# MATHEMATICAL PROGRAMMING TECHNIQUES FOR DESIGNING MINIMUM COST PIPELINE NETWORKS FOR $CO_2$ SEQUESTRATION

H. Y. Benson[1] and J. M. Ogden[2]

[1] Princeton Environmental Institute and Department of Operations Research and Financial Engineering, Princeton University, Princeton, NJ, 08544, USA
[2] Princeton Environmental Institute, Princeton University, Princeton, NJ, 08544, USA

## ABSTRACT

It has been proposed that the $CO_2$ produced at sources such as fossil energy conversion facilities and industrial process plants could be captured, compressed to supercritical pressures, transported via pipelines and stored in underground geologic formations such as depleted oil and natural gas reservoirs and deep saline aquifers. In this paper, we describe the initial phases of a project using mathematical programming techniques to find the minimum cost network for transporting the $CO_2$ from the sources to the sinks. A non-linear model of the $CO_2$ pipeline system is described, with appropriate constraints, and the lowest cost system is found using a state-of-the-art nonlinear optimization software called LOQO. We also present ways to model the behavior of the system over time, and to model the impact of uncertainties. In future work, we plan to extend these methods to look at the larger system of fossil energy complexes with $CO_2$ capture and sequestration and distribution of hydrogen.

## INTRODUCTION

The greenhouse gas emissions from direct combustion of fuels can be greatly reduced with the use of hydrogen produced from fossil sources, when hydrogen production is coupled with capture and secure sequestration of the resulting $CO_2$ at storage sites such as aquifers and depleted hydrocarbon reservoirs. The transportation of $CO_2$ from fossil hydrogen plants to sequestration sites can be accomplished via a pipeline network. The design of this pipeline is an important aspect of the overall energy network.

In this paper, we describe a mathematical model for a pipeline infrastructure for $CO_2$ sequestration and utilize various mathematical programming techniques to find minimum cost strategies for building and operating this pipeline network. Following previous work in the literature, the techniques we investigate are:

- Nonlinear Programming: The capital cost of the network depends nonlinearly on the diameter and the length of each pipe and its flow rate. We use a nonlinear programming solver to find the lowest cost solution, subject to constraints on the flow rates over time from the sources and to the sinks.
- Stochastic Programming: Some of the sequestration site characteristics, such as maximum flow rate, permeability and storage capacity, may have large uncertainties. Moreover, there is some probability of leakage out of the reservoir. In order to handle these uncertainties, we investigate work with a stochastic framework, whenever probability distributions for the uncertain variables can be identified.
- Dynamic Programming: This is another technique that allows us to answer questions about how the flow over the network evolves over time. For example, we could specifically consider the optimum order in which to bring the sequestration sites online (should nearby, small sites be used first, or larger more distant sites), and backup strategies for rerouting $CO_2$ flows in the case of leakage in the network.

As mentioned above, these approaches have been widely discussed in literature. Nonlinear programming formulations of models for water distribution networks and oil and natural gas pipelines are given in [8], [3], and [4]. In fact, problems that arise in fossil fuel networks are, in principle, the reverse of designing a $CO_2$ sequestration network. For a $CO_2$ sequestration network, the sources provide the "demand" for sequestration, and the sinks provide the finite "supply" of capacity. The major difference between the models in previous literature and the one we consider is the cost function, which is customized for the case of $CO_2$ sequestration. The formulation for the cost function will be given in Section 3, and further details can be found in [9].

There are different approaches presented in [6], [3], and [5] to solve the pipeline optimization problems. [6] uses a linearization of the model, [3] uses a Sequential Unconstrained Minimization technique (SUMT), and [5] uses a global optimization approach. In this paper, we are proposing using a state-of-the-art nonlinear optimization software called LOQO [10], which uses an infeasible interior-point method approach for finding local optima for nonconvex, nonlinear optimization problems. As shown in [1], it is particularly suited for solving large-scale sparse problems, and it is quite robust at working with complicated objective functions. The network problems that arise in $CO_2$ sequestration are indeed quite sparse and the objective (cost) function is nonconvex and nonlinear. As we consider uncertainties in the system and propose a stochastic framework consisting of many possible scenarios, the problems become quite large in order to represent as many situations as possible. Therefore, we feel that LOQO is a good tool for approaching $CO_2$ sequestration pipeline design problems, and the examples we present in the upcoming sections will underline this idea.

We have also included a discussion of a dynamic programming approach for designing the optimal pipeline network for $CO_2$ sequestration. This technique has been widely used in pipeline design, and a good survey can be found in [2]. It is particularly useful for answering specific questions such as which of the sinks to bring online first. For the current state of our model, however, we found that a nonlinear programming approach was quite sufficient and include the discussion of dynamic programming solely for completeness.

In future work, we plan to extend these methods to look at the larger system of fossil energy complexes with $CO_2$ capture and sequestration and distribution of hydrogen.

## DESIGN PROBLEMS IN OPTIMIZATION

Many optimization problems are termed *design problems*. In such problems, the constraints would allow infinitely many designs. The objective function that is optimized (in our case, the cost of the $CO_2$ pipeline network) selects one design out of the design space (in our case, the lowest cost $CO_2$ pipeline network). Many of the design variables are zero in the optimal solution. These represent components that are not built. Variables that are non-zero in the optimal solution are "built" and the values of these variables give appropriate "sizing" of the pipeline segments. Our model starts with all of the possible connections between the sources and the sinks in place (e.g. Figure 1(a)), and uses mathematical programming methods to find an optimal solution that indicates which connections should actually be built and utilized to achieve the lowest overall cost of $CO_2$ disposal (e.g. Figure 1(b)).

## A NONLINEAR OPTIMIZATION MODEL FOR $CO_2$ SEQUESTRATION

The optimization model that we propose is one that considers the pipeline capital cost for a $CO_2$ sequestration network. The cost function is minimized subject to four sets of constraints:

1. conservation of mass at the sources,
2. requirement that the total flow into each sink during any one time period cannot exceed a maximum amount allowed by the physical characteristics of that sink,
3. flow balance at the intermediate nodes, and
4. requirement that the total flow into each sink throughout the lifetime of the system cannot exceed the total capacity of the sink.

This approach takes into account the lifetime of the pipeline and the sequestration site and allows for trunk lines to consolidate flow on certain routes. The model has the following form:

$$
\begin{aligned}
&\min NPV(s,L) \\
&\text{s.t.} \sum_{j\in D} s_{ijt} + \sum_{k\in R} s_{ikt} = Q_i, i \in O, t \in T \\
&\sum_{i\in O} s_{ijt} + \sum_{k\in R} s_{kjt} \le smax_j, j \in D, t \in T \\
&\sum_{i\in O} s_{ikt} - \sum_{j\in D} s_{kjt} = 0, k \in R, t \in T \\
&\sum_{t\in T}\left(\sum_{i\in O} s_{ijt} + \sum_{k\in R} s_{kjt}\right) \le scap_j, j \in D.
\end{aligned} \tag{1}
$$

The variable $s_{ijt}$ is the flow from source $i$ to sink $j$ at time $t$, $L_{ij}$ is the distance between source $i$ and sink $j$, $Q_i$ is the amount of $CO_2$ at source $i$, $smax_j$ is the maximum flow into sink $j$, $scap_j$ is the total capacity of sink $j$, O is the set of sources, $D$ is the set of sinks, and $T$ is a discretization of the lifetime of the system. The objective function, $NPV(s,L)$, uses a net present value of the cost of the pipeline over the lifetime of the system. Further details about the cost function appear below.

In the above model, $R$ is the set of intermediate nodes. Using intermediate nodes allows several smaller flows to be combined into a larger (and possibly cheaper) flow at these nodes. We can then accommodate trunk lines in our pipeline network. We will take the approach of determining beforehand the number of intermediate points to be used, and the locations of these intermediate points are variables in our model. Note that especially for a large network, determining the number of intermediate nodes is an intractable problem—in fact, in its most general form, it is the Steiner network problem which is NP-complete [7].

In the model presented above, the lifetime of the pipeline and especially the useful capacity of the reservoir have important effects on the system economics. For example, if a given reservoir has a small capacity, it will be used up in only a few years, and no more $CO_2$ can be sent there. The pipeline connected to this sink will be useless for $CO_2$ disposal. When the sink capacity is considered, we need to be able to balance the cost spent on a certain pipe with the length of time it is in use. It may be more advantageous for a network to utilize a more distant, but larger capacity, sink if it leads to an increased lifetime for the pipeline network.

In this model, the linear constraints can be easily handled by any solver. The cost function, presented below, is nonlinear and nonconvex and requires the use of a robust nonlinear solver.

### *The Single-Period Cost Function*

The cost function that we use depends nonlinearly on the flow over the pipe and its length:

$$Cost_t(s,L) = C_0 \sum_{i\in O}\sum_{j\in D}\left(\frac{s_{ijt}}{s_0}\right)^{0.48}\left(\frac{L_{ij}}{L_0}\right)^{1.24} \Big/ \sum_{i\in O} Q_i ,$$

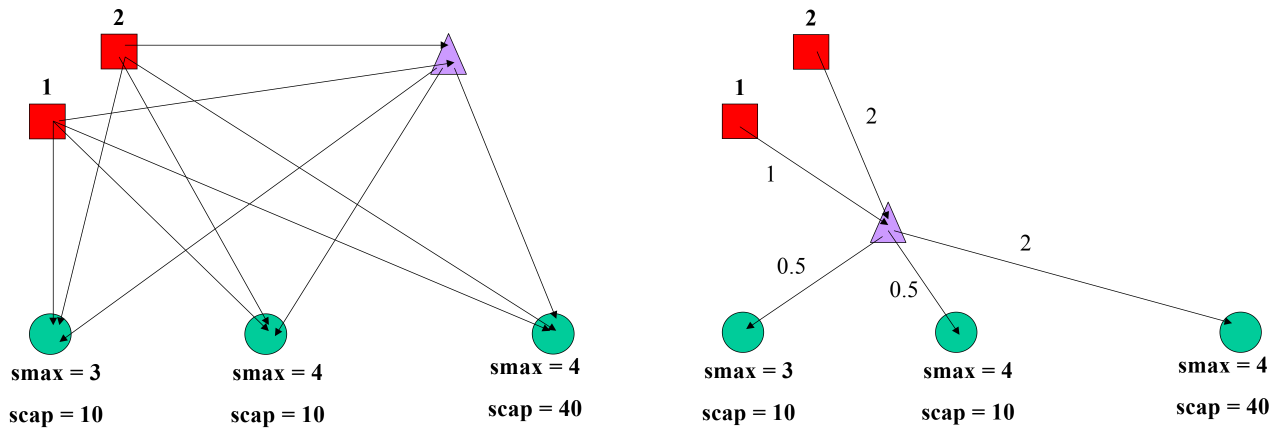
where $C_0$, $s_0$, and $L_0$ are constants. This cost function represents the cost for one time period, t. Further details on how this cost function was formulated and the particular constants involved can be found in [9].

### *Multi-period Cost Function for Sinks with Capacities*

We use a Net Present Value (NPV) approach in our cost function to compute total cost over the lifetime of the system:

$$NPV(s,L) = \sum_t \left(\frac{1}{1+d}\right)^t Cost_t(s,L),$$

where $d$ is the discount rate (taken to be 0.12 in our models). Note that by using this objective function, we are favoring using the capacity of a sink over as long a period of time as possible.




**Figures 1(a) and (b).** Setup and solution of the network model with trunk lines and sink capacities. The squares represent the sources, the circles represent the sinks, and the triangle represents the additional node where multiple flows come together for consolidation. The numbers above the squares represent the output from each source and the numbers below the sinks represent the maximum daily flow into a sink (smax) and the total capacity of each sink (scap). On the left, all of the possible links are drawn to setup the problem. On the right, the optimal solution is shown with the flows listed next to the links. For this example, we have assumed that the flow is constant over the total system lifetime of 20 years of the system.

## SOLVING THE NONLINEAR OPTIMIZATION PROBLEM

As previously mentioned, we are using a state-of-the art solver called LOQO [10] to solve the problem given by (1). LOQO is a primal-dual interior-point code, which uses Newton's Method to solve the first-order conditions of the optimization problem. It is designed to work with sparse data structures, so it can handle large-scale problems. It requires the use of first and second derivatives. Further details on the algorithm can be found in [10].

There are three important issues to note here. First, the cost function in our model is nonlinear. This requires the use of a solver that can handle nonlinearities when working with our models. LOQO is such a solver. It is based on an interior-point method that can solve nonconvex nonlinear optimization problems to a local optimum. Second, the cost function is nonconvex. This means that there may not be a unique global optimum to the problem. If a good guess can be provided as an initial solution, LOQO should be able to reach the global optimum. However, if such a guess is not available, it is also possible to use LOQO as a subroutine in a global optimization approach such as Tabu search or a genetic algorithm. The third issue to note is that the objective function is nondifferentiable at solutions where there are pipes with no flow. For an optimization algorithm that depends on the computation of derivatives, as is the case with LOQO, nondifferentiability can cause a problem. In the case of a design problem, we expect there to be some pipes with no flow, and therefore, it is quite natural to run into the nondifferentiability problem. However, this is easily overcome by using a small perturbation for the numerators of the fractions in the cost function. Such a technique has been used for other types of nondifferentiable optimization problems, as discussed in [11].

In Figures 1(a) and (b), we provide the setup and the solution of a small problem with 3 sources, 3 sinks, and 1 trunk line. This problem also asserts that the flows remain constant over the lifetime of the system. The sinks have capacities, and the system has a lifetime of 20 years in the optimal solution. When the assertion is removed, 1 unit of flow is sent to sink 1, depleting it in the first 10 years. 2 units of flow are sent to sink 2 in the first two years, followed by 1 unit of flow for the next 6 years. Therefore, this sink is depleted in 8 years. The rest of the flow is handled by sink 3, which is nearly used up at the end of the 20$^{th}$ year.

## HANDLING UNCERTAINTIES IN THE SYSTEM

In a system, there could be uncertainties about aquifer characteristics such as maximum flow and total capacity, or there could be a possibility of a leak for a particular sink. If we have a probability distribution that can characterize these uncertainties, we can generate possible scenarios and find the optimal solution for each scenario while minimizing the expected cost over all of them. Clearly, the more scenarios generated, the better the decisions will be on how to design the optimal network. The model we work with in this stochastic framework is

$$
\begin{aligned}
&\min \mathrm{E}\big(NPV(s,L)\big) \\
&\text{s.t.} \sum_{j\in D} s_{ijt}^{\omega} + \sum_{k\in R} s_{ikt}^{\omega} = Q_i^{\omega}, i \in O, t \in T, \omega \in \Omega \\
&\sum_{i\in O} s_{ijt}^{\omega} + \sum_{k\in R} s_{kjt}^{\omega} \le smax_j^{\omega}, j \in D, t \in T, \omega \in \Omega \\
&\sum_{i\in O} s_{ikt}^{\omega} - \sum_{j\in D} s_{kjt}^{\omega} = 0, k \in R, t \in T, \omega \in \Omega \\
&\sum_{t\in T}\left(\sum_{i\in O} s_{ijt}^{\omega} + \sum_{k\in R} s_{kjt}^{\omega}\right) \le scap_j^{\omega}, j \in D, \omega \in \Omega,
\end{aligned}
\tag{2}
$$

where E($NPV(s,L)$) is the *expected* Net Present Value computed over all scenarios, $\Omega$ is the set of all scenarios, and all the flow variables and all the data are now indexed also by scenario.

The specific example we consider here is that the sinks in Figure 1(a) have uncertain total capacities. We will assume that the capacities of sinks one and two are uniformly distributed between 0 and 15 and the capacity of sink 3 is uniformly distributed between 0 and 30. We stipulate that the lifetime of the system should be at least 10 years and generate 100 scenarios, each with a total sequestration capacity of at least 30 units. Solving this model gives that the scenarios with larger capacities for sinks one and two send very little flow to sink three as it is farther away, and scenarios with smaller capacities for sinks one and two rely heavily on sink three.

The example above is fairly simple, but illustrates the power of the optimization technique used. With 100 scenarios, the model has 6000 variables and 2519 constraints (after preprocessing). The solution time for this model is 7 seconds on a machine with a 400MHz CPU. In fact, the optimization approach presented here has been successfully used to solve similar problems in the financial sector that have millions of variables and constraints. In all of these models, more complicated probability distributions can be used or the nonlinear solver can be connected to simulation software.

## USING DYNAMIC PROGRAMMING TO MODEL FLOW OVER TIME

Another way to think about the problem of sinks with capacities is to consider the allocation of available capacity to yearly usage over the lifetime of the pipeline network. Such a problem can be handled using the standard optimization technique of Dynamic Programming. The resulting problem is recursive, and it is solved starting at the end of the lifetime and working back to time period 1.

$$
\begin{aligned}
&f_t(scap) = \min Cost(s,L) + f_{t+1}(scap - s^j) \\
&\text{s.t.} \sum_{j\in D} s_{ijt} + \sum_{k\in R} s_{ikt} = Q_i, i \in O \\
&\sum_{i\in O} s_{ijt} + \sum_{k\in R} s_{kjt} \le \min(smax_j, scap_j), j \in D \\
&\sum_{i\in O} s_{ikt} - \sum_{j\in D} s_{kjt} = 0, k \in R
\end{aligned}
\tag{3}
$$

where $Cost(s,L)$ is the cost for the current time period $t$, with flows $s_{ijt}$, and $s^j$ is the total flow into sink $j$.

Using dynamic programming is a fairly classical approach in pipeline design, and a good survey of the progress of the last several decades is provided in [2]. At this stage in our study, we prefer using a straightforward nonlinear programming approach as it seems to work quite well. In fact, we wrote a piece of code to implement the recursion given by (3) and were easily able to verify the results from LOQO as being the global optimum on this small problem presented in Figures 1(a) and 1(b).

## CONCLUSION

We have presented here mathematical programming techniques that can be used to build and solve models representing a $CO_2$ sequestration network. Preliminary numerical results were also presented. The work that is done follows closely on the footsteps of previous work in pipeline design, as outlined in [2], [3], [4], [5], [6] and [8]. There are two main contributions of this study: (1) we bring the state-of-the-art modeling and optimization technology into solving such problems. The solver used in our study, LOQO, has been documented [1] as a leader in large-scale nonconvex nonlinear optimization, and (2) we incorporate a nonconvex nonlinear objective function to represent the cost of building and operating a pipeline for $CO_2$ sequestration.

An important future consideration here is to bring real-world constraints into our problem. Besides the uncertainties that were mentioned in the previous section, we also need to consider incorporating geographical concerns into the mathematical model. For example, the location of existing rights of way may determine the actual length and location of the pipeline. This issue can be handled by using a look up table for the distances between sources and sinks along rights of way, replacing the Euclidean length computation that is currently used in the model. Pipeline capital costs reported in the literature vary widely depending on the terrain and location. These variations could be added to a geographic specific model.

There are two further steps in our study. The first is to apply our techniques to actual data for possible $CO_2$ emission sites and sequestration sites. The second is to introduce the concept of larger scale fossil energy complexes with hydrogen production and $CO_2$ capture into the model and look at a more general framework as described in [9].


## ACKNOWLEDGEMENTS

The authors wish to thank British Petroleum and the Ford Motor Company for their support of the Carbon Mitigation Initiative where this research was conducted. We also thank Robert Socolow and Robert Vanderbei of Princeton University for helpful comments throughout this work.


## REFERENCES


[1] H.Y. Benson, D.F. Shanno, and R.J. Vanderbei. A Comparative Study of Large Scale Nonlinear Optimization Algorithms. *To appear in the Proceedings of the Workshop on High Performance Algorithms and Software for Nonlinear Optimization, Erice, Italy, 2001.*

[2] R.G. Carter. Pipeline Optimization: Dynamic Programming after 30 years. *Technical Report 9803, Pipeline Simulation Interest Group, 1998.*

[3] B. Djebedjian, A. Herrick, and M.A. Rayan. Modeling and Optimization of a Potable Water Network. *Proceedings of the 2000 International Pipeline Conference* 2:1227-1233, 2000.

[4] T.F. Edgar, D.M. Himmelblau, and T.C. Bickel. Optimal Design of Gas Transmission Networks. *Society of Petroleum Engineers Journal* (April 1978), 96-104.

[5] D.E. Goldberg, and C.H. Kuo. Genetic Algorithms in Pipeline Optimization. *Journal of Computing in Civil Engineering* 1(2):128-141, 1987.

[6] D.C. McClure. Using Linear Programming to Optimize Pipeline Analysis. *Paper presented at the annual PSIG conference, 1982.*

[7] Z.A. Melzak. On the problem of Steiner. *Canadian Mathematical Bulletin*, 16:143-148, 1961.

[8] M. Mohitpour, H. Golshan, and A. Murray. *Pipeline design & construction : a practical approach* New York : ASME Press, 2000.

[9] J.M. Ogden. Modeling infrastructure for a fossil hydrogen energy system With $CO_2$ sequestration. *To appear in the Proceedings of the Sixth InternationalConference on Greenhouse Gas Control Technologies, Kyoto, Japan, 2002.*

[10] R.J. Vanderbei and D.F. Shanno. An interior-point algorithm for nonconvex nonlinear programming. *Computational Optimization and Applications* 13:231-252, 1999.

[11] R.J. Vanderbei and H. Yurttan. Using LOQO to solve second-order cone programming problems. Technical Report, SOR-98-09, Statistics and Operations Research, Princeton University, 1998.