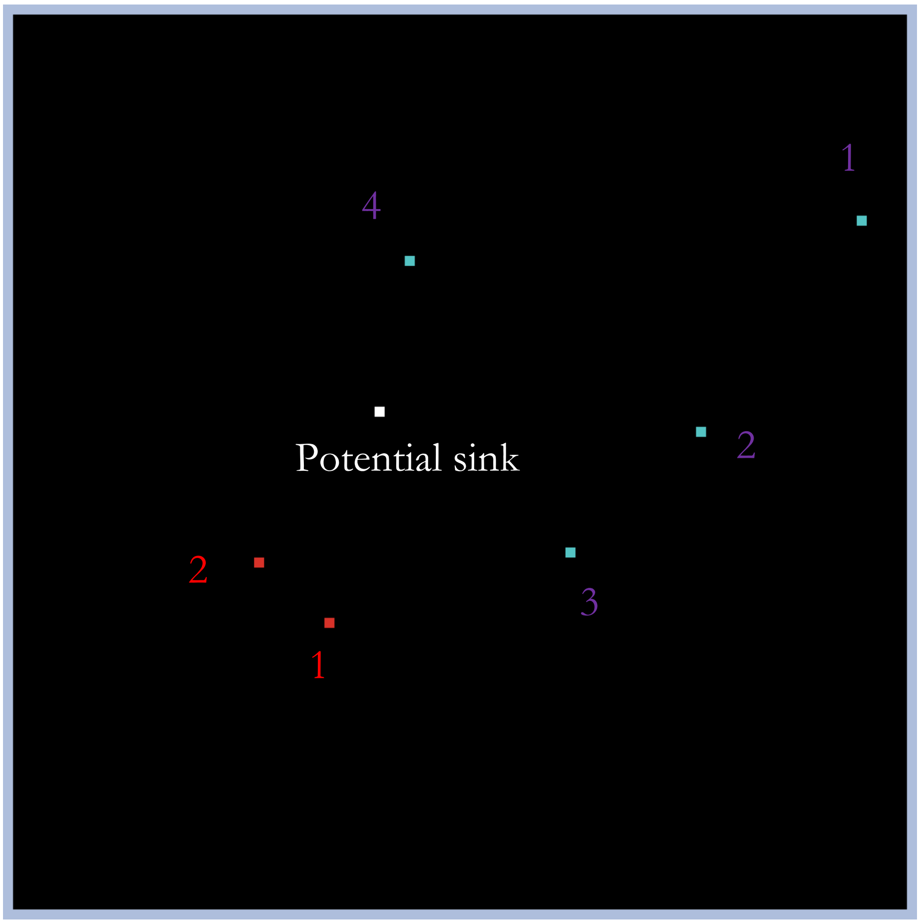


Figure 20. The potential sink is inputted from the start

2. The ants behave exactly the same as they do in the deterministic n:n ACO model to find the paths connecting all the sinks including the potential sinks to the sources. By including the potential sinks into the model at the first stage, the ants somehow are influenced by these sinks and the network paths that lead to them. In the example, the path between sink 4 and source 2 is pulled towards the potential sink. Figure 21 below shows the difference of with and without the potential sink.

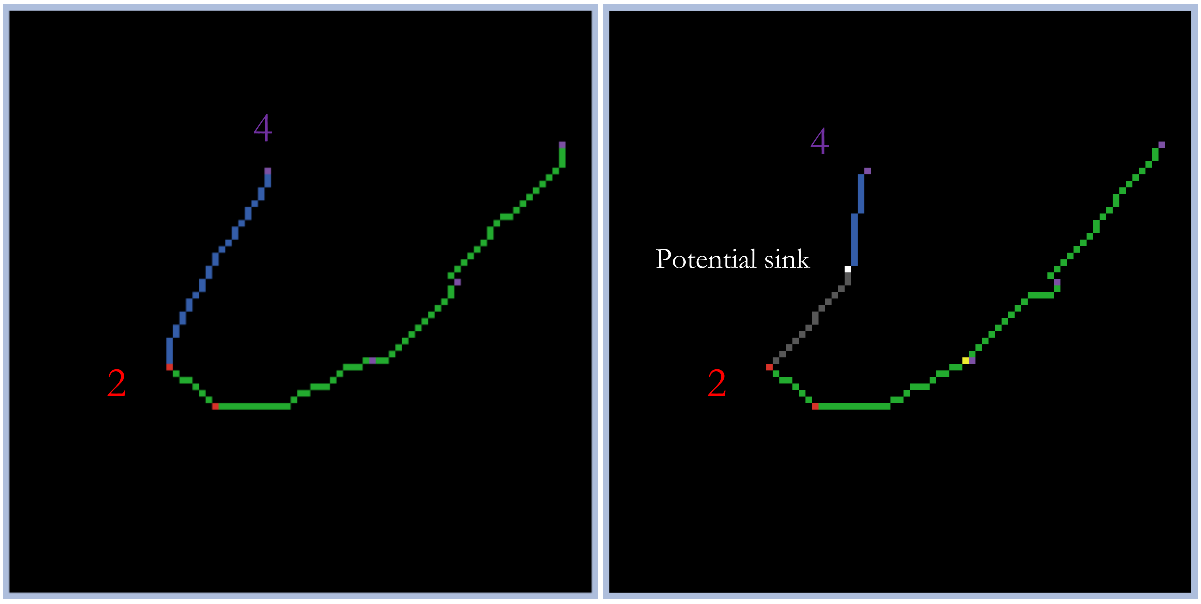


Figure 21. The influence of the potential sink on the path of sink 4 and source 2

3. Right after all the sinks are connected, all the paths of the network that start from the fixed sources and end at any potential sinks are reduced in capacity accordingly to the demand from the potential sinks. It can happen that the capacity 10 of a path is reduced by 5 because that path leads to a potential sink which has a potential demand of 5; it can also happen that a path of capacity 10 is entirely removed from the network if that path only leads to the potential sink with demand of 10. In the example, the capacity of the gray path between the potential sink and the source 2 is reduced by the demand of the potential sink (figure 22).

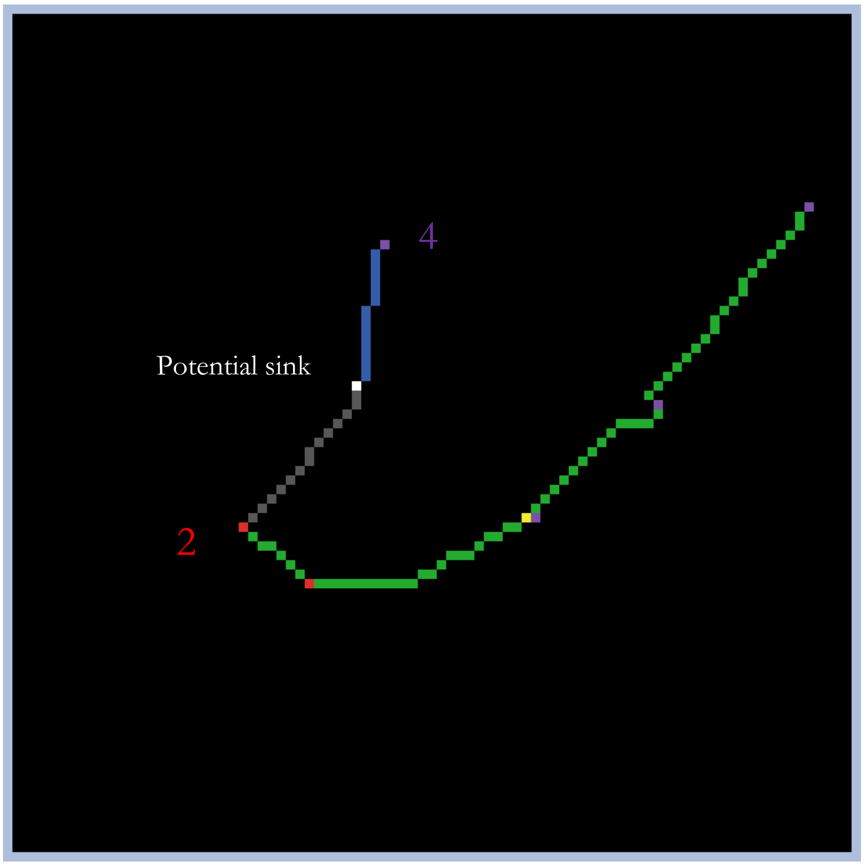


Figure 22. Path connected to the potential sink

4. In terms of network requirements, the output dynamic layout after all the paths related to the potential sinks are capacity-reduced or removed in step 3 is as the same as the output from the deterministic n:n ACO models. However, depending on the locations of the potential sinks, the output dynamic layout can either run through or pass close by the potential sinks. This provides the ease to connect these potential nodes to the network in future. In the example, the network runs through the potential sink. The cheapest building cost of the dynamic layout is 358, which is only slightly more expensive than the cheapest deterministic layout of 357. Yet in future there is no need to build a new path to connect this potential sink to the network. However, there is still a need for investing in extra-capacity as the next step.

5.1.2 The extra-capacity ACO sub-model

As indicated in the previous paragraph, the investment in the dynamic layout requires the investment in the extra-capacity as the next step. The proximity could not bring the cost-effectiveness without the available capacity to connect the future sinks. However, the input of the extra-capacity ACO sub-model can also be the network layout from either the dynamic layout ACO model or the deterministic n:n ACO model. This feature enables the network planner to opt between implementing the complete package of the two flexibility options or only implementing the extra-capacity based flexibility option. In this stage, the extra-capacity ACO sub-model will ask the ants to locate the parts of the network that need to have extra capacity.

As discussed in chapter 4, the uncertainty is on the location, the connection timing, and the demand of the potential sinks. According to these uncertainties, the ants may come up with different solutions: either investing extra-capacity in a specific part of the network, or rather building a separately new connection in the future. To enable the ants to do so, a new metric is applied instead of the total building cost as used in the deterministic n:n ACO model.

The new metric: Present Value (PV) of building cost

In this stage, the ants look for the potential sinks and return to the sources in exactly the same way they do in the deterministic n:n ACO model. However, the path connecting the potential sinks to the sources consists of 2 parts: the future new connection part and the existing network part that needs to be invested for a larger capacity. For this reason, the future cost needs to be converted to the present cost. The cost of the path is calculated as following.

$$\textit{Cost of the path in present} = \textit{Cost to invest the network part with larger capacity in present}$$

$$+ \textit{PV(cost to build the new connection part in future)}$$

For example, with timing = t years, discount-rate in present value calculation = r, the current capacity of the network part is Q with the length of L and the potential demand is q with the length of new connect is l, the cost of the path is:

$$\textit{Cost of the path in present} = [\ L(Q+q)^{\beta} - LQ^{\beta}\ ] + [\ lq^{\beta} / (1+r)^{t}\ ]$$

For the same input network displayed by the upper image in figure 23, the decision can change from investing extra-capacity in present to building an entirely new path as the timing and discount rate change. In the context in the lower image on the left, it suggests the network planner not to invest in the extra-capacity whereas in the lower image on the right, a part of the network should be equipped with extra-capacity to prepare for future connection.

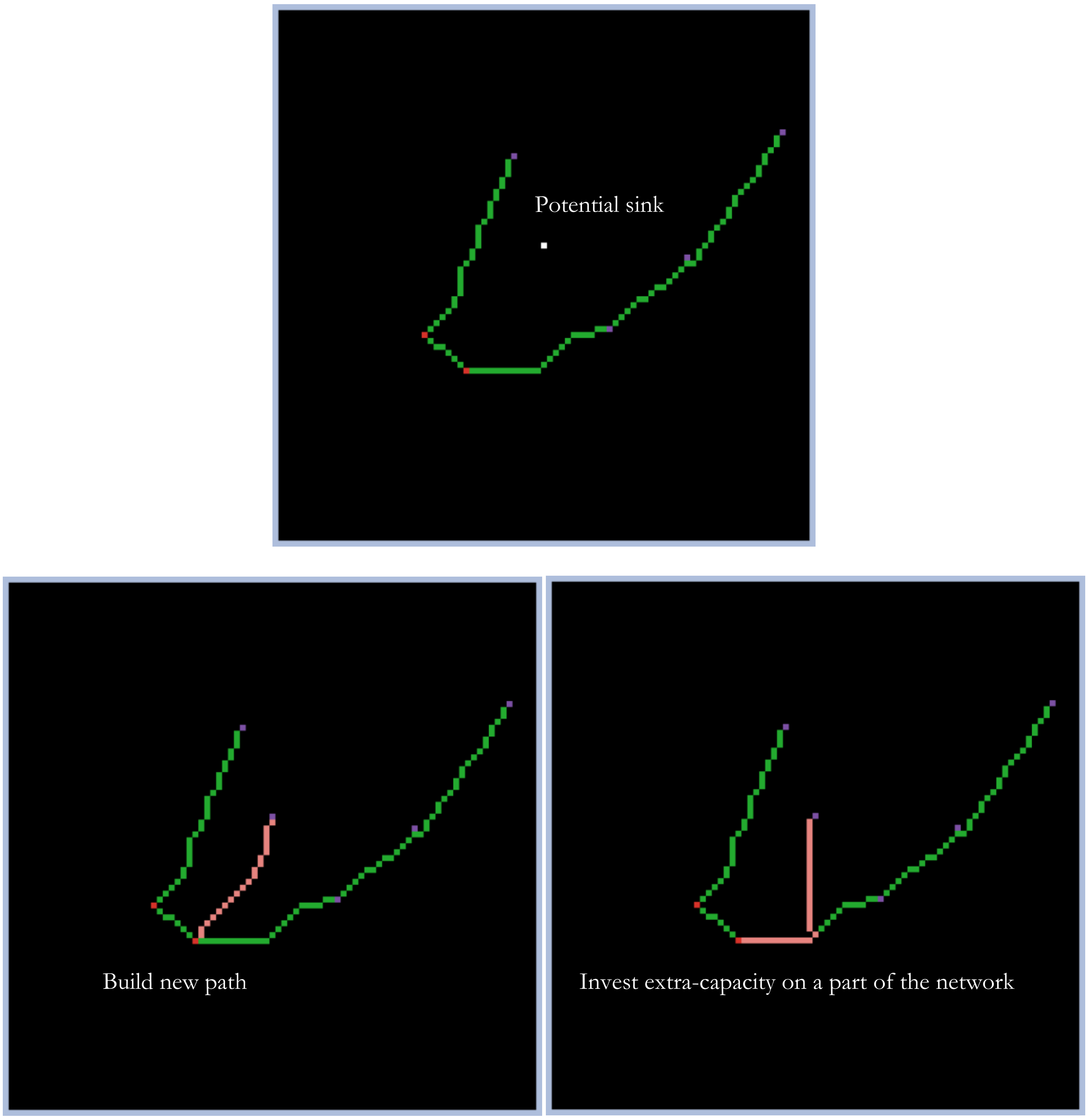


Figure 23. The left image: t = 10 years, r = 0.07; the right image: t = 5 years, r = 0.05; the pink network is the path connecting the potential sink to the source.

The output of the extra-capacity ACO sub-model helps the network planners make decision on which part of the network to invest in extra-capacity and how much it will cost more for this capacity investment.

Recalling the example in the previous sub-section, because the dynamic layout runs through the potential sink (see figure 24), it is easy to understand that the choice of investing extra-capacity is better even when the sink is connected after a long time of 10 years with a high discount rate of 0.07.

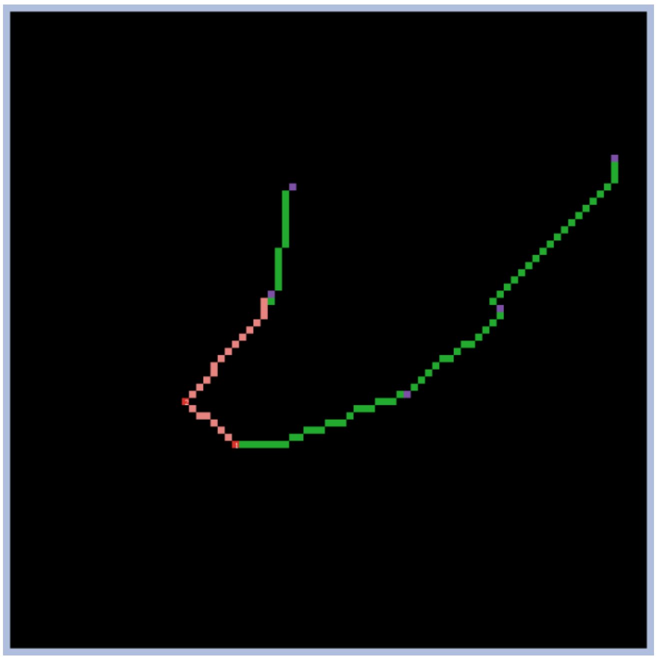


Figure 24. Investing extra-capacity on a (pink) part of the network

5.2 Model validation on comparison with the non-deterministic n:n Central Link model

For fair comparison, the *non-deterministic n:n Central Link model* is developed similarly to the development of the non-deterministic n:n ACO model as shown in figure 25. Yet the output network layouts from the non-deterministic n:n Central Link model are not expected to differ much from the deterministic n:n Central Link's output due to the fact that the central link is built at the start under no influence from the potential sinks and it pulls the ants afterwards.

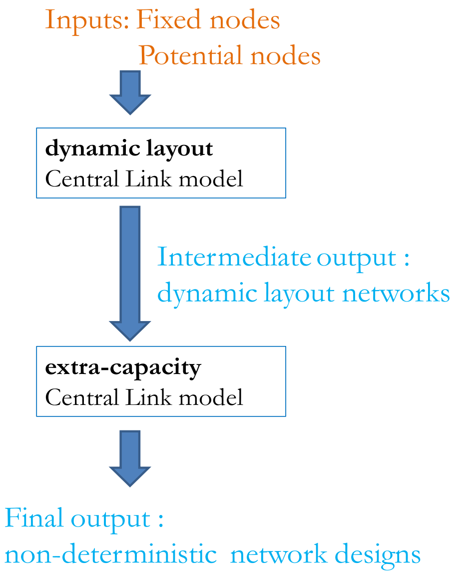


Figure 25. The non-deterministic Central Link model

While the optimal network design model for multiple source - multiple sink networks are not available, the comparison between the non-deterministic n:n ACO model and the non-deterministic n:n Central Link model makes a valid validation for the non-deterministic n:n ACO model.

5.2.1 Case description

The hypothetical case study's setup is shown in figure 26 with 2 fixed sources (red labels), 4 fixed sinks (violet labels) and 1 potential sink. Regarding the potential sink:

- Location of the potential sink is set at *inside the area and close to the nest* (case 1 - the left one) and *outside area and far from the nest* (case 2 - the right one).
- Potential demand is fixed at 5 (approx. 20% of the total demand). It is assumed that at the time of connection, the supply capacity of the biggest nest increases by 5.
- Timing: the potential sink is connected after 1 year, 5 years or 10 years.
- Discount rate r: 5% or 7%
- Capacity cost exponent β: 0 (low), 0.5 (medium) and 0.9 (high)

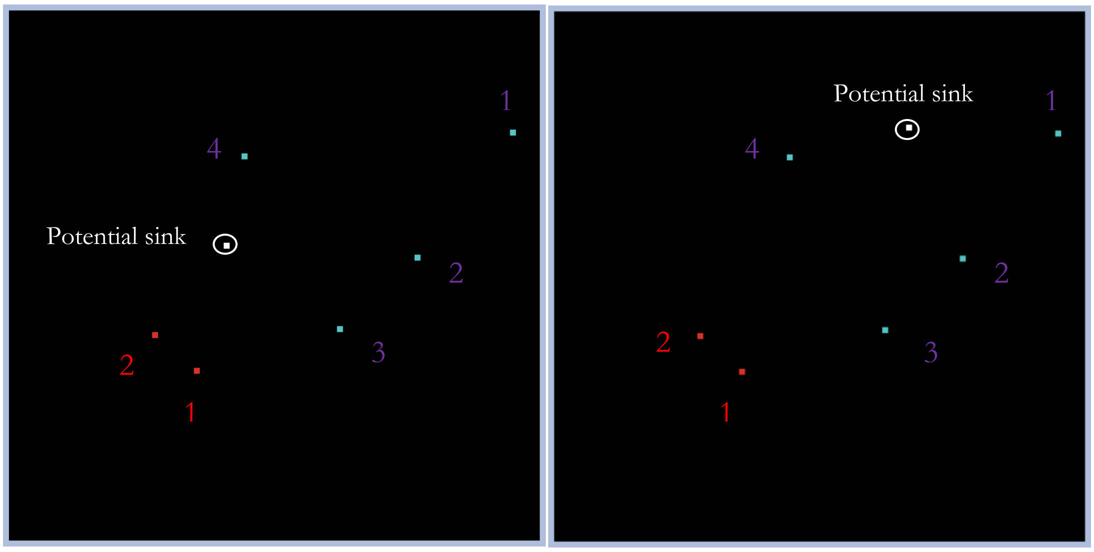


Figure 26. Two locations of the potential sink

5.2.2 The experiment setup

Since the expected returns are considered to be the same for both methods, the comparison is based on the total investment cost in present including fixed cost to connect the fixed sinks to the sources, the exercise cost of the flexibility options and the connection cost for the potential sink.

An important assumption regarding the cost of underutilization is made as the cost of underutilization is *negligible* in both methods. Even though the underutilization is known as a major discussion to the oversized design because it can incur new costs, in some cases it can also reduce the costs. For instance, in the study on the CO_2 network by Wiley et al. (2013) on the economics of using oversized and parallel pipelines, it is stated that actually the underutilization will lead to a lower compression cost for the flow of CO_2 in the pipelines, therefore it reduces the total cost.

Experiment steps are as follows.

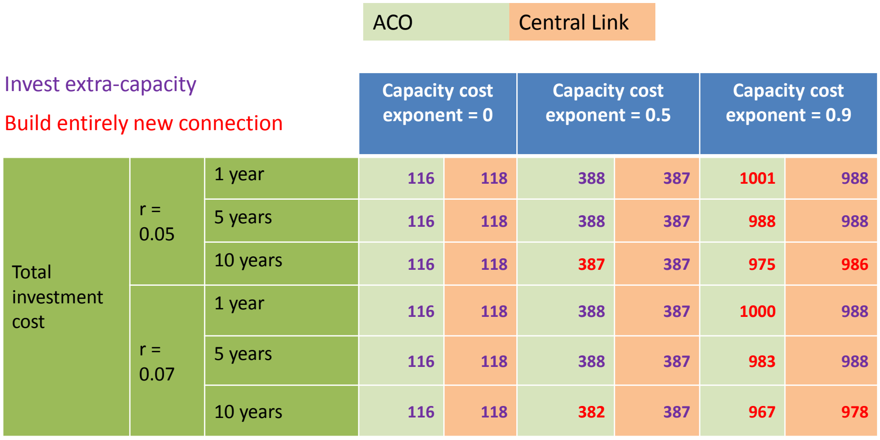
- Step 1: in the Dynamic Layout stage, 30 runs for each case were executed for both models. The best-cost network layouts for each case are selected to be the inputs for the next step.

- Step 2: in the Extra-capacity stage, 5 repetition runs were executed for each combination of timing, discount rate, capacity cost exponent. The comparison is based on the best-cost network layouts.

5.2.3 Results

Detailed results are shown below. In the total investment cost, the red numbers indicate that the decision is made to defer the investment and build an entirely new connection in future; the purple numbers indicate that the decision is made to invest extra-capacity in a part of the network.

The potential sink is close to the nest and inside the network area

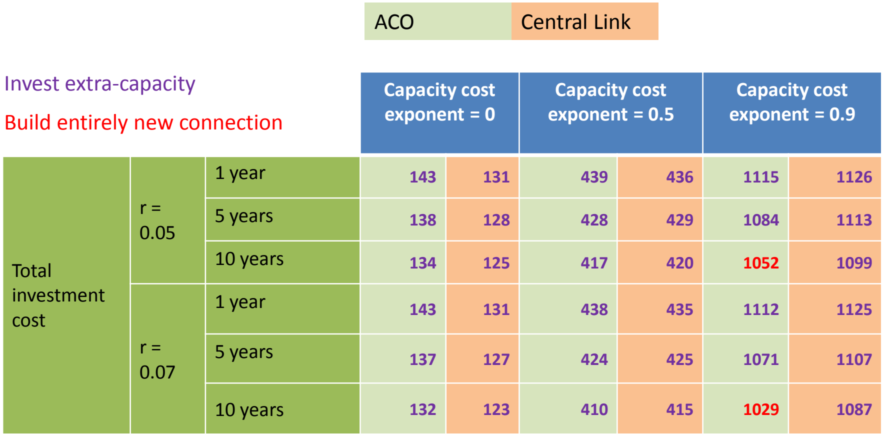


ACO | Central Link

Invest extra-capacity
Build entirely new connection

			Capacity cost exponent = 0		Capacity cost exponent = 0.5		Capacity cost exponent = 0.9	
Total investment cost	r = 0.05	1 year	116	118	388	387	1001	988
		5 years	116	118	388	387	988	988
		10 years	116	118	387	387	975	986
	r = 0.07	1 year	116	118	388	387	1000	988
		5 years	116	118	388	387	983	988
		10 years	116	118	382	387	967	978

The potential sink is far from the nest and outside the network area



ACO | Central Link

Invest extra-capacity
Build entirely new connection

			Capacity cost exponent = 0		Capacity cost exponent = 0.5		Capacity cost exponent = 0.9	
Total investment cost	r = 0.05	1 year	143	131	439	436	1115	1126
		5 years	138	128	428	429	1084	1113
		10 years	134	125	417	420	1052	1099
	r = 0.07	1 year	143	131	438	435	1112	1125
		5 years	137	127	424	425	1071	1107
		10 years	132	123	410	415	1029	1087

Figure 27. Results on the comparison of the non-deterministic ACO method and the non-deterministic Central Link method

The results per each respect are each follows.

- *The location of the potential sink*

  When the potential sink is far from the nests and outside the network area, the best dynamic network layouts are hardly different from the best network layouts from the deterministic n:n ACO model (review figure 21). In this case, the reason is that the potential sink has hardly any influence on the ants.

  When the potential sink is close to the nests and inside the network area, some paths of the network are pulled towards the potential sink. In this case study, there is one path running through the potential sink (review figure 21).

- *The timing of connection*

  When the timing of connection is estimated to be far in the future, the choice to build an entirely new connection is preferred. When the timing is set at 10 years, both methods mostly resulted in the option of not investing extra-capacity but building a new path to connect the sink after 10 years. The cost of implementing extra-capacity is rather higher than the cost of building a new path in future due to the discount rate.

  When the timing of connection is estimated to be far in future, the ACO approach is more cost-effective than the Central Link approach. The ACO network layout generally covers a wider network area. Therefore it has advantage on the cost to build the new connections.

- *The capacity cost exponent*

  The higher the capacity cost exponent is, the more the ACO approach is preferred over the Central Link approach. This result is in line with the result from the comparison between the deterministic ACO approach and the deterministic Central Link approach. The same reasoning can be applied here as well: the capacity cost exponent increases the building cost substantially and the central link's capacity is often as high as the total demand capacity from the sinks.

  The higher the capacity cost exponent is, the more the choice to build an entirely new connection is preferred over the choice to invest extra-capacity. Building a link with large capacity is very costly when the capacity cost exponent is high. For this reason, it is a good option to defer the investment to future.

- *General performance*

  In this case study, the ACO approach is more cost-effective than the Central Link approach. This result shows that the non-deterministic n:n ACO model can possibly provide better non-deterministic designs in terms of cost in comparison with the common practice network design in reality. Nonetheless there are a number of examples that the Central Link approach seems to work better. By here, the validation of the non-deterministic n:n ACO model is accomplished.

5.3 Design output evaluation

The output of the non-deterministic model have been assessed solely on the initial building cost with built-in flexibility. Although it is of the paramount importance to the network planners, there are other factors needed to be taken into account in the evaluation of the designs. Each of the infrastructure sectors requires further specific criteria besides the initial capital expenditures. In this chapter, two common criteria regarding the network construction are discussed in details along with an example of a CO_2 network for demonstration. The trade-off between the building cost and these two criteria may change the final preference of the network planners.

Taking the example introduced in the last section under the uncertainty of network participants, two best-cost network layouts with clearly different network topologies among the non-deterministic outputs are selected for the discussion as shown in the figure below.

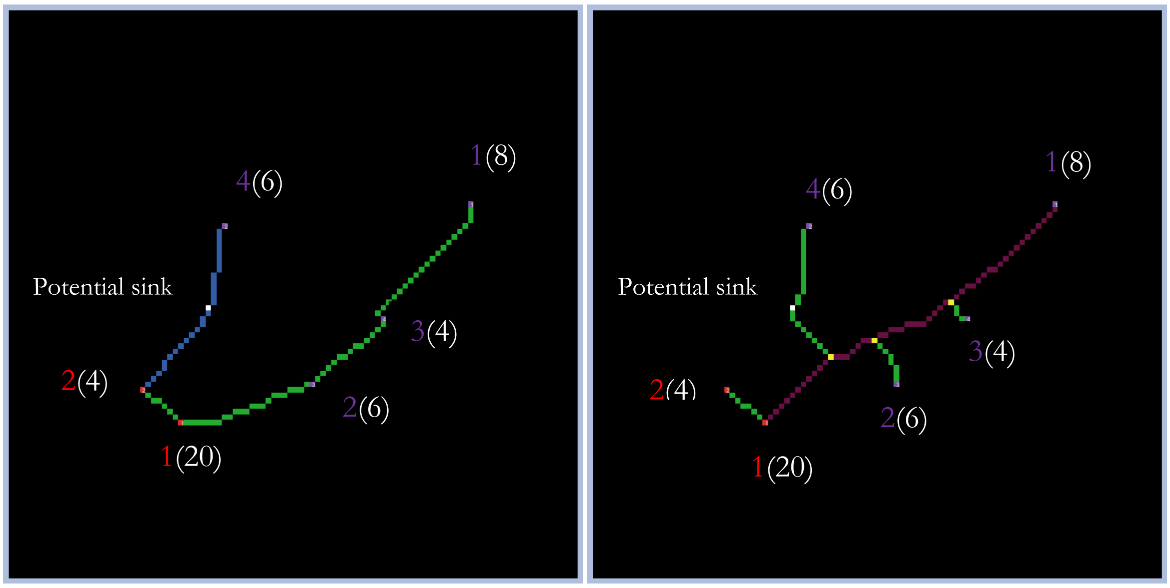


Figure 28. On the left: the ACO layout; on the right: the Central Link layout.

Both layouts run through the potential consumer (the white node), which is a great stimulation to the potential consumer to join the network. If the extra-capacity is invested, the network part from the white node to the biggest source will be enlarged accordingly to the expected potential demand. Assume that the potential demand is 5 in future, the cost for each layout are shown in the table 4.

Table 4. Cost details of two best layouts

Cost	ACO layout	Central Link layout	Cost difference (% of the lowest)
Deterministic layout (figure 9)	357	364	2%
Dynamic layout (with fixed participants)	358	373	4%
Final cost after extra-capacity investment	382	390	2%

In this case, while the cost differences after investing extra-capacity are more or less negligible (the Central Link layout is slightly more expensive). The following criteria contribute to the decision making by the network planners.

- *From the operational perspective*

  The output layouts are the inputs for further modifications to obtain the operational network layouts. Locations and the installation of pressure stations have to be added for example. Because the flow of the commodity in the Central Link layout is unidirectional whereas the commodity runs to two directions in the ACO layout, it is relatively easier to add the stations to the Central Link layout, but not certainly cheaper due to the high capacity required.

  In regard to the operational flexibility, the Central Link layout provides a better controlling configuration to implement various operational strategies in low seasons, peak seasons. The consumer 4 in the ACO layout are in some extent separated from the rest of consumer areas, which requires a duplicate effort to manage the flow towards this area.

- *From the multi-actor decision making context perspective*

  As the consumers, they are interested in having a reliable access to the network of high quality at an affordable cost. As the suppliers, they want to be connected with long-term consumers at an acceptable cost. As the network planners, their goal is to minimize the initial capital expenditure and smoothly finalize the network planning within a specified time-frame. The coordination between the actors is the prerequisite requirement to plan, build and operate the network.

  The Central Link layout requires the attendance of all the actors to the only one decision-making arena on the construction cost. It may also include the potential consumers. It is the simplest situation when the network planners are responsible for all the construction costs. In reality, the calculation can be more complicated due to the decision on how much each of the actors must contribute to the construction of the network. Taking the OCAP's CO_2 network as an example, OCAP was entirely responsible for the construction of the central pipe, yet the segments connecting the central pipe to each of the greenhouse areas were paid partly by the farmers.

  With the ACO layout, there are possibly three decision-making arenas according to the connections between the sinks and the sources. The actor break-down may lead to a longer or shorter decision-making process and likely change the bargaining power of each actor compared to the centralized arena with the Central Link layout. The sink 4 is supplied by both the supplier 1 and the supplier 2 while the other sinks are supplied only by the supplier 1. The arrangements and the contributions by the actors to the construction of the network and the future (operational) coordination among them are of the two major topics in the decision-making process.

  - The first arena includes the consumer 4, the supplier 1, the supplier 2 and the network planner
  - The second arena consists of the supplier 1, the supplier 2 and the network planner
  - The last arena includes the consumer 1, the consumer 2, the consumer 3, the supplier 1 and the network planner.

Summary of chapter 5

The non-deterministic n:n ACO model consists of two sub-models: the dynamic layout ACO sub-model to implement the dynamic layout based flexibility option and the extra-capacity ACO sub-model to implement the extra-capacity based flexibility option. The comparison between the non-deterministic n:n ACO model and the non-deterministic n:n Central-Link model resulted in the out-performance of the ACO approach. Nonetheless, the network planners need to take the other factors such as the operational perspective and the required coordination among the involved actors into account before making the final decision on the network design.

Chapter 6. The design approach, conclusions and discussions

This chapter first concludes the research, starting with the final design approach as the research goal in section 6.1 and followed by the evaluation of the approach in section 6.2. Section 6.3 provides recommendations for future research. Finally, section 6.4 closes the thesis with reflections on the limitation of the research methodology and assumptions.

6.1 The final design approach

The final design approach is visualized by the flow diagram below, which is also the research flow. This is a generic design approach. Its application to a specific networked infrastructure requires minor changes to fit the specific requirements of the network.

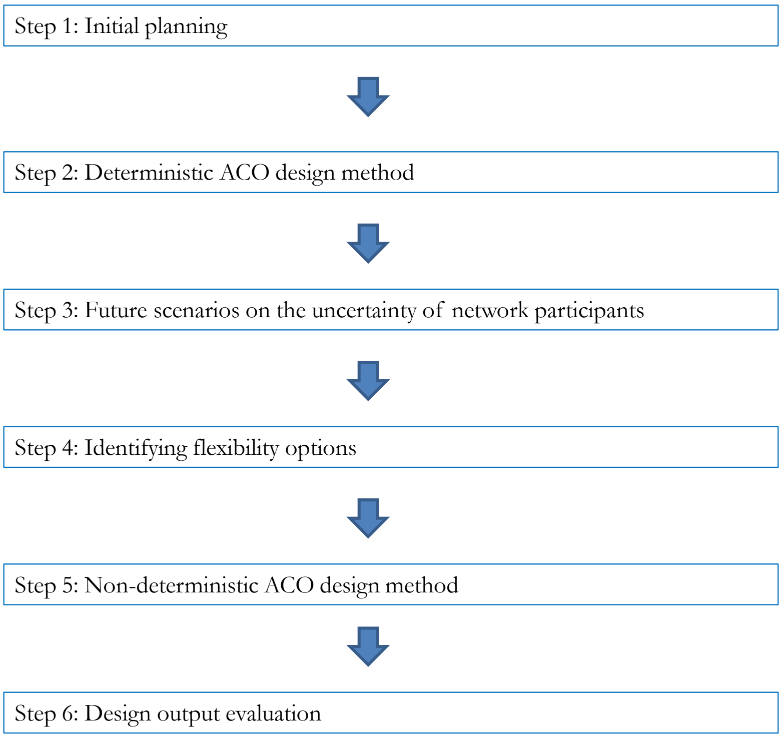


Figure 29. The design approach

The approach is to design a network connecting multiple sources and multiple sinks under the uncertainty of network participants. The initial planning such as the number of nodes, their locations, the required capacities and the capacity cost exponent are the inputs of the design model. With the deterministic ACO design method, cost-minimized rigid network layouts are generated. The next step is to identify the uncertainties regarding to the network participants and the flexibility options. With the non-deterministic ACO design method, the initial building cost is minimized while providing space for real options to deal with the future changes of network participants. The network planners are given a multitude of design outputs with different network topologies and investment costs. They include potential dynamic architectural layouts to prepare for new connections in the future, and they also support the decision making on whether or not it is beneficial to invest extra-capacity in several parts of the network. Finally, the design output evaluation needs to be executed so that criteria other than the initial building cost are covered.

6.2 Evaluation

The proposed approach is the first design approach for generic networked infrastructures. In addition, it is the first available non-deterministic design approach that deals with the uncertainty on network participants in multiple source - multiple sink networks. It is clearly applicable to solve the design problems in one source - multiple sink networks as well. The approach is evaluated on the following criteria.

- *Methodology*
  The Ant Colony Optimizations approach proved to be effective in solving the routing problem in the design of multiple source - multiple sink networks. While there are several studies applying the Top-Down approach for the design of *n : n* networks, this study is the *first attempt* to apply a Bottom-Up approach to deal with the same problem. During the extension from the *1 : n* model to the *n : n* model, the agent-based implementation of Ant Colony Optimization is experienced to be intuitive and easily extensible. The extension to add multiple sources to the models and the embedding flexibility options into the models were done without too much effort and time. The implementation of the Central Link approach into the Central-Link models is a clear example for the great customizability of the approach.

- *Quality of design solution*
  The deterministic n:n ACO model was developed based on the original 1:n ACO model introduced by Heijnen et al. (2014). The results of the comparison between the deterministic ACO design method and both the theoretical GG method and the practical Central-Link method validated the feasibility in finding cost-minimized network layouts of the deterministic ACO design method. The GG method works slightly better in locating cheaper networks, yet the deterministic ACO method is rich in providing numerous network topologies. The results also indicate that in a few cases, the Central Link approach is preferable and in most of the cases, the ACO approach can find more cost-effective network layouts.

  The non-deterministic ACO method is an extension of the deterministic ACO method with the incorporation of the flexibility options. The non-deterministic n:n ACO model was built on the two flexibility options identified in *n : n* networks: dynamic layout based flexibility and extra-capacity based flexibility. The model was validated by comparing its performance with the practical Central Link approach's model. In most of the cases, the non-deterministic n:n ACO model generated better cost-minimized network layouts. The output design of the model provides supportive visual answers to the network planners on the decision about the network's architectural layout and the extra capacity required for future connections.

- *Decision-making support*
  Reality often shows that the judgments of the involved actors on a design are likely to differ widely to accord with their interests, set-values and predilections (Rittel & Webber, 1973). Their assessments of the design are not definite but more likely expressed as *good enough* or *bad*. In this study, the initial investment cost is the main metric to evaluate the output designs. In order to bridge the design problem sphere and the decision making sphere, the design output evaluation step of the approach guides the network planners to take into account further criteria, which are the operational perspective and the necessary coordination among the actors in the multi-actor decision making context. Eventually, a design that covers a broader set of criteria has more chance of being selected.

Another supportive feature of this approach is that it equips the network planners with both the deterministic and the non-deterministic designs. Available approaches before this study either deal with only one type among the deterministic and the non-deterministic design or do not draw a clear line between the two designs. Due to the barriers for the implementation of the real options (Lander & Pinches, 1998; Herder, et al., 2011), the non-deterministic design faces the risk of being forsaken by the decision makers. Besides the intuitiveness, the proposed approach alleviates this risk by developing the non-deterministic design based on the deterministic design. It means that the trade-off between the flexibility to deal with the uncertainty and the increase in the initial capital expenditure to embed the flexibility options into the deterministic design are made apparent. In addition, the network planners are also able to choose between adopting only the capacity based flexibility or adopting both the capacity based flexibility and the dynamic-layout based flexibility. This adds more freedom for the network planner in applying the approach. Equipped with the visually intuitive outputs from the approach, the network planners have more chance to facilitate a smooth communication and negotiation with the involved actors at the planning stage.

6.3 Future research

Based on the main findings of this research, three recommendations for future research are formulated.

- *A user-friendly application tool*
  Scientific approaches are not always used by real-world decision makers partly because of their complicatedness and partly because they lack of realistic requirements. The ACO models are equipped with intuitiveness, extensibility and customizability. By embedding these ACO models to the proposed design prototype by Mattaparthi (2015), it is feasible to create a multi-phase user-friendly application tool with which the network planners can customize from the initial inputs of locations and supply/demand capacities to the practical decision-making criteria according to each specific infrastructure sector. A large data set on the design criteria for each infrastructure types can be collected through surveys and expert-interviews. Upon the achievement of such a tool, the gap between scientific knowledge and practical decision making is bridged and the network planners can directly bring it to the decision-making table for negotiation with the involved actors.

- *The design of network including hybrid nodes*
  In this study, every node is either a supplier or a consumer. If their roles change during the operation of the network, e.g. there are special nodes that are suppliers in this moment but will be consumers in other moments, how can these special nodes be connected to the rest of the network? It is possible that these nodes require multiple paths connecting them to the network for in-flows as well as for out-flows. Connecting the sinks to one of these nodes will not guarantee the continuity of the supply of the commodity. Not connecting the sinks to any of these nodes is a safe solution, but it is not efficient, especially if the total supply cannot satisfy the total demand at some periods. In this case, there is a need to set up the priority to be connected to the network for each node. Moreover, the links in the network need to become directional in order to distinguish between the in-flow to and out-flow from a node.

- *Including the expected return in the modeling*
  This study, within the available time limit, assumed that the expected returns are the same when comparing the deterministic ACO method and the deterministic Central Link method.

However important issues such as the underutilization cost should not be ignored. Taking the expected return into calculation gives a more accurate evaluation of the different network layouts. One example of the expected return formulation can be referred to (Heijnen, et al., 2013).

6.4 Reflection

Most of the currently proposed design models and approaches are either not well known or understood by real-world network planners and decision makers. For example, Gasunie is in charge of planning the pipelines to connect multiple off-shore small fields to the Dutch gas grid as introduced in chapter 1. The tools used by the company's network planners are capacity and flow related linear optimization models such as SIMONE[6] (Chung, 2015). The physical layout of the pipelines was designed mainly based on experience. Because of this, the possible future benefits, if the flexible layout had been implemented. were overlooked. Lander & Pinches (1998) identified the primary reasons why real-option models are not widely used in practice in many areas such as manufacturing, natural resources and real estate, and Herder et al. (2011) identified major barriers for the implementation of a real option approach focusing on infrastructure projects. Based on the findings from these two studies, this last section discusses the most relevant problems to the future applicability of the proposed approach in this research.

6.4.1 Problems of modeling assumptions

Regardless whether the Top-Down or Bottom-Up approach is applied in the modeling, the famous quote by George Box in 1979, "all models are wrong, some are useful", is always true. The required assumptions for the modeling may be breached in practical applications. However, the extensibility of the ACO method is a great advantage to fit the models to real-world requirements.

- *The cost function*
  The (non)deterministic ACO models were created to solve the (non)deterministic design problems, which mainly relate to the initial building costs of a networked infrastructure. As discussed in the design output evaluation, there are many factors other than the initial building costs that should be taken into account. The generic cost function used in the model, which is a function of the length and the capacity, serves the generic infrastructural networks well but could not capture the realistic requirements by the decision makers. E.g. in CO_2 networks, the total cost considerably depends on the compression cost, which in turn depends on the diameter and the material of the pipeline. In addition, whereas the cost function favors the merge of small paths into a bigger path when the capacity cost exponent is small in order to reduce the building cost, it may incur extra costs due to a higher compression cost. Actually, a big pipeline with high-pressure CO_2 is another concern of the decision makers because it carries high risk of leakage. To sum up, an extended cost function incorporating extra practical criteria are required to make the models more useful. In addition, the decision-making rules influencing the behaviors of the ants need to be changed accordingly.

- *The multi-actor system*
  The developed methods in the approach assumed that all the suppliers are equal and all the consumers are equal. The layout is optimized for all the network participants. In real-world contexts, the sources (sinks) can be differentiated on many aspects. E.g. in the penetration of

---

[6] http://www.simone.eu/simone-simonesoftware-onlinesimulation.asp

renewable energy sources, hydro, wind or solar sources with assumed similar supply capacities could be prioritized in different orders according to the current government energy policies. Some of the sources may have a higher priority to be connected to the electricity grid than the others. Therefore, the network needs to be optimized not globally but for those network participants. These real-world characteristics can be reflected in the method with additional efforts to modify the parameter settings of the ACO models. The influence of the nests on the ants can be easily changed by modifying the setting of nest scents. The food sources also can have influence on the ants if similar settings with food scents are added to the models.

6.4.2 Problems of capacity

The network's planning team is constrained in available data, quantity and quality of personnel, time and money. Due to the intuitiveness of the methodology, it can be surmised that the network planner would not need a long time to understand the essence of the approach and to get used to the models. The cost of utilizing the approach mainly incur from the cost of data gathering activities. As the inputs, the approach first requires the basic set of data including the locations of the certain nodes and their demand and supply capacities. Any design model/approach would require at least the same set of data. The research on the locations and the demand and/or supply capacities of the potential network nodes is an extra burden to the team. Recalling the case of off-shore gas small fields, the exploration of the reserves of the potential fields is costly in terms of time and personnel. The most difficult situation would arise if the decision makers are not convinced of the additional benefits of the approach in the beginning. As a consequence, the extra data collecting activities would face the risk to be aborted.

Appendix A. Path trimming algorithms

Due to the influence from multiple nests, the ants wiggle multi-directionally when returning to one of the nests. Therefore the paths may include useless patches or loops. A simple example is shown in the figure below.

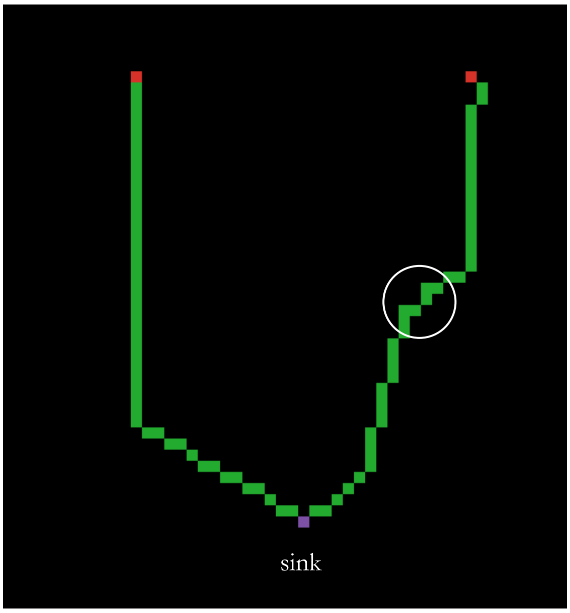


Figure 30. Useless patches (in the white circles) included in the path

- The procedure to remove the useless patches is as follows

```
1. i = 1                          ; i is the counter
2. WHILE (i  <= n - 2)            ; n is the total length of the path
3.   IF (patch[i] AND patch[i + 1] are neighbors) AND (patch[i] AND patch[i + 2] are neighbors) THEN
         Remove patch[i + 1] from the path
         Increment i by 1
4. END WHILE
```

Appendix B. Model agents and Parameterizing the ACO models

The agents of the agent-based Netlogo models are the ants and the patches. The parameters of the environment have a strong influence on the internal states and the behaviors of the agents.. Below is the descriptions of the internal variables of the agents and the summary of important environment parameters beside the *small-margin-get-on-network* and the *small-margin-get-off-network*. A good understanding about the influences of these parameters on the agents gives the model users freedom to customize the modeling environment to fit their own preferences.

Table 5. Important internal state variables of the agent

Model agent	*State variable(s)*	*Description*
ants	path	It is a list of the patches that the ants walk on in their ways returning to a nest from the food source. After reaching the nest, the cost of the path will be calculated and compared with the current cheapest path.
	food-source-found	The label number of the food source that the ants found.
	carrying-food?	This Boolean variable indicates if the ants are on their way returning to the nests (true) or walking to find a food source (false).
patches	nest?, nest-number, food-source	This group of variables indicates the role of the patch: a nest or a food source, with its label number.
	network?, which-nests, which-food-sources, network -capacity	This group of variables indicates the properties of a patch if it belongs to a network path. Is it a network patch? If so, the network to which nests, to which food sources and with which capacity.
	nest-scent	It is a list of the strength of multiple nest scents from multiple nests at the patch.
	central-link-scent	It is the strength of the central-link scent at the patch.

Table 6. Important environment parameters of the ACO models

Parameter	*Description*	*Influence*
test-case	This parameter has three values representing for 3 different options: ACO approach to build the network from scratch, ACO approach to build the first random path between a source and a sink, and the Central Link approach	This parameter indicates the approach that the user wants to implement to design the network.
patience	The number of ants that need to return from a food source without finding a cheaper path before the cheapest path is fixed as network path	Increasing the patience would improve the cost of the network paths, yet slows down the simulation.
nest-scent-power	This variable represents for the full strength of nest scent right at the nest	This parameter is chosen to be the basic parameter to which other parameters are adjusted. It is fixed at 200 when the maximum x-coordinate is 100 and the maximum y-coordinate is 100.
population	The total number of ants in the simulation	A high number of ants is used to balance the computing time.
potential-sink-x	The x-coordinate of the potential sink	This parameter allows the user to test all the possibility of the location/demand/connecting timing of the potential sink.
potential-sink-y	The y-coordinate of the potential sink	
potential-sink-demand	The demand of the potential sink	
year-time	The connection timing of the potential sink	
discount-rate	The value of the discount rate in net present value calculation	This parameter is to test the all the possibility of the discount rate.

References

Ajah, A. N. & Herder, P. M., 2005. Addressing flexibility during Process and Infrastructure Systems Conceptual Design: Real Options Perspective. *Systems, Man and Cybernetics, 2005 IEEE International Conference,* Volume 4, pp. 3711 - 3716.

Berghout, N., 2014. Designing cost-optimal CCS configurations for an industrial cluster. In: R. de Vos, ed. *Linking The Chain.* Zutphen: CPI, pp. 91 - 93.

Bijker, W., Hughes, T. & Pinch, T., 1987. *The social construction of technological systems: new directions in the sociology and history of technology.* Cambridge: MIT Press.

Chung, J.-W., 2015. *Gas-transport network planning optimization models* [Interview] (06 05 2015).

Correljé, A. F., 2015. *The multiple source - multiple sink networks characteristics* [Interview] (18 2 2015).

Correljé, A., van der Linde, C. & Westerwoundt, T., 2003. *Natural Gas in the Netherlands: from cooperation to competition?.* Amsterdam: Drukkerij Hooiberg bv.

de Neufville, R. & Scholtes, S., 2011. *Flexibility in Engineering Design.* 1 ed. Cambridge, Massachusetts: The MIT Press.

Escudero, L. F., Quintana, F. J. & Salmeron, J., 1999. CORO, a modeling and an algorithmic framework for oil supply, transformationa and distribution optimization under uncertainty. *European Journal of Operational Research,* Volume 114, pp. 638 - 656.

Feitelson, E. & Salomon, I., 2000. The implications of differential network flexibility for spatial structures. *Transportation Reearch Part A,* Volume 34, pp. 459 - 479.

Gil, N., 2007. On the value of project safeguards: Embedding real options in complex products and systems. *Research policy,* Volume 36, pp. 980 - 999.

Goel, V., Grossmann, I. E., El-Bakry, A. S. & Mulkay, E. L., 2006. A noval branch and bound algorithm for optimal development of gas fields under uncertainty in reserves. *Computers and Chemical Engineering,* Volume 30, pp. 1076 - 1092.

Heijnen, P., Chappin, E. & Nikolic, I., 2014. Infrastructure network design with a multi-model approach: Comparing Geometric Graph Theory with an Agent-Based Implementation of Ant Colony Optimization. *Journal of Artificial Societies and Social Simulation,* 17(4), pp. 1 - 21.

Heijnen, P. W., 2014. *Expert interview: On the literature review of infrastructure network design* [Interview] (28 Nov 2014).

Heijnen, P. W., Stikkelman, R. M., Ligtvoet, A. & Herder, P. M., 2013. Maximising the Worth of Nascent Networks. *Networks and Spatial Economics,* 14(1), pp. 27 - 46.

Heijnen, P. W., Stikkleman, R. M., Ligtvoet, A. & Herder, P. M., 2011. Using Gilbert networks to reveal uncertainty in the planning of multi-user infrastructures. *International Conference on Networking, Sensing and Control,* pp. 371 - 376.

Herder, P. M. et al., 2011. Buying real options - Valuing uncertainty in infrastructure planning. *Futures,* Volume 43, pp. 961 - 969.

Herder, P. & Pulles, K., 2011. Flexible Gas Infrastructures. In: *Engineering Asset Management.* London: Springer, pp. 655 - 663.

Hu, Y. et al., 2006. ACO-Steiner: Ant Colony Optimization Based Rectilinear Steiner Minimal Tree Algorithm. *Journal of Computer Science and Technology,* 21 (1), pp. 147 - 152.

Kabirian, A. & Hemmati, M. R., 2007. A strategic planning model for natural gas transmission networks. *Energy Policy,* 35(11), pp. 5656 - 5670.

Kazmierczak, T., Brandsma, R., Neele, F. & Hendriks, C., 2009. Algorithm to create a CCS low-cost pipeline network. *Energy Procedia,* Volume 1, pp. 1617 - 1623.

Lander, D. M. & Pinches, G. E., 1998. Challenges to the Practical Implementation of Modeling and Valuing Real Options. *The quarterly review of economics and finance,* 38(Special Issue), pp. 537 - 567.

Lin, J., de Weck, O., de Neufville, R. & Yue, H. K., 2013. Enhancing the value of offshore developments with flexible subsea tiebacks. *Journal of Petroleum Science and Engineering,* Volume 102, pp. 73 - 83.

Lin, J., Yue, H. K., Neufville, R. d. & Weck, O. d., 2013. Enhancing the value of offshore developments with flexible subsea tiebacks. *Journal of Petroleum Science and Engineering,* Volume 102, pp. 73 - 83.

Liu, Q. & Wang, C., 2012. Multi-terminal pipe routing by Steiner minimal tree and particle swarm optimization. *Enterprise Information Systems,* 6(3), pp. 315 - 327.

Maier, H. R. et al., 2003. Ant colony optimization for design of water distribution systems. *Journal of water resources planning and management,* 129 (3), pp. 200 - 209.

Marcoulaki, E. C., Papazoglou, I. A. & Pixopoulou, N., 2012. Integrated framework for the design of pipeline systems using stochastic optimisation and gis tools. *Chemical Engineering Research and Design,* Volume 90, pp. 2209 - 2222.

Martinez-Anido, B. C., 2013. The future needs for European Cross-Border Transmission Capacity. In: *Electricity Without Borders.* Delft: Next Generation Infrastructure Foundation, pp. 39 - 54.

Mattaparthi, G. S., 2015. *Decision support system for energy network design: A socio-technical analysis* [To be published]

Melese, Y., Heijnen, P. W. & Stikkelman, R. M., 2014. Designing networked energy infrastructures with architectural flexibility. *Procedia Computer Science,* Volume 28, pp. 179-186.

Middleton, R. S. & Bielicki, J. M., 2009. A scalable infrastructure model for carbon capture and storage: SimCCS. *Energy Policy,* Volume 37, pp. 1052 - 1060.

Morbee, J., Serpa, J. & Tzimas, E., 2011. Optimal planning of CO2 transmission infrastructure: The JRC Infra CCS tool. *International Conference on Greenhouse Gas Control Technologies,* Volume 4, pp. 2772 - 2777.

Ramirez, A., Hendriks, C. & van Engelenburg, B., 2014. Integrating the CCS chain. In: R. d. Vos, ed. *Linking the chain - Integrated CATO2 knowledge prepares for the next step in CO2 Capture & Storage.* Zutphen: CPI, pp. 81 - 94.

Rittel, H. W. J. & Webber, M. M., 1973. Dilemmas in a general theory of planning. *Policy Sciences,* Volume 4, pp. 155 - 169.

Thomas, D. & Weng, J., 2006. Minimum cost flow-dependent communication networks. *Networks,* 48(1), pp. 39 - 46.

Ukkusuri, S. V., Mathew, T. V. & Waller, S. T., 2007. Robust Transportation Network Design Under Demand Uncertainty. *Computer-Aided Civil and Infrastructure Engineering,* Volume 22, pp. 6 - 18.

van Tol, P. J., 2015. *TuDelft Institutional Repository.* [Online]
Available at: http://repository.tudelft.nl/view/ir/uuid%3A4ce5064b-681e-410b-b909-937d8a6f9b05/
[Accessed 04 05 2015].

Veenstra, J. W. & Limbeek, J. T., 2014. *OCAP - the story of industry-greenhouse CO2 network in the Netherlands* [Interview] (16 12 2014).

Veltin, J. & Belfroid, S., 2013. *Flow assurance for studies for CO2 transport,* s.l.: CATO2 Deliverable.

Weihs, F. G. A., Wiley, D. E. & Ho, M., 2011. Steady-state optimisation of CCS pipeline networks for cases with multiple emission sources and injection sites: south-east Queensland case study. *Energy Procedia,* Volume 4, pp. 2748 - 2755.

Wiley, D. E., Wang, Z., Cardenas, G. & Weihs, G. F., 2013. Optimal pipeline design with increasing CO2 flow rates. *Energy Procedia,* Volume 37, pp. 3089-3096.

Zhao, T. & Tsend, C. L., 2003. Valuing flexibility in Infrastructure Expansion. *Journal of Infrastructure systems,* Volume 9, pp. 89 - 97.