1-1-2020

Carbon-Dioxide Pipeline Infrastructure Route Optimization And Network Modeling For Carbon Capture Storage And Utilization

Karthik Balaji

CARBON-DIOXIDE PIPELINE INFRASTRUCTURE ROUTE OPTIMIZATION AND NETWORK MODELING FOR CARBON CAPTURE STORAGE AND UTILIZATION

by

Karthik Balaji

Bachelor of Engineering, University of Mumbai, 2014

Master of Science, University of Southern California, 2016

A Dissertation

Submitted to the Graduate Faculty

of the

University of North Dakota

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Grand Forks, North Dakota

December

2020

Name: Karthik Balaji

Degree: Doctor of Philosophy

This document, submitted in partial fulfillment of the requirements for the degree from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

Dr. Minou Rabiei

Dr. Hui Pu

Dr. Kengang Ling

Dr. Vamegh Rasouli

Dr. Gregory Vandeberg

This document is being submitted by the appointed advisory committee as having met all the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

Chris Nelson
Dean of the School of Graduate Studies

12/1/2020

Date

iii

PERMISSION

Title: Carbon-Dioxide Pipeline Infrastructure Route Optimization and Network Modeling For Carbon Capture Storage And Utilization

Department: Petroleum Engineering

Degree: Doctor of Philosophy

In presenting this dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my dissertation work or, in her absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Karthik Balaji

12/01/2020

ACKNOWLEDGEMENTS

The dissertation being as complicated as it is to write, required the help and support of several parties. The foremost acknowledgment goes to my advisor, Dr. Minou Rabiei, whose support and inputs made the research and writing of the dissertation a possibility. Her belief and leadership have always been a significant driving force. Further I would also like to acknowledge the support of the Petroleum Engineering department at the University of North Dakota for providing me with the appropriate resources for research and the North Dakota Industrial Commission for their financial support of the work (Grant G-51-02). Acknowledgement also extends to the department of Geography and Geographic Information Sciences for the technical and software support. I would also like to thank Dr. Vamegh Rasouli, Dr. Hui Pu, Dr. Kengang Ling and Dr. Gregory Vandeberg for providing their valuable inputs and time for editing the work.

My education and research would not be possible without the moral support of my mom (Jayanthi Balaji) , dad (Balaji Ramadurai) and sister (Aishwarya Balaji). They made sure that despite any hardship or ailment, I continued my education. While in Grand Forks Anai Caparo Bellido and Pavankumar Challa Sasi made sure I never waivered from my goals. I would also like to thank my old friends Ramesh Kamath, Vignesh Pai, Venkatesh Naik and Karan Chheda. Lastly, the support and enthusiasm shown by all my friends, colleagues and mentors during the dissertation has been nothing short of spectacular

TABLE OF CONTENTS

Contents

TABLE OF FIGURES

TABLE OF TABLES

ABSTRACT

Carbon capture, utilization, and storage (CCUS) is a technology value-chain which can help reduce CO_2 emissions while ensuring sustainable development of the energy and industrial sectors. However, CCUS requires large-scale deployment of infrastructure for capturing feasible amounts of CO_2 that can be capital intensive for stakeholders. In addition, CCUS deployment leads to the development of extensive pipeline corridors, which can be inconsistent with the requirements for future CCUS infrastructure expansion.

With the implementation and growth of CCUS technology in the states of North Dakota, Montana, Wyoming, Colorado and Utah in mind, this dissertation has two major goals: (a) to identify feasible corridors for CO_2 pipelines; and (b) to develop a CCUS infrastructure network which minimizes project cost. To address these goals, the dissertation introduces the CCSHawk methodology that develops pipeline routes and CCUS infrastructure networks using a variety of techniques such as multi-criteria decision analysis (MCDA), graph network algorithms, natural language processing and linear network optimization. The pipeline route and CCUS network model are designed using open-source data, specifically: geo-information, emission quantities and reservoir properties.

The MCDA of the study area reveals that North Dakota, central Wyoming and Eastern Colorado have the highest amount of land suitable for CO_2 pipeline corridors. The optimized graph network routing algorithm reduces the overall length of pipeline routes by an average of 4.23% as compared to traditional routing algorithms while maintaining low environmental impact. The linear optimization of the CCUS infrastructure shows that the cost for implementing the technology in the study area can vary between \$24.05/tCO_2 to \$42/tCO_2 for capturing 20 to 90MtCO_2. The analysis also reveals that there would be a declining economic impact of existing pipeline infrastructure on the future growth of CCUS networks ranging between 0.01 to 1.62\$/tCO_2 with increasing CO_2 capture targets.

This research is significant, as it establishes a technique for pipeline route modeling and CCUS economic analysis highly adaptable to various geographic regions. To the best of the author's knowledge, it is also the first economic analysis that considers the effect of pre-existing infrastructure on the growth of CCUS technology for the region. Furthermore, the pipeline route model establishes a schema for considering not only environmental factors but also ecological factors for the study area.

CHAPTER 1

CCUS Value Chain and Network Analysis: Introduction

1.1 Introduction

Life in the 21st century is marked by significant usage of electronic devices and mass-produced goods. The comfort of the usage of products and devices is enabled due to the energy and industrial sectors. However, these sectors are also significant contributors to overall emissions worldwide (Intergovernmental Panel on Climate Change [IPCC], 2018). The advancement of more efficient systems and minimization of emissions (land, water, noise and air) is a critical factor influencing the development of technology and policy in the energy and industrial sectors. Amongst the emission problems, Greenhouse Gases (GHG) leading to climate change are of particular interest worldwide. The 2015 United Nations Climate Change Conference, COP 21, held in Paris, France, stated that several countries acknowledged the importance of emission control and its' significant negative impacts, making the control of these GHG emissions even more important.

GHG emissions are composed of several different gases, with the biggest proportion consisting of Carbon Dioxide (CO_2), Methane (CH_4), Nitrous Oxide (NO_x) and Fluorinated Gases (United States Environmental Protection Agency [USEPA], 2020a). The United States Environmental Protection Agency (USEPA) database on GHG emissions and sinks states that of the total emissions in 2018, 81% of the total emissions were composed of CO_2 equivalent to 6677 Million tons. This increasing concern over CO_2 emission levels has led to increasing awareness of the potential for sustainable development of the energy and industrial sectors (International Energy Agency [IEA], 2011). Despite these heightened concerns, CO_2 emission levels have increased from 521 Mt of CO_2 in 1990 to 840 Mt of CO_2 globally; an increase of 61.22% in CO_2 emissions (IEA, 2019). The per capita emission of CO_2 despite worldwide concerns have risen from 3.9-ton CO_2/capita in 1990 to 4.4-ton CO_2/capita in 2017 (IEA, 2019a). In the United States of America (USA), the overall emission increase has been lower as compared to worldwide statistics by 81.21% (IEA, 2019a). Amongst all the contributing factors to emissions, the energy sector in 2017 contributed 41.64% of overall CO_2 emissions as compared to the transportation sector at 24.61%, industrial sector at 19.06% and residential emissions at 5.91% (IEA,2019a).

The current CO_2 capture goals set by various international conventions aim at reducing emission levels by 20% by 2025 and by 30% by 2030 (UN, 2015). Various strategies have been utilized to reduce CO_2 emissions including the use of many new technologies. Some of the strategies for CO_2 reduction mentioned in the United States Department of Energy (DOE) report (US DOE, 2017) include Carbon Capture, Utilization and Storage technology (CCUS), Natural Gas Combined Cycle technology (NGCC), Nuclear Light Water Reactors, Land-Based Wind Turbines, Offshore Wind Turbines, Utility-scale Solar Photovoltaic (PV) technology and Concentrating Solar Power (CSP) technology. Amongst them, CCUS has received significant attention in recent years especially following the Intergovernmental Panel on Climate Change (IPCC) report of 2005. Even in more recent reports credence has been provided to the potential of CCUS to reduce carbon dioxide emissions (IPCC, 2018)

Carbon Capture and Storage (CCS) is a technology, composed of capturing CO_2 at the source (energy/industrial) site, separating the CO_2 from other gases, transporting the CO_2 over large distances and further storing the CO_2 in underground geological storage sites over several years. CCUS has an additional stage known as Utilization, where the captured CO_2 is used in the oil and gas (O&G) industry for CO_2 Enhanced Oil Recovery (CO_2 EOR) projects. CO_2 can also be used as raw material for refrigeration, food processing, welding amongst other applications. CO_2 also increases productivity of product such as in the production of Urea, carbonates, and acids . This technology does not have a long history and has only recently transitioned to commercial projects after several field test opportunities (IEA, 2019b). Nevertheless, the development of CCUS has been steady, with the earliest test project beginning in 1996 at the Sleipner, Norway O&G offshore field, storing 22 million tons of CO_2 till 2017 (Ringrose, 2018). A second project was started at the Salah field, Algeria in 2004 with an estimated 17 million tons equivalent of CO_2 storage capacity (Ringrose et al., 2013). Currently, only a few commercially viable active CCUS projects exist in the world including the Petro-Nova project (USA) and the Boundary Dam project (Canada) (IEA, 2019b). However, many commercially viable CCUS projects are under development or recently developed worldwide including projects in the United Kingdom, Australia, Brazil, Netherlands, the United Arab Emirates, China, and South Korea (IEA, 2019b).

A CCUS project involves several processes, with different stakeholders, and various factors that interact with each other in multiple fashions (International Risk Governance Council [IRGC], 2008). Multiple countries are developing a dedicated regulatory system capable of handling CCUS-scale multi-decade projects. Such regulatory frameworks need to satisfy certain requirements such as dealing with post CO_2 injection site management, CO_2 tax incentives and handling community acceptance (IRGC, 2008). Considering that the economic scale of these projects run in the several billion USD in terms of both capital expenditure and annual operational expenditure, CCUS processes require a proper decision making workflow to plan and develop the infrastructure in a timely and efficient manner. The major components in CCUS infrastructure are sources, sinks and pipeline. The points at which CO_2 can be captured in known as a source, the points at which CO_2 is either used or stored is known as sinks and transportation of CO_2 is carried out through pipelines. In this complex network of sources, sinks and pipelines, each source and sink have a specific goal and pipelines are the means of balancing out these goals. An efficient decision-making workflow should be able to answer fundamental questions related to the types of the sources to be deployed and utilization processes to be used as well as the location of the sinks and logistics of the CO_2 transportation. Such an integrated decision-making workflow would enable stakeholders to make informed decisions related to feasible deployment of CCUS infrastructure in a given region in a suitable fashion.

In this work, we develop an integrated workflow, called CCSHawk, for CCUS infrastructure planning using graph theory and network analysis. The core aspect of this workflow lies with the generation of potential pipeline routes which needs to be analyzed to find the safest and most economical means to get from one point to another. CCSHawk uses graph analysis in order to delineate the best combination of sources, sinks and pipelines to enable the set-up of a viable CCUS infrastructure network for North-Central USA. The optimization problem we consider in this study can be stated as: industry A needs to capture X amount of CO_2 annually and deliver it to site B for usage through a pipeline of diameter F with certain technical specifications.

For this purpose, we use mixed-integer linear programming to determine the best combination of these features in a quantitative sense to achieve the goals set by the user.

1.2 CCUS Value Chain

The CCS technology includes the concept of capturing CO_2 from emissions at the source (pre or post combustion), transporting the CO_2 via pipelines and then injecting the carbon dioxide CO_2 through wells into subsurface reservoirs (IEA, 2013). Although, this concept theoretically holds true, economic viability of a pure storage facility is not feasible in many cases, prompting geologic storage of carbon to be paired with other means of CO_2 utilization (G.C. Institute, 2016), including but not limited to CO_2 EOR activities, chemical synthesis, methanol fuel and biofuel generation. Each component of the CCUS value-chain, as depicted in Figure 1.1, is crucial to the economic and technical viability of the CCUS procedure.

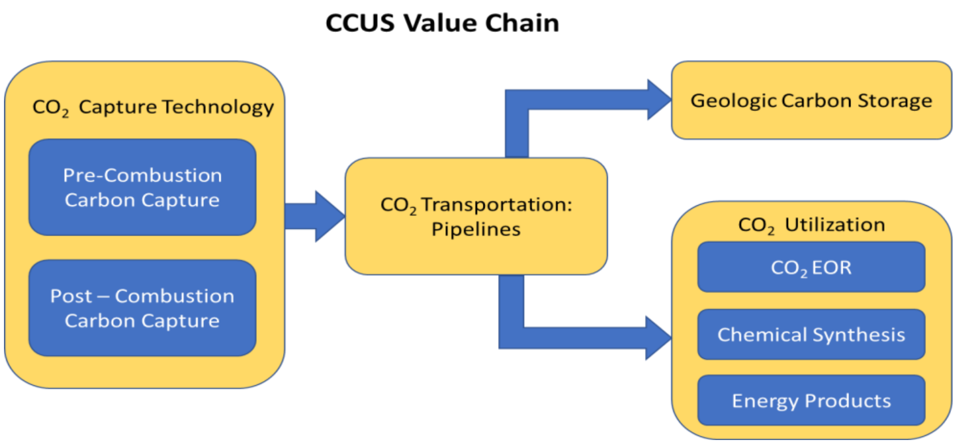


Figure 1.1: CCUS value chain.

1.2.1 Pre-Combustion Carbon Capture

Pre-combustion carbon capture technology refers to capturing of CO_2 from the fossil fuel or biomass streams prior to combustion (Global CCS Institute, 2012). Figure 1.2 provides the general schema related to pre-combustion carbon capture procedure. Pre-combustion carbon capture techniques are usually associated with higher CO_2 concentrations, elevated pressures, and higher temperature ranges (Wall, 2007). CO_2 pre-combustion separation techniques include physical absorption, where the gas is contacted with counter-current solvent stream; adsorption, where the gas is contacted with solid adsorption beds; cryogenic separation, where a series of cooling and compression cycles separate the gas stream; and membrane technology (Theo et al, 2016).

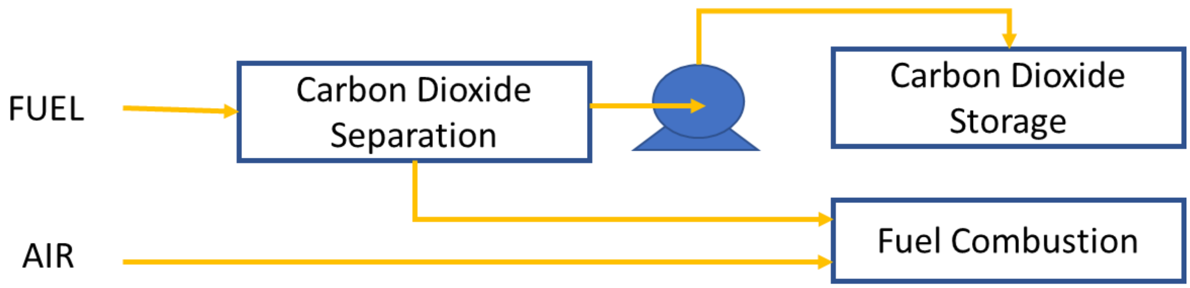


Figure 1.2: Pre-combustion carbon capture schema.

1.2.2 Post-Combustion Carbon Capture

Post-combustion carbon capture refers to the technology related to capturing CO_2 after the combustion process. The advantage of post-combustion is that this technology could be retrofitted to most energy and industrial facilities (Zhao et al., 2016). However, there exists several challenges to this post-combustion technique such as the low flue gas outlet pressures, low CO_2 output concentration streams and low size difference between the captured gas molecules (D'Alessandro, 2010). Liquid absorbent-based capture techniques are the leading methods in the post-combustion-based carbon capture technology, amongst which, amine-based absorptions are the most dominant technique with a capture efficiency of 90% . In this type of absorption, flue gas is pretreated for removal of Sulphur Oxides (SO_x) and Nitrogen Oxides (NO_x) components and brought in direct contact with the absorbent stream to form a rich CO_2 stream (Kohl et al., 1997). The general schema for post-combustion carbon capture is shown in Figure 1.3.

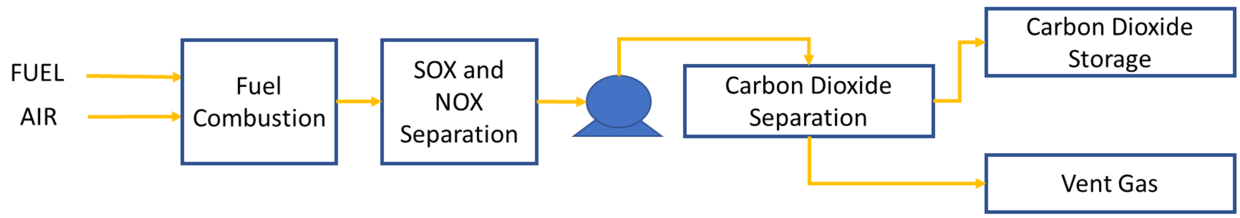


Figure 1.3: Post-combustion carbon capture.

1.2.3 Geologic Carbon Storage

Geologic storage of CO_2 involves the injection of the rich CO_2 stream into the subsurface formations (Raza et al., 2016). The CO_2 injection procedure utilizes several of the long-standing practices employed by the O&G industry for Enhanced Oil Recovery (EOR). Important control parameters related to gas injection include formation storage capacity, formation mineral composition, injectivity, trapping mechanisms, containment and formation stability (Hosseini et al., 2013). Testing of potential candidate formations for storage of CO_2 include numeric simulation, lab-scale testing of multi-flow fluid, fluid-fluid interactions and rock-fluid interactions. These studies are important to understand the long-term storage potential of the candidate reservoirs and to understand the risk of potential CO_2 leakage to aquifers or other mineral rights regions. Table 1.1 shows some of the advantages and disadvantages of geologic storage sites used for CO_2 storage.

Table 1.1: Advantages and Disadvantages of Geologic CO_2 Storage Sites (Raza et al., 2016).

Geological Setting	Advantages	Disadvantages
Coal Seams	Capacity	High-Cost
	Enhanced Hydrocarbon Recovery	Geographically limited
Salt Domes	Safety	High-Cost
	Ideal Design	Geographically limited
Saline Aquifers	Capacity	Safety
	Geographically widespread	
Depleted Hydrocarbon Fields	Proven Safety	Geographically limited
	Enhanced Hydrocarbon Recovery	Timely Availability
	Infrastructure in-place	Problems with multi-phase flow

CO_2 geologic storage is classified as a multi-decade long-term project. In order to maintain safety and assurance of long-term storage without leakage, it is essential to conduct several monitorings during and post-injection of the CO_2 stream . There are many measurement techniques to monitor the movement of CO_2 in the geological structure used for storage including the use of monitoring wells, which measure rate of injection, pressure and temperature variations, and CO_2 plume composition. Some of the factors affecting CO_2 migration into geological formation post injection include, pressure and hydraulic gradients, buoyancy, diffusion and dispersion, dissolution, mineralization and phase trapping (Chadwik et al. 2014).

1.2.4 CO_2 Utilization

Carbon utilization includes the usage of captured CO_2 in various processes instead of being permanently stored in an underground formation. CO_2 can be utilized in the chemical, oil, power, food and pharmaceutical sectors as well as the paper and steel industries. The use of CO_2 can be categorized into resource recovery (e.g. enhanced oil and gas recovery and enhanced coal-bed methane recovery), captive use (process integrated) of CO_2 as an intermediate product in the manufacturing chain without external sources, and the non-captive or merchant use (Styring et al., 2011; Fortes et al., 2014; Aresta et al., 2007). The highest usage of these resources comes from the oil and chemical industry, followed by the cement industry and the food industry.

1.2.4.1 CO_2 Enhanced Oil Recovery

Amongst the utilization techniques of CO_2, EOR is the most used. This is due to the direct application of CO_2 without much processing. EOR also has a higher associated value (cost/tCO_2) and larger overall quantity of CO_2 involved in the process.

Enhanced Oil Recovery is an activity classified as a tertiary oil recovery procedure in O&G production, used for improved sweep efficiency and production from residual oil zones. Such practices of oil recovery are usually ensued after the reservoirs' potential for natural production of oil is depleted and other efforts of pressure maintenance (such as water flooding) is no longer a viable option (Bachu, 2016). The main categories of EOR in O&G include Thermal EOR, Chemical EOR and Gas Injection-based EOR. Amongst the gas Injection-based EOR techniques, CO_2 EOR is one of the most commonly used strategies amongst the gas-injection based EOR techniques.

The scale of CO_2 EOR in the last two decades have spanned the globe with the USA at the forefront. Depending on the miscibility of CO_2 with other reservoir fluids and reservoir properties can be used for miscible or immiscible operation (Kuuskraa et al., 2013). The operations involved with either method is different. In miscible flooding operation the CO_2 dissolves in the reservoir fluid, reducing viscosity and decreasing interfacial tension within the reservoir. The major problem with such activity involves phase separation and viscous fingering; wherein bypassing of fluid front occurs between wells. On the other hand, immiscible flooding is a topic of much interest as many shale reservoirs. The operations can be conducted in multiple manners such as Water Alternating Gas operations, where CO_2 is cycled with water for better sweep efficiency or huff-n-puff operations, where the same well is used to both inject CO_2 and produce oil (Sheng, 2017).

The ideal formations for CO_2 enhanced EOR are generally at a range of 1600-11950 ft below the surface at a temperature range of 82-260℉ with a permeability range of 1-4500 mD (Koottungal, 2014). This type of recovery procedure generally works best with oil of 27-45 API and 0.4-6 cP viscosity.

1.2.5 CO_2 Pipelines

CO_2 pipelines have gotten a great impetus due to various driving factors such as EOR activity, carbon reduction strategies, enhanced coal bed methane recovery and industrial production. In 2007, there were 2414 km of CO_2 pipeline, with majority of the length of pipeline located in North America, majorly focusing on transportation of CO_2 for EOR projects in O&G fields (Towler et al, 2007). This number dramatically increased in 2010, with the USA alone having about 3862 km of CO_2 pipeline laid out transporting 30 million tons of CO_2 annually (International Energy Agency Greenhouse Gas Program [IEAGHG], 2010). This number further increased to 6437 km worldwide in 2014 (IEAGHG, 2015) and 8046 km in 2018 (Peletiri et al., 2018). Table 1.2 compiles a list of trunkline CO_2 pipeline worldwide.

Table 1.2: Pre-existing and planned CO_2 pipeline worldwide (IEAGHG, 2015; Peletiri et al., 2018; USEPA, 2020b).

Project Name	Country	Status	Length (km)	Capacity (MtCO_2 /yr)
Quest	Canada	Planned	240	1.8
Alberta Trunkline	Canada	Planned	240	14.6
Weyburn	Canada	Operational	330	2
SaskPower Boundary Dam	Canada	Planned	66	1.2
Shute Creek	USA	Operational	142	4.5
Monell	USA	Operational	53	1.6
Bairoil	USA	Operational	258	23
West Texas	USA	Operational	204	1.9
Transpetco	USA	Operational	193	7.3
Salt Creek	USA	Operational	201	4.3
Sheep Mountain	USA	Operational	656	11
Val Verde	USA	Operational	130	2.5
Slaughter	USA	Operational	56	2.6
Cortez	USA	Operational	808	24
Central Basin	USA	Operational	232	27
Canyon Reef Carriers	USA	Operational	225	1.1
NEJD	USA	Operational	294	7
Dectaur	USA	Operational	1.9	1.1
Eastern Shelf	USA	Operational	91	1.1
GreenCore	USA	Operational	232	2.65
GreenLine	USA	Operational	314	9.30
Delta	USA	Operational	108	2.2
Snohvit	Norway	Operational	153	0.7
OCAP	Netherlands	Operational	97	0.4
Lacq	France	Operational	27	0.06
Rhourde Nouss	Algeria	Planned	30	0.5
Qinshui	China	Planned	116	0.5

Gorgon	Australia	Planned	8.4	0.5
Bravo	USA	Operational	350	7.3
Bati Raman	Turkey	Operational	90	1.1
Este	USA	Operational	191	4.8

CO_2 pipelines are quite similar to natural gas pipeline networks in their design and technical specifications. This similarity between pipeline networks has let pipeline developers to draw feasible conclusions on CO_2 pipeline standards such as their steel grade, pipeline diameter, frequency of booster stations as well most importantly risk and economics (Knoope et al., 2013). This has also led to the development of regulatory standards of CO_2 pipelines as a product carrier. In the USA, CO_2 pipelines are treated similar to a non-volatile hazardous liquid carrier (USA CFR Section 49, 2019).

1.3 Motivation and Objectives

The need for better CO_2 management is a worldwide issue, where every country is trying to reduce emissions in a sustainable manner of maintaining economic growth and achieving energy independence while ensuring a safe environment for future generations. Countries are approaching this problem through retrofitting traditional power plants and industries (cement, natural gas, fertilizer and so on) with lower emission technology (British Petroleum [BP], 2019). The North-Central region of the USA comprising of 5 states of North Dakota, Montana, Wyoming, Utah, and Colorado, has a large amount of conventional energy sources and industries. The energy transition challenge in this region is due to sparse and distributed population centers and existence of many traditional natural resource-based industrial sectors. Despite these factors, it must be noted that this region has shown tremendous potential for renewable energy and adapting well to the new lower CO_2 emission environment.

CCUS proves to be a feasible technology that enables stakeholders to transition to lower CO_2 emission rates at a convenient pace and more energy-efficient operations. CCUS technology lets stakeholders take advantage of government tax credits for reducing CO_2 emissions while also converting the CO_2 to value added products or as a substance in enhancing production in other sectors such as O&G. The major problem related to commercial implementation of CCUS projects is the large upfront capital and labor charges along with significant annual operating costs.

CCUS comes with many inherent decision-making points related to selection of CO_2 sources, sinks and pipelines. Problems related to CO_2 source selection in CCUS projects includes the sources to be selected for a CO_2 operation, the technology these sources need to be retrofitted with as well as the amount of CO_2 to be captured at these sources. Problems related to sinks for CCUS projects include the type of utilization process to be used, the sink sites to be selected and the quantity of CO_2 that can be utilized/stored in the selected site. Problems related to transportation includes the best routes for transporting CO_2 and the technical requirements of the said pipeline.

With this in mind, the main objective of this body of work is to determine the cheapest means of deployment of a CCUS network in North-Central USA in order to capture and store/utilize CO_2 to keep up with emission reduction goals over a given period of time. The following objectives are defined in order to achieve this goal:

- Studying the impact of terrain, ecology, and environment on pipeline corridors.
- Identifying CO_2 pipeline corridors in the study area.
- Determining the CO_2 sources to be added to the CCUS infrastructure network to capture a set minimum level of CO_2.

- Calculating the amount of CO_2 that should be captured in the capture locations
- Determining the CO_2 sink that should be added to the CCUS infrastructure network to store/utilize produced CO_2.
- Calculating the amount of CO_2 that should be stored/utilized at the selected sink sites.
- Defining the best and safest pipeline trunk routes in the region related to CCUS infrastructure.
- Establishing a president for identifying regulations into pipeline network modeling
- Defining a method to incorporate regulations into decision-making

CCSHawk methodology is developed to materialize these objectives in a manner that is easy-to-implement and easy to understand.

1.4 Methodology

CCSHawk as described in the previous section provides a workflow to be used for preliminary decision-making to obtain a specified level of carbon capture in the North-Central region of USA. CCSHawk is composed of several steps of which the most important steps are briefly mentioned:

- ***Mapping Study Area:*** The study area is characterized and mapped using 19 thematic map layers conveying a variety of information which characterize each tract of land according to their physical, environmental, and infrastructure-based features.
- ***Generation of Cost Map:*** The mapped region is analyzed according to their acceptance towards building and sustaining a safe pipeline section. The region is analyzed using a multi-criteria decision analysis (MCDA) technique known as Analytic Hierarchy Process (AHP) which indicates the parts of the study area best suited for CO_2 pipelines.
- ***Generation of Candidate Pipeline Network:*** This part consists of several sub-steps eventually leading to the generation of multiple candidate pipeline routes between the various sources and

sinks of CO_2. Major sub-steps include the creation of suitable pipeline route pairs using Delaunay algorithm, tracing of the least cost path using a graph network technique known as A-star (A*) algorithm and refining the pipeline network.

- ***Techno-Economic Modeling:*** Each source, sink and candidate pipeline routes is then fitted with a suitable standardized techno-economic model commonly used in the industry to obtain technical features of the components as well as their associated costs.
- ***Optimization Model:*** The optimization model is the tool used for decision making in CCSHawk. The model used here utilizes a mixed-integer linear programming (MILP) formulation to obtain the most economic and safe CCUS infrastructure network to meet specified CO_2 capture goals.
- ***Text Mining of Regulations:*** The regulations related to CO_2 pipeline are extracted and incorporated into the network model. This step includes extraction of regulations from a XML format database and utilization of text mining and Natural Language Processing for extracting the regulation.
- ***Visualization:*** Visualization and tabulation of the most economically optimal combination of sources, sinks and CO_2 pipeline to meet emission reduction goals.

1.5 Significance

The contribution of this body of work is as follows:

- First comprehensive study of CCUS infrastructure in the North-Central region of CCUS. This study will benefit the planning and execution of future CCUS projects in the region.
- Development of a feasible pipeline route panning schema for the region. This schema can be further extended for other application including natural gas and crude oil pipelines.

- This study will further the knowledge on the impact of existent infrastructure on the planning of future CCUS related infrastructure in terms of economics and overall feasibility.
- This study would be a first of kind to incorporate insights from regulatory texts into the CO_2 pipeline model.
- Though the methodology provided in this study is specific to the development of a CCUS infrastructure network, however, it can easily be adopted for analysis of other utility infrastructure such as power lines, telecommunications and so on.
- This study can be utilized to analyze the maximum recommendable CO_2 capture quantities in the study area in order to ensure sustainable development of CCUS infrastructure in the future.

1.6 Dissertation Structure

This work consists of 6 chapters.

Chapter 1 provides a brief introduction to the concept of Carbon Capture, Utilization, and Storage. The chapter also introduces CCSHawk methodology along with the objectives and significance of this thesis.

Chapter 2 reviews the worldwide studies related to CCUS network analysis. A brief summary of the work in pipeline routing algorithms as well as multi-criteria decision analysis is also provided. Also, the work related to CCUS decision-analysis is explored along with it's knowledge gaps.

Chapter 3 presents the various steps related to the CCSHawk methodology in detail including generation of potential pipeline routes, development of decision-making framework, text-mining framework and cost map generation. The chapter also details the data sources and the various pre-processing required for using the data in the methodology.

Chapter 4 discusses the implementation of study area analysis and route generation process. Further, the chapter details the results of both static and dynamic versions of CCSHawk in the study area along with the visualization of the results.

Chapter 5 presents the arguments related to implementation of CCSHawk. The impact of various factors on CCUS infrastructure are explored. This chapter further explores the impact of variation in control parameters on the CCUS network along with the effect of having existent infrastructure included in the CCUS network.

Chapter 6 provides a summary of major finding of the work along with discussion of avenues to better answer questions still pending after this study.

1.7 Summary

This chapter introduces CCUS technology and the components of the CCUS value-chain. It is concluded that the processes related to the CCUS value-chain are capital intensive and great consideration while establishing infrastructure. To help with the decision-making process and reduce economic burden of CCUS, the CCSHawk methodology is introduced which focuses on selecting suitable capture and utilization sites for CO_2 and optimizing CO_2 pipeline routes. This chapter also establishes the objectives, significance and structure of the thesis.

CHAPTER 2

Literature Review

2.1 Introduction

CCUS is a relatively new technology and is being tested for implementation in several parts of the world. It has proven to be commercially feasible in some regions and is being adopted by several countries including the USA, Canada, the Netherlands, China, Brazil and the United Arab Emirates (IEA, 2019b).

Although, CCUS is a feasible solution for CO_2 emission reduction, it is quite capital intensive. To reduce cost, countries are employing strategies like cost-sharing amongst multiple stakeholders, deployment of common carriers of commodities and technology transfer (IRGC, 2008). Choosing the right combination of sources and sinks of CO_2 to be retrofitted with CCUS infrastructure can also make CCUS more affordable and viable. Choosing the right locations for CCUS infrastructure deployment is not limited to choosing the correct sources and sinks, but also developing appropriate pipeline corridors that would pass through areas which are safe for pipeline construction.

In this chapter, the study area, North-Central USA, is explored in greater detail. The chapter also explores literature and previous work related to CCUS network analysis. The intention of this chapter is to give a background to the work done in this thesis. The aspects explored in this chapter related to previous works in CCUS infrastructure networks, include various geo-information used for mapping pipeline routes, development of "Cost Maps", pipeline route generation, and CCUS decision support systems.

2.2. Study Area: North Central USA (North Dakota, Montana, Wyoming, Colorado, and Utah)

The area of study is the north-central part of the United States of America, which includes the states of North Dakota, Montana, Wyoming, Utah, and Colorado as shown in Figure 2.1 (Balaji, 2020). The total area of study comes to about 1.32 million km^2. This region is not heavily populated as compared to many other regions of the USA with a 10.02 million people (according to the 2010 census). However, there are pockets of heavily populated regions along metropolitan areas such as Denver and Salt Lake City. The region has a diverse topography varying from the steep regions of the Rocky Mountains to large plains in parts of Montana and most of North Dakota. The study area also has various terrestrial and ecological factors that could affect pipeline construction. There are several river systems and lakes as well as national and state parks scattered through the region. It must also be noted that various reservations including the Forth Berthold, Standing Rock, Wind River, Fort Peck, Uintah and Ouray are within the study area, which need special consideration in this study.

Despite the lower population density of the area of study, there are several industries and power plants located in this region, making it a net exporter of power and energy. The total CO_2 emission of the region amounts to 239 $MtCO_2$ (million tons of CO_2) annually (US EPA, 2020b). The current largest stationary source of CO_2 emissions is the Jim Bridger coal power plant (14.6 million tons) in Wyoming, followed by the Colstrip power plant in Montana (13.57 million tons).

The region also has several hydrocarbon basins, which has led to the development of many EOR projects in Wyoming, Eastern Montana and Western North Dakota.

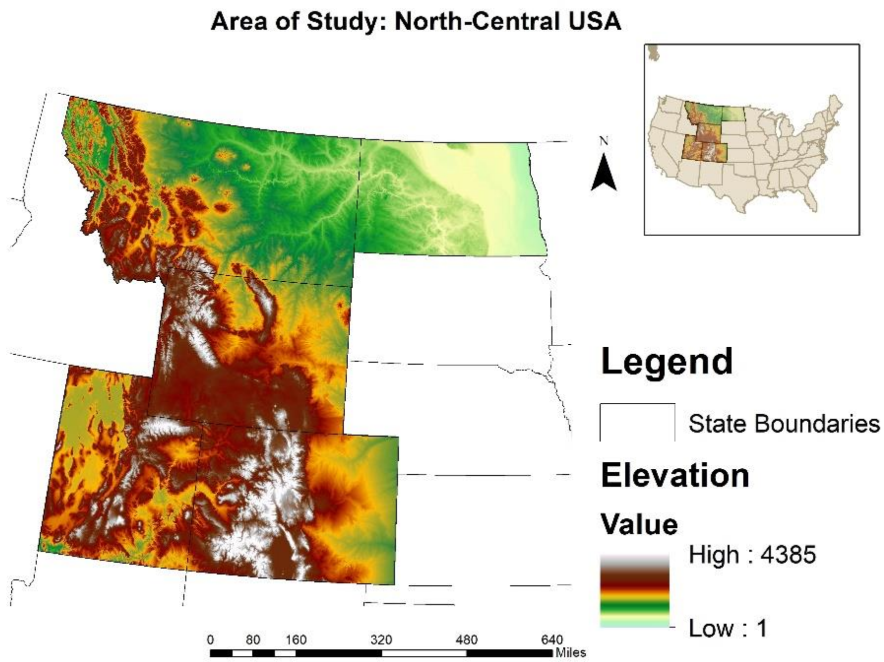


Figure 2.1: Area of study, North-Central USA - North Dakota, Montana, Wyoming, Utah and Colorado (Balaji et al., 2020).

With stricter regulations and increasing environmental responsibility, several industries are looking for alternatives to reduce CO_2 emissions in the study area and CCUS is seen as a viable technology for this purpose. There are several projects under consideration for permanent geologic storage of CO_2 including Red Trail Energy project and the Milton R. Young plant, North Dakota (Energy and Environmental Research Centre [EERC], 2019) along with the existent CO_2 EOR projects. Pilot geologic CO_2 storage projects are also being tested in Williston Basin (North Dakota) and Kevin Dome Formation (Montana). The region has a long history of transportation of CO_2, dating back to late 1980s with CO_2 EOR projects in the Bairoil field. The development of EOR projects and storage opportunities in the region, has led to the development of four major pipeline networks including the Green Core pipeline (373 km), Exxon pipeline (229 km), FDL pipeline (257 km) and the Dakota Gasification pipeline (531 km) (US DOE, 2015).

2.3. Mapping the Study Area: Factors Affecting Pipeline Corridors

Mapping the study area helps determine pipeline parameters related to (Menon, 2011; Huseynli, 2015; Potter et al., 2013):

- Technical feasibility
- Economic feasibility
- Regulatory compliance

These are parameters which can affect the project objectives and goals, and can influence future operation and maintenance of the pipeline facilities. The important factors that are usually considered during the route planning procedure for pipelines may include, but are not limited to:

- *Population Density*: The distribution of human population is one of the most important criteria to be considered while planning pipeline routes (Menon, 2011). The high importance given to populated regions, is because of the increased risk to human life and property related to pipeline failures. Regulations also put restrictions on certain types of CO_2 pipelines near densely populated regions. For instance, in the USA, federal regulation prohibits the presence of a CO_2 trunk pipeline near population centers (USA. CFR., Section 49, 2019). Population density is commonly used in various studies related to CCUS network planning (Middleton et al, 2012a; Towler et al., 2007; Berry, 2004), however, population density can also be replaced with factors such as the location of urban regions and towns.

- *Right-of-Way/Existent Pipeline Routes*: Right-of-way refers to the rights related to pipeline passing through a specific piece of land and a means for physical access to the pipeline (Menon, 2011; ESRI, 2012). There is a need of purchasing these right when passing through lands without existent rights-of-way, greatly affecting the cost of projects (Callan, 2008). The process for obtaining rights-of-way near existent pipeline is easier and more cost-effective (Menon, 2011). Due to this factor, construction of CO_2 pipelines near other pipelines such as natural gas and crude oil is preferred.

- *Elevation and Slope*: Elevation is an important factor related to the technical feasibility of pipeline construction which refers to the relative altitude to the region. Slope refers to the change in altitude with respect to its surrounding regions which can be calculated from elevation (ESRI, 2012). Steep slopes are not preferred for pipeline construction due to increased load on booster stations. Usually either slope or altitude is considered during analysis of a study area (Towler, 2007; Menon, 2011; Berry, 2004; Potter, 2013; Middleton et al, 2012a).

- *Soil, Geology and Faults*: Soil and/or geology are factors that can affect construction of pipelines, especially due to their effect on burrowing and corrosion (Menon, 2011; Berry, 2004; Potter, 2013). Subsurface faults are factors that may lead to future issues of maintenance and probable risk of leaks, thus pipelines should avoid areas (if possible) with major faults (Menon, 2011; Potter, 2013).

- *Land Use/Land Cover*: Land use refers to usage of land for various types human activities, while land cover refers to natural foliage cover of land surfaces (ESRI, 2012). However, the terms are used interchangeably in the literature. Land use/ land cover is an important factor in the establishment of a pipeline routes for several technical and economic reasons. For example, lands designated as wetlands have several construction problems due to potential future leak risks which can disturb bird and reptilian habitats (Menon, 2011; Berry, 2004).

- *Lakes, Rivers and Distance to Water*: Lakes and rivers pose higher construction cost for pipelines and potential future risk in the case of pipeline leakage. Thus, most pipelines go around lakes and minimize river/stream crossing (Menon, 2011; Berry, 2004). Distance to water is used as an alternative to rivers and lakes (Potter, 2013).

- *Physical Boundaries (cities, towns, roads, railways, state parks, archaeological sites, protected lands and other places of interest)*: Cities and towns and other settlement regions are avoided for reasons of regulation and safety in case of future pipeline leaks (Menon, 2011; Potter, 2013). These regions also have larger densities of population. Similarly, archeological sites, protected lands and other places of interest are avoided due to regulations and safety (Menon, 2011). Crossing infrastructure networks such as roads and railways lead to measures such as burrowing, which is not preferred. However, laying a pipeline parallel to roads and railway paths, is encouraged as it gives easier access to the pipeline.

- *Wildlife and Protected Species*: Pipelines can affect the habitats of protected species and wildlife (Potter, 2013). Pipeline regulations have been developed to discourage development harmful to wildlife. These regions which are considered important for wildlife and endangered species are classified as unusually sensitive areas in the USA (USA. CFR., Section 49).

One of the most important work related to impact of terrain and environment on CCUS infrastructure decision-analysis was done by Herzog et al. (2009), forming a reference point for geo-information considerations in other work. The authors mapped the study area using 1000-by-1000m resolution geo-information layers related to slope, populated regions, wetlands, national parks, state parks, railroads and roads. Herzog et al. (2009) also considered a few regulatory aspects such as EPA underground storage class as additional layers for future usage if needed. In addition, they considered classification of wilderness areas (part of federal lands) and federal lands as additional layers to be chosen for possible usage.

Middleton et al. (2012a) added considerations for pipeline rights-of-way to the list of geo-information layers used by Herzog et al. (2009). In the paper, geo-information layers such as federal lands, natural gas and crude oil pipelines and land ownership characterized the pipeline rights-of-way. The work is done as part of the data development description of a software called SimCCS. SimCCS is an open-source CCUS infrastructure network decision-making tool which is used extensively by several CCUS partnership's (NETL) across the USA. In this software, the layers to be considered as input are provided by the user, however, the authors have described several layers that should be taken into consideration while mapping study area including slope, railways, roadways, rivers, federal lands, land use, rights-of-way (ROW) and population density (Middleton et al., 2012a). In a similar study, Fritze et al. (2009) routed pipelines from source to sink for *5* sources and 2 sinks in the Gulf of Mexico region of the USA using the mapping considerations provided by Middleton et al. (2012a). The work done by Herzog et al. (2009) also inspired the geo-information considerations for a study of CO_2 pipeline routing in China known as ChinaCCS (Chen et al., 2010). ChinaCCS included factors such as digital elevation maps (DEM), slope, rivers, cities, highways and railways. Sun et al. (2013) improved ChinaCCS by considering environmental factors in CCUS decision analysis by adding soil information and wetland locations.

The MARKAL-NL-UU model used current and future geographic considerations in generating a resource allocation analysis for CCUS infrastructure in the Netherlands (Broek et al., 2010a). In this work, Broek analyzed the study area using land usage and population distribution as the only criteria. The maps used for the study had land usage and population estimates for the years of 2025, 2040 and 2050. These maps were not generated as part of the study, but were sourced from the Dutch government. Kanudia et al. (2013) established the TIMES COMET model, an extension of the MARKAL-NL-UU model (Broek et al., 2010a) with the inclusion of databases from the countries of Spain, Portugal, and Morocco. The TIMES-COMET model added geo-information layers such as slope, railways, roadways and existent pipelines to the considerations made by the MARKAL-NL-UU model.

The work done till date in CCUS decision-analysis and CO_2 pipeline planning have some common features such as slope, population areas, roadways, railways and to a lesser extent rights-of-way and wetlands. Although, these factors were considered in the analysis of the region of interest, they are not sufficient to map the effect of environmental and ecological factors on pipeline routes. An important aspect lacking in these studies, involves the utilization of local factors such as snow cover, frost action, and corrosion factor. These factors have been used in pipeline planning for water and crude oil and serve as a hinderance to pipeline development (Cevik, 2003; Potter et al., 2013). It is also important to use ecological factors in planning pipelines as they can affect the trajectory of pipeline development and could also lead to future geopolitical issues. The work done till date in CCUS network formation, has not considered these ecological factors. Herzog et al. (2009) touches upon some ecological factors, however, they have not been used in generation of the “cost map”, essential to mapping pipeline routes. In this thesis, both ecological and local factors are used to map the study area in addition to other factors such as roads, railways, slope, and waterbodies. The analysis of these geo-information layers is discussed in Section 3.3.

2.4 Multi-Criteria Decision Analysis in Cost Map Generation

Multi-decision criteria analysis (MCDA) for quantifying importance of each layer of information in the region of interest can be classified mainly into two categories of subjective and objective weighting schema (Ozcan et al., 2011). Subjective weighting schema relies on the prognosis and knowledge of the user, or expert knowledge and bare no relation to the quantitative aspects of the criterion being considered (Ozcan et al., 2011; Zardari et al., 2015). Objective weighting schema takes into account the differential analysis in the quantitative values of the criterion being compared (Zardari et al., 2015; Al-Aomar, 2010). In the case of pipeline route analysis, the heterogeneity/variation in data may not be significant and the variation in data can be misleading, as it can lead to higher importance to factors which minimally affect pipelines. Thus, subjective weighting schema are preferred over objective weighting schemes in pipeline network generation.

Some of the most popular subjective weighting schema are as follows:

- *Analytic Hierarchy Process (AHP):* The AHP method (Saaty, 1980) is the most commonly used and a popular methodology in pipeline routing applications. This method considers a set of evaluation criteria, and a set of alternative options among which the best decision is to be made. It is important to note that, since some of the criteria could be contrasting, the best option is generally not the one which optimizes each single criterion, rather the one which achieves the most suitable trade-off among different criteria. The AHP generates a weight for each evaluation criterion according to the decision maker's pairwise comparisons of the criteria. The higher the weight, the more important the corresponding criterion (Kolios et al., 2016; Liu et al., 2016; Al-Aomar, 2016). The global score for a given option is a weighted sum of the scores is obtained with respect to all the criteria. AHP can be executed in three steps: a) a hierarchical structure is created for more justifiable comparison between hierarchies, b) the vector of criterion weight is computed , in which each criterion (in the same node of the hierarchy tree) is weighted against the other and given a score between 0 and 9, depending on their relative importance, and c) a weighted summation function is applied to the criterion according the score obtained and normalized to form a uniform cost raster for the specific pipeline routing task.

- *Fuzzy Analytic Hierarchy Process (Fuzzy AHP):* The fuzzy AHP utilizes a statistical standard variation to delineate categories in a continuous fashion rather than the categorical system used in traditional AHP technique. A fuzzification methodology is utilized to convert the various opinions of the experts into a fuzzy range, to reduce uncertainty (ESRI, 2012; Torfi et al., 2016). This method is used for the purpose of routing; however, it is not used commonly for pipeline routing applications.

- *Technique for Order Preference by Similarity to the Ideal Solution (TOPSIS):* TOPSIS is a subjective technique similar to AHP in execution. TOPSIS penalizes those options that are farther from the ideal option creating a scoring system on suitability (Kolios et al., 2016).

In CCUS infrastructure decision-making and CO_2 pipeline routing literature MCDA techniques are used for weighing geo-information layers. However, explanation related to the MCDA techniques is not available. There are a few exceptions such as Herzog et al. (2009) where it is briefly mentioned that the various geo-information layers were processed using the AHP methodology and weightages were provided, however, no further explanation on the AHP implementation has been provided. Similar studies, such as Middleton et al. (2009; 2012a), Sun et al. (2013) and Kanudia et al. (2015) mention the usage of a MCDA method but have not provided the means for generating the cost maps.

The weightages used for creation of "cost map" is essential to map the effect of terrain, ecology and environment on pipelines. Hence, fair explanation needs to be provided on the creation of the "cost map" as it determines the orientation of pipeline routes. In this work, Section

3.4 will discuss the implementation of the AHP methodology to obtain the relative importance of each geo-information layers.

2.5 Routing of CO_2 Pipelines

Path finding techniques, known as "Least Cost Path" (LCP) algorithms, are used for computationally routing pipelines. In computer systems, the way of getting from point A to B is usually through the usage of these LCP algorithms by analyzing the relative location of the starting and ending points and any obstacle in the path. The most common and easily executed LCP method/algorithm is the straight-line path (Callan, 2008; ESRI 2012). However, this routing technique is most suited for measuring distances and is not suitable for pipeline route selection, due to the fact that straight-line paths do not reflect the actual pipeline paths. Despite this shortcoming several studies have utilized straight-line LCP for pipeline routing (Kobos et al., 2006; Broek et al., 2010a; Knoope et al., 2014).

Kobos et al. (2006) used straight line LCP methods to map pipeline routes between sources and sinks. This method was used as part of a source-sink matching software known as "String of Pearls". The authors used the straight-line LCP to join linked up nodes matched by the tool. The distance between two points was measured using the Euclidean distance and the "Proximity tool" on ArcGIS.

Unlike the String of Pearls theory, Broek et al. (2010a) used straight line LCP in their MARKAL-NL-UU model to calculate distances between sources and sinks of CO_2. The authors of MARKAL-NL-UU acknowledged that no pipeline can have a straight-line distance and used a factor of 1.4 to increase the length of pipeline for further cost analysis. They also used distance factors with land use and population density to increase the cost of pipeline in case of crossing through water or urban areas. Similar to the MARKAL-NL-UU model, Knoope et al. (2014) used the straight-line LCP for calculation of distances in their work which focused on technical optimization of pipeline in terms of diameter, thickness and steel grade rather than pipe routing. This type of straight-line heuristic for measuring distance between two points as a proxy for pipeline length was used by Lee et al. (2019), Kazmierczak et al., (2009), Guo (2020) and Ravi et al. (2017).

Dijkstra Algorithm (Dijkstra, 1959) is a prominent method used to track routes between any two points and is used for several applications such as traffic routing, migratory pattern analysis as well as utility mapping (ESRI, 2011; ESRI, 2012). Dijkstra algorithm is a graph network search algorithm that provides the shortest path from a node to every other nodes in a graph (Cormen, 2001). Consequently, it can also be used for finding the shortest path from a single source vertex (node) to a single destination vertex (Souissi et al., 2017).

Dijkstra algorithm visualizes any two-dimensional map/picture, as a grid. The algorithm starts by establishing the start point as the starting vertex, and then considers the surrounding 4 or 8 cells (depending on the sequencing of the algorithm as seen in Figure 2.2) and calculates the distances to move from the starting vertex to these adjoining cells. These adjoining cells are moved to a list known as “Open List”. The distance and cost calculation for the movement depends on the distance heuristics such as Euclidian distance or Manhattan distance. Once the movement to the next cell is traced, it moves the initial cell into a “Closed List” and then places all the adjacent cells of the newly selected current position into the “Open List”. This is a recursive algorithm which visits all cells in the cost map or reiterates the selection of new cells till the destination cell is reached (depending on the algorithm). Once the sink/destination cell is identified, a path of least cost is traced back to the starting node. Dijkstra algorithm is executed using GIS software suites like ArcGIS or QGIS, or through programming platforms such as Python, R, C++ and Java (ESRI, 2011; Middleton, 2020; Broek, 2010a; Morbee, 2012; Sun, 2013).

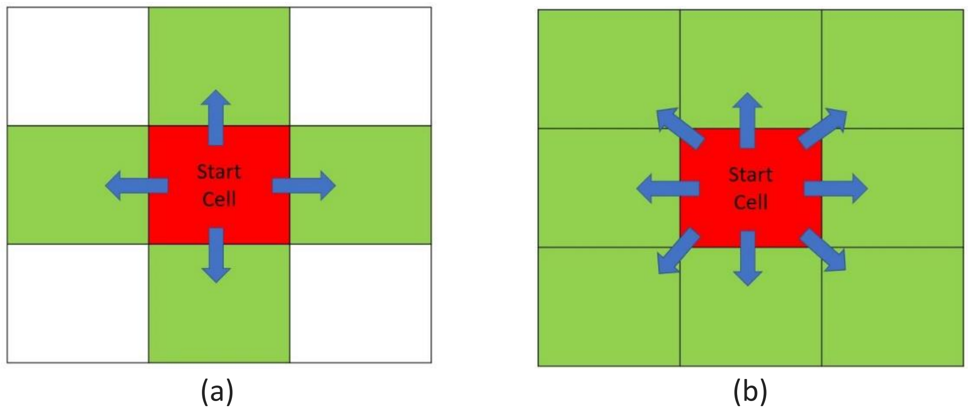
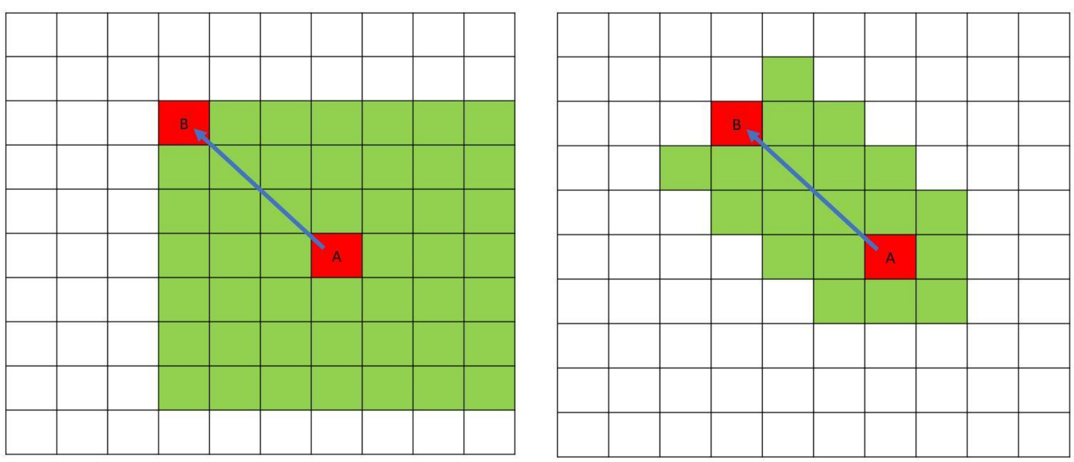


Figure 2.2: Movement pattern used by Dijkstra algorithm and A* algorithm for neighboring cell selection (a) 4-cell movement (b) 8-cell movement.

A-star algorithm (A*) is a technique derived from graph theory which is commonly utilized in pre-planned path-finding problems in the gaming and unmanned aerial vehicle industries (Hart et al., 1968) and is an extension of the Dijkstra algorithm. Dijkstra algorithm expands in every direction equally and can take a long execution time. A* algorithm overcomes these shortcomings by creating a more goal-oriented search approach (Reddy, 2013). This goal-oriented approach is due to biased search of the algorithm in the direction of the destination using a distance heuristic. A* thus enables quicker completion of the task for single source - single sink routing problems. However, in the case of a single source - multiple sink routing problem, the same algorithm needs to be reused multiple times. Figure 2.3 illustrates the difference in growth pattern in A* algorithm as compared to Dijkstra algorithm, while Figure 2.4 illustrates the difference in heuristics. The details for the implementation of A* is further elaborated in Section 3.5. To check the usefulness of one algorithm over the other, an analysis needs to be done to check the number of connections each node needs to make with other nodes as well as the distance of each node point from other nodes. If the node points are distant from each other and the area of interest is large, Dijkstra algorithm will lead to longer computational times. However, if a node must be connected with many other nodes in a relatively small region of interest, Dijkstra algorithm will have a relatively low computational load.

Dijkstra and A* are both deterministic algorithms, however, there exist various other stochastic algorithms which may be used for path generation. Generally, authors in road network generation and ocean path finding problems use genetic algorithms and particle swarm optimization to optimize the path between start and end node dynamically. These techniques are also popularly used in unmanned aerial vehicles (UAVs) for optimizing flight paths using movement pattern based on elevation data. Genetic algorithms and particle swarm optimization are examples of probabilistic pathing algorithms, which may not necessarily provide the same solution for the problem on multiple executions (Souissi et al., 2013). These probabilistic algorithms are computationally expensive and provide minimal gains in path generation problems with multiple destinations. The reason for not using these algorithms in this work is due to the increased computational load as well as the redundancy of the procedure due to the cost layer being static and not dynamic in terms of cost of movement. In static environments the map/network does not change through multiple executions, while dynamic environments vary with time. For static environments like the cost map used in pipeline networks, deterministic algorithms perform as well as probabilistic algorithms (Souissi et al., 2013).

One of the early examples of usage of graph network techniques for solving CO_2 pipeline routing problem is the work of Herzog et al. (2009) where Dijkstra algorithm was used through ArcGIS in order to route sources to sinks. This algorithm was used only to connect sources to sinks and not nodes of a similar kind (source-to-source or sink-to -sink). The algorithm used in their model was run in a way that the cost of reaching from a source node to every other point in the raster was calculated for each node used in the CCUS network. A similar approach was adopted in ChinaCCS (Chen et al. 2010; Sun et al. 2013) where the authors used Dijkstra algorithm through ArcGIS LCP toolbox (ESRI, 2012) to route sources to sinks. The straight-line LCP used by Broek et al. (2010a) was replaced by the usage of the Dijkstra algorithm in the improved TIMES-COMET model (Kanudia et al., 2013). Similar approaches of using Dijkstra algorithm have also been reported by Fritze et al. (2009) and Weihs et al. (2012) for routing CO_2 pipelines. In order to shorten distances and enable quicker computation of routes, these studies used a minimum cell size of 1000-by-1000 meter. However, using these large map resolutions in path finding algorithm tend to negatively affect narrow features in the study area such as roads or railways leading to information loss. Also, for large study regions such as in the case of ChinaCCS (the country of China), the repetitive Dijkstra algorithm implementation is computationally intensive (Soussisi et al, 2017).



(a) (b)

Figure 2.3: The growth pattern (a) Dijkstra algorithm grows in every direction (b) A* algorithm which grows in the direction of the destination node.

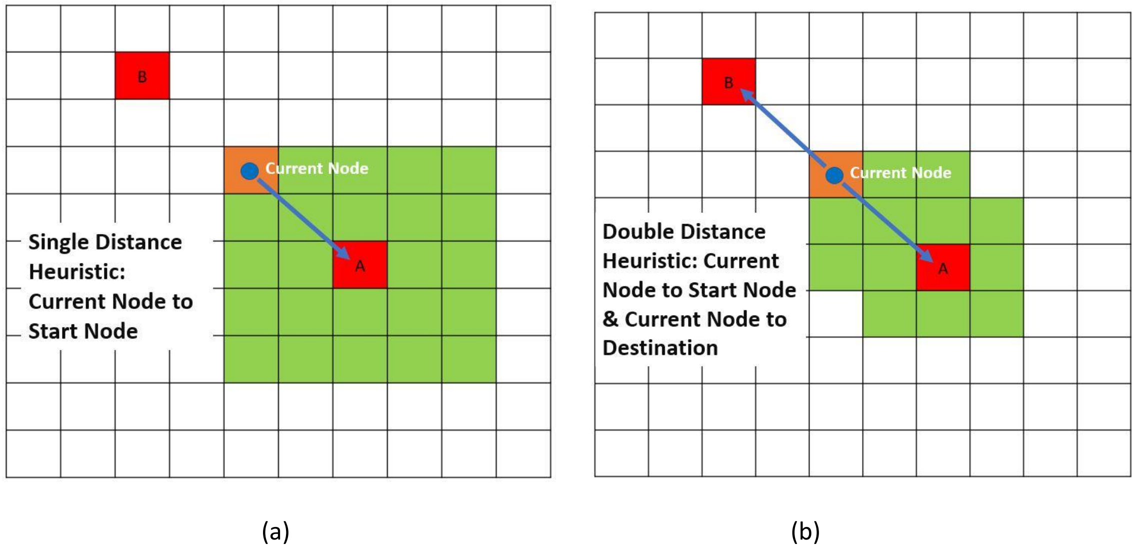


Figure 2.4: Path generation heuristics (a) Dijkstra algorithm considered cost of movement only from the start point (b) A* algorithm considers distance from start node to current node and current node to destination node.

Middleton et al. (2012a) improved the implantation of Dijkstra algorithm by changing its means of application and refinement. Middleton et al. (2012a) defined a five-step CO_2 pipeline network generation methodology. First, the raster inputs are overlaid on each other using an overlay function with established weights for each layer to generate the cost map. The second step includes the use of the Delaunay algorithm to choose the best combination between the nodes for pipeline routes to be generated. The Delaunay triangulation algorithm is a method to connect points in a plain such that every three points form an empty circumcircle. This step is used to generate the basic network for path analysis. The third step is to use the Dijkstra algorithm to find the LCP between the nodes. The fourth and fifth steps include changes in data format and removal of redundancies. This five-step procedure has the inherent problem related to large cell sizes similar to the work of Herzog et al. (2009). Another drawback of the process is the large number of nodes generated in the process of combining paths which are close to each other. The many new nodes generated in the technique leads to expansion of the possible CCUS network and large computational time for network analysis.

On the other end, Morbee et al. (2012) improved the Dijkstra algorithm's functioning rather than the workflow. The authors developed the InfraCCS model, which covers the premise of source to sink pipeline routing and resource management. The model utilizes Dijkstra algorithm in order to generate a LCP between various node points. However, InfraCCS tweaks this algorithm by changing the mode of movement from the general 8-cell pattern in the standardized Dijkstra algorithm to a 16-cell movement pattern, with the purpose of reducing the overall cost of movement (Figure 2.5). Despite the improvements made by the 16-cell movement pattern, the algorithm neglects the possibility of moving into a high cost cell, when it skips over the immediate eight cells in the vicinity of the current cell.

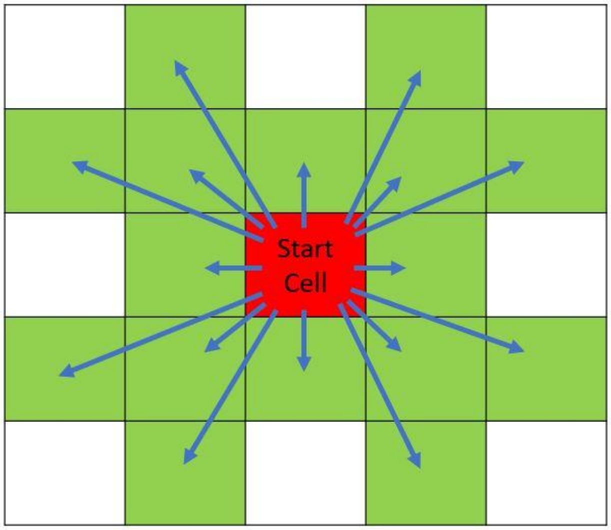


Figure 2.5: The 16-cell movement pattern for Dijkstra algorithm used by Morbee et al. (2012).

The analysis of paths between nodes in CCUS networks in the literature have used two major techniques: straight-line LCP analysis and Dijkstra algorithm. Amongst these techniques, straight-line LCP is better for analytical purposes and authors using this technique have noted that the straight line representation of pipelines is not realistic. Other studies have used Dijkstra algorithm for path finding with large map resolutions. The consequence of using this technique is larger computation times and loss of information due to large map resolutions. In this study, A* algorithm is used for path finding between the nodes, as the nodes are evenly distributed through the region of interest and also to avoid loss of information.

2.6. CCUS Infrastructure Analysis and Decision Support

Various authors have contributed to the decision analysis of CCUS infrastructure using tools such as numerical optimization, probabilistic analysis, and pinch analysis. These tools are used to make decision related to matching of various nodes (sources and sinks) using different modes of transport (mostly pipelines) and planning of a cost-effective CCUS infrastructure network.

A technique commonly used for network analysis is the one-to-one matching of sources to sinks, by pairing sources and sinks of equal CO_2 capture/storage capacity or according to their CO_2 inventory requirement. This is a technique based on resource allocation and is not usually the most economically feasible solution (Herzog et al, 2009; Middleton et al., 2009). The most utilized technique in CCUS network analysis is mathematical programming through numerical optimization. The analysis techniques used for mathematical programming of the network include linear programming, mixed integer linear programming (MILP) and multi-integer non-linear programming (MINLP). These models are optimized to either minimize risk, cost or technical liabilities. The analysis of the models is controlled through usage of constraints, which determine the level of detailing related to a network. The more the number of constraints related to the model, the more complicated the model gets which then requires more computational efficiency. In a few cases, the model is used to optimize more than one objective problem at a time. The studies involving economic and risk analysis, are characterized by techno-economic analysis of the sources, sinks and pipelines. The base techno-economic analysis used in the studies influence the results of each optimization study. However, these calculations are not provided in most studies, thus a fair comparison between each optimization study is not feasible. Another common technique utilized in CCUS network analysis to prepare source-sink combinations, is the pinch analysis. This method involves material balance using a graphical solution for CO_2 at various vertices and indicates the best pairing between the vertices. This analysis is done by carefully mapping the requirements of each node and then matching the deficit or excess CO_2 with another node which can compensate for the difference in CO_2 requirements (Ooi, 2013). The core difference between pinch analysis and one-to-one matching is the ability for a single node to be paired with more than one other nodes at the same time and the ability to constrain a few core parameters related to the sources and sinks. Probabilistic analysis is another technique used in optimization of CCUS network and includes methods such as genetic algorithm, decision tree analysis (economic) and particle swarm optimization. These techniques are based on statistical analysis of the parameters but can handle only a few parameters depending on the formulation of the network before the functioning of algorithm gets computationally intensive (Tian et al., 2017).

In CCUS network analysis, few studies generate custom CO_2 pipeline routes using LCP techniques. These studies tend to be more focused on CO_2 transportation infrastructure. Kobos et al. (2006) introduced one of the first decision-making frameworks for CCUS infrastructure known as “String of Pearls” which matched sources and sinks one-to-one according to their capacities. The sources, sinks and existent pipelines are ranked in the order of the amount of CO_2 emission, CO_2 storage capacities, and CO_2 transport quantity, respectively. The sources, sinks and existent pipelines are matched by their capacities, and as sinks fill up, the sources are connected to the next sink/nearest existent pipeline closest to the source. The system is tuned to optimize the cost of individual nodes rather than the whole CCUS network. This type of optimization and capacity matching serves as a good example for decision making related to matching sources and sinks according to the preferences of an individual stakeholder rather than the whole region. The model also does not designate the destination for an amount of CO_2 being transported in case the fluid is routed to a nearby existent pipeline and considers only the cost of capture and storage in the analysis leaving out the cost related to transportation of CO_2.

Another example of one-to-one matching of source and sink without accounting for costs or risks associated with the CCUS network was proposed by Kazmierczak et al. (2009). The author came up with a recursive algorithmic methodology for creating a CCUS network. The algorithm takes into consideration all sources and sinks in a region and matches them one-to-one based on their respective capture and storage capacities. The algorithm starts by matching the largest source and sink to one another. Then it iteratively chooses its target sources and sinks by their decreasing capacity and checks each connection one by one to spot if a node is nearer to another node or to the existent pipeline connection from previous iteration. Thus, using material balance and distance heuristic, pipelines are created to make sure most of the captured CO_2 has a destination sink. The algorithm also recursively reiterates pipeline diameter sizes to accommodate for extra flow added in each iteration. This model is an example of the application of “minimum spanning tree” graph search technique, in which pipeline routes are represented by straight line connections. However, it must be noted that this process is computationally intensive, and the combination of routes generated may not be the most optimized model between the sources and sinks.

Amongst the one-to-one matching tools in literature, the work done by Broek et al. (2010a; 2010b) is considered the most important (Middleton et al., 2012b; Sun et al., 2013; Morbee et al., 2012). Broek et al. (2008) introduced the MARKAL (Market Analysis) model to estimate the effect of CCUS prices on the power sector in the Netherlands. This paper takes into consideration cost of capture technology, operation of storage facilities and pipelines along with probable variations in cost of power in a comprehensive capture model to estimate the cost of a CCUS network. Reservoirs, both onshore and offshore, are considered in the study with major focus only on the storage capacity of each reservoir. This model uses a linear one-to-one connection between sources and sinks of appropriate capacities to find the optimal economic scenario of CCUS development over multiple review years (dynamic setting). The focus of the paper is on CO_2 capture costs (Euros/kWh) and its potential influence over the power bills in Europe. The major drawbacks of this method are several including the absence of the real cost of pipeline construction and no limits on the injection rate into reservoirs. The model also links only a single source to a single sink at a time.

The MARKAL model was improved upon in Broek et al. (2010a, 2010b). The new model was named MARKAL-NL-UU and included the power sector market factors as mentioned in Broek (2008) along with better pipeline modeling and source-sink matching. The improvement in decision-analysis is based on better control mechanism in injection operations. The major focus in this new improved model was on the Utsira formation and the usage of the formation to reduce cost by avoiding offshore storage. However, the major drawback of this work is the one to one connection between sources and sinks. By enabling one-to-many connections, the feasibility of the scenario could be more realistic. Another problem with the model is the lack of integration of current infrastructure into the model development. The MARKAL-NL-UU model was adopted for Portugal, Spain, and Morocco by Kanudia et al. (2013) and renamed as the TIMES-COMET model. The network analysis in the TIMES-COMET model remained the same as MARKAL-NL-UU and it used one-to-one matching of sources and sinks according to their respective capacities.

One of the most prominent and widely referred work in CCUS network analysis and CO_2 pipeline routing is SimCCS (Middleton et al., 2020). SimCCS is a top-down feasibility model for various development scenarios in CCUS, comprising of sources such as coal-fired power plants and varying sinks such as saline aquifers, oil fields and EOR fields. SimCCS has 3 different modules. The first module related to SimCCS is described in Middleton et al. (2009). This paper employs most of the techniques used in SimCCS and talks about the MILP formulation related to the decision-making process for capturing a minimum amount of CO_2 in a region within a time period. The objective function in this paper is the reduction of the economic cost of the entire CCUS operation including capture, storage, and transportation. This is controlled by a number of constraints including limiting the capacity on each component, the number of pipelines between two points as well as the mass balance constraints in order to maintain proper flow. It must be noted here that the most significant control variable in the formulation is the amount of CO_2 to be captured in a given period of time (static model). Middleton et al. (2020) used California as the study area to demonstrate the applicability of SimCCS, and they called it as $\text{SimCCS}^{\text{CAP}}$.

The second module of SimCCS is described by Kuby et al. (2009). The paper discusses the MILP formulation for SimCCS which aims at identifying the appropriate development strategy for a given capture amount of CO_2 at a fixed tax rate (for each ton of CO_2 emission) in a region in a static setting. The formulation is similar to Middleton et al. (2009) except for fixing the tax rate related to reduction of CO_2 emissions. Kuby et al. (2009) uses a case study in California and shows slight changes in the infrastructure plan as compared to Middleton et al. (2009). This model is later referred to as $\text{SimCCS}^{\text{PRICE}}$ in Middleton et al. (2020). The third module of SimCCS is provided in Middleton et al. (2012b) which describes a MILP formulation for a given CO_2 capture target in a study area, however, in this case the setting is dynamic. The objective functions as well as the constraints are a function of an additional time variable. The time-based formulation provides better economic growth of the CCUS operation over multiple periods of time with the goal of prioritizing overall economic optimization rather than optimization in every individual time period. Middleton et al. (2012b) applied this time-based formulation in Texas panhandle region, and they later refer to this module in Middleton et al. (2020) as $SimCCS^{TIME}$.

The pipeline arcs used in these models are uni-directional which can increase computational load when solving the optimization problem. Further, these problems do not consider the cost of purchase of CO_2 for EOR activities. The techno-economic model shown in Kuby et al. (2009), does not incorporate cost of monitoring, verification and abandonment related to CCS activity. Further, the model does not have the means to capture the interaction of existent CO_2 transportation infrastructure with new potential CCUS infrastructure, which could make some cases unrealistic. However, SimCCS makes these techno-economic parameters as inputs, where the user has the ability to vary the cost related to storage and capture of CO_2.

SimCCS has additionally been used to check the effect of the CCUS regional network on overall economy and wellsite operations. Middleton et al. (2012c) demonstrated the importance of modeling CCUS operations from pore scale to regional scale by estimating the performance of reservoirs right from the atomic to regional scale. Further, the authors demonstrated how the output of each lower scale-level study affects the operation of immediate higher-level operations. The effect of reservoir properties such as porosity, permeability, and saturation on the overall regional scale development of CCUS infrastructure is explored in Middleton et al. (2012d). In this paper CO_2PENS, an integrated asset management tool is used to vary reservoir properties and show their effect on the overall capacity of reservoirs, which is in turn fed to the SimCCS software. As expected, it is demonstrated that a decrease in the capacity of reservoirs and well injectivity rates increases the overall cost of development of CCUS infrastructure increases. A similar study is conducted in Pawar et al. (2016) where SimCCS is partnered with the National Risk Assessment

Partnership (NRAP) tools (improved version of CO_2PENS) to demonstrate the effect of variation of reservoir properties on the overall CCS infrastructure development. As can be seen from these examples the SimCCS software has been used in multiple setting, however, the base strategy of SimCCS remains the same and only the inputs change in each scenario.

A study based on the work by Middleton et al. (2009) was done by Morbee et al. (2012), which improves the static setting laid out by SimCCS to create a dynamic framework for CCUS networks. In this work, InfraCCS covers the premise of source to sink pipeline routing and resource management. The model utilizes MILP to make decisions related to resource utilization according to time of implementation. The objective function utilized in Morbee et al., (2012) minimizes total discounted pipeline investment cost. Constraints were put on the number of nodes used, pipeline capacity, pipeline construction cost, sink injectivity and sink capacity. The key inputs utilized in the InfraCCS model include the capital expenditure (CAPEX) and operational expenditure (OPEX) for transportation of CO_2, without including the costs related to capture and storage. One issue with this model is the calculation of custom pipeline diameters depending on flow rates after optimization despite the usage of predetermined pipeline diameters as a control variable. In addition, the model does not incorporate the cost related to capturing and storing of pipelines. This is an important flaw in the model, considering the cost of capturing CO_2 emission is several times the cost of transportation, thus pipeline development is influenced more by type of capturing mechanism rather than pipeline costs (Middleton et al., 2009; Herzog et al., 2009; Sun et al., 2017, D'Amore et al., 2020).

An example of a model that started out as a one-to-one matching algorithm and turned into a MILP formulation-based analysis is ChinaCCS developed by Chen et al. (2010). Sun et al. (2013,2017) provided an improved ChinaCCS decision support system with linear optimization model rather than a one-to-one matching pattern. The initial one-to-one matching algorithm is similar to the work done by Broek et al. (2010a) where the source with the largest capacity is matched with a sink with the largest capacity. The MILP formulation is similar to the model developed by Middleton et al. (2009) for fixed capture quantities of CO_2 in a static setting. The objective function in the newer ChinaCCS model is to reduce the pipeline net present value under the constraints of capacity of sink, capture limits of sources, pipeline capacity and mass balance. ChinaCCS uses inputs such as OPEX and CAPEX costs for capture, transport, and storage exclusively for CO_2 EOR projects. ChinaCCS does not incorporate existent infrastructure into the CCUS infrastructure and the model generated by ChinaCCS is static in nature. The arcs used in this model are uni-directional and the model is computationally intensive given the nature of connections.

One of the first examples of usage of numerical optimization tools for the development of CCUS network was introduced by Herzog et al. (2009). The model uses a simple MINLP with optimization for the CCUS network decision analysis. The model developed by Herzog influenced the development of SimCCS, however, there is a distinction in the definition of pipeline diameter between the model developed by Herzog et al. (2009) and SimCCS. In the former case, a custom diameter is calculated according to the flow requirements of each line, while in SimCCS pre-defined pipeline diameters are used in the decision analysis. The calculation of custom pipeline diameter for each line in the model leads to the non-linear nature of the formulation. The model forms a particularly good basis for constraints and inputs for other optimization models especially MILP formulations. However, the interaction of derivative pipelines with existent pipelines is not explored. Another feature lacking in the model is a means of clustering various closely located vertices to reduce candidate networks which is adopted by other works including Middleton et al.

(2012a) and Morbee et al. (2012). The MINLP model developed in their work is computationally heavy and can lead to inconsistent results varying with each run. Jensen, et al. (2013) utilized the MIT model to calculate the future CO_2 pipeline needs in the mid-western and Central Canada over several years, utilizing projected values of CO_2 emission levels. In the model Jensen et al. (2013) explored to some degree the effect of existent Dakota gasification pipeline between the states of North Dakota and Saskatchewan on future pipeline development. However, the interaction of the existent pipeline with potential future networks has been explored as extensions to the existent Dakota Gasification pipeline or duplicate lines to the existent pipeline.

A change from the trend of utilizing mathematical programming for network optimization was presented by Weihs et al. (2012). The authors utilized genetic algorithm to minimize the cost of the CO_2 pipeline optimization problem for Queensland, Australia. The model focuses on an established set of sources and sinks, with focus on criteria such as capture quantity, pipeline transport parameters and injection quantities in order to obtain the configuration of parameters under uncertainty with the lowest cost per ton of CO_2 emitted. The costs related to transportation and storage are predetermined using an established techno-economic model, however, the parameters are varied to estimate the factors which have the most effect on CCUS costs. The study was limited to the analysis of costs related to pipelines and injection only, without considering the cost of capturing CO_2. In addition, no limits were put on the injection of CO_2 in each well. The analysis done by Weihs et al. (2012) in mapping parametric sensitivity of projects cost was extended by Wang et al. (2016), where the authors explored the effects of pipeline distance, well injectivity (constant rate for each well), CO_2 pipeline capacity under different CO_2 flow rates, pipeline lengths and storage site properties for future CO_2 storage development in Australia. The methodology utilized decision trees to analyze the sensitivity of these parameters for minimizing the objective function of the overall pipeline cost. It must be noted that the base optimization calculation done by Wang et al. (2016) still relied on the genetic algorithm developed by Weihs et al. (2012).

Network building is not the extent of usage of mathematical programming in CCUS as shown by Knoope et al. (2014). The authors developed a model for estimation of CO_2 pipeline cost and technical features for a single pipeline from source to sink. The paper indicates a systematic flowchart of an algorithm to determine the approximate engineering properties to design a CO_2 pipeline for a stream in either dense or supercritical phase. The study was extended to consider multiple pipelines in Knoope et al. (2015) by sacrificing a few technical parameters. Knoope et al. (2015) developed the source-sink matching model for CO_2 pipelines, which involved the quantification of risk and uncertainty in terms of economic parameters. Though, the focus of the model is more on sensitivity of market factors related to CO_2 pipelines, the appendix of the paper details a decision-making optimization tool. The optimization tool is a dynamic MINLP to match source nodes to the sink nodes utilizing inputs such as CAPEX and OPEX of CO_2 storage, capture and transportation. The model performs dynamic optimization over the entire period of the operation rather than step-by-step incremental optimization. The non-linear nature of the formulation is due to calculation of custom diameters for each flowline. The optimization model also incorporates variation of pipeline material which is not seen in other models. The model uses Euclidean distances as a measure for pipelines and also does not incorporate existent pipeline networks in estimation of CCUS network development scenarios. Another shortcoming of the model like Herzog et al. (2009) is the increased computational time related to MINLP formulation which can easily be avoided by using a set of pre-determined pipeline diameters.

In a deviation from statistical and numerical analysis of CCUS networks, Diamante et al. (2013) utilized pinch analysis for making source-sink combination over a single time period, for better and quick mass flow rate-based analysis. The model presented by Diamante et al. (2013) uses a static condition for a few sources and sinks. The major limitation of such a technique is the necessity of individual capacity analysis of each node and lack of constrains in the capture and storage systems. Another shortcoming of pinch analysis is the limit on the number of nodes considered for the study. The work on pinch analysis was expanded to a dynamic setting by Ooi et al. (2013) to make the analysis valid for source-sink matching over multiple periods. However, this improved model lacks factors such as distance between source and sink and any technical specifications related to the CCUS network. Further, Diamante et al. (2014) used a unified graphical pinch analysis technique which overcame the limitation of previous models by addressing injection rate constraints, thus limiting the individual annual capacity of each reservoir. The pinch analysis in Diamante et al. (2014) provides a quick and easy resource management tool suited for source-sink matching.

A different type of network analysis as compared to models developed by Herzog et al. (2009), Middleton et al. (2009) and Broek et al. (2010a) was introduced by D'Amore (2017). The authors used equal dimension sectorial division of every part of Europe and summarized the levels of storage and capture for each of the rectangular sectors. The size of each of rectangular sector is 1000-by-1000 kms. The model for decision-making framework developed by D'Amore is a MILP formulation to reduce overall cost of capture, transport, and storage in the network. All the storage costs and capture costs in each individual sector is averaged out and used as a cluster in the MILP formulation. The MILP formulation generated by D'Amore was based on minimization of economic cost related to the CCUS network. The inputs related to OPEX and CAPEX of the network is predetermined and proportional to the quantity of CO_2 captured, stored, and transported. D'Amore (2018) expanded on the study by introducing risk associated to CCUS into the MILP formulation instead of cost of the network. The authors quantified risk using the probability of spillage while transportation and leakage while injecting into subsurface. This formulation naturally reduces the distance of transportation networks and prefers EOR networks over long-term saline storage. This is because the risk used in the study is proportional to the distance of transport and quantity of CO_2 transported and stored. Social acceptance was introduced into the framework by using the level of acceptance of CCUS by the populace in each sector (D'Amore, 2020). The analysis done by D'Amore is a regional scale analysis and does not give perspective related to finer details of study area. The study also does not focus on aspects related to CO_2 capture.

Probabilistic analysis for finding the best CCUS network combination was explored further by Tian et al. (2017), serving as one of the few instances of its usage in addition to Weihs et al. (2012). The authors developed a linear matrix inequality model for different time steps to optimize CCUS infrastructure networks using various parameters such as flow rate, operation time, maintenance and labor charges and land prices. However, despite the optimization of thermodynamic and infrastructural properties, the model focuses majorly on cost optimization purely for pipelines and not the CCUS value-chain. The model also uses linear distances for pipeline measurements without accounting for deviations in the pipeline route. The model developed by Tian et al. (2017) is further varied to check the effectiveness of evolutionary algorithms as optimization tools, namely: co-evolutionary particle swarm optimization (Tian et al., 2018) and genetic algorithms (Tian et al., 2020).

In addition to CCUS network analysis, supply chain analysis related to the CCUS value-chain are also reviewed by various studies. A study involving supply chain analysis of CCUS networks, was developed by Ravi et al. (2017) for the Netherlands focusing on different CO_2 capture types. The model uses a MILP formulation with the explicit aim of choosing appropriate capture sources for a single mega-sink. The model analyzes the development of the network, if all captured CO_2 from the sources are deposited in a single sink and further analyzes the best combination of sources for further development in case the sink reaches capacity in the form of a back-up sink. It must be noted that this model does not develop any pipeline routes or use any techno-economic analysis related to transportation. The scale of supply chain analysis was significantly increased by Leonzio et al. (2019; 2020) by developing a MILP network in order to match different types of CO_2 capture technologies with various utilization pathways including EOR, greenhouses, fertilizer industry and food industry. The model used by Leonzio et al. (2019; 2020) does not focus on development of feasible pipeline corridors, nonetheless the study helps in analyzing the potential of a variety of utilization avenues. The model uses both costs related to implementation of the CCUS network and profits earned from the utilization pathways in the analysis of the networks in the United Kingdom and Germany. A similar supply chain analysis was conducted by Zhang et al. (2018; 2020). The authors utilized MILP to minimize the annualized net cost of the project where the model is dependent on the entire supply chain from flue gas dehydration to carbon utilization in terms of subsurface storage. The model adds multiple new aspects in the decision-making framework by including 16 different types of capture technology and several utilization destinations.

In terms of optimization objectives, CCUS network analysis can be extended to other parameters like the work done Lee et al. (2017; 2019), which described the formulation of the pipeline optimization problem based on a MILP methodology using four objective functions: maximizing profit, minimizing environmental impact, and minimizing financial or environmental risk due to uncertainty (according to preference). The problem is designed in two phases: phase 1 is formulated as a multi-objective problem that maximizes the total annual profit and minimizes environmental impact; Phase 2 is a two-phase stochastic multi-objective problem designed to reduce the financial risk based on preferred risk levels (Lee et al., 2019). Despite the variations available in such formulation, the formulation can only optimize by one factor at a time. The analysis does not include cost/parameters related to utilization and storage.

From the analysis of the various work done on CCUS networks, it is clear that CCUS infrastructure planning is important and help with the assessment of the potential of this technology. The most commonly optimized parameter in CCUS network is the cost related to the carbon capture and utilization, followed by risk related to the environment. However, very few studies do a complete evaluation of the value of CCUS, leaving out factors related to capture, utilization, or transportation. The review of the available literature also indicates a trade-off between the breadth of model (elements of the value-chain) and depth of model (details related to each component of value chain analyzed), where more complete evaluations of the CCUS value-chain have fewer parameters analyzed. A variety of methodologies have been used to optimize the CCUS network, and each has its advantages and disadvantages. The most commonly used optimization technique is mathematical analysis like MILP or MNILP, which provides a wider scope of the CCUS network at the cost of the depth of each component in the value-chain. On the other hand, analyses such as probabilistic analysis perform well for a few components of the CCUS value-chain but are computationally expensive. It is also seen that even though many studies have tried to optimize networks, they use one-to-one matching of sources and sinks leading to inefficient pairing and network generation. Pinch analysis serves as a good solution for supply-demand relations between sources and sinks, but at the cost of technical detailing. Many studies have included pipeline analysis in their evaluation of CCUS networks but have limited their scope to minimal detailing of their pipeline routes. In this work, a MILP network will be used to evaluate CO_2 capture, storage, and transportation in the study area. The study will include power plants and commercial CO_2 suppliers in its evaluation along with saline aquifers and EOR activities. The analysis of the network will be made under both static and dynamic settings for a specified CO_2 capture limit.

2.7. Knowledge Gaps

There are several potential areas to be further explored in CCUS network infrastructure decision-analysis. These areas include:

- Most of the models reviewed in the literature ignore important environmental factors including frost cover, soil conditions, seismic activity, and ecological factors such as protected areas and endangered species. It would not be possible to have a comprehensive feasible network without these environmental considerations in the analysis.
- Many studies evaluating the cost and technical aspects of the CCUS value chain ignore CO_2 pipelines and even if included in the study, measure the pipeline distances using Euclidean distance. Euclidean distance cannot reflect real pipeline lengths as pipelines are not built in a straight line.
- Most models have not explored any options for pipeline routing beyond straight-line LCP and Dijkstra algorithm.

- The models have not taken into consideration regulatory factors which could affect the CCUS network. Some models have taken into consideration a few aspects such as material of pipeline (Knoope et al., 2015) or parks (Herzog et al., 2009).
- The role of existent pipeline in the development of future CCUS networks is seldom explored except for a few studies (Kobos et al, 2006; Jensen et al., 2013).
- Many models have used simple one-to-one source-sink matching formulas or used all the sources and sinks in the study area without making the hard choice of which sources and sinks are best suited to be in the CCUS network. Such analysis does not give economically feasible options.
- The weightages used for preparing "cost maps" are not provided in the literature apart from Herzog et al. (2009).

2.8. Summary

This chapter explored the literature related to CCUS infrastructure network and decision-analysis system. A review of the available literature revealed that very few studies used realistic pipeline routes in their analysis. It was also clear that consideration of local and ecologic factors, which may affect pipelines were not included in these studies. Further, it was also identified that none of the papers explored methods for route creation beside the Dijkstra algorithm or straight-line LCP. This chapter also explored the tools used in the decision analysis of CCUS infrastructure. The tools and methods utilized aim at many different objectives including economic optimization and risk reduction in the CCUS infrastructure. Nonetheless, every one of them emphasis the importance of mapping out CCUS infrastructure for planning and future implementation.

CHAPTER 3

CCSHawk – A Methodology

3.1 Introduction

Carbon capture, utilization, and Storage (CCUS) can play a pivotal role in reducing CO_2 emissions and several countries are adopting it as an active means to combat climate change in addition to development of energy storage and efficiency technologies. The undertaking related to CCUS is massive and involves several stakeholders. There is a need for active decision-making procedures which enables the stakeholders to make fundamental choices related to the deployment of CCUS infrastructure. In view of this need, scientists have developed tools to help in this analysis related to CCUS technology. However, as discussed in Chapter 2, these studies still lack perspective in respect to the effect of environment, ecology, and existent infrastructure on the development of CCUS networks.

CCSHawk is a methodology which aims to be an aid in the decision-making process related to CCUS infrastructure planning and CO_2 pipeline route optimization for the North-Central region of USA. CCSHawk answers the following questions: a) which regions are the best candidates for pipeline development?, (b) which types of CO_2 sources should be included in CCUS infrastructure?, (c) which CO_2 sinks are the most effective for storage?, (d) which pipeline route between two points is the best and safest for transporting CO_2?, and (e) how much CO_2 can be captured/stored safely in a given period of time economically?

This chapter looks into the techniques and methods related to the development and optimization of a CCUS infrastructure network which forms the basis for CCSHawk. Section 3.3 describes the geo-information layers used in the development of the Cost Map and also discusses the sources and sinks of CO_2 considered in the study. Section 3.4 provides the implementation and discussions on the AHP technique and preparation of a cost map. The pipeline route generation methods are discussed in Section 3.5. The techno-economic model employed for the sources, sinks and pipelines are provided in Section 3.6. The last section describes the formulation of a mixed-integer problem for decision-making. Section 3.7 describes the formulation of a mixed-integer problem for decision-making. Finally, an additional feature of the methodology in terms of CO_2 pipeline regulatory analysis using text mining procedure used to extract obligations/norms is described in Section 3.8.

3.2 Major Constituents of the CCSHawk Methodology

CCSHawk methodology consists of 6 major sections: 1) Cost Map where the study area is mapped and regions which are suitable for pipeline corridors are identified; 2) Candidate Route Generation where the pipeline routes are generated in an optimized fashion from various nodes points; 3) Cost Model where the related techno-economic properties of the source, sinks and candidate pipeline are identified; 4) Mixed Integer Linear Programming where all the inputs are combined to form a decision-making framework which is optimized to reduce the cost given certain control parameters and; 5) text mining of regulatory conditions related to pipelines. The major inputs related to CCSHawk include geo-information layers of the study area (Section 3.3), relevant source and sinks of CO_2 along with their respective CO_2 specifications (Section 3.3 and 3.6), CO_2 capture goals (in terms of $MtCO_2$/yr) and a timeframe (Section 3.7). The general workflow related to the methodology involved in CCSHawk is depicted in Figure 3.1.

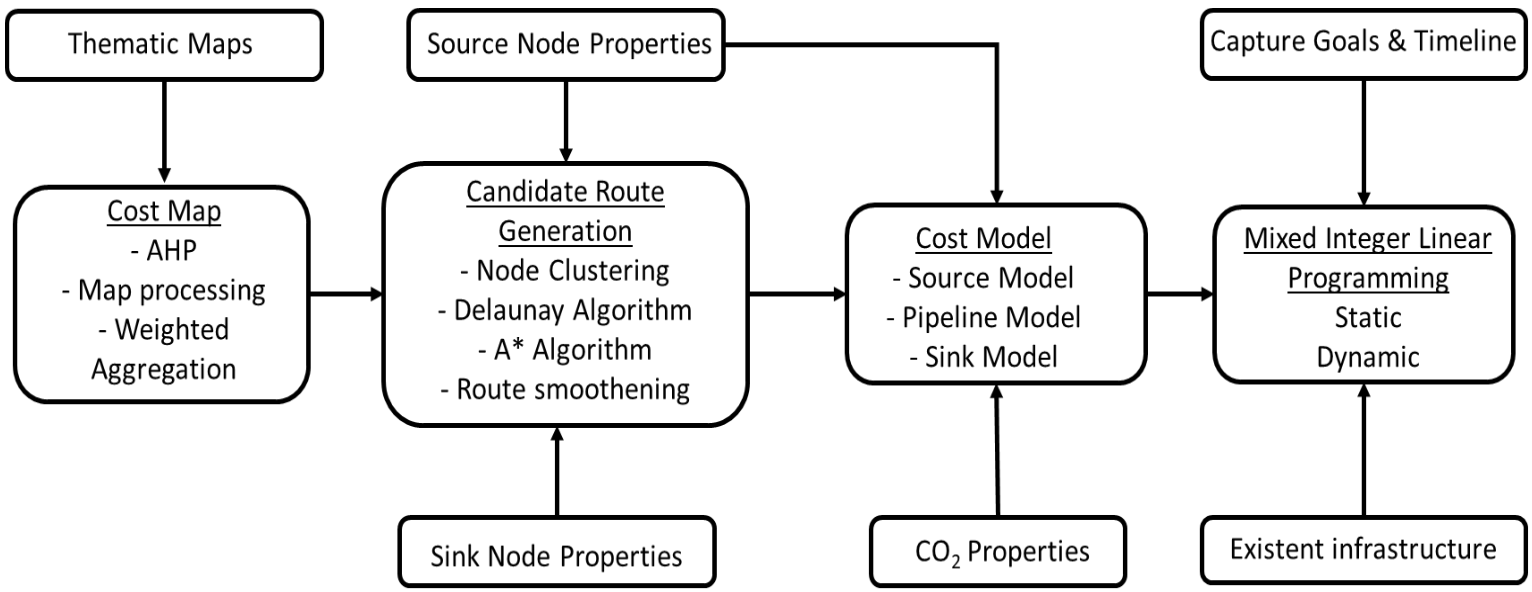


Figure 3.1: Workflow for the CCSHawk methodology.

There are several sub-tasks within the CCSHawk methodology. Some of these sub-tasks are elaborated below in order of their usage:

- *Mapping the Study Area:* The North-Central region of USA is mapped using 19 thematic geo-information layers providing information on roads and trains, waterways, lakes, parks, slope, and so on. These maps require pre-processing and standardization described in Section 3.3.
- *Analytic Hierarchy Process (AHP):* The processed geo-information maps are further analyzed and classified into various categories and levels of importance based on literature using a multi criteria decision analysis process known as AHP (Section 3.4).
- *Cost Map Generation:* The geo-information layers are combined with the weights obtained from AHP to create a pipeline suitability map using weighted overlay. (Section 3.4)
- *Tracking Properties of Capture and Storage Points:* A complete list of capture and storage areas within the study area is accumulated and filtered out along with the appropriately required properties. (Section 3.3).
- *Tracking Convergence into Existent pipeline (Only when tracking the influence of existent pipeline):* Mapping and identification of possible points of convergence between existent pipeline networks and new source and sink nodes (Section 3.3).
- *Clustering:* There are several nodes which are extremely close to one another and many are clustered together to reduce computational load (Section 3.5).
- *Delaunay Pairs:* The clustered node points are then processed using Delaunay triangulation to find feasible pairs suitable as candidates for pipeline routes. (Section 3.5)
- *Pipeline Pathing and Smoothening:* A* algorithm is used to create paths between the Delaunay pairs based on the generated cost map (Section 3.5).

- *Techno-Economic Model:* An analysis of the selected sources and sinks for CO_2 along with the probable pipeline routes is conducted according to a selected techno-economic model (Section 3.6).
- *Mixed-Integer Programming:* A problem is formulated with an objective of reducing the annualized economic cost of setting up the overall CCUS infrastructure for the region provided certain time constraints and capture goals. This problem is solved using a mixed-integer program for both static and dynamic outcomes. (Section 3.7).
- *Text-Mining regulations:* The regulations from CFR Section 49 Title 195 are extracted using text mining and Natural Language Processing (NLP) as the regulatory basis for establishing any pre-conditions for generation of pipeline corridors and CO_2 pipeline technical modeling.

The details about the sub-tasks mentioned above are provided in the following sections

3.3 Mapping the Study Area- Environment, Ecology, and Infrastructure

The study area is the North-Central region of USA consisting of five states including North Dakota, Montana, Wyoming, Utah, and Colorado (further description in Chapter 2). The study area is a net exporter of power (USEPA, 2020b) and has many critical industries based on energy, chemicals, and mining. It is also a fact that this region has significant CO_2 emissions, which industry experts are trying to mitigate using different strategies including the usage of CCUS (EERC, 2019). CCUS requires large scale infrastructure deployment, where the components of the network will be connected by pipelines. Environmental and ecological factors along with existent infrastructure influence pipeline routes to a large extent (Menon, 2011).

3.3.1 Factors Influencing CO_2 Pipelines

Mapping regions suitable for pipeline development is essential. The factors used in mapping the regions could affect the future economic feasibility of pipelines and it's related risk. The geo-information factors used considered in this study include: roads and railways, waterways, lakes, parks, slope, corrosion susceptibility, soil, frost effect on topsoil, faults, towns, the Western Hemisphere Shorebird Reserve Network (WHSRN), areas of critical environmental concern, protected land, federal lands, existent energy pipelines (natural gas, crude oil, crude products) and land use (Balaji, 2020). For purposes of better classification and further processing seventeen geo-information layers are classified into three categories: technical barriers, regulatory barriers, and right-of-way barriers. There are two additional layers of information sometimes used in CCUS studies, namely reservations and cities, which are blocked from consideration in the study area. These regions are removed from the study area as reservations have had geopolitical issues in recent history related to pipelines, while cities are high risk regions and extremely regulated. ArcGIS Desktop 10.7.1 is used for processing each of the nineteen geo-information latyers. The geo-information layers mentioned here are depicted in Appendix A.

3.3.1.1 Natural Barriers

Natural barriers are features of a piece of land that could have a significant effect on the technical feasibility and cost related to construction and maintenance of pipelines (Balaji, 2020).

Waterways and Lakes: There are several waterways, river systems, lakes, and associated water bodies in the study area. It is possible for pipelines to cross rivers and waterways, although, there are several technical problems related to construction of such pipelines including: sinking, higher rates of corrosion and borrowing (Menon, 2011). Hence, it is preferable for pipelines to avoid river and waterway crossings (Middleton et al., 2012a; Sun et al., 2013; Kanudia et al., 2013;

Macharia et al., 2015; Yousefi-Sahzabi et al., 2011; Wan et al., 2011). Lake systems form natural barriers to pipelines. It is possible to cross lakes, however it would to a significant increase in costs as well as risk, which results in most pipelines diverting around the edges of lakes (Menon, 2011; Middleton et al., 2012; Sun et al., 2013; Herzog et al., 2009; Yousefi-Sahzabi et al, 2011 Wan et al., 2011). The data for these hydrologic bodies is obtained from the national hydrography dataset (NHD) (U.S.G.S, 2019) developed by the United States Geologic Survey (USGS). The vector data obtained needed to be clipped to contain features within the study area only, as the dataset consists of hydrologic data for the entire continental United States.

Slope: Slope refers to the continuous change in elevation of land and is an important technical consideration for the construction of pipeline. The cost of pipeline construction increases with higher degrees of elevation (Menon, 2011; Middleton et al., 2012a; Sun et al., 2013; Yousefi-Sahzabi et al., 2011). Slope is a feature obtained by processing the elevation data over a moving window using the "Slope" function on ArcGIS 10.7.1. The elevation data used to prepare the slope raster is known as a digital elevation map (DEM) (90-by-90m resolution) and is obtained from the NASA Shuttle Radar Topography Mission (SRTM) Version 3 (Farr, 2007).

Fault: Fault lines refer to major subsurface discontinuities and these lines could have a negative impact on the maintenance of pipelines (Menon, 2011; Cevik, 2003). The fault line data is obtained from USGS quaternary fault database of 2018 (USGS, et al., 2018a). The vector features need to be buffered with a 1.6km radius using the "Buffering" tool on ArcGIS 10.7.1. The fault data is characterized by the slip rate where large slip rates indicate a larger chance of a major seismic occurrence.

Soil: The study area has a complicated heterogeneity of soil types. In the analysis soil particle size is considered, as it affects retention of water. It is observed that larger particle sizes facilitate quicker water drainage and higher evaporation rates which is better for pipeline quality maintenance (Menon, 2011; He et al., 2015; Norhailan et al., 2012). The soil vector data is obtained from the gSSURGO database of the United States Department of Agriculture (USDA) (Soil Survey Staff, 2015). The data for each state is separate and needs to be combined together using the "Mosaic" function on ArcGIS 10.7.1.

Corrosion and Frost Susceptibility: Corrosion susceptibility of steel and frost action on topsoil are two important study area specific factors, which affect the maintenance of pipelines. Lower corrosion susceptibility and frost action on topsoil is better for maintenance of pipelines (Menon, 2011; Potter, 2013; Cevik, 2003). Both of these geo-information layers are also provided with the gSSURGO dataset (Soil Survey Staff, 2015) and are classified as zones with high, medium or low corrosion susceptibility to steel or frost action.

3.3.1.2 Regulatory Barriers

Regulatory barriers are regions indicated by the federal, state, or local regulatory bodies where the risk of pipeline related problems will have a higher effect on the general populace, wildlife or the environment. This category involves regions with higher population or regions which are ecologically or environmentally important known as Unusually Sensitive Areas (USAs) (Balaji et al., 2020).

Towns: Towns are regions with medium density of population and are distributed through the study area. Locating pipelines near populated regions are not favorable due to security and United States regulations (USA CFR 49, 2019). For the purposes of capturing small towns distributed through the region, schools were considered as population centers and buffered with a 1.6 km radius using the “Buffer” tool on ArcGIS 10.7.1. The shapefile for schools is obtained from the United States National Centre for Educational Statistics for 2019 (Geverdt, 2019).

Protected Land, Areas of Critical Environmental Concern (ACEC) and Western Hemisphere Shorebird Reserve Network (WHSRN): Pipelines are encouraged by regulations to avoid areas of critical ecological value and even if they do pass through such region, the costs related to leasing of land is quite capital intensive (Menon, 2011). Further, regulation identifies regions which are unusually sensitive areas for ecologic development and these regions must be avoided by pipelines (USA CFR 49, 2019). Unfortunately, this aspect is not incorporated into any previous work. The list of regions considered as areas of critical environmental concern (ACEC) is obtained from the Bureau of Land management (Bureau of Land Management, 2020), which indicates areas important for wildlife as well as regions important as places of historic/cultural value. This region is buffered using a 100m radius through the “Buffer” tool on ArcGIS 10.7.1. Protected land data is obtained from the USGS database of PAD-US 2.0 (U.S.G.S, 2018b), which are buffered with a radius of 100m. Protected land entails regions such as wilderness areas, conservation zones and marine protected areas. Further, the list of WHRSN is obtained from the association’s website and is buffered with a 100m radius. The WHRSN sites are regions which are important migratory bird zones which are not covered under protected land parcels.

3.3.1.3 Right-of Way Barriers

The right to have access to pipelines and setting up of related infrastructure is known as right-of-way. In the study area, right-of0way has to be purchased and can significantly impact pipelines costs (Balaji et al., 2020).

Roads and Railways: Roadways and railways are extensively spread throughout the study area and are an essential part of right-of-way system. If pipelines need to cross roads or railways, they need to burrow below them which increases costs, however, it is highly suitable for pipelines to be along the periphery of these networks (Middleton et al., 2012a; Herzog et al., 2009, Callan, 2008; Menon, 2011; Macharia et al., 2015; Yildrim et al., 2014; Wan et al., 2011). This data is obtained from the 2010 TIGER/Line data collected by the United States Census Bureau (U.S. Census Bureau, 2012). The road and rail networks are buffered with a 100m diameter to indicate probable width of infrastructure and another 100m is provided to indicate the periphery. The buffering is done using the "Buffer" tool on ArcGIS 10.7.1.

Federal Lands and Parks: Federal lands are regions in the USA owned by the federal government with the exception of parks and the outer-continental shelf. The cost related to right-of0way is higher in federal lands (USA CFR Section 185, 2019; Middleton 2012a). The data for the overall federal lands is obtained from the National Atlas of the United States (2006); while the data for parks is obtained from the National Park Services database for park systems in the USA (National Park Service, 2006).

Existent Pipeline: Existent pipeline consist of established pipeline routes for resources such as natural gas and crude oil products. These routes are essential, as they mark regions which have pre-existing pipeline rights-of-way, dictating development of future pipelines by providing low-cost pipeline corridors (Potter et al., 2013; Menon, 2011; Middleton et al., 2012a). The data for the pipeline networks are available from the Environmental Information Agency (EIA, 2019). These pipelines are buffered with a 300m radius in order to depict pipeline corridors that are favorable for pipeline development using the "Buffer" tool on ArcGIS 10.7.1.

Land Cover: Land cover defines the vegetation and usage of land. The type of vegetation and usage of land influences the relative cost of acquisition of land. Shrub and herb covered regions would have less economic impact on the land as compared to regions under agricultural occupation (Menon, 2011; Potter et al., 2013; Middleton et al., 2012a; Sun et al., 2013; Broek et al., 2010a). The dataset for land cover is obtained from the USGS GAP/LANDFIRE National Terrestrial Ecosystems dataset for 2011(U.S.G.S Gap Analysis Program, 2016).

3.3.1.4 Absolute Barriers

Some regions in the study area are blocked from usage for trunk pipelines due to regulatory reasons or general geopolitical disparities. These regions would be excluded from the study area (Balaji et al., 2020).

Cities: Cities have population density and can pose significant risk provided any mishaps. The cost related to rights-of-way in urban regions is extremely high. Most pipeline literature encourage distance between trunk pipelines (those that carry large amounts of the product) and high population regions (Middleton et al., 2012a; Broek et al., 2010a; Potter et al., 2013; Sun et al., 2013; Baufune et al., 2013; Nonis et al., 2007; Yousefi-Sahzabi et al., 2011, USA CFR Section 49, 2019). The data for major urban regions are obtained from the 2010 TIGER/Line shapefiles of the US Census Bureau (U.S. Census Bureau, 2012). These regions are provided with a buffer zone of 1.6 km and blocked from the study area using the “Buffer” tool on ArcGIS 10.7.1.

Reservations: In the USA, reservation are regions which are autonomously governed usually presenting Native American tribes. In recent history, pipeline development within reservations have led to geopolitical issues. Hence, in this study, these territories will be removed from the study area to avoid any issues. The Reservation data is obtained from US Census Bureau’s 2010 TIGER/Line database (U.S. Census Bureau, 2012).

3.3.2 Mapping Sources, Sinks and Existent Pipelines

The study area has several possible sources and sinks for CO_2, which need to be processed to select a few nodes. There are some existent CO_2 pipeline routes in the study area too, which need to be marked as part of the study.

Source nodes refer to those entities which have the potential for releasing large amounts of CO_2 emissions which commercially supply CO_2 for EOR purposes. The source nodes are geographically distributed throughout the entire study area and are chosen according to their emission quantities, corresponding to existent pipeline systems as well as industry type. The initial list of source nodes was taken from the EIA's Inventory of U.S. Greenhouse Gas Emissions and Sinks (2020). The sources that emit at least one million tons of CO_2 annually were short-listed. This large quantity of emission was taken as a cut-off as many EOR operations as well as storage sites need a steady supply of CO_2 to necessitate a trunk pipeline being built. Further, sources other than power plants were excluded from the study. It must be noted that the other major sources in the region such as those related to chemical and cement production had large costs for capturing CO_2. Three additional sources of CO_2 were added in the form of commercial suppliers as they were already providing CO_2 for EOR activities. There were a few other sources such as the Comanche (470) power plant (Colorado) and Bonanza power plant (Utah) which were eligible for the study, however, these plants were already part of the CCUS networks which reached into the southern part of USA centered out of New Mexico and Texas. Table 3.1 depicts the 26 sources used in this analysis.

Table 3.1: List of CO_2 sources considered for analysis in CCUS Network

Sr No.	Longitude	Latitude	Source Name	State	Source Type
1	-106.61	45.88	Colstrip	MT	Coal Fired Power Plant
2	-104.88	42.11	Laramie River	WY	Coal Fired Power Plant
3	-108.79	41.74	Jim Bridger	WY	Coal Fired Power Plant
4	-101.16	47.38	Coal Creek	ND	Coal Fired Power Plant
5	-101.32	47.28	Leland Olds	ND	Coal Fired Power Plant
6	-101.84	47.37	Antelope Valley	ND	Coal Fired Power Plant
7	-101.84	47.36	Great Plains Gasification Plant	ND	Chemicals and CO_2 Supplier
8	-111.03	39.17	Hunter	UT	Coal Fired Power Plant
9	-111.08	39.38	Huntington	UT	Coal Fired Power Plant
10	-112.58	39.51	Intermountain	UT	Coal Fired Power Plant
11	-107.59	40.46	Craig	CO	Coal Fired Power Plant
12	-101.21	47.07	Milton R. Young	ND	Coal Fired Power Plant
13	-105.78	42.84	Dave Johnston	WY	Coal Fired Power Plant
14	-110.60	41.76	Naughton	WY	Coal Fired Power Plant
15	-103.68	40.22	Pawnee	CO	Coal Fired Power Plant
16	-110.22	42.24	Shute Creek Facility	WY	Natural Gas and CO_2 Supplier
17	-105.39	44.29	Wyodak	WY	Coal Fired Power Plant
18	-105.38	44.29	Wygen 1-2-3	WY	Coal Fired Power Plant
19	-105.38	44.29	Neil Simpson 2	WY	Coal Fired Power Plant
20	-105.46	44.39	Dry Fork Station	WY	Coal Fired Power Plant
21	-101.81	47.22	Coyote	ND	Coal Fired Power Plant
22	-107.19	40.49	Hayden	CO	Coal Fired Power Plant
23	-105.03	40.86	Rawhinde Energy Station	CO	Coal Fired Power Plant
24	-104.88	40.24	Fort St. Vrain	CO	Coal Fired Power Plant
25	-107.60	43.27	Lost Cabin	WY	Natural Gas and CO_2 Supplier
26	-110.42	42.50	Riley Ridge	WY	Natural Gas and CO_2 Supplier

Sink nodes in this thesis refer to those entities which have the potential for geologic storage of CO_2 in sufficient quantities. This thesis includes both saline aquifers and CO_2 EOR operations sites which could eventually turn into depleted oilfields in the study. These reservoirs and their properties have been identified from various discrete sources of Regional Carbon Sequestration Partnership Initiative of the National Energy Technology Laboratories (NETL, 2015), the

NATCARB Viewer 2.0 database (2020) and the University of Wyoming's WYRIT database (2020). Table 3.2 depicts the list of sink nodes used in this thesis for analysis.

Table 3.2: List of CO_2 sinks considered for analysis in the CCUS network

Sr No.	Longitude	Latitude	Source Name	State	Source Type
1	-101.84	47.36	Great Plains Gasification Plant	ND	Saline Aquifer
2	-107.59	40.46	Craig	CO	Saline Aquifer
3	-101.21	47.07	Milton R. Young	ND	Saline Aquifer
4	-104.93	45.35	Bell Creek	MT	EOR
5	-108.81	44.87	Elk Basin	WY	EOR
6	-107.00	42.74	Grieve	WY	EOR
7	-108.57	43.78	Hamilton Dome	WY	EOR
8	-107.04	42.10	Bairoil Fields	WY	EOR
9	-108.32	42.85	Beaver Creek	WY	EOR
10	-108.54	41.57	Patrick Draw	WY	EOR
11	-108.88	40.10	Rangeley Weber	CO	EOR
12	-109.08	44.23	Spring Creek	WY	EOR
13	-108.91	44.36	Oregon Basin	WY	EOR
14	-109.06	44.14	Pitchfork Field	WY	EOR
15	-106.31	43.43	Salt Creek	WY	EOR
16	-102.30	46.88	Red Trail Energy	ND	Saline Aquifer
17	-111.64	48.70	Kevin Dome	MT	Saline Aquifer
18	-110.87	39.62	Gordon Creek	UT	Saline Aquifer
19	-110.21	43.26	Moxa arch	WY	Saline Aquifer
20	-110.35	39.27	Woodside Dome	UT	Saline Aquifer
21	-104.42	46.61	Cedar Creek Anticline	MT	EOR
22	-103.83	46.21	Cedar Hill	ND	EOR
23	-103.07	47.38	Little Knife	ND	EOR
24	-103.57	47.34	Rough Rider	ND	EOR
25	-102.95	48.30	Beaver Lodge	ND	EOR

The study region also has 4 major pipeline systems carrying CO_2 for EOR projects majorly in Colorado, Wyoming and North Dakota with parts of the pipeline extending into Canada. The 4 pipeline networks include: The Green Core pipeline, Exxon pipeline, FDL pipeline and the Dakota Gasification pipeline which runs to the Weyburn Field at Canada. The routes for these pipelines were obtained from the North Dakota Pipeline Authority database (2020), the University of

Wyoming's WYRIT database (2020) and NETL report on CO_2 pipeline infrastructure (2015). It must be noted here that other than Wyoming, no other state provides digitized copies of their CO_2 pipeline routes due to reasons of security. Those parts of the pipelines which are in regions outside Wyoming have been digitized from PDF files using ArcGIS 10.7.1. Figure 3.2 depicts the sources and sinks of CO_2 along with the existent pipeline infrastructure in the region.

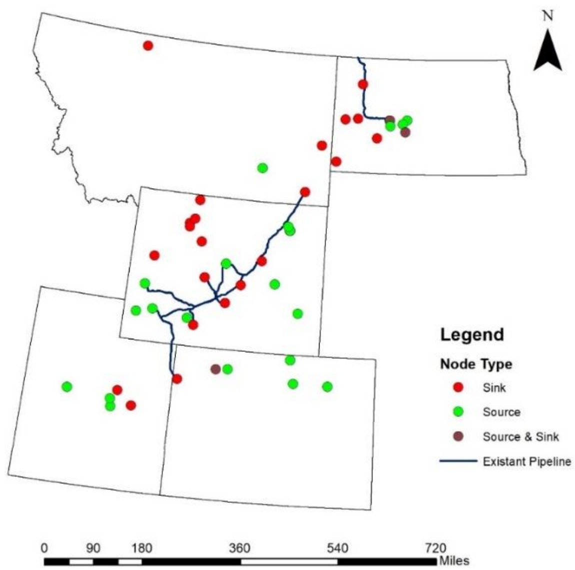


Figure 3.2: Sources and sinks under consideration along with existent CO_2 pipeline in study area.

3.3.3 Mapping Connectors

When analyzing the effect of existent pipeline infrastructure on the growth of future resources and infrastructure for a CCUS network, there needs to be a point where future and current infrastructure commingle. To analyze this, connector nodes are used. These nodes are points on existent pipelines where nearby source or sink nodes can join/connect to existent pipelines. These nodes also consist of points in the map where existent pipelines split or merge leading to different destinations.

- A proximity analysis is done to check the nearest point of connection between every node (source and sink) and existent pipeline using the "Proximity" tool on ArcGIS 10.7.1. These points on the existent pipelines are chosen as connector nodes.

- Connector points which are more than 160 km away from their respective source/sink nodes are removed from the list, while the rest are saved as a new feature consisting of their geographic coordinates.
- Further, forks and splits in existent pipeline are mapped and added to the existent connector node dataset.

Connector nodes are not sources or sinks of CO_2 and hence will not add or store any CO_2 to the CCUS system.

3.4 Cost Map Generation

The study area has a diverse environment and ecology with pre-existing transportation infrastructure. The three barrier criteria: natural barriers, regulatory barriers and right-of-way barriers play an important role in influencing the development of new potential pipeline routes through cost of construction and maintenance as well as restrictions on zones of pipeline development. Thus, it critical to consider these geo-information layers in greater detail while mapping areas best suited for pipelines. The development of a “Cost Map” enables the mapping of these varied factors in an amenable fashion and analyze the study area according to its feasibility towards development of infrastructure.

The process of development of the Cost Map has three important aspects detailed in Figure 3.3; where initially the geo-information layers are identified and processed and weighted by the AHP technique. Further, these layers are combined using a weighted overlay (summation) and the few of the geo-information layer zones are excluded from the study area.

3.4.1 Analytic Hierarchy Process (AHP)

The AHP technique is used to evaluate the importance of each geo-information layer compared to the others to quantify their effect on the CO_2 pipeline corridors. Utilizing a set of evaluation criteria, the AHP techniques chooses the best alternatives from a set of varying options. In this work the AHP technique is shaped to discern the best combination of weightages to be assigned to each criterion based on pairwise comparison between the geo-information layers in a hierarchical structure. The respective weightages of each criteria defines the importance related to pipeline suitability (Kolios et al., 2016). The measurement of inconsistency is obtained using a measure known as consistency ratio (CR). The steps involved in the implementation of AHP are as follows according to Balaji et al. (2020):

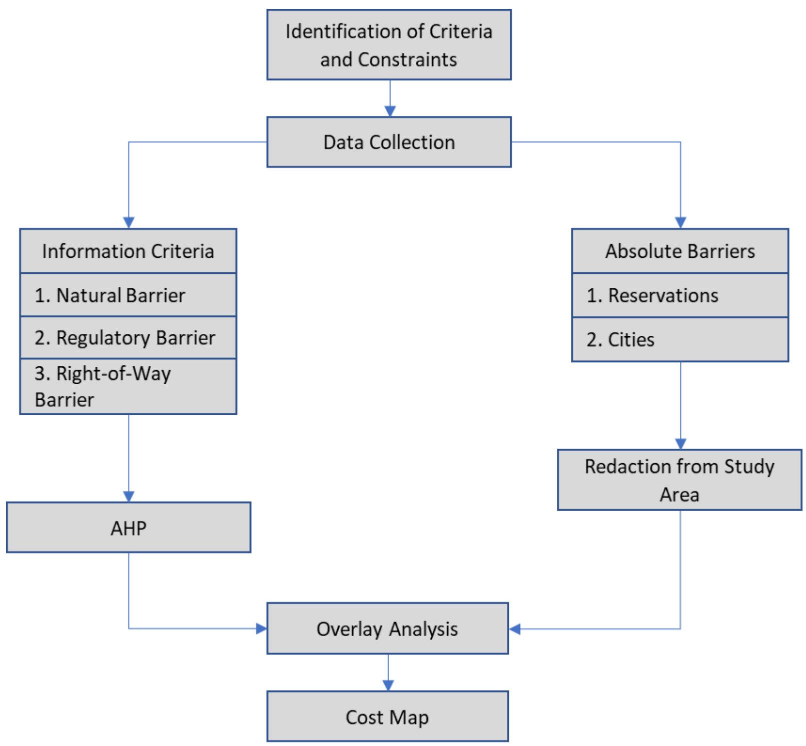


Figure 3.3: Process for generation of cost map using AHP and weighted overlay (Balaji et al., 2020).

- A multi-hierarchy structure is established to categorize the geo-information layers used to map the study area.

- A pairwise matrix for each hierarchical category is generated of dimensions $n \times n$, where n is number of criteria in each category. Let P_{ij} be the preference score of criteria i to criteria j for a specific hierarchy category as suggested by Saaty (1980). A score is provided for each criteria as compared to another in the same category (0-9) indicating the relative importance (Saaty, 1980). According to Saaty: 9 indicates extreme importance, 8: very, very strong, 7: very to extreme importance, 6: strong plus, 5: strong importance, 4: moderate plus, 3: moderate importance, 2: weak and 1: equal importance. P_{ij} denotes the entry in the i th row and the j th column of matrix m. The entries of preference score P_{ij} and P_{ji} must satisfy Equation (3.1):

$$P_{ij} \times P_{ji} = 1 \tag{3.1}$$

- A normalized pairwise comparison matrix m is established where, the sum of all values in each row is 1. Equation (3.2) is used to calculate P_{ij} for each entry of the matrix m.

$$\overline{P_{ij}} = \frac{P_{ij}}{\sum_{i=1}^{n} P_{ij}} \tag{3.2}$$

- Equation (3.3) is used to calculate the average of each row. The overall weight of each criteria used in the cost map is obtained by multiplying the respective weight of criteria in their category with the weight of the category as compared to the other.

$$w_i = \frac{\sum_{i=1}^{n} \overline{P_{ij}}}{n} \tag{3.3}$$

- Equation (3.4) is used to calculate the consistency index (CI), where λ_{max} is the maximum eigenvalue of the comparison matrix.

$$CI = \frac{\lambda_{max} - n}{n-1} \tag{3.4}$$

- The level of bias in the comparison matrix is indicated by the consistency Ratio (CR) which is obtained from Equation (3.5). RI is the random consistency index indicated in Table 3.3.

$$CR = \frac{CI}{RI} \tag{3.5}$$

- Bias is indicated in the study if the CI is less than or equal to 0.1, and study needs to be repeated.

Table 3.3: Random Index (RI) for different number of elemnts(n) (Saaty, 1980)

n	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0
RI	0.0	0.6	0.9	1.1	1.2	1.3	1.4	1.5	1.5

3.4.2 Processing and Overlay

The 19 geo-information layers discussed in Section 3.3 are first processed to make the following adjustments for the purpose of preparing a proper Cost Map on ArcGIS Desktop 10.7.1:

- All the raster and vector information layers are first brought to a uniform coordinate system: NAD 1983 UTM Zone 14N.
- All raster and vector information are brought to a uniform datum: D_north_america_1983.
- All the vector data are converted to raster information layers by using the "to Raster toolbox" based on varied cell assignment methods. For conversion of datasets from vector to raster, the type of information conveyed needs to be determined. For instance, in case of "discrete" data such as slip rate in mapping of Faults, the raster conversion is done by using the feature, "Slip Rate" for classification while area with no faults is defined with "No Value". In the case of "continuous uniform" data which conveys if a feature such as Reservations exists, an additional feature with uniform value is created to be the basis of conversion to a Raster file.
- Layers are adjusted to the resolution (cell size/pixel size) of 100-by-100m (to scale). If the cell size does not match the intended format, the "Resample tool" is used to bring all the raster files to the target resolution. The "resample Tool" works on the concept of interpolation/aggregation for restructuring the resolution of raster files. In this study, "Bilinear

Interpolation" is used as the default resample method for continuous datasets like Slope, which uses the weighted distance average of the four nearest cells; while the "Nearest Neighbor" interpolation method is used for resampling discrete data such as Lakes which uses the value of the nearest or most dominant overlying cell as the default assignment of the cell. An additional feature of this "Resample tool" is the option to "snap" the information to another layer, which makes sure that all the resized cells overlie each other accurately.

- A vector data file of the boundary of the study area with the uniform coordinate system and datum is created to show the extent of the study area.
- All the geo-information layers are overlaid on the vector layer containing the study area to extract features which lie within the boundaries of the study area using the "Clip function". This tool is modified to break features to smaller features in case certain data are present both inside and outside of the study area by modifying their length or area, but retaining the underlying identifying information.
- The layers are reclassified to a numeric scale of 1 to 10 using the tool "reclassify" on ArcGIS 10.7.1. For categorical data, the maps are scaled according to their favorability towards pipelines, with 1 being the most favorable and 10 being the least favorable; while in the case of continuous data, they are standardized from a scale of 1 to 10.

The AHP method provides the weighted values for sampling each of the geo-information layers considered. The various layers are first reclassified according to their relative importance determined through the AHP method. After which the individual weights of each layer are used in a linear overlay to a form the Initial Cost Map (except the layer for "Reservations" and "Cities").

$$Cost = \sum_{i=1}^{n} x_i w_i \tag{3.6}$$

Equation (3.6) details the basic overlay formula used in the preparation for the "Overlay" function of ArcGIS 10.7.1, where $Cost$ is the favorability of each cell towards the positioning of a pipe in that cell; x_i is the individual value of the layer in the specific cell (1-9); and w_i is the weight of the specific layer obtained from the AHP method. Further, the Cost Map is standardized to a scale of 0 to 100, where lower cell cost indicates higher favorability for CO_2 pipeline construction and vice versa. It must be noted that the obtained geo-information layer is in raster format. The two geo-information layers considered under the "Absolute Barrier" category are excluded from the study area using the "Mask" function. The reason for exclusion was stated in Section 3.3 of this chapter. The output of the "Mask" function is taken as the final Cost Map and used for further processing.

3.5 Candidate Route Generation

This step of the methodology uses the geographic information related to source and sink nodes and establishes legitimate paths between them using clustering, pairing using a mesh algorithm and a deterministic graph technique to create routes between node pairs. The goal of this step (in simple terms) is to establish the means for getting from point A to point B in the most efficient fashion. When multiple geographically distributed points are considered, this step generates a network of interconnected routes between the points to get from one location to another.

3.5.1 Clustering

The first step in the pipeline network generation process is the clustering of nodes currently being considered in the study area. The study area has several nodes, which are in the immediate vicinity of each other. During the clustering process, both sources and sinks of CO_2 are considered together. The number of possible candidate arcs between the nodes is given by $\sum_{i=1}^{n-1} x$ where x is the number of nodes being considered. Thus, by clustering the nodes, which are close to each other, it is possible to control the extent of the network. This type of clustering aims at reducing the computational load of the networks. The number of nodes exponentially increases the computational burden of future operations by increasing the number of probable candidate pipeline routes under consideration.

A clustering operation consisting of a constant radius is used on the geographic coordinates of the nodes. This means, that the generated clusters can represent any number of elements, however, the radius of the clusters will remain constant. The clustering uses a constant radius of 24.14kms (15 miles) on geodesic coordinates. The following steps are used in the generation of clusters:

- Initially the geographic coordinates of all the nodes in the study area are considered and a proximity analysis is conducted to find the distance of each point from one another using the "Proximity" toolset on ArcGIS 10.7.1.
- The points which have a distance less than 24.14 km from each other (15 miles) in the proximity study are separated and grouped into a new vector file.
- All points in the new vector file are buffered using the "Buffer" tool on ArcGIS 10.7.1 with a radius of 24.14kms.
- The points within the radius of each other's buffer are collected as a cluster.
- The geographic coordinates of the points to be clustered together are first converted to x-y coordinates and their unweighted aggregate is taken to find the centroid of the cluster.
- The points which are part of the cluster are replaced by the centroid as a representative in the analysis.

The source properties are reevaluated by summing up the total emissions from all the sources in the cluster into a single entity. For sink nodes, the total storage capacity is added up while, the other properties such as depth, reservoir thickness, permeability and porosity are replaced by a mean value. It must be noted here that the nodes are aggregated together despite being sources or sinks, thus leading to generation of nodes which may be both source and sink of CO_2. This aggregation of sources and nodes into a single entity is done in order to reduce the size of CCUS network for computational efficiency.

A similar clustering technique is done to check for proximity between the connector nodes (this clustering is exclusive to only the connector nodes, that is they are not paired with sink or source nodes). Such a clustering is done because the point of connection between nodes and existent pipeline can result in points which are "coincidental" or close to one another (24.24km of each other). These connector nodes are used only while analyzing the effect of existent infrastructure on new infrastructure.

3.5.2 Delaunay Pairs

It is necessary to create a network of pipeline routes between the various nodes after the clustering operations. This can easily be done by assuming that a route exists between each node. The networks generated after clustering still can consist of many potential routes. Some of these routes may coincide with existent paths between other nodes and be quite cumbersome and baseless. Thus, to reduce the number of pipeline routes or "candidate arcs" between nodes and to make more meaningful routes, the Delaunay triangulation algorithm is used. Delaunay triangulation is a method of combining points in a plain (Delaunay, 1934), in which, an edge of a Delaunay triangle is such that a circumcircle drawn through the edge points coincide with a third point from the overall set of points. Delaunay triangulation is extensively used in network design and numerical simulation for mesh design and in determination of the Voronoi plains (Fang et al., 1993). A set of corollaries related to Delaunay triangulation (Fang et al., 1993) include:

- The set of all the points generate edges, such that the boundary of the network forms a convex hull, if not triangulation needs to be repeated.
- Every point in the node set needs to be a part of at least one triangle.
- Generated triangles should be such that the circumscribed circle is empty, that is, no other points should be within the generated circle.
- The triangulated points may share common edges amongst themselves.
- Generated triangles will maximize angles between the contained edges.
- Every set of 3 points will have a triangle generated between them except, if they are co-planar in nature.
- Every generated triangle between a set of points will be unique unless the set of points are circumferentially equidistant.

For the work in this thesis, the node set used as input for the Delaunay algorithm is 2-dimensional in nature, and hence uses a divide and conquer technique. Initially the points are laid out in x-y coordinates. The dataset is divided into planes, each containing maximum three data points and edges are generated. Further, the edges are created between the various different planes keeping in mind the above-mentioned corollaries. The process is reiterated such that the edges generated by the algorithm reflect the maximum possible angles of vertices between the edges of the triangle while maintaining the corollary conditions.

The algorithm used for obtaining the Delaunay triangulation is provided in Appendix B and is scripted in Python 3.2. The generated algorithm utilizes the “Delaunay” package available under the Scipy spatial analysis toolbox (Millman et al., 2011). The function uses the aforementioned "Divide and Conquer" strategy to formulate the triangulations and provides an output of generated points for each triangle. The algorithm further goes on to extract edges from each of the provided triangles and removes duplicate edges from consideration. The edges obtained from this algorithm is referred to in this work as "Delaunay Pairs" and is used as the basis for generating pipeline routes. The reduction in number of potential "Candidate Arcs" utilizing the Delaunay algorithm is not fixed and is dependent on the geographic location and the Euclidean distance between every node.

In addition to the Delaunay pairs, in the case where the effect of existent pipeline is being analyzed, connector nodes are also considered for the Delaunay triangulation in addition to the clustered nodes (sink and source nodes). Further, existent pipeline routes between the connector points are also added to the Delaunay pairs and if there are any duplication, they are removed. It should be noted that, the existent pipeline-based Delaunay pairs do not need a new route generated between them, as they have an existent pipeline between them.

3.5.3 Route Generation – A star Algorithm (A*)

A graph network algorithm known as A* algorithm is used to generate a path between two points. This algorithm generates a Least Cost Path (LCP) between the points in a directed fashion such that the movement of the algorithm along the network is always towards the destination while measuring the cost of moving from one cell to another.

The A* algorithm is coded in R mathematical package due to its quick nature of data management and straightforward implementation. The A* algorithm in a two-dimensional setting tends to visualize a map as a graph. To visualize a raster file as a graph, each cell (pixel) of the map needs to be considered as a graph point. Thus, raster files are converted to a two-dimensional matrix format due to the ease and accessibility of stored data. The size of the matrix used in the algorithm dictates the actual computational time to transverse between points using the algorithm. Since the A* algorithm basically finds a LCP to the destination (from the source), increasing the size of matrix increases the distance between the points, leading to larger computational loads. There are two inputs required for running the A* algorithm in this thesis: The cost raster containing the cost map obtained from previous steps/section; and the source and the sink locations. The cost map from the previous steps is in a raster file format, which is first changed to a resolution of 200-by-200 m to enable quicker computation of routing. The source and sink locations which are stored as vector points on a map file, are also converted into raster format using the dimensions, extent and resolution similar to the Cost Map file before using it as an input for the execution of the A* algorithm.

A* algorithm is a directional graph search algorithm, which means the algorithm does not wander about the matrix blindly, but, rather uses the aid of two distance heuristics: cost of movement from the source; and distance to destination. The heuristic used to calculate distance is the Euclidean distance formula shown in Equation (3.7), where x_i and y_i are the row and column numbers of the first cell and x_j and y_j are the row and column number of the second cell:

$$Distance = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{3.7}$$

Although, Equation (3.7) provides the distance between two points, it does not provide the cost of movement. The "cost" mentioned here is the difficulty of locating a pipeline section in that specific region, in other words, Cost Map value related to that cell. Equation (3.8) provides the formula for cost of movement to the immediate next cell, where $cost$ is the cost value of the child cell; x_i and y_i are the row and column number of the parent cell; and x_j and y_j are the row and column number of the child cell . This cost of movement can be calculated only one step at a time and hence, the total cost of moving to current cell from starting point is provided by Equation (3.9) where $accumulated\ cost$ is the overall cost (cumulative cost from source) of moving to the parent cell.

$$Cost\ of\ Movement = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} * Cost \quad (3.8)$$

$$Overall\ Cost = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} * Cost + accumulated\ cost \quad (3.9)$$

Figure 3.4 explains the “Cost of Movement” in more details.

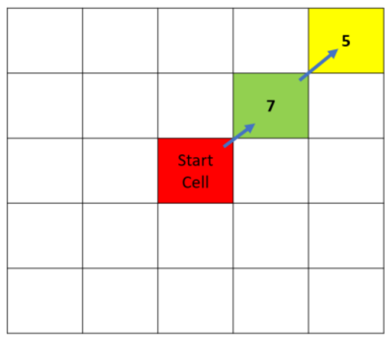


Figure 3.4: Example for cost of movement.

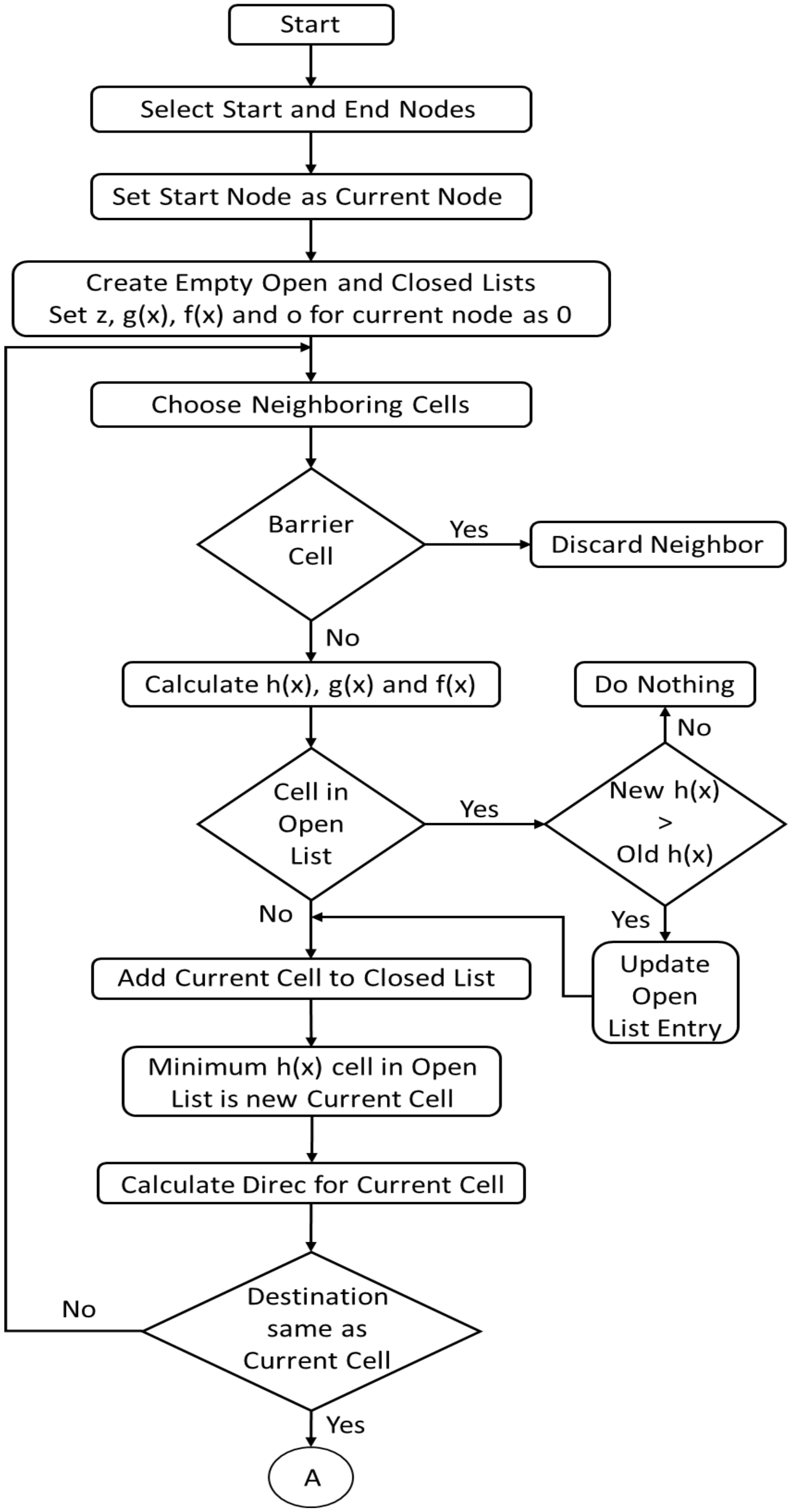
To find the cost of movement from the red cell to the yellow cell, the cost of movement from red cell to green cell needs to be calculated first. The distance from red cell to green cell is 1.414 units and the cost of green cell is 7, thus the cost of movement to green cell from red is 9.899 using Equation (3.8). This value of movement to green cell is stored as the accumulated cost, and further the cost of movement from green cell to yellow cell is computed. The distance from green to yellow cell is in 1.414 units and cost of yellow cell is 5. Thus, the cost of movement from green to yellow is 16.97, referred to as the overall cost of movement to the yellow cell from the red cell calculated using Equation (3.9).

The final heuristic for guiding the movement depends on the addition of the cost of movement from the source cell to current cell and the distance of the destination cell to the current cell depicted by Equation (3.10), where g(x) is the total of the distance heuristics; h(x) is distance to destination cell calculated using Equation (3.8) and f(x) is the cost of movement to current cell from start node calculated using Equation (3.9).

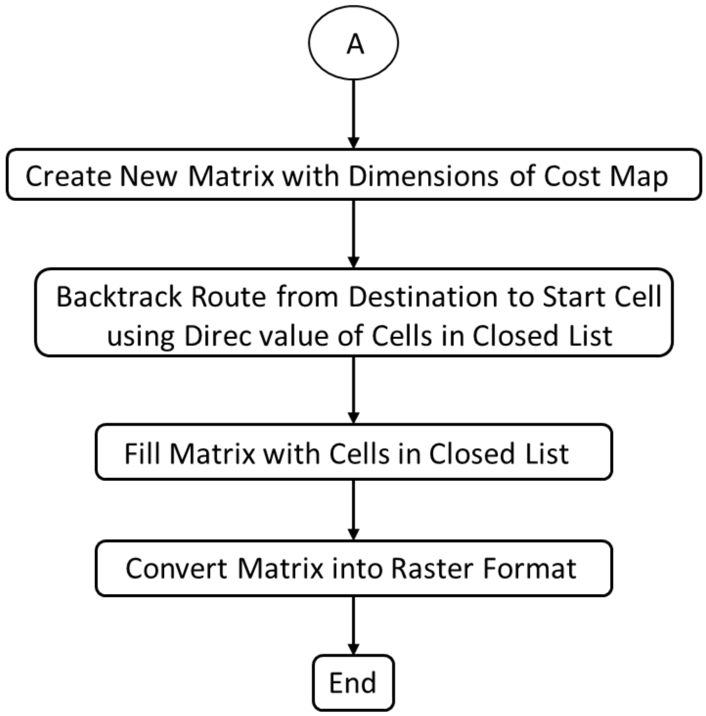
$$g(x) = h(x) + f(x) \tag{3.10}$$

The algorithm used for modeling the A* procedure is provided in Appendix B. Figure 3.5 depicts the general workflow for implementing A* algorithm. In the workflow, z is the accumulated cost of moving to the parent cell; and $Direc$ is the direction of the parent cell as compared to the neighboring cell (depending on g(x)). The workflow is as follows:

- Select the start and end nodes using an identifier related to the nodes (source and sink IDs Table 3.2 and 3.3) and set the start node as the *current cell*. Create 2 empty lists: *open list* and *closed list*. *Open list* is the list of cells which have been processed and waiting for selection as potential "*current cell*"; while *closed list* is an archive of cells already used as *current cell.*
- Set the $h(x)$, $f(x)$ and $Direc$ with a default value of 0 for the initial *current cell*.
- A function for selecting neighboring cells is used, which selects the immediate 8 cells surrounding the *current cell* and checks if any of the cells are *barrier cells*, that have no cost value. Neighboring cells which are *barrier cells* are discarded while the remaining cells are considered for addition to the *open list* and their $g(x)$ and $f(x)$ values are calculated.
- If any of the neighboring cells are already part of the *open list*, then the newly calculated $g(x)$ value is compared to the old $g(x)$ value. If the new $g(x)$ value is lower than the old $g(x)$, then the information of the cell is updated with the newer values, else the old values are retained.
- The *current cell* is archived as part of the *closed list* and the cell with the least $g(x)$ value from the whole *open list* is chosen as the new *current cell.* Once the new cell is selected, its direction ($direc$) as compared to its parent (previous *current cell*) is calculated.


Start
Select Start and End Nodes
Set Start Node as Current Node
Create Empty Open and Closed Lists
Set z, g(x), f(x) and o for current node as 0
Choose Neighboring Cells
Barrier Cell
Yes
Discard Neighbor
No
Calculate h(x), g(x) and f(x)
Do Nothing
No
Cell in Open List
Yes
New h(x) > Old h(x)
No
Yes
Update Open List Entry
Add Current Cell to Closed List
Minimum h(x) cell in Open List is new Current Cell
Calculate Direc for Current Cell
No
Destination same as Current Cell
Yes
A

(a)



(b)

Figure 3.5: Workflow for A* algorithm. (b) is a continuation of (a).

- If the *current cell* is not the destination, then the process reiterates several times till the destination node is chosen as the *current cell*.
- A new empty matrix is created with similar dimensions as the cost matrix.
- The *closed list* values are backtracked using the recorded *direc* values and recorded as entries into the new empty matrix.
- The new matrix is then converted to raster data format with coordinates and extent similar to the cost map.
- In addition to the length of the route, the average cost of traveling through the cost map is calculated.

It must be noted that there are several special conditions built into the algorithm for dealing with troubles related to the choosing of neighboring cells. For instance, if there are no neighboring cells or just one neighboring cell to the *current cell*, then the algorithm backtracks by one iteration and assigns the last evaluated *current cell* as *barrier cell*. To record the direction of the parent cell in comparison to the cell in consideration, a numerical numbering system is employed to assign a categorical value for each direction, which helps in the backtracking process.

The numbers used in the numeric system are categories defining the direction rather than continuous values. In addition, the process uses a dynamic algorithm structure, where every new value calculated is stored in a pre-defined data structure, with the explicit purposes of quicker computation. If such a structure is not created, the values for $g(x)$ and $f(x)$ must be recalculated for each iteration, increasing the time complexity of the algorithm. However, such dynamic processing has its limits when the number of stored values becomes exponentially large, creating problems in recovering data, which is the main reason for restructuring the input layers to a resolution of 200-by-200 m.

The output obtained from the A* algorithm is not in an ideal format for constructing networks. The following process is used to make the routes suitable for the required network:

- A* algorithm data is obtained in the format of float point rasters, which indicates that the value stored in the raster are in the float format using decimal digits. Hence, the "Raster Calculator" tool is used to convert the float format rasters to an integer format raster in ArcGIS 10.7.1.
- The raster integer format routes are converted into line vector datasets using the "to Polyline" tool in the "Conversion" toolbox of ArcGIS 10.7.1, which converts the data to line features by connecting the centers of each raster cell representing the route.
- The line features in each individual file can consist of multiple different line features. These features are merged together to a single entity in each route file. The "Simplify Line" tool of ArcGIS 10.7.1 is used to reduce unnecessary bends and turns in the route generated due to conversion of feature from raster to vector data format. The setting for the tool is based on the

Douglas-Peucker algorithm which retain critical points while eliminating unnecessary points representing the feature (ESRI, 2011). The algorithm begins by connecting the endpoints of a line with a trend line. The distance of each vertex to the trend line is measured perpendicularly. Vertices closer to the line than the tolerance are eliminated. The line is then divided by the vertex farthest from the trend line, which makes two new trend lines. The remaining vertices are measured against these lines, and the process continues until all vertices within the tolerance are eliminated.

The candidate routes generated from this study are analyzed for their lengths as well as their connections to be used in further analysis.

3.6 Cost Model

The cost model refers to the techno-economic analysis of the edges and nodes related to the CCUS network. Here, edges refer to the candidate arcs generated in the previous section; while the nodes refer to the source and sink nodes (and connector nodes) which have to be analyzed for this study.

CO_2 in pipelines can be transported as a liquid, supercritical fluid or in two phase flow (liquid and gas). CO_2 is considered to be in supercritical state at a temperature of at least 31.1^o C and a pressure above 7.3 MPa; while the fluid is called dense if the temperature of the fluid falls below 31.1^o C but the pressure remains over 7.3MPa (Det Norske Veritas Germanischer Lloyd [DNVGL], 2010). It is recommended that the CO_2 flow in pipelines remain in the supercritical or dense phase to reduce friction losses (Knoope et al., 2013). Here it is assumed that the temperature of CO_2 will be at a constant temperature (ambient) of 11.66 ^{o} C with a minimum pressure of 8.27 MPa and maximum pressure of 15.16MPa (NETL, 2018) in order to maintain the fluid in a dense phase.

The density and viscosity of CO_2 are obtained from the transport model developed by McCollum et al. (2006). The correlations were generated through regression analysis and curve fitting of observed density and viscosity at varied temperature and pressure intervals. Equation (3.11) and (3.12) are used for measuring the density and viscosity of CO_2 respectively, where ρ is the density of CO_2 in the system; μ is the viscosity of CO_2 in the system; a_n and b_n is temperature dependent constant defined by McCollum et al. (2006); and $P_{op}$is the pressure in the system. Appendix C lists the coefficient tables related to Equation (3.11) and (3.12).

$$\rho = a_1 P_{op}^6 + a_2 P_{op}^5 + a_3 P_{op}^4 + a_4 P_{op}^3 + a_5 P_{op}^2 + a_6 P_{op} + a_7 \quad (3.11)$$

$$\mu = b_1 P_{op}^6 + b_2 P_{op}^5 + b_3 P_{op}^4 + b_4 P_{op}^3 + b_5 P_{op}^2 + b_6 P_{op} + b_7 \quad (3.12)$$

The regression analysis done by McCollum et al. (2006) is only for certain fixed pressures and temperatures. If the ρ and μ needs to be analyzed for a different pressure and temperature not listed in Appendix C, they need to be interpolated using Equation (3.13) and (3.14) where T is temperature; the suffix $high$, low and op stand for value at higher scale, value at lower scale and operating point.

$$\rho = \frac{(\rho_{high} - \rho_{low}) * (T_{op} - T_{low})}{(T_{high} - T_{low})} + \rho_{low} \quad (3.13)$$

$$\mu = \frac{(\mu_{high} - \mu_{low}) * (T_{op} - T_{low})}{(T_{high} - T_{low})} + \mu_{low} \quad (3.14)$$

3.6.1 Capture Cost Model

The source nodes were obtained from the US EPA's Inventory of U.S. Greenhouse Gas Emissions and Sinks (2020). The capture cost model for CO_2 sources is basic and follows a simple formula of evaluating cost of capture related to every ton of CO_2 captured. From the EPA's above-mentioned database, the emission levels as well as the production levels (for commercial outlets) for the years from 2010 to 2018 are available. The study for this thesis has taken 2019 as the start year of the analysis. Thus, the CO_2 emission levels for each source needs to be adjusted for 2019, and this is done using regression analysis for each individual source based on the previous years' CO_2 capture levels.

According to the Global CCS Institute (2017), the cost of CO_2 avoidance is shown in Table 3.4 for First-of-Kind and N^{th}-of-Kind capture technology. The total emission of a power plant supplier cannot be captured fully, hence, as indicated by global CCS institute, a factor of 0.8 is multiplied with the emission rate in order to indicate the maximum value of possible CO_2 capture available for CCUS purposes as shown in Equation (3.15). The cost of production per ton of CO_2 as well as the maximum capture quantity from a source node is used for decision analysis in choosing sources.

$$Maximum\ CO_2\ from\ Source = 0.8 * CO_2\ Emission \quad (3.15)$$

Table 3.4: Cost to Avoid a ton of CO_2 (Global CCS Institute, 2017).

Technology	**FOAK (USD/ton of CO_2)**	**NOAK (USD/ton of CO_2)**
PC Supercritical	67.5	44
Oxy-Combustion Supercritical	59.5	41
IGCC	86	35
NGCC	78	32
Iron and Steel	66	54
Cement	113	92
Natural Gas	10.5	9.4
Fertilizer	14.4	12.8
Biomass	10.5	9.4

3.6.2 Pipeline Cost Model

The pipeline cost model applies to all probable pipeline routes obtained from the A* algorithm. The objective of this section is to establish the technical and economic details related to pipeline routes for varying diameters. The reason for choosing different diameter pipelines in the analysis is to account for the varying flow rates of CO_2 between the nodes. Another reason for using different diameters is that the cost for pipeline construction increases with the diameter.

NETL (2018) calculates cost of a pipeline route between two fixed points for five different diameter pipelines, namely 12, 16, 24, 32 and 40in diameter pipes along with their associated flow capacities. In this study, 12, 16, 24 and 32in pipeline are used because the EPA's database of U.S. Greenhouse Gas Emissions and Sinks (2020) does not indicate any usage of 40in diameter pipes in the region. The pipeline is assumed to be of AMSE X-70 grade with a minimum yield strength of 70,000 psi (Middleton, et al, 2012a; Knoope et al., 2014; Morbee et al., 2012; McCollum et al., 2006; NETL et al., 2018). ASME X-70 specification is utilized as it is the most cited steel grade used in CO_2 pipeline literature. As related to the ASME X-70 steel grade, the following assumptions are also made (NETL, 2018): pipeline elastic factor (E) is taken as 1, the safety factor (F) is taken as 0.72, the maximum operating pressure (P_{MOP}) is taken as 15.3 MPa and the strength of pipe (S) is taken as 482.633 MPa.

The mass flow rate ($\dot{m}$) for the pipeline diameters (outer diameter, D_o) is obtained from NETL (2018) and listed in Table 3.5. The mass flow rates listed in table are used to calculate the maximum flow rates through these pipelines using Equation (3.16), where Q_{mx} is the maximum flow rate through pipe; and $Capacity\ Factor$ is taken as 0.8.

$$Q_{mx} = \frac{\dot{m}}{Capacity\ Factor} \tag{3.16}$$

Table 3.5: Mass Flow Rate according to Pipeline Diameters (NETL, 2018).

$\mathbf{D_o}$ (in)	12	16	24	32
$\dot{\mathbf{m}}$ ($MtCO_2$/yr)	2.5	5	15	30

The minimum flow rate for each pipeline configuration is the maximum flow rate of the pipeline of the immediately preceding pipeline configuration (the maximum flow rate of pipeline with 12in diameter is the minimum flow rate of pipeline with 16in diameter). The thickness of the pipe is calculated using Equation (3.17), where t is pipe thickness. Equation (3.18) is used calculate the inner diameter of the pipeline where D_i is the inner diameter of the pipeline.

$$t = \frac{P_{MOP} * D_o}{2 * S * E * F} \quad (3.17)$$

$$D_i = D_o - 2 * t \quad (3.18)$$

Reynold's number and the fanning friction factor are calculated from the diameter and average velocity (1.47m/s obtained from Knoope et al., 2013) using Equation (3.19) and (3.20), where, m is the average velocity; μ is viscosity of the fluid; Re is the Reynold's number, ε is pipe roughness taken as 0.0457mm (NETL, 2018) and f_D is fanning's friction factor. f_D and Re were calculated iteratively using a tolerance factor of 0.001 using Equation (3.19) and (3.20) on Microsoft Excel.

$$Re = \frac{4 * m}{\mu * \pi * D_i} \quad (3.19)$$

$$\frac{1}{\sqrt{f_D}} = -2 * \log_{10}\left(\frac{\varepsilon/D}{3.7} + \frac{2.51}{Re\sqrt{f_D}}\right) \quad (3.20)$$

The cost model developed by Rui et al. (2011) for natural gas pipelines is used for estimation of the cost of each candidate arc. It is assumed that the cost of natural gas pipelines is remarkably similar to the cost of CO_2 pipelines (NETL, 2018). This cost model is also cited by NETL (2018) as an ideal pipeline cost analysis. Rui's model is dependent on both pipeline length and diameter unlike other pipeline models. Rui et al (2011) used regression on the NETL natural gas pipeline database for a time period from 1998 to 2007 to obtain a breakdown of pipeline related costs including the cost for materials, labor, ROW and miscellaneous categories. The general cost equation used by Rui et al (2011) for the pipeline capital cost breakdown is given by Equation (3.21), where $C_{category}$ is the capital cost indexed for materials, labor, ROW and miscellaneous categories; L is the length of pipeline in feet obtained from pipeline analysis in the previous section; SA is cross-sectional diameter of the pipeline in feet^2; and a_{i0}, $a_{i-reg},$ a_{i1} and a_{i2} are geographic and categorical constants provided in Table 3.6.

$$C_{category} = e^{(a_{i0}-a_{i-reg})} * L^{a1} * SA^{a2} \quad (3.21)$$

The capital costs for materials, labor, ROW and miscellaneous charges were calculated for all 4 pipeline diameters for every candidate pipeline in the study. The cost indices provided in Table 3.6 were for 2011. To correct the costs for 2019 prices, the Upstream Capital Cost Index (Information Handling Services [IHS Markit], 2019) is used for material's cost, the GDP chain type price index (US Bureau of Economics, 2019) is used for ROW charges and the producer price index (US Bureau of Labor Statistics, 2019) is used for labor and miscellaneous charges. The Upstream Capital Cost Index, the GDP chain type index and the producer price index are documented in Appendix D for 2000 to 2019. An additional cost correction factor is used for the different pipeline diameters as indicated by the Rui's model (Table 3.7) to adjust for diameter dependent cost variations.

It is essential to compute, the annual operating costs for a pipeline too. Operating and maintenance charges are estimated to be $5000/mile (1mile=1.6km) according to NETL (2018) in 2011 prices. The operating costs are corrected using the Upstream Operating Capital Index (IHS Markit, 2019). The operating prices are assumed to be independent of pipeline diameters. The Upstream Operating Capital Index is documented in Appendix D for 2000 to 2019.

Table 3.6: Cost Indices for Equation (3.21) (Rui et al.2011).

Parameter	Materials	Labor	ROW	Miscellaneous
a_{i0}	4.814	5.697	1.259	5.58
$a_{i\text{-}NE}$	0	0.784	0.645	0.704
$a_{i\text{-}SE}$	0.176	0.772	0.798	0.967
$a_{i\text{-}MW}$	-0.098	0.541	1.064	0.547
$a_{i\text{-}CEN}$	0	0	0	0
$a_{i\text{-}SW}$	0	0.498	0.981	0.699
$a_{i\text{-}west}$	0	0.653	0.778	0
$a_{i\text{-}CAN}$	-0.196	0	-0.83	0
a_{i1}	0.873	0.808	1.027	0.765
a_{i2}	0.734	0.459	0.191	0.458

Table 3.7: Cost Correction Factor dependent on pipeline diameter (NETL, 2018).

D_o	(in)	12	16	24	32
eCO_2	(fraction)	1	1.12	1.18	1.25

This thesis also accounts for costs related to booster stations for ensuring constant pressure of the CO_2 stream. A fixed cost of $83,851 and a variable cost of $1325 per booster station is provided by NETL (2018). The booster station capital cost is also corrected with the Upstream Capital Cost Index and the variable cost is adjusted using the Upstream Operating Cost Index for 2019 (IHS Markit, 2019). A booster station is estimated to be located per 80.46 km of pipeline length (NETL, 2018).

For the existent pipeline no extra cost in terms of capital cost is added, however, the techno-economic analysis needs to be conducted irrespectively. This is since the operating and maintenance costs related to these pipelines and booster stations must be included to the overall costs of the CCUS network and considered in decision analysis.

3.6.3 Storage Cost Model

CO_2 storage nodes include saline aquifers and CO_2 EOR sites. These storage sites are characterized by various reservoir properties such as thickness, permeability, porosity, depth, reservoir pressure (initial) and target formation. Injection rates for individual wells in each reservoir was calculated using the injectivity formula, Equation (3.22), obtained from Law et al. (1997), where Q is rate of injection; k_h and k_v are the horizontal and vertical permeability ($k_h$is assumed to be 80% of k_v); D is the depth of reservoir; ΔP is the pressure difference in reservoir; and μ is viscosity of fluid (CO_2). Here the reservoir pressure is taken to be represented by the hydrostatic gradient and the CO_2 viscosity is same as the fluid stream running through the pipeline system.

$$Q = 0.0208 * (k_h * \ k_v)^{0.5} * D * \frac{\Delta P}{\mu} \qquad (3.22)$$

The cost model used for defining the cost of preparing and operating a reservoir is obtained from Rubin et al. (2008). The cost model for CO_2 EOR is different than the model used for saline aquifers but both models are dependent on number of wells. However, CO_2 EOR is also dependent on number of existent wells as opposed to number of new wells. Here we have taken the assumption that 60% of wells are reworked for the purposes of CO_2 EOR while the rest are new wells (Azzolina et al., 2015).

The model used by Rubin et al. (2008) is an exponential correlation developed by regression analysis of various capital and operating costs obtained from upstream oil and gas operators in different regions in the USA from 1998 to 2004. The model has individual exponential regression correlations for components related to EOR activity including cost for leasing equipment, cost for production equipment, cost of injection equipment, drilling and completion costs and operation and maintenance costs. Of these categories, only operation and maintenance costs are a recurring annual charge. Equation (3.23) shows the correlation for cost estimation of various components related to CO_2 EOR cost model , where C_{type} indicates the cost of the 5 different factors related to setting up new wells in a reservoir mentioned above; D is the average depth of the reservoir; and a_1, a_2 are geographic constants. Table 3.8 is a summary of the geographic constants for the sates related to the study area which have been used in cost estimation related to Equation (3.23). Equation (3.24) indicates the capital costs for reworking existing wells, where C_{wo} is the workover cost for a well, C_{DC} is the cost for drilling and completion of the well and C_{PE} is the cost for production equipment:

$$C_{type} = a_1 * e^{a_2 * D} \tag{3.23}$$

$$C_{WO} = 0.48 * C_{DC} + 0.5 * C_{PE} \tag{3.24}$$

Table 3.8: Geographic Constants for CO_2 EOR Cost Estimation in Equation (3.23) (Rubin et al., 2008).

States	**Lease Equipment**		**Production Equipment**		**Injection Equipment**		**Drilling and Completion**		**Operations and Maintenance**	
	a_1	a_2	a_1	a_2	a_1	a_2	a_1	a_2	a_1	a_2
UT	50362	0.00003	48328	0.00011	185819	0.00032	28577	0.00011	28577	0.00011
CO	34774	0.00003	31130	0.00015	74808	0.00028	26878	0.00011	26878	0.00011
WY, MT, ND	34774	0.00003	31130	0.00015	74808	0.00028	26878	0.00011	26878	0.00011

The cost model for the design of saline aquifer is defined by Rubin et al. (2008) and is a well based cost analysis. The assumption for saline aquifers is that all wells are new, and no well is reworked for the purposes of CO_2 injection. The model defines a reservoir site characterization charge of \$100,000/m^2 (USD 2008) and test well charge of \$3,000,000/25m^2 (USD 2008). The model also estimated a data processing charge for the field at 30% of the sum of the site characterization and test well costs. The charges assumed by the Rubin et al. (2008) uses three categories: cost of injection equipment, drilling and completion cost and cost for operation and maintenance. Of these categories, only operation and maintenance costs are a recurring annual charge. The charges related to saline aquifer development is in the form of an exponential equation shown by Equation (3.23). Table 3.9 is a summary of the geographic constants in Equation (3.23) related to the study area for saline aquifer permanent CO_2 storage. The capital cost for saline aquifer involves all outlined costs except for charges related to the operations and maintenance of the well.

Table 3.9: Geographic Constants for CO_2 geologic storage in Saline Aquifers (Rubin et al., 2008)

Lease Equipment		**Injection Equipment**		**Operations and Maintenance**	
a_1	a_2	a_1	a_2	a_1	a_2
80086	0.0003	29611	0.00008	32893	0.00009

The outlined cost model for both CO_2 EOR and saline aquifers need to be adjusted for inflation using the Upstream Capital Cost Index for capital costs and the Operation and Maintenance costs are adjusted using the Upstream Operating Cost Index (IHS Markit, 2019).

3.7 Optimization Tool

The steps discussed before in this chapter, deal with preparation of data and procedures to prepare a CCUS network capable of reflecting the probable infrastructure in the study area. This section deals with the formulation of an optimization tool which takes into consideration the economic and technical parameters related to various candidate arcs, source nodes and sinks nodes and establishes a common ground for decision-making. The aim of this step is to weigh all these factors together to choose the best combination of pipelines, sources and sinks to generate the most economically feasible CCUS infrastructure network amongst the options, given a set of defined constraints. The defined constraints are based on various material balance and flow parameters which make the network accountable for inconsistencies. The optimization tool answers the following questions: (a) which type of CO_2 sources should be included in CCUS infrastructure? (b) which are the most cost-effective CO_2 sinks available for storage? (c) Which are the best pipeline routes to transport CO_2? (d) how much CO_2 can be captured/stored safely in a given period of time economically? (e) how much CO_2 needs to be captured at each source and stored at each sink? and (f) what should be the diameter of each pipeline?

This section has two parts, namely, static and dynamic time period decision analysis. In static time period analysis, an infrastructure system is generated that will ensure continued unhindered deliverance of CO_2 from source to sink for a given period of time. Dynamic decision analysis optimizes routes through the passage of time for varied rates of CO_2 capture. The optimization tool in both cases relies on MILP formulations to support decisions. The mixed integer tool varies from linear tools as it combines binary variables with regular float point variables, enabling the representation of choice in the system. This tool avoids the usage of non-linear optimization for reduced time complexity for problem solving. The modeling is done on the General Algebraic Modeling Software (GAMS) platform supported by the IBM CPLEX linear solver.

3.7.1 Static Decision-Making Module

The static decision module of the optimization tool focuses on the economic optimization of a set of pipeline candidates, sources and sinks to identify optimal capture rates and injections rates along with the best sites to set-up CCUS infrastructure for a user-defined CO_2 annual capture goal. The analysis takes place with four different core components in the system: source nodes, sink nodes, pipeline arcs and nodes which are both sources and sinks (dual nodes). The detailed formulation implemented on GAMS is provided in Appendix C.

Objective Function:

Minimize

$$Cost = \sum_i a_i * Ccap_i * X_i + \sum_p c_p * Ccapb_p * Xb_p + \quad (3.25a)$$

$$\frac{int(1+int)^{time}}{(1+int)^{time}-1} * \sum_{arc}\sum_t Ctran_{arc,t} * b_{arc,t} + \sum_{arc}\sum_t Otran_{arc,t} * b_{arc,t} + \quad (3.25b)$$

$$\frac{int(1+int)^{time}}{(1+int)^{time}-1} * \sum_j \left(r_j * Cres_j + h_j * Cwell_j\right) + \sum_j \begin{pmatrix} h_j * Owell_j + Y_j * MVA_j * MVAC \\ +Y_j * l_j * costCO_2 \end{pmatrix} +$$

$$\frac{int(1+int)^{time}}{(1+int)^{time}-1} * \sum_j \left(c_p * Cresb_p + hb_p * Cwellb_p\right) +$$

$$\sum_j \begin{pmatrix} hb_p * Owellb_p + Yb_p * MVAb_p * MVAC \\ +Yb_p * lb_p * costCO_2 \end{pmatrix} \quad (3.25c)$$

Constraints:

$$Q_{arc,t} \geq -Qmx_t * b_{arc,t} \qquad \forall arc \in arc;\ \forall t \in T \quad (3.26)$$

$$Q_{arc,t} \leq Qmx_t * b_{arc,t} \qquad \forall arc \in arc;\ \forall t \in T \quad (3.27)$$

$$\sum_t b_t \leq 1 \qquad \forall arc \in arc \quad (3.28)$$

$$X_i \leq a_i * Xmx_i \qquad \forall i \in I \quad (3.29)$$

$$Xb_p \leq c_p * Xmxb_p \qquad \forall p \in P \quad (3.30)$$

$$Y_j \leq r_j * Yresmax_j / time \qquad \forall j \in J \quad (3.31)$$

$$Yb_p \leq c_p * Yresmaxb_p / time \qquad \forall p \in P \quad (3.32)$$

$$h_j * Ywellmx_j = Y_j \qquad \forall j \in J \quad (3.33)$$

$$hb_p * Ywellmxb_p = Yb_p \qquad \forall p \in P \quad (3.34)$$

$$\sum_i X_i + \sum_p Xb_p \geq CapCO_2 \quad (3.35)$$

$$\sum_{arc}\sum_t Q_{arc(m,n),t} - \sum_{arc}\sum_t Q_{arc(n,m),t} = X_i \qquad \forall i \in I \quad (3.36)$$

$$\sum_{arc}\sum_t Q_{arc(m,n),t} - \sum_{arc}\sum_t Q_{arc(n,m),t} = -Y_j \qquad \forall j \in J \quad (3.37)$$

$$\sum_{arc}\sum_t Q_{arc(m,n),t} - \sum_{arc}\sum_t Q_{arc(n,m),t} = Xb_p - Yb_p \qquad \forall p \in P \quad (3.38)$$

Bounds:

$$a_i, b_{arc,t}, r_j, c_p \in 0,1 \tag{3.39}$$

$$X_i, Y_j, Xb_p, Yb_p \ \geq 0 \tag{3.40}$$

$$h_j, hb_p \in 0,1,2,3,4,5, \dots \dots .., n \tag{3.41}$$

***Decision Variables**:*

a_i	Capture node binary decision variable
$b_{arc,t}$	Transport arc binary decision variable
r_j	Sink node binary Decision Variable
X_i	Amount captured at each source node in Mton of CO_2
Y_j	Amount stored at reservoir node annually in Mton of CO_2
h_j	Number of wells in sink nodes
$Q_{m,n,t}$	Amount transported through pipe in Mton of CO_2
$cost$	Cost of CCUS network
c_p	Dual node binary decision variable
Xb_p	Amount captured at dual node in Mton of CO_2
Yb_p	Amount stored at capture and sink annually in Mton of CO_2
hb_p	Number of wells

Sets:

N, I, J, P	Set of all nodes, Set of source nodes, Set of sink nodes, set of dual nodes
T	Set of all pipeline diameter configurations
Arc	Set of all probable pipeline arcs

Input Variables:

$Ccap_i$, $Ccapb_p$	Cost for capturing 1MtCO_2 at source nodes (cap) and dual nodes (capb)
$Ctran_{arc,t}$, $Cres_j$, $Cresb_p$, $Cwell_j$, $Cwellb_p$	Capital cost for opening pipeline (tran), reservoirs (res), dual node reservoirs (resb), wells in reservoir nodes (well) and wells in dual nodes (wellb)
$Otran_{arc,t}$, $Owell_j$, $Owellb_p$	Annual operating cost for pipelines (tran), wells in reservoir nodes (well) and wells in dual nodes (wellb)
l_j, lb_p	Recycle ratio for sink nodes (l) and dual nodes (lb)
MVA_j, $MVAb_p$	Reservoir type indicator (0-EOR site or 1-Saline aquifer)
Qmx_t	Maximum flow rate through pipes in MtCO_2/yr
Xmx_i, $Xmxb_p$	Upper limit to annual capture quantity for source nodes (x) and dual nodes (xb)

$Ywellmx_j$, $Ywellmxb_p$	Upper limit to annual well injection quantity for source nodes (x) and dual nodes (xb)
$Yresmax_j$, $Yresmaxb_p$	Upper limit to reservoir quantity for source nodes (x) and dual nodes (xb)
$CostCO_2$	Cost of purchasing 1 MtCO_2 for EOR purposes
$Time$	Time period of operations
int	Inflation rate

The objective function aims to minimize the overall annualized cost of the infrastructure model over the time period of operations. The objective function depicted by Equation (3.25) consists of three parts namely, the cost of capture Equation (3.25a); the cost related to transportation of CO_2 through pipelines Equation (3.25b); and the cost related to injection of CO_2 in a geologic reservoirs Equation (3.25c). Equation (3.24a) depicts the cost of capturing of CO_2 annually at various capture nodes as a function of the total probable capture capacity of each individual node. Binary decision variable a_i and c_p control the usage of each source node. It must be noted that the two terms involved in Equation (3.25a) refer to capture costs related to nodes which only capture CO_2 and dual nodes (terms suffixed with -b). The terms do not need an annualization function as the capture quantity is dependent on annual capture amounts. Equation (3.25b) deals with the annualized cost related to deployment of selected candidate pipeline arcs of specific diameter. The first term in Equation (3.25b) refers to the capital cost of constructing a pipeline which is a function of binary control variable $b_{arc,t}$ and is annualized over the time-period of the project. The second term of Equation (3.25b) refers to the annual operational costs related to maintenance of the pipeline and is also a function of the binary control variable $b_{arc,t}$. Equation (3.25b) is iterated over all possible candidate arcs for all pipeline diameters. Equation (3.25c) calculates the annualized capital cost related to opening each sink node by taking into consideration the capital costs at reservoir level as well as the cost related to wells. Equation

(3.25c) has four terms: the first term deals with the annualized capital costs of opening the reservoir and wells; the second terms is the annual operating cost for operating wells in the open reservoirs; the third term is the annualized capital costs for opening dual node reservoirs; and the fourth term is the annual operating costs for running wells in an open dual node reservoir. The reservoir level charges are controlled by binary decision variable r_j and c_p, which determines the usage of specific reservoirs while well costs are determined by the integer decision variable h_j and hb_p which determines the number of wells if a reservoir is open. Further, the variable costs related to deployment of each reservoir is dependent on the number of wells used as well as the type of reservoir as it determines the cost for monitoring, verification and abandonment (MVA) as well as the cost of purchasing CO_2 for EOR projects. The charge for MVA is taken as 5cents/ton of CO_2 injected (can be varied) (Rubin et al., 2008) and the cost of CO_2 is an input variable.

Equation (3.26) and (3.27) are flow constraints related to the annual flow of CO_2 in $MtCO_2$/yr through each pipeline arc of a specific diameter. The pipeline arcs utilized in this paper are bi-directional in nature and thus flow rates generated in the study could be negative in value implying flow of CO_2 in the opposite direction. The control of flow rate through each pipeline is done by the objective function, as the lower diameter pipes have lower costs. Conversely, it is possible to use unidirectional pipeline arcs too, however, the related decision variables double in quantity, thus increasing the computational complexity of the formulation. The maximum flow rate through the pipeline is obtained from the cost model for CO_2 pipelines. Equation (3.28) is the constraint which ensures that only a single pipeline of a specific diameter can be built on a given pipeline route. By removing this constraint, it is possible to build multiple pipelines (different diameters) along the same route. Construction of multiple pipelines through the same arc has been analyzed by a few authors like Jensen et al. (2013).

Equation (3.29) and (3.30) ensure that the amount of CO_2 captured annually at each source is below the specified maximum capture capability of the source which is determined by the cost model for capture nodes, where Equation (3.29) is exclusively for source nodes, while Equation (3.30) is for dual nodes. The constraint ensuring that the total injected CO_2 in each reservoir is below the total capacity of the reservoir during the entire analysis period of the case study is provided by Equation (3.31) and (3.32), where Equation (3.31) is for reservoir nodes and Equation (3.32) is for dual nodes. Similarly, well specific injection constraints are provided by Equation (3.33) and (3.34), where Equation (3.33) is for sink nodes and Equation (3.34) is for dual nodes. The annual capture and storage goals for the infrastructure network and is a critical input parameter is determined by Equation (3.35). Equation (3.36), (3.37) and (3.38) are material balance constraints and ensure that all CO_2 captured in the network find a destination and no CO_2 in the system is left unaccounted. Equation (3.39), (3.40) and (3.41) are bounds for the decision variable used in the MILP formulations.

It must also be noted that this formulation has multiple extra constraints to accommodate for nodes that are dual nodes. In the case, where nodes can only be either a sink or a source, the objective function will have lower amount of constraints and Equation (3.30), (3.32), (3.34) and (3.38) would be eliminated from the formulation.

In the static formulation, if existent pipelines were to be analyzed for checking their effect on the CCUS network, then the list of connecting nodes need to be considered in the MILP formulation. There will be fixed usage of certain sources and sinks nodes along with the fixed usage of the existent pipeline arcs (details will be provided in Section 4.5). This is done by setting a minimum capacity on each respective node and pipeline arc. Further, a new set of node points must be included in the formulation called "Connector Nodes" represented by set K. The role of set K elements is just to serve as connections between existent and new infrastructure and thus is avoided in analysis related to the objective function. Yet, the byproducts of the connector nodes in terms of pipeline arcs still play a crucial role in network analysis. Further, to maintain fluidity in the network, an additional constraint must be added to the formulation to ensure that no CO_2 is either stored or captured at these connector nodes given by Equation (3.42).

$$\sum_{arc}\sum_t Q_{arc(m,n),t} - \sum_{arc}\sum_t Q_{arc(n,m),t} = 0 \qquad \forall k \in K \qquad (3.42)$$

3.7.2 Dynamic Decision-Making Module

The dynamic decision module of the optimization tool focuses on the economic optimization of a set of pipeline candidates, sources and sinks to identify optimal capture and injections rates for varying annual CO_2 capture goals such that the infrastructure set-up is gradual through the time periods. The defined CO_2 capture goals is defined for individual sub time periods (of equal duration) in the problem. In the usual step-wise optimization, each pipeline is gradually increased in diameter, however, in this thesis the diameters are optimized for the maximum flow rate when it is set-up and not reworked to increase the pipeline diameter at a later time. The formulation is similar to the static decision-making module except for the dependency of constraints on time period of operations. Another important consideration in dynamic optimization is the time duration ($time$) and the length of each time period ($time_i$). The number of time durations/periods (d_n) is given by $d_n = \frac{time}{time_i}$. The detailed formulation implementation on GAMS is in Appendix C.

Objective Function:

Minimize

$$Cost = \frac{int(1+int)^{time}}{(1+int)^{time}-1} \left(\sum_d \frac{(1+int)^{time_i-1}}{int(1+int)^{time_i}} * \frac{1}{(1+int)^{time-n*time_i}} \left(\sum_i a_i * Ccap_i * X_i + \right.\right.$$
$$\left. + \sum_p c_p * Ccapb_p * Xb_p\right) + \sum_d \frac{1}{(1+int)^{time-n*time_i}} \left(\sum_{arc}\sum_t Ctran_{arc,t} * b_{arc,t} + \right.$$
$$\frac{(1+int)^{time_i-1}}{int(1+int)^{time_i}} *$$

$$\sum_{arc}\sum_t Otran_{arc,t} * \sum_{d-1}^{0} b_{arc,t}) + \sum_d \frac{1}{(1+int)^{time-n*time_i}} * (\sum_j \big(r_j * Cres_j + h_j * Cwell_j\big) +$$

$$+ \frac{(1+int)^{time_i-1}}{int(1+int)^{time_i}} * \sum_j \begin{pmatrix} h_j * Owell_j + Y_j * MVA_j * MVAC \\ +Y_j * l_j * costCO_2 \end{pmatrix} + \sum_j * \big(c_p * Cresb_p + hb_p *$$

$$Cwellb_p\big) + \frac{(1+int)^{time_i-1}}{int(1+int)^{time_i}} \sum_j \begin{matrix} (hb_p * Owellb_p + Yb_p * MVAb_p * MVAC \\ +Yb_p * lb_p * costCO_2) \end{matrix})) \quad (3.45)$$

Constraints:

$$Q_{arc,t,d} \geq -Qmx_t * \sum_{d-1}^{0} b_{arc,t} \qquad \forall arc \in arc;\ \forall t \in T;\ \forall d \in D \qquad (3.44)$$

$$Q_{arc,t,d} \leq Qmx_t * \sum_{d-1}^{0} b_{arc,t} \qquad \forall arc \in arc;\ \forall t \in T;\ \forall d \in D \qquad (3.45)$$

$$\sum_d \sum_t b_{arc,t,d} \leq 1 \qquad \forall arc \in arc \qquad (3.46)$$

$$X_{i,d} \leq a_{i,d} * Xmx_i \qquad \forall i \in I \qquad (3.47)$$

$$X_{i,d=2,3,\ldots,n} \geq Xmx_{i,d=1,2,\ldots,n} \qquad \forall i \in I \qquad (3.48)$$

$$Xb_{p,d} \leq c_{p,d} * Xmxb_p \qquad \forall p \in P \qquad (3.49)$$

$$Xb_{p,d=2,3,\ldots,n} \geq Xbmx_{p,d=1,2,\ldots,n} \qquad \forall p \in P \qquad (3.50)$$

$$Y_{j,d} \leq r_j * Yresmax_j / time \qquad \forall j \in J;\ \forall d \in D \qquad (3.51)$$

$$Yb_{j,p} \leq c_p * Yresmaxb_p / time \qquad \forall p \in P;\ \forall d \in D \qquad (3.52)$$

$$\sum_d Y_{j,d} \leq Yresmax_j \qquad \forall j \in J \qquad (3.53)$$

$$\sum_p Y_{p,d} \leq Yresmax_p \qquad \forall p \in P \qquad (3.54)$$

$$h_{j,d} * Ywellmx_j = Y_{j,d} \qquad \forall j \in J \qquad (3.55)$$

$$hb_{p,d} * Ywellmxb_p = Yb_{p.d} \qquad \forall p \in P \qquad (3.56)$$

$$\sum_i X_{i,d} + \sum_p Xb_{p,d} \geq (CapCO_2)_d \qquad \forall d \in D \qquad (3.57)$$

$$\sum_{arc}\sum_t Q_{arc(m,n),t,d} - \sum_{arc}\sum_t Q_{arc(n,m),t,d} = X_{i,d} \qquad \forall i \in I;\ \forall d \in D \qquad (3.58)$$

$$\sum_{arc}\sum_t Q_{arc(m,n),t,d} - \sum_{arc}\sum_t Q_{arc(n,m),t,d} = -Y_{j,d} \qquad \forall j \in J;\ \forall d \in D \qquad (3.59)$$

$$\sum_{arc}\sum_t Q_{arc(m,n),t,d} - \sum_{arc}\sum_t Q_{arc(n,m),t,d} = Xb_{p,d} - Yb_{p,d} \qquad \forall p \in P;\ \forall d \in D \qquad (3.60)$$

Bounds:

$$a_i, b_{arc,t}, r_j, c_p \in 0,1 \qquad (3.61)$$

$$X_i, Y_j, Xb_p, Yb_p \geq 0 \qquad (3.62)$$

$$h_j, hb_p \in 0,1,2,3,4,5, \dots \dots .., n \tag{3.63}$$

Decision Variables:

$a_{i,d}$	Capture node binary decision variable
$b_{arc,t,d}$	Transport arc binary decision variable
$r_{j,d}$	Sink node binary decision variable
$X_{i,d}$	Amount captured at each source node in Mton of CO_2
$Y_{j,d}$	Amount stored at reservoir node annually in Mton of CO_2
$h_{j,d}$	Number of wells in sink nodes
$Q_{m,n,t,d}$	Amount transported through pipe in Mton of CO_2
$cost$	Cost of CCUS network
$c_{p,d}$	Dual node binary decision variable
$Xb_{p,d}$	Amount captured at dual node in Mton of CO_2
$Yb_{p,d}$	Amount stored at capture and sink annually in Mton of CO_2
$hb_{p,d}$	Number of wells

Sets:

N, I, J, P	Set of all nodes, Set of source nodes, Set of sink nodes, Set of dual nodes
T	Set of all pipeline diameter configurations
D	Set of all time periods
Arc	Set of all probable pipeline arcs

Input Variables:

$Ccap_i$, $Ccapb_p$	Cost for capturing 1MtCO_2 at source nodes (cap) and dual nodes (capb)
$Ctran_{arc,t}$, $Cres_j$, $Cresb_p$, $Cwell_j$, $Cwellb_p$	Capital Cost for opening pipeline (tran), reservoirs (res), dual node reservoirs (resb), wells in reservoir nodes (well) and wells in dual nodes (wellb)
$Otran_{arc,t}$, $Owell_j$, $Owellb_p$	Annual operating cost for pipelines (tran), wells in reservoir nodes (well) and wells in dual nodes (wellb)
l_j, lb_p	Recycle ratio for sink nodes (l) and dual nodes (lb)
MVA_j, $MVAb_p$	Reservoir type indicator (0-EOR site or 1-Saline aquifer)
Qmx_t	Maximum flow rate through pipes in MtCO_2/yr
Xmx_i, $Xmxb_p$	Upper limit to annual capture quantity for source nodes (x) and dual nodes (xb)
$Ywellmx_j$, $Ywellmxb_p$	Upper limit to annual well injection quantity for source nodes (x) and dual nodes (xb)
$Yresmax_j$, $Yresmaxb_p$	Upper limit to reservoir quantity for source nodes (x) and dual nodes (xb)
$CostCO_2$	Cost of purchasing 1 MtCO_2 for EOR purposes

$(CapCO_2)_d$	Annual Capture mgoal for each time period
$time$	Time Period of Operations
$time_i$	Duration of each individual period
int	Inflation Rate

The formulation of the dynamic optimization module is very similar to the static formulation and hence, only the difference between the two formulations will be discussed in greater details. The objective function is provided by Equation (3.43) and accounts for both capital and variable costs related to CO_2 capture, transport and storage. The objective function still minimizes the annualized costs related to the CCUS network, however, in this case straight implantation of annualized cost is not possible. In each part of the CCUS value chain, all capital costs and the total operating costs for the entire $time_i$ needs to be calculated and brought to year 2019 using discounting. Once the overall cumulative cost for the project for period $time$ is calculated, the annualized cost for the operating period is calculated. Each individual term is only valid for the specified time duration and the cost accumulated over each of these time durations are summarized separately. Here the optimization of economic cost is done over the entire duration of operation and hence, even though the infrastructure construction in one period is not economically optimal, when all time periods are taken together, the decision is more justifiable. This is the main difference between the objective functions of the two methods.

Equation (3.44) and (3.45) are flow constraints ensuring that pipeline arcs in every time period are within the capacity of the selected pipe configurations. Equation (3.46) makes sure that only one pipe of a certain configuration is built along a route. However, the dynamic module also makes sure that pipe is built only once during the entire duration of the operations. Equation (3.48) and (3.50) are new constraints which ensure continued supply of CO_2 from sources in certain volumes equal or above supply of previous period. The accounting to ensure minimum annual CO_2 capture is represented by Equation (3.57), where each period has a different capture goal. In the case where all capture goals are the same, the resultant solution would be like the static formulation. Equation (3.46), (3.47) and Equation (3.51) to (3.63) are similar to the constraints in the static formulation but modified to ensure the consistency of the overall MILP operation in each individual time periods.

For checking the effect of existent pipelines and infrastructure on the overall dynamic CCUS network, connector nodes need to be considered in the formulation using Equation (3.61). Further, certain lower bounds need to be provided for capture and storage nodes to necessitate the usage of current nodes used in the networks. Similar bounds would be provided to pipeline arcs representing existent pipelines. These limits were discussed in Section 3.5.

$$\sum_{arc} \sum_{t} Q_{arc(m,n),t,d} - \sum_{arc} \sum_{t} Q_{arc(n,m),t,d} = 0 \qquad \forall k \in K \qquad (3.61)$$

3.8 Text Mining of Regulations

Regulatory compliance is an aspect often left out of initial decision analysis due to the complexity of source material, lack of domain knowledge, and general lack of connective media between regulations and decision analysis. Regulations can disrupt initial ideas significantly by generating unforeseen barriers towards the development of infrastructure. However, regulations are important as they protect the health and safety of the general populous and environment over capital interests. In this work, federal regulations will be analyzed to extract clauses related to the placement, construction and maintenance of pipelines.

In order to extract regulations in a format suitable for decision analysis, it is necessary to process and classify regulations. Text Mining analysis of regulations have the ability to provide the crucial link between technical and economic analysis (Conrad et al., 2018). This works analyzes USA CFR title 49 Section 195 for the purpose of extracting regulations. Tile 49 of the regulations is named “Transportation-Volume 1”, while Section is “Transportation of Hazardous Liquids by Pipelines” (USA CFR Title 49, 2019). Section 195 covers any pipeline related to transportation of CO_2 (inter-state and intra-state) including in-land pipelines, offshore pipelines (including those in the outer continental shelf of USA) and gathering pipelines. USA CFR title 49 Section 195 is available publicly in the XML data format.

Text mining is often used to analyze the importance of a concept to a document or to measure the frequency of terms in an internet search. However, the literature related to text analysis for regulations is very scant. Most work related to regulatory compliance-based text mining is in the area of medical regulations, software compliance laws and building construction regulations (Zhang et al., 2016; Breaux et al., 2006; Hjelseth et al., 2011; Song et al., 2018; Zeni et al., 2015; Conrad et al., 2018; Lim et al., 2011). Taking inspiration from previous work, the workflow for regulatory text mining will follow the procedures utilized in building construction management. Natural language processing (NLP) is a sub-category of text mining. Usually the analysis of text will follow either semantic or syntactic analysis. Semantic analysis in NLP deals with a document in terms of logic and statistics. An example of semantic analysis of a text would be finding the importance of a given word in relation to a document by measuring its frequency. Syntactic analysis in NLP refers to analysis of the structure of a sentence in terms of grammar (Breaux et al., 2006). An example of syntactic analysis is the breakdown of words in a sentence to understand their role like nouns, adverbs, verbs, and pronouns. In this work, we utilize the syntactic analysis to analyze the regulations relevant to CO_2 pipelines due to the following factors (Zeni et al., 2015; Hjelseth et al., 2011; Zhang et al., 2016):

- The unstructured and heterogenous nature of regulations, where keywords are uniformly dispersed through the body of the document (example: word "construction" is used uniformly throughout the text, even though a clause may not pertain to the concept).
- Identification of constraints and relationship terms (example: less than, more than) from the text is not possible without syntactic analysis of a statement.
- Identification of exceptions and conjectures to regulatory clauses require syntactic analysis.

The workflow related to text mining is shown in Figure 3.6. The text mining workflow consist of 5 main parts: XML parsing where the raw text data is categorically extracted for text analysis; the text-preparation stage uses the extracted text and converts it to an actionable format; the text processing phase divides the actionable text to derive structured information for the text; the information representation stage involves manual tagging of the structured information; the information extraction phase provides output in the form of structured information. The entire execution of the text mining workflow is done on Python 3.6 in coordination with a spreadsheet tool (Microsoft Excel). An example of the python code relating to execution of the individual text mining of clauses is provided in Appendix B.

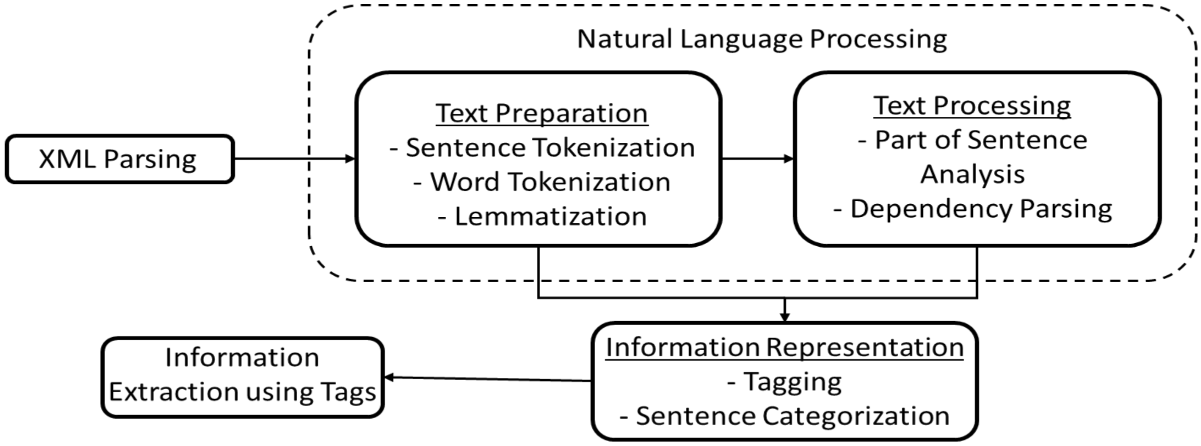


Figure 3.6: Text mining workflow for regulatory information.

3.8.1 XML Parsing

The raw data obtained from USA CFR Title 49 Section 195 is in the XML data format. The data format is a specially annotated form of text used for input into database systems. The specialty of the data format is the hierarchical format storage of data. However, for the purpose of textual mining such data needs to be extracted in a plain text without annotations or hierarchy. This is done in the workflow using a code on Python 3.7 and an in-built package named *xml.etree.elementree* which converts the input data into a tree structure. In this data tree, each entry is part of a clause, and the clauses are distributed unevenly in levels.

The output of this xml tree is a csv format data spreadsheet containing all the regulatory clauses as rows and classified by their relevant clause number. For instance, clause §195.3.1.1 becomes section 195 sub-section 3 sub-section(2) 1 Subsection(3) 1 subsection(4) 0 which acts as a unique classifier each clause. The code for conversion of the raw xml data to an actionable data format is provided in Appendix B.

3.8.2 Text Preparation

This step of the workflow deals with converting the plain text obtained from the XML parsing into a machine readable and actionable format. To properly annotate and classify parts of sentences, it is necessary to break down the text into its minimal form, which is achieved in this section. There are three basic steps: Sentence tokenization which breaks a text into its core sentences; Word tokenization which is necessary to split each word into disconnected entities; and Stemming which revert each word back to its root word (with no grammatical tense). Generally, in many other pre-processing NLP workflows, an additional step of removing *stop words* (example: and; is; the; a; an) is used (Zhang et al., 2016; Song et al., 2018; Lin et al., 2011). However, in this work stop words are required for retaining grammatical meaning to the text.

Sentence Tokenization: Sentence tokenization is the process of splitting a body of text into its respective constitutive sentences (Zhang et al., 2016; Hjelseth et al., 2011; Song et al., 2018; Zeni et al., 2015; Lim et al., 2011). The sentences are recognized on basis of normal sentence boundaries such as punctuation marks including periods (.), exclamation marks (!), question marks (?) and hyphens. The topic is more complicated as it needs to recognize symbols as natural boundaries and not as part of numeric values or words, for example, the period (.) in 3.142 is part of a numeric value and not an end of a sentence. These cases make sentence tokenization a special extension of NLP and the code needs to be pre-trained on several examples before execution. Sentence Tokenization in this work is executed using the python library "nltk" with sub package "sent_tokenize" which is pre-trained by the developers on extensive databases. An example of sentence tokenization is given below with its derivative results:

Original text: "An operator may make arrangements with another person for the performance of any action required by this part. However, the operator is not thereby relieved from the responsibility for compliance with any requirement of this part."

Derivative Tokenized Sentences:

- An operator may make arrangements with another person for the performance of any action required by this part.
- However, the operator is not thereby relieved from the responsibility for compliance with any requirement of this part.

Word Tokenization: Word tokenization is the process of breaking down each component of a document into their respective constituent words and symbols (Zhang et al., 2016; Hjelseth et al., 2011; Song et al., 2018; Zeni et al., 2015; Lim et al., 2011). Individual words, symbols and numeric values are recognized using blank spaces or punctation marks like periods and exclamation points. Word tokenization much like sentence tokenization is an extended version of NLP as it requires pre-trained knowledge by the classifier to accurately identify the component words, values, and symbols in a sentence. Word tokenization can be carried out with before or after sentence tokenization, however, sentence tokenization must be carried before word tokenization. Word tokenization in this work is carried out using the python library "nltk" with the package "word-tokenize". An example of a tokenized sentence is as follows:

Original text: "However, the operator is not thereby relieved from the responsibility for compliance with any requirement of this part."

Derivative Tokenized Sentence: "However", ",", "the", "operator", "is", "not", "thereby", "relieved", "from", "the", "responsibility", "for", "compliance", "with", "any", "requirement", "of", "this", "part", "."

Stemming: Stemming is the process of deriving the root word for any given tokenized word or symbol (Zhang et al., 2016; Hjelseth et al., 2011; Song et al., 2018; Zeni et al., 2015; Lim et al., 2011). Individual words are broken down to its root component used a pre-trained set and are thus considered an extended part of NLP. For instance, the word "jumping", "jumped" or "jumper" will be reverted to "jump". This step is a pre-step essential for to recognize of the grammatical type of each word. In this work, stemming is performed using the python library "nltk" through the package "SnowballStemmer".

3.8.3 Text Processing

Text processing analyzes the pre-processed text to understand and categorize the text grammatically. This section forms the core syntactic analysis of the regulations. The grammatical analysis of text provides the dependency of one part of the text as compared to another. This section has two parts: Part-of-Speech analysis (POS analysis) where each word is categorized as verbs, nouns, and adverbs; and Dependency parsing where the relationship of each word is provided.

Part of Speech (POS) Analysis: POS tags/analysis assigns tags to each word in a sentence according to their lexical and functional categories in the language of analysis (English). Here *lexical* refers to the meaning of the word according to the vocabulary of the language used, while *functional category* refers to grammatical nature of the word in the sentence. (Zhang et al., 2016; Hjelseth et al., 2011; Zeni et al., 2015). Thus, each word in the statement depending on it's relative role is classified as a verb, noun, adjective, proposition and so on. In the English language words can play different roles in a statement, for instance, the word 'play; can be a noun or verb depending on it's usage. Thus, the POS tagging needs to be pre-trained on a large dataset making it an essential part of NLP. In this work, the POS tagging is done using the StandfordCoreNLP servers through Python, which refers to the pre-training dataset utilized from massive cloud network based in Palo Alto, California, USA. An example of POS tagging of a clause is as follows:
Original text: "However, the operator is not thereby relieved from the responsibility for compliance with any requirement of this part."
Derivative POS Tagged Sentence: "However : adverb", ", : ,", "the : Determiner", "operator : Noun", "is : Verb (third person singular)", "not : adverb", "thereby : adverb", "relieved : Verb (past participle)", "from : Preposition", "the : Determiner", "responsibility : Noun", "for : Preposition", "compliance : Noun", "with : Preposition", "any : Determiner", "requirement : Noun", "of : Preposition", "this : Determiner", "part : Noun", ". : ."

Dependency parsing: Dependency parsing elaborates on the POS tags and stemmed words in order to create phrasal tags (Zhang et al., 2016; Hjelseth et al., 2011). This kind of phrasal tagging is done by using multiple different combinations of POS tag to analyze a relative structure to each part of the sentence. This kind of phrasal tagging helps determine clauses, exceptions and addendums of the regulations. The dependency parsing thus combines a group of words around nouns and verbs to create structure. The dependency parsing in the work is executed using StanfordCoreNLP servers on Python 3.6. The dependency structure of a sentence is derived from pre-training large datasets on multiple documents and forms a core part of syntactic NLP. An example of such depednancy parsing is as follows:

Original text: “However, the operator is not thereby relieved from the responsibility for compliance with any requirement of this part.”

Parsed Statement: (S (ADVP (RB However))(, ,) (NP (DT the) (NN operator)) (VP (VBZ is) (RB not) (ADVP (RB thereby)) (VP (VBN relieved) (PP (IN from) (NP (NP (DT the) (NN responsibility)) (PP (IN for) (NP (NN compliance))))) (PP (IN with) (NP (NP (DT any) (NN requirement)) (PP (IN of) (NP (DT this) (NN part))))))) (. .))

Where, S is Subject; ADVP is adverb phrase; RB is adverb; NP is noun phrase; DT is determiner; VP is verb phrase; VBZ is verb in present tense; VBN is verb in past participle; PP is pronoun phrase; and NN is noun.

3.8.4 Information Representation

It necessary to represent the regulatory clauses in a manner which is suitable towards usage and meaningful representation for CCSHawk. To transform the syntactically processed text into a suitable format, it is essential to classify parts of the overall regulatory clause into 7 parts (or less): header, clause, exception, addendum, condition, consequence and coreference.

To classify the parts of the clause into these categories, the dependency parsing and sentence tokenization of the text is used. Manual classification of the dependencies, sentences and parts of statements need to be carried out for each clause individually to obtain a feasible result.

The classification of statements within clause is managed through a tree structure developed in the dependency parsing. The python library ‘nltk’ is utilized for management of the statement depednancy using the package “element.tree”. The reason for manual classification is due to the limited training instances available for classification and lack of a general unified structure amongst the classes. Another major barrier towards automation of the classification process is the lack of a proper *ontology* for the classification of the text. Here *ontology* refers to a set hierarchy and classification of key terms to determine the distribution of text into relevant classes. For instance, in pipeline technology a hierarchical structure for key activities could include *construction, maintenance, testing, design, reporting, personnel, scope* and *location*. Similarly, an ontology of overarching categories of key elements relevant to a pipeline system could include components such as *pipelines, breakout tanks, valves, flanges, closures* and *welds*. These categories are attached to each other in a hierarchical form such that usage of these terms in the text automatically classifies a statement according to their criteria. The development of an ontology for a text corpus requires rigorous dissection of text and complimentary additional documentation of regulatory clauses, which is out of scope of the work in this thesis.

Tagging of information related to the classified text is another important feature relevant to this thesis as it is required for proper querying of clauses. Each statement needs to be tagged to the relevant main topic they relate to. Further, each clause is also tagged with a subtag to classify the clauses according to specific subcategories. Each clause is provided with only a single tag, however, clauses can have multiple subtags. An example of the classification of a regulatory clause is given below:

Original Clause: § 195.246 Installation of pipe in a ditch.

(a) All pipe installed in a ditch must be installed in a manner that minimizes the introduction of secondary stresses and the possibility of damage to the pipe.

(b) Except for pipe in the Gulf of Mexico and its inlets in waters less than 15 feet deep, all offshore pipe in water at least 12 feet deep (3.7 meters) but not more than 200 feet deep (61 meters) deep as measured from the mean low water must be installed so that the top of the pipe is below the underwater natural bottom (as determined by recognized and generally accepted practices) unless the pipe is supported by stanchions held in place by anchors or heavy concrete coating or protected by an equivalent means.

Represented Text:

Header	Installation of pipe in a ditch.
Tag	Pipeline Construction
Sub Tag	Ditches
Clause	All pipe installed in a ditch must be installed in a manner that minimizes the introduction of secondary stresses and the possibility of damage to the pipe.
Condition	all offshore pipe in water at least 12 feet deep -LRB- 3.7 meters -RRB- but not more than 200 feet deep -LRB- 61 meters -RRB- deep as measured from the mean low water
Condition-Lower Quantity	12
Condition-Lower Quantity Unit	feet
Condition-Lower Quantity Relation	at least
Condition-Upper Quantity	200
Condition-Upper Quantity Unit	feet
Condition-Upper Quantity Relation	not more than
Consequence	must be installed so that the top of the pipe is below the underwater natural bottom -LRB- as determined by recognized and generally accepted practices -RRB- unless the pipe is supported by stanchions held in place by anchors or heavy concrete coating or protected by an equivalent means

Exception	Except for pipe in the Gulf of Mexico and its inlets in waters less than 15 feet deep

3.8.5 Information Extraction

The final step of the text mining workflow is the development of a python code based on querying the classified regulations dependent on tags developed in the previous. The python code requires no specific dependencies as the output of the information representation step is in the form of a spreadsheet. By querying the tags or subtags, it is possible to retrieve relevant regulations. The code related to the querying regulations is provided in Appendix B.

3.9 Conclusion

This chapter provided the methodology related to the preparation and decision-analysis of CCUS infrastructure networks for different CO_2 capture goals along with the process of generating feasible pipeline routes. The methodology has five steps including mapping the study area, preparation of the cost map, pipeline route generation, cost modeling and decision-making analysis. The methodology uses multiple platforms including Microsoft Excel, ArcGIS 10.7.1, Python, R and GAMS. The techniques demonstrated in this chapter answer many critical questions related to the deployment of CCUS such as which sites are best suited for development of infrastructure to meet a regions strategic CO_2 reduction goals, the cost related to deployment of CCUS infrastructure as well as displaying the potential of CO_2 pipelines. The effect of environment, ecology, and existent infrastructure were also investigated on potential future pipeline routes. The final goal of the study is to provide a detailed analysis of the potential of CCUS in the study area, which will be discussed in Chapter 4.

CHAPTER 4

Results: Pipeline Corridor Mapping and Network Analysis

4.1 Introduction

CO_2 emissions in traditional industries and power plants is a point of rising concern amongst experts around the world which led to the implementation of regulations and policies aiming to curb emissions. CCUS, one of the leading means of emission reductions in the industrial and energy sector, is in its infancy and needs feasible strategies for growth. The large capital expenditures and the multi-stakeholder nature of CCUS is a hindrance to its rapid expansion and implementation. Proper decision-making based on existent resources and a systematic economic evaluation of the CCUS value-chain can help with better implementation and faster growth of CCUS.

As was discussed in Chapter 2, a number of key knowledge gaps were identified during the literature review on CCUS decision-making and pipeline routing. These gaps include the lack of analysis on the effect of existent infrastructure, environment, and ecology on the future growth of CCUS networks along with a lack of pipeline route planning in North-Central USA. With the view of developing a decision-making platform to answer some of the shortcomings of previous models, this thesis puts forth a CO_2 pipeline corridor and infrastructure network analysis model, dubbed as CCSHawk. The aim of the research is to generate a feasible and economically sound CCUS infrastructure development strategy for the North-Central region of the USA with special focus on conventional power plants, storage in saline aquifers and usage of CO_2 for EOR activities. The study also aims to identify the cost of implementing CCUS per ton of CO_2 emission prevented.

This chapter explores the implementation of the methodology described in Chapter 3 for the North-Central region of USA. Section 4.2 details the implementation of AHP on geo-information layers and preparation of the cost map. Section 4.3 discusses the preparation and implementation of the CCUS networks and pipeline routing algorithms. The effects of generating a CCUS network to capture specific amounts of CO_2 are discussed in Section 4.4 and 4.5. While Section 4.6 explores the effect of incremental changes to the CO_2 capture goals on CCUS systems. Section 4.7. reviews the established text analysis workflow.

4.2 Analytic Hierarchy Process and Preparation of the Cost Map

The geo-information utilized in the analysis of the study area includes 19 layers: roads and trains, rivers, lakes, parks, slope, corrosion susceptibility, soil, frost effect on topsoil, faults, towns, Western Hemisphere Shorebird Reserve Network (WHRSN), areas of critical environmental concern (ACEC), protected land, federal lands, existent energy pipelines (natural gas, crude oil, crude products) and land use. Amongst these, 17 layers are classified into three categories for further processing: technical barriers, regulatory barriers, and right-of-way barriers (exclusions: reservations and urban regions) as shown in Table 4.1. These categories will be at higher levels of hierarchy in the AHP analysis compared to the base geo-information layers.

The base layers are modified to have a uniform coordinate system: NAD 1983 UTM Zone 14N; and datum: D_north_america_1983. The layers are processed to add in buffers and clipped to include the features just within the study area (details provided in Section 3.3). For better handling, the layers are all converted into raster format with a 100-by-100 m resolution and snapped on to each other (in this case all layers are snapped to the slope map).

Table 4.1: Hierarchical Structure of data layers for implementation of AHP (Balaji et al., 2020).

	Cost Map		
Absolute Block	***Technical***	***ROW***	***Regulatory***
Cities	Slope	Federal Land	Towns
Reserves	Corrosion	Pipe	WHRSN
	Waterways	Land Use	Protected Land
	Lake	Parks	ACEC
	Soil	Highway	
	Frost	Train	
	Fault		

The processed geo-information layers are reclassified, as mentioned in Chapter 3, according to their favorability towards pipeline development and maintenance. The reclassification for all maps are based on a continuous numeric scale from 1 to 10, with 1 being most favorable and 10 being the least. The reclassification of the 17 layers are based on the type of data: categorical or continuous. The data for the reclassification of each base layer is provided in appendix A.

After the reclassification of the raster data into uniform comparable layers, the AHP technique is applied. First, the pairwise comparison of the content within each barrier criteria is carried out, which measures the importance of each factor with respect to the other. This is followed by the preparation of a normalized pairwise comparison matrix and the relative weightage computation (details in Section 3.4). The consistency index and consistency ratio serve as a validity check for the weightage obtained from the analysis. The results obtained from the AHP methodology is reproduction from Balaji et al., (2020).

The factors included in the pairwise comparative analysis of technical barriers are slope, corrosion susceptibility of steel, rivers, lakes, soil particle size, frost action and faults. Lakes rae given the most importance as they form natural barriers and raise the cost of construction of maintainence and construction of pipeline. Further tlaskes also serve as the natural habitat of several anaimal and bird species (Middleton et al., 2012a; Herzog et al., 2009). High weight is assigned to slope to it's steep impact on pipeline construction costs (Menon, 2011). Rivers have a comapritive alrge weightage assigned, although several of streams used in the study include sub-streams which can be burrowed through (Yousefi-Sahzabi et al., 2011). Soil particle size, frost action and corrosion susceptibility are local factors affecting the long-terms maintainance of pipelines and thus given lower weightages (Pelitere et al., 2018; He et al., 2015; Cevik et al, 2003, USA CFR 195). Fault zones are a liability in security of pipelines, however most faults have lower slip ratesand not active, hence warranting a low weightage (Potter, 2013). Table 4.2 indicates the pairwise comparison for natural barriers along with their respective weights and consistency values. The λ_{max} value, which is the maximum eigen value of the comparitive table, is taken as 7.35 from Table 3.3. RI, which is the random consistency index, is obtained as 1.32 for seven elements. Using the λ_{max} and RI gives the analysis a CR (consistency ratio) of 0.044 which is well below 0.1. The consistency ratio indicates the bias in the analysis. The symbology used in this section is adopted from Section 3.4.

The pairwise comparison of regulatory barriers include small towns, WHRSN sites, protected lands and ACEC zones. Towns and urban regions have been given highest weightage of the criteria because of regulatory and risk considerations (Menon, 2011; Middleton et al., 2012a; Herzog et al., 2000; Sun et al, 2013). According to federal regulations, various ecological considerations related to unusually sensitive areas need to be assessed (USA CFR Section 49, 2019). Pipeline and Hazardous Material Safety Administration [PHMSA] (2019) suggests the usage of both protected lands as well as ACEC zones as probable ecological zones for cosnideration as a unusually sensitive area. Towns, ACEC zones and protected lands are given equal weightage in the study, only WHRSN zones are given lower weightage, as WHRSN zones are not a suggested ecological unusually senesitve area. Table 4.3 indicates the pairwise comparison for regulatory barriers. From the pairwise comparision, the CR is obtained as 0.047 which is below the cut-off point of 0.1 using the constants obatined from Table 3.4.

Table 4.2: Pairwise comparison of technical barriers (Balaji et al., 2020).

	Consistency ratio (CR) = 0.044							
	Slope	**Corrosion**	**River**	**Lake**	**Soil**	**Frost**	**Fault**	**Weight**
Slope	1.00	5.00	3.00	0.20	5.00	5.00	7.00	0.24
Corrosion	0.20	1.00	0.33	0.14	1.00	1.00	3.00	0.06
River	0.33	3.00	1.00	0.33	3.00	3.00	5.00	0.14
Lake	5.00	7.00	3.00	1.00	7.00	7.00	9.00	0.43
Soil	0.20	1.00	0.33	0.14	1.00	1.00	3.00	0.06
Frost	0.20	1.00	0.33	0.14	1.00	1.00	3.00	0.06
Fault	0.14	0.33	0.20	0.11	0.33	0.33	1.00	0.03

Table 4.3: Pairwise comparison of regulatory barriers (Balaji et al., 2020).

	Consistency ratio (CR) = 0.047				
	Towns	**WHRSN**	**Protected Land**	**ACEC**	**Weights**
Towns	1	3	1	1	0.3
WHRSN	0.33	1	0.5	0.5	0.12
Protected Land	1	2	1	1	0.27
ACEC	1	3	1	1	0.3

The pairwise comparison of the right-of-way barriers consist of federal lands, existent pipelines, land use, parks, highways and trains. Exisiting pipeline regions have a siginificant positive effect on the cost of right-of-way and also encourages quicker processing and rights acquisition (Menon, 2011; Middleton et al., 2012a; Kanudia et al., 2013). High weights are provided to parks and federal lands as these regions have a higher ROW cost (Menon et al., 2011; Middleton et al., 2012a; Potter et al., 2013). Along with land use, federal lands and parks are given equal weightage as the development and land costs are related to the type of land cover/use (Broek et al., 2010; Middleton et al., 2012a). The road and rail networks are given equal but low weightages as the cost increase related to ROW by these factors is lower(Menon et al., 2011; Middleton et al., 2012a; Herzog et al., 2009). Table 4.4 indicates the pairwise comparisons for right-of-way barriers. For the case of the pairwise comparision of rights-of-way barrier, the CR is calculated as 0.042 which is below 0.1.

Table 4.4: Pairwise comparison of right-of-way barriers (Balaji et al., 2020).

	Consistency ratio (CR) = 0.042						
	Federal Land	**Pipe**	**Land Use**	**Parks**	**Highway**	**Train**	**Weights**
Federal Land	1	0.33	1	1	2	2	0.15
Pipe	3	1	3	3	3	3	0.36
Land Use	1	0.33	1	1	2	2	0.15
Parks	1	0.33	1	1	3	3	0.18
Highway	0.5	0.33	0.5	0.5	1	1	0.08
Train	0.5	0.33	0.5	0.5	1	1	0.08

The first level of hierarchy deals with the different categories of geo-information layers: natural barriers, regulatory barriers and rights-of-way barriers. Natural barriers have the most significant effect on pipeline costs and are provided the most weightage as compared to regulatory and ROW barriers (Base case). The reasoning for the higher weightage of natural barriers also is encouraged by literature review, which indicate that the cost related to pipeline construction is slarger in comparision to costs related to ROW and regulatiory compliance (McCollum et al., 2006; Rubin et al., 2008; Rui et al., 2011). Table 4.5 shows the pairwise comparisons for the above categories. The λ_{max} value is taken as 3.05 and RI is obtained as 0.58 for three elements, giving a CR value of 0.046 which is well below 0.1.

The overall weightages from the different geo-information layers are obtained by multiplying their respective indivdual weights from Table 4.2, 4.3 and 4.4 with the weights of their respective barrier criteria from Table 4.5. Table 4.6 gives the weightages of each of the 17 geo-information layers to be used in the formation of the cost map.

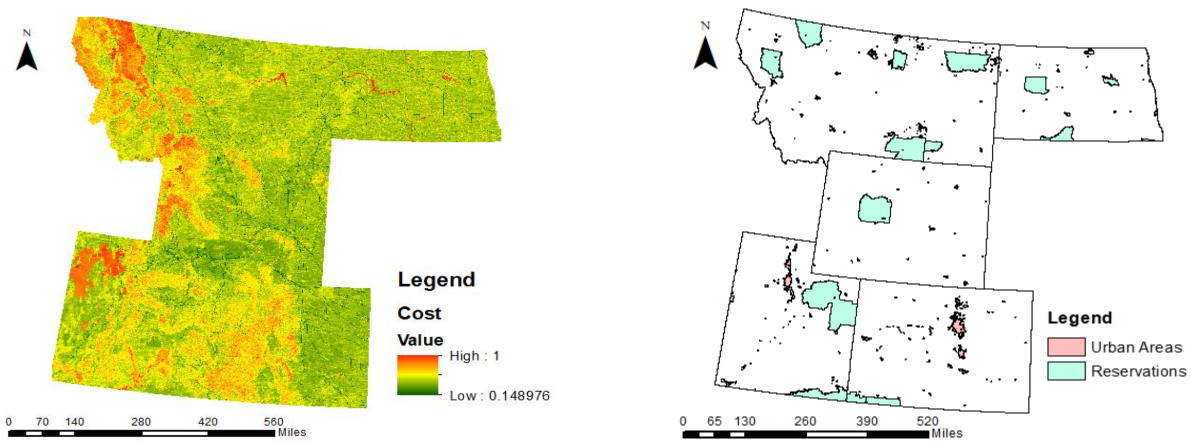
Table 4.5: Pairwise comparison of barrier criteria (Balaji et al., 2020).

	Consistency ratio (CR) = 0.046			
	Natural	**Regulatory**	**ROW**	**Weights**
Natural	1	2	2	0.49
Regulatory	0.5	1	0.5	0.2
ROW	0.5	2	1	0.31

Table 4.6: Overall weightage of geo-information layers for usage in weighted summation.

Technical Barriers		Regulatory Barriers		Rights-of-way Barriers	
Layer	**Weight**	**Layer**	**Weight**	**Layer**	**Weight**
Slope	0.12	Towns	0.06	Federal Land	0.05
Corrosion	0.03	WHRSN	0.02	Pipe	0.11
River	0.07	Protected Land	0.05	Land Use	0.05
Lake	0.21	ACEC	0.06	Parks	0.05
Soil	0.03			Highway	0.03
Frost	0.03			Train	0.03
Fault	0.01				

Figure 4.1(a) is the generated cost map from the weighted overlay (Balaji et al., 2020). The geo-information layers categorized as *Absolute Block* (cities and reservations) are removed from the cost map, depicted in Figure 4.1(b). The reformed cost map without cities and reservations is shown in Figure 4.2. The generated cost map has a value range of 0 to 1, with 0 representing the areas that are most suitable and 1 representing the areas that are least suitable for placing pipelines. From the analysis of the overall cost map, the initial assessment notes that no region is perfect for pipeline placement.



(a) North-Central USA CO_2 Pipeline Cost Map

(b) North Central USA Blocked Zones

Figure 4.1: (a)North-Central USA CO_2 pipeline cost map (without barriers), depicting the overall aggregated AHP-derived values, (b)North-Central USA exclusion zones left out from area of study. (Balaji et al., 2020)

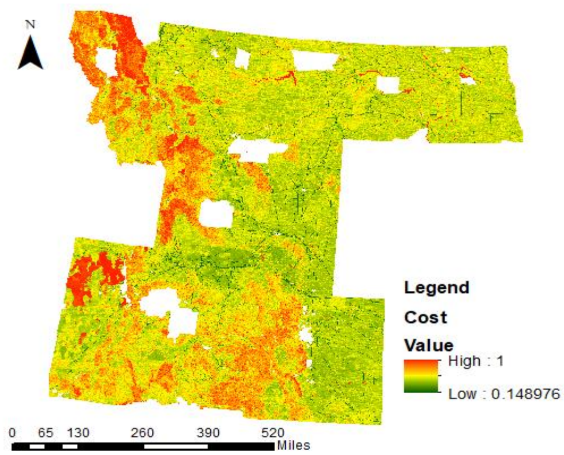


Figure 4.2: CO_2 pipeline cost map with barriers indicating potential cost related to building and maintenance of pipelines (Balaji et al., 2020).

North Dakota has large amounts of area which is sutable for pipeline dvelopment as seen in the cost map (Balaji et al., 2020). But traditional few pipeline pass through many parts of the state due to lower amount of demand. Pipeline unsuitibility rises in the central part of North Dakota near Lake Sakakawea as it is also a natural bird habitat. Such regions of unfavorable for pipeline development can also be found near Devil's Lake and the Spirit Lake reservation. North Dakota in recent hisotry has had several geopolotical issues with pipeline development especially related to reservations (New York Times, 2016).

Eastern Montana is also suitable for pipeline development except for the regions surrounding the Missouri River and Fort Peck Indian Reservation (Balaji et al., 2020). Western Montana has steep slopes and many national Parks including the Flathead national park and Kootenai national parl. The south western part of Montana is also unfavotable due to the combination of having many towns, moderate slopes and ecologic habitats.

Wyoming is a region having large amounts of development in terms of CO_2 EOR as seen in Section 3.2, leading to several pre-exisitng crude oil and natural gas pipelines (Balaji et al., 2020). Western part of Wyoming has several tract unfavorable towards pipeline developmentdue to steep slopes, several national parks and federal lands. Utah overall can be considered as a region of lower pipeline suitability due to large tracts of lands being federally owned along with the presence of the Salt Lake which is a ecological hotspot. In addition, Utah also several reservation sand national parks within it's bounds (Balaji et al., 2020).

Colorado has a long history related to oil and gas devlopment thus having many existaing pipelines especially due to deman, making the eastern part of the state favorable for pipelines (Balaji et al., 2020). But, the central and western parts of the state has several national parks, ecological hotspots as well as mountainous regions making in unfavorable for pipelines.

From the development and analysis of the cost map it is seen that exisiting pipelines regions are most influential to regions for future development of pipelines (Balaji et al., 2020). This is compunded by the existance of these routes near the vicinity of rail and road networks. A unique observation is the overlap of agricultural regions with low cost zones for pipelines, which is pro oted by the large distance between farmlands and parks and ecological zones. Regions with high cost for pipelines have high coincidence with lakes, since lakes also serve as ecological hotspots (Balaji et al., 2020). Other regiona with high cost include regions with higher elevation, parks and polar vegetation. The high weightages associated with higher elevation regions exist despite low human population and acceptable soil particle size.

Geo-information layers such as existent pipelines, slope, lakes, waterways, protected land, ACEC zones and land use have high coincidence with low cost regions in the cost map (Balaji et al., 2020). Whereas factors such as population distribution, soil particle size, corrosion susceptibility, roads and rail networks and fault zones coincide with regions of high cost in the cost map (Balaji et al., 2020). Altough, it is also seen that no sible factor affects the overall affinity of a piece of land for pipelines, but, rather the cost is determined by commingling of several different geo-information layers.

The region has four pre-exisitng major CO_2 pipeline systems: The Green Core pipeline, Exxon pipeline, FDL pipeline and the Dakota Gasification pipeline which runs to the Weyburn Field at Canada as shown in Figure 4.3. These pipeline system are all located in regions of low cost in the cost map marked by other types of pipeline networks, low slope areas, low population density areas and regions with open rock or sparse vegetations. None of the pipeline systems lie in regions with ACEC zones, protected land or WHRSN areas. But it must be noted that, these systems do pass through large regions of low to medium corrosion suscpetiibility and frost cover. Further analysis of the cost map in terms of weightages and land suitability will be discussed in Chapter 5.

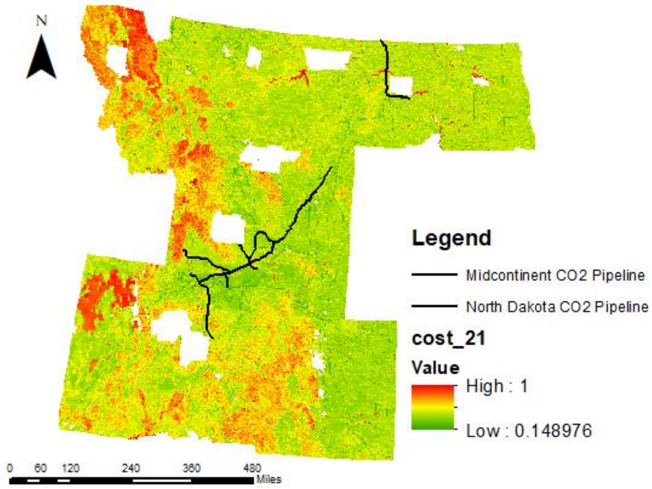


Figure 4.3: Existent CO_2 pipeline in the study area (Balaji et al., 2020).

4.3. CCUS Network Generation

This section discusses the process of route creation and overall CCUS network generation. The results of clustering, generation of Delaunay pairs as well as route generation using A* algorithm are included in this section. The combination of clustering, route generation and cost analysis together form a complete CCUS network.

4.3.1. Clustering

Several nodes in the study area are geographically close to one another. Initially the study considers a total of 30 sources and 26 sinks. The nodes that are within 24.24 km (15 miles) of each other are clustered together to reduce the number of nodes under consideration and to reduce the computational complexity of the study. The nodes are clustered together based on their geodesic coordinates as discussed in Section 3.4. It must be noted here, that in previous studies only nodes of similar types were clustered together, in other words, sources were only clustered with other sources and sinks were only clustered with other sinks. However, in this study sources and sinks can be clustered together, due to the nature of the bi-directional arc formulation in the optimization problem. After the clustering, the study area contains 20 sources and 23 sinks (40 in total, 3 nodes are both sources and sinks). An example of such clustering is seen in the case of Wyodak power station, Wygen1 and Wygen 2 power stations, Neil Simpsons and Dry Fork Station. These sources are all within 24.24km of their cluster center and are taken together as a single stationary source. The overall reduction of node points also reduces the probable number of interconnected potential routes between the various points from 1326 arcs in the case of un-clustered nodes to 780 arcs post-clustering. Figure 4.4 depicts the various clustered nodes under consideration in the study after the clustering operations. Table 4.7 shows the geographic details and description of the clustered nodes. The details of the clustered nodes in terms of emission capacity, storage capacity and reservoir properties are provided in Appendix C.

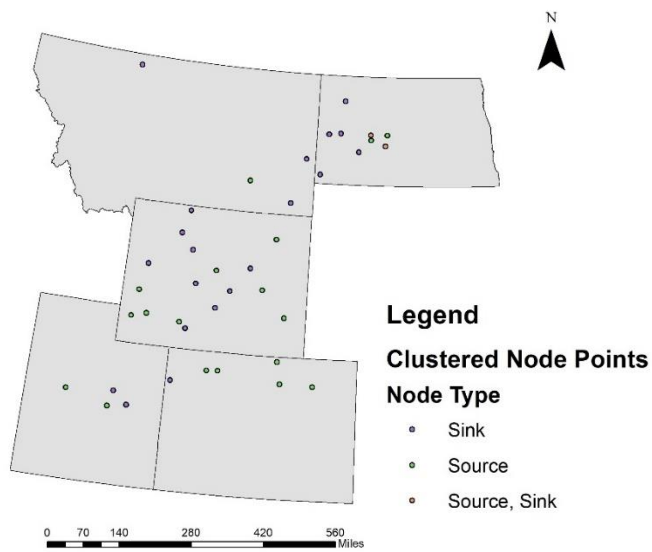


Figure 4.4: Clustering of node points upto a radius of 24.24km of each other.

Table 4.7: Clustered nodes include sinks, sources and dual nodes obtained from clustering operation.

Node ID	Longitude	Latitude	Name	Location	Node Type
0	-106.61	45.88	Colstrip	MT	Source
1	-104.88	42.11	Laramie River	WY	Source
2	-108.79	41.74	Jim Bridger	WY	Source

3	-101.16	47.38	Coal Creek & Leland Olds	ND	Source
4	-111.03	39.17	Hunter & Huntington	UT	Source
5	-112.58	39.51	Intermountain	UT	Source
6	-101.84	47.37	Great Plains Gasification & Antelope Valley	ND	Source & Sink
7	-107.59	40.46	Craig	CO	Source & Sink
8	-101.21	47.07	Milton R. Young	ND	Source & Sink
9	-105.78	42.84	Dave Johnston	WY	Source
10	-110.60	41.76	Naughton	WY	Source
11	-103.68	40.22	Pawnee	CO	Source
12	-110.04	41.87	Shute Creek	WY	Source
13	-105.39	44.29	Wyodak, Wygen, Neil Simpson, Dry Fork Station	WY	Source
14	-101.81	47.22	Coyote	ND	Source
15	-107.19	40.49	Hayden	CO	Source
16	-105.03	40.86	Rawhide Energy	CO	Source
17	-104.88	40.24	Fort St. Vrain	CO	Source
18	-107.60	43.27	Lost Cabin	WY	Source
19	-110.42	42.50	Riley Ridge	WY	Source
20	-104.93	45.35	Bell Creek Facility	MT	Sink
21	-108.81	44.87	Elk Basin Gas Plant	WY	Sink
22	-107.00	42.74	Grieve Facility	WY	Sink
23	-108.57	43.78	Hamilton Dome	WY	Sink
24	-107.50	42.23	Lost Soldier & Wertz	WY	Sink
25	-108.32	42.85	Beaver Creek	WY	Sink
26	-108.54	41.57	Monell (Patrick Draw)	WY	Sink
27	-108.88	40.10	Rangeley Field	CO	Sink
28	-109.08	44.23	Spring Creek Field, Oregon Basin, Pitchfork Field	WY	Sink
29	-106.31	43.43	Salt Creek	WY	Sink
30	-102.30	46.88	Red Trail Energy	ND	Sink
31	-111.64	48.70	Kevin Dome	MT	Sink
32	-110.87	39.62	Gordon Creek	UT	Sink
33	-110.21	43.26	Moxa Arch	WY	Sink

34	-110.35	39.27	Woodside Dome	UT	Sink
35	-104.42	46.61	Cedar Creek Anticline	MT	Sink
36	-103.83	46.21	Cedar Hill	ND	Sink
37	-103.07	47.38	Little Knife	ND	Sink
38	-103.57	47.34	Rough Rider	ND	Sink
39	-102.95	48.30	Beaver Lodge	ND	Sink

4.3.2 Delaunay Pairs

The clustered node points are processed through the Delaunay algorithm, which determines the best possible combination of pipeline arcs between multiple points. The Delaunay algorithm takes into consideration the geographic distribution of nodes and reduces the number of probable redundant pipeline arcs between the nodes. Figure 4.5 shows the generated pipeline arcs after processing them through the Delaunay algorithm. It must be noted that the straight line arcs indicated in Figure 4.5 are not real pipeline arcs between various nodes, as the probability of deviation of arcs from straight lines are quite high. For example, instead of creating a pipeline arc that stretches from the Naughton power plant, Utah to the Dakota Gasification plant at North Dakota, which are 1250 km apart, multiple interconnected nodes in between them can be used to make the connection. Also the probability of the need for CO_2 in a facility much closer to the Naughton power plant is quite high. Using the Delaunay algorithm the probable number of routes reduce from 780 to 104, helping to reduce the computational burden of the algorithm.

Similar to the case without pre-existent pipeline networks, the Delaunay triangulation helps to produce routes between points with existent pipeline infrastructure in the study area. The main difference between the two generated networks is the number of routes generated. In the case with pre-existent pipeline networks, the number of probable routes increases. This is due to the fact, that routes crossing pre-existent pipelines are excluded, but instead connections were made between the pre-existent pipeline routes and probable candidate arcs. This minor adjustment takes advantage of the Delaunay triangulation's nature of not circumscribing points in the network, to make a convenient network which does not cross pre-existing pipelines.

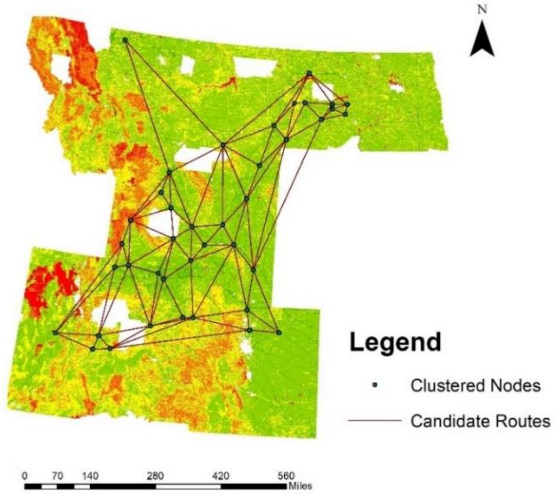


Figure 4.5: Pipeline route arc pairs generated by Delaunay algorithm represented by straight lines.

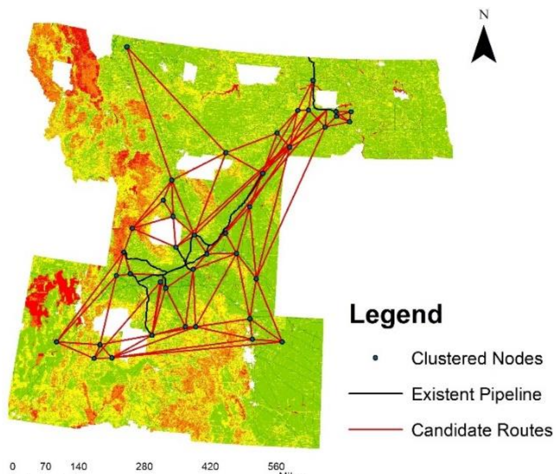


Figure 4.6: Pipeline route arc pairs generated by Delaunay algorithm represented by straight lines with existent pipeline networks.

4.3.3. Pipeline Routing

The process of development of a feasible path from one node to another is executed using the A* algorithm which was discussed in Section 3.4.3. The outputs from Delaunay algorithm as well as the cost map are used together in the execution of the A* algorithm in R, the output of which is a path in the raster data format. The path is in float point format, where every cell in the file is considered as an extension of the path, shown by Figure 4.7(a). This path is converted to an integer raster where float number (data capable of having a decimal point) is converted to integers, thus reducing the file size, shown by Figure 4.7(b). This integer raster file is converted into a vector file format polyline, which help in further processing and network building, shown by Figure 4.7(c). Finally, the path is simplified to remove deviations from the path as shown by Figure 4.7(d).

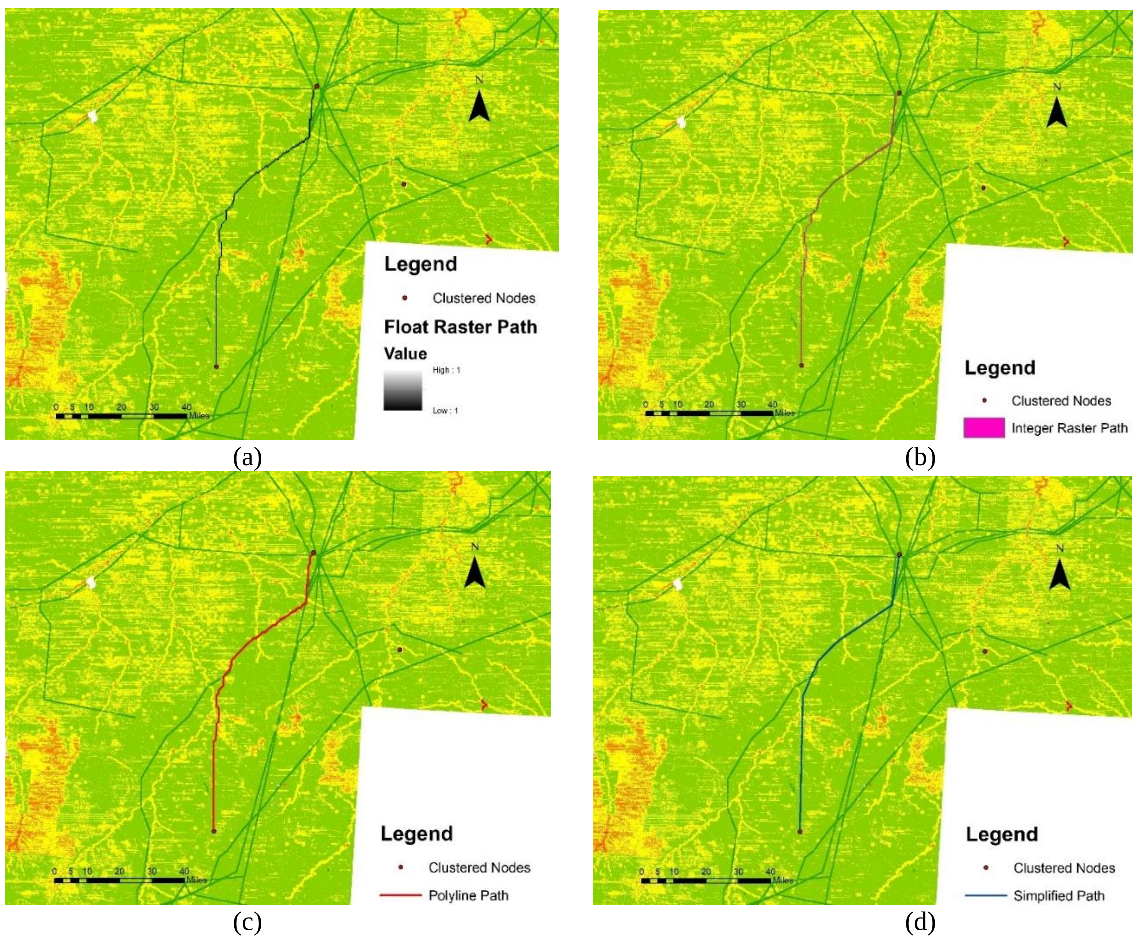


Figure 4.7: Paths between two node points: (a) Raster path generated by A* algorithm in float format; (b) Raster path converted to integer format, (note the change in legend representation); (c) Path converted to vector format; (d) Path simplified to reduce abrupt deviations.

The A* algorithm processes each pair of nodes individually and hence the A* algorithm is run in a loop to generate pipeline routes between each Delaunay pair and obtain the resultant routes in an efficient fashion. The output of the A* algorithm delivers a pipeline route from start node to end node, the average cost to transverse through each cell that the path passes through and the overall length of the path. The A* algorithm's Least Cost Path (LCP) is based on the weighted cost map. The average cost of movement through the path is basically the average of the cost of each cell contained in the route. Figure 4.8 depicts the generated pipeline arcs through the Delaunay pairs. It can be observed that none of the paths follow an exact straight route. It must be noted that there is an average increase of 15.76% in the length of the A* generated path as compared to the straight-line path. The application of the smoothening algorithm reduces the length of each path as it discards points that do not contribute to the overall shape of the pipeline route. There is an average reduction of 7.23% in the resultant pipeline arc lengths resulting from the application of the smoothening as compared to the A* generated routes. Even with this relative reduction in route lengths upon the application of smoothening algorithms, these final paths are still 7.06% longer than straight-line paths. Similar to analysis done in Figure 4.8, the Delaunay pairs obtained by processing clustered pairs with pre-existent pipeline are considered for route analysis. The A* algorithm is used iteratively to generate a path between all Delaunay pairs and converted to vector format. The simplified pipeline network is showcased in Figure 4.9 where the new probable pipeline arcs are shown commingled with pre-existent pipelines.

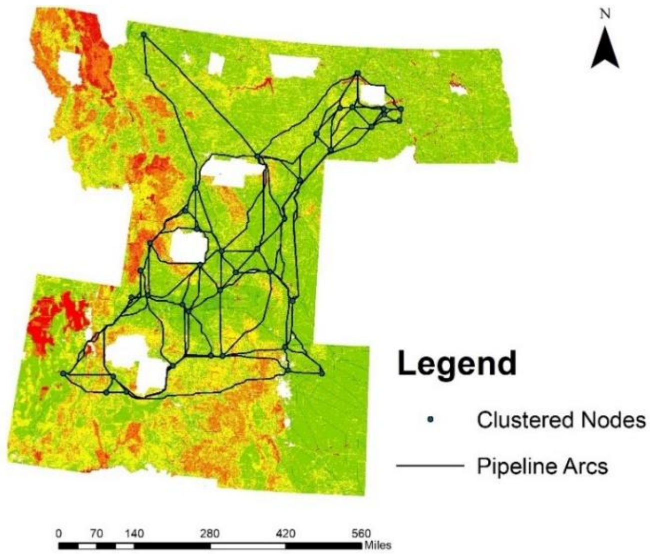


Figure 4.8: Pipeline route arcs generated by A* star algorithm.

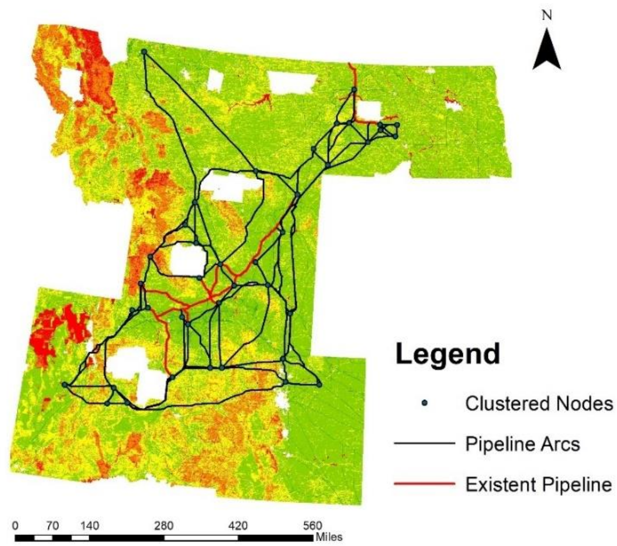


Figure 4.9: Simplified pipeline arcs obtained by using Decker-Peucker smoothening.

4.3.4. CCUS Network Generation

The CCUS network consists of the nodes and the candidate pipeline arcs as edges. From the clustering analysis and route generation technique, it is possible to generate the CCUS network for the study area. However, to populate the network with data, it is essential to analyze the nodes and edges with the techno-economic analysis shown in Section 3.5. The analyzed clustered source nodes and dual nodes are first used to calculate the maximum CO_2 emissions and supply rates along with the cost to capture every ton of CO_2 emitted from each individual source. The summary of the source nodes is provided in Appendix C. Further, the sink nodes are analyzed for their storage capacities and individual injection rates along with the associated cost for storage/EOR activities for each individual reservoir point. This analysis is further elaborated in appendix C. The candidate arcs generated from the route generation procedure are used for estimation of pipeline technical properties. Also, the cost of pipeline construction and maintenance is estimated for four different diameters for both cases: without pre-existent pipelines and with pre-existent pipelines. The cost analysis and properties are enumerated in Appendix C.

4.4 Static Decision Analysis

The static analysis of the CCUS network utilizes the Mixed Integer Linear Programming (MILP) formulation discussed in Section 3.6.1. The formulation optimizes the network and configures it such that a CCUS operation is built for the least cost to capture a fixed specified amount of CO_2 annually. The major inputs related to the analysis include the network model consisting of nodes and pipeline arcs. The other major input related to analysis is the result of the cost analysis of sinks, sources and pipelines conducted in the previous section, which populates the technical and economic parameters related to each component. Other key components include the estimated amount of CO_2 to be captured annually by the network, the timeframe for operations, the interest rate used in the study and the cost for purchasing CO_2 for EOR related operations.

The major consideration controlling the overall decision analysis is done by the control variables which indicate if a node/arc is being utilized. These control variables are driven by a linkage between material balance, flow, and economic constraints. The analysis done in this study utilizes annualized cost of operation, that is, the equal breakdown of capital and operational costs through the duration of operations. It is also important to note that the static analysis deploys the infrastructure for the network at the beginning of the operations. This type of analysis gives a fair indication of choices between the best operations in the CCUS network equipping stakeholders with a fair consideration of the economic and environmental impact of their choice.

In this section, the analysis of the study area using the network displayed in Figure 4.8 is conducted. The cost and technical inputs for the case study are provided in Appendix C. The analysis is conducted on a fixed level of CO_2 capture in the study area for an operational period of 30 years (NETL, 2018; IRGC, 2008). An annual interest rate of 3% is used for inflation, with the cost of CO_2 being at 20$/t of CO_2 purchased (NETL, 2018). Although the analysis in this section only checks the effect of varying CO_2 capture targets on the development of infrastructure and cost, further analysis related to the effect of time of operations, cost of CO_2, rate of recycle, monitoring, verification and abandonment (MVA) charge by regulatory body and other factors can be analyzed. Such analysis will be provided in Chapter 5. The details related to the amount of capture and storage at each node along with the transport quantities in the pipelines used in the static analysis is provided in appendix E.

Figure 4.10 depicts the static infrastructure deployment scenario for the case where 20 Mt CO_2/yr is to be captured for an operating period of 30 years, dubbed as scenario 1. In the figure, the capture quantities are represented by the beige bar, while the storage quantities are represented by the red bars. The pipelines are depicted according to their diameters in black lines proportional to the diameter of the suggested candidate pipeline to support the network. In this scenario a lot of capture occurs in North Dakota, Western Wyoming, and Colorado. The state-wise infrastructure development in scenario 1 Is provided in Table 4.8. As can be seen, most development in Scenario 1 is in North Dakota and Wyoming, with no feasible locations is Montana and Utah. In terms of pipeline development, most of the suggested networks consist of 12-inch pipelines except for a few 16-inch pipelines between Dakota Gasification and Milton R. Young facility; and Moxa Arch Storage facility and Riley Ridge. From the development of this scenario it can be noted that most of the utilized CO_2 sources are natural gas plants, as the cost of capture at these points are lower compared to other facilities. Similarly, saline aquifers are preferred as storage sites compared to EOR facilities as they have a cost related to purchasing each ton of CO_2. If the profits attained from sale of oil & gas were to be included in the analysis, the results would reflect the preference of EOR activities to permanent storage in saline aquifers. Another important factor to be noted is that the distance of pipeline networks in the network development is less important than the cost of CO_2 capture at the sources. The economics of scenario 1 is shown in Table 4.12.

Table 4.8: State-wise capture and sink statistics for scenario 1 (static, no pre-existing infrastructure, 20 $MtCO_2$/yr).

State	Capture Location	Capture ($MtCO_2$/yr)	Storage Location	Storage ($MtCO_2$/yr)
CO	Rawhinde Plant, Fort St. Vrain	2.59	Craig	4.06
ND	Great Plains Gasification, Milton R. Young	9.77	Great Plains Gasification, Milton R. Young, Red Trail Energy	10.08
WY	Jim Bridger, Shute Creek, Lost cabin, Riley Ridge	7.90	Moxa Arch	5.40

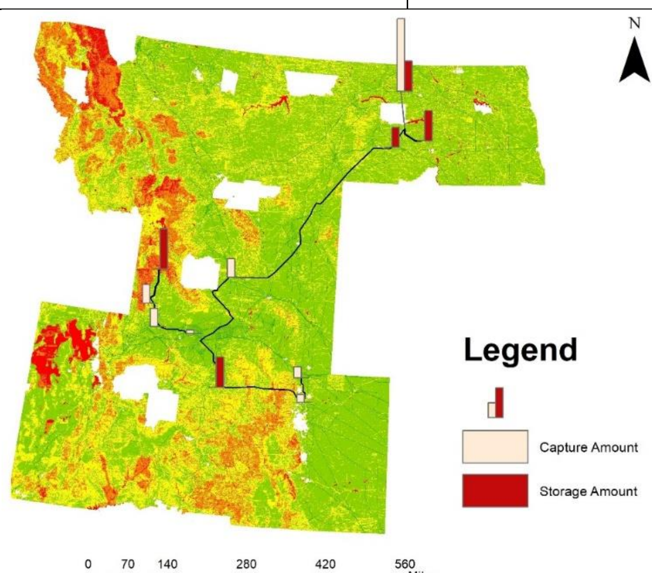


Figure 4.10: CCUS infrastructure for capturing 20 Mt CO_2/yr for 30 years in static environment (Scenario 1).

Figure 4.11 shows scenario 2 where 40 Mt CO_2/yr is captured for an operating period of 30 years. This scenario involves heavy development in North Dakota, Wyoming, Utah, and Colorado and the development is detailed in Table 4.9.

In terms of pipeline development, most routes are still dominated by low flow, thus using just 12-inch pipelines. However, a few 16-inch pipelines have been developed between Hunter, Gordon Creek and Woodside dome; Moxa Arch and Riley Ridge; Jim Bridger facility and Lost Cabin; and Dakota Gasification plant and Red Trail Energy saline aquifer. As can be observed in this scenario, most storage is still occurring around saline aquifers and capture operation are still focused on commercial suppliers and natural gas facilities. The economics of Scenario 2 is shown in Table 4.12.

Table 4.9: State-wise capture and sink statistics for scenario 2(static, no pre-existing infrastructure, 40 MtCO_2/yr).

State	Capture Location	Capture (MtCO_2/yr)	Storage Location	Storage (MtCO_2/yr)
CO	Craig, Rawhinde Facility, Fort St. Vrain	4.60	Craig, Rangeley Field	4.60
MT	-	0.00	Cedar Creek Anticline	1.08
ND	Great Plains Gasification, Milton R. Young	13.42	Great Plains Gasification, Milton R. Young, Red Trail Energy	12.34
UT	Hunter	8.10	Gordon Creek, Woodside Dome	8.10
WY	Jim Bridger, Shute Creek, Lost Cabin, Riley Ridge	14.40	Bairoil Field, Beaver Creek, Moxa Arch	0.00

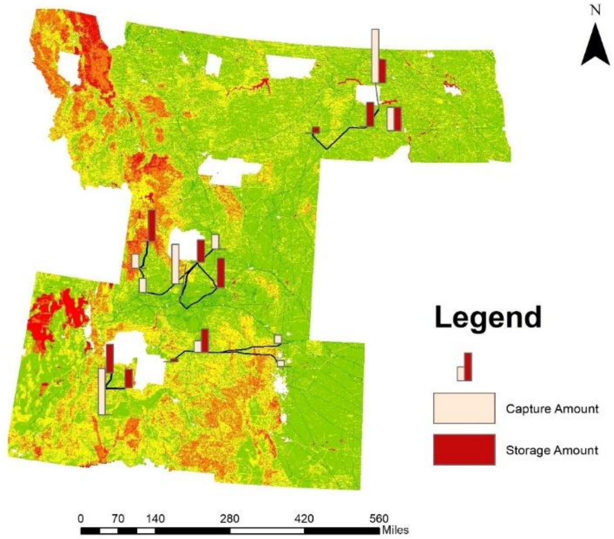


Figure 4.11: CCUS infrastructure for capturing 40Mt CO_2/yr for 30 years in static environment (Scenario 2).

Scenario 3 involves the capture of 60 Mt CO_2/yr for 30 years shown in Figure 4.12. The development in this scenario reveals usage of resources in all five states of the study area and is depicted in Table 4.10.

In scenario 3, the usage of EOR sites for storage has significantly been increased tending towards reservoirs that have higher injectivity. Further, coal-fired plants are used in significant quantities majorly governed by their proximity to convenient storage sites. In terms of pipeline, most connections are still 12- and 16- inch pipelines, however, the routes where multiple lines commingle, see a significant increase of volume, demanding usage of 24-inch pipelines as seen between Cedar Creek Anticline and Wyodak/Wygen power plants; and Bairoil EOR site and Laramie River. The economics of Scenario 3 is shown in Table 4.12.

Table 4.10: State-wise capture and sink statistics for scenario 3(static, no pre-existing infrastructure, 60 $MtCO_2$/yr).

State	Capture Location	Capture ($MtCO_2$/yr)	Storage Location	Storage ($MtCO_2$/yr)
CO	Craig, Hayden Facility, Fort St. Vrain	8.11	Craig, Rangeley Field	8.11
MT	Colstrip	1.82	Cedar Creek Anticline, Kevin Dome	9.92
ND	Great Plains Gasification, Milton R. Young	12.84	Great Plains Gasification, Milton R. Young, Red Trail Energy, Rough Rider field	12.84
UT	Hunter	8.10	Gordon Creek, Woodside Dome	8.10
WY	Laramie River, Jim Bridger, Shute Creek, Lost Cabin, Riley Ridge, Dave Johnston, Naughton Plant, Wyodak	29.18	Bairoil Field, Beaver Creek, Moxa Arch	21.08

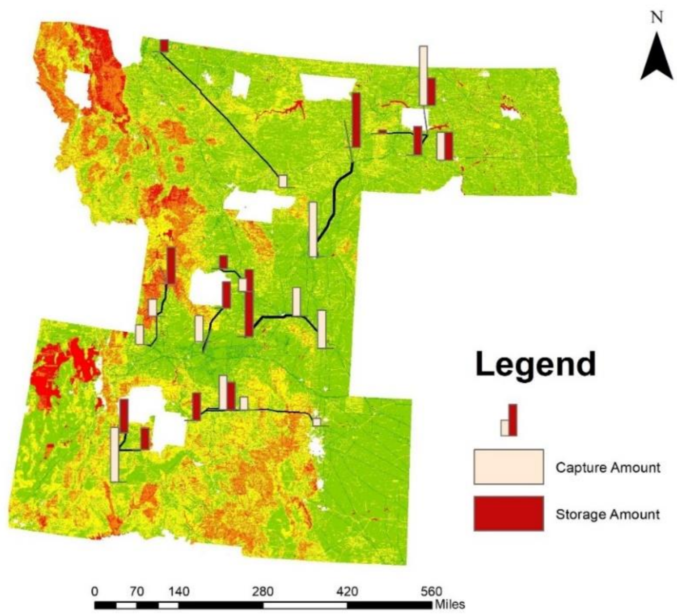


Figure 4.12: CCUS infrastructure for capturing 60$MtCO_2$/yr for 30 years in static environment (Scenario 3).

Figure 4.13 describes the development of scenario 4, where 80 Mt CO_2/yr is captured for an operating period of 30 years. The development adds on the infrastructure seen in scenario 3 significantly, depicted in Table 4.11.

Table 4.11: State-wise capture and sink statistics for scenario 4(static, no pre-existing infrastructure, 80 $MtCO_2$/yr).

State	Capture Location	Capture ($MtCO_2$/yr)	Storage Location	Storage ($MtCO_2$/yr)
CO	Hayden Facility, Rawhinde Facility, Fort St. Vrain	4.52	Craig, Rangeley Field	8.11
MT	Colstrip	11.39	Cedar Creek Anticline, Kevin Dome	9.92
ND	Great Plains Gasification, Milton R. Young, Coal Creek, Coyote Station	23.78	Great Plains Gasification, Milton R. Young, Red Trail Energy, Cedar Hill, Rough Rider field, beaver Lodge	22.85
UT	Hunter	8.10	Gordon Creek, Woodside Dome	8.10
WY	Laramie River, Jim Bridger, Shute Creek, Lost Cabin, Riley Ridge, Dave Johnston, Naughton Plant, Wyodak	32.21	Bairoil Field, Beaver Creek, Moxa Arch, Elk Basin, Grieve Field, Hamilton Dome, Oregon Basin, Salt Creek	31.02

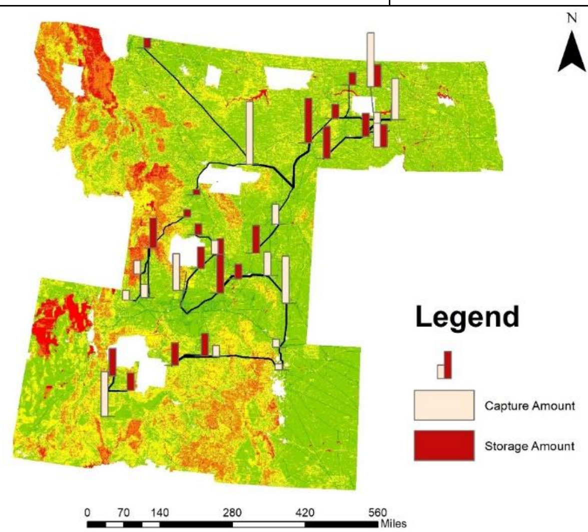


Figure 4.13: CCUS infrastructure for capturing 80$MtCO_2$/yr for 30 years in static environment (Scenario 4).

There is interconnection between the networks in Montana and North Dakota as well as the network in Central Wyoming and Colorado. In terms of pipelines, most connections are still 12- and 16-inch pipelines, however, there exist several 24-inch trunklines in Central Wyoming, Utah, and Western North Dakota. The economics of Scenario 3 is shown in Table 4.12.

The general trend as seen from the analysis of varying capture targets in the CCUS network indicates the preference of commercial vendors of CO_2 and natural gas plants over other types of power sources. In addition, saline aquifer-based sites are used over EOR fields due to lack of cost for purchasing CO_2 for storage. However, a strong sub-factor influencing decisions on storage sites is based on the number of wells and injectivity of individual reservoirs, as capital and operating costs are heavily based on the number of wells developed in each reservoir. In terms of capture nodes, choice between coal-fired power plants for CCUS development depends on connectivity of capture nodes to type of reservoirs. For instance, the development of the Hunter and Huntington plants is consistent for most cases beyond 30Mt CO_2/yr capture levels, due to the proximity of these plants to two convenient storage sites (Gordon Creek and Woodside Dome). In terms of pipeline development, most one-sided connection between sources and sinks do not have a pipeline more than 16-inch in diameter. However, commingling of fluid from several sources, can lead to the need for larger diameter pipelines as can be seen in Western Wyoming and Eastern Montana in scenarios 3 and 4.

The economic comparison between scenarios is done by analyzing the cost of capture, transport and storage per ton of CO_2 stored. This done by calculating the cumulative costs of capture, transport and storage in each scenario and dividing by the total captured CO_2 in the scenario. This is depicted by Table 5.12.

Table 4.12: Summary of cost analysis for static optimization without existing infrastructure (scenario 1-4).

Scenario	**Amount Captured** ($MtCO_2$/yr)	**Overall Cost** ($/ton)	**Capture Cost** ($/ton)	**Transport Cost** ($/ton)	**Storage Cost** ($/ton)
1	20.26	24.05	21.83	1.48	0.74
2	40.52	34.90	33.10	0.70	1.11
3	60.05	40.88	38.65	0.68	1.55
4	80.00	41.28	38.39	0.93	1.97

4.5. Static Analysis with Existent Infrastructure

Pre-existent CCUS infrastructure will affect the functioning and deployment of future potential CCUS networks. The setting for analyzing the effect of existent infrastructure is static in nature and follows an additional constraint for material balance in connector nodes as discussed in Section 3.6. The additional variations in the formulation include lower limits on the CO_2 capture levels of certain source nodes and lower limits on the injection into certain sink nodes. The lower limits of capture and injection levels ensure the definite usage of the nodes to the levels indicated by the US Inventory of Greenhouse Gas Emissions and Sinks (2020). These limits are indicated along with their respective node ID numbers in Table 4.13.

Table 4.13: The constraints on source and sink nodes considered in the static analysis with existent infrastructure.

Node ID	Name	Minimum Capture Quantity ($MtCO_2$/yr)	Minimum Injection Quantity ($MtCO_2$/yr)
6	Great Plains Gasification & Antelope Valley	3.5	0
12	Shute Creek	2	0
18	Lost Cabin	1.8	0
19	Riley Ridge	1	0
20	Bell Creek Facility	0	1.2
22	Grieve Facility	0	0.3
24	Lost Soldier & Wertz	0	1
25	Beaver Creek	0	0.6
26	Monell (Patrick Draw)	0	0.5
27	Rangeley Field	0	0.6
29	Salt Creek	0	0.6
39	Beaver Lodge (Lignite)	0	0.5
51	To Canada	0	3

The pre-existent pipeline networks have also been adjusted with the view of ensuring deliverance of CO_2 stream as shown in Table 4.14 and their respective pipeline parameters are obtained from the US Department of Transportation Pipeline and Hazardous Materials Safety Administration (2020). The limits on pipeline flow rates force the usage of existing pipelines. The removal of limits may cause the usage of alternative pipeline routes, which are not in current usage. This happens because the existing pipeline infrastructure network may not be the most cost-effective solution for the CCUS infrastructure.

Table 4.14: The constraints considered in existent pipelines for CO_2 transportation.

Start	End	Operator	Quantity ($MtCO_2$/yr)	Diameter
26	42	Exxon Mobil	0.5	12
6	43	Dakota Gasification	3.5	16
43	44	Dakota Gasification	3.5	16
44	45	Dakota Gasification	3.5	16
45	51	Dakota Gasification	3	24
27	46	Exxon Mobil	0.6	16
12	46	Chevron	2	24
40	46	Exxon Mobil	1.4	16
40	42	Exxon Mobil	1.4	16
42	47	Exxon Mobil	0.9	16
19	47	Denbury Resources	1	24
24	47	FDL	1	16
25	47	Denbury Resources	0.6	12
18	47	Denbury Resources	0.6	24
22	47	Exxon Mobil	0.9	16
18	41	Denbury Resources	1.2	24
22	29	FDL	0.6	16
20	41	Denbury Resources	1.2	24
39	45	Dakota Gasification	0.5	12

The analysis consists of additional node points essential for connection between new potential CCUS sites and the existent pipelines, for generation of interconnected lines and flow channels. The new nodes also consist of forks in the existent pipelines where pipelines converge or diverge. Although these additional nodes increase the computational time, they are essential to simulate existent networks and to obey proper flow patterns. This section explores the development of new CCUS infrastructure for four sets of minimum annual CO_2 capture goals similar to the analysis presented in Section 4.4 (20, 40, 60 and 80 Mt CO_2/yr). The detailed summary of the capture, storage, and transport quantities of CO_2 with the related costs are provided in appendix E. The operating period for all cases is taken as 30 years with an interest rate and cost of CO_2 at 3% and 20$/t CO_2 respectively. It can be derived from Table 4.13 that 8.3 MtCO_2/yr is already being captured and stored at existent locations using pipelines, which will be incorporated in every scenario in this section.

Figure 4.14 depicts scenario 5 for infrastructure deployment where 20 Mt CO_2/yr needs to be captured with the presence of existent infrastructure for an operational period of 30 years. The development in scenario 5 more than pre-existing infrastructure is provided in Table 4.15. All new pipelines in the case study are 12-inch in diameter. The development in terms of storage capacity of pre-existent sinks, were not expanded during this scenario. The cost for capture, transport and storage are provided in Table 4.19. Like in the cases without pre-existent infrastructure it can be seen that there is high favorability of aquifers over sinks when storing excess capture capacity despite pre-existent capacity in pipelines for carriage of CO_2.

Table 4.15: State-wise capture and sink statistics for scenario 5(static, pre-existing infrastructure, 20 $MtCO_2/yr$).

State	Capture Location	Capture ($MtCO_2/yr$)	Storage Location	Storage ($MtCO_2/yr$)
CO	Pawnee, Fort St. Vrain	2.80	Rangeley Field	0.21
ND	Great Plains Gasification	5.80	Great Plains Gasification, Beaver Lodge	2.48
UT	Hunter	1.35	Woodside Dome, Gordon Creek	4.15
WY	Shute Creek, Lost cabin, Riley Ridge	2.56	Grieve Field, Bairoil Field, Monell Field, Moxa Arch	2.30
External	-	-	Canada	2.60

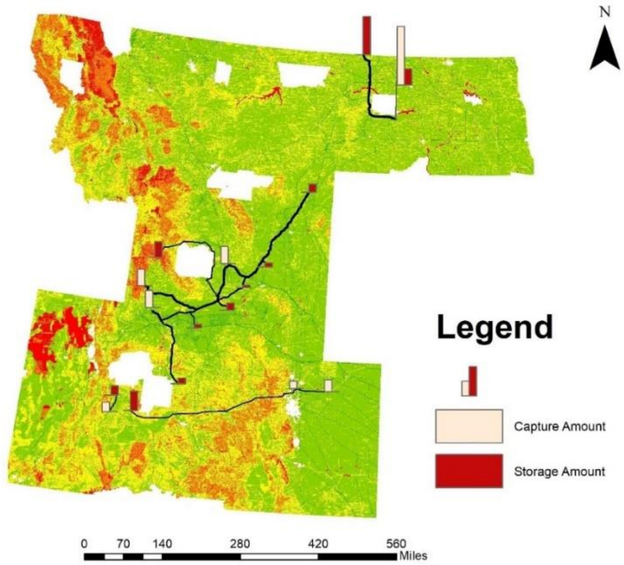


Figure 4.14: CCUS infrastructure network for capturing 20Mt CO_2/yr for 30 years in static environment with existent infrastructure (Scenario 5).

Scenario 6 depicted by Figure 4.15 shows the development of CCUS infrastructure to capture 40 Mt CO_2/yr for a 30-year operational period. 8.3 Mt CO_2/yr is captured with pre-existent infrastructure. Table 4.16 shows the development in each stage in addition to existing infrastructure. Pipeline development is limited to 12-inch pipelines across the study area except for the 16-inch pipeline connection between Hunter, Gordon Creek and Woodside dome. The economics related to scenario 6 is provided in Table 4.19. Yet again the development of traditional coal power plants is dependent on their proximity to convenient storage locations. There are fewer 16-inch pipelines developments as the Bairoil field and Dakota Gasification plant are already equipped with pre-existent pipelines to handle excess capacities as compared to scenario 2. In terms of cost, even though scenario 6 is more costly, the excess cost can be attributed to the high usage of coal-fired plants near convenient storage locations as compared to routing lower capture cost CO_2 from natural gas plants straight to saline aquifers.

Table 4.16: State-wise capture and sink statistics for scenario 6(static, pre-exiting infrastructure, 40 $MtCO_2$/yr).

State	Capture Location	Capture ($MtCO_2$/yr)	Storage Location	Storage ($MtCO_2$/yr)
CO	Craig	4.06	Craig, Rangeley Field	4.27
ND	Great Plains Gasification, Milton R.Young, Coyote Station	10.86	Great Plains Gasification, Milton R. Young, Beaver Lodge	8.26
UT	Hunter	8.10	Woodside Dome, Gordon Creek	8.10
WY	Jim Bridger, Naughton, Shute Creek, Lost cabin, Riley Ridge	8.77	Grieve Field, Bairoil Field, Beaver Creek, Monell Field, Moxa Arch	8.56
External	-	-	Canada	2.60

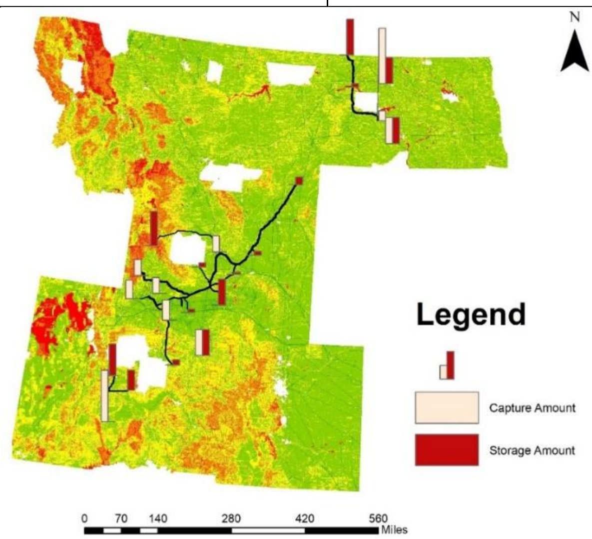


Figure 4.15: CCUS infrastructure network for capturing 40Mt CO_2/yr for 30 years in static environment with existent infrastructure (Scenario 6).

Scenario 7 is the infrastructure development formulated for a minimum CO_2 capture quantity of 60 Mt CO_2/yr as shown by Figure 4.16 and Table 4.17. There is also a 24-inch pipeline developed between Bell Creek facility and Cedar Creek anticline, similar to scenario 3. The cost analysis for scenario 7 is briefly showing in Table 4.19. This case shows that the pipeline developments made in the region is an extension of pre-existing pipelines, especially the lines built to the Wygen power plant region and the Cedar Creek Anticline. Also, the cost of executing scenario 7 is surprisingly lower than that of scenario 3, as the natural gas facilities in scenario 3 are underutilized.

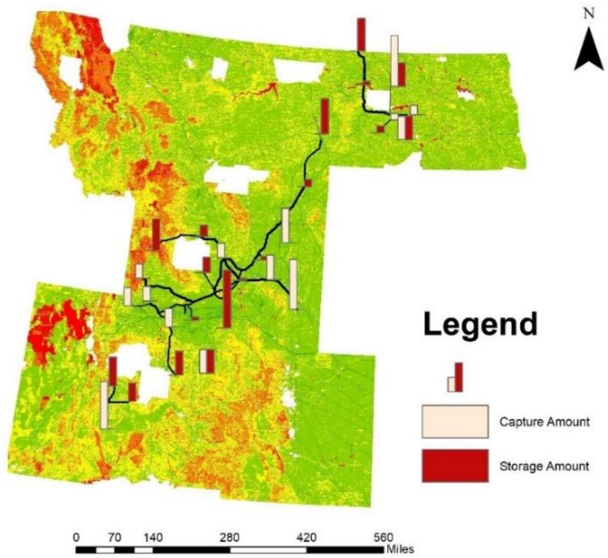


Figure 4.16: CCUS infrastructure network for capturing 60Mt CO_2/yr for 30 years in static environment with existent infrastructure (Scenario 7).

Table 4.17: State-wise capture and sink statistics for scenario 7(static, pre-exiting infrastructure, 60 MtCO_2/yr).

State	Capture Location	Capture (MtCO_2/yr)	Storage Location	Storage (MtCO_2/yr)
CO	Craig	4.06	Craig, Rangeley Field	7.51
MT	-	0.00	Cedar Creek Anticline	5.94
ND	Coal Creek, Great Plains Gasification, Milton R. Young, Coyote Station	11.96	Great Plains Gasification, Milton R. Young, Red Trail Energy, Beaver Lodge	9.36
UT	Hunter	8.10	Woodside Dome, Gordon Creek	8.10
WY	Laramie River, Jim Bridger, Dave Johnston, Naughton, Shute Creek, Wygen, Lost Cabin, Riley Ridge	27.60	Grieve Field, Hamilton Dome, Bairoil Field, Beaver Creek, Monell Field, Moxa Arch	18.21
External	-	-	Canada	2.60

Figure 4.17 shows the development scenario 8 where a minimum capture quantity of 80 Mt CO_2/yr is set for a 30-year operational period with pre-existent facilities. The development is briefly summarized in Table 4.18. There are several new 24- and 16-inch pipeline development throughout Wyoming and North Dakota. Overall, on comparison between scenario 4 and 8, the major differences are at the Colstrip power plant, Naughton power plant and Rough Rider field. The average costs related to scenario 8 are provided in Table 4.19. It seems that the effect of pre-existent networks on infrastructure diminishes as the CO_2 capture goals increase in this scenario. Also, the adoption of new pipelines in this scenario shows them as extension to pre-existent pipelines, leading to an outstretched webbed pattern in the pipeline network.

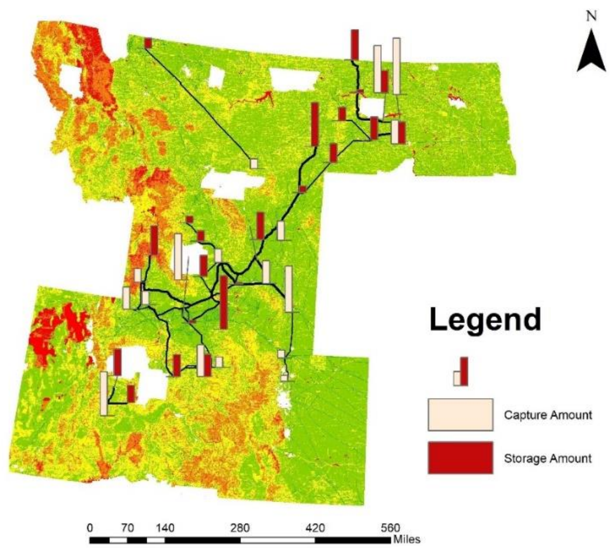


Figure 4.17: CCUS infrastructure network for capturing 80Mt CO_2/yr for 30 years in static environment with existent infrastructure (Scenario 8).

The general trend of infrastructure development in case of scenario 5 through 8 follows similar trends as seen in cases without pre-existent infrastructure developments. The costs for scenario 5 through 8 is provided in Table 4.19. The overall cost of the operations has gone up a little, due to the early usage of EOR reservoirs as compared to scenarios 1 through 4. The decrease in pipeline spending does not off put the spending made in terms of storage requirements. Despite these similarities there are several differences in CCUS infrastructure in scenarios with lower CO_2

capture targets. However, with higher CO_2 targets, these differences start to reduce and start becoming similar to cases without pre-existent networks. It can also be seen that with increased capture goals, a single CCUS corridor stretching from south-western Wyoming to central North Dakota is formed.

Table 4.18: State-wise capture and sink statistics for scenario 8(static, pre-exiting infrastructure, 80 $MtCO_2$/yr).

State	Capture Location	Capture ($MtCO_2$/yr)	Storage Location	Storage ($MtCO_2$/yr)
CO	Craig, Hayden, Rawhinde Plant, Fort St. Vrain	10.36	Craig, Rangeley Field	7.51
MT	Colstrip	1.82	Cedar Creek Anticline, Kevin Dome	9.92
ND	Coal Creek, Great Plains Gasification, Milton R.Young	19.94	Great Plains Gasification, Milton R. Young, Red Trail Energy, Cedar Hill, Little Knife, Rough Rider, Beaver Lodge	18.42
UT	Hunter	8.10	Woodside Dome, Gordon Creek	8.10
WY	Laramie River, Jim Bridger, Dave Johnston, Naughton, Shute Creek, Wygen, Lost cabin, Riley Ridge	31.52	Grieve Field, Hamilton Dome, Bairoil Field, Beaver Creek, Monell Field, Oregon Basin, Salt Creek, Moxa Arch	25.19
External	-	-	Canada	2.60

Table 4.19: Summary of cost analysis for static optimization with existing infrastructure (scenario 5-8).

Scenario	Amount Captured (MtCO2/yr)	Overall Cost ($/ton)	Capture Cost ($/ton)	Transport Cost ($/ton)	Storage Cost ($/ton)
5	20.06	1.45	25.57	22.552	1.56
6	40.09	1.11	35.38	33.71	0.57
7	60.02	1.55	39.27	37.12	0.60
8	80.04	1.78	41.27	38.62	0.86

4.6 Dynamic Decision Analysis

The dynamic decision analysis is based on a MILP formulation to reduce the annualized cost of operations and construction of the CCUS network. The project has multiple different CO_2 capture goals which gradually increases along with time. The entire duration of operations is divided into equal time intervals dependent on the number of time steps. A major variation from the static formulation in Section 4.4 is the modification of several flow and material balance constraints to maintain the balance of the network at each time interval on an annual basis.

The analysis of the network in a dynamic setting is more computationally expensive as each constraint and variable in the formulation is dependent on an additional time parameter. The dependence of each parameter on the time variable increases the number of constraints by a factor equal to the number of time intervals to be analyzed by the system. In other words, if there are three time intervals, the number of constraints in the equation increases by a factor of three. The number of individual capture, transport and storage cost terms in the objective function also increases by a factor equal to the number of time intervals used in the analysis. This type of staggered deployment of CCUS infrastructure to capture increasing amounts of CO_2 is closer to a real world scenario as many countries set their carbon-dioxide emission control goals in such a staged manner (IPCC, 2013).

In the analysis, the period of operation is taken as 30 years divided into three time steps of 10 years each (2019-2049). The rate of interest and cost for CO_2/ton is taken as 3% and 20$/ton, respectively. Further the cost for MVA is fixed at 5cents/ton of CO_2 stored in reservoir. The reservoirs are utilized in a manner that the capacity of each reservoir lasts for the entire duration of CCUS operations. The detailed capture, storage and transportation quantities of s case study is provided in Appendix E.

Figure 4.18 shows scenario 9, where the capture goals are 20Mt CO_2/yr in the 2019-2029 period, 30 Mt CO_2/yr in 2029-2039 and 40 Mt CO_2/yr from 2039 to 2049. A brief description of the infrastructure development for scenario 9 is provided in Table 4.20. The economics related to scenario 9 for implementing CCUS is provided in Table 4.22. The development for capture capacity in this case is seen around natural gas plants and commercial vendors, while development of storage capacity is at saline aquifers. The development of the CCUS network is separate at each region and does not interconnect. Also, the pipelines developed in a single period of development may not be the best route economically for the next period of development. Most developed pipeline routes are 12- and 16-inch in diameter.

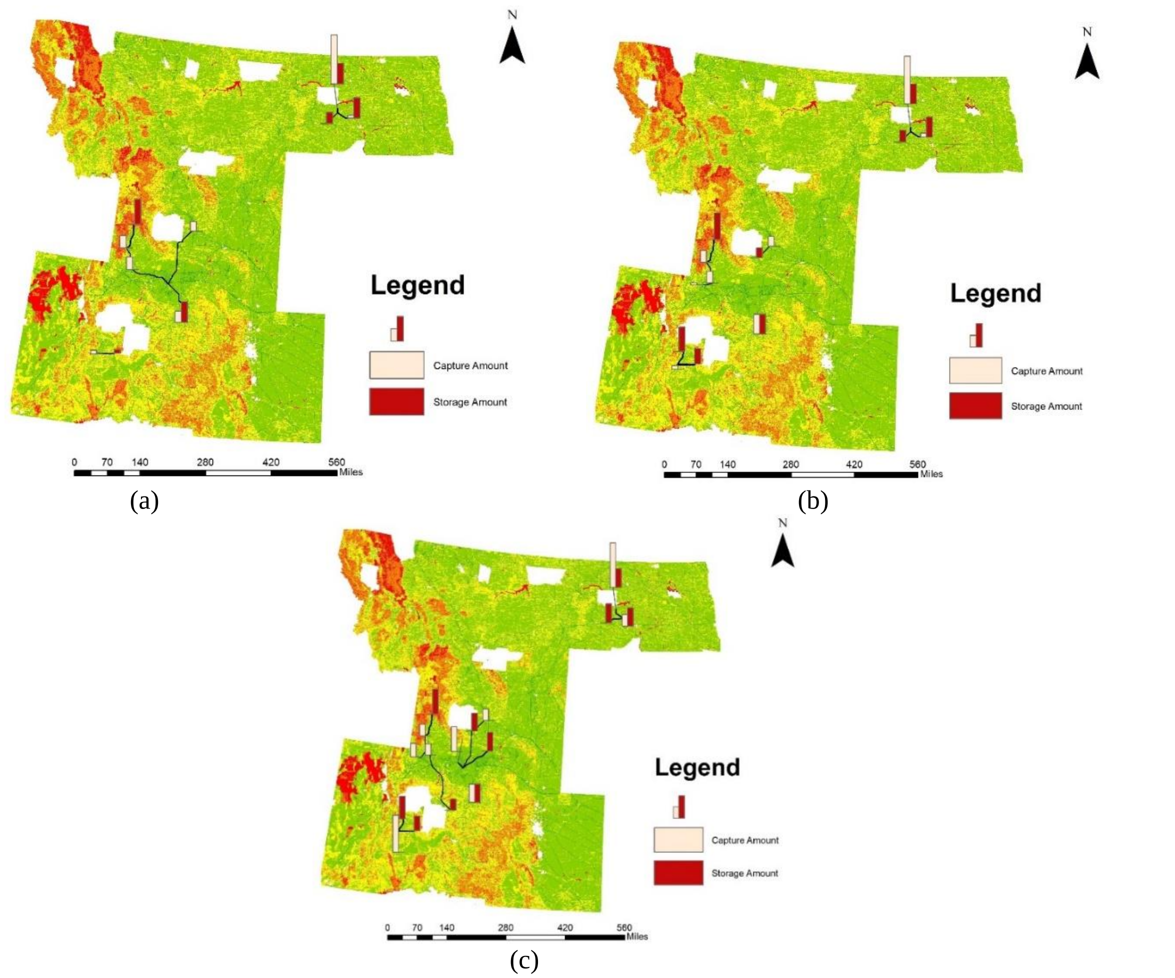


Figure 4.18: CCUS infrastructure network for 30 years in dynamic environments for scenario 9 with graduated CO_2 capture goals: (a) 20 Mt CO_2/yr for 2019-2029; (b) 30 Mt CO_2/yr for 2029-2039; (c) 40Mt CO_2/yr for 2039-2049.

Table 4.20: State-wise statistics for scenario 9(dynamic, 20-30-40 $MtCO_2$/yr).

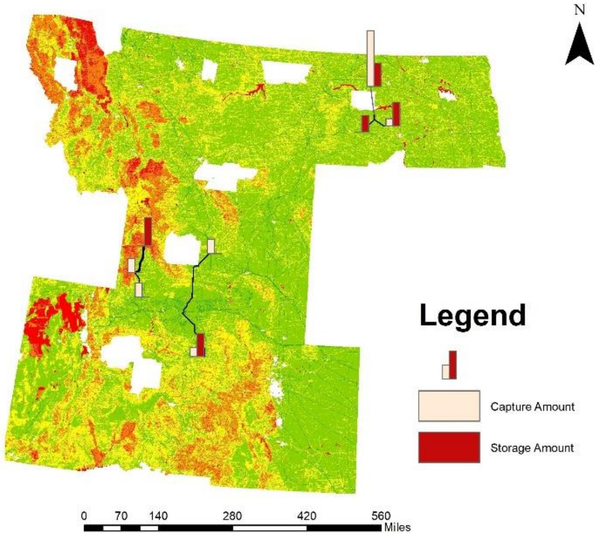
	State	Capture Location	Capture ($MtCO_2$/yr)	Storage Location	Storage ($MtCO_2$/yr)
2019-2029	*CO*	Craig	2.13	Craig	4.06
	ND	Great Plains Gasification, Milton R. Young	10.36	Great Plains Gasification, Milton R. Young, Red Trail Energy	10.36
	UT	Hunter	0.70	Woodside Dome	0.70
	WY	Shute Creek, Lost Cabin, Riley Ridge	6.88	Moxa Arch	4.95
2029-2039	*CO*	Craig	4.06	Craig	4.06
	ND	Great Plains Gasification, Milton R. Young	10.58	Great Plains Gasification, Milton R. Young, Red Trail Energy	10.58
	UT	Hunter	8.10	Woodside Dome, Gordon Creek	8.10
	WY	Shute Creek, Lost Cabin, Riley Ridge, Naughton	7.35	Beaver Creek, Moxa Arch	7.35
2039-2049	*CO*	Craig	4.06	Craig, Rangeley Field	6.49
	ND	Great Plains Gasification, Milton R. Young	12.34	Great Plains Gasification, Milton R. Young, Red Trail Energy	12.34
	UT	Hunter	8.10	Woodside Dome, Gordon Creek	8.10
	WY	Shute Creek, Lost Cabin, Riley Ridge, Naughton, Jim Bridger	15.81	Beaver Creek, Moxa Arch, Bairoil Field	13.38

Scenario 10 shows the case for capturing 20, 40 and 60 Mt CO_2/yr for the periods of 2019-2029, 2029-2039 and 2019 -2049 respectively. The brief description of infrastcrture in scenario 10 is provided in Table 4.21 and the related economics is provided in Table 4.22. For the period of 2029 to 2039, major 16-inch pipelines can be seen utilized in Utah between Hunter, Gordon Creek and Woodside dome; Riley Ridge and Moxa Arch along with development of a 24-inch pipeline between Bairoil fields and Laramie River facility, Montana; and Dakota Gasification plant and Red Trail Energy aquifer site. During this period (2039-2049) major 24-inch pipeline development can be seen between Colstrip plant and Cedar Creek anticline and between Hayden plant and Rangeley field. The developments each region in Scenario 9 are disconnected from one another.

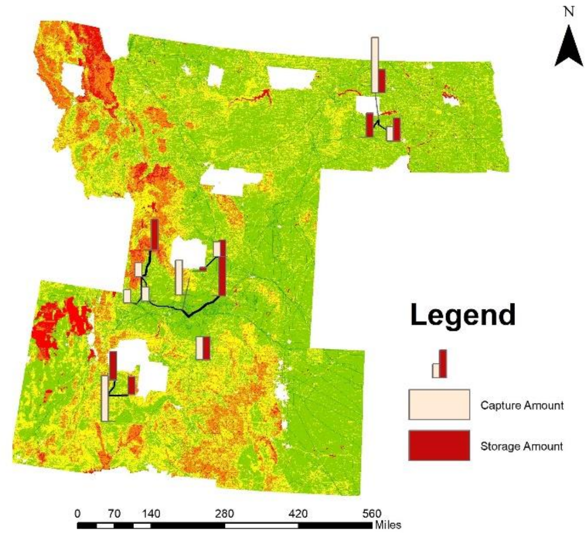
However, it can be seen that the overall cost of development of the scenario is much lower compared to static optimization despite having to constantly evolve/change pipelines and add new facilties. This phenomenon is mainly attributed to the distribution of capital expenditure through various time periods and development of infrastructure in a consistent incremental fashion.

Table 4.21: State-wise statistics for scenario 10(dynamic, 20-40-60 $MtCO_2$/yr).

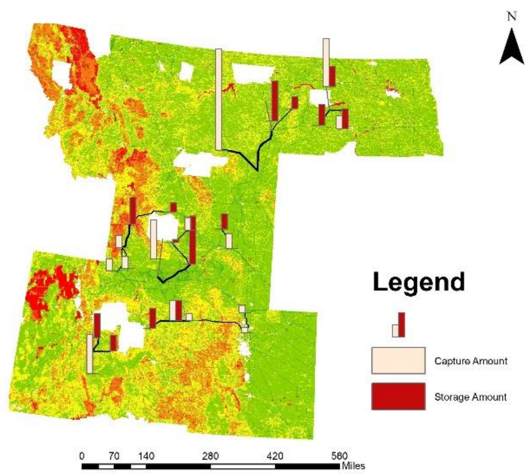
	State	Capture Location	Capture ($MtCO_2$/yr)	Storage Location	Storage ($MtCO_2$/yr)
2019-2029	*CO*	Craig	1.56	Craig	4.06
	ND	Great Plains Gasification, Milton R. Young	11.02	Great Plains Gasification, Milton R. Young, Red Trail Energy	11.02
	WY	Shute Creek, Lost Cabin, Riley Ridge	7.45	Moxa Arch	4.95
2029-2039	*CO*	Craig	4.06	Craig	4.06
	ND	Great Plains Gasification, Milton R. Young	12.34	Great Plains Gasification, Milton R. Young, Red Trail Energy	12.34
	UT	Hunter	8.10	Woodside Dome, Gordon Creek	8.10
	WY	Shute Creek, Lost Cabin, Riley Ridge, Naughton, Jim Bridger	15.91	Beaver Creek, Moxa Arch, Bairoil Field	15.91
2039-2049	*CO*	Craig, Hayden, Rawhinde Energy, Fort St.Vrain	8.11	Craig, Rangeley Field	8.11
	MT	Colstrip	10.60	Cedar Creek Anticline	8.10
	ND	Great Plains Gasification, Milton R. Young	12.34	Great Plains Gasification, Milton R. Young, Red Trail Energy, Rough Rider	14.84
	UT	Hunter	8.10	Woodside Dome, Gordon Creek	8.10
	WY	Shute Creek, Lost Cabin, Riley Ridge, Naughton, Jim Bridger, Dave Johnston	20.88	Beaver Creek, Moxa Arch, Bairoil Field, Hamilton Dome, Salt Creek	20.88



(a)



(b)



(c)

Figure 4.19: CCUS infrastructure network for 30 years in dynamic environment for scenario 10 with graduated CO_2 capture goals: (a) 20 Mt CO_2/yr for 2019-2029; (b) 40 Mt CO_2/yr for 2029-2039; (c) 60Mt CO_2/yr for 2039-2049.

Table 4.22: Summary of cost analysis for dynamic optimization without existing infrastructure (scenario 9 and 10).

Scenario	Average Amount Captured ($MtCO_2$/yr)	Overall Cost ($/ton)	Capture Cost ($/ton)	Transport Cost ($/ton)	Storage Cost ($/ton)
9	30.16	28.95	27.33	0.59	1.03
10	40.16	30.90	28.88	0.70	1.31

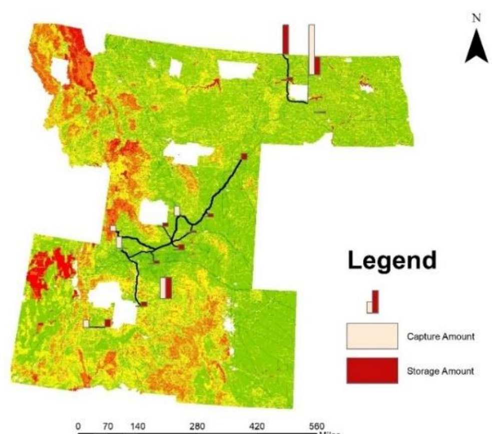
In scenarios 11 and 12, a dynamic decision analysis is made for graduated CO_2 capture targets for 2019-2029, 2029-2039 and 2039-2049 with pre-existent infrastructure. The aim of scenario 11 and 12 is to see the differences in development of CCUS network with and without the pre-existent infrastructure. The information for existent infrastructure is provided in Table 4.8 and 4.9.

Scenario 11 discusses the development of a CCUS infrastructure with graduated CO_2 capture goals of 20, 30 and 40 Mt CO_2/yr shown in Figure 4.20. A summary of the infrastructure in addition to the existent infrastructure for 8.3 MtCO_2/yr is provided in Table 4.23 and the overall cost summary is provided in Table 4.25. For the period from 2029 to 2039, two 16-inch pipelines are constructed between Hunter facility, Woodside Dome and Gordon Creek in Utah and between Riley Ridge and Moxa Arch. In 2039 to 2049, a 16-inch pipeline is built between Riley Ridge and Naughton facility. The major differences between scenarios 9 and 11 can be seen at level of deployment at Utah and North Dakota in terms of Hunter facility, Milton R. Young and Red Trail Energy. In addition, in Wyoming there is a huge difference in the usage of conventional coal fired power plants reflected by Laramie River and Naughton power plants. There are no major differences in pipelines except for a few additional 16in pipelines developed for scenario 9. The cost of the venture is also higher in scenario 11 as compared to scenario 9 due to the utilization of EOR sites. This behavior is similar to the comparison between the static case with and without pre-existent infrastructure, where low CO_2 capture goals resulted in higher costs in the former case.

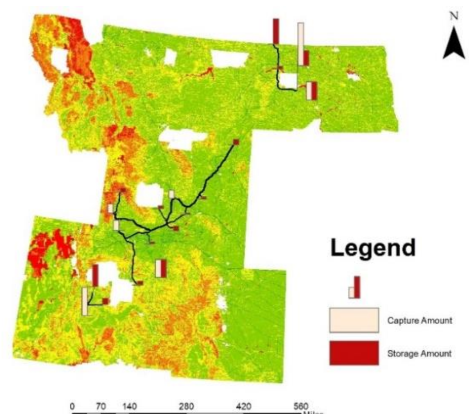
Table 4.23: State-wise statistics for scenario 11(dynamic, existing infrastructure, 20-30-40 $MtCO_2$/yr).

	State	Capture Location	Capture ($MtCO_2$/yr)	Storage Location	Storage ($MtCO_2$/yr)
2019 - 2029	*CO*	Craig	4.06	Craig, Rangeley Field	4.27
	ND	Great Plains Gasification, Milton R. Young	6.34	Great Plains Gasification, Milton R. Young, Beaver Lodge	3.74
	UT	Hunter	1.40	Woodside Dome	1.40
	WY	Shute Creek, Riley Ridge	0.31	Grieve Field, Bairoil Field, Beaver Creek, Monell Field	0.31
	Ext	-	0.00	Canada	2.60
2029 - 2039	*CO*	Craig	4.06	Craig	4.27
	ND	Great Plains Gasification, Milton R. Young	10.18	Great Plains Gasification, Milton R. Young, Beaver Lodge	7.58
	UT	Hunter	6.35	Woodside Dome, Gordon Creek	6.35
	WY	Shute Creek, Riley Ridge	1.21	Grieve Field, Bairoil Field, Beaver Creek, Monell Field, Moxa Arch	1.21
	Ext	-	0.00	Canada	2.60
2039 - 2049	*CO*	Craig, Hayden	4.54	Craig, Rangeley Field	4.54
	ND	Great Plains Gasification, Milton R. Young, Coyote Station	10.86	Great Plains Gasification, Milton R. Young, Beaver Lodge	8.26
	UT	Hunter	8.10	Woodside Dome, Gordon Creek	8.10
	WY	Shute Creek, Riley Ridge, Dave Johnston, Naughton, Lost Cabin	8.22	Grieve Field, Bairoil Field, Beaver Creek, Monell Field, Moxa Arch	8.70
	Ext	-	0.00	Canada	2.60

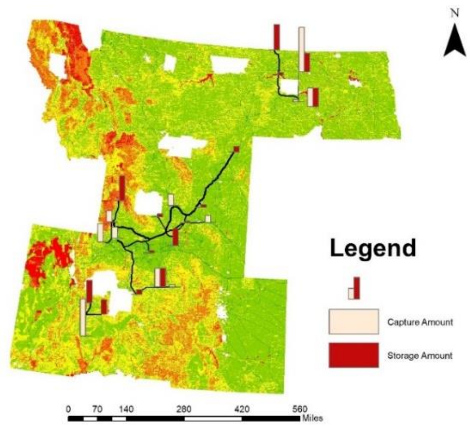
Scenario 12 summarizes the infrastructure development with pre-existing infrastructure development for the case where CO_2 capture goals are 20, 40 and 60 Mt CO_2/yr for 2019-2029, 2029-2039 and 2039-2049 respectively. Table 4.24 provides a brief summary of the infrastructure in scenario 12 and the costs are provided in Table 4.25. For the first period of time between 2019 and 2029, a 16-inch pipeline is developed between Dakota Gasification and Milton R. Young plants, North Dakota and Hunter facility and Woodside Dome, Utah. This development is similar to scenario 11. For the second period, between 2029 and 2039, additional 16-inch pipeline is built between Hunter facility and Gordon Creek, Utah; Grieve facility and Jim Bridger; and Naughton facility and Riley Ridge, Wyoming. In third time period, new 16-inch pipelines are built between Red Trail Energy aquifer and Milton R. Young facility; and Bairoil fields and Laramie river. The major differences between scenarios 10 and 12 can be seen in the development of Montana with the heavy usage of Colstrip power plant and Cedar Creek Anticline as well as the Laramie River power plant in Wyoming in scenario 10. Further in scenario 12, there is a network formed along the northern stretches of Colorado connecting Craig site, Fort St.Vrain, Rawhide power plant and Rangeley field, which is not seen in scenario 10. In terms of pipelines, scenario 10 sees many more 24in pipeline developments along Montana and eastern Wyoming as compared to scenario 12, due to the lack of pre-existent transportation routes. In terms of cost, both scenarios 10 and 12 are nearly equal.



(a)



(b)



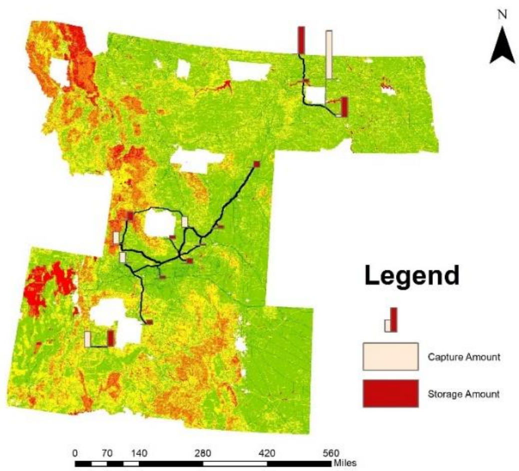
(c)

Figure 4.20: CCUS infrastructure network for 30 years in dynamic environments for scenario 11 with graduated CO_2 capture goals with existent infrastructure: (a) 20 $MtCO_2$/yr for 2019-2029; (b) 30 $MtCO_2$/yr for 2029-2039; (c) 40$MtCO_2$/yr for 2039-2049.

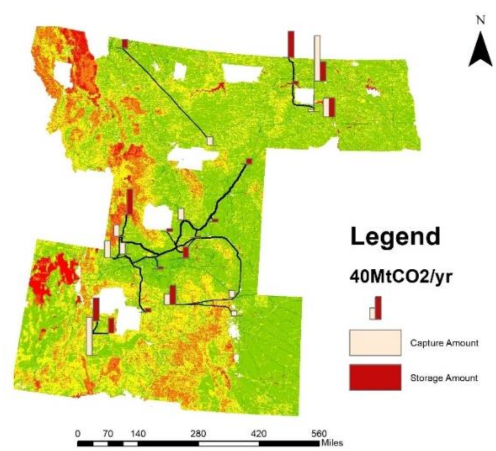
Table 4.24: State-wise statistics for scenario 12(dynamic, existing infrastructure, 20-40-60 $MtCO_2$/yr).

	State	**Capture Location**	**Capture ($MtCO_2$/yr)**	**Storage Location**	**Storage ($MtCO_2$/yr)**
2019 - 2029	*CO*	-	0.00	Rangeley Field	0.21
	ND	Great Plains Gasification, Milton R. Young	6.78	Milton R. Young, Beaver Lodge	4.18
	UT	Hunter	3.15	Woodside Dome	3.15
	WY	Shute Creek, Riley Ridge, Lost Cabin	2.11	Grieve Field, Bairoil Field, Beaver Creek, Monell Field, Moxa Arch	1.90
	Ext	-	0.00	Canada	2.60
2029 - 2039	*CO*	Craig, Rawhinde Plant, Fort St. Vrain	4.81	Craig, Rangeley Field	4.27
	MT	Colstrip	1.82	Kevin Dome	1.82
	ND	Great Plains Gasification, Milton R. Young, Coyote Station	10.86	Great Plains Gasification, Milton R. Young, Beaver Lodge	8.26
	UT	Hunter	8.10	Woodside Dome, Gordon Creek	8.10
	WY	Shute Creek, Riley Ridge, Naughton, Lost Cabin	6.32	Grieve Field, Bairoil Field, Beaver Creek, Monell Field, Moxa Arch	7.07
	Ext	-	0.00	Canada	2.60
2039 -	*CO*	Craig, Rawhinde Plant, Fort St. Vrain, Hayden	9.01	Craig, Rangeley Field	7.51

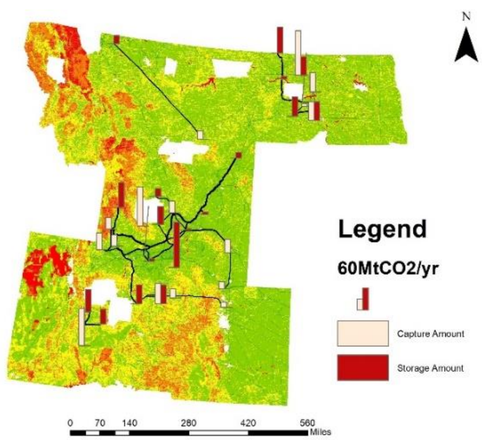
	MT	Colstrip	1.82	Kevin Dome	1.82
	ND	Great Plains Gasification, Milton R. Young, Coyote Station, Coal Creek	15.04	Great Plains Gasification, Milton R. Young, Beaver Lodge, Red Trail Energy	12.44
	UT	Hunter	8.10	Woodside Dome, Gordon Creek	8.10
	WY	Shute Creek, Riley Ridge, Naughton, Lost Cabin, Laramie River, Jim Bridger	17.77	Grieve Field, Bairoil Field, Beaver Creek, Monell Field, Moxa Arch, Hamilton Dome	19.27
	Ext	-	0.00	Canada	2.60



(a)



(b)



(c)

Figure 4.21: CCUS infrastructure network for 30 years in dynamic environments for Scenario 12 with graduated CO_2 capture goals with existent infrastructure: (a) 20 Mt CO_2/yr for 2019-2029; (b) 40 Mt CO_2/yr for 2029-2039; (c) 60Mt CO_2/yr for 2039-2049.

From the overall analysis of dynamic optimization, it can be seen that the overall costs of development of each CCUS network development is relatively cheaper as compared to their static counterparts. This is majorly attributed to scattered capital investment and slower development of operational capacities ensuring a better economic return. The comparison of cases with and without pre-existent networks reveals a similar financial trend in terms of cost per ton of CO_2 captured.

Table 4.25: Summary of cost analysis for dynamic optimization with existing infrastructure (scenario 11 and 12).

Scenario	Amount Captured ($MtCO_2$/yr)	Overall Cost ($/ton)	Capture Cost ($/ton)	Transport Cost ($/ton)	Storage Cost ($/ton)
11	30.18	30.68	28.71	0.61	1.36
12	40.20	31.37	28.99	1.05	1.32

4.7 text Analysis

Text analysis is the part of the CCSHawk workflow which deals with regulations. It extracts and processes each clause within the regulatory text in order to represent the clause in a meaningful manner. The goal of the analysis is to present the regulatory data in manner suitable for cognizant understanding for the user. The text analysis is done for CFR 195 Section 45 of EPA regulations, which relates to transportation of hazardous liquids by pipeline. This section covers transportation of CO_2 by pipeline on federal land and the outer-continental shelf of the USA.

The workflow related to regulation text mining has to be conducted separately on each regulatory clause due to the heterogenous structure of the data. This section will emphasis the work done in text mining for the clause 250 of CFR 195 Section 4.5 which describes the details for the clearance between pipelines and underground structures. The first step of the text analysis is the XML parsing. XML parsing has to be done over the whole document where each clause is extracted into a comma delimited spreadsheet. This is done due to nature of the xml document being hierarchically categorized for text parsing and further application. XML parsing is shown is Figure 4.22, where the structure of the clause is extracted as *header* and *text* for further processing.


```
<DIV8 N="§ 195.250" NODE="49:3.1.1.2.11.4.20.21" TYPE="SECTION">

<HEAD>§ 195.250   Clearance between pipe and underground structures.</HEAD>

<P>Any pipe installed underground must have at least 12 inches (305 millimeters) of clearance between the outside
of the pipe and the extremity of any other underground structure, except that for drainage tile the minimum
clearance may be less than 12 inches (305 millimeters) but not less than 2 inches (51 millimeters). However, where
12 inches (305 millimeters) of clearance is impracticable, the clearance may be reduced if adequate provisions are
made for corrosion control.

</P>

<CITA TYPE="N">[Amdt. 195-22, 46 FR 38360, July 27, 1981, as amended by Amdt. 195-63, 63 FR 37506, July
13, 1998]Original

</CITA>

</DIV8>
```


Original Text

Section	sub 1	sub 2	sub 3	sub 4	Header	Text
250	0	0	0	0	Clearance between pipe and underground structures.	Any pipe installed underground must have at least 12 inches (305 millimeters) of clearance between the outside of the pipe and the extremity of any other underground structure, except that for drainage tile the minimum clearance may be less than 12 inches (305 millimeters) but not less than 2 inches (51 millimeters). However, where 12 inches (305 millimeters) of clearance is impracticable, the clearance may be reduced if adequate provisions are made for corrosion control.

Figure 4.19: XML Parsing to convert Original XML text to actionable spreadsheet format.

The actionable spreadsheet format data obtained from XML parsing is further used to extract individual clauses to prepare the text for syntactic natural language Processing (NLP). The text preparation steps consists of three steps: sentence tokenization, word tokenization and stemming. Sentence tokenization recognizes individual sentences from a text corpus (body of text). Figure 4.23 shows text tokenization where part A, original text, is converted to part B, individual sentences. Word tokenization separates each word, symbol, and numerical value in the sentence. In Figure 4.23, word tokenization is indicated by conversion of part B to part C. Finally stemming is the conversion of a body of text into its constituent base words indicated by the conversion of Part C to D in Figure 4.23. All 3 sections of text preparation require extensive training and hence is conducted using a python library known as ‘nltk’ on Python 3.2. The libraries are pre-trained on large datasets, which provides a highly accurate execution of text preparation process.

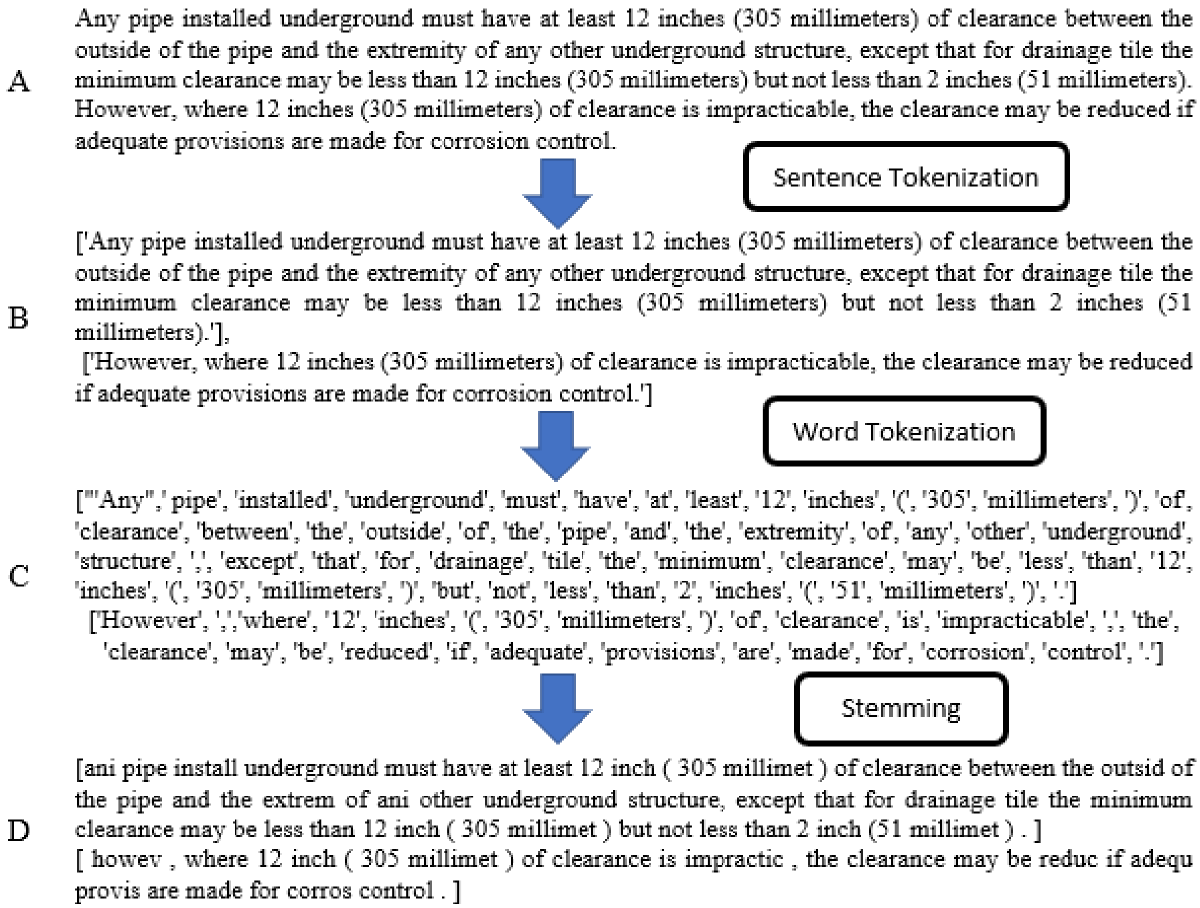


Figure 4.20: Text preparation. A: Original text; B: Text with sentence tokenization; C: Text with word tokenization; D: Text with stemming.

The prepared text is used for syntactic analysis in the text processing process where the grammatical analysis of a sentence is conducted using NLP procedures. The text processing has two parts: Part-of-speech analysis (POS analysis) and depednancy parsing. In POS analysis each word in a sentence is analyzed to find their grammatical implication in a statement in terms of components such as verbs, nouns and pronouns. Figure 4.24 shows POS analysis from part A, original text, to part B, POS analyzed. Dependency parsing takes the POS analysis and check the syntactic depednancy of each word on one another in the context of the statement. Figure 4.24 shows the depednancy parsing of POS analyzed text from part B to part C. POS analysis uses *nltk* library on Python, which is heavily pre-trained by the developers to recognize the context of each word, while depednancy parsing uses *StanfordCoreNLP* library on Python.

A

['Any pipe installed underground must have at least 12 inches (305 millimeters) of clearance between the outside of the pipe and the extremity of any other underground structure, except that for drainage tile the minimum clearance may be less than 12 inches (305 millimeters) but not less than 2 inches (51 millimeters).'],
['However, where 12 inches (305 millimeters) of clearance is impracticable, the clearance may be reduced if adequate provisions are made for corrosion control.']

B

[("'Any", 'POS'),('pipe', 'NN'),('installed', 'VBD'),('underground', 'JJ'),('must', 'MD'),('have', 'VB'),('at', 'IN'),('least', 'JJS'),('12', 'CD'),('inches', 'NNS'),('(', '('),('305', 'CD'),('millimeters', 'NNS'),(')', ')'),('of', 'IN'),('clearance', 'NN'),('between', 'IN'),('the', 'DT'),('outside', 'NN'),('of', 'IN'),('the', 'DT'),('pipe', 'NN'),('and', 'CC'),('the', 'DT'),('extremity', 'NN'),('of', 'IN'),('any', 'DT'),('other', 'JJ'),('underground', 'JJ'),('structure', 'NN'),(',', ','),('except', 'IN'),('that', 'DT'),('for', 'IN'),('drainage', 'NN'),('tile', 'NN'),('the', 'DT'),('minimum', 'JJ'),('clearance', 'NN'),('may', 'MD'),('be', 'VB'),('less', 'JJR'),('than', 'IN'),('12', 'CD'),('inches', 'NNS'),('(', '('),('305', 'CD'),('millimeters', 'NNS'),(')', ')'),('but', 'CC'),('not', 'RB'),('less', 'JJR'),('than', 'IN'),('2', 'CD'),('inches', 'NNS'),('(', '('),('51', 'CD'),('millimeters', 'NNS'),(')', ')'),('.', '.')],
[("'However", 'RB'),(',', ','),('where', 'WRB'),('12', 'CD'),('inches', 'NNS'),('(', '('),('305', 'CD'),('millimeters', 'NNS'),(')', ')'),('of', 'IN'),('clearance', 'NN'),('is', 'VBZ'),('impracticable', 'JJ'),(',', ','),('the', 'DT'),('clearance', 'NN'),('may', 'MD'),('be', 'VB'),('reduced', 'VBN'),('if', 'IN'),('adequate', 'JJ'),('provisions', 'NNS'),('are', 'VBP'),('made', 'VBN'),('for', 'IN'),('corrosion', 'NN'),('control', 'NN'),('.', '.')]

C

{(ROOT(S(NP(NP (DT Any) (NN pipe))(VP (VBN installed)(NP (NN underground))))(VP (MD must)(VP (VB have)(NP(NP(QP (IN at) (JJS least) (CD 12))(NNS inches))(PRN (-LRB- -LRB-)(NP (CD 305) (NNS millimeters))(-RRB- -RRB-))(PP (IN of)(NP (NN clearance))))(PP (IN between)(NP(NP(NP (DT the) (NN outside))(PP (IN of)(NP (DT the) (NN pipe))))(CC and)(NP(NP (DT the) (NN extremity))(PP (IN of)(NP (DT any) (JJ other) (JJ underground) (NN structure))))))(, ,)(PP (IN except)(NP(NP (DT that))(PP (IN for)(NP(NP (NN drainage) (NN tile))(SBAR(S(NP (DT the) (JJ minimum) (NN clearance))(VP (MD may)(VP (VB be)(NP(NP(NP(QP (JJR less) (IN than) (CD 12))(NNS inches))(PRN (-LRB- -LRB-)(NP (CD 305) (NNS millimeters))(-RRB- -RRB-)))(CONJP (CC but)(RB not))(NP(NP(QP (JJR less) (IN than) (CD 2))(NNS inches))(PRN (-LRB- -LRB-)(NP (CD 51) (NNS millimeters))(-RRB- -RRB-))))))))))))))(. .)))},
{(ROOT(S(ADVP (RB However))(, ,)(SBAR(WHADVP (WRB where))(S(NP(NP(NP (CD 12) (NNS inches))(PRN (-LRB- -LRB-)(NP (CD 305) (NNS millimeters))(-RRB- -RRB-)))(PP (IN of)(NP (NN clearance))))(VP (VBZ is(ADJP (JJ impracticable)))))(, ,)(NP (DT the) (NN clearance))(VP (MD may)(VP (VB be)(VP (VBN reduced)(SBAR (IN if)(S(NP (JJ adequate) (NNS provisions))(VP (VBP are)(VP (VBN made)(PP (IN for)(NP (NN corrosion) (NN control))))))))))(. .)))}

Figure 4.21: Text Processing. A: Original text; B: Text with POS analysis; C: Text with dependency parsing.

Information repetition involves the representation of the analyzed text in a manner suitable for analysis. This step involves tagging of clauses with *tags* and *subtags* which are manually assigned. Along with the representation of each clause in a fashion which makes it easy to understand and utilize for future purposes. The list of all tags and respective sub tags assigned to the clauses of the Subsection 195 is provided in Table 4.26.

Table 4.20: List of tags and sub-tags used in information extraction of the regulations.

List of Tags	List of Sub Tags
Scope, Pipelines covered, Definition, Pipe Compatibility, Qualification for regulation, Pipeline Construction, Pipeline Qualification, Pipeline Responsibility, Subpart Scope, Pipeline Reporting, Pipeline Assignment, Pipeline Information, Pipeline Design, Valve Design, Fitting Design, Flange connection, Station Piping, Fabricated assemblies, Breakout Tank Design, System requirement, Pipeline Inspection, Pipe Transportation, Welding, Pipeline Corridor, Pipe Bends, Welding Joint, Welding Arc Burn, Surface Equipment, Valves, Pumping Station, Pipeline Testing, Breakout Tank Testing, Tie-in Testing, Pipeline Operations, Breakout Tanks Operations, Valve Maintenance, Pipeline Repairs, Safety Device Operation, Pumping Station Safety, Breakout Tank Inspection, Breakout Tank Safety, Pipeline Systems, Pipeline Coating, Pipeline Corrosion, Breakout Tank Corrosion	Material, Test, In-line Inspection, Offshore, Mapping, Accidents, Safety, Submission, Abandonment and Deactivation, Obligation, Temperature, Pressure, External Load, Fracture Propagation, Fabricated branch connections., Closures., Capacity, CPM, Diameter, Personnel, Records, Weather, Repairs, Ditches, Excavation, Crossings, Alternatives, Medium, Training, Communications, Line Markers, Right-of-way, In-line Inspection, Signs., Public awareness, Damage Prevention Program, Control Room, SCADA, High Consequence Regions, Qualification, Test Leads, Corrosion, Electric Current

4.8 Conclusion

This chapter reveals the effect of various pre-existent infrastructure, environment and ecology-based parameters on the development of CO_2 pipeline corridors and CCUS infrastructure networks. It can be seen that no single environmental or ecological factor can heavily sway the orientation of pipeline corridors, but rather it is the effect of the commingling of different parameters. Factors such as existent pipeline routes, land use, slope and lakes have a higher correlation with the presence of pipeline corridors, rather than certain factor such as urban areas, roads, railways, and waterways.

Further, from the network analysis, the cost of the CCUS network per ton of CO_2 is lower initially due the utilization of natural gas plants and commercial vendors. However, with the increasing capture goals related the CCUS networks, the cost per ton of CO_2 emission increases with the usage of coal-fired plants. It is also seen that saline aquifers are preferred over EOR sites because of the cost related to purchasing CO_2 for utilization. Sites with higher injectivity of CO_2 per well are preferred to those with lower injection rate due to the costs related to individual wells. Further, it is seen that transportation networks do not have a large effect on the overall CCUS operations due to relative cost of CO_2 capture being higher than the rest of the CCUS value chain. There are further effects of individual factors such as inflation rates and cost of CO_2 that are checked in the following chapters.

CHAPTER 5

Discussions: Variations in Input Parameters

5.1 Introduction

Carbon Capture, Utilization, and Storage is one of the leading techniques utilized for transitioning the energy sector to a low carbon economy. It can help the conservative energy sector meet the regulatory requirements for CO_2 emissions. However, CCUS has several components in its value chain and each component is capital intensive. Proper planning and decision-making are critical to evaluate the economic feasibility of each component in the CCUS value chain. With the use of CCSHawk, this thesis puts forth a decision-making framework for CO_2 pipeline corridor selection and CCUS infrastructure deployment.

The goal is to determine the optimal economic scenarios related to the deployment of CCUS infrastructure in North-Central USA with a focus on coal-fired plants, natural gas plants, saline aquifers and EOR activities. In addition, the development of pipeline corridors is investigated with special attention to the effects of environment and ecology on pipelines. To achieve these goals, Chapter 3 put forth a multi-step workflow which develops an infrastructure network to capture the economic impact of CCUS. Chapter 4 shows that it is possible to calculate the cost of preventing CO_2 emissions through CCUS . This chapter checks the economic impact of varying factors in the CCUS network. This chapter also explores the effect of changing AHP weightages on the identification of pipeline corridors.

In this chapter various factors related to the creation of cost map will be varied to check the effect on pipeline corridors in Section 5.2. Section 5.3 will check the effect of variations in parameters related to clustering, creation of Delaunay pairs and least cost path (LCP) generation. Section 5.4 will discuss the economic effect of parameter variation on the static and dynamic formulation of the network problem. As tax savings and increased earning is the driving motivator for adoption of CCUS technology, the effect of addition of these factors in the network formulation will be investigated in Section 5.5.

5.2 Effect of Parameter Variation on Cost Map Generation

Cost map analysis is an important part of the CCSHawk workflow, as it accounts for terrain, environment, and ecology of the study area. The results of this analysis influence the route generation process between two points in the study area. A degree of measure needs to be established to understand the suitability of a tract of land for pipeline development derived from the cost map.

A Tract Suitability Index (TSI) was generated with the cost map to indicate lands most suitable for pipeline development as seen in Balaji et al., (2020). The TSI index was assigned according to classification of cells according to the *cost* of each cell, as indicated by Table 5.1 into 5 categories. The reclassification of the raster file was performed using the “Reclassify” tool in ArcGIS 10.7.1. According to the developed TSI index, 1 is the most suitable and 5 is the least suitable for development of pipeline corridors. The study area classified by the TSI index is shown in Figure 5.1. The map utilized for generation of Figure 5.1 is obtained from the cost map developed from Table 4.5. The TSI index classification developed from Table 4.5 is called the base case. From the TSI analysis, it can be seen that only 2.68% of the land within the study area is classified as highly suitable for pipeline development. These regions are distributed through the central regions of Wyoming which have low slopes, many pre-existing pipelines and few ecological hotspots In addition, 51.82% of the whole study area is favorable overall towards pipeline development mainly in North Dakota, Eastern Montana and Eastern Colorado. Central Montana, western North Dakota and Utah have lands moderately suited for pipelines amounting to 20.12% of the whole study area. 11.76% of the study area has low suitability to pipelines which mostly surrounding regions deemed as extremely poor regions for pipeline development (at 3.63%). These unsuitable regions are mostly around lakes and regions of high altitudes near western Colorado, western Montana, and north-western Utah (Balaji et al., 2020).

Table 5.1: Cost ranges for development of tract suitability index (TSI) map.

Cost Map	Tract Suitability Index Value	Tag
0 - 0.4	5	Highly Unsuitable
0.4 - 0.5	4	Unsuitable
0.5 - 0.6	3	Moderately Suitable
0.6 - 0.7	2	Suitable
0.7 - 1	1	Highly Suitable

To check the effects of the analysis, a sensitivity study was done by changing the weightages related to Natural, Right-of-way and regulatory barriers (Balaji et al., 2020). The weightages for the base case, were shown in Table 4.5, in the previous Chapter. For case 2, equal weights were given to all criteria according to Table 5.2. In Case 2 significant changes were seen in the areas with high slope (Figure 5.2). Regions around lakes were generally unsuitable for pipelines, however with lower weight assignments these regions became more favorable for pipeline construction. Due to the increased weightages to protected land, the region around the Salt Lake became more unsuitable for pipelines.

Case 3 involves giving higher priority to right-of-way as compared to natural and regulatory barriers as shown in Table 5.3 (Balaji et al., 2020). Figure 5.3 depicts the Case3, where it can be seen that most regions are classified as moderately suitable for pipelines. This change in behavior is due to the higher weightage provided to regulatory consideration, increasing the effects of land use. In this case too the most suitable regions of pipelines is in the central region of Wyoming.

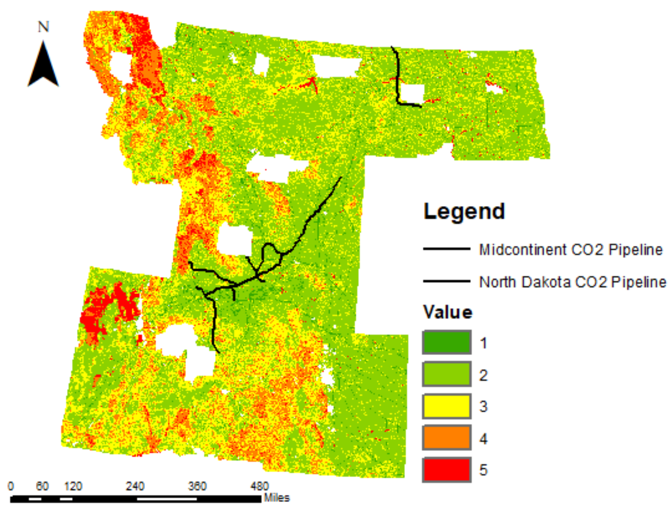


Figure 5.1: CO_2 pipeline tract suitability index (TSI) – base case (Balaji et al., 2020).

Table 5.2: Pairwise comparison of barrier criteria for case 2.

	Natural	Regulatory	ROW	Weights
Natural	1	1	1	0.34
Regulatory	1	1	1	0.34
ROW	1	1	1	0.34

Table 5.3: Pairwise comparison of barrier criteria for case 3.

	Natural	Regulatory	ROW	Weights
Natural	1	1	0.5	0.25
Regulatory	1	1	0.5	0.25
ROW	2	2	1	0.5

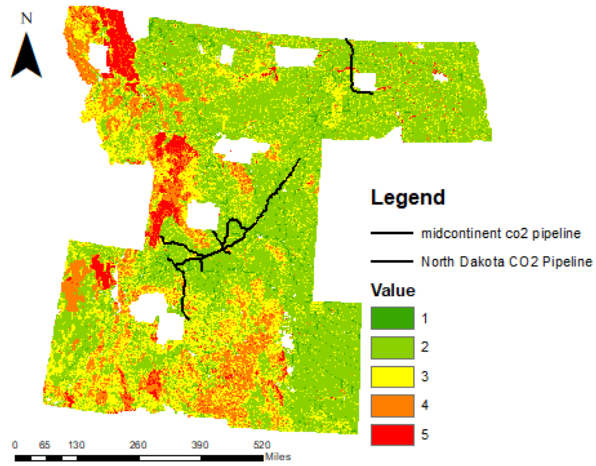


Figure 5.2: CO_2 pipeline tract suitability index (TSI) – case 2 (Balaji et al., 2020).

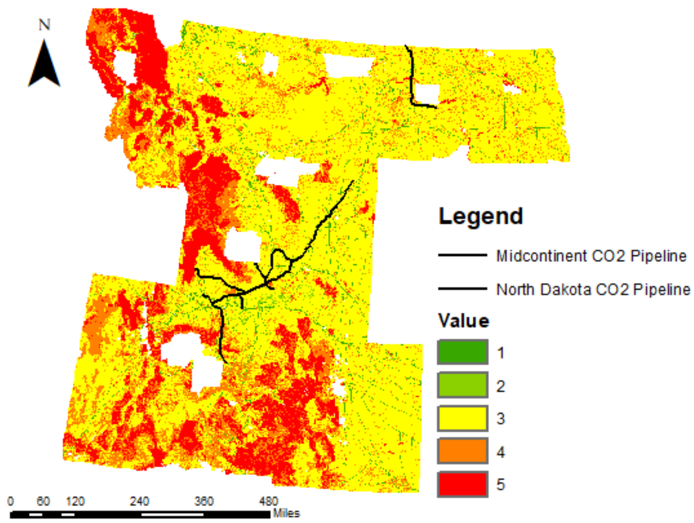


Figure 5.3: CO_2 pipeline tract suitability index (TSI) – case 3 (Balaji et al., 2020).

A comparison between the three case studies, reveals that the results does not change much in terms of general regions which favor the development of pipelines: North Dakota, central Wyoming and eastern Montana. The following geo-information layers had a good correlation with regions suitable for pipelines: existing pipelines, slope and land use. Similarly, geo-information layers with correlation with low suitability for pipelines include lakes, Areas of Critical Environmental Concern (ACEC zones), parks, slopes, fault zones and protected land. These correlations are found despite the relative weightage of the geo-information layers. Of the three cases, case 3 had a major difference because of the increased weightage of land use. Parks too have a significant effect in case 3, however due to parks laying in the same region as unfavorable spots for pipeline, the increased weightage to parks is not reflected in the analysis. For case 2, the changes in terms of regions suitable for pipeline development remain virtually the same as the base case, except that the regions surrounding lakes are more favorable for pipe, while on the other hand, regions near parks and unusually sensitive areas, become more unfavorable (Balaji et al., 2020).

If regulatory barriers were given a higher weightage as compared to natural or right-of-way barriers, the analysis would be similar to Case 2, except for having a greater importance for population centers, leading to small regions of un-favorability distributed throughout the study area. These small regions are dispersed in nature due to population centers being represented by schools. There are regions where the density of schools would increase, however, these regions would be around cities and urban regions, which are left out of the study area. Similarly, ecological hotpsots (ACEC zones, WHRSN sites and protected areas) would be extremely unfavorable for pipelines which are already classified as high cost regions in other cases (Balaji et al., 2020).

The analysis and generation of the cost map for CO_2 pipeline suitability is specific to the study area. If a similar cost map analysis for pipeline suitability were to be conducted in a different region of the USA, many of the geo-information layers to be considered would remain the same (Balaji et al., 2020). The layers to differ would be local factors such as topsoil frost susceptibility. The importance of factors such as population distribution would vary significantly according to the region of study, for instance in the coastal regions of USA would have much higher population concentration, reducing the importance of the geo-information layer. Some factors such as fault zones would play an important role in the Western region of USA, which sees more seismic activity overall. The south-eastern region of USA would have to consider local environmental factors known as "Act of God" which includes hurricanes and other natural disasters (Menon, 2011). The role of ecological factors should remain constant through the USA. Similarly, the cost of land would vary with the region in consideration with heavier costs seen near the coastal regions of USA, affecting the suitability of the land towards pipelines. Thus, individual analysis needs to be done according to varying study areas, as there is no constant considerations.

5.3 Effect of Parameter Variations on Route Analysis

CO_2 pipelines are an important part of the CCUS networks, serving as the medium for material flow. However, it is essential to check for parameters which would most likely influence the route generation process. The most important factors affecting pipeline routes are the radius of clustering, number of node points, and location of node points. Another important factor influencing the location of pipeline routes is the cost map generation, which was discussed in the previous section.

In the case of clustering, the computational load and the number of probable pipeline arcs between the nodes significantly increase, as the number of nodes in the study increase. The clustering technique reduces the number of nodes from 56 original nodes points to 40 nodes, using a clustering diameter of 24.24km. As seen in Figure 5.4, the number of nodes in the study area reduces with the increase in cluster diameter. It should be noted, that increasing the cluster diameter will decrease computational load, but it would lead to deviation from the reality of the case study, especially in terms of reservoir properties for fields that are distant from each other.

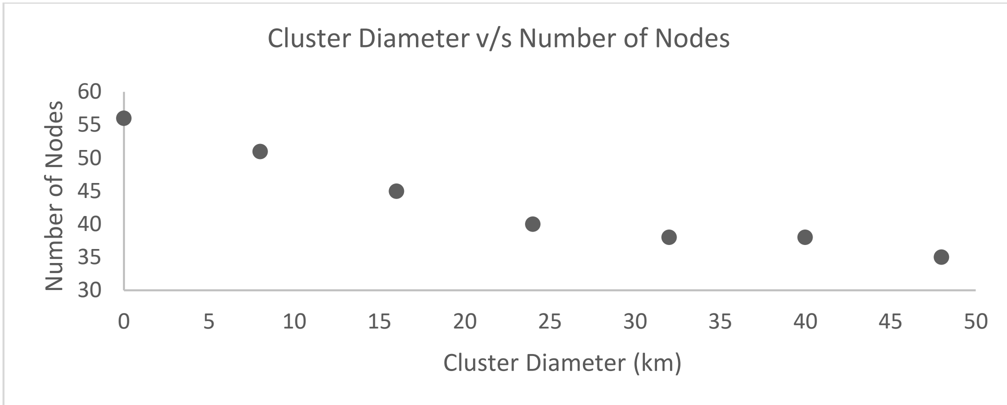


Figure 5.4: Decrease in number of nodes with cluster diameter.

Clustering also affects the number of probable pipeline routes as the number of clustered nodes influences the selection of Delaunay pairs. The number of Delaunay pairs is the number of probable pipeline arcs in the study area between the nodes. By analyzing the effect of cluster diameter on number of nodes, it is possible to see its effect on the number of probable pipeline arcs in the study area. From Figure 5.5, which shows the trend between number of clustered nodes and probable pipeline arcs in the study area, it is clear that as the number of clustered nodes increases, the number of pipeline arcs steadily increases. However, the rate of increase of number of candidate pipeline arcs is not homogeneously affected by the number of clustered nodes. The main reason for this is the effect of relative position of each node to the other. Delaunay algorithm ensures that no node would be circumscribed within another Delaunay triangle. This effect is highly influenced by the relative position of the nodes. However, if the clustered points were to be equidistant and homogeneously distributed in the study area, a linearly increasing trend between the number of nodes and probable pipeline arcs would be derived.

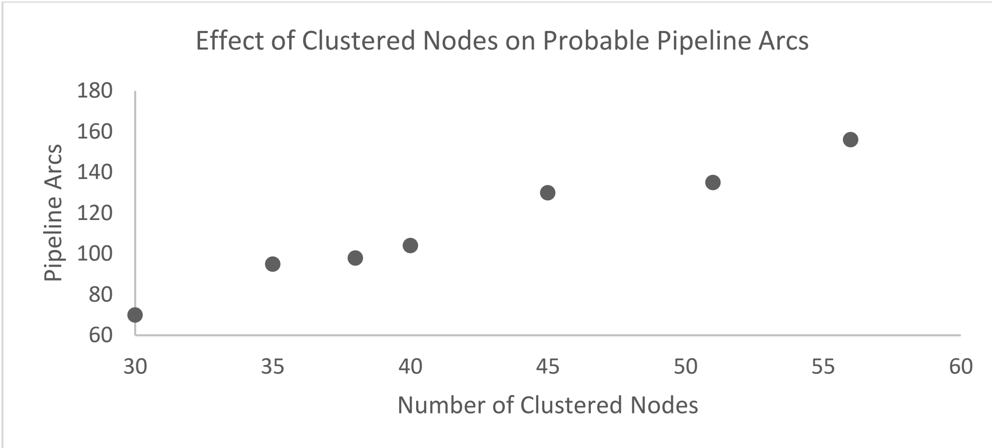


Figure 5.5: Effect of number of clustered nodes on number of probable pipeline arcs.

The routing algorithm used in CCSHawk is the A-star algorithm. However, most studies in literature have used the Dijkstra algorithm or the straight-line LCP to route pipelines from one point to another. In terms of length of pipeline and reduction of cost, straight line LCP is the most conducive pathing algorithm. However, as discussed in Chapter 2, no pipeline can have a perfectly straight pipeline from source to destination. The A-star algorithm is a goal-oriented version of the Dijkstra algorithm. Figure 5.6 shows a comparison of execution of Dijkstra algorithm compared to A-star algorithm to generate a path between two points. As can be seen from the figure the paths take different routes to their destination due to variation in pathing priorities. However, there is a difference between the length of two path, where A-star algorithm generated LCP is shorter than the Dijkstra algorithm generated LCP by 6.23%. Overall, there is a reduction in pipeline length between the A-star algorithm as compared to the Dijkstra algorithm in most cases. The reduction in length is not constant and varies from case to case, however, the average reduction in length is about 4.23%. The amount of reduction in lengths is due to a combined effect of cost raster and location of source and sink.

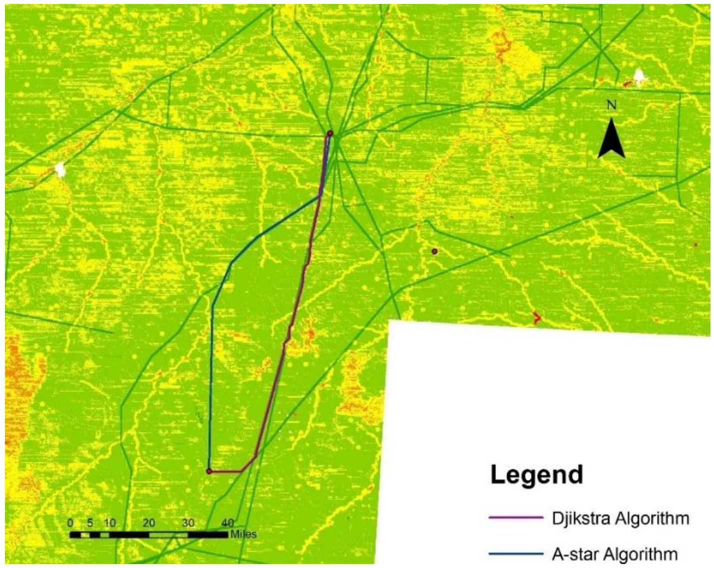


Figure 5.6: Dijkstra algorithm route as compared to A-Star algorithm generated route.

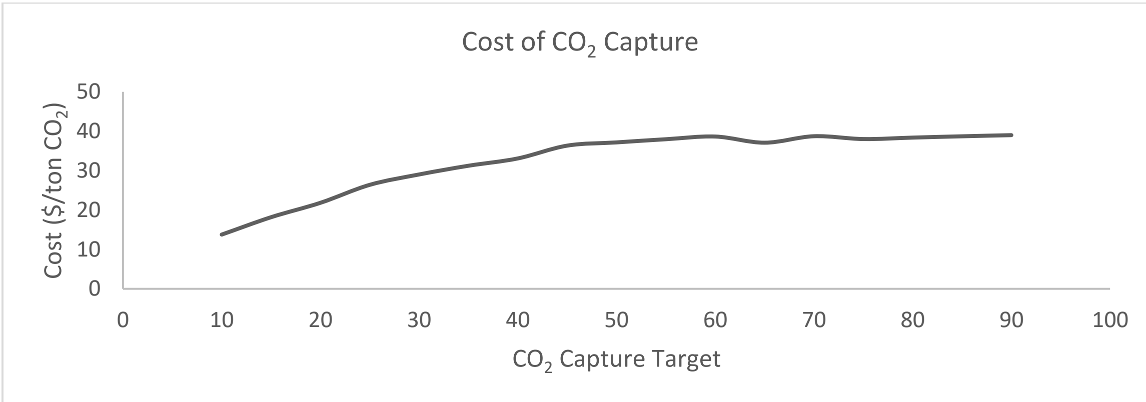
In regards to control parameters in the route generation process, the major influencer is the cost map and location of source and destination. However, from the sensitivity analysis from Section 5.1, it can be seen that despite changes in the weightages related to generation of cost map, the trend of regions having favorability towards pipeline sections remains uniform. Routes generated by varying weightages of cost map would be insignificant, as the comparative weightages between cells, do not change significantly. Thus, route generation process between different cost maps would vary minimally as the routing is also heavily dependent on the relative location of the nodes.

5.4 Impact of Parameters on Decision-Analysis

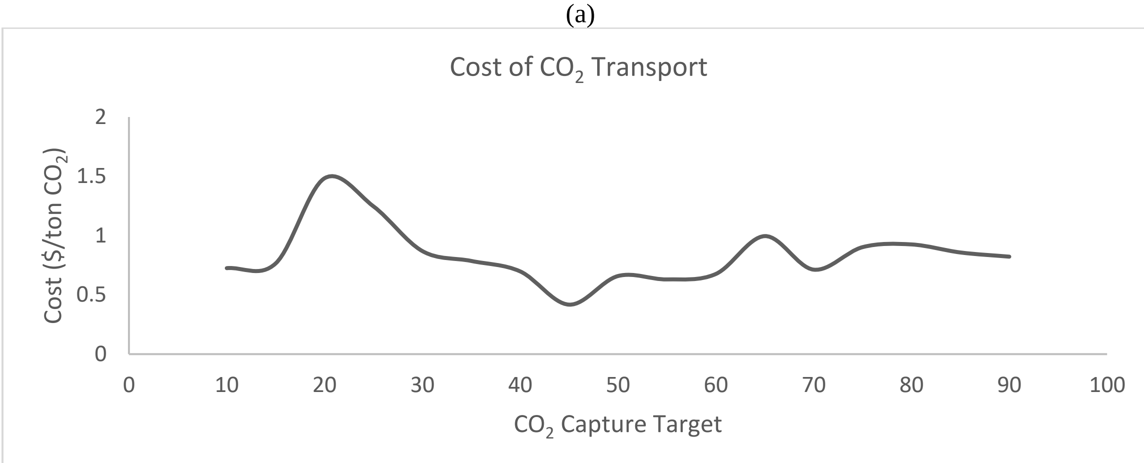
From the optimization of the CCUS network several queries related to selection of specific elements of the network for deployment, levels of activity and cost for deployment and operation can be answered. Chapter 4 discussed the deployment of CCUS network under both static and dynamic deployment formulations. The effect of variation of various control parameters in the optimization models is important to see their effect on the overall CCUS infrastructure cost.

5.4.1 Cost Variation with CO_2 Capture Goals

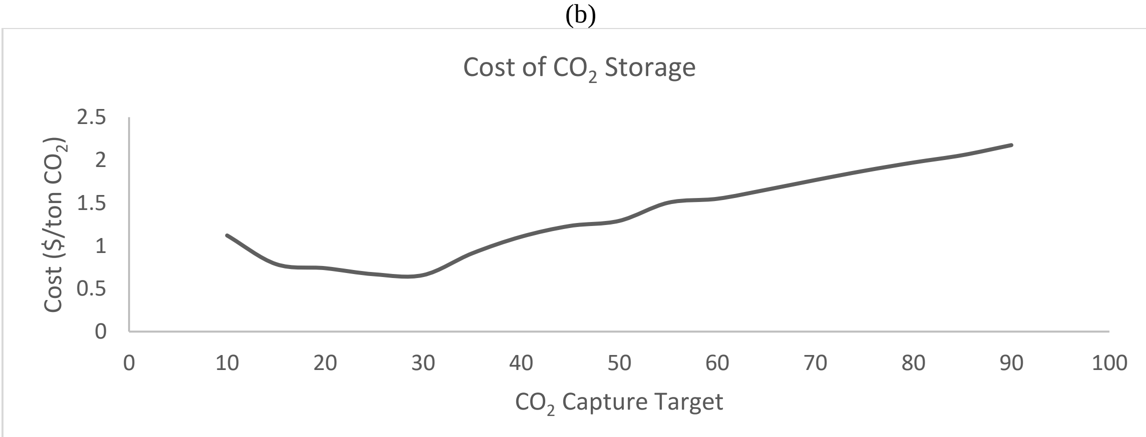
It is crucial to analyze the costs related to development of CCUS networks, as it can reveal several interesting trends related to choices made by the optimization tool in terms of sources, sinks and pipelines in the CCUS network. From Figure 5.7a, it can be observed that the overall cost of CO_2 capture (per ton of CO_2) increases with increasing capture goals from 10 to 90 $MtCO_2$/yr. The major influence for such behavior is related to the transition of capturing CO_2 from only natural gas fired plants to plants involving other type of capture methods. Figure 5.7b displays the cost of transporting a ton of CO_2 in relation to varying capture goals. It can be observed that there is no general trend related to the costs of transporting CO_2, however, there is a slight unsteady increase in costs. The sudden fluctuations in transport cost can be attributed to increasing CCUS infrastructure dimensions. Initially with lower CO_2 capture volumes, there are many 12-inch pipelines that are underused. As the amount of CO_2 in the system increases, there is a better utilization of the 12-inch pipeline infrastructure. But as the capture goals further increases beyond 40 $MtCO_2$/yr, the number of pipelines with higher diameter increase and the cycle continues. Similar to the cost related to capture, the cost of storage/utilization also increases with the capture goals, due to increasing proportion of usage of reservoir with low injectivity, encouraging the usage of larger number of wells. However, there is a spot in figure 5.7c where there is a dip in the cost related to CO_2 storage (near 30 $MtCO_2$/yr) due to increased usage of saline aquifers.



(a)



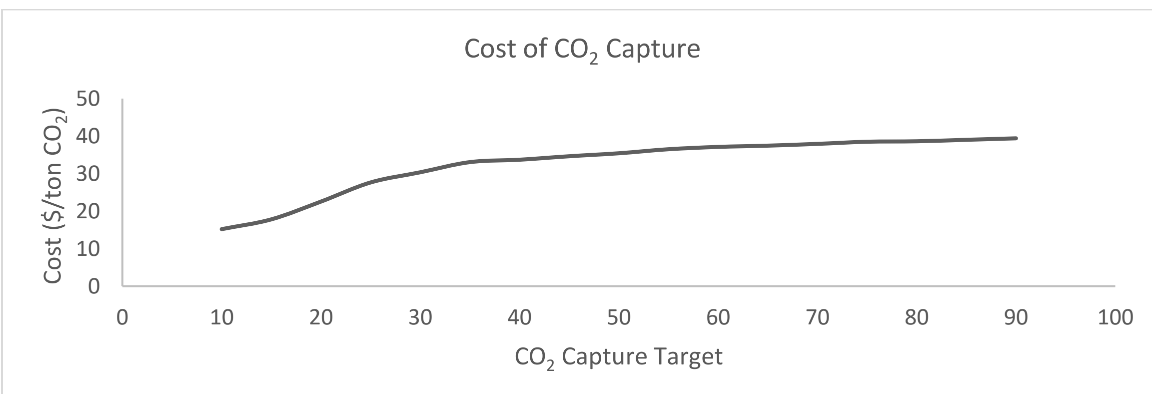
(b)



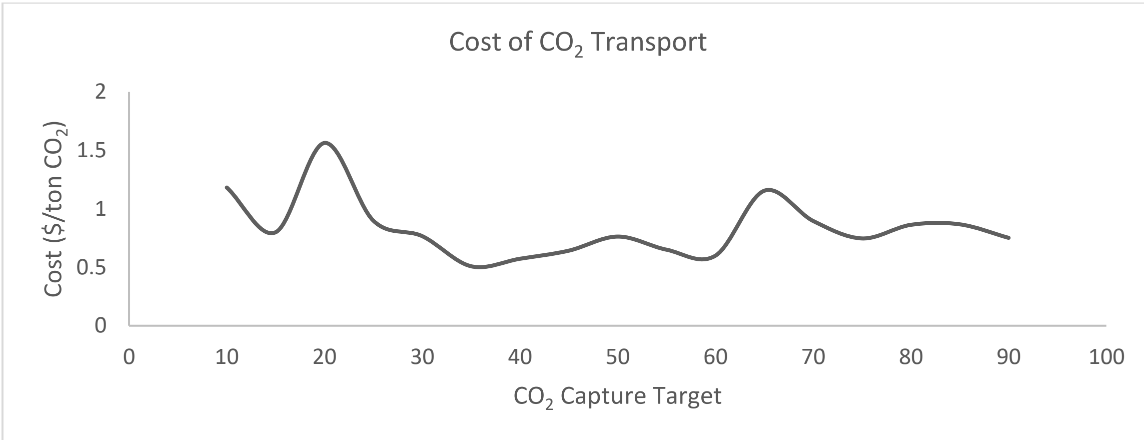
(c)

Figure 3.7: Cost per ton of CO_2 avoidance with no pre-existent infrastructure (a) Capture (b)Transport (c)Storage.

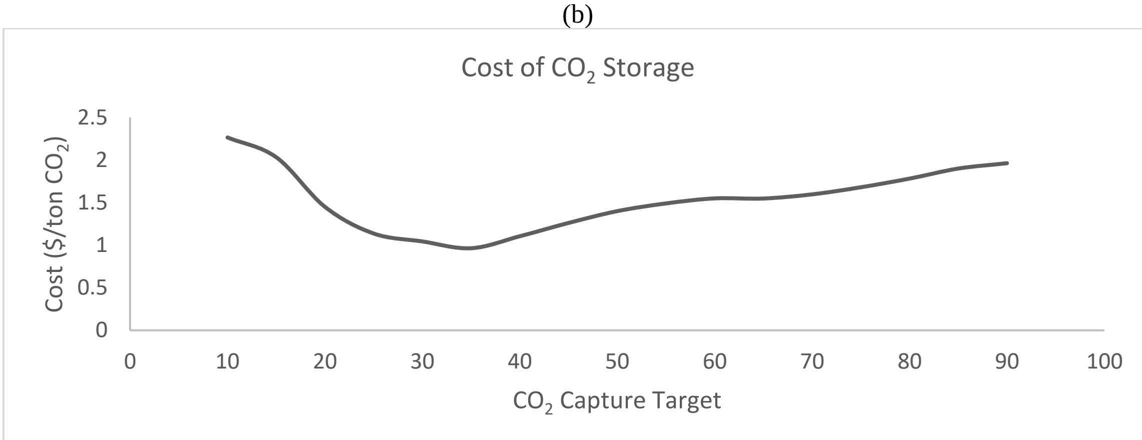
It is essential to analyze CCUS networks with pre-existing CCUS infrastructure for their cost trends with increasing CO_2 capture goals. Figure 5.8 depicts the various trends related to the CCUS value chain with pre-existing infrastructure. Figure 5.8a shows the increasing capture cost of CO_2 with increasing capture goals similar to Figure 5.7a. There is no significant difference in type of capture goals as the pre-existent infrastructure is based around natural gas plants and CO_2 commercial vendors. Figure 5.8b reveals the trend of pipeline cost per ton of CO_2 avoided, where there is a slight trend of increasing cost, however, the trend is cyclic in nature. The initial large increase in pipeline cost is due to the usage of resources which are not related to pre-existent infrastructure; however, the general cost of pipeline is not as high as those seen in Figure 5.7b. This can majorly be attributed to the larger utilization of pre-existing pipelines in later stages of CCUS network development with higher CO_2 capture goals. Similarly, we can see an initial high cost related to CO_2 storage in Figure 5.7c as the pre-existent infrastructure uses EOR for utilization rather than saline aquifers. Due to the cost related to the storage of each ton of CO_2 and eventual transition to saline aquifers as the capture goals increases, there is a decrease in cost related to CO_2 storage/utilization. However, the cost of storage again starts increasing around the 35MtCO_2/yr capture goal, as the usage of EOR activities starts to rise.



(a)



(b)

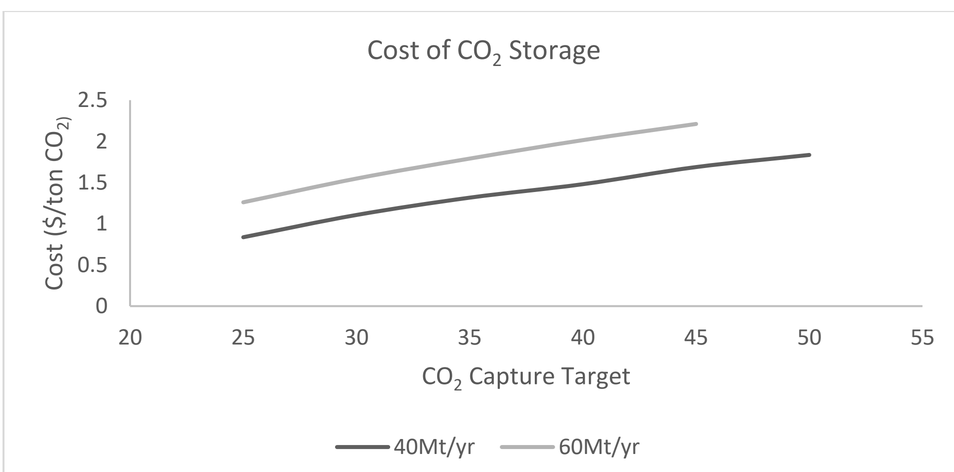


(c)

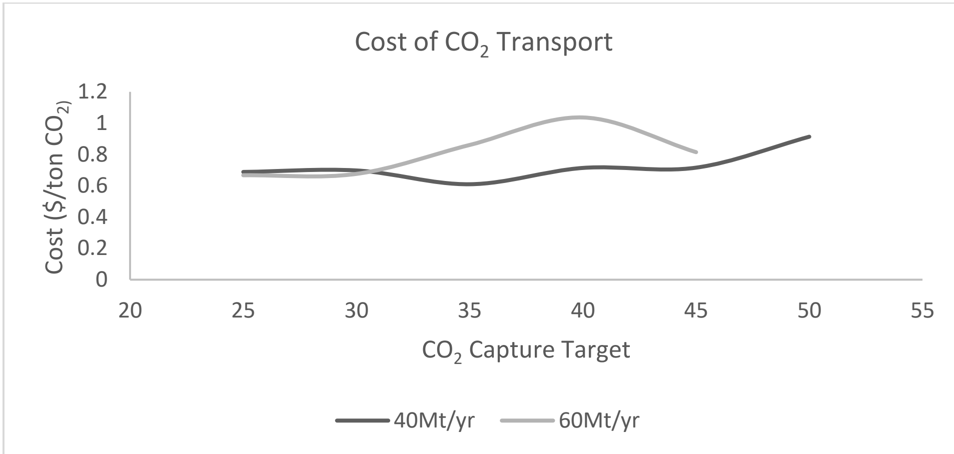
Figure 5.8: Cost per ton of CO_2 avoidance for CCUS network with pre-existent infrastructure (a) Capture (b) Transport; (c) Storage.

5.4.2 Cost Variation with Operating Periods

Operating period of the CCUS infrastructure is an aspect that can affect the overall cost related to the network. To investigate the effect of operating period on the overall cost, the period was varied from 25 to 50 years of operation for two different capture goals: 40MtCO_2/yr and 60MtCO_2/yr. On overall analysis of annual cost variation of CCUS network deployment, there was no significant trend except for a slight upward trend caused due to an increase in storage costs related to the network. Figure 5.9 depicts the cost variation with time period for the two different annual capture amounts. As can be seen from Figure 5.9a, there is an increase in cost related to storage with the extended operating period for both capture goals. This is due to the annual limitation of CO_2 being inversely proportional to the period of operations, as the reservoirs in the formulations are deisgned to last throughout the operating period. The goal of the constraint on reservoir capacity is to make each reservoir last for the entire period of operations. Hence, as the time period of operations increases, the relative annual storage allocation (in $MtCO_2$) for each reservoir decreases. The decreased annual storage for each reservoir, increases the number of reservoirs used in the study area and increases the usage of EOR reservoirs, which are more costly. This trend of increased usage of reservoirs is reflected in the overall cost increase for the network. Figure 5.9b shows the cost of transportation related to increasing periods of time. There is a slight increase in cost per ton due to the increased number of utilized pipelines due to the usage of a larger number of reservoirs. The effect of variation of operation period is mainly observed on the choice of storage/utilization reservoirs, under the condition that the term for repayment of capital expenditure is held constant at 30 years. If period of repayment on capital expenditure varies along with the operating period, then annual capital recovery would decrease with increasing operating period.



(a)



(b)

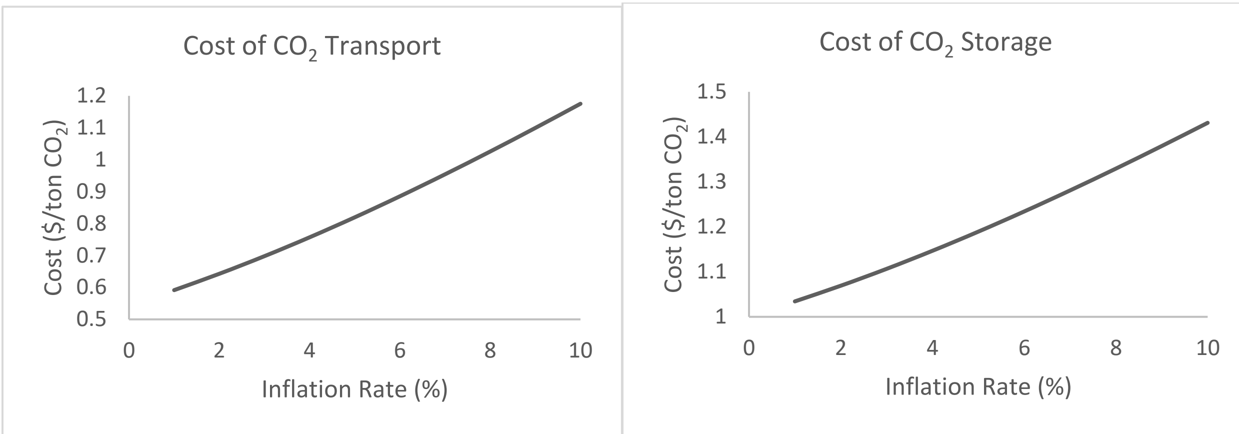
Figure 5.9: Effect of operating period on CCUS infrastructure cost for: (a) Storage; (b) Transportation.

5.4.3. Cost Variation with Inflation Rates

Inflation rate is a parameter which affects the capital costs related to operations, as the optimization formulation uses annualized cost of capital expenditure. The annualized cost is provided by Equation 5.1, where A is the annualized amount; P is the total capital expenditure; n is the time period and i is the rate of inflation.

$$A = P \frac{i*(1+i)^n}{(1+i)^n-1} \quad (5.1)$$

Figure 5.10a depicts the variation of cost related to CO_2 transportation at varying inflation rates for a CCUS infrastructure network aimed at capturing 40 $MtCO_2$/yr for a 30 year operating period. From Figure 5.10a, it is clear that the cost of pipelines increases with the increasing inflation rate as the time value of money increases. The effect of the inflation rate is limited to the capital costs related to the pipeline. Similarly, Figure 5.10b depicts the increasing cost of CO_2 storage/utilization with increasing inflation rate. The effect of inflation rate is only on cost and it does not affect the operations of the CCUS network.



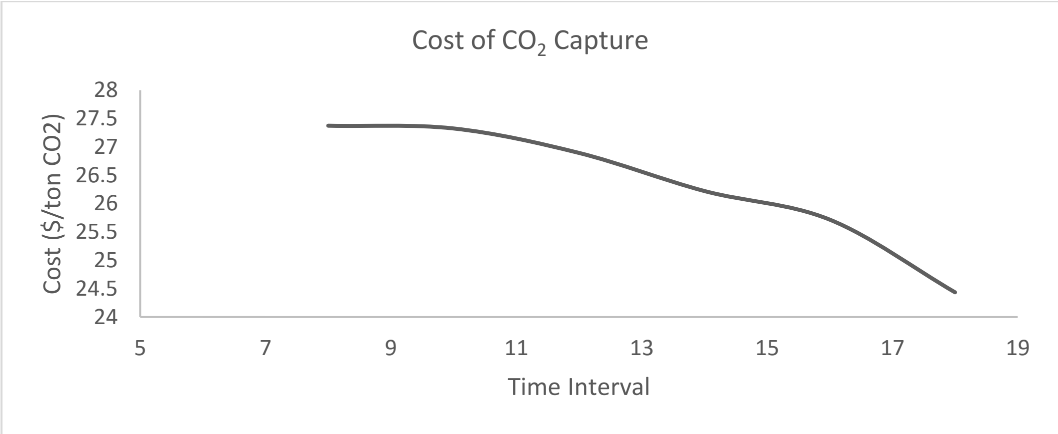
(a) (b)

Figure 5.10: Effect of variation in inflation rate on (a) cost of CO_2 Transport/ton (b) Cost of CO_2 Storage/ton.

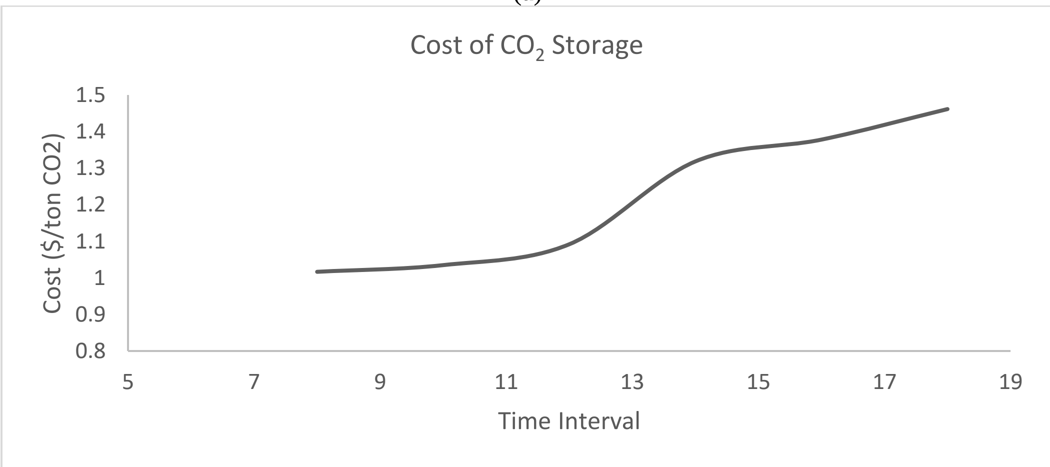
5.4.4 Cost Variation with Time Intervals

The period of operations can have an alternate effect on the dynamic formulation of CCSHawk. As the time of operation increases, the time intervals between each successive capture goal increases. Time interval variation can have a significant effect on annualized costs, as the capital expenditure is distributed across different time periods, and the storage-based limitations on each reservoir increases the number of reservoirs in usage.

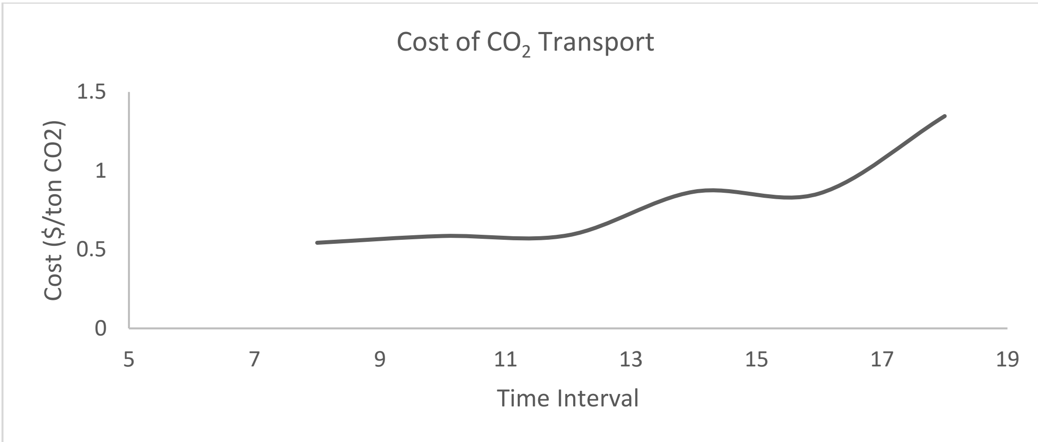
The operating period is varied from 24 to 54 years, which leads to individual time period variations between 8 to 16 years for the dynamic formulation to capture three different targets of 20, 30 and 40 $MtCO_2$/yr respectively. Figure 5.11a shows the cost variation in capturing CO_2 in the CCUS network with increasing time periods. As time period between intervals of the capture targets increases, the capture cost decreases mainly due to the effect of time value of money. Also, as the target period for capturing 40 $MtCO_2$/yr increases from current year, the relative value of each USD in 2019 Equivalent would reduce. However, in Figure 5.11b, the storage/utilization cost increases with increasing operating period, as the overall capacity of each reservoir decreases with time. There is the effect of time value of money even on the storage/utilization aspect of the formulation with increasing time periods, however, the effect of storage capacity has greater effect on the overall storage/utilization cost. Similarly, as seen in previous cases, the cost of transportation through pipeline also increases with increasing time intervals (Figure 5.11c). This increase in cost related to pipeline transportation is due to increased deployment of pipelines with utilization of more storage options. Thus, it can be concluded from Figure 5.11c that reservoir capacity influences the cost of pipeline transportation for CO_2.



(a)



(b)



(c)

Figure 5.11: Cost per ton of CO_2 avoidance for CCUS network with variation in time intervals in the dynamic formulation (a) Capture (b) Transport; (c) Storage.

5.4.5 Cost Variation with CO_2 Purchase Cost

A major element which can affect the cost of storage related to the CCUS infrastructure is cost of purchase of CO_2. The cost related to the purchase of CO_2 in USD/ton is dependent on supply and demand of the commodity. In Chapter 4, an assumed value of \$20/tCO_2 was used. The effect of cost of CO_2 ideally would not affect the choice related to source and storage sites. Figure 5.12 depicts the effect of CO_2 purchase cost on storage costs and it shows that the cost of CO_2 storage increases with the purchase cost of CO_2. The effect of the cost as shown in the figure is not as significant as expected due to the nature of formulation, where recycle rate is multiplied by the cost of CO_2. Further, the CO_2 purchase cost effect is reduced due to the usage of saline aquifers where CO_2 does not have to be purchased.

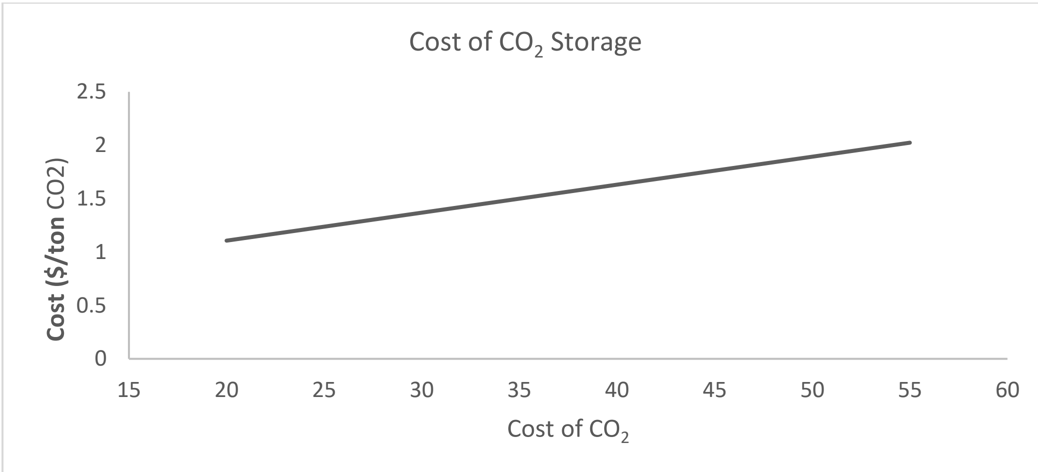


Figure 5.12: Cost per ton of CO_2 storage/utilization for CCUS network with variation of CO_2 cost.

5.5 Effect of Tax Incentives and Profits from Oil

CCUS technology is important to reduce the environmental effects of the power sector by reducing CO_2 emissions. However, in most cases reduction of CO_2 emission is not a sufficient incentive for adopting CCUS technology. This is mostly due to the high costs associated with usage of CCUS as seen in this work. As experts and institutions are mindful of such developments, the United States government has put forth tax incentives to support the development of CO_2 emission prevention technology. The tax incentives related to implementation of CCUS is known as 45Q incentives in the USA (US DOE, 2019). These 45Q incentives provide relief in taxation related to each ton of CO_2 stored or utilized through the CCUS process. According to the US Department of Energy, the tax incentive to store one tCO_2 in a permanent geologic storage reservoir like saline aquifers or depleted reservoirs will be $50 by 2026. Similarly, the tax incentive for using one tCO_2 for EOR purposes or any other form of utilization is rated at $35 by 2026. (US DOE, 2019). Like tax incentives earned from usage of CCUS technology, oil sales from EOR activities need to be used for profit calculations. EOR technology as was mentioned in Chapter 1 is used to enhance oil production in mature reservoirs. The profits earned from the additional sales of oil due to EOR technology can be quite substantial.

In order to track the savings and earnings related to implementation of CCUS technology, the tax incentives related to storage/utilization of CO_2 along with the sale of incremental oil from EOR is modeled into the CCUS network. In this work, the rate for tax incentive related to storage in saline aquifers is taken as \$50/tCO_2, while the tax incentive for using CO_2 in EOR activities is taken as \$35/tCO_2 (US DOE, 2019). The tax incentives are assumed to be valid in 2020, even though these rates are suggested to be utilized in 2026 by the US Department of Energy. Further to incorporate the earnings from incremental oil production through EOR, an incremental oil production rate of 2bbl/tCO_2 (2 incremental barrels of oil produced per ton of CO_2 injected) is assumed (Cooney et al., 2015). The incremental production rate varies in studies, but Azzolina et al. (2016) suggests that such variation is aimed at different conducive regions for CO_2 EOR. Further, the study suggests that the incremental oil production varies from 2 to 5.6 bbl/tCO_2. However, higher incremental oil production rates account for only 20% of the fields used in CO_2 enhanced oil operations. Azzolina et al. (2016) indicates that the conservative assumption of 2bbl/tCO_2 suggested by Cooney et al. (2015) is more appropriate (Azzolina et al. used 2.1 bbl/tCO_2 in their studies). The price of oil in the analysis is assumed to be \$40/bbl of crude oil.

The tax incentives and sale of incremental oil can be included into the CCUS network model through few changes in formulation of the objective function. Thus, Equation 3.25 from Chapter 3 is changed to Equation 5.2, where, $taxinc_j$ and $taxincb_p$ are the tax incentives for sink nodes and dual nodes respectively (\$50 for saline aquifers and \$35 for EOR); $price_j$ and $priceb_p$ are oil prices for reservoir nodes and dual nodes respectively (\$0 for saline aquifers and \$40 for EOR); and PR is the incremental oil production ratio taken as 2 bbl/tCO_2. The rest of the notations are the same as Chapter 3. The major difference in the formulation is the maximization of the objective function as compared to minimization used in Equation 3.24, due to the intention of maximizing profits. Further, all costs related to the CCUS network are subtracted from the sales and savings of the network. In Equation 5.2, the savings from incentives are calculated by multiplying annual storage/utilization in each reservoir with the respective tax incentive. Similarly, the sales from incremental oil production is calculated by multiplying the annual injection rates in each reservoir with the incremental oil production ratio and the assumed price of oil.

$Maximize$

$$Cost = \sum_j (Y_j * taxinc_j + Y_j * price_j * PR) + \sum_p (Yb_p * taxincb_p + Yb_p * priceb_p * PR)$$
$$- \sum_i a_i * Ccap_i * X_i + \sum_p c_p * Ccapb_p * Xb_p +$$
$$\frac{int(1+int)^{time}}{(1+int)^{time}-1} * \sum_{arc} \sum_t Ctran_{arc,t} * b_{arc,t} + \sum_{arc} \sum_t Otran_{arc,t} * b_{arc,t} -$$
$$\frac{int(1+int)^{time}}{(1+int)^{time}-1} * \sum_j \left(r_j * Cres_j + h_j * Cwell_j\right) - \sum_j \begin{pmatrix} h_j * Owell_j + Y_j * MVA_j * MVAC \\ +Y_j * l_j * costCO_2 \end{pmatrix} -$$
$$\frac{int(1+int)^{time}}{(1+int)^{time}-1} * \sum_j \begin{pmatrix} c_p * Cresb_p + \\ hb_p * Cwellb_p \end{pmatrix} + \sum_j \begin{pmatrix} hb_p * Owellb_p + Yb_p * MVAb_p * MVAC \\ +Yb_p * lb_p * costCO_2 \end{pmatrix} \qquad (5.2)$$

An additional variation is required in the formulation to change the constraint related to CO_2 capture goals (Equation 3.35) to Equation 5.3. The major difference in Equation 5.3 is that the total capture needs to be below the set goal. This is due to the nature of the objective function, as Equation 3.35 will maximize profit such that all the reservoirs will be completely filled/utilized.

$$\sum_i X_i + \sum_p Xb_p \le CapCO_2 \qquad (5.3)$$

In order to track the differences caused by the variation of the objective function to include tax incentives and sales from oil, two cases were analyzed for a CO_2 capture goal of 20 (alternative 1 as shown by Figure 5.13a) and 40 $MtCO_2$/yr (alternative 2 as shown by Figure 5.13b) for a 30-year operational period. The summary of CCUS infrastructure development in alternative 1 and 2 is shown in Table 5.4 and Table 5.5 respectively. The details of the development in terms of quantities, cost, savings, and earnings are provided in Appendix E. The summary of the costs, tax incentives and sales of oil for alternative 1 and 2 is provided in Table 5.6.

Table 5.4: State-wise capture and sink statistics for alternative 1(static, no pre-existing infrastructure, 20 $MtCO_2$/yr).

State	Capture Location	Capture ($MtCO_2$/yr)	Storage Location	Storage ($MtCO_2$/yr)
MT	-	0.00	Cedar Creek Anticline	8.10
ND	Great Plains Gasification	9.60	Rough Rider Field	1.50
WY	Jim Bridger, Shute Creek, Lost cabin, Riley Ridge	10.02	Bairoil Field, Beaver Creek	10.02

Table 5.5: State-wise capture and sink statistics for alterbative 2(static, no pre-existing infrastructure, 40 $MtCO_2$/yr).

State	Capture Location	Capture ($MtCO_2$/yr)	Storage Location	Storage ($MtCO_2$/yr)
CO	Rawhinde Energy, Fort St. Vrain, Hayden	4.05	Rangeley Field	4.05
MT	-	0.00	Cedar Creek Anticline	8.10
ND	Great Plains Gasification	9.76	Rough Rider Field, Cedar Hill	5.86
WY	Jim Bridger, Shute Creek, Lost cabin, Riley Ridge, Dave Johnston, Wygen	26.16	Bairoil Field, Beaver Creek, Hamilton Dome, Oregon Basin, Salt Creek	21.96

Table 5.6: Summary of cost analysis for static optimization without existing infrastructure (alternative 1-2).

Scenario	Amount Captured (MtCO2/yr)	Capture Cost ($/ton)	Transport Cost ($/ton)	Storage Cost ($/ton)	Tax Incentives and Oil Sales ($/ton)	Profits ($/ton)
1	19.62	22.15	0.79	2.60	115.00	89.47
2	39.97	34.34	1.06	2.69	115.00	76.91

The CCUS network used to model both alternatives remains the same as the previously analyzed models. From alternative 1, it can be seen that all storage/utilization is done at locations which use CO_2 for EOR. The development is concentrated at western North Dakota and Wyoming, similar to case 1 from Chapter 4. There are no changes in the sources of CO_2, however, the larger pipeline development is observed, with lower number of total pipelines. This change is quite significant because of the deployment of 24-inch pipeline quite early. Similarly, in alternative 2, all development in terms of storage/utilization is around EOR sites. The capture related to alternative 2 is still centered around natural gas plants and commercial vendors, however, the choice between the coal-fired power plants is influenced by proximity to convenient EOR sites. The choice amongst EOR sites is based on the injectivity of each reservoir. Like alternative 1, pipeline development is more prominent in terms of larger diameter pipes. This leads to the development of a corridor stretching from central North Dakota to southern Wyoming and some development in northern Colorado. Overall, the major differences between the alternatives seen in this chapter as compared to scenario 1 and 2 from Chapter 4, is the choice of utilization sites. In the current alternatives, EOR sites are preferred to saline aquifers due to the combined effect of tax incentives as well as sales from incremental oil. The overall development of capture nodes is still favored towards lower cost natural gas plants and commercial vendors rather than coal-fired plants. Another major change is the shift in pipeline development trends, where larger diameter pipelines are preferred over smaller diameter pipelines. This shift in pipeline trends indicates a more centralized and concentrated distribution network in some regions as compared to the distributed networks seen in scenario 1 and 2 of Chapter 4.

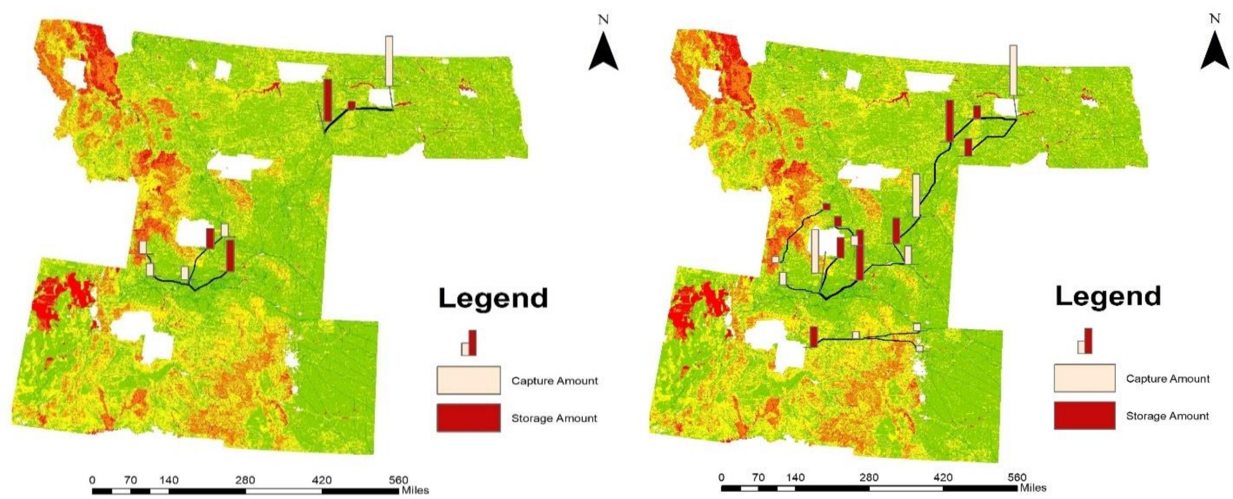
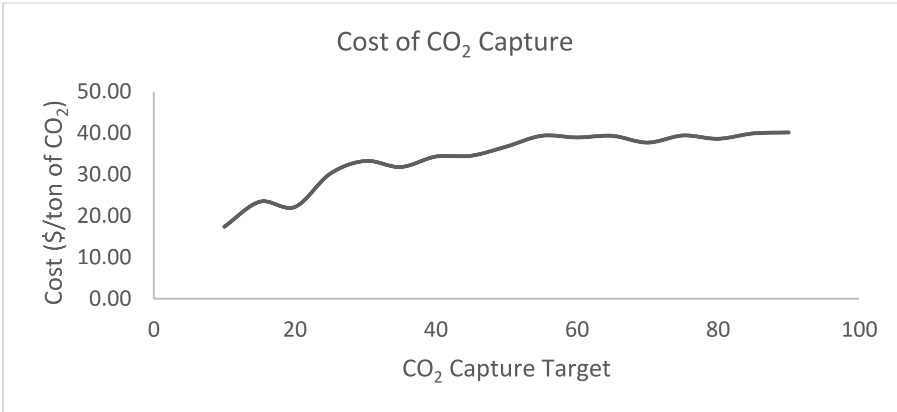
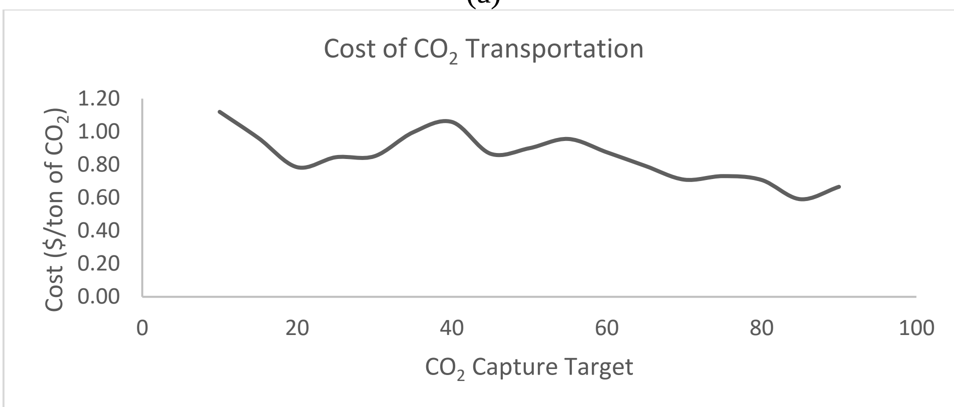


Figure 5.13: CCUS infrastructure deployment scenarios with tax incentives and oil sales for varying CO_2 capture targets: (a) 20 $MtCO_2$/yr (alternative 1) (b) 40 $MtCO_2$/yr (alternative 2).

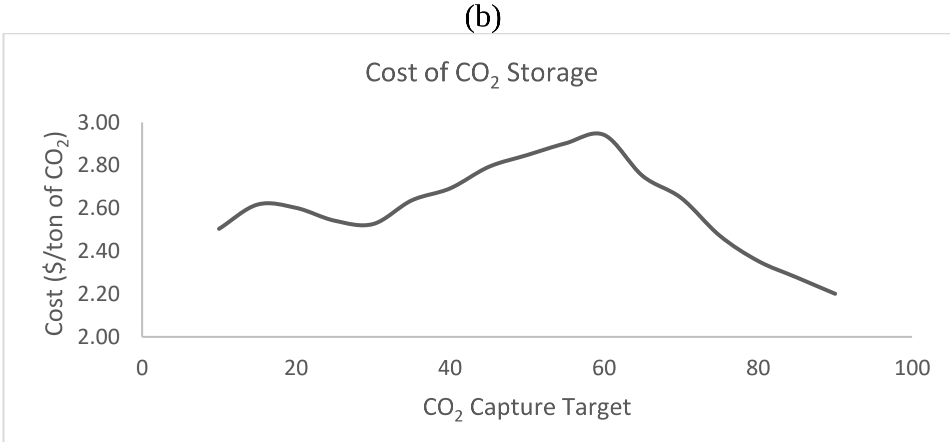
Figure 5.14 shows the variation of costs, savings and profits per ton of CO_2 with varying capture goals from 10 to 90 $MtCO_2$/yr. Figure 5.14a indicates the increase in capture costs with increasing CO_2 capture goals, similar to the case without tax incentives and oil sales. This trend is due to the preference of natural gas plants and commercial vendors over coal-fired power plants. Figure 5.14b shows a varying decrease in transportation cost per ton of CO_2 captured. The initial high costs are due to the deployment of larger diameter pipelines early on, as compared to distributed small size pipelines seen in the cases without tax incentives and oil sales. There is a slight increase in transportation cost between 20 to 40 $MtCO_2$/yr due to deployment of some distributed pipelines, however, these costs further decrease steadily between 40 to 90 $MtCO_2$/yr due to utilization of the higher capacity of large diameter pipelines. Figure 5.14c shows an interesting storage cost pattern with increasing CO_2 targets. Initially the costs related to storage steadily increase due to the usage of EOR sites having lower injectivity, which increases cost. However, after a capture target of 60 $MtCO_2$/yr the cost of storage steadily decreases. This steady decrease coincides with the utilization of saline aquifers for storage, with their respectively lower storage costs and lack of the need for purchasing CO_2. Figure 5.14d shows the savings and sales from using the CCUS network per ton of CO_2. It can be seen that savings/earnings from the CCUS network is steady at \$115/tCO_2 till a capture target of 60 $MtCO_2$, after which the savings reduce per ton of CO_2 captured. This coincides once again with the usage of saline aquifers as they do not have oil profits. Finally, Figure 5.14e shows the decreasing profits per ton of CO_2 stored. This figure indicates two trends: one trend from 10 to 60 $MtCO_2$/yr capture target is due to the increasing capture and storage costs related to the shift towards coal-fired plants and lower injectivity reservoirs; the trend from 60 to 90 $MtCO_2$/yr, coincides with decreasing profits due to the usage of saline aquifers despite lower costs of storage associated with saline aquifers.


Cost of CO2 Capture
Cost ($/ton of CO2)
50.00
40.00
30.00
20.00
10.00
0.00
0
20
40
60
80
100
CO2 Capture Target

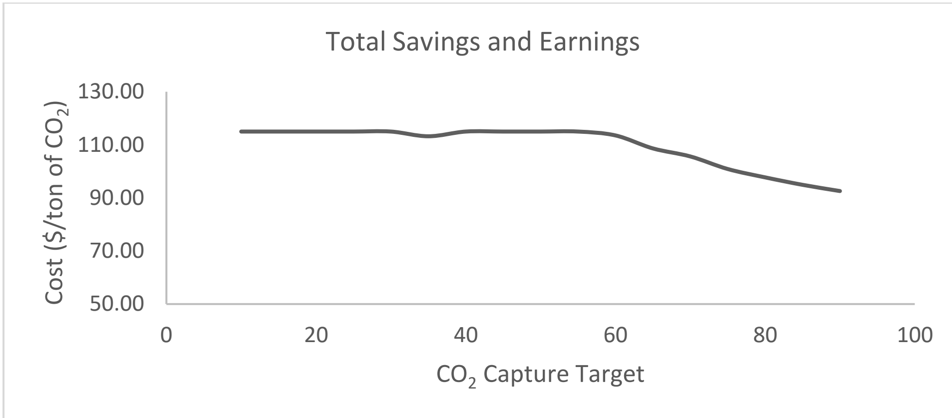
(a)


Cost of CO2 Transportation
Cost ($/ton of CO2)
1.20
1.00
0.80
0.60
0.40
0.20
0.00
0
20
40
60
80
100
CO2 Capture Target

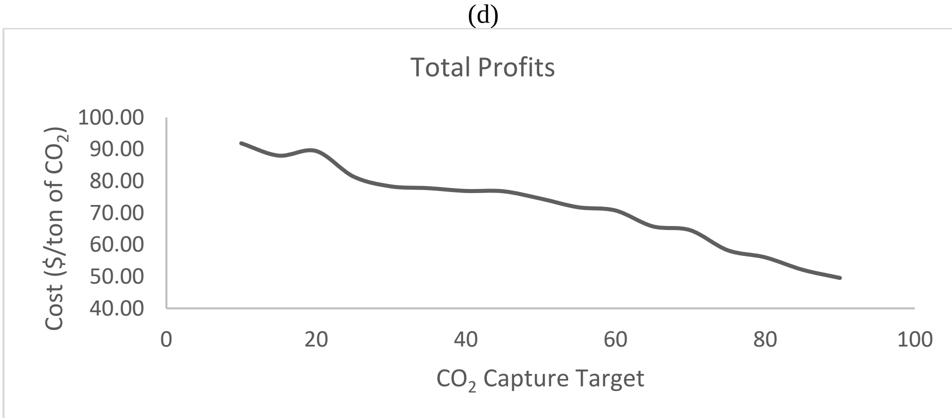
(b)


Cost of CO2 Storage
Cost ($/ton of CO2)
3.00
2.80
2.60
2.40
2.20
2.00
0
20
40
60
80
100
CO2 Capture Target

(c)



(d)



(e)

Figure 5.14: Economics of CCUS network for varying CO_2 capture targets per ton of CO_2: (a) CO_2 capture costs (b) CO_2 transportation costs (c) CO_2 storage costs (d) total savings and sales (e) total profits.

Further, in order to see the effect of including tax incentives and oil sales in the formulation of CCUS network with pre-existent infrastructure, two scenarios are visualized with varying CO_2 capture targets for an operating period of 30 years: 20 and 40 $MtCO_2$/yr (Figure 5.15). The summary of CCUS infrastructure development in addition to the pre-existing infrastructure in alternative 3 and 4 is shown in Table 5.7 and 5.8 respectively. The details of the development in terms of quantities, cost, savings, and earnings are provided in Appendix E. The summary of the costs, tax incentives and sales of oil for alternative 3 and 4 is provided in Table 5.9.

Table 5.7: State-wise capture and sink statistics for alternative 3(static, pre-existing infrastructure, 20 $MtCO_2$/yr).

State	Capture Location	Capture ($MtCO_2$/yr)	Storage Location	Storage ($MtCO_2$/yr)
CO	-	0.00	Rangeley Field	0.21
MT	-	0.00	Cedar Creek Anticline	2.52
ND	Great Plains Gasification, Coyote Station	5.22	Beaver Lodge	0.10
WY	Dave Johnston, Shute Creek, Lost Cabin, Riley Ridge	6.43	Grieve Field, Bairoil Field, Monell Field, Beaver Creek	6.22
External	-	-	Canada	2.60

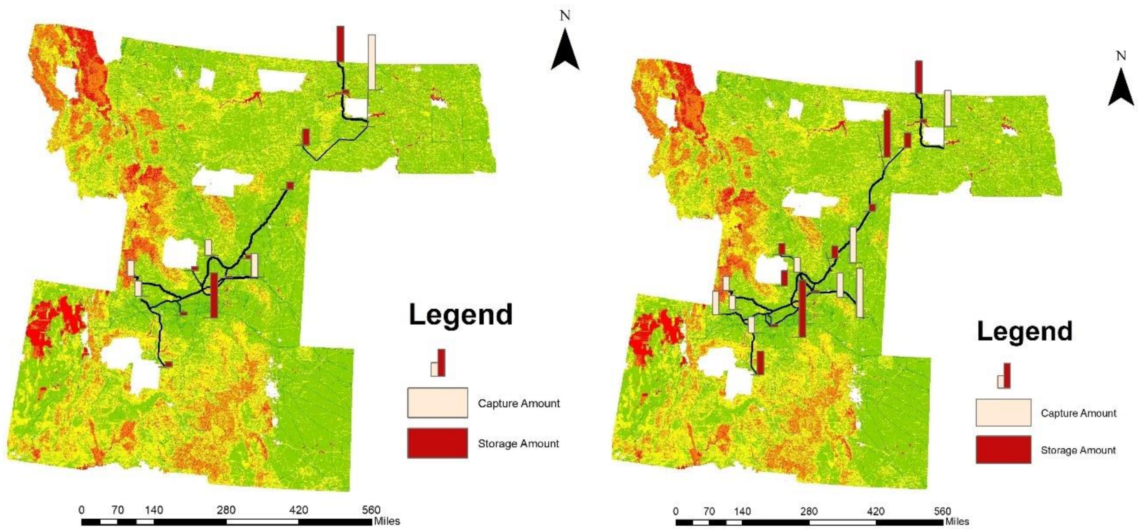
Table 5.8: State-wise capture and sink statistics for alternative 4(static, pre-existing infrastructure, 40 $MtCO_2$/yr).

State	Capture Location	Capture ($MtCO_2$/yr)	Storage Location	Storage ($MtCO_2$/yr)
CO	-	0.00	Rangeley Field	3.45
MT	-	0.00	Cedar Creek Anticline	8.10
ND	Great Plains Gasification	2.70	Beaver Lodge, Rough Rider	2.60
WY	Laramie River, Jim Bridger, Dave Johnston, Naughton, Shute Creek, Wygen, Lost Cabin, Riley Ridge	28.46	Grieve Field, Bairoil Field, Monell Field, Beaver Creek, Hamilton Dome, Salt Creek	14.41
External	-	-	Canada	2.60

Table 5.9: Summary of cost analysis for static optimization without existing infrastructure (alternative 3-4) .

Scenario	**Amount Captured** (MtCO2/yr)	**Capture Cost** ($/ton)	**Transport Cost** ($/ton)	**Storage Cost** ($/ton)	**Tax Incentives and Oil Sales** ($/ton)	**Profits** ($/ton)
3	19.95	23.32	0.97	2.11	115.00	88.61
4	39.46	34.64	0.81	2.37	115.00	77.18

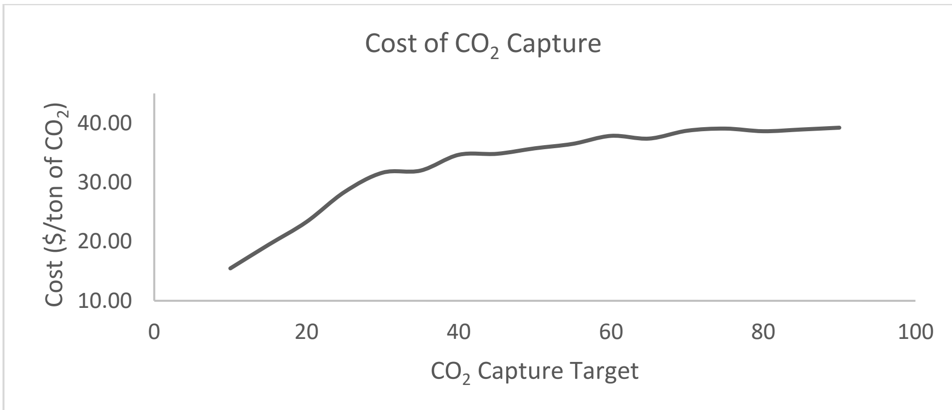
For the capture target of 20 $MtCO_2$/yr (alternative 3) the development is around existing infrastructure targeted towards commercial vendors and natural gas plants. The storage of CO_2 is in EOR fields which have higher injectivity rates, which is not necessarily the case in reservoirs already developed as part of the existing CCUS networks. The additional pipelines built, are low diameter 12- and 16-inch pipes. For the capture target of 40 $MtCO_2$/yr (alternative 4), the development is more prominent as the sources chosen are spread along existing pipeline routes and convenient EOR sites. Pre-existing storage sites are used to a higher degree in alternative 4 compared to its counterpart in Chapter 4 (scenario 6). The pipelines mostly consist of 16- and 24-inch pipes developed as an extension of existing pipelines. The most significant difference between scenario 6 and alternative 4 is the higher utilization of existing pipelines in the flow of CO_2.



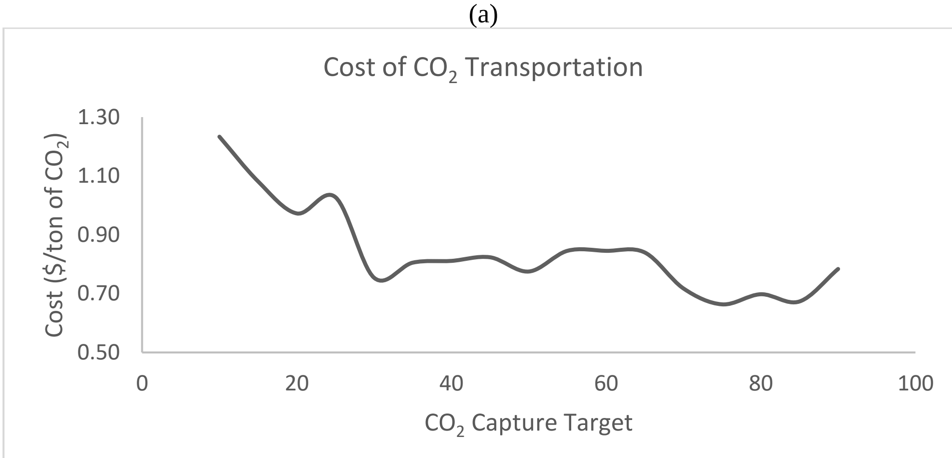
(a) (b)

Figure 5.15: CCUS infrastructure deployment scenarios with pre-existing infrastructure, tax incentives and oil sales for varying CO_2 capture targets: (a) 20 $MtCO_2$/yr (alternative 1) (b) 40 $MtCO_2$/yr (alternative 2).

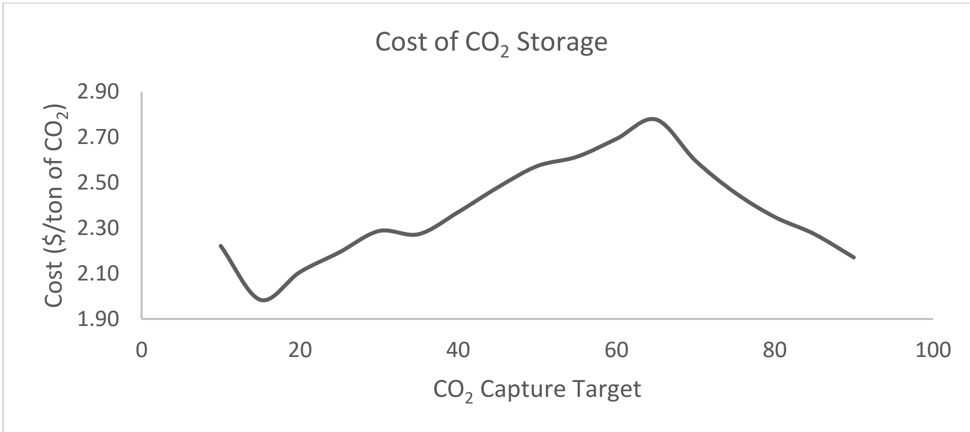
Figure 5.16 shows the variation of costs, savings, and profits per ton of CO_2 with varying capture goals from 10to 90 $MtCO_2$/yr for a network with pre-existing infrastructure. Figure 5.16a shows the increasing cost of capturing CO_2 with increasing capture goals due to the contribution of coal-fired power plants to the capture total. Figure 5.16 b depicts the trend in pipeline cost. The cost of pipeline decreases initially due to higher usage of existing infrastructure, but the cost stabilizes around the capture target of 30 $MtCO_2$/yr with the steady inclusion of new pipelines with increasing capture goals. As expected the cost of CO_2 storage/utilization increases with the capture goals up to 60 $MtCO_2$/yr as the EOR activities shift to lower injectivity fields. But after the 60 $MtCO_2$/yr capture goal threshold, saline aquifers are increasingly used which decreases the storage costs as shown by Figure 5.16c. A familiar trend is seen in the total savings and earnings related to the CCUS network with a steady income of \$115/tCO_2 up to the capture goal of 60MtCO_2/yr, followed by a steady decline in earnings/savings due to the usage of saline aquifers. Figure 5.16e indicates that the profits decrease with increasing capture goals majorly influenced by rising cost for capturing CO_2 and declining profits due to the usage of saline aquifers.



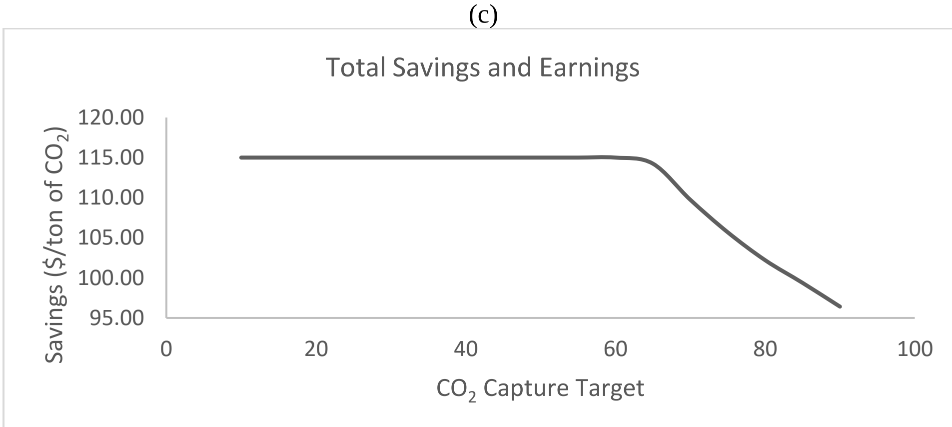
(a)



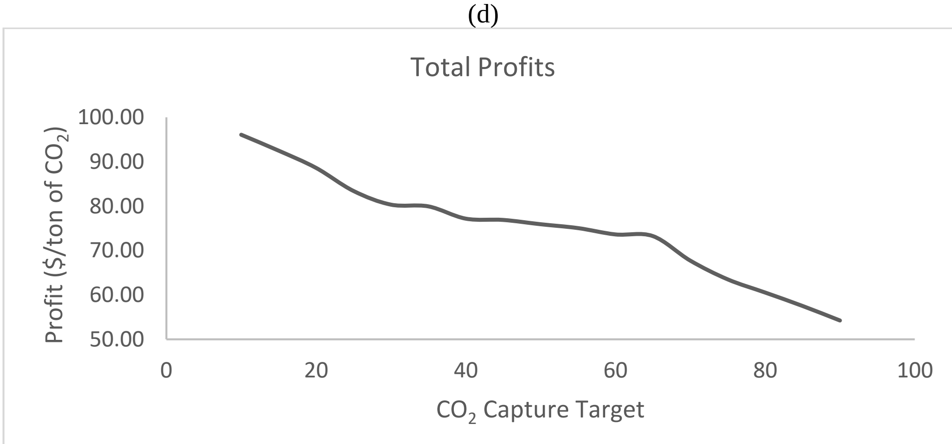
(b)



(c)



(d)



(e)

Figure 5.16: Economics of CCUS network for varying CO_2 capture targets per ton of CO_2 for a network with pre-existing infrastructure: (a) CO_2 capture costs (b) CO_2 transportation costs (c) CO_2 storage costs (d) total savings and sales (e) total profits.

The variation of parameters in the formulation can affect the overall performance of the CCUS network even with tax incentives and oil sales. In this case, a few additional parameters could be considered for their effect on the overall decision-making including oil price and tax incentives. The effect of variation of incentives on the decision-making would be quite unilateral and will not affect the choice in sinks or sources. However, the oil price could affect the overall decision between choice of saline aquifer and EOR sites. If the price of oil falls below 7.5USD/bbl in this model, it will render saline aquifers more profitable as an investment choice compared to EOR sites. It must be understood however, that even though oil prices are set at a standard rate, oil sales do not equate to profits in the market. Thus, the tipping over point between EOR and saline aquifers is highly dependent on the oil price and an associated cost for extraction, transportation, and processing of oil before metering at pipelines.

5.6 Conclusion

This chapter looked into the details related to varying parameters in aspects of the CCSHawk workflow. We see that that in the preparation of the cost map, that variations in the weightages related to the barriers criteria's does not affect the relative cost of the cells. We also see the effect of clustering in overall network creation due to it's effect on number of nodes and the candidate pipeline arcs. Further the economic impacts of varying the control parameters in the optimization formulation is investigated. Also the objective function is varied in order to check the effect of tax incentives and oil sales on overall decision-making process.

CHAPTER 6

Conclusion and Future Work

6.1 Conclusions

Carbon capture, utilization, and storage (CCUS) technology is a key part of several countries' strategy for a sustainable transition to lower CO_2 emission economy. However, CCUS is costly and needs proper management for large-scale deployment. To define the growth and development of CCUS in north-central USA, this thesis has two objectives: firstly, to identify appropriate CO_2 pipeline corridors and routes; secondly, to minimize the cost related to CCUS deployment. Through the implementation of the CCSHawk methodology, the evaluation of CO_2 pipeline corridors as well as the network analysis for CCUS infrastructure was achieved. Some of the key findings from the analysis include:

- North Dakota, central Wyoming, and eastern Colorado have the largest areas of land suitable for CO_2 pipeline corridors. The suitability of a region for pipeline construction is influenced by a combination of several different factors. However, it is seen that existing pipeline routes, slope, lakes, and land use have a high correlation with the land suitability for pipelines using the tract suitability index.
- The modified A* algorithm reduces the length of pipeline routes by 4.23% on average as compared to the conventionally used Dijkstra algorithm. The reduction in route length results in cost reduction, as pipeline cost is dependent on length.

- The cost related to development of the CCUS network varies from 24.05 to 42 \$/tCO_2 for a target capture range of 20 to 90 MtCO_2/yr in the study area. The cost of CCUS is most affected by the type of CO_2 capture technology. Thus, natural gas power plants and combined cycle plants are preferred for CO_2 capture rather than coal-fired power plants due to the lower cost of CO_2 capture. Further, the choice between similar types of capture sources is influenced by proximity to higher injectivity reservoirs. Amongst CO_2 sinks, saline aquifers are preferred over EOR fields because of the cost savings related to purchasing CO_2. However, inclusion of tax incentives and oil sales make EOR fields more attractive than saline aquifers, when the sale price of oil is higher than 7.5 USD/bbl. Amongst sinks of similar categories (saline, EOR), preference of sink utilization inclines towards higher injectivity reservoirs.
- Pre-existing networks can cause a cost difference between 0.01 to 1.62 \$/tCO_2 on CCUS networks. Expansion on networks with pre-existing CCUS infrastructure is more costly up to a capture target of 30 MtCO_2/yr. Beyond the 30 MtCO_2/yr threshold, networks with pre-existing pipeline infrastructure become cheaper than networks without pre-existing infrastructure.

From the dissertation it can be concluded that, environmental factors together with human infrastructures, regulations, and location play a crucial role in determining the regions and routes suitable for CO_2 pipelines. Furthermore, costs related to the CCUS can be optimized by considering different the different components of the CCUS value chain together (sources, sinks and pipelines) together. The CCSHawk methodology and the developed models can be easily used for analysis of any other geographic regions too. The results obtained in this analysis, furthers the knowledge from previous literature by incorporating information related to pre-existing infrastructure as well as shows an increased usage of regulatory considerations.

6.2 Future Work

The work in this thesis can be taken forward in different directions, such as:

- Pipeline corridors and pathing can be improved through the usage of LiDAR based imaging of the terrain. The LiDAR imaging confirm path locations and improve the accuracy of pathing.
- The economic analysis related to CCUS networks can be improved with the inclusion of the detailed analysis of the processing and separation of CO_2 during capture and detailed utilization pathways for CO_2. Such detailed analysis for a network would increase the number of variables in the formulation significantly, increasing the computational load.
- The link between reservoir simulation packages with CCSHawk, would enable the system to predict the overall economic impact of individual reservoir properties on the CCUS network. This level of detailing would help in forming a link between reservoir level modeling and regional level economic impact analysis.
- Another direction for the future work, could be the estimation of the impact of cost saving strategies in the CCUS network analysis such as the usage of bi-directional pipeline systems or usage of multi-commodity pipeline systems.

APPENDIX A

Mapping Study Area

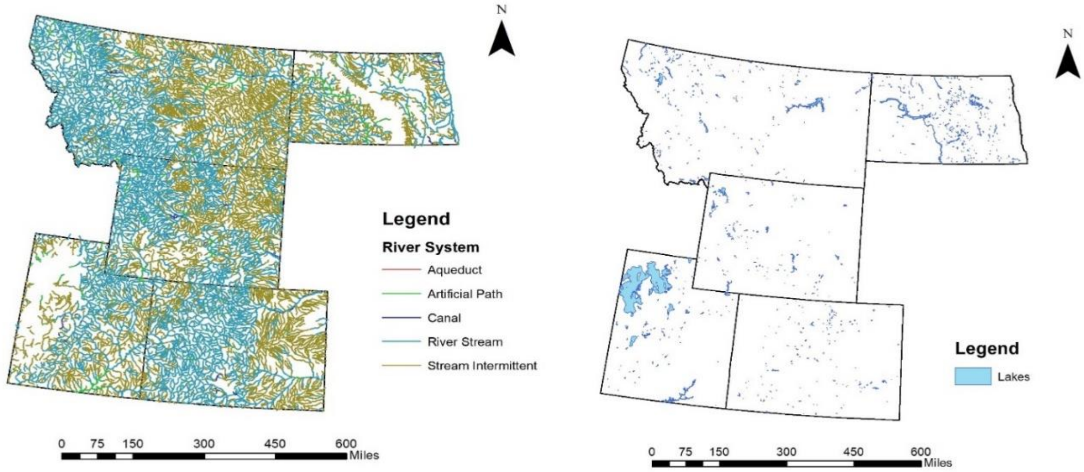
A.1 Introduction

This Appendix showcases the 19 thematic maps used in mapping the study area of North-Central USA (North Dakota, Montana, Wyoming, Utah and Colorado).Section A.2 of this appendix is linked to geo-information layers described in Section 3.3. Section A.3 of this appendix is linked to the reclassification information for reforming the 17 thematic maps for implementation of AHP and cost map preparation. The maps and weightages used in this analysis is a reproduction of the artwork and weightages used in Balaji et al. (2020)

A.2 Maps

Waterways: Figure A.1a depicts a map of river systems contained with the study area consisting of smaller prominent features such as aqueducts, artificial river paths, canals, river streams and stream intermittent where most value is given to river streams and canals, intermediate importance is given to stream intermittent, artificial river paths and aqueducts. The vector features are buffered with a 100m radius around them.

Lakes: Figure A.1b depicts a map of lakes and major waterbodies in the study area. It must be noted that these bodies are buffered using 100m radius around them.

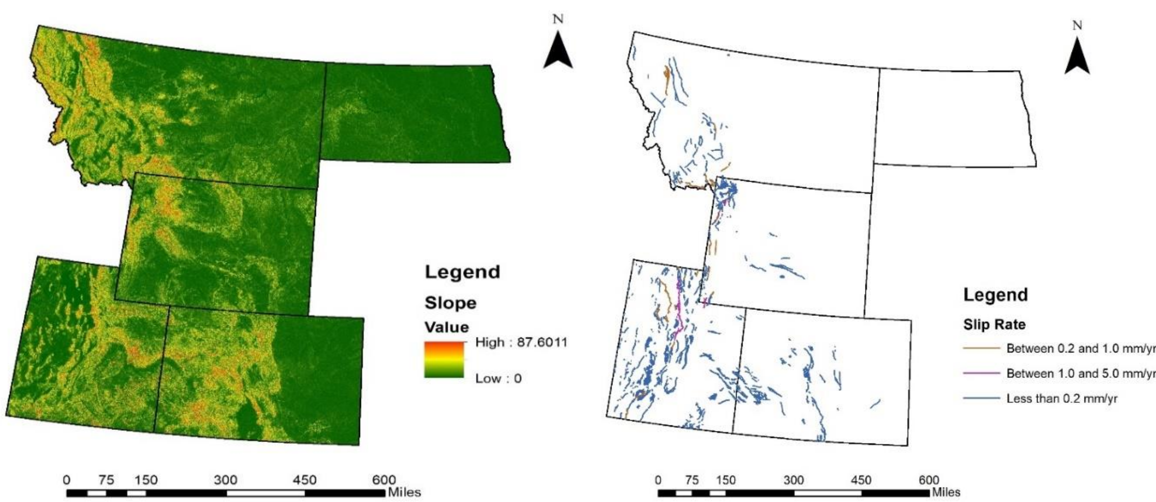


(a) (b)

Figure A.1: (a) Waterways (b) Lakes and major waterbodies.

Slope Map: Figure A.2a depicts the slope map of the study area which is in raster format. The value ranges from 0 to 90-degree elevation change.

Fault Map: Figure A.2b depicts quaternary fault map for the study area where the faults are characterized by slip rate. The slip rate fault map is in vector format.

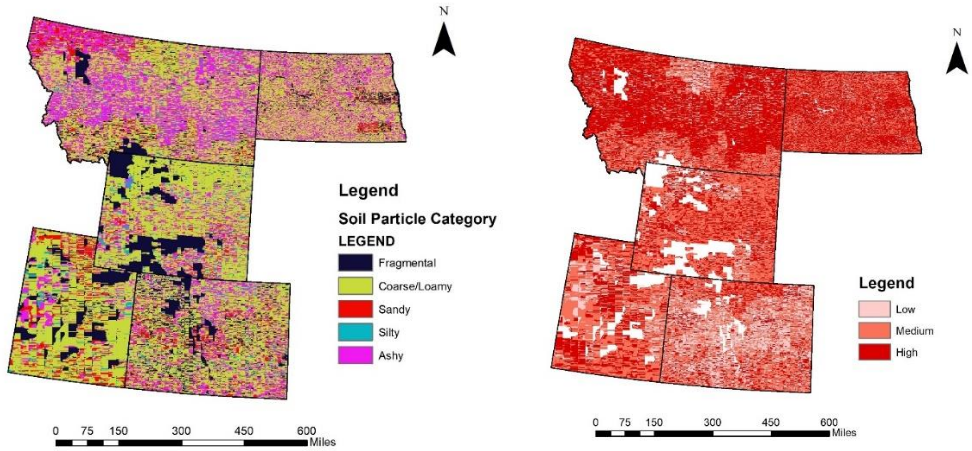


(a) (b)

Figure A.2: (a) Slope map of in degrees (b) Quaternary fault map.

Soil: Figure A.3a depicts soil particle size for the study area. The soil particle size is classified as ashy, medial, coarse, silty, loamy, sandy and fragmental arranged from smallest particle size to the largest respectively.

Corrosion Susceptibility of Steel: Figure A.3b depicts the susceptibility of steel to climate in local region in study area classified ass low, medium, and high.



(a) (b)

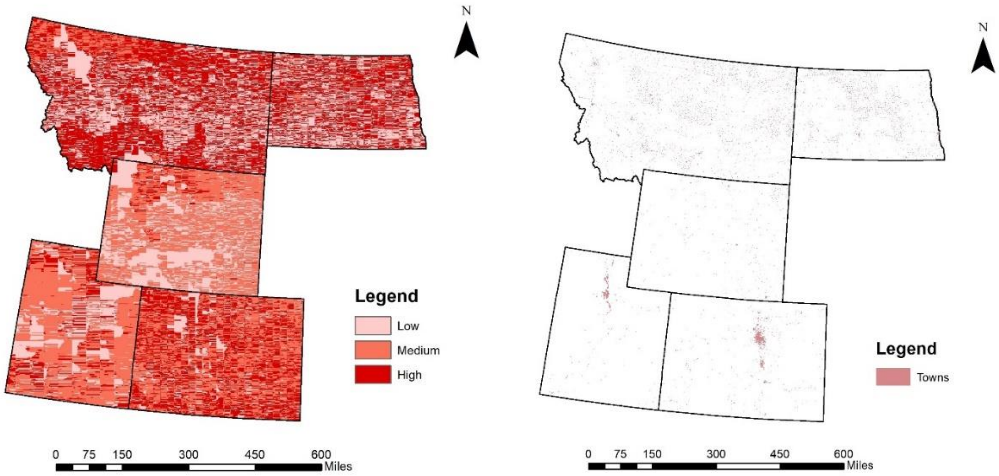
Figure A.3: (a) Soil particle size category map (b) Steel corrosion susceptibility map.

Frost Cover on Topsoil: Figure A.4a depicts the susceptibility of topsoil to frost cover in regions in study area, where the regions are classified as low, medium and high levels of susceptibility.

Towns: Figure A.4b depicts the population region distribution in the study area, which is assumed to be around schools buffered by 1.6km.

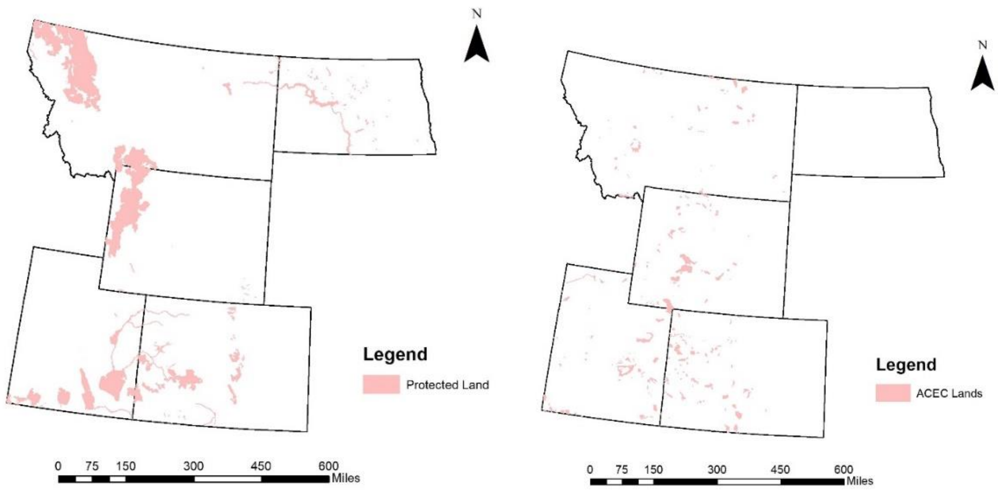
Protected Land: Figure A.5a depicts protected lands distributed throughout the study area, which includes wildlife reserves and historical sites.

Areas of Critical Environmental Concern (ACEC): Figure A.5b depicts ACEC regions in the study area.



(a) (b)

Figure A.4: (a) Topsoil susceptibility to frost (b) Towns.

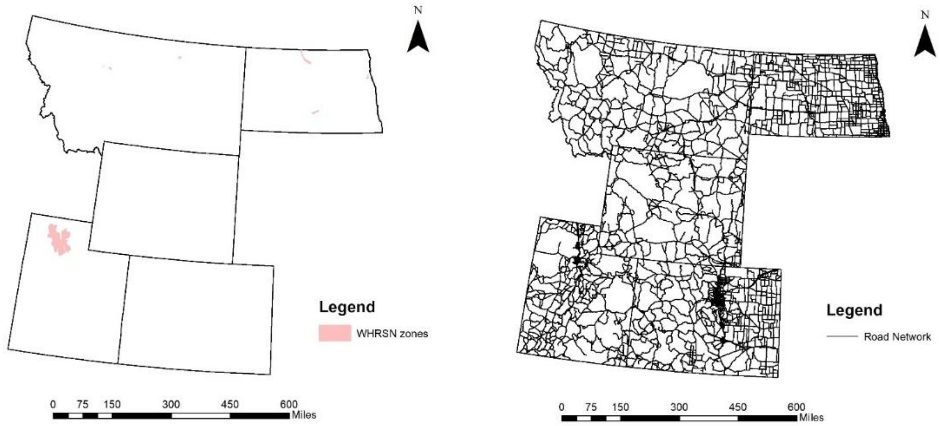


(a) (b)

Figure A.5: (a) Protected land (b) Areas of critical environmental concern.

Western Hemisphere Shorebird Reserve Network (WHRSN): Figure A.6a depicts regions classified as WHRSN zones for migratory birds.

Roads: Figure A.6b depicts road networks in the study area. The roads include interstates, state highways and county roads. The vector data is buffered with a 100m radius.

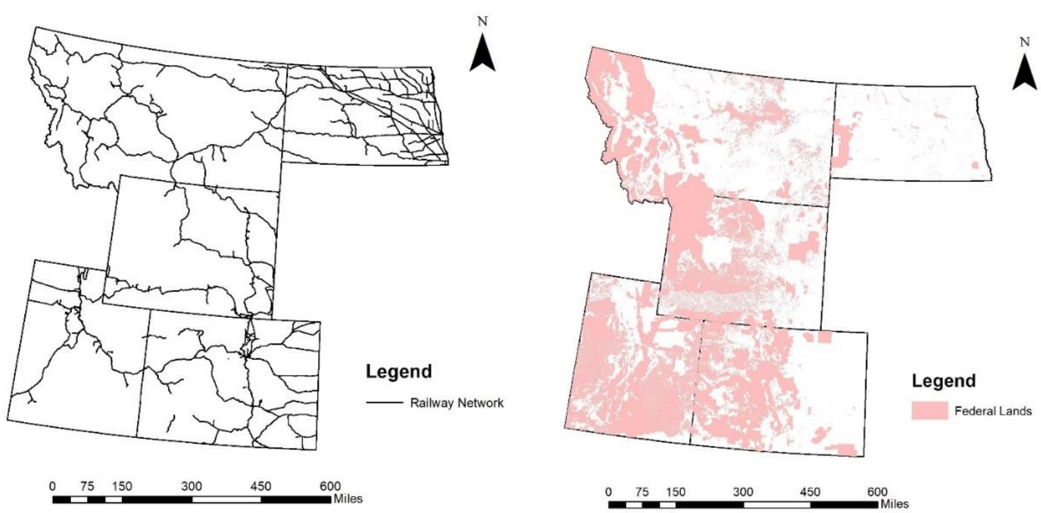


(a) (b)

Figure A.6: (a) WHRSN sites (b) Roads.

Railroad: Figure A.7a depicts railway network through the study area. The rail network includes commercial rail lines and military rail lines. The vector data is buffered with a 100m radius.

Federal Lands: Figure A.7b depicts federal lands in the study area which includes various facilities.

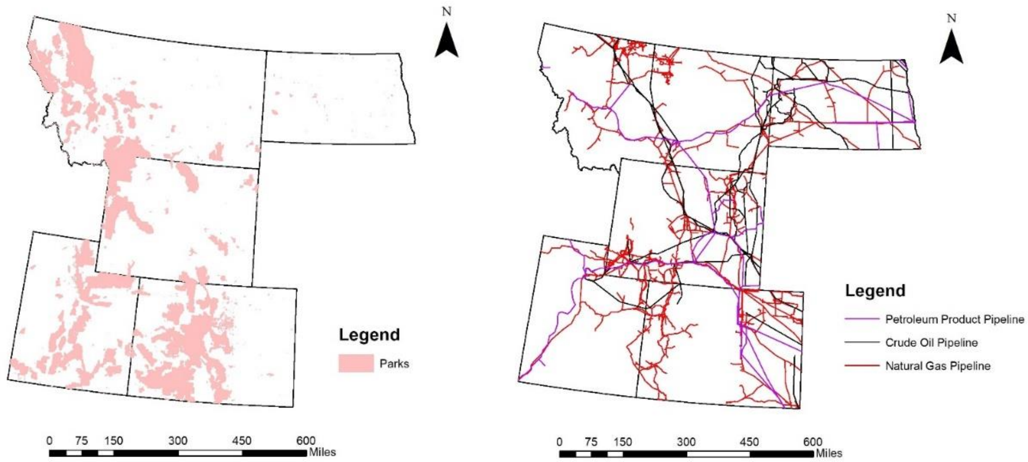


(a) (b)

Figure A.7: (a) Railway Network (b) Federal Lands.

Parks: Figure A.8a depicts parks distributed throughout the study area which includes both national parks and state parks.

Existent Pipeline: Figure A.8b depicts existent pipeline routes in the study area. The layer includes pipelines for Natural Gas, Crude Oil and Processed Crude Oil products.



(a) (b)

Figure A.8: (a) Parks (b) Existent pipeline in the study area.

Land Cover: Figure A.9 depicts land cover distribution in the study area. The categories in the study include forest, shrub vegetation, desert, polar, open rock vegetation, nonvascular, agricultural, introduced vegetation, recently developed, open water and developed.

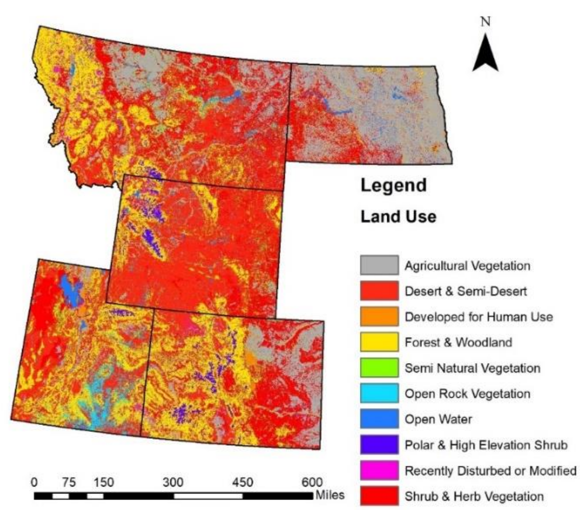
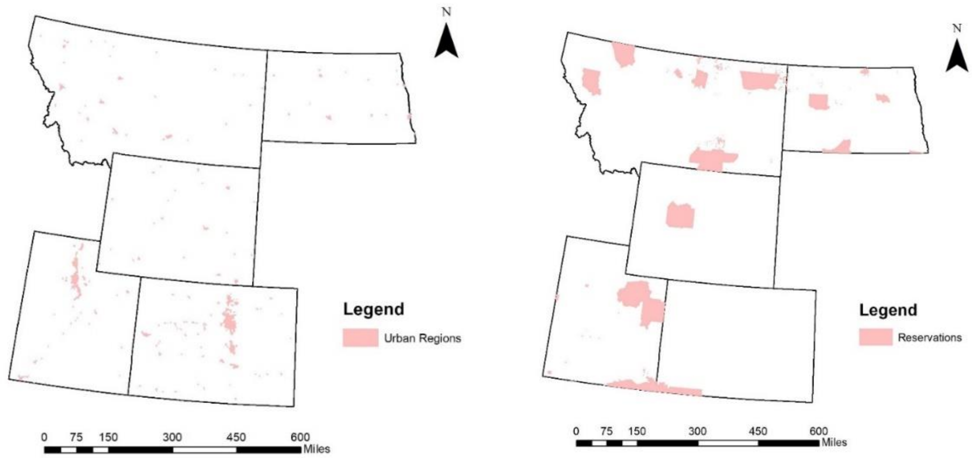


Figure A.9: Land use distribution in the study area.

Cities: Figure A.10a depicts the various city and urban regions in the study area. The areas displayed under this layer are excluded from the study area.

Reservations: Figure A.10b depicts the reservations around the study area. The areas displayed under this layer are excluded from the study area.



(a) (b)

Figure A.10: (a) Cities (b) Reservations in the study area.

A.2 Reclassification Tables

This section will provide the reclassification information for both discrete and continuous data in the 17 geoinformation layers. The data used for reclassification is a continuous scale from 1 to 10, where 1 is most favorable towards pipeline construction, while 10 is least favorable.

Waterways: Waterway system's data is categorical, and reclassification of the information is provided in Table A.1

Table A.1: Reclassification values for waterway systems.

Original Value Categorical	**Reclassified Value** Continuous
River Stream	10
Stream Intermittent	7
Aqueduct	5
Artificial Path	5
Canal	5
No Value	1

Fault: The data for fault map is categorical in terms of ranges of slip rate per year and the reclassification is in Table A.2.

Table A.2: Reclassification values for faults.

Original Value Categorical	**Reclassified Value** Continuous
0.2 – 1.0 mm/yr	10
1.0 – 5.0 mm/yr	7
Less than 0.2 mm/yr	5
No Value	1

Soil: The data for soil is in terms of particle size and is categorical and the reclassification is in Table A.3.

Table A.3*:* Reclassification values for soil particle size.

Original Value Categorical	**Reclassified Value** Continuous
Ashy	10
Coarse	7
Silty	5
Sandy	3
Fragmental	1
No Value	3

Corrosion Susceptibility of Steel and Frost Cover of Topsoil: The data for corrosion map and frost cover is categorical and the reclassification is in Table A.4.

Table A.4: Reclassification values for corrosion susceptibility of steel and frost topsoil.

Original Value Categorical	**Reclassified Value** Continuous
High	10
Medium	5
Low	1
No Value	1

Land Use: The data for land use categorical and the reclassification is in Table A.5.

Slope: The data for slope is continuous and standardized and reclassified from 1 to 10. The original data anges from 0 to 90 degrees in angles.

Categorical Geo-information Layers: Lakes, towns, protected lands, ACEC zones, WHRSN zones, road networks, railway networks, federal lands, parks and existent networks are geo-information layers which are categorical in nature. These layers are categorized by existence of the entity in the map and are reclassified as 1 or 10. The reclassification of these layers are detailed in Table A.6.

Table A.5: Reclassification values for land use.

Original Value Categorical	**Reclassified Value** Continuous
Developed for Human Use	10
Open Water	10
Forest & woodland	7
Recently Disturbed or Modified	7
Polar & High Elevation Shrubs	5
Agricultural Vegetation	5
Semi Natural Vegetation	3
Shrub & Herb Vegetation	1
Desert	1
Open Rock Vegetation	1

Table A.6: Reclassification value for categorical geo-information layers.

Geo-information Layer	**Original Value**	**Reclassified Value**
Lakes	Lake No Value	10 1
Towns	Town No Value	10 1
Protected Land	Protected Land No Value	10 1
ACEC	ACEC zone No Value	10 1
WHRSN	WHRSN zone No Value	10 1
Roads-	Road No Value	10 1
Railway Network	Railway No Value	10 1
Federal Lands	Federal Land No Value	10 1
Parks	Park No Value	10 1
Existent Pipelines	Pipeline No Value	1 10

APPENDIX B

Algorithms

B.1 Introduction

This purpose of this appendix is to detail the algorithms used in the implementation of various section of the dissertation. This includes the algorithms for: Creation of Delaunay pairs on Python (Section B.2); Route generation using A* algorithm on R (Section B.3); MILP formulation for static optimizations on GAMS (Section B.4); MILP formulation for dynamic optimization on GAMS (Section B.5); and Text analysis on Python (Section B.6).

B.2 Delaunay Pairs

The Delaunay pairs algorithm is referenced in Section 3.5.2. The code is executed on Python and has general in-built dependacies which are usually comes with Anaconda based Python interface. The dependencies for this algorithm are the Numpy, Pandas and Scipy libraries. The required inputs for this algorithm is a csv format spreadsheet with the clustered nodes, connector nodes and joints in pre-existing pipes. Each node is labelled with a attached node type, namely: Source, Sink, Source and Sink, Blank and Pipe. Here *blank* refers to connector nodes while *pipe* refers to wedges in existing pipelines. In the code "#" indicates comments.

```
#Import Libraries
import numpy as np
import pandas as pd
from scipy.spatial import Delaunay

#Reading of Input Clustered nodes
initial_data=pd.read_excel("exist_node_excel.xlsx")
df=initial_data[['longitude_wgs84','latitude_wgs84']]
x=list(df.iloc[:,0])
y=list(df.iloc[:,1])
#Generation of Delaunay Triangles
tri=Delaunay(df)
triangles=tri.simplices
#Extract Edges of Triangle
outX= []
outY=[]
for i in range(len(triangles)):
    outX.extend([triangles[i,0], triangles[i,0], triangles[i,1]])
    outY.extend([triangles[i,1], triangles[i,2], triangles[i,2]])
d={'X':outX,'Y':outY}
out=pd.DataFrame(d)
#Rearrangment on nodes according to ID number
```

```
for i in range(len(out)):
    a=out.iloc[i,0]
    b=out.iloc[i,1]
    mini=min(a,b)
    maxi=max(a,b)
    out['X'][i]=mini
    out['Y'][i]=maxi
#Removing duplicate edges
out_try=out.drop_duplicates(keep="first")
#Removing exist Connection between Connector Nodes
blank=list(np.where(initial_data["Type_node"]=="blank")[0])
for a in range(len(blank)):
    b=blank[a]
    out_try=out_try[out_try.X != b]
    out_try=out_try[out_try.Y != b]
#Removal reuncadncies and pre-existant network nodes
node=list(initial_data.node)
del_node=[x for x in node if str(x) != 'nan']
pipe=list(np.where(initial_data["Type_node"]=="Pipe")[0])
for a in range(len(pipe)):
    b=pipe[a]
    out_try=out_try[out_try.X != b]
    out_try=out_try[out_try.Y != b]
```

```
for a in range(len(pipe)):
    b=pipe[a]
    c=del_node[a]
    out_try=out_try.append({'X' : c, 'Y' : b}, ignore_index=True)
```

B.3 A* Algorithm: Routing

The A* algorithm is used pipeline routing and is referred to and discussed in detail in Section 3.5.3. The algorithm is executed through R in the R Studio interface. The major dependencies to execute the algorithm is only the raster package. The input for this algorithm is a raster file of the cost map and a raster file with node location details in the dimensions of referenced cost map. In the code "#" indicates comments.

```
#Import Library
library(raster)
#Import Rasters of Cost and Source-Sink
CostMat <- as.matrix(raster("R/ArcTest/north_central/cost_200.tif"))
IOMat <- as.matrix(raster("R/ArcTest/north_central/io200v3"))
source=which(IOMat==30, arr.ind = TRUE)
sink=which(IOMat==47, arr.ind = TRUE)
sourcei=source[1,2]
sourcej=source[1,1]
sinki=sink[1,2]
sinkj=sink[1,1]
```

```
#Create Function for Calculating Euclidian Distance between any point in raster and its Parent * Cost -> Accumuated COst for Cell
distance_calculator <- function(x1,y1,x2,y2,z) {((((((x1-x2)^2)+((y1-y2)^2))^0.5)*CostMat[x1,y1])+z}

#Create function for making matrix of locations of neighboring cells without blocked or redundant cells
pa <- matrix(NA,nrow=8,ncol=2)
fiter=0
get_neighbors <- function(i,j) {
 if((i==1 || j==1)==TRUE) {pa <-na.omit(pa)
 return(pa)} else {
  if (!is.na(CostMat[(i+1), (j+1)])==TRUE) {fiter=fiter+1
  pa[fiter,1]=(i+1)
  pa[fiter,2]=(j+1)}
  if (!is.na(CostMat[(i+1), j])==TRUE) {fiter=fiter+1
  pa[fiter,1]=(i+1)
  pa[fiter,2]=j}
  if (!is.na(CostMat[(i+1), (j-1)])==TRUE) {fiter=fiter+1
  pa[fiter,1]=(i+1)
  pa[fiter,2]=(j-1)}
  if (!is.na(CostMat[i, (j+1)])==TRUE) {fiter=fiter+1
  pa[fiter,1]=i
```

```
    pa[fiter,2]=(j+1)}
    if (!is.na(CostMat[i, (j-1)])==TRUE) {fiter=fiter+1
    pa[fiter,1]=i
    pa[fiter,2]=(j-1)}
    if (!is.na(CostMat[(i-1), (j+1)])==TRUE) {fiter=fiter+1
    pa[fiter,1]=(i-1)
    pa[fiter,2]=(j+1)}
    if (!is.na(CostMat[(i-1), j])==TRUE) {fiter=fiter+1
    pa[fiter,1]=(i-1)
    pa[fiter,2]=j}
    if (!is.na(CostMat[(i-1), (j-1)])==TRUE) {fiter=fiter+1
    pa[fiter,1]=(i-1)
    pa[fiter,2]=(j-1)}
    pa <-na.omit(pa)
    return(pa)}}

# Distance Heuristic
hfactor <- function(x,y){2.2*((((x-sinkj)^2)+((y-sinki)^2))^0.5)}


#Function to backtrack to Source Cell
get_source <- function(x1,y1,x2,y2) {
  if ((x2==(x1+1) && y2==(y1+1))==TRUE) {accx=1}
  else if ((x2==(x1+1) && y2==y1)==TRUE) {accx=2}
```

```
  else if ((x2==(x1+1) && y2==(y1-1))==TRUE) {accx=3}
  else if ((x2==x1 && y2==(y1+1))==TRUE) {accx=4}
  else if ((x2==x1 && y2==y1)==TRUE) {accx=5}
  else if ((x2==x1 && y2==(y1-1))==TRUE) {accx=6}
  else if ((x2==(x1-1) && y2==(y1+1))==TRUE) {accx=7}
  else if ((x2==(x1-1) && y2==y1)==TRUE) {accx=8}
  else if ((x2==(x1-1) && y2==(y1-1))==TRUE) {accx=9}
  return(accx)}

#List Declarations
open_list <- cbind.data.frame(sourcej,sourcei,0,0,0)
open_list <- open_list[-c(1),]
closed_list <- matrix(nrow =10000000, ncol =4)
current_list <- rbind(c(sourcej,sourcei,0,0,0))
iteration=1
acc_last = 0
#1st iteration
neighbors_list <- get_neighbors(current_list[1,1], current_list[1,2])
for (f in 1:nrow(neighbors_list)) {
  acc <-
distance_calculator(current_list[1],current_list[2],neighbors_list[f,1],neighbors_list[f,2],current_list[3])
  direc <- get_source(current_list[1],current_list[2],neighbors_list[f,1],neighbors_list[f,2])
```

```
  fact <- acc + hfactor(neighbors_list[f,1],neighbors_list[f,2])
  open_list=rbind(open_list, c(neighbors_list[f,1], neighbors_list[f,2], acc, direc,fact))}
names(open_list) <- c('x', 'y', 'acc','dir','fact')
open_list <- open_list[order(open_list$fact),]
closed_list[iteration,1:4]=c(current_list[1],current_list[2],current_list[3],current_list[4])
CostMat[current_list[1],current_list[2]]=NA
current_list[1:5] = c(open_list[1,1],open_list[1,2],open_list[1,3],open_list[1,4],open_list[1,5])
open_list <- open_list[-c(1),]
iteration=iteration+1
#While loop
while (open_list>=1) {
  neighbors_list <- get_neighbors(current_list[1], current_list[2])
  print(iteration)
  if (length(neighbors_list)==2) {
    acc <- distance_calculator(current_list[1],current_list[2],neighbors_list[1],neighbors_list[2],
current_list[3])
    direc <- get_source(current_list[1],current_list[2],neighbors_list[1],neighbors_list[2])
    fact <- acc + hfactor(neighbors_list[1],neighbors_list[2])
    to_continue_or_not <- FALSE
    g=intersect(which(neighbors_list[1]==open_list$x),which(neighbors_list[2]==open_list$y))
    if (length(g)>=1) {
      if ((acc >= open_list[g,3])==TRUE) {open_list[g,3] = acc
      open_list[g,4] = direc
```

```
      open_list[g,5]=fact}
      to_continue_or_not <- TRUE}
    if (to_continue_or_not==FALSE) {open_list=rbind(open_list, c(neighbors_list[1], neighbors_list[2], acc, direc,fact))}
    open_list <- open_list[order(open_list$fact),]
    closed_list[iteration,1:4]=c(current_list[1],current_list[2],current_list[3],current_list[4])
    CostMat[current_list[1],current_list[2]]=NA
    if (current_list[1]==sinkj && current_list[2]==sinki) {break}
    current_list[1:5] = c(open_list[1,1],open_list[1,2],open_list[1,3],open_list[1,4],open_list[1,5])
    open_list <- open_list[-c(1),]
    iteration=iteration+1
    next }
  if (length(neighbors_list)==0) {
    open_list <- open_list[order(open_list$fact),]
    closed_list[iteration,1:4]=c(current_list[1],current_list[2],current_list[3],current_list[4])
    CostMat[current_list[1],current_list[2]]=NA
    if (current_list[1]==sinkj && current_list[2]==sinki) {break}
    current_list[1:5] = c(open_list[1,1],open_list[1,2],open_list[1,3],open_list[1,4],open_list[1,5])
    open_list <- open_list[-c(1),]
    iteration=iteration+1
    next}
  for (f in 1:nrow(neighbors_list)) {
```

```
    acc <- distance_calculator(current_list[1],current_list[2],neighbors_list[f,1],neighbors_list[f,2],
current_list[3])
    direc <- get_source(current_list[1],current_list[2],neighbors_list[f,1],neighbors_list[f,2])
    fact <- acc + hfactor(neighbors_list[f,1],neighbors_list[f,2])
    to_continue_or_not <- FALSE
g=intersect(which(neighbors_list[f,1]==open_list$x),which(neighbors_list[f,2]==open_list$y))
    if (length(g)>=1) {
      if ((acc >= open_list[g,3])==TRUE) {open_list[g,3] = acc
      open_list[g,4] = direc
      open_list[g,5]=fact}
      to_continue_or_not <- TRUE
    }
    if (to_continue_or_not==FALSE) {open_list=rbind(open_list, c(neighbors_list[f,1],
neighbors_list[f,2], acc, direc,fact))}
  }
  open_list <- open_list[order(open_list$fact),]
  closed_list[iteration,1:4]=c(current_list[1],current_list[2],current_list[3],current_list[4])
  CostMat[current_list[1],current_list[2]]=NA
  if (current_list[1]==sinkj && current_list[2]==sinki) {break}
  current_list[1:5] = c(open_list[1,1],open_list[1,2],open_list[1,3],open_list[1,4],open_list[1,5])
  open_list <- open_list[-c(1),]
  iteration=iteration+1
}
```

```
#Backtrack Procedure
closed_list <- na.omit(closed_list)
AccMat <- matrix(data=NA, nrow =2909, ncol =3038)
for (b in 1:nrow(closed_list)) {
  AccMat[closed_list[b,1],closed_list[b,2]]=closed_list[b,4]
}
path_list <- matrix(nrow =1000000, ncol =2)
curcell=c(sinkj,sinki)
itercell=1
while (curcell>=1){
  path_list[itercell,1:2]=c(curcell[1],curcell[2])
  if (AccMat[curcell[1],curcell[2]]==1) {i=curcell[1]-1
  j=curcell[2]-1
  curcell[1:2]=c(i,j)
  itercell=itercell+1
  next}
  if (AccMat[curcell[1],curcell[2]]==2) {i=curcell[1]-1
  j=curcell[2]
  curcell[1:2]=c(i,j)
  itercell=itercell+1
  next}
  if (AccMat[curcell[1],curcell[2]]==3) {i=curcell[1]-1
```

```
j=curcell[2]+1
curcell[1:2]=c(i,j)
itercell=itercell+1
next}
if (AccMat[curcell[1],curcell[2]]==4) {i=curcell[1]
j=curcell[2]-1
curcell[1:2]=c(i,j)
itercell=itercell+1
next}
if (AccMat[curcell[1],curcell[2]]==6) {i=curcell[1]
j=curcell[2]+1
curcell[1:2]=c(i,j)
itercell=itercell+1
next}
if (AccMat[curcell[1],curcell[2]]==7) {i=curcell[1]+1
j=curcell[2]-1
curcell[1:2]=c(i,j)
itercell=itercell+1
next}
if (AccMat[curcell[1],curcell[2]]==8) {i=curcell[1]+1
j=curcell[2]
curcell[1:2]=c(i,j)
itercell=itercell+1
```

```
  next}
  if (AccMat[curcell[1],curcell[2]]==9) {i=curcell[1]+1
  j=curcell[2]+1
  curcell[1:2]=c(i,j)
  itercell=itercell+1
  next}
  if (AccMat[curcell[1],curcell[2]]==0) {break}
  if (curcell[1]==sourcej && curcell[2]==sourcei) {break}
}
path_list <- na.omit(path_list)
plot(path_list[,1], path_list[,2], pch=19)
#Generation of output matrix
CostMat <- as.matrix(raster("R/ArcTest/north_central/cost_200.tif"))
cost =0
for (s in 1:nrow(path_list)) {cost=cost+CostMat[path_list[s,1],path_list[s,2]]}
avgcost =cost/itercell
move=0
for (s in 2:nrow(path_list)) {move=move+((((path_list[s,1]-path_list[(s-1),1])^2)+((path_list[s,2]-
path_list[(s-1),2])^2))^0.5)}
total_length=move*200
demo <- matrix(nrow =2909, ncol =3038)
for (m in 1:nrow(path_list)) {demo[path_list[m,1], path_list[m,2]]=1}
#Generation of output raster
```

```
rst <- raster(demo)
coord<-raster("R/ArcTest/north_central/cost_200.tif")
rst <- raster(demo, xmn=-833913.6 , xmx=685086.4, ymn=4098519, ymx=5553019)
crs(rst) <- "+proj=utm +zone=14 +datum=NAD83 +units=m +no_defs +ellps=GRS80 +towgs84=0,0,0"
writeRaster(rst, 'trial_30_47.tif')
```

B.4 Static MILP Formulation

This section provides the code developed on GAMS interface to model a MILP formulation to determine the static optimization of a pipeline and CCUS infrastructure network for North-Central USA.

B.4.1 Static MILP formulation without Pre-Existent Infrastructure

The code provided below has no pre-existent infrastructure. This model is related to Section 3.6.1. No dependencies are required for the execution as the trial version of IBM CPLEX is available on GAMS for execution. The major inputs of this model include the cost models for source, sinks and pipelines and CO_2 capture goals. In the code "*" indicates comments. The results of the execution and fixed inputs are provided in Section 4.4. the results for the execution is provided in Appendix E.

```
$title Cost Reduction of Static CCUS network with no pre-existent infrastructure
$onText
This problem solves the transportation of co2 from various sources and sinks provides fixed cost and variable OPEX cost dependent on production of CO2
```

and injection rates of CO2 into both saline aquifers and Co2 EOR stations.

$offText

*Declare Node Sets

Set

n 'nodes in the CO2 pipeline network'

/ s0 'Colstrip', s1 'Laramie River', s2 'Jim Bridger', s3 'Coal Creek',

………….

s32 'Gordon Creek', s33 'Moxa Arch', s34 'Woodside Dome', s35 'Cedar Creek',

s36 'Cedar Hill', s37 'Little Knife', s38 'Rough Rider', s39 'Beaver Lodge' /

* arc(n,n) is a dynamic set and only dynamic sets can have double superset. Such definition is required to connect nodes in a proper fashion

i(n) 'Source Nodes' / s0, s1, s2, s3, s4, s5, s9, s10, s11, s12, s13, s14, s15, s16, s17, s18, s19 /

j(n) 'Sink Nodes' /s20, s21, s22, s23, s24, s25, s26, s27, s28, s29, s30, s31, s32, s33, s34, s35, s36, s37, s38, s39/

p(n) 'Both Source & Sink Nodes' /s6, s7, s8/

t 'Pipeline Configuration type' / t12, t16, t24, t32/;

Alias (n,m);

*Declare Variables

Variable

a(i) 'Capture Node binary decision variable'

```
   b(m,n,t) 'Transport arc binary decision variable'
   r(j)     'Sink Node binary Decision Variable'
   X(i)     'Amount captured at each source in Mton of Co2'
   Y(j)     'Amount stored at reservoir annually in Mton of Co2'
   h(j)     'Number of wells'
   Q(m,n,t) 'Amount of CO2 transported through pipe in Mton'
   cost     'Cost of Co2 infrastructure'
   c(p)     'Dual Node Binary'
   Xb(p)    'Amount captured at each dual node in Mton of Co2'
   Yb(p)    'Amount stored at dual  node annually in Mton of Co2'
   hb(p)    'Number of wells'
;
*Declare Variable Type
Binary Variable a,b,r,c;
Positive Variable X,Y,Xb,Yb;
Integer Variable h,hb;
*Fixed Value Inputs
Scalar
   costco2    'Cost of Co2 in million per million ton of co2' /20/
   int        'Interest Rate'                    /0.03/
   time       'Timescale'                        /30/
   MVAC       'MVA Cost per million ton of CO2'   /0.02/
   CapCO2     'Total storage goal in Mton of Co2 yearly' /20/;
```

*Input Variable values for Nodes

Parameter

Ccap(i) 'Capture cost per million ton of Co2 in million$'

/s0 44, s1 44, s2 44, ……., s19 9/

Ccapb(p) 'Dual Capture cost per million ton of Co2 in million$'

/s6 26.5, s7 44, s8 44/

Ctran(m,n,t) 'Transport CAPEX fixed dependent on arc and pipe design in million $'

/s0.s39.t12 60.74, …………… s37.s39.t32 0.52, s38.s39.t32 0.6 /

arcexist(m,n) 'Indicates if arc exists in pre-model'

/s0.s39 1, s0.s31 1, s0.s35 1, s0.s20 1, s0.s21 1, s0.s18 1, s0.s29 1, s0.s13 1, s1.s30 1, ……… s35.s39 1, s35.s36 1, s35.s38 1, s36.s37 1, s36.s38 1, s37.s38 1, s37.s39 1, s38.s39 1 /

Cres(j) 'Reservoir CAPEX in million $'

/s20 0, s21 0, s22 0, s23 0, ……., s36 0, s37 0, s38 0, s39 0 /

Cresb(p) 'Dual Reservoir CAPEX in million $'

/s6 14.87, s7 41.52, s8 14.87/

Cwell(j) 'Well CAPEX in million $'

/s20 0.43, s21 0.44, s22 0.76, s23 0.27, ….., s36 1.21, s37 1.35, s38 1.35, s39 1.79/

Cwellb(p) 'Dual Well CAPEX in million $'

/s6 0.6, s7 3.11, s8 0.6/

Owell(j) 'Well OPEX in million $'

/s20 0.06, s21 0.06, s22 0.08, s23 0.05, …….., s36 0.1, s37 0.1, s38 0.1, s39 0.12 /

Owellb(p) 'Dual Well OPEX in million $'

/s6 0.07, s7 0.13, s8 0.07/

l(j) 'Recycle Rate factor (1-recycle rate)'

/s20 0.1, s21 0.1, s22 0.1, s23 0.1, …….., s36 0.1, s37 0.1, s38 0.1, s39 0.1 /

lb(p) 'Dual Recycle Rate factor (1-recycle rate)'

/s6 0, s7 0, s8 0/

MVA(j) 'Monitoring, Verification and Abandonment cost'

/s20 0, s21 0, s22 0, s23 0, …….., s36 0, s37 0, s38 0, s39 0 /

MVAb(p) 'Dual Monitoring, Verification and Abandonment cost'

/s6 1, s7 1, s8 1/

Qmx(t) 'Maximum flow capacity in MTCO2'

/t12 3.13, t16 6.25, t24 18.75, t32 37.50/

Xmx(i) 'Maximum Capture quantity in million ton of Co2'

/s0 11.39, s1 8.52, s2 8.54, s3 10.38, ……., s16 1.5, s17 1.09, s18 2.5, s19 2.5/

Xmxb(p) 'Dual Maximum Capture quantity in million ton of Co2'

/s6 9.76, s7 5.84, s8 4.38/

Ywellmx(j) 'Maximum storage quantity per well in million ton of Co2'

/s20 0.05, s21 0.06, s22 0.08, s23 0.12, …….., s36 0.12, s37 0.08, s38 0.25, s39 0.15 /

Ywellmxb(p) 'Dual Maximum storage quantity per well in million ton of Co2'

/s6 0.17, s7 0.58, s8 0.24/

Yresmx(j) 'Maximum storage quantity per reservoir in million ton of Co2'

/s20 38, s21 245, s22 81, s23 60, ……….., s36 173, s37 83, s38 76, s39 71 /

Yresmxb(p) 'Dual Maximum storage quantity per reservoir in million ton of Co2'

```
        /s6 126, s7 124, s8 126/;


*Set up of bi-directional pipeline arcs
Set
   arc(m,n) 'arc between two nodes';
arc(m,n)$arcexist(m,n) = yes;


*Objective Function and Constraint declarations
Equations
   Cons1(m,n,t)        'Flow constraint in Pipe for Lower Limit'
………..
   Cons13(p)           'Maximum Input in each dual reservoir node'
   obj                 'Objective Function';
*Detailed Objective functions and Constraints
Cons1(arc(m,n),t)..         Q(arc,t) =g= -Qmx(t)*b(arc,t);
Cons2(arc(m,n),t)..         Q(arc,t) =l= Qmx(t)*b(arc,t);
Cons3(arc(m,n))..           sum(t,b(arc,t)) =l= 1;
Cons4(i)..                X(i) =l= a(i)*Xmx(i);
Cons5(j)..                h(j)*Ywellmx(j) =e= Y(j);
Cons6(j)..                Y(j) =l= r(j)*Yresmx(j)/time;
Cons7..                   sum(i,X(i))+sum(p,Xb(p)) =g=CapCO2;
Cons8(i(n))..              sum((arc(m,n),t),Q(arc,t))-sum((arc(n,m),t),Q(arc,t)) =e= X(i);
Cons9(j(n))..              sum((arc(m,n),t),Q(arc,t))-sum((arc(n,m),t),Q(arc,t)) =e= -Y(j);
```

```
Cons10(p(n))..        sum((arc(m,n),t),Q(arc,t))-sum((arc(n,m),t),Q(arc,t)) =e= Xb(p)-Yb(p);
Cons11(p)..           Xb(p) =l= c(p)*Xmxb(p);
Cons12(p)..           hb(p)*Ywellmxb(p) =e= Yb(p);
Cons13(p)..           Yb(p) =l= c(p)*Yresmxb(p)/time;
obj..                 cost =e= sum(i(n), Ccap(i)*X(i)) + sum(p(n), Ccapb(p)*Xb(p)) +
((int*((1+int)**time))/(((1+int)**time)-1))*sum((arc(m,n),t),Ctran(arc,t)*b(arc,t)) +
sum((arc(m,n),t),Otran(arc,t)*b(arc,t)) + ((int*((1+int)**time))/(((1+int)**time)-
1))*sum(j,r(j)*Cres(j)+h(j)*Cwell(j)) +
sum(j,h(j)*Owell(j)+Y(j)*l(j)*costCO2+Y(j)*MVA(j)*MVAC) +
((int*((1+int)**time))/(((1+int)**time)-1))*sum(p,c(p)*Cresb(p)+hb(p)*Cwellb(p))
+ sum(p,hb(p)*Owellb(p)+Yb(p)*lb(p)*costCO2+Yb(p)*MVAb(p)*MVAC);
*Model Execution using MILP based on CPLEX
Model pipebase /all/;

solve pipebase minimizing cost using mip;
```

B.4.2 Static MILP formulation with Pre-Existent Infrastructure

The code provided below has pre-existent infrastructure. This model is related to Section 3.6.1. No dependencies are required for the execution as the trial version of IBM CPLEX is available on GAMS for execution. The major inputs of this model include the cost models for source, sinks and pipelines and CO_2 capture goals. In the code "*" & "$" indicate comments. The model is similar to the code provided in B.4.1, however with the existence of some constraints on the flow and capture variables. The results of the execution and fixed inputs are provided in Section 4.5. the results for the execution are provided in Appendix E.

```
$title Cost Reduction of Static CCUS networks with Pre-Existent  infrastructure
$onText
This problem solves the transportation of co2 from various sources and sinks provides fixed cost and variable OPEX cost dependent on production of CO2
and injection rates of CO2 into both saline aquifers and Co2 EOR stations.
$offText
*Declaring Nodes and Sets
Set
   n       'nodes in the CO2 pipeline network'
           /  s0 'Colstrip', ……..  s51 'Sink at Canada'/

* arc(n,n) is a dynamic set and only dynamic sets can have double superset. Such definition is required to connect nodes in a proper fashion
   i(n)    'Source Nodes' / s0, s1, s2, s3, s4, s5, s9, s10, s11, s12, s13, s14, s15, s16, s17, s18, s19 /
………
   t       'Pipeline Configuration type' / t12, t16, t24, t32/;
Alias (n,m);

*Declaring Variables
Variable
   a(i)    'Capture Node binary decision variable'
……….
```

```
    hb(p)     'Number of wells';


*Variable Type Definition

Binary Variable a,b,r,c;

Positive Variable X,Y,Xb,Yb;

Integer Variable h,hb;

*Constant Inputs

Scalar

    costco2    'Cost of Co2 in million per million ton of co2' /20/

    ……..

    CapCO2      'Total storage goal in Mton of Co2 yearly' /80/;


*Variable Inputs for Nodes

Parameter

    Ccap(i)     'Capture cost per million ton of Co2 in million$'

            /s0 44, s1 44, s2 44, s3 44, ……………, s16 41, s17 32, s18 9, s19 9/

…………

    Yresmxb(p)     'Dual Maximum storage quantity per reservoir in million ton of Co2'

          /s6 126, s7 124, s8 126/;

Set

    arc(m,n) 'arc between two nodes';

arc(m,n)$arcexist(m,n) = yes;

*Flow and Material Balance Constraints
```

```
Xb.lo('s6')=3.5;

………….

Q.lo('s39','s45','t12')=0.5;


*Objective function and constraint Declaration

Equations

   Cons1(m,n,t)         'Flow constraint in Pipe for Lower Limit'

   ………….

   obj                  'Objective Function';


*Objective Function and constraint equations

Cons1(arc(m,n),t)..          Q(arc,t) =g= -Qmx(t)*b(arc,t);

Cons2(arc(m,n),t)..          Q(arc,t) =l= Qmx(t)*b(arc,t);

Cons3(arc(m,n))..            sum(t,b(arc,t)) =l= 1;

Cons4(i)..                   X(i) =l= a(i)*Xmx(i);

Cons5(j)..                   h(j)*Ywellmx(j) =e= Y(j);

Cons6(j)..                   Y(j) =l= r(j)*Yresmx(j)/time;

Cons7..                      sum(i,X(i))+sum(p,Xb(p)) =g=CapCO2;

Cons8(i(n))..                sum((arc(m,n),t),Q(arc,t))-sum((arc(n,m),t),Q(arc,t)) =e= X(i);

Cons9(j(n))..                sum((arc(m,n),t),Q(arc,t))-sum((arc(n,m),t),Q(arc,t)) =e= -Y(j);

Cons10(p(n))..               sum((arc(m,n),t),Q(arc,t))-sum((arc(n,m),t),Q(arc,t)) =e= Xb(p)-Yb(p);

Cons14(k(n))..               sum((arc(m,n),t),Q(arc,t))-sum((arc(n,m),t),Q(arc,t)) =e= 0;

Cons11(p)..                  Xb(p) =l= c(p)*Xmxb(p);
```

```
Cons12(p)..                    hb(p)*Ywellmxb(p) =e= Yb(p);

Cons13(p)..                    Yb(p) =l= c(p)*Yresmxb(p)/time;

obj..                          cost =e= sum(i(n), Ccap(i)*X(i)) + sum(p(n), Ccapb(p)*Xb(p)) +
((int*((1+int)**time))/(((1+int)**time)-1))*sum((arc(m,n),t),Ctran(arc,t)*b(arc,t))      +
sum((arc(m,n),t),Otran(arc,t)*b(arc,t))
+        ((int*((1+int)**time))/(((1+int)**time)-1))*sum(j,r(j)*Cres(j)+h(j)*Cwell(j))    +
sum(j,h(j)*Owell(j)+Y(j)*l(j)*costCO2+Y(j)*MVA(j)*MVAC)                                  +
((int*((1+int)**time))/(((1+int)**time)-1))*sum(p,c(p)*Cresb(p)+hb(p)*Cwellb(p))
                               +
sum(p,hb(p)*Owellb(p)+Yb(p)*lb(p)*costCO2+Yb(p)*MVAb(p)*MVAC);


*Model Execution
Model pipebase /all/;
solve pipebase minimizing cost using mip;
```

B.5 Dynamic MILP Formulation

This section provides the code developed on GAMS interface to model a MILP formulation to determine the dynamic optimization of a pipeline and CCUS infrastructure network for North-Central USA for 3 different capture goals in fixed time intervals. The code provided below has pre-existent infrastructure. This model is related to Section 3.6.2. No dependencies are required for the execution as the trial version of IBM CPLEX is available on GAMS for execution. The major inputs of this model include the cost models for source, sinks and pipelines and CO_2 capture goals. In the code "*" indicates comments. The results of the execution and fixed inputs are provided in Section 4.6. the results for the execution are provided in Appendix E.

```
$title Cost Reduction of CO2 pipeline Infrastructure

$onText

This problem solves the transportation of co2 from various sources and sinks provides fixed cost
and variable OPEX cost dependent on production of CO2
and injection rates of CO2 into both saline aquifers and Co2 EOR stations.

$offText

*Node and Set Declarations

Set

   n      'nodes in the CO2 pipeline network'

          /  s0 'Colstrip',        ………

             s51 'Sink at Canada'/


* arc(n,n) is a dynamic set and only dynamic sets can have double superset. Such definition is
required to connect nodes in a proper fashion

   i(n)    'Source Nodes' / s0, s1, s2, s3, s4, s5, s9, s10, s11, s12, s13, s14, s15, s16, s17, s18, s19 /

…………..

   k(n)    'Blank Nodes'  /s40, s41, s42, s43, s44, s45, s46, s47/;

Alias (n,m);


*Variable Declarations

Variable

a(i,d)    'Capture Node binary decision variable'

…………….
```

```
hb(f,d)     'Number of wells';


*Variable Type
Binary Variable a,b,r,c,u;
Positive Variable X,Y,Xb,Yb, Capd1, Capd2, Capd3, Trand1, Trand2, Trand3, Injd1, Injd2, Injd3;
Integer Variable h,hb;


*Constant Inputs
Scalar
   costco2    'Cost of Co2 in million per million ton of co2' /20/
   int        'Interest Rate'                    /0.03/
   time        'Timescale'                         /30/
   MVAC        'MVA Cost per million ton of CO2'   /0.02/
   Cap1      'Total storage goal in Mton for first 10 years' /20/
   Cap2      'Total storage goal in Mton for years 10-20' /40/
   Cap3      'Total storage goal in Mton for years 20-30' /60/;


*Variable Inputs for Nodes
Parameter
   ts(d)      'Timeline Period'
         /d1 0, d2 10, d3 20/
   ……………
   Yresmxb(f)     'Dual Maximum storage quantity per reservoir in million ton of Co2'
```

```
        /s6 126, s7 124, s8 126/;


Set
   arc(m,n) 'arc between two nodes';
arc(m,n)$arcexist(m,n) = yes;


*Constraints on Flow and Material balance to mimic pre-existent infrastcrture
Xb.lo('s6',d)=3.5;
……………
Q.lo('s39','s45','t12',d)=0.5;


*Objective Function and Constraint Declaration
Equations
   Cons1(m,n,t,d)        'Flow constraint in Pipe for Lower Limit'
   …………..
   obj                   'Objective Function';


*Objective Function and Constraint equations
Cons22a(j)..                  Y(j,'d2') =g= Y(j,'d1');
Cons22b(j)..                  Y(j,'d3') =g= Y(j,'d2');
Cons23a(f)..                  Yb(f,'d2') =g= Yb(f,'d1');
Cons23b(f)..                  Yb(f,'d3') =g= Yb(f,'d2');
Cons1(arc(m,n),t,d)..           Q(arc,t,d) =g= -Qmx(t)*b(arc,t,d);
```

```
Cons2(arc(m,n),t,d)..          Q(arc,t,d) =l= Qmx(t)*b(arc,t,d);

Cons3a(arc(m,n),d)..           sum((t),b(arc,t,d)) =l= 1;

Cons4a(i,d)..                  X(i,d) =l= a(i,d)*Xmx(i);

Cons4b(i)..                    X(i,'d2') =g= X(i,'d1');

Cons4c(i)..                    X(i,'d3') =g= X(i,'d2');

Cons5(j,d)..                   h(j,d)*Ywellmx(j) =e= Y(j,d);

Cons6a(j,d)..                  Y(j,d) =l= Yresmx(j)/time;

Cons6b(j)..                    sum(d,Y(j,d)) =l=Yresmx(j);

Cons7a..                       sum(i,X(i,'d1'))+sum(p,Xb(p,'d1')) =g=Cap1;

Cons7b..                       sum(i,X(i,'d2'))+sum(p,Xb(p,'d2')) =g=Cap2;

Cons7c..                       sum(i,X(i,'d3'))+sum(p,Xb(p,'d3')) =g=Cap3;

Cons8(i(n),d)..                sum((arc(m,n),t),Q(arc,t,d))-sum((arc(n,m),t),Q(arc,t,d)) =e= X(i,d);

Cons9(j(n),d)..                sum((arc(m,n),t),Q(arc,t,d))-sum((arc(n,m),t),Q(arc,t,d)) =e= -Y(j,d);

Cons10(p(n),f(n),d)..          sum((arc(m,n),t),Q(arc,t,d))-sum((arc(n,m),t),Q(arc,t,d)) =e=
Xb(p,d)-Yb(f,d);

Cons11a(p,d)..                 Xb(p,d) =l= c(p,d)*Xmxb(p);

Cons11b(p)..                   Xb(p,'d2') =g= Xb(p,'d1');

Cons11c(p)..                   Xb(p,'d3') =g= Xb(p,'d2');

Cons12(f,d)..                  hb(f,d)*Ywellmxb(f) =e= Yb(f,d);

Cons13a(f,d)..                 Yb(f,d) =l= Yresmxb(f)/time;

Cons13b(f)..                   sum(d,Yb(f,d)) =l=Yresmxb(f);

Cons14(j)..                    sum(d,r(j,d)) =l= 1;

Cons16(f)..                    sum(d,u(f,d)) =l= 1;
```

```
Cons18a..          Capd1 =e= 8.53*(sum(i(n), Ccap(i)*X(i,'d1')) + sum(p(n),
Ccapb(p)*Xb(p,'d1')));
Cons18b..          Capd2 =e= 8.53*(sum(i(n), Ccap(i)*X(i,'d2')) + sum(p(n),
Ccapb(p)*Xb(p,'d2')));
Cons18c..          Capd3 =e= 8.53*(sum(i(n), Ccap(i)*X(i,'d3')) + sum(p(n),
Ccapb(p)*Xb(p,'d3')));
Cons19a..          Trand1 =e=
sum((arc(m,n),t),Ctran(arc,t)*b(arc,t,'d1'))+19.6*(sum((arc(m,n),t),Otran(arc,t)*b(arc,t,'d1')));
Cons19b..          Trand2 =e=
sum((arc(m,n),t),Ctran(arc,t)*b(arc,t,'d2'))+14.87*(sum((arc(m,n),t),Otran(arc,t)*b(arc,t,'d2')));
Cons19c..          Trand3 =e=
sum((arc(m,n),t),Ctran(arc,t)*b(arc,t,'d3'))+8.53*(sum((arc(m,n),t),Otran(arc,t)*b(arc,t,'d3')));
Cons20a..          Injd1 =e= sum(j,r(j,'d1')*Cres(j)+h(j,'d1')*Cwell(j)) +
8.53*sum(j,h(j,'d1')*Owell(j)+Y(j,'d1')*l(j)*costCO2+Y(j,'d1')*MVA(j)*MVAC) +
sum(f,u(f,'d1')*Cresb(f)+hb(f,'d1')*Cwellb(f))
+ 8.53*sum(f,hb(f,'d1')*Owellb(f)+Yb(f,'d1')*lb(f)*costCO2+Yb(f,'d1')*MVAb(f)*MVAC);
Cons20b..          Injd2 =e= (sum(j,r(j,'d2')*Cres(j)+h(j,'d2')*Cwell(j)) +
8.53*sum(j,h(j,'d2')*Owell(j)+Y(j,'d2')*l(j)*costCO2+Y(j,'d2')*MVA(j)*MVAC) +
sum(f,u(f,'d2')*Cresb(f)+hb(f,'d2')*Cwellb(f))
+ 8.53*sum(f,hb(f,'d2')*Owellb(f)+Yb(f,'d2')*lb(f)*costCO2+Yb(f,'d2')*MVAb(f)*MVAC));
Cons20c..          Injd3 =e= (sum(j,r(j,'d3')*Cres(j)+h(j,'d3')*Cwell(j)) +
8.53*sum(j,h(j,'d3')*Owell(j)+Y(j,'d3')*l(j)*costCO2+Y(j,'d3')*MVA(j)*MVAC) +
sum(f,u(f,'d3')*Cresb(f)+hb(f,'d3')*Cwellb(f))
```

```
 + 8.53*sum(f,hb(f,'d3')*Owellb(f)+Yb(f,'d3')*lb(f)*costCO2+Yb(f,'d3')*MVAb(f)*MVAC));
Cons21a..                CapCost =e= Capd1+Capd2/1.34+Capd3/1.81;
Cons21b..                TranCost =e= Trand1+Trand2/1.34+Trand3/1.81;
Cons21c..                InjCost =e=  Injd1+Injd2/1.34+Injd3/1.81;
Cons24(k(n),d)..           sum((arc(m,n),t),Q(arc,t,d))-sum((arc(n,m),t),Q(arc,t,d)) =e= 0;
obj..                      cost =e= 0.051*((8.53*(sum(i(n), Ccap(i)*X(i,'d1')) + sum(p(n),
Ccapb(p)*Xb(p,'d1'))))
+(8.53*(sum(i(n), Ccap(i)*X(i,'d2')) + sum(p(n), Ccapb(p)*Xb(p,'d2'))))/1.34
+(8.53*(sum(i(n), Ccap(i)*X(i,'d3')) + sum(p(n), Ccapb(p)*Xb(p,'d3'))))/1.81

+Trand1+Trand2/1.34+Trand3/1.81+Injd1+Injd2/1.34+Injd3/1.81);

*Model Execution
Model pipebase /all/;
solve pipebase minimizing cost using mip;
```

B.6 Text Analysis

Text analysis is a set of individual codes for each regulatory clause in USA CFR Title 49 Section 195. The first code utilized under this section is to convert the XML format Section 195 to csv format for easier access to utilization of text analytical tools. The code has been executed on Python 3.7 and utilizes the packages ‘xml.etree.ElementTree’ and ‘Pandas’. The output of the code is a delimited spreadsheet to be utilized as input to NLP.

```
#Import Library and parse xml
import xml.etree.ElementTree as ET
```

```
tree = ET.parse('section195mod.xml')
root = tree.getroot()
import pandas as pd

#Creating relevant clause list
biglist =["§ 195.0", "§ 195.1", "§ 195.2", "§ 195.3", "§ 195.4", "§ 195.5", "§ 195.6", "§ 195.8", "§ 195.9", "§ 195.10", "§ 195.11", "§ 195.12", "§ 195.48", "§ 195.49", "§ 195.50", "§ 195.52", ……….. "§ 195.571", "§ 195.573", "§ 195.575", "§ 195.577", "§ 195.579", "§ 195.581", "§ 195.583", "§ 195.585", "§ 195.587", "§ 195.588", "§ 195.589", "§ 195.591"]

#Looping for extraction of individual hierarchical clauses
testlist=[]
for child in root:
    for a in child:
        for b in a:
            for c in b:
                for d in c:
                    for e in d:
                        for f in e:
                            for g in f:
                                if g.tag=='DIV8':
                                    testlist.append(g)
```

```
#test loop
for child in root:
    for a in child:
        for b in a:
            for c in b:
                for d in c:
                    for e in d:
                        for f in e:
                            for g in f:
                                for h in g:
                                    if g.tag=='DIV8':
                                            print (h.tag)


#Loop for getting header information
headlist=[]
for x in range(134):
   for child in root:
       for a in child:
           for b in a:
               for c in b:
                   for d in c:
                       for e in d:
                           for f in e:
```

```
                    for g in f:
                        for h in g:
                            if g.tag=='DIV8':
                                if g.attrib['N']==biglist[x]:
                                    if h.tag=='HEAD':
                                        headlist.append(h.text)


#Loop for getting P and other information
bodylist=[]
for x in range(134):
    for child in root:
        for a in child:
            for b in a:
                for c in b:
                    for d in c:
                        for e in d:
                            for f in e:
                                for g in f:
                                    for h in g:
                                        if g.tag=='DIV8':
                                            if g.attrib['N']==biglist[x]:
                                                if h.tag=='P' or h.tag=='FP-2' or h.tag=='FP':
                                                    bodylist.append(h.text)
```

```
bodylist2=[]
for val in bodylist:
    if val != None:
        bodylist2.append(val)


#Loop for getting Citations
citlist=[]
for x in range(135):
   for child in root:
       for a in child:
           for b in a:
               for c in b:
                   for d in c:
                       for e in d:
                           for f in e:
                               for g in f:
                                   for h in g:
                                       if g.tag=='DIV8':
                                           if g.attrib['N']==biglist[x]:
                                               if h.tag=='CITA':
                                                   citlist.append(h.text)
```

```
listnumb=[0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2,
2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
2, 2, …………588, 588, 588, 588, 588,  588, 588, 588, 588, 588, 588, 588, 588, 588, 588, 588,
588, 588, 588, 588, 588, 588, 588, 588, 588, 588, 588, 588, 588, 588, 588, 588, 588, 588, 588,
588, 588, 588, 588, 588, 588, 588, 588, 588, 588, 589, 589, 589, 589, 589, 589, 591]

regulation = pd.DataFrame({'Section': listnumb, 'Text': bodylist})
#regulation.to_excel(r'EditingFile.xlsx')
edit_reg=pd.read_excel(r'EditingFile.xlsx')

headnumb = [0, 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 48, 49, 50, 52, 54, 55, 56, 58, 59, 60, 61, 63, 64,
100, 101, 102, 104, 106, 108, 110, 111, 112, 114, 116, 118, 120, 122, 124,
…………………………………
        446, 501, 503, 505, 507, 509, 551, 553, 555, 557, 559, 561, 563, 565, 567, 569, 571, 573,
575, 577, 579, 581, 583, 585, 587, 588, 589, 591]
header = pd.DataFrame({'Section': headnumb, 'Header': headlist })
trial = edit_reg.merge(header, how='left', left_on='Section', right_on='Section')
regulation_DIV8_list = pd.concat([trial,trial7], ignore_index=True)
regulation_DIV8_list = regulation_DIV8_list[['Section', 'sub1', 'sub2', 'sub3', 'sub4', 'Header',
'Text']]
regulation_DIV8_list.to_excel(r'Div8regs195.xlsx')
```

The code for analysis of regulatory clauses needs to be unique for each statement for proper extraction of tags and classification of the clause corpus. An example of the analytical code is provided in the section. The major input for this section is the delimited spreadsheet obtained from parsing the XML format regulatory data. The code is executed on python 3.7 and relies on several depednancies including ‘nltk’, ‘pandas’, ‘numpy’ and ‘StanfordCoreNLP’. Amongst these, the stanfordCoreNLP relies on external servers for operations and hence needs a special *command prompt* based setup of accessing the installed execution location for the package and establishing a stable internet connection to create a host IP for processing texts.

```
#Import Libraries
import nltk
from nltk import sent_tokenize, word_tokenize
from nltk.stem.snowball import SnowballStemmer
import pandas as pd
import re
#Setup of libraries
snowball = SnowballStemmer(language = 'english')
from pycorenlp import StanfordCoreNLP
nlp_wrapper = StanfordCoreNLP('http://localhost:9000')
from nltk.tree import Tree

#Import Data
df=pd.DataFrame(columns=['Section', 'sub1', 'sub2', 'sub3', 'sub4', 'Title', 'Description'])
```

```
regs = pd.read_csv("Div8regs195.csv", encoding = 'utf-8')
relevant=regs.loc[regs['Section']==8]
Title=relevant.iloc[0]['Header']
df=df.append({'Section': 8, 'sub1':0, 'sub2': 0, 'sub3':0, 'sub4':0, 'Title': 'Header', 'Description': Title}, ignore_index=True)
df=df.append({'Section': 8, 'sub1':0, 'sub2': 0, 'sub3':0, 'sub4':0, 'Title': 'Tag', 'Description': 'Restriction'}, ignore_index=True)

#Perform core NLP analysis
sentok=sent_tokenize(relevant.iloc[0]['Text'])
pars=nlp_wrapper.annotate(sentok[0], properties={'annotators': 'tokenize,parse','outputFormat': 'json'})
parsorder=pars["sentences"][0]["parse"]
parsetree=Tree.fromstring(parsorder)

#Setup classification and tagging
df=df.append({'Section': 8, 'sub1':0, 'sub2': 0, 'sub3':0, 'sub4':0, 'Title': 'Clause', 'Description': " ".join(parsetree[0][0].leaves()+parsetree[0][1][0].leaves()+parsetree[0][1][1][0].leaves()+parsetree[0][1][1][1].leaves()+parsetree[0][1][1][2].leaves())}, ignore_index=True)
df=df.append({'Section': 8, 'sub1':0, 'sub2': 0, 'sub3':0, 'sub4':0, 'Title': 'Restriction hazardous fluid pipeline deadline', 'Description': '10-01-1970'}, ignore_index=True)
df=df.append({'Section': 8, 'sub1':0, 'sub2': 0, 'sub3':0, 'sub4':0, 'Title': 'Restriction Carbon-Dioxide pipeline deadline', 'Description': '06-12-1991'}, ignore_index=True)
```

```
df=df.append({'Section': 8, 'sub1':0, 'sub2': 0, 'sub3':0, 'sub4':0, 'Title': 'Exmption', 'Description': " ".join(parsetree[0][1][1][3].leaves())}, ignore_index=True)
df=df.append({'Section': 8, 'sub1':0, 'sub2': 0, 'sub3':0, 'sub4':0, 'Title': 'Exmption Notice Period', 'Description': 90}, ignore_index=True)
df=df.append({'Section': 8, 'sub1':0, 'sub2': 0, 'sub3':0, 'sub4':0, 'Title': 'Exmption Notice Period Unit', 'Description': 'days'}, ignore_index=True)
df=df.append({'Section': 8, 'sub1':0, 'sub2': 0, 'sub3':0, 'sub4':0, 'Title': 'Exmption Notice Content', 'Description': sentok[1]}, ignore_index=True)
df=df.append({'Section': 8, 'sub1':0, 'sub2': 0, 'sub3':0, 'sub4':0, 'Title': 'Exmption Denial Clause', 'Description': sentok[2]}, ignore_index=True)
df=df.append({'Section': 8, 'sub1':0, 'sub2': 0, 'sub3':0, 'sub4':0, 'Title': 'Exmption Denial Clause Period', 'Description': 90}, ignore_index=True)
df=df.append({'Section': 8, 'sub1':0, 'sub2': 0, 'sub3':0, 'sub4':0, 'Title': 'Exmption Denial Clause Period Unit', 'Description': 'days'}, ignore_index=True)
#Generate output spreadsheet
df.to_csv(r'195_8.csv')
```

The code for taking user input in terms of Tags and using these tahs to ffetch relevant clasues.

```
import pandas as pd

import numpy as np

regs = pd.read_excel("Compliation.xlsx")

searchword=input("Enter Tag to be searched")
```

```
regslist=[]

for i in range(len(regs)):

    if regs['Title'][i]=='Tag' and regs['Description'][i]==searchword:

        regslist.append(regs['Section'][i])

print(regslist)

selectsection=input("Enter Clause number from displayed search list")

df=regs.loc[regs['Section']==int(selectsection)]

out=df.drop(['Section','sub1','sub2','sub3','sub4'],axis=1)

print(out)
```

APPENDIX C

Techno-Economic Analysis

C.1 Introduction

This Appendix showcases data tables related to models related to CO_2 properties as well as the cost analysis generated for sources, pipeline, and sinks.

C.2 CO_2 Properties

This section showcases the data tables used to iterate the high and low values of CO_2 density (ρ) and viscosity (μ).Table C.1 shows the regression coefficients for CO_2 density calculation, while table C.2 shows the regression coefficients for CO_2 viscosity calculations. These calculations are referred to in Chapter 3 and 4 of the theses.

C.3 CO_2 Source Technical and Cost Analysis

Analysis of CO_2 sources here enumerate the various clustered sources described in Section 4.3. In addition, details related to the type of source and it's relative maximum emission rate/ supply rate is provided. Another additional parameter provided is cost related to capturing a ton of CO_2.

Table C.1: CO_2 Density regression coefficients (Ogden et al., 2006).

Temperature (°C)	Regression Equation Coefficient						
	a_1	a_2	a_3	a_4	a_5	a_6	a_7
-1.1	-3.128E-07	3.248E-05	-1.439E-03	3.675E-02	-6.572E-01	1.205E+01	8.988E+02
4.4	-9.548E-08	1.979E-05	-1.414E-03	5.070E-02	-1.077E+00	1.771E+01	8.428E+02
10.0	-6.993E-07	8.561E-05	-4.412E-03	1.255E-01	-2.199E+00	2.820E+01	7.686E+02
15.6	-2.930E-07	6.573E-05	-4.755E-03	1.676E-01	-3.320E+00	4.211E+01	6.706E+02
21.1	-7.864E-06	8.728E-04	-4.028E-02	9.977E-01	-1.428E+01	1.218E+02	3.842E+02
26.7	-4.149E-05	4.437E-03	-1.954E-01	4.550E+00	-5.961E+01	4.302E+02	-5.263E+02
32.2	-1.103E-03	1.135E-01	-4.767E+00	1.045E+02	-1.261E+03	7.948E+03	-1.971E+04
37.8	-5.429E-04	5.981E-02	-2.708E+00	6.445E+01	-8.509E+02	5.926E+03	-1.632E+04
43.3	9.609E-04	-9.444E-02	3.735E+00	-7.541E+01	8.076E+02	-4.212E+03	8.422E+03
48.9	1.030E-03	-1.052E-01	4.362E+00	-9.331E+01	1.077E+03	-6.233E+03	1.427E+04
54.4	4.919E-04	-5.207E-02	2.329E+00	-5.290E+01	6.487E+02	-3.972E+03	9.613E+03
60.0	1.783E-05	-5.256E-03	3.796E-01	-1.200E+01	1.862E+02	-1.322E+03	3.607E+03
65.6	-2.014E-04	1.793E-02	-6.142E-01	9.953E+00	-7.502E+01	2.483E+02	-1.205E+02
71.1	-2.273E-04	2.177E-02	-8.255E-01	1.563E+01	-1.537E+02	7.788E+02	-1.492E+03
76.7	-1.723E-04	1.711E-02	-6.760E-01	1.343E+01	-1.399E+02	7.578E+02	-1.563E+03
82.2	-1.040E-04	1.071E-02	-4.387E-01	9.024E+00	-9.704E+01	5.475E+02	-1.158E+03

Table C.2: CO_2 viscosity regression coefficients (Ogden et al., 2006).

Temperature (°C)	Regression Equation Coefficient						
	b_1	b_2	b_3	b_4	b_5	b_6	b_7
-1.1	-3.765E-14	4.427E-12	-2.219E-10	6.353E-09	-1.201E-07	3.212E-06	9.699E-05
4.4	-4.132E-11	5.058E-12	-2.672E-10	8.101E-09	-1.597E-07	3.686E-06	8.534E-05
10.0	-1.801E-13	1.969E-11	-9.099E-10	2.333E-08	-3.708E-07	5.353E-06	7.071E-05
15.6	-3.837E-13	4.250E-11	-1.974E-09	4.999E-08	-7.544E-07	8.426E-06	5.178E-05
21.1	-9.835E-13	1.085E-10	-4.979E-09	1.227E-07	-1.751E-06	1.586E-05	2.015E-05
26.7	-4.043E-12	4.324E-10	-1.907E-08	4.457E-07	-5.877E-06	4.396E-05	-6.756E-05
32.2	2.278E-10	-2.271E-08	9.154E-07	-1.899E-05	2.122E-04	-1.197E-03	2.684E-05
37.8	9.445E-11	-9.374E-09	3.753E-07	-7.700E-06	8.444E-05	-4.576E-04	9.694E-04
43.3	4.615E-11	-4.645E-09	1.894E-07	-3.983E-06	4.499E-05	-2.504E-04	5.508E-04
48.9	2.174E-11	-2.273E-09	9.721E-08	-2.167E-06	2.624E-05	-1.573E-04	3.810E-04
54.4	1.751E-11	-1.839E-09	7.909E-08	1.776E-06	2.178E-05	-1.329E-04	3.320E-04
60.0	1.594E-11	-1.663E-09	7.090E-08	1.580E-06	1.929E-05	-1.179E-04	2.991E-04
65.6	1.331E-11	-1.382E-09	5.864E-08	-1.301E-06	1.587E-05	-9.746E-05	2.523E-04
71.1	9.596E-12	-9.946E-10	4.212E-08	-9.351E-07	1.148E-05	-7.098E-05	1.905E-04
76.7	4.940E-12	-5.141E-10	2.193E-08	-4.944E-07	6.233E-06	-3.935E-05	1.154E-04
82.2	8.355E-13	-9.235E-11	4.291E-09	-1.102E-07	1.664E-06	-1.168E-05	4.941E-05

Table C.3: Source nodes considered in the analysis for CO_2 emissions and supply along with the type of source, CO_2 emission/supply rate (EPA), maximum supply quantity and cost of capture (Section 3.5).

Node ID	Longitude (degrees)	Latitude (degrees)	Source	Type of Source	CO_2 Output (Mton/yr)	Maximum Supply Quantity (Mton)	Cost of Capture (USD 2019/ton)
0	-106.61	45.88	Colstrip	PC Supercritical	14.24	11.3939423	44
1	-104.88	42.11	Laramie River	PC Supercritical	10.65	8.52159198	44
2	-108.79	41.74	Jim Bridger	PC Supercritical	10.67	8.535412461	44
3	-101.16	47.38	Coal Creek & + Leland Olds	PC Supercritical	12.98	10.38174108	44
4	-111.03	39.17	Hunter & Huntington	PC Supercritical	12.98	10.38463835	44
5	-112.58	39.51	Intermountain	PC Supercritical	6.75	5.398066622	44
6	-101.84	47.37	Great Plains Gasification & Antelope Valley	Natural Gas & PC Supercritical	10.84	9.760060631	26.5
7	-107.59	40.46	Craig	PC Supercritical	7.30	5.841617408	44
8	-101.21	47.07	Milton R. Young	PC Supercritical	5.47	4.379899012	44
9	-105.78	42.84	Dave Johnston	PC Supercritical	5.26	4.210483561	44
10	-110.60	41.76	Naughton	PC Supercritical	5.08	4.061791829	44
11	-103.68	40.22	Pawnee	PC Supercritical	3.91	3.127931168	44
12	-110.22	42.24	Shute Creek	Natural Gas	3.07	2.454264255	9
13	-105.39	44.29	Wyodak, Wygen, Neil Simpson, Dry Fork Station	PC Supercritical	10.47	8.37478152	44
14	-101.81	47.22	Coyote	PC Supercritical	2.77	2.213280477	44
15	-107.19	40.49	Hayden	PC Supercritical	2.42	1.933333067	44
16	-105.03	40.86	Rawhide Energy	Oxy Combustion	1.87	1.497770404	41
17	-104.88	40.24	Fort St. Vrain	NGCC	1.36	1.089395209	32
18	-107.60	43.27	Lost Cabin	Natural Gas	2.50	2.50	9
19	-110.42	42.50	Riley Ridge	Natural Gas	2.50	2.50	9

Note: In this table the node ID used is similar to the one indicated in Table 4.7. The type of source is based on emission generation/supply type, where PC Supercritical stands for Pulverized Coal supercritical used commonly in coal-fired power plants, Natural Gas is natural gas power plant/sourced from natural gas and NGCC is Natural Gas Combined Cycle.

C.4 Pipeline Technical and Cost Analysis

Technical Analysis related to pipe configurations and related calculations can be found in Section 3.5.3. Table C.4 describes the different pipeline diameters used as configuration in this study along with other technical parameters related to individual diameters. These technical parameters are generalized throughout the analysis of pipelines in this study and is used as inputs in the optimization study in Sections 4.4, 4.5 and 4.6.

Table C.4: Technical specification for different pipeline diameters.

Pipe Outer Diameter (in)	Mass Flow Rate (Mt/yr)	Maximum Flow Rate (Mt/yr)	Minimum Flow Rate (Mt/yr)	Thickness (in)	Pipe Inner Diameter (in)	Reynolds Number	Fanning's Friction Factor
12	2.50	3.13	0.00	0.26	11.47	0.00	0.05
16	5.00	6.25	3.13	0.35	15.30	0.01	0.07
24	15.00	18.75	6.25	0.53	22.94	0.05	0.14
32	30.00	37.50	18.75	0.70	30.59	0.12	0.25
40	50.00	62.50	37.50	0.88	38.24	0.26	0.40

The cost analysis for the pipeline analysis adopted from Rui et al. 92011) is shown in Section 3.5. Table C.5 depicts the cost analysis for pipelines of diameter 12in, 16in, 24in and 32in. Table C.6 shows cost analysis for pipeline with existent infrastructure.

Table C.5: Cost analysis of pipeline routes for different pipeline diameters.

Start Node	End Node	Pipe Length (miles)	12'' Pipeline CAPEX (Million USD 2019)	12'' Pipeline OPEX (Million USD 2019)	16'' Pipeline CAPEX (Million USD 2019)	16'' Pipeline OPEX (Million USD 2019)	24'' Pipeline CAPEX (Million USD 2019)	24'' Pipeline OPEX (Million USD 2019)	32'' Pipeline CAPEX (Million USD 2019)	32'' Pipeline OPEX (Million USD 2019)
0	39	251.02	60.74	2.56	89.86	3.11	151.00	5.33	225.70	8.65
0	31	308.77	73.10	3.07	108.41	3.66	183.88	6.05	276.84	9.63
0	35	141.77	37.21	1.64	54.98	2.15	91.77	4.20	136.37	7.26
0	20	94.70	25.67	0.75	37.58	0.75	60.72	0.75	87.87	0.75
0	21	158.85	40.84	1.78	60.28	2.29	100.33	4.33	148.78	7.40
0	18	218.23	54.21	2.30	80.33	2.85	135.64	5.07	203.44	8.39
0	29	186.49	46.60	2.00	68.69	2.51	113.91	4.55	168.44	7.62
0	13	134.06	35.55	1.58	52.55	2.09	87.85	4.13	130.69	7.20
1	30	404.54	92.21	3.87	136.78	4.51	232.54	7.07	350.85	10.91
1	36	312.31	73.78	3.09	109.40	3.69	185.46	6.08	279.14	9.66
1	11	157.07	40.47	1.77	59.73	2.28	99.45	4.32	147.49	7.38
1	13	164.47	42.02	1.82	62.01	2.33	103.12	4.38	152.81	7.44
1	9	71.83	20.35	0.57	29.80	0.57	48.15	0.57	69.66	0.57
1	16	88.71	24.30	0.71	35.57	0.71	57.48	0.71	83.17	0.71
2	26	17.16	6.15	0.14	9.01	0.14	14.54	0.14	20.99	0.14
2	25	83.25	23.04	0.66	33.73	0.66	54.50	0.66	78.85	0.66
2	27	125.08	33.59	1.51	49.70	2.02	83.24	4.06	124.02	7.13
2	12	69.39	19.77	0.55	28.95	0.55	46.77	0.55	67.66	0.55
3	39	118.24	32.09	1.46	47.51	1.97	79.70	4.01	118.88	7.07
3	6	32.01	10.34	0.26	15.15	0.26	24.47	0.26	35.36	0.26
3	8	22.36	7.66	0.18	11.23	0.18	18.13	0.18	26.19	0.18
3	14	34.27	10.95	0.27	16.04	0.27	25.90	0.27	37.44	0.27
4	32	32.97	10.60	0.26	15.53	0.26	25.08	0.26	36.25	0.26
4	5	92.65	25.20	0.74	36.90	0.74	59.62	0.74	86.27	0.74
4	34	38.09	11.96	0.30	17.52	0.30	28.30	0.30	40.91	0.30
5	19	267.91	64.06	2.69	94.69	3.25	158.80	5.46	236.99	8.78
5	10	207.43	52.03	2.21	77.15	2.77	130.50	4.98	196.01	8.30
5	32	94.23	25.56	0.75	37.42	0.75	60.47	0.75	87.50	0.75
6	39	106.49	29.48	1.36	43.69	1.87	73.53	3.91	109.95	6.98
6	37	58.94	17.24	0.47	25.25	0.47	40.79	0.47	59.00	0.47
6	14	10.27	4.01	0.08	5.88	0.08	9.48	0.08	13.68	0.08
6	30	40.42	12.57	0.32	18.41	0.32	29.74	0.32	43.00	0.32
7	15	21.45	7.40	0.17	10.85	0.17	17.51	0.17	25.29	0.17
7	27	77.93	21.79	0.62	31.91	0.62	51.56	0.62	74.59	0.62

7	26	97.42	26.29	0.78	38.48	0.78	62.18	0.78	89.99	0.78
8	30	53.94	16.00	0.43	23.44	0.43	37.87	0.43	54.77	0.43
8	14	31.48	10.20	0.25	14.94	0.25	24.13	0.25	34.88	0.25
9	13	110.83	30.45	1.40	45.11	1.91	75.82	3.95	113.27	7.01
9	29	48.98	14.76	0.39	21.62	0.39	34.93	0.39	50.51	0.39
9	16	159.26	40.93	1.78	60.41	2.29	100.54	4.34	149.07	7.40
9	15	193.26	47.99	2.05	70.72	2.56	117.18	4.61	173.18	7.67
9	24	106.97	29.59	1.37	43.85	1.88	73.79	3.92	110.33	6.98
9	22	64.23	18.53	0.51	27.13	0.51	43.84	0.51	63.41	0.51
10	19	61.33	17.82	0.49	26.10	0.49	42.17	0.49	61.00	0.49
10	32	175.68	44.37	1.91	65.43	2.42	108.64	4.47	160.80	7.53
10	12	30.15	9.83	0.24	14.41	0.24	23.27	0.24	33.63	0.24
11	17	63.93	18.45	0.51	27.03	0.51	43.67	0.51	63.17	0.51
11	16	91.20	24.87	0.73	36.41	0.73	58.83	0.73	85.13	0.73
12	27	151.01	39.18	1.72	57.86	2.23	96.42	4.27	143.11	7.33
12	32	209.28	52.40	2.23	77.70	2.78	131.38	4.99	197.28	8.32
12	19	51.41	15.37	0.41	22.52	0.41	36.38	0.41	52.61	0.41
12	33	99.20	26.69	0.79	39.07	0.79	63.13	0.79	91.36	0.79
12	25	120.02	32.48	1.47	48.08	1.98	80.63	4.02	120.23	7.09
13	29	75.43	21.20	0.60	31.05	0.60	50.16	0.60	72.58	0.60
13	20	80.07	22.29	0.64	32.64	0.64	52.74	0.64	76.31	0.64
13	36	162.94	41.70	1.81	61.54	2.32	102.36	4.36	151.72	7.43
14	30	32.96	10.59	0.26	15.52	0.26	25.07	0.26	36.24	0.26
15	34	230.36	56.64	2.40	83.88	2.95	141.36	5.16	211.73	8.48
15	27	97.81	26.38	0.78	38.61	0.78	62.39	0.78	90.29	0.78
15	17	130.36	34.74	1.55	51.38	2.06	85.96	4.11	127.95	7.17
15	16	118.34	32.11	1.46	47.54	1.97	79.75	4.01	118.96	7.07
15	24	122.56	33.04	1.49	48.89	2.00	81.94	4.04	122.13	7.11
16	17	46.43	14.11	0.37	20.67	0.37	33.40	0.37	48.30	0.37
17	34	319.15	75.08	3.15	111.30	3.75	188.53	6.13	283.58	9.71
18	25	51.63	15.42	0.41	22.59	0.41	36.50	0.41	52.79	0.41
18	23	67.75	19.37	0.54	28.37	0.54	45.84	0.54	66.32	0.54
18	29	67.80	19.39	0.54	28.39	0.54	45.87	0.54	66.36	0.54
18	21	128.12	34.26	1.53	50.67	2.05	84.81	4.09	126.29	7.15
18	24	72.89	20.60	0.58	30.17	0.58	48.75	0.58	70.52	0.58
18	22	49.13	14.80	0.39	21.68	0.39	35.02	0.39	50.64	0.39
19	33	57.56	16.90	0.46	24.75	0.46	39.99	0.46	57.84	0.46
20	35	98.44	26.52	0.78	38.82	0.78	62.73	0.78	90.78	0.78
20	36	83.63	23.12	0.67	33.85	0.67	54.70	0.67	79.15	0.67
21	33	139.01	36.62	1.62	54.11	2.13	90.37	4.17	134.35	7.24

21	31	313.58	74.02	3.10	109.75	3.70	186.04	6.09	279.97	9.67
21	28	47.24	14.32	0.38	20.98	0.38	33.89	0.38	49.00	0.38
21	23	77.32	21.65	0.62	31.70	0.62	51.22	0.62	74.10	0.62
22	29	59.40	17.35	0.47	25.41	0.47	41.05	0.47	59.38	0.47
22	24	44.02	13.50	0.35	19.77	0.35	31.94	0.35	46.18	0.35
23	25	88.25	24.19	0.70	35.42	0.70	57.23	0.70	82.81	0.70
23	33	97.65	26.34	0.78	38.56	0.78	62.31	0.78	90.17	0.78
23	28	44.35	13.58	0.35	19.90	0.35	32.14	0.35	46.48	0.35
24	25	60.99	17.74	0.49	25.98	0.49	41.97	0.49	60.71	0.49
24	26	71.33	20.23	0.57	29.62	0.57	47.87	0.57	69.25	0.57
25	26	92.93	25.27	0.74	36.99	0.74	59.77	0.74	86.49	0.74
25	33	113.48	31.04	1.42	45.97	1.93	77.21	3.97	115.29	7.03
26	27	113.20	30.98	1.42	45.88	1.93	77.07	3.97	115.07	7.03
27	32	165.21	42.18	1.83	62.24	2.34	103.49	4.38	153.35	7.45
27	34	130.49	34.77	1.55	51.42	2.06	86.02	4.11	128.05	7.17
28	33	96.21	26.01	0.77	38.08	0.77	61.53	0.77	89.04	0.77
30	36	92.77	25.23	0.74	36.93	0.74	59.68	0.74	86.36	0.74
30	37	50.54	15.15	0.40	22.19	0.40	35.86	0.40	51.85	0.40
32	34	37.62	11.83	0.30	17.34	0.30	28.01	0.30	40.49	0.30
35	39	138.30	36.46	1.62	53.89	2.13	90.01	4.17	133.82	7.23
35	36	40.74	12.65	0.32	18.53	0.32	29.94	0.32	43.28	0.32
35	38	64.92	18.69	0.52	27.38	0.52	44.23	0.52	63.98	0.52
36	37	92.39	25.14	0.74	36.81	0.74	59.48	0.74	86.07	0.74
36	38	82.15	22.78	0.65	33.35	0.65	53.89	0.65	77.97	0.65
37	38	23.33	7.94	0.19	11.64	0.19	18.79	0.19	27.14	0.19
37	39	65.81	18.91	0.52	27.69	0.52	44.74	0.52	64.72	0.52
38	39	74.76	21.04	0.60	30.82	0.60	49.79	0.60	72.04	0.60

Table C.6: Cost Analysis of pipeline routes for different pipe diameters in the case with existent pipeline.

Start Node	End Node	Pipe Length (miles)	12'' Pipeline CAPEX (Million USD 2019)	12'' Pipeline OPEX (Million USD 2019)	16'' Pipeline CAPEX (Million USD 2019)	16'' Pipeline OPEX (Million USD 2019)	24'' Pipeline CAPEX (Million USD 2019)	24'' Pipeline OPEX (Million USD 2019)	32'' Pipeline CAPEX (Million USD 2019)	32'' Pipeline OPEX (Million USD 2019)
0	18	218.23	54.21	2.77	82.75	2.85	140.45	5.07	211.79	8.39
0	20	94.70	25.67	0.75	38.77	0.75	63.11	0.75	92.04	0.75
0	21	157.54	40.57	2.24	61.71	2.28	103.34	4.32	154.19	7.38
0	31	308.77	73.10	3.53	111.67	3.66	190.34	6.05	288.02	9.63
0	35	142.16	37.29	2.11	56.77	2.16	95.32	4.20	142.49	7.26
1	9	71.83	20.35	0.57	30.74	0.57	50.05	0.57	72.98	0.57
1	11	156.79	40.41	2.23	61.47	2.27	102.95	4.32	153.63	7.38
1	13	164.13	41.95	2.29	63.80	2.33	106.74	4.37	159.15	7.44
1	16	88.74	24.31	0.71	36.71	0.71	59.76	0.71	87.15	0.71
1	30	404.31	92.17	4.34	140.82	4.51	240.58	7.07	364.74	10.91
1	36	312.58	73.83	3.56	112.76	3.69	192.12	6.08	290.61	9.66
2	26	17.16	6.15	0.14	9.29	0.14	15.12	0.14	22.02	0.14
2	27	124.12	33.38	1.97	50.89	2.01	85.74	4.06	128.52	7.12
2	4	15.26	5.58	0.12	8.43	0.12	13.72	0.12	19.97	0.12
3	6	32.01	10.34	0.26	15.63	0.26	25.44	0.26	37.07	0.26
3	8	22.36	7.66	0.18	11.58	0.18	18.85	0.18	27.47	0.18
4	5	92.65	25.20	0.74	38.06	0.74	61.96	0.74	90.36	0.74
4	32	34.53	11.02	0.28	16.65	0.28	27.11	0.28	39.50	0.28
4	34	38.09	11.96	0.30	18.07	0.30	29.42	0.30	42.88	0.30
5	10	207.13	51.97	2.68	79.37	2.76	134.97	4.98	203.79	8.30
5	19	267.51	63.98	3.16	97.45	3.24	164.34	5.46	246.63	8.78
5	32	94.24	25.57	0.75	38.61	0.75	62.85	0.75	91.66	0.75
6	14	10.27	4.01	0.08	6.06	0.08	9.87	0.08	14.36	0.08
7	15	21.45	7.40	0.17	11.19	0.17	18.21	0.17	26.53	0.17
7	24	126.15	33.83	1.99	51.55	2.03	86.83	4.07	130.10	7.13
7	26	97.42	26.29	0.78	39.70	0.78	64.63	0.78	94.26	0.78
7	27	78.23	21.86	0.62	33.02	0.62	53.76	0.62	78.39	0.62
8	14	31.48	10.20	0.25	15.41	0.25	25.09	0.25	36.57	0.25
8	30	54.29	16.09	0.43	24.31	0.43	39.58	0.43	57.70	0.43
9	13	110.91	30.47	1.86	46.49	1.91	78.59	3.95	118.08	7.01
9	15	193.26	47.99	2.52	72.89	2.56	121.53	4.61	180.73	7.67
9	16	158.95	40.86	2.25	62.15	2.29	104.07	4.33	155.25	7.40

9	22	63.73	18.40	0.51	27.80	0.51	45.27	0.51	66.01	0.51
9	24	106.65	29.52	1.83	45.06	1.87	76.26	3.92	114.68	6.98
9	29	48.98	14.76	0.39	22.30	0.39	36.31	0.39	52.93	0.39
10	19	61.80	17.94	0.49	27.10	0.49	44.12	0.49	64.33	0.49
10	32	177.33	44.71	2.39	67.95	2.44	113.48	4.48	168.99	7.54
10	12	30.08	9.82	0.24	14.84	0.24	24.16	0.24	35.20	0.24
11	16	91.29	24.89	0.73	37.59	0.73	61.19	0.73	89.24	0.73
11	17	64.29	18.54	0.51	28.01	0.51	45.60	0.51	66.49	0.51
12	19	51.22	15.32	0.41	23.15	0.41	37.69	0.41	54.95	0.41
13	20	80.29	22.35	0.64	33.75	0.64	54.95	0.64	80.13	0.64
13	29	75.43	21.20	0.60	32.02	0.60	52.14	0.60	76.03	0.60
13	36	162.94	41.70	2.28	63.42	2.32	106.13	4.36	158.26	7.43
13	41	13.93	5.17	0.11	7.81	0.11	12.71	0.11	18.50	0.11
14	30	32.96	10.59	0.26	16.01	0.26	26.07	0.26	37.99	0.26
15	16	118.54	32.16	1.93	49.04	1.97	82.74	4.01	124.13	7.07
15	17	131.02	34.89	2.03	53.15	2.07	89.43	4.11	133.90	7.17
15	24	122.69	33.07	1.96	50.41	2.00	84.97	4.04	127.39	7.11
15	27	97.82	26.38	0.78	39.84	0.78	64.85	0.78	94.58	0.78
15	34	230.21	56.61	2.86	86.36	2.95	146.33	5.16	220.36	8.48
16	17	46.71	14.18	0.37	21.43	0.37	34.90	0.37	50.87	0.37
17	34	320.52	75.34	3.63	115.04	3.76	195.81	6.14	296.00	9.72
18	21	128.25	34.29	2.00	52.25	2.05	87.95	4.09	131.74	7.15
18	22	49.13	14.80	0.39	22.36	0.39	36.40	0.39	53.07	0.39
18	23	67.23	19.25	0.54	29.08	0.54	47.34	0.54	69.03	0.54
18	25	51.63	15.42	0.41	23.31	0.41	37.95	0.41	55.32	0.41
19	33	58.72	17.18	0.47	25.96	0.47	42.27	0.47	61.63	0.47
20	35	98.17	26.46	0.78	39.95	0.78	65.05	0.78	94.86	0.78
20	36	83.63	23.12	0.67	34.92	0.67	56.85	0.67	82.91	0.67
21	23	77.32	21.65	0.62	32.70	0.62	53.23	0.62	77.63	0.62
21	28	47.56	14.40	0.38	21.76	0.38	35.43	0.38	51.64	0.38
21	31	314.90	74.27	3.58	113.43	3.71	193.20	6.10	292.19	9.68
21	33	139.19	36.65	2.09	55.81	2.13	93.76	4.18	140.21	7.24
22	24	44.02	13.50	0.35	20.39	0.35	33.21	0.35	48.40	0.35
23	25	87.65	24.05	0.70	36.33	0.70	59.14	0.70	86.25	0.70
23	28	44.35	13.58	0.35	20.52	0.35	33.42	0.35	48.71	0.35
23	33	98.76	26.59	0.79	40.16	0.79	65.37	0.79	95.34	0.79
24	26	71.33	20.23	0.57	30.56	0.57	49.75	0.57	72.55	0.57
25	33	113.68	31.08	1.89	47.42	1.93	80.10	3.97	120.29	7.04
26	27	113.20	30.98	1.88	47.26	1.93	79.84	3.97	119.90	7.03

26	42	30.57	0.00	0.24	0.00	0.24	0.00	0.24	0.00	0.24
27	32	165.48	42.24	2.30	64.22	2.34	107.43	4.38	160.16	7.45
27	34	130.35	34.74	2.02	52.93	2.06	89.07	4.10	133.37	7.17
28	33	95.80	25.92	0.76	39.14	0.76	63.72	0.76	92.93	0.76
30	36	92.78	25.23	0.74	38.10	0.74	62.03	0.74	90.46	0.74
30	37	50.54	15.15	0.40	22.89	0.40	37.27	0.40	54.34	0.40
32	34	37.62	11.83	0.30	17.88	0.30	29.12	0.30	42.44	0.30
35	36	40.74	12.65	0.32	19.12	0.32	31.12	0.32	45.37	0.32
35	38	64.92	18.69	0.52	28.24	0.52	45.98	0.52	67.04	0.52
36	37	92.39	25.14	0.74	37.97	0.74	61.82	0.74	90.15	0.74
36	38	82.15	22.78	0.65	34.40	0.65	56.01	0.65	81.68	0.65
37	38	23.33	7.94	0.19	12.00	0.19	19.54	0.19	28.47	0.19
37	39	66.15	18.99	0.53	28.69	0.53	46.71	0.53	68.10	0.53
37	44	13.98	5.18	0.11	7.83	0.11	12.75	0.11	18.56	0.11
38	39	74.76	21.04	0.60	31.79	0.60	51.75	0.60	75.47	0.60
38	44	37.88	11.90	0.30	17.99	0.30	29.29	0.30	42.68	0.30
39	45	2.49	1.24	0.02	1.87	0.02	3.04	0.02	4.42	0.02
6	8	36.23	11.47	0.29	17.33	0.29	28.21	0.29	41.12	0.29
30	47	38.08	11.95	0.30	18.07	0.30	29.41	0.30	42.87	0.30
6	43	20.96	0.00	0.17	0.00	0.17	0.00	0.17	0.00	0.17
43	44	30.13	0.00	0.24	0.00	0.24	0.00	0.24	0.00	0.24
44	45	57.82	0.00	0.46	0.00	0.46	0.00	0.46	0.00	0.46
45	51	56.50	0.00	0.45	0.00	0.45	0.00	0.45	0.00	0.45
27	46	132.12	0.00	2.03	0.00	2.08	0.00	4.12	0.00	7.18
12	46	24.64	0.00	0.20	0.00	0.20	0.00	0.20	0.00	0.20
40	46	49.63	0.00	0.40	0.00	0.40	0.00	0.40	0.00	0.40
40	42	8.28	0.00	0.07	0.00	0.07	0.00	0.07	0.00	0.07
42	47	52.73	0.00	0.42	0.00	0.42	0.00	0.42	0.00	0.42
19	47	159.90	0.00	2.26	0.00	2.30	0.00	4.34	0.00	7.40
24	47	19.71	0.00	0.16	0.00	0.16	0.00	0.16	0.00	0.16
25	47	45.5117	0.00	0.36	0.00	0.36	0.00	0.36	0.00	0.36
18	47	70.4001	0.00	0.56	0.00	0.56	0.00	0.56	0.00	0.56
22	47	52.7774	0.00	0.42	0.00	0.42	0.00	0.42	0.00	0.42
18	41	166.89	0.00	2.31	0.00	2.35	0.00	4.40	0.00	7.46
22	29	72.1946	0.00	0.58	0.00	0.58	0.00	0.58	0.00	0.58
20	41	75.3226	0.00	0.60	0.00	0.60	0.00	0.60	0.00	0.60

C.5 CO_2 Storage sites

Analysis of CO_2 sink here enumerate the various clustered sinks described in Section 4.3. In addition, details related to the type of sink, reservoir properties and cost for operating a single well in the facility is provided. The cost analysis is obtained from Section 3.5.

Table C.7: Sink nodes in analysis with reservoir properties and associated well costs.

Node ID	Sink	Type	Pressure (psi)	Thickness (ft)	Depth (ft)	Permeability (mD)	Capacity ($MtCO_2$)	Injection rate (tCO_2/well/yr)	CAPEX (M$/well)	OPEX (M$ /well)
6	Great Plains Gasification & Antelope Valley	Saline	2372	124	5100	315.1	126	0.17	0.60	0.07
7	Craig	Saline	5400	185	11528	50	124	0.58	3.11	0.13
8	Milton R. Young	Saline	2372	169	5100	315.1	126	0.24	0.60	0.07
20	Bell Creek	EOR	1572	37.5	4500	662.5	38	0.05	0.43	0.06
21	Elk Basin	EOR	2264	46	4600	207.5	245	0.06	0.44	0.06
22	Grieve Field	EOR	2068	45	6900	176	81	0.08	0.76	0.08
23	Hamilton Dome	EOR	750	184	2400	60	60	0.12	0.27	0.05
24	Lost Soldier & Wertz	EOR	3500	223	5600	25.5	300	0.34	0.56	0.07
25	Beaver Creek	EOR	5300	212	11235	9	120	0.65	2.27	0.13
26	Monell (Patrick Draw)	EOR	1800	25	4500	30	43	0.03	0.43	0.06
27	Rangeley Field	EOR	2750	189	5300	8	126	0.27	0.52	0.07
28	Spring Creek Field, Oregon Basin & Pitchfork Field	EOR	1560	80	3557	125	40	0.08	0.35	0.06
29	Salt Creek	EOR	1000	85	2000	52	227	0.05	0.25	0.05
30	Red Trail Energy	Saline	2976	124	6400	315.1	126	0.22	0.83	0.08
31	Kevin Dome	Saline	1488	300	3200	20	60	0.26	0.38	0.06
32	Gordon Creek	Saline	3900	250	6585	50	150	0.45	0.87	0.08
33	Moxa arch	Saline	4000	170	9800	100	171	0.45	1.98	0.11
34	Woodside Dome	Saline	4992	280	4539	100	103	0.35	0.52	0.07
35	Cedar Creek Anticline	EOR	4400	75.25	9000	175	245	0.18	1.28	0.10
36	Cedar Hill	EOR	4085	49	8785	10	173	0.12	1.21	0.10
37	Little Knife	EOR	4400	31	9200	30	83	0.08	1.35	0.10
38	Rough Rider	EOR	4278	100	9200	0.6	76	0.25	1.35	0.10
39	Beaver Lodge	EOR	4790	55	10300	100	71	0.15	1.79	0.12

APPENDIX D

Market Cost Indices

D.1 Introduction

This Appendix showcases data tables related to market cost indices, in order to bring the costs of various factors in the techno-economic analysis of this thesis to 2019 USD equivalent.

D.2 Upstream Capital Cost Index

Upstream Capital Cost Index is related to the cost indices for capital project in the upstream and downstream sector of energy industry. This index is prepared by IHS Markit (IHS, 2019) and shown in Table D.1.

D.3 Upstream Operating Cost Index

Upstream Operating Cost Index is related to the cost indices for operating & maintenance costs for project in the upstream and downstream sector of energy industry. This index is prepared by IHS Markit (IHS, 2019) and is shown in Table D.2.

Table D.1: Upstream Capital Cost Index (IHS, 2019).

Year	Quarter	UCCI
2000	Annual	100
2001	Annual	102
2002	Annual	104
2003	Annual	106
2004	Annual	109
2005	Q1	115
2005	Q3	126
2006	Q1	148
2006	Q3	167
2007	Q1	179
2007	Q3	198
2008	Q1	210
2008	Q3	230
2008	Q4	221
2009	Q1	210
2009	Q2	205
2009	Q3	202
2009	Q4	201
2010	Q1	201
2010	Q2	205
2010	Q3	207
2010	Q4	209
2011	Q1	218
2011	Q2	222
2011	Q3	220
2011	Q4	220
2012	Q1	227
2012	Q2	228
2012	Q3	230
2012	Q4	230
2013	Q1	231
2013	Q2	230
2013	Q3	229
2013	Q4	232
2014	Q1	232
2014	Q2	233
2014	Q3	233
2014	Q4	229
2015	Q1	198
2015	Q2	195
2015	Q3	184
2015	Q4	176
2016	Q1	167
2016	Q2	172
2016	Q3	170
2016	Q4	170
2017	Q1	172
2017	Q2	172
2017	Q3	175
2017	Q4	177
2018	Q1	181
2018	Q2	182
2018	Q3	183
2018	Q4	182
2019	Q1	183
2019	Q2	184
2019	Q3	183
2019	Q4	181

D.4 GDP Chain Type Index

The GDP chain type index is as the market cost index for labor and ROW statistics. This index is prepared by US Bureau of Economics (2019) and is shown in Table D.3.

Table D.2: Upstream Operating Cost Index (IHS, 2019).

Year	Quarter	UOCI
2000	Q1	100
2000	Q2	100
2000	Q3	100
2000	Q4	100
2001	Q1	100
2001	Q2	100
2001	Q3	101
2001	Q4	101
2002	Q1	101
2002	Q2	102
2002	Q3	103
2002	Q4	103
2003	Q1	105
2003	Q2	107
2003	Q3	108
2003	Q4	110
2004	Q1	113
2004	Q2	116
2004	Q3	119
2004	Q4	123
2005	Q1	125
2005	Q2	127
2005	Q3	130
2005	Q4	132
2006	Q1	137
2006	Q2	141
2006	Q3	146
2006	Q4	151
2007	Q1	155
2007	Q2	159
2007	Q3	164
2007	Q4	168
2008	Q1	182
2008	Q2	182
2008	Q3	183
2008	Q4	172
2009	Q1	167
2009	Q2	169
2009	Q3	168
2009	Q4	171

Year	Quarter	UOCI
2010	Q1	172
2010	Q2	172
2010	Q3	174
2010	Q4	174
2011	Q1	178
2011	Q2	179
2011	Q3	185
2011	Q4	184
2012	Q1	189
2012	Q2	188
2012	Q3	190
2012	Q4	192
2013	Q1	194
2013	Q2	196
2013	Q3	196
2013	Q4	198
2014	Q1	199
2014	Q2	202
2014	Q3	201
2014	Q4	199
2015	Q1	191
2015	Q2	190
2015	Q3	180
2015	Q4	174
2016	Q1	165
2016	Q2	167
2016	Q3	166
2016	Q4	167
2017	Q1	168
2017	Q2	169
2017	Q3	172
2017	Q4	173
2018	Q1	176
2018	Q2	176
2018	Q3	175
2018	Q4	172
2019	Q1	173
2019	Q2	174
2019	Q3	173
2019	Q4	173

Table D.3: GDP Chain Type Index (US Bureau of Economics, 2019).

Year	Quarter	GDPCTI
2000	Q1	77.39
2000	Q2	77.84
2000	Q3	78.32
2000	Q4	78.73
2001	Q1	79.23
2001	Q2	79.76
2001	Q3	80.01
2001	Q4	80.28
2002	Q1	80.50
2002	Q2	80.83
2002	Q3	81.18
2002	Q4	81.64
2003	Q1	82.05
2003	Q2	82.29
2003	Q3	82.74
2003	Q4	83.20
2004	Q1	83.82
2004	Q2	84.52
2004	Q3	85.06
2004	Q4	85.71
2005	Q1	86.37
2005	Q2	86.98
2005	Q3	87.79
2005	Q4	88.49
2006	Q1	89.10
2006	Q2	89.85
2006	Q3	90.51
2006	Q4	90.85
2007	Q1	91.78
2007	Q2	92.34
2007	Q3	92.73
2007	Q4	93.15
2008	Q1	93.57
2008	Q2	93.94
2008	Q3	94.65
2008	Q4	94.90
2009	Q1	94.96
2009	Q2	94.86
2009	Q3	94.91
2009	Q4	95.27
2010	Q1	95.49
2010	Q2	95.91
2010	Q3	96.25
2010	Q4	96.78
2011	Q1	97.28
2011	Q2	97.98
2011	Q3	98.52
2011	Q4	98.68
2012	Q1	99.28
2012	Q2	99.69
2012	Q3	100.30
2012	Q4	100.73
2013	Q1	101.12
2013	Q2	101.43
2013	Q3	101.99
2013	Q4	102.55
2014	Q1	102.96
2014	Q2	103.54
2014	Q3	104.01
2014	Q4	104.08
2015	Q1	104.07
2015	Q2	104.68
2015	Q3	105.00
2015	Q4	105.00
2016	Q1	104.93
2016	Q2	105.62
2016	Q3	105.99
2016	Q4	106.54
2017	Q1	107.04
2017	Q2	107.39
2017	Q3	108.03
2017	Q4	108.72
2018	Q1	109.34
2018	Q2	110.21
2018	Q3	110.77
2018	Q4	111.21
2019	Q1	111.50
2019	Q2	112.17
2019	Q3	112.68
2019	Q4	113.04

D.5 Producer Price Index

The producer Price Index (PPI) is the market cost index for ROW charges. This index is prepared by US Bureau of Labor Statistics (2019) and is shown in Table D.4.

Table D.4: Producer Price Index (US Bureau of Labor Statistics, 2019).

Year	Quarter	PPI
2000	Q1	129.60
2000	Q2	132.0
2000	Q3	133.80
2000	Q4	135.50
2001	Q1	137.80
2001	Q2	136.20
2001	Q3	133.40
2001	Q4	129.40
2002	Q1	128.90
2002	Q2	130.80
2002	Q3	131.70
2002	Q4	133.10
2003	Q1	138.00
2003	Q2	137.20
2003	Q3	138.10
2003	Q4	139.20
2004	Q1	142.20
2004	Q2	146.30
2004	Q3	147.70
2004	Q4	150.50
2005	Q1	152.10
2005	Q2	154.50
2005	Q3	158.70
2005	Q4	164.30
2006	Q1	162.80
2006	Q2	165.40
2006	Q3	166.70
2006	Q4	164.10
2007	Q1	166.70
2007	Q2	172.80
2007	Q3	173.70
2007	Q4	177.40
2008	Q1	183.90
2008	Q2	196.00
2008	Q3	200.50
2008	Q4	178.00
2009	Q1	169.50
2009	Q2	171.30
2009	Q3	173.90
2009	Q4	176.90

Year	Quarter	PPI
2010	Q1	182.10
2010	Q2	184.20
2010	Q3	184.60
2010	Q4	188.00
2011	Q1	195.90
2011	Q2	203.70
2011	Q3	203.80
2011	Q4	200.80
2012	Q1	202.20
2012	Q2	201.80
2012	Q3	202.40
2012	Q4	202.30
2013	Q1	203.60
2013	Q2	204.00
2013	Q3	204.20
2013	Q4	201.90
2014	Q1	205.50
2014	Q2	208.20
2014	Q3	207.10
2014	Q4	200.40
2015	Q1	191.50
2015	Q2	193.00
2015	Q3	191.60
2015	Q4	185.60
2016	Q1	182.00
2016	Q2	185.40
2016	Q3	187.10
2016	Q4	187.10
2017	Q1	191.30
2017	Q2	193.10
2017	Q3	194.00
2017	Q4	195.70
2018	Q1	198.80
2018	Q2	202.60
2018	Q3	203.80
2018	Q4	202.60
2019	Q1	199.70
2019	Q2	201.40
2019	Q3	199.40
2019	Q4	198.90

APPENDIX E

Example Result Summary

E.1. Introduction

This Appendix displays the data related to the cases run in the result section including Section 4.4, 4.5 and 4.6. The intention of placement of these data tables is not to diminish their value, but rather just to save space and time for the reader.

E.2. Static Decision Analysis

This section is linked to Section 4.4. It describes the in greater details the parameters related to the nodes and pipelines in the analysis. Table E.1 and E.2 describe the scenario for CCUS infrastructure development for capturing and storage of 60 $MtCO_2$/yr for 30 years (Scenario 3).

Table E.1: Capture and storage amounts and related costs (total in annualized form) for scenario 3.

Node ID	Capture Amount ($MtCO_2$/yr)	Storage Amount ($MtCO_2$/yr)	Number of Wells	Capture Cost (M$/yr)	Storage Cost (M$/yr)
0	1.82	0.00	0.00	80.08	0.00
1	5.65	0.00	0.00	248.60	0.00
2	3.90	0.00	0.00	171.60	0.00
3	0.00	0.00	0.00	0.00	0.00
4	8.10	0.00	0.00	356.40	0.00
5	0.00	0.00	0.00	0.00	0.00
6	8.76	4.08	24.00	232.14	3.31
7	5.09	4.06	7.00	223.96	4.20
8	4.08	4.08	17.00	179.52	2.59

9	4.21	0.00	0.00	185.24	0.00
10	2.90	0.00	0.00	127.60	0.00
11	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00
13	8.10	0.00	0.00	356.40	0.00
14	0.00	0.00	0.00	0.00	0.00
15	1.93	0.00	0.00	84.92	0.00
16	0.00	0.00	0.00	0.00	0.00
17	1.09	0.00	0.00	34.88	0.00
18	1.92	0.00	0.00	17.28	0.00
19	2.50	0.00	0.00	22.50	0.00
20	0.00	0.00	0.00	0.00	0.00
21	0.00	0.00	0.00	0.00	0.00
22	0.00	0.00	0.00	0.00	0.00
23	0.00	1.92	16.00	0.00	4.85
24	0.00	9.86	29.00	0.00	22.57
25	0.00	3.90	6.00	0.00	9.28
26	0.00	0.00	0.00	0.00	0.00
27	0.00	4.05	15.00	0.00	9.51
28	0.00	0.00	0.00	0.00	0.00
29	0.00	0.00	0.00	0.00	0.00
30	0.00	4.18	19.00	0.00	2.77
31	0.00	1.82	7.00	0.00	2.26
32	0.00	4.95	11.00	0.00	1.86
33	0.00	5.40	12.00	0.00	3.00
34	0.00	3.15	9.00	0.00	1.71
35	0.00	8.10	45.00	0.00	23.72
36	0.00	0.00	0.00	0.00	0.00
37	0.00	0.00	0.00	0.00	0.00
38	0.00	0.50	2.00	0.00	1.35
39	0.00	0.00	0.00	0.00	0.00

Table E.2: Transportation amounts and related costs (total in annualized form) for 60MtCO_2/yr for scenario 3.

Start Node	End Node	Length (miles)	Quantity Transported (MtCO_2/yr)	Pipeline Diameter (in)	Cost (M$/yr)
0	31	308.77	1.82	12.00	6.79
1	9	71.83	5.65	16.00	2.09
2	25	83.25	3.90	16.00	2.38
4	32	32.97	4.95	16.00	1.05
4	34	38.09	3.15	16.00	1.20
6	37	58.94	0.50	12.00	1.35
6	30	40.42	4.18	16.00	1.26
7	27	77.93	1.03	12.00	1.73
9	22	64.23	9.86	24.00	2.75
10	19	61.33	2.90	12.00	1.40
13	20	80.07	8.10	24.00	3.33
15	27	97.81	3.02	12.00	2.12
15	17	130.36	1.09	12.00	3.32
18	23	67.75	1.92	12.00	1.53
19	33	57.56	5.40	16.00	1.72
20	35	98.44	8.10	24.00	3.98
22	24	44.02	9.86	24.00	1.98
37	38	23.33	0.50	12.00	0.59

E.3. Static Decision Analysis with Existent Infrastructure

This section is linked to Section 4.5. It describes the in greater details the parameters related to the nodes and pipelines in the analysis. Table E.3 and E.4 describe the scenario for CCUS infrastructure development for capturing and storage of 60 $MtCO_2$/yr for 30 years (Scenario 7).

Table E.3: Capture and Storage amounts and related costs (total in annualized form) for scenario 7 with existent infrastructure.

Node ID	Capture Amount ($MtCO_2$/yr)	Storage Amount ($MtCO_2$/yr)	Number of Wells	Capture Cost (M$/yr)	Storage Cost (M$/yr)
0	0.00	0.00	0.00	0.00	0.00
1	8.52	0.00	0.00	0.00	0.00
2	3.15	0.00	0.00	139.04	0.00
3	1.60	0.00	0.00	0.00	0.00
4	8.10	0.00	0.00	356.40	0.00
5	0.00	0.00	0.00	0.00	0.00
6	8.68	4.08	24.00	230.02	3.31
7	4.06	4.06	7.00	178.64	4.20
8	4.08	4.08	17.00	179.52	2.59
9	4.21	0.00	0.00	0.00	0.00
10	3.13	0.00	0.00	130.24	0.00
11	0.00	0.00	0.00	0.00	0.00
12	2.45	0.00	0.00	22.05	0.00
13	5.94	0.00	0.00	0.00	0.00
14	1.10	0.00	0.00	70.40	0.00
15	0.00	0.00	0.00	0.00	0.00
16	0.00	0.00	0.00	0.00	0.00
17	0.00	0.00	0.00	0.00	0.00
18	2.50	0.00	0.00	22.50	0.00
19	2.50	0.00	0.00	22.50	0.00
20	0.00	1.20	24.00	0.00	4.42
21	0.00	0.00	0.00	0.00	0.00
22	0.00	0.32	4.00	0.00	1.12
23	0.00	1.92	16.00	0.00	4.85
24	0.00	9.86	29.00	0.00	22.57
25	0.00	2.60	4.00	0.00	6.18
26	0.00	0.51	17.00	0.00	2.45
27	0.00	4.05	15.00	0.00	9.51
28	0.00	0.00	0.00	0.00	0.00
29	0.00	0.60	12.00	0.00	1.92
30	0.00	1.10	5.00	0.00	0.98
31	0.00	0.00	0.00	0.00	0.00
32	0.00	4.95	11.00	0.00	1.86
33	0.00	5.40	12.00	0.00	3.00
34	0.00	3.15	9.00	0.00	1.71
35	0.00	5.94	33.00	0.00	17.39
36	0.00	0.00	0.00	0.00	0.00
37	0.00	0.00	0.00	0.00	0.00
38	0.00	0.00	0.00	0.00	0.00
39	0.00	0.60	4.00	0.00	2.03
51	0.00	5.60	28.00	0.00	2.91

Table E.4: Transportation amounts and related costs (total in annualized form) for scenario 7 with existent infrastructure.

Start Node	End Node	Length (miles)	Quantity Transported ($MtCO_2$/yr)	Pipeline Diameter (in)	Cost (M$/yr)
1	9	71.835	8.52	24	3.13
2	40	15.262	3.15	16	0.55
3	6	32.007	1.6	12	0.78
4	32	34.531	4.95	16	1.12
4	34	38.091	3.15	16	1.23
6	43	20.955	6.2	24	0.17
9	22	63.73	12.73	24	2.82
10	12	30.084	3.13	12	0.74
12	46	24.642	5.58	24	0.20
13	41	13.926	5.94	16	0.51
14	30	32.956	1.1	12	0.80
18	22	49.13	9.1	24	2.25
18	23	67.231	7.32	24	2.95
18	47	70.4	3.08	24	0.56
18	41	166.89	1.2	24	2.31
19	47	159.9	2.5	24	2.26
20	35	98.169	5.94	16	2.82
20	41	75.323	7.14	24	0.60
22	24	44.017	3.61	16	1.39
22	47	52.777	0.9	16	0.42
22	29	72.195	0.6	16	0.58
23	33	98.759	5.4	16	2.84
24	47	19.714	6.25	16	0.16
25	47	45.512	2.6	12	0.36
26	42	30.574	0.51	12	0.24
27	46	132.12	4.05	16	2.08
39	45	2.4855	0.6	12	0.08
40	46	49.633	1.53	16	0.40
40	42	8.2816	4.68	16	0.07
42	47	52.732	4.17	16	0.42
43	44	30.135	6.2	24	0.24
44	45	57.817	6.2	24	0.46
45	51	56.498	5.6	24	0.45

E.4 Dynamic Decision Analysis

This section is linked to Section 4.6. It describes the in greater details the parameters related to the nodes and pipelines in the analysis. Table E.5, E.6 and E.7 describe the scenario for CCUS infrastructure development for capturing and storage of 20 $MtCO_2$/yr for 2019 to 2019, 40 $MtCO_2$/yr from 2019 to 2039 and 60 $MtCO_2$/yr from 2039 to 2049 (Scenario 10).

Table E.5: Capture amounts and related costs (total of CAPEX and OPEX) for scenario 10.

Node ID	Capture Amount 2020-2029 ($MtCO_2$/yr)	Capture Amount 2030-2039 ($MtCO_2$/yr)	Capture Amount 2040-2049 ($MtCO_2$/yr)	Cost 2020-2029 (M$)	Cost 2030-2039 (M$)	Cost 2040-2049 (M$)
0	0.00	0.00	10.60	0.00	0.00	466.4
1	0.00	0.00	0.00	0.00	0.00	0
2	0.00	6.09	8.01	0.00	267.96	352.44
3	0.00	0.00	0.00	0.00	0.00	0
4	0.00	8.10	8.10	0.00	356.40	356.4
5	0.00	0.00	0.00	0.00	0.00	0
6	9.76	9.76	9.76	258.64	258.64	258.64
7	1.56	4.06	4.06	68.64	178.64	178.64
8	1.26	2.58	2.58	55.44	113.52	113.52
9	0.00	0.00	3.05	0.00	0.00	134.2
10	0.00	2.37	2.37	0.00	104.28	104.28
11	0.00	0.00	0.00	0.00	0.00	0
12	2.45	2.45	2.45	22.05	22.05	22.05
13	0.00	0.00	0.00	0.00	0.00	0
14	0.00	0.00	0.00	0.00	0.00	0
15	0.00	0.00	1.46	0.00	0.00	64.24
16	0.00	0.00	1.50	0.00	0.00	61.5
17	0.00	0.00	1.09	0.00	0.00	34.88
18	2.50	2.50	2.50	22.50	22.50	22.5
19	2.50	2.50	2.50	22.50	22.50	22.5
Total	20.03	40.41	60.03	449.77	1346.49	2192.19

Table E.6: Storage amounts and related costs (total of CAPEX and OPEX) for scenario 10.

Node ID	Storage Amount 2020-2029 ($MtCO_2$/yr)	Storage Amount 2030-2039 ($MtCO_2$/yr)	Storage Amount 2040-2049 ($MtCO_2$/yr)	Wells 2020-2029	Wells 2030-2039	Wells 2040-2049	Cost 2020-2029 (M$)	Cost 2030-2039 (M$)	Cost 2040-2049 (M$)
6	4.08	4.08	4.08	24	24	24	44.81	29.93	29.93
7	4.06	4.06	4.06	7	7	7	71.60	30.08	30.08
8	4.08	4.08	4.08	17	17	17	36.28	21.41	21.41
20	0	0	0	0	0	0	0.00	0.00	0.00
21	0	0	0	0	0	0	0.00	0.00	0.00
22	0	0	0	0	0	0	0.00	0.00	0.00
23	0	0	1.92	0	0	16	0.00	0.00	43.82
24	0	9.86	9.86	0	29	29	0.00	201.70	201.70
25	0	0.65	0.65	0	1	1	0.00	14.47	14.47
26	0	0	0	0	0	0	0.00	0.00	0.00
27	0	0	4.05	0	0	15	0.00	0.00	85.55
28	0	0	0	0	0	0	0.00	0.00	0.00
29	0	0	3.05	0	0	61	0.00	0.00	91.75
30	2.86	4.18	4.18	13	19	19	26.94	29.64	29.64
31	0	0	0	0	0	0	0.00	0.00	0.00
32	0	4.95	4.95	0	11	11	0.00	25.36	18.15
33	4.95	5.4	5.4	11	12	12	40.11	35.90	35.90
34	0	3.15	3.15	0	9	9	0.00	26.04	10.54
35	0	0	8.1	0	0	45	0.00	0.00	234.91
36	0	0	0	0	0	0	0.00	0.00	0.00
37	0	0	0	0	0	0	0.00	0.00	0.00
38	0	0	2.5	0	0	10	0.00	0.00	65.00
39	0	0	0	0	0	0	0.00	0.00	0.00
Total	20.03	40.41	60.03	72	129	276	219.7365	414.5222	912.8416

Table E.7: Transport amounts and related costs (total of CAPEX and OPEX) for scenario 10.

Start Node	End Node	Length (miles)	Transport Amount 2020-2029 ($MtCO_2$/yr)	Transport Amount 2030-2039 ($MtCO_2$/yr)	Transport Amount 2040-2049 ($MtCO_2$/yr)	Pipeline Diameter (in)	Cost 2020-2029 (M$)	Cost 2030-2039 (M$)	Cost 2040-2049 (M$)
0	20	94.703	0	0	10.6	24	0.00	0.00	67.16
2	26	17.158	0	8.01	8.01	24	0.00	15.70	1.17
2	12	69.386	0	1.92	0	12	0.00	24.48	0.00
3	6	32.007	0	0	1.5	12	0.00	0.00	12.51
3	8	22.355	0	0	1.5	12	0.00	0.00	9.18
4	32	32.973	0	4.95	4.95	16	0.00	17.77	2.24
4	34	38.091	0	3.15	3.15	16	0.00	20.11	2.59
6	14	10.271	5.68	1.5	0	16	6.58	0.70	0.00
6	30	40.419	0	4.18	4.18	16	0.00	21.16	2.75
7	26	97.421	2.5	0	0	12	32.91	0.00	0.00
8	14	31.485	2.82	1.5	0	12	12.34	2.14	0.00
9	29	48.975	0	0	3.05	12	0.00	0.00	18.09
10	19	61.329	0	2.37	2.37	12	0.00	21.99	4.17
12	19	51.412	2.45	0.53	2.45	12	18.87	3.50	3.50
14	30	32.956	2.86	0	0	12	12.83	0.00	0.00
15	27	97.81	0	0	4.05	16	0.00	0.00	45.26
15	17	130.36	0	0	2.59	12	0.00	0.00	47.99
16	17	46.433	0	0	1.5	12	0.00	0.00	17.27
18	25	51.627	2.5	0.65	2.5	12	18.93	3.51	3.51
18	24	72.894	0	1.85	0	12	0.00	25.56	0.00
19	33	57.559	4.95	5.4	7.32	24	43.90	3.91	3.91
20	35	98.438	0	0	10.6	24	0.00	0.00	69.42
23	33	97.652	0	0	1.92	12	0.00	0.00	32.98
24	25	60.985	0	0	1.85	12	0.00	0.00	21.88
24	26	71.33	0	8.01	8.01	24	0.00	52.72	4.85
25	26	92.93	2.5	0	0	12	31.58	0.00	0.00
35	38	64.921	0	0	2.5	12	0.00	0.00	23.11

Table E.8, E.9 and E.10 describe the scenario for CCUS infrastructure development for capturing and storage of 20$MtCO_2$/yr for 2019 to 2029, 40 $MtCO_2$/yr from 2029 to 2039 and 60 $MtCO_2$/yr from 2039 to 2049 with existent infrastructure (Scenario 12).

Table E.8: Capture amounts and related costs (total of CAPEX and OPEX) for scenario 12.

Node ID	Capture Amount 2020-2029 ($MtCO_2$/yr)	Capture Amount 2030-2039 ($MtCO_2$/yr)	Capture Amount 2040-2049 ($MtCO_2$/yr)	Cost 2020-2029 (M$)	Cost 2030-2039 (M$)	Cost 2040-2049 (M$)
0	0.00	1.82	1.82	0.00	80.08	80.08
1	0.00	0.00	2.91	0.00	0.00	128.04
2	0.00	0.00	8.54	0.00	0.00	375.76
3	0.00	0.00	4.18	0.00	0.00	183.92
4	3.15	8.10	8.10	138.60	356.40	356.40
5	0.00	0.00	0.00	0.00	0.00	0.00
6	9.76	9.76	9.76	258.64	258.64	258.64
7	0.00	2.22	4.49	0.00	97.68	197.56
8	0.52	4.08	4.08	22.88	179.52	179.52
9	0.00	0.00	0.00	0.00	0.00	0.00
10	0.00	3.67	3.67	0.00	161.48	161.48
11	0.00	0.00	0.00	0.00	0.00	0.00
12	2.22	2.45	2.45	19.98	22.05	22.05
13	0.00	0.00	0.00	0.00	0.00	0.00
14	0.00	0.52	0.52	0.00	22.88	22.88
15	0.00	0.00	1.93	0.00	0.00	84.92
16	0.00	1.50	1.50	0.00	61.50	61.50
17	0.00	1.09	1.09	0.00	34.88	34.88
18	2.19	2.50	2.50	19.71	22.50	22.50
19	2.50	2.50	2.50	22.50	22.50	22.50
Total	20.34	40.21	60.04	482.31	1320.11	2192.63

Table E.9: Storage amounts and related costs (total of CAPEX and OPEX) for scenario 12.

Node ID	Storage Amount 2020-2029 ($MtCO_2$/yr)	Storage Amount 2030-2039 ($MtCO_2$/yr)	Storage Amount 2040-2049 ($MtCO_2$/yr)	Wells 2020-2029	Wells 2030-2039	Wells 2040-2049	Cost 2020-2029 (M$)	Cost 2030-2039 (M$)	Cost 2040-2049 (M$)
6	0	4.08	4.08	0	24	24	0.00	44.81	29.93
7	0	4.06	4.06	0	7	7	0.00	71.60	30.08
8	4.08	4.08	4.08	17	17	17	36.28	21.41	21.41
20	1.2	1.2	1.2	24	24	24	43.53	43.53	43.53
21	0	0	0	0	0	0	0.00	0.00	0.00
22	0.32	0.32	0.32	4	4	4	11.26	11.26	11.26
23	0	0	1.68	0	0	14	0.00	0.00	38.34
24	1.02	2.38	9.86	3	7	29	20.87	48.69	201.70
25	0.65	0.65	3.9	1	1	6	14.47	14.47	86.84
26	0.51	0.51	0.51	17	17	17	25.03	25.03	25.03
27	0.81	0.81	4.05	3	3	15	17.11	17.11	85.55
28	0	0	0	0	0	0	0.00	0.00	0.00
29	0.6	0.6	0.6	12	12	12	18.05	18.05	18.05
30	0	0	4.18	0	0	19	0.00	0.00	36.30
31	0	1.82	1.82	0	7	7	0.00	39.11	6.62
32	0	4.95	4.95	0	11	11	0.00	25.36	18.15
33	1.8	5.4	5.4	4	12	12	19.17	35.90	35.90
34	3.15	3.15	3.15	9	9	9	26.04	10.54	10.54
35	0	0	0	0	0	0	0.00	0.00	0.00
36	0	0	0	0	0	0	0.00	0.00	0.00
37	0	0	0	0	0	0	0.00	0.00	0.00
38	0	0	0	0	0	0	0.00	0.00	0.00
39	0.6	0.6	0.6	4	4	4	21.38	21.38	21.38
51	5.6	5.6	5.6	28	28	28	24.84	24.84	24.84
Total	20.34	40.21	60.04	126	187	259	278.0373	473.0798	745.4431

Table E.10: Transport amounts and related costs (total of CAPEX and OPEX) for scenario 12.

Start Node	End Node	Length (miles)	Transport Amount 2020-2029 ($MtCO_2$/yr)	Transport Amount 2030-2039 ($MtCO_2$/yr)	Transport Amount 2040-2049 ($MtCO_2$/yr)	Pipeline Diameter (in)	Cost 2020-2029 (M$)	Cost 2030-2039 (M$)	Cost 2040-2049 (M$)
0	31	308.77	0	1.82	1.82	12	0	103.2401	30.13908
1	9	71.835	0	0.75	4.41	16	0	35.62187	4.883712
1	16	88.744	0	0.75	1.5	12	0	30.33971	6.033304
2	26	17.158	0	0	3.91	16	0	0	10.45752
2	40	15.262	0	0	4.63	16	0	0	9.465975
3	6	32.007	0	0	0.52	12	0	0	12.51449
3	8	22.355	0	0	3.66	16	0	0	13.10184
4	32	34.531	0	4.95	4.95	16	0	18.99576	2.347602
4	34	38.091	3.15	3.15	3.15	16	20.66029	2.589611	2.589611
6	14	10.271	0	0.52	0	12	0	4.70983	0
6	8	36.226	3.56	0	0	16	19.79122	0	0
6	43	20.955	6.2	6.2	6.2	16	1.424639	1.424639	1.424639
7	15	21.446	0	1.84	0	12	0	8.859538	0
7	27	78.227	0	0	0.43	12	0	0	27.17925
8	30	54.292	0	0	3.66	16	0	0	27.99927
9	22	63.73	0	0.75	4.41	16	0	32.13499	4.332701
10	19	61.801	0	3.9	3.9	16	0	31.29659	4.201534
10	12	30.084	0	0.23	0.23	12	0	11.86221	2.045253
12	46	24.642	2.22	2.22	2.22	24	1.675316	1.675316	1.675316
14	30	32.956	0	0	0.52	12	0	0	12.83473
15	16	118.54	0	0.75	0	12	0	48.58317	0
15	17	131.02	0	1.09	1.09	12	0	52.16274	17.27417
15	27	97.823	0	0	3.02	12	0	0	33.0324
18	22	49.13	0	0	1.75	12	0	0	18.13708
18	23	67.231	0.3	0	1.68	12	23.82015	0	4.570732
18	25	51.627	0	0	0.77	12	0	0	18.93464
18	47	70.4	0.69	1.3	0.6	24	4.78617	4.78617	4.78617
18	41	166.89	1.2	1.2	1.2	24	19.71254	19.71254	19.71254
19	33	58.722	1.5	5.4	5.4	16	29.95139	3.992209	3.992209
19	47	159.9	1	1	1	24	19.23734	19.23734	19.23734
20	41	75.323	1.2	1.2	1.2	24	5.120827	5.120827	5.120827
22	24	44.017	0	0.73	2.64	12	0	16.48835	2.992479
22	47	52.777	0.92	0.9	0.9	16	3.588085	3.588085	3.588085
22	29	72.195	0.6	0.6	0.6	16	4.90817	4.90817	4.90817
23	33	98.759	0.3	0	0	12	33.30828	0	0
24	26	71.33	0	0	3.9	16	0	0	35.40614
24	47	19.714	1.02	1.65	3.32	16	1.340278	1.340278	1.340278
25	47	45.512	0.65	0.65	3.13	12	3.094125	3.094125	3.094125
26	42	30.574	0.51	0.51	0.5	12	2.078578	2.078578	2.078578
27	46	132.12	0.81	0.81	0.6	16	17.34848	17.34848	17.34848
39	45	2.4855	0.6	0.6	0.6	12	1.409042	0.168976	0.168976
40	46	49.633	1.41	1.41	1.41	16	3.374288	3.374288	3.374288
40	42	8.2816	1.41	1.41	1.41	16	0.563029	0.563029	0.563029
42	47	52.732	0.9	0.9	0.9	16	3.585001	3.585001	3.585001
43	44	30.135	6.2	6.2	6.2	16	2.048711	2.048711	2.048711
44	45	57.817	6.2	6.2	6.2	16	3.930727	3.930727	3.930727
45	51	56.498	5.6	5.6	5.6	16	3.841001	3.841001	3.841001

E.5 Static Analysis with Oil Sales and Tax Incentives

This section is linked to Section 5.4. Table E.11, E.12 and E.13 describe the scenario for CCUS infrastructure development for capturing and storage of 40 $MtCO_2$/yr for 30 years with tax incentives and oil sales (Alternative 2).

Table E.11: Capture and Storage amounts and related costs (total in annualized form) for alternative 2.

Node ID	Capture Amount ($MtCO_2$/yr)	Storage Amount ($MtCO_2$/yr)	Number of Wells	Capture Cost (M$/yr)	Storage Cost (M$/yr)
0	0.00	0.00	0.00	0.00	0.00
1	0.00	0.00	0.00	0.00	0.00
2	8.54	0.00	0.00	375.76	0.00
3	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	0.00	0.00
6	9.76	0.00	0.00	258.64	0.00
7	0.00	0.00	0.00	0.00	0.00
8	0.00	0.00	0.00	0.00	0.00
9	3.60	0.00	0.00	158.40	0.00
10	0.00	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	0.00	0.00
12	2.45	0.00	0.00	22.05	0.00
13	8.37	0.00	0.00	368.28	0.00
14	0.00	0.00	0.00	0.00	0.00
15	1.46	0.00	0.00	64.24	0.00
16	1.50	0.00	0.00	61.50	0.00
17	1.09	0.00	0.00	34.88	0.00
18	1.92	0.00	0.00	17.28	0.00
19	1.28	0.00	0.00	11.52	0.00
20	0.00	0.00	0.00	0.00	0.00
21	0.00	0.00	0.00	0.00	0.00
22	0.00	0.00	0.00	0.00	0.00
23	0.00	1.92	16.00	0.00	4.85
24	0.00	9.86	29.00	0.00	22.57
25	0.00	3.90	6.00	0.00	9.28
26	0.00	0.00	0.00	0.00	0.00
27	0.00	4.05	15.00	0.00	9.51
28	0.00	1.28	16.00	0.00	3.74
29	0.00	5.00	100.00	0.00	15.98
30	0.00	0.00	0.00	0.00	0.00
31	0.00	0.00	0.00	0.00	0.00
32	0.00	0.00	0.00	0.00	0.00
33	0.00	0.00	0.00	0.00	0.00
34	0.00	0.00	0.00	0.00	0.00
35	0.00	8.10	45.00	0.00	23.72
36	0.00	3.36	28.00	0.00	11.24
37	0.00	0.00	0.00	0.00	0.00
38	0.00	2.50	10.00	0.00	6.73
39	0.00	0.00	0.00	0.00	0.00

Table E.12: Oil sales and tax incentives for alternative 2.

Node ID	Tax incentive (M$/yr)	Sale of Oil (M$/yr)	Income (M$/yr)
6	0.00	0.00	0.00
7	0.00	0.00	0.00
8	0.00	0.00	0.00
20	0.00	0.00	0.00
21	0.00	0.00	0.00
22	0.00	0.00	0.00
23	67.20	153.60	220.80
24	345.10	788.80	1133.90
25	136.50	312.00	448.50
26	0.00	0.00	0.00
27	141.75	324.00	465.75
28	44.80	102.40	147.20
29	175.00	400.00	575.00
30	0.00	0.00	0.00
31	0.00	0.00	0.00
32	0.00	0.00	0.00
33	0.00	0.00	0.00
34	0.00	0.00	0.00
35	283.50	648.00	931.50
36	117.60	268.80	386.40
37	0.00	0.00	0.00
38	87.50	200.00	287.50
39	0.00	0.00	0.00

Table E.13: Transportation amounts and related costs (total in annualized form) for 40 $MtCO_2$/yr for alternative 2.

Start Node	End Node	Length (miles)	Quantity Transported ($MtCO_2$/yr)	Pipeline Diameter (in)	Cost (M$/yr)
2	26	17.16	7.09	24	0.88
2	25	83.25	3.9	16	2.38
2	12	69.39	2.45	12	1.56
6	37	58.94	6.4	24	2.55
6	30	40.42	3.36	16	1.26
9	29	48.98	0.83	12	1.14
9	22	64.23	2.77	12	1.46
13	29	75.43	4.17	16	2.18
13	20	80.07	4.2	16	2.30
15	27	97.81	4.05	16	2.75
15	17	130.36	1.09	12	3.32
15	16	118.34	1.5	12	3.09
18	23	67.75	1.92	12	1.53
19	33	57.56	1.28	12	1.32
20	35	98.44	4.2	16	2.76
22	24	44.02	2.77	12	1.04
24	26	71.33	7.09	24	3.01
28	33	96.21	1.28	12	2.09
30	36	92.77	3.36	16	2.62
35	38	64.92	3.9	16	1.91
37	38	23.33	6.4	24	1.14

Table E.14, E.15 and E.16 describe the scenario for CCUS infrastructure development for capturing and storage of 40MtCO_2/yr for 30 years with tax incentives, oil sales and pre-existing infrastructure (Alternative 4).

Table E.14: Capture and storage amounts and related costs (total in annualized form) for alternative 4.

Node ID	**Capture Amount** (MtCO_2/yr)	**Storage Amount** (MtCO_2/yr)	**Number of Wells**	**Capture Cost** (M$/yr)	**Storage Cost** (M$/yr)
0	0.00	0.00	0.00	0.00	0.00
1	8.52	0.00	0.00	374.88	0.00
2	2.77	0.00	0.00	121.88	0.00
3	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	0.00	0.00
6	6.20	0.00	14.00	164.30	0.00
7	0.00	0.00	0.00	0.00	0.00
8	0.00	0.00	0.00	0.00	0.00
9	4.21	0.00	0.00	185.24	0.00
10	4.06	0.00	0.00	178.64	0.00
11	0.00	0.00	0.00	0.00	0.00
12	2.45	0.00	0.00	22.05	0.00
13	6.25	0.00	0.00	275.00	0.00
14	0.00	0.00	0.00	0.00	0.00
15	0.00	0.00	0.00	0.00	0.00
16	0.00	0.00	0.00	0.00	0.00
17	0.00	0.00	0.00	0.00	0.00
18	2.50	0.00	0.00	22.50	0.00
19	2.50	0.00	0.00	22.50	0.00
20	0.00	1.20	24.00	0.00	4.42
21	0.00	0.00	0.00	0.00	0.00
22	0.00	0.32	4.00	0.00	1.12
23	0.00	1.92	16.00	0.00	4.85
24	0.00	9.86	29.00	0.00	22.57
25	0.00	2.60	4.00	0.00	6.18
26	0.00	0.51	17.00	0.00	2.45
27	0.00	4.05	15.00	0.00	9.51
28	0.00	0.00	0.00	0.00	0.00
29	0.00	2.20	44.00	0.00	7.03
30	0.00	0.00	0.00	0.00	0.00
31	0.00	0.00	0.00	0.00	0.00
32	0.00	0.00	0.00	0.00	0.00
33	0.00	0.00	0.00	0.00	0.00
34	0.00	0.00	0.00	0.00	0.00
35	0.00	8.10	45.00	0.00	23.72
36	0.00	0.00	0.00	0.00	0.00
37	0.00	0.00	0.00	0.00	0.00
38	0.00	2.50	10.00	0.00	6.73
39	0.00	0.60	4.00	0.00	2.03
51	0.00	5.60	28.00	0.00	2.91

Table E.15: Oil sales and tax incentives for alternative 4.

Node ID	Tax incentive (M$/yr)	Sale of Oil (M$/yr)	Income (M$/yr)
6	0.00	0.00	0.00
7	0.00	0.00	0.00
8	0.00	0.00	0.00
20	42.00	96.00	138.00
21	0.00	0.00	0.00
22	11.20	25.60	36.80
23	67.20	153.60	220.80
24	345.10	788.80	1133.90
25	91.00	208.00	299.00
26	17.85	40.80	58.65
27	141.75	324.00	465.75
28	0.00	0.00	0.00
29	77.00	176.00	253.00
30	0.00	0.00	0.00
31	0.00	0.00	0.00
32	0.00	0.00	0.00
33	0.00	0.00	0.00
34	0.00	0.00	0.00
35	283.50	648.00	931.50
36	0.00	0.00	0.00
37	0.00	0.00	0.00
38	87.50	200.00	287.50
39	21.00	48.00	69.00
51	196.00	448.00	644.00

Table E.16: Transportation amounts and related costs (total in annualized form) for 40 $MtCO_2$/yr for alternative 4.

Start Node	End Node	Length (miles)	Quantity Transported ($MtCO_2$/yr)	Pipeline Diameter (in)	Cost (M$/yr)
1	9	71.83	8.52	24	3.13
2	26	17.16	3.62	16	0.61
2	40	15.26	0.85	12	0.41
6	43	20.96	6.2	24	0.17
9	22	63.73	12.73	24	2.82
10	12	30.08	4.06	16	1.00
12	46	24.64	6.51	24	0.20
13	41	13.93	6.25	16	0.51
18	22	49.13	11.11	24	2.25
18	23	67.23	1.92	12	1.52
18	47	70.40	6.14	24	0.56
18	41	166.89	5.55	24	2.35
19	47	159.90	2.5	24	2.26
20	35	98.17	10.6	24	4.10
20	41	75.32	11.8	24	0.60
22	47	52.78	0.9	16	0.42
22	29	72.19	2.2	16	0.58
24	26	71.33	3.61	16	2.13
24	47	19.71	6.25	16	0.16
25	47	45.51	2.6	12	0.36
26	42	30.57	0.5	12	0.24
27	46	132.12	4.05	16	2.08
35	38	64.92	2.5	12	1.47
39	45	2.49	0.6	12	0.08
40	46	49.63	2.46	16	0.40
40	42	8.28	1.61	16	0.07
42	47	52.73	1.11	16	0.42
43	44	30.13	6.2	24	0.24
44	45	57.82	6.2	24	0.46
45	51	56.50	5.6	24	0.45

REFERENCES

- Intergovernmental Panel on Climate Change. (2018). Global Warming of 1.5ºC. An IPCC Special report on the impacts of global warming of 1.5 ºC above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty.
- United Nations. (2015). Report on the Conference of the Parties on its twenty-first session, held in Paris from 30 November to 13 December 2015.
- United States Environmental Protection Agency (2020a). Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990–2018.
- International Energy Agency. (2011). Technology Roadmap- Carbon Capture and Storage in Industrial Applications, IEA, Paris.
- International Energy Agency. (2019a). World Energy Balances: Overview. IEA, Paris.
- United States Department of Energy. (2017). Energy CO_2 Emissions Impacts of Clean Energy Technology Innovation and Policy.
- Intergovernmental Panel on Climate Change (2005). Carbon Dioxide Capture and Storage.
- International Energy Agency. (2019b). Tracking Report – Energy.
- Ringrose, P. S. (2018). The CCS hub in Norway: some insights from 22 years of saline aquifer storage. Energy Procedia 146: 166-172.

- Ringrose, P. S., Mathieson, A. S., Wright, I. W., Hansen, O., Bissell, R., Saoula, N., Midgley, J. (2013). The Salah CO_2 Storage Project: Lessons Learned and Knowledge Transfer. Energy Procedia 37: 6226-6236.
- International Risk Governance Council. (2008). Regulation of Carbon Capture and Storage.
- G.C. Institute. (2016) The Global status of CCS. 2016 Summary Report: 9-10.
- Global CCS Institute (2012). CO_2 capture technologies: pre-combustion capture.
- Wall, T. F. (2007). Combustion processes for carbon capture. Process Combustion Institute 31: 31-47.
- Theo, W. L., Lim, J. S., Hashim, H., Mustaffa, A. A., Ho, W. S. (2016). Review of pre-combustion capture and ionic liquid in carbon capture and storage. Applied Energy 183: 1633-1663.
- Zhao, S., Feron, P. H. M., Deng, L., Chabanon, E., Yan, S., Hou, J., Chen, J., Qi, H. (2016). Status and progress of membrane contactors in post-combustion carbon capture: A state-of-the art review of new developments. Journal of Membrane Sciences 511: 180-206.
- D'Alessandro, D. M., Smit, B., Long, J. R. (2010). Carbon-dioxide capture: prospects for new materials. Chem Inst Ed 49: 6058-6082.
- Kohl, A L., Nielsen, R. B. (1997). Gas Purification: Fifth Edition. Gulf Professional Publishing, Houston.
- Raza, A., Rezaee, R., Gholami, R., Bing, C. H., Nagarajan, R., Hamid, M. A. (2016). A screening criterion for selection of suitable CO_2 storage sites. Journal of Natural Gas Science and Engineering 28: 317-327.

- Hosseini, S. A., Lashgari, H., Choi, J. W., Nicot, J. P., Lu, J., Hovorka, S. D. (2013). Static and dynamic reservoir modeling for geological CO_2 sequestration at Cranfield, Mississippi, U.S.A. International Journal of Greenhouse Gas Control 18:449-462.
- Chadwik, R. A., Marchant, B. P., Williams, G. A. (2014) CO_2 storage monitoring: Leakage detection and measurement in subsurface volumes from 3-D seismic data at Sleipner. Energy Procedia Vol 63: 4224-4239.
- Styring, P., Jansen, D., Coninck, H., Reith, H., Armstrong, K. (2011). Carbon Capture and Utilization in the Green Economy, Using CO_2 to manufacture Fuels, Chemicals and Materials. http://co2chem.co.uk/wp-content/uploads/2012/06/CCU%20in%20the%20green%20economy%20report.pdf.
- Fortes, M. P., Dumitriu, A. B., Tzimas, E. (2014). CO_2 utilization pathways: Techno-Economic Assessment and Market Opportunities. Energy Procedia 63: 7968-7975.
- Aresta, M., Dibenedetto, A. (2007). Utilization of CO_2 as a chemical feedstock: Opportunities and Challenges. Dalton Transcripts: 2975-2992.
- Bachu, S. (2016). Identification of oil reservoirs suitable for CO_2-EOR and CO_2 storage (CCUS) using reserves databases, with application to Alberta, Canada. International Journal of Greenhouse Gas Control 44: 152-165.
- Kuuskraa, V. A., Godec, M. L., Dipietro, P. (2013). CO_2 Utilization from "Next Generation" CO_2 Enhanced Oil Recovery Technology. Energy Procedia 37: 6854-6866.
- Sheng, J. J. (2017). Critical review of field EOR projects in shale and tight reservoirs. Journal of Petroleum Science and Engineering 159: 654-665.
- Koottungal, L., (2014). 2014 worldwide EOR survey, Special report. Oil Gas Journal 112(4): 78-97.

- Towler, B. F., Agarwal, D., Mokhatab, S. (2007). Modeling Wyoming's carbon-dioxide pipeline network. Energy Source Part A Recovery Utilization Environmental Efficiency. 30: 259-270.
- International Energy Agency Greenhouse Gas Program. (2010). Development of a Global CO_2 Pipeline Infrastructure.
- International Energy Agency Greenhouse Gas Program. (2015). CO_2 Pipeline Infrastructure.
- Peletiri, S. P., Rahmanian, N., Mujtaba, I. M. (2018). CO_2 Pipeline Design: A review. Energies 11:2184-2209.
- United States Environmental Protection Agency. (2020b). Greenhouse Gas Reporting Program (GHGRP). Accessed February, 2020. https://www.epa.gov/ghgreporting/ghg-reporting-program-data-sets.
- Knoope, M. M. J., Ramirez, A., Faaij, A. P. C. (2013). A state-of-the-art review of techno-economic models predicting the costs of CO_2 pipeline transport. Ineternational Journal of Greenhouse Gas Control 16: 241-270.
- USA. CFR., Section 49, Subsection B, Part 195, USA Regulations. (2019).
- British Petroleum. (2019). BP Energy Outlook: 2019.
- Energy and Environmental Research Centre (2019). CCS Project Fact Sheet: Red Trail Energy CCS Project.
- Menon, E. S. (2011). Pipeline Planning and Construction Field Manual. Gulf Professional Publishing, Houston.
- Huseynli, S. (2015). Determination of the most suitable oil pipeline route using GIS Least Cost Path Analysis. Doctoral Dissertation, Universitat Jaume I, Valencia, Spain.

- Potter, E., Dorshow, W. (2013). Optimizing Pipeline Routing in the 21st Century. http://www.geogathering.com/presentations/2013/GeoGathering%202013%20Pipeline%20Routing_M3Midstream.pdf.
- Middleton, R.S., Kuby, M.J., Bielicki, J. M. (2012a) Generating candidate networks for optimization: The CO_2 capture and storage optimization problem. Computers, Environment and Urban Systems 36: 18-29.
- ESRI (2012). THE ESRI Guide to GIS Analysis Volume 3: Modeling Suitability, Movement, and Interaction. Redlands, CA: Environmental Systems Research Institute.
- Towler, B. F., Agarwal, D., Mokhatab, S. (2007). Modeling Wyoming's carbon-dioxide pipeline network. Energy Source Part-A Recovery Utilization Environmental Efficiency 30: 259-270.
- Berry, J.K. , King, M.D., and Lopez, C. (2004). A web-based application for identifying and evaluating alternative pipeline routes and corridors. GITA Oil and Gas Conference, Houston, Texas.
- Callan, T. (2008). Pipeline Technology Today and Tomorrow. Oil Gas European Magazine 3
- Herzog, H., Javedan, H. (2009). Development of a Carbon Management Geographic Information System (GIS) for the United States.
- Fritze K. (2009). Modeling CO2 storage pipeline routes in the United States. Dissertation, Duke University, USA.
- Chen, W., Nindre, Y. M. L., Xu, R., Allier, D., Teng, F., Dompatail, K., Xiang, X., Guillon, L., Chen, J., Huang, L., Zeng, R. (2010). CCS scenarios optimization by spatial multi-criteria analysis: Application to multiple source sink matching in Hebei province. International Journal of Greenhouse Gas Control 4: 341-350.

- Sun, L., Chen, W. (2013). The improved ChinaCCS decision support system: A case study for Beijing-Tianjin-Hebei Region of China. Applied Energy 112: 793-799.
- Van den Broek M., Brederode, E., Ramirez, A., Kramers, L., van der Kuip, M., Wildenborg, T., Turkenburg, W., Faaij, A. (2010a). Designing a cost-effective CO_2 storage infrastructure using a GIS based linear optimization energy model. Environmental Modeling & Software 25: 1754-1768.
- Kanudia, A., Berghout, N., Boavida, D., van den Broek, M., Cabal, H., Carneiro, J., Fortes, P., Gargiulo, M., Pedro, J., Labriet, M., Lechon, Y., Martinez, R., Mesquita, P., Rimi, A., Seixas, J., Toasato, G. (2013). CCS infrastructure development scenarios for the integrated Iberian Peninsula and Morocco energy system. Energy Procedia 37: 2645-2656.
- Cevik, E., Topal T. (2003). GIS-based landslide susceptibility mapping for a problematic segment of the natural gas pipeline, Hendek (Turkey). Environmental Geology 44: 949.
- Ozcan, T., Celebi, N., Esnaf, S. (2011). Comparative Analysis of multi-criteria decision-making methodologies and implementation if a warehouse location selection problem. Expert Systems with Application 38: 9773-9779.
- Zardari, N. H., Ahmed, K., Shirazi, S. M., Yusop, Z. B. (2015). Weighting Methods and their Effects on Multi-Criteria Decision-Making Model Outcomes in Water Resources Management. Springer, Malaysia: 2194-7244.
- Al-Aomar, R.(2010). A combined AHP-Entropy Method for Deriving Subjective and Objective Criteria Weights. International Journal of Industrial Engineering 17(1): 12-24.
- Saaty, T.L. (1980) The Analytic Hierarchy Process. McGraw-Hill, New York.
- Kolios, A., Mytilinou, V., Zozano-Minguez, E., Salonitis, K. A. (2016). Compartive Study of Multi-Criteria Decision-Making methods under stochastic Inputs. Energies 9: 566.

- Liu, S., Chan, F. T. S., Ran, W. Decision making for the selection of cloud Vendor: An improved approach under group decision-making with integrated weights and objective/subjective attributes. Expert Systems with Applications 55 (2016) 37-47.
- Torfi, F., Farahani, R. Z., Mahdavi, I. (2016). Fuzzy MCDM for weight of object's phrase in location routing problem. Applied Mathematical Modelling 40: 526-541.
- Middleton, R. S., Bielicki, J. M. (2009). A scalable infrastructure model for carbon capture and storage: SimCCS. Energy Policy 37 pp: 1052-1060.
- Kobos, P. H., Malczynski, L. A., Borns, D. J. (2006). Employing the 'String of Pearls' Integrated Assessment Model: A Carbon Sequestration Systems Analysis Tool. 26th USAEE/IAEE North American Conference, Ann Arbor, USA.
- Knoope, M. M. J., Ramirez, A., Faaij, A. P. C. (2014). Improved cost models for optimizing CO_2 pipeline configuration for point-to-point pipelines and simple networks. International Journal of Greenhouse Gas Control 22: 25-46.
- Lee, S. Y., Lee, I. B., Han, J. (2019). Design under uncertainty of carbon capture, utilization and storage infrastructure considering profit, environmental impact, and risk preference. Applied Energy 238: 34-44.
- Kazmierczak, T., Brandsma, R., Neele, F., Hendricks, C. (2009). Algorithm to create a CCS low-cost pipeline network. Energy Procedia 1: 1617-1623.
- Guo, J. X. (2020). Integrated optimization model for CCS hubs and pipeline network design., Computers & Chemical Engineering 132.
- Ravi, N. K., Annaland, M. V. S., Fransoo, J. C., Grievink, J., Zondervan, E. (2017). Development and implementation of supply chain optimization framework for CO_2 capture and storage in the Netherlands. Computers and Chemical Engineering 102: 40-51.

- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. Numerische Mathematik 1: 269–271.
- ESRI (2011). ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.
- Cormen, Thomas H., Leiserson, Charles E., Rivest, Ronald L., Stein, Clifford. (2001). Introduction to Algorithms (Second ed.). MIT Press and McGraw–Hill.
- Souissi, O., Benatitallah, R., Duvivier, D., Artiba, A., Belanger, N., Feyzeau, P. (2013). Path planning: A 2013 survey. *Proceedings of 2013 International Conference on Industrial Engineering and Systems Management (IESM),* Rabat, Morocco:1-8.
- Middleton, R. S., Yaw, S. P., Hoover, B. A., Ellett, K. M. (2020). SimCCS: An open-source tool for optimizing CO_2 capture, transport, and storage infrastructure. Environmental Modeling and Software 124: 1-8.
- Morbee, J., Serpa, J., Tzimas, E. (2012). Optimized deployment of a European CO_2 transport network. International Journal of Greenhouse Gas Control 7: 48-61.
- Hart, P. E., Nilsson, N. J., Raphael, B. (1968). A formal Basis for the Heuristic Determination of Minimum Cost Paths. IEEE trans on Sys Sci and Cybernet 4(2): 100-107.
- Reddy, H. (2013). Path Finding – Dijkstra and A* Algorithm.
- Weihs, G. A. F., Wiley, D. E. (2012). Steady-state design of CO_2 pipeline networks for minimal cost per tonne of CO_2 avoided. International Journal of Greenhouse Gas Control 8:150-168.
- Ooi, R. E. H., Foo, D. C.Y., Ng, D. K. S., Tan, R. R. (2013). Planning of carbon capture and storage with pinch analysis techniques. Chemical Engineering Resources and Design 91: 2721-2731.

- Tian, Q., Zhao, D., Li, Z., Zhu, Q. (2017). Robust and stepwise optimization design for CO_2 pipeline transportation. International Journal of Greenhouse Gas Control 58: 10-18.
- Van den Broek M., Ramirez, A., Groenenberg, H., Neele, F., Viebahn, P., Turkenburg, W., Faaij, A. (2010b). Feasibility of storing CO_2 in the Utsira formation as part of a long-term Dutch CCS strategy: An evaluation based on a GIS/MARKAL toolbox. International Journal of Greenhouse Gas Control 4(2): 351-366.
- Van den Broek, M., Faaij, A., Turkenburg, W. (2008). Planning for an electricity sector with carbon capture and storage. International Journal of Greenhouse Gas Control 2: 105-129.
- Kuby, M. J., Bielicki, J. M., Middleton, R. S.(2009) Optimal Spatial Deployment of CO_2 given a price on Carbon. Portland GSA Annual Meeting, USA.
- Middleton, R. S., Kuby, M. J., Wei, R., Keating, G. N., Pawar, R. J. (2012b). A dynamic model for optimally phasing in CO_2 capture and storage infrastructure. Environmental Modeling & Software 37: 193-205.
- Middleton, R. S., Keating, G. N., Stauffer, P. H., Jordan, A. B., Viswanathan, H. S., Kang, Q. J., Carey, W., Mulkey, M . L., Sullivan, E. J., Chu, S. P., Esposito, R., Meckel, T. A. (2012c). The cross-scale science of CO_2 capture and storage: from pore scale to regional scale. Energy and Environmental Science 5: 7328-7345.
- Middleton, R. S., Keating, G. N., Viswanathan, H. S., Stauffer, P. H., Pawar, R. J. (2012d). Effects of geologic reservoir uncertainty on CO_2 transport and storage infrastructure. International Journal of Greenhouse Gas Control 8: 132-142.
- Pawar, R. J., Bromhal, G. S., Chu, S., Dilmore, R. M., Oldenburg, C. M., Stauffer, P. H., Zhang, Y., Guthrie, G. D. (2016). The National Risk Assessment Patnership's Integrated

assessment model for carbon storage: A tool to support decision making amidst uncertainty. International Journal of Greenhouse Gas Control 52: 175-189.

- Sun, L., Chen, W. (2017). Development and application of a multi-stage CCUS source-sink matching model. Applied Energy 185: 1424-1432.
- D'Amore, F., Lovisotto, L., Bezzo, F. (2020). Introducing social acceptance into the design of CCS supply chains: A case study at a European level. Journal of Cleaner Production: 249.
- Jensen, M. D., Pe, P., Snyder, A. C., Heebink, L. V., Botnen, L. S., Gorecki, C. D., Steadman, E. N., Harju, J. A. (2013). Methodology for phased development of a hypotheical pipeline network for CO_2 Transport during Carbon Capture, Utilization, and Storage. Energy Fuels 77: 4175-4182.
- Wang, Z., Weihs, G. A. F., Neal, P. R., Wiley, D. E. Effects of pipeline distance, injectivity and capacity on CO_2 pipeline and storage site selection. International Journal of Greenhouse Gas Control 51:95-105.
- Knoope, M. M. J., Ramirez, A., Faaij, A. P. C. (2015). The influence of uncertainty in the development of a CO_2 infrastructure network. Applied Energy 158: 332-347.
- Diamante, J. A. R., Tan, R. R., Foo, D. C. Y., Ng, D. K. S., Aviso, K. B., Bandyopadhyay, S. (2013). A Graphical Approach for Pinch-based Source-Sink Matching and Sensitivity Analysis in Carbon Capture and Storage Systems. Industrial Engineering Chemical Resource 52: 7211-7222.
- Diamante, J.A.R., Tan, R.R., Foo, D.C.Y., Ng, D.K.S., Aviso, K.B., Bandyopadhyay, S. (2014). Unified pinch approach for targeting of carbon capture and storage (CCS) systems with multiple time periods and regions. Journal of Cleaner Production 71: 67–74.

- D'Amore, F., Bezzo, F. (2017). Economic optimization of European supply chains for CO_2 capture, transport, and sequestration. International Journal of Greenhouse Gas Control 65: 99-116.
- D'Amore, F., Mocellin, P., Vianello, C., Maschio, G., Bezzo, F. (2018). Economic optimization of European supply chains for CO_2 capture, transport and sequestration, including societal risk analysis and risk mitigation measures. Applied Energy 223: 401-415.
- Tian, Q., Zhao, D., Li, Z., Zhu, Q. (2018). A two-step co-evolutionary particle swarm optimization approach for CO_2 pipeline design with multiple uncertainties. Carbon Management 9(4): 1-14.
- Tian, Q., Zhao, D., Li, Z., Zhu, Q. (2020). CO_2 pipeline transportation System Optimization Design Based on Multiple Population Genetic Algorithm. Proceedings of 11th International Conference on Modelling, Identification and Control: 643-651.
- Leonzio, G., Foscolo, P. U., Zondervan, E. (2019). An outlook towards 2030: Optimization and design of a CCUS supply chain in Germany. Computers and Chemical Engineering 125: 499-513.
- Leonzio, G., Bogle, D., Foscolo, P. U., Zondervan, E. (2020). Optimization of CCUS supply chains in the UK: A strategic role for emissions reduction. Chemical Engineering & Research 155: 211-228.
- Zhang, S., Liu, L., Zhang, L., Zhuang, Y., Du, J. (2018). An optimization model for carbon capture utilization and storage supply chain: A case study for Northeastern China. Applied Energy 231: 194-206.

- Zhang, S., Zhuang, Y., Tao, R., Liu, L., Zhang, L., Du, J. (2020). Multi-objective optimization for the deployment of carbon capture utilization and storage supply chain considering economic and environmental performance. Journal of Cleaner Production 270.
- Lee, S. Y., Lee, J. U., Lee, I. B., Han, J. (2017). Design under uncertainty of carbon capture and storage infrastructure considering cost, environmental impact and preference on risk. Applied Energy 189: 725-738.
- Lee, S. Y., Lee, I. B., Han, J. (2019). Design under uncertainty of carbon capture, utilization and storage infrastructure considering profit, environmental impact, and risk preference. Applied Energy 238: 34-44.
- Yousefi-Sahzabi, A., Saski, K., Djamaluddin, I., Yousefi, H., Sugai, Y. (2011). GIS modeling of CO_2 emission sources and storage possibilities. Energy Procedia 4: 2831-2838.
- Wan, J., Qi, G., Zeng, Z., Sun, S. (2011). The application of AHP in oil & gas pipeline route selection. (2011). Presented on 19th International Conference of Geoinformatics, Shanghai, China.
- Balaji, K., Rabiei, M. (2020). Effect of terrain, environment and infrastructure on potential CO_2 pipeline corridors: a case-study from North-Central USA. Energy, ecology and Environment. https://doi.org/10.1007/s40974-020-00194-y
- United States Geologic Survey (2019). USGS TNM Hydrography (NHD). Accessed February 23, 2020. https://hydro.nationalmap.gov/arcgis/rest/services/nhd/MapServer.
- Farr, T. G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Pallar, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D. (2007), The Shuttle Radar Topography Mission. Reviews of Geophysics 45(2).

- United States Geologic Survey; New Mexico Bureau of Mines and Mineral Resources. (2018a). Quaternary fault and fold database for the United States. Accessed February 23, 2020, at: https://www.usgs.gov/natural-hazards/earthquake-hazards/faults.
- He, B., Han, P., Lu, C., Bai, X. (2015). Effect of soil particle size on the corrosion behavior of natural gas pipeline. Engineering Failure Analysis 58: 19-30.
- Norhazilan, M. N., Nordin, Y., Lim, K. S., Siti, R. O., Safuan, A. R. A., Norhamimi, M. H. (2012). Relationship between Soil Properties and Corrosion of Carbon Steel. Journal of Applied Science and Resources 8(3): 1739-1747.
- Soil Survey Staff. (2015). Gridded Soil Survey Geographic (gSSURGO) Database for the United States of America and the Territories, Commonwealths, and Island Nations served by the USDA-NRCS. United States Department of Agriculture, Natural Resources Conservation Service. Accessed on 23 February 2020. https://gdg.sc.egov.usda.gov/.
- Geverdt, D. (2019). Education Demographic and Geographic Estimates Program (EDGE): Composite School District Boundaries File Documentation, 2018 (NCES 2017-035). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved 23 February 2019.
- Bureau of Land Management. (2020). BLM National Designated Areas of Critical Environmental Concern Polygons. Accessed on 23, February. https://gis.blm.gov/EGISDownload/LayerPackages/BLM_National_ACEC.zip.
- United States Geologic Survey. (2018). Protected Areas Database of the United States (PAD-US). Accessed on 23, February. https://www.sciencebase.gov/catalog/item/5b030c7ae4b0da30c1c1d6de.

- Macharia, P. M., Mundia, C. N., Wathuo, M. W. (2015). Experts' Responses Comparison in a GIS-AHP Oil Pipeline Route Optimization: A Statistical Approach. American Journal of GIS 4(2): 53-63.
- Yildrim, V., Yomralioglu, T. (2007). GIS Based Pipeline Route Selection by ArcGIS in Turkey. 27th Annual ESRI International User Conference, Sand Diego, USA.
- U.S. Census Bureau. (2012). 2010 TIGER/Line Shapefiles. Accessed February 23, 2020. https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-geodatabase-file.html.
- National Alas of the United States (2006). Federal Lands of the United States. Accessed on: 23 February, 2020. https://www.sciencebase.gov/catalog/item/4f4e4b32e4b07f02db6b47a9.
- National Park Service. (2006). National Parks. Accessed on: 23 February 2020. https://www.bts.gov/geospatial/national-transportation-atlas-database.
- Environmental Information Agency. (2019). Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990–2017.
- United States Geologic Survey Gap Analysis Program. (2016). GAP/LANDFIRE National Terrestrial Ecosystems 2011: U.S. Geological Survey. Accessed on 23, February. https://doi.org/10.5066/F7ZS2TM0.
- Baufune, S., Gruger, F., Grube, T., Kreig, D., Linssen, J., Weber, M., Hake, J. F., Stolten, D. (2013). GIS-based scenario calculations for a nationwide German hydrogen pipeline infrastructure. International Journal of Hydrogen Energy 38(10): 3813-3829.
- Nonis, C. N., Varghese, K., Suresh, K. S. (2007). Investigation of an AHP based Criteria weighting scheme for GIS routing of Cross-Country Pipeline Projects. 24th International Symposium on Automation & Robotic Construction, Kochi, India.

- National Energy Technology Laboratory. (2015). A review of the CO2 Pipeline Infrastructure in the U.S . DOE/NETL-2014/1681. https://www.energy.gov/sites/prod/files/2015/04/f22/QER%20Analysis%20-%.
- National Energy Technology Laboratories. (2020). NATCARB Viewer 2.0. Accessed on: 23rd February 2020. https://edx.netl.doe.gov/geocube/#natcarbviewer.
- Enhanced Oil Recovery Institute. (2020). WYRIT. Accessed on: 23rd February 2020. https://www.eoriwyoming.org/maps/wyrit.
- North Dakota Pipeline Authority. (2020). ND CO2 Pipeline Map. Accessed on: 23rd February 2020. https://ndpipelines.files.wordpress.com/2012/05/nd-co2-map.pdf.
- Boris, D. (1934). "Sur la sphère vide". Bulletin de l'Académie des Sciences de l'URSS, Classe des Sciences Mathématiques et Naturelles 6: 793–800.
- Fang, T., Piegl, L. A. (1993). Delaunay triangulation using a uniform grid. IEEE Computer Graphics and Applications 13(3): 36-47.
- Millman, K. J., Aivazis. A. (2011). Python for Scientists and Engineers. Computing in Science & Engineering, 13: 9-12.
- Det Norske Veritas Germanischer Lloyd (2010). Recommended practice. Design and operation of CO_2 pipelines. DNVRP-J202: 1–42.
- National Energy Technology Laboratory. (2018). FE/NETL CO_2 Transport Cost Model: Description and User’s Manual. DOE/NETL-2018/1877.
- McCollum, D. L., Ogden, J. M. (2006). Techno-Economic Models for Carbon-Dioxide Compression, Transport, and Storage & Correlations for Estimating Carbon Dioxide Density and Viscosity. University of California, Davis. UCD-ITS-RR-06-14.
- Global CCS Institute. (2017). Global Costs of Carbon Capture and Storage.

- Rui, Z., Metz, P., Reynolds, D., Chen, G., and Zhou, X. (2011). Regression models estimate pipeline construction costs. Oil and Gas Journal 109(27): 120-127.
- Information Handling Services Markit. (2019). Upstream Capital Cost Analysis Service.
- United States Bureau of Economics. (2019). Gross Domestic Product: Chain-type Price Index.
- United States Bureau of Labor Statistics. (2019). Producer Price Index.
- Law, D. H. S., Bachu, S. (1997). Hydrogeological and Numerical analysis of CO2 Disposal in deep aquifers in the Alberta sedimentary Basin. Energy Conversion Management 37(6): 1167-1174.
- Rubin, E. S., Berkenpas, M. B., McCoy, S. (2008). Development and Application of Optimal Design Capability for coal gasification Systems. Carnegie Mellon University.
- Azzolina, N. A., Nakles, D. V., Gorecki, C. D., Peck, W. D., Ayash, S. C., Melzer, L. S., Chatterjee, S. (2015). CO_2 storage associated with CO_2 enhanced oil recovery: A statistical analysis of historical operations. International Journal of Greenhouse Gas Control 37: 384-397.
- GAMS Development Corporation. General Algebraic Modeling System (GAMS) Release 27.1.0, Fairfax, VA, USA, 2019.
- CPLEX, IBM ILOG. (2009). V12. 1: User's Manual for CPLEX. International Business Machines Corporation 46(53) 157.
- Conrad, J. G., Branting, L. K.(2018). Introduction to the special issue on legal text analytics. Artificial Intelligence in Law 26:99-102.
- Zhang, J., El-Gohary, N. M. (2016). Semantic NLP-based Information Extraction from Construction Regulatory Documents for Automated Compliance Checking. Journal of Computing in Civil Engineering 30(2): 04015014.

- Breaux, T. D., Vail, M. W., Anton, A. I. (2006). Towards Regulatory Compliance: Extracting Rights and Obligations to Align Requirements with Regulations. 14th IEEE International Requirements Engineering Conference, Minneapolis, USA.
- Hjelseth, E., Nisbeth, N. (2011). Capturing Normative Constraints by use of the Semantic mark-up RASE Methodology. Proceedings of the CIB International Conference, Antipolis, France.
- Song, J., Kim, J., Lee, J. K. (2018). NLP and Deep Learning based Analysis of Building Regulations to support Automated Rule Checking Systems. 35th International Symposium on Automation and Robotics in Construction, Red Hook, USA.
- Zeni, N., Kiyavitskaya, N., Mich, L., Cordy, J. R., Mylopoulos, J. (2015). GaiusT: supporting the extraction of rights and obligations for regulatory compliance. Requirements Engineering 20: 1-22.
- Lim, J. J., Rebholz-Schuhmann, D. (2011). Improving the extraction of complex regulatory events from scientific text by using ontology-based inference. Journal of Biomedical Semantics 2(5): 1-13.
- El-Gohary, N. M., El-Diraby, E. (2010). Domain Ontology for Processes in Infrastructure and Construction. Journal of Computing in Civil Engineering 136(7): 730-744.
- Steven, B., Loper, E., Klein, E. (2009). Natural Language Processing with Python. O’Reilly Media Inc.
- Manning, C. D., Mihai, S., Bauer, J., Finkel, J., Bethard, S. J., McClosky, D. (2014). The Stanford CoreNLP Natural Language Processing Toolkit. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations.

- New York Times. (2016). North Dakota Oil Pipeline Battle: Who's Fighting and Why. Accessed: 28th September 2020. https://www.nytimes.com/2016/11/02/us/north-dakota-oil-pipeline-battle-whos-fighting-and-why.html.
- US Department of Transportation Pipeline and Hazardous Materials Safety Administration. (2020). US DOE Pipeline and Hazardous Material. Accessed on: 23rd February 2020.
- Pipeline and Hazardous Material Safety Administration. (2019). Ecological Unusually Sensitive Areas (Eco USA). [Accessed on 23rd February 2019]. https://www.npms.phmsa.dot.gov/USAEcoData.aspx.
- United States Department of Energy. (2019). Internal Revenue Code Tax Fact Sheet. Accessed on 23rd February 2020. https://www.energy.gov/sites/prod/files/2019/10/f67/Internal%20Revenue%20Code%20Tax%20Fact%20Sheet.pdf.
- Cooney, G., Littlefield, J., Marriot, J., Skone, T. J. (2015). Evaluating the climate benefits of CO_2-enhanced oil recovery using life cycle analysis. Environmental Science and Technology 49(12): 7491-7500.
- Azzolina, N. A., Peck, W. D., Hamling, J. A., Gorecki, C. D., Ayash, S. C., Doll, T. E., Nakles, D. V., Melzer, L. S. (2016). How green is my oil? A detailed look at greenhouse gas accounting for CO_2-enhanced oil recovery (CO_2-EOR) sites. International Journal of Greenhouse Gas Control 51: 369-379.