Milk Collection Problem: Integrating the Traveling Salesman and Set Covering Problem - A Case Study in West Virginia, USA

Md Rabiul Hasan
West Virginia University

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Part of the Business Analytics Commons, Industrial Engineering Commons, Operational Research Commons, Operations and Supply Chain Management Commons, Other Operations Research, Systems Engineering and Industrial Engineering Commons, Systems Engineering Commons, and the Transportation and Mobility Management Commons

Recommended Citation
Hasan, Md Rabiul, "Milk Collection Problem: Integrating the Traveling Salesman and Set Covering Problem - A Case Study in West Virginia, USA" (2024). Graduate Theses, Dissertations, and Problem Reports. 12468.
https://researchrepository.wvu.edu/etd/12468

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.
Milk Collection Problem: Integrating the Traveling Salesman and Set Covering Problem - A Case Study in West Virginia, USA

Md Rabiul Hasan

Thesis submitted to the Benjamin M. Statler College of Engineering and Mineral Resources at West Virginia University in partial fulfillment of the requirements for the degree of

Master of Science

in

Industrial Engineering

Thorsten Wuest, PhD, Committee Chair
John Saldanha, PhD
Zeyu Liu, PhD
Department of Industrial and Management Systems Engineering

Morgantown, West Virginia
2024

Keywords: Traveling Salesman Problem, Dairy, Set Covering Problem, MILP, Multi-Vehicle Routing

Copyright 2024 Md Rabiul Hasan
ABSTRACT

Milk Collection Problem: Integrating the Traveling Salesman and Set Covering Problem - A Case Study in West Virginia, USA

Md Rabiul Hasan

Route determination for perishable products is complex due to its unique characteristics, such as limited shelf-life regulatory requirements, or possibility of getting damaged. This research investigates a novel problem of collecting raw milk from a rural network of dairy farms. The research problem is grounded in a real scenario of milk collection in West Virginia, USA. The milk in this scenario is produced by small farms incapable of realizing transportation economies of density out in mostly rural areas throughout the state. Maximum coverage area and milk processing overhead costs are used to identify suitable locations for intermediate milk collection centers or depots to store the milk to realize economies of density and reduce transportation costs. Each depot needs to be established within a certain maximum distance, for the refrigeration time of the multi-stop vehicle to not exceed the allowable time limit set to maintain the quality of the collected milk. This problem provides the unique opportunity to incorporate two separate classical optimization problems: the Set Covering Problem (SCP) (identifying depot locations) and the Traveling Salesman Problem (TSP) (routing). The SCP involves identifying the optimal number of service facilities required for the aggregation operations to maintain milk quality in transit. The TSP determines the most optimal route between farms and the depot. The milk collection problem requires solving both the SCP and TSP. However, the problem also becomes more difficult to solve when combining TSP constraints with the SCP. This study proposes a novel Mixed Integer Linear Programming (MILP) model to address this problem. The objective of the proposed model is to minimize the depot assignment cost, overhead cost of the depot, cleaning cost of the vehicle, and the vehicle distance traveled to reduce the fuel cost. The exact algorithm has been analyzed and we use sensitivity analyses to determine the model’s reliability and robustness to changes in the problem scenario. The model was tested for different scenarios for the dairy industry in West Virginia. However, it has to be noted that other applications can be developed for similar structured problems based on this study given the flexible working path created by the Application Programming Interface (API) of Google Maps. We evaluate the proposed model by comparing its result with the steepest ascent Hill climbing algorithm, a mathematical optimization problem in Artificial Intelligence (AI) and Nearest Neighbor heuristics. The two algorithms are compared regarding solution quality and computational efficiency to determine the better heuristic algorithm for the developed model. The Hill climbing algorithm has given significantly better results than the Nearest Neighbor heuristics. The Hill climbing algorithm ended up in near-optimal results with an efficient computational time. Further, the Multi-Vehicle Routing (MVR) model is analyzed for the transportation part of the model and found that MVR scenario shows the potential over other scenario (TSP-based) based on vehicle cycle time, still, there is a door to future research to incorporate the heterogeneous fleet and multi-depot constraints in the developed model.
ACKNOWLEDGEMENT

It is my pleasure to acknowledge all the people who have helped me to prepare this thesis. First and foremost, I would like to thank my research supervisor, Dr. Thorsten Wuest, for his valuable guidance and advice at different stages of this research study. It would be impossible to carry on this thesis work and make it into the final shape of a thesis without his guidance and sympathetic encouragement. I am truly blessed for his valuable suggestions and support.

I am grateful to the committee members, Dr. John Saldanha and Dr. Zeyu Liu, who motivated me and gave their valuable insights and feedback over the time of this research work.

Finally, my heartiest thanks go to my family members, whose love, affection, and encouragement have always inspired me, and I could not come to this stage of my life without their constant prayer and support. Last, I am grateful to the Almighty Allah for enabling me to bear all the stress and complete the research in due time.
# Table of Contents

1 Introduction ............................................................................................................. 1  
  1.1 Background of the study ................................................................. 1  
  1.2 Research Gaps ................................................................................. 6  
  1.3 Objectives ....................................................................................... 6  

2 Traveling Salesman Problem (TSP) ..................................................................... 7  
  2.1 Classification of TSP ...................................................................... 11  
    2.1.1 STSP ................................................................. 11  
    2.1.2 ATSP ................................................................. 11  
    2.1.3 MTSP ................................................................. 11  

3 Set Covering Problem (SCP) ............................................................................ 13  

4 Problem formulation .......................................................................................... 14  
  4.1 Mathematical Model ...................................................................... 16  
  4.2 Parameters ....................................................................................... 17  
  4.3 Variables ......................................................................................... 18  

5 Results and Discussion ...................................................................................... 20  
  5.1 Scenario 1 ..................................................................................... 20  
  5.2 Scenario 2 ..................................................................................... 23  
  5.3 Scenario 3 ..................................................................................... 26  

6 Multi-vehicle Routing (Scenario 4) ..................................................................... 30  

7 Economic Feasibility Analysis .......................................................................... 35  
  7.1 For scenario 1 ................................................................................. 35  
  7.2 For scenario 4 ................................................................................. 35  

8 Conclusion, limitation, and future works ......................................................... 36
List of Tables

Table 1: Database search keyword terminology ................................................................. 1
Table 2: Summary of selected paper references ................................................................... 5
Table 3: All possible solutions (Brute force) ....................................................................... 9
Table 4: Transportation cost ............................................................................................... 22
Table 5: Total cost of the optimal scenario ......................................................................... 23
Table 6: Variable hours in the MCD .................................................................................. 25
Table 7: Different overhead cost of the depot based on the state ......................................... 26
Figure 11: Route of Nearest neighbor ................................................................................ 28
Table 8: Multi-vehicle routes and durations ....................................................................... 32
Table 9: Computational time of different scenarios ............................................................ 34
Table 10: Economic feasibility of the transportation models .............................................. 36
List of Figures

Figure 1: Number of publications related to TSP is dairy supply chain .............................................. 2
Figure 2: Hamilton path .......................................................................................................................... 8
Figure 3: Hamilton cycle .......................................................................................................................... 8
Figure 4: Application of MTSP (Cheikhrouhou & Khoufi, 2021) ............................................................. 12
Figure 5: Dairy farms’ location at West Virginia ...................................................................................... 15
Figure 6: Visual representation of the vehicle routing problem ................................................................. 16
Figure 7: Optimal scenario of the depot location ...................................................................................... 21
Figure 8: Vehicle route for the optimal scenario ....................................................................................... 23
Figure 9: Sensitivity analysis of depot location ....................................................................................... 24
Figure 10: Nearest neighbor algorithm ..................................................................................................... 27
Figure 11: Route of Nearest neighbor ....................................................................................................... 28
Figure 12: Route durations of Steepest Ascent Hill climbing (SAHC) algorithm within a time frame .......................................................................................................................... 29
Figure 13: Routing path of steepest ascend Hill climbing ........................................................................ 30
Figure 14: Capacitated multi vehicle route path ......................................................................................... 34
1 Introduction

1.1 Background of the study

A literature review is conducted as the initial stage to identify the research gaps and the research questions for this master thesis project. Literature review is helpful to address the practical measures and know-how the research studies will be effective for a certain field (Knopf, 2006). A scientific literature review is a methodological process to explore the database search and help to track the current research topics and themes. A scientific literature review can be classified in two broad ways: i) Narrative literature (theoretical and contextual point of view); and ii) Systematic literature review (systematic and explicit methodology to critically evaluate a research topic and its trends and gaps) (Rother, 2007) (Hasan et al., 2024) (Rahman et al., 2024) (Bin Syed et al., 2023). The narrative and systematic literature review has been conducted for this research study to explore the previous literature and identify the research opportunities. The Scopus database has been used to analyze high quality published research articles that are peer-reviewed. A search string was developed to adequately reflect the research intent, and lead to a return of the relevant research articles, as shown in the Table 1 below:

<table>
<thead>
<tr>
<th>No.</th>
<th>Search terms</th>
<th>Focus area</th>
<th>No. of papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>i.</td>
<td>(&quot;Dairy*&quot;) OR (&quot;Milk&quot;)</td>
<td>Dairy Industry</td>
<td>393,551</td>
</tr>
<tr>
<td>ii.</td>
<td>(&quot;Vehicle routing problem*&quot;) OR (&quot;Capacitated Vehicle routing problem*&quot;) OR (&quot;Travelling salesman problem*&quot;)</td>
<td>TSP</td>
<td>23,422</td>
</tr>
<tr>
<td>iii.</td>
<td>((&quot;Vehicle routing problem*&quot;) OR (&quot;Capacitated Vehicle routing problem*&quot;) OR (&quot;Travelling salesman problem*&quot;)) AND ((&quot;Dairy*&quot;) OR (&quot;Milk&quot;) OR (&quot;Liquid&quot;)) AND ((&quot;Set covering problem&quot;) OR (&quot;facility location problem&quot;))</td>
<td>TSP+Dairy Industry + Set Covering Problem</td>
<td>None</td>
</tr>
<tr>
<td>iv.</td>
<td>(&quot;Vehicle routing problem*&quot;) OR (&quot;Capacitated Vehicle routing problem*&quot;) OR (&quot;Travelling salesman problem*&quot;)</td>
<td>Combined i and ii search terms</td>
<td>99</td>
</tr>
</tbody>
</table>
From Table 1, it can be derived that none of the returned Scopus-indexed article covers the Travelling Salesman Problem (TSP), dairy, and Set Covering Problem (SCP) related search words simultaneously based on their title and abstract. Therefore, the number iv search string has been selected to identify the related previous research articles closely aligned with this thesis work. As the Figure 1 shows, there has been an increase in the number of research articles published on the topics of TSP in the dairy supply chain industry during the past 20 years.

![Figure 1: Number of publications related to TSP is dairy supply chain.](image)

This study only explores top peer-reviewed articles in operations research field to ensure quality and transparency to identify the research gaps. Most of the dairy farms' locations are far from the milk processing center, often in rural areas. This situation is similar globally. Hence, it is necessary to collect the milk from the dairy production locations and transport it to the closest milk processing center by tanker (Polat et al., 2022). Enzymatic degradation happens in raw milk due to the growth of heat-sensitive organisms; hence, it is important to process it after collection (Marth & Steele, 2001). Griffiths et al. (1987) found that raw milk storage lives can be increased for more than 72 hours at two degrees Celsius after a 65-degree thermization treatment for fifteen seconds.
Time-dependent variables and costs are important for the perishable food product supply chain, such as milk. Therefore, it is necessary to design and develop a cost-effective service network comprising dairy farms and milk processing centers (Polat & Topaloğlu, 2022). The main decision is to determine the collection tanker’s travel sequence, known as the Milk Collection Problem (MCP) (Polat et al., 2022). Paredes-Belmar et al. (2022) reported three main challenges associated with MCP. They are:

1. Milk production in very remote places
2. Limited shelf life of raw milk
3. Milk production varies from farm to farm (Paredes-Belmar, Montero, & Leonardini, 2022).

The transportation challenges and their solution have yet to be explored among the researchers (Nadal-Roig & Plà-Aragonés, 2015).

One of the earliest works about the MCP has been found in the work of Sankaran and Ubgade (1994) about a case study in India (Sankaran & Ubgade, 1994). It is very common to find research on the types of milk in the MCP among the researchers. In 2009, Dooley et al. published a paper in which they considered two milk types in different trucks for each milk for a production network in New Zealand (Dooley et al., 2005). Caramia and Guerriero (2009) analyzed four different types of milk in a separated compartment using trucks and trailers (Caramia & Guerriero, 2010). In an analysis, Paredes-Belmar et al. (2016) calculated the profit of blending the different milk qualities in the same truck (Paredes-Belmar et al., 2016). Later, a recent study by Belmar et al. (2022) further proposed a vehicle routing model with graded milk that is graded in three categories based on the somatic cell percentage of milk. In their study, they considered different qualities of milk blended in one truck (Paredes-Belmar, Montero, Lüer-Villagra, et al., 2022). Mason et al. (2015) analyzed the MCP based on the weekly demand and production variation of milk transportation (Masson et al., 2016). On Belenguer et al. (2016) study examined the milk collection strategy where a large vehicle cannot be reached. It detached the trailer in an appropriate parking place for 50 customers with ten trailer points (Belenguer et al., 2016). Nguyen et al. (2022) researched the lower bound of the vehicle capacity on the vehicle routing problem with the efficiency of the proposed method, decreasing the solution time within two hours, which previously needed one day (Nguyen et al., 2022). Tarantilis and Kiranoudis (2007) explored and compared the vehicle routing results of heterogeneous fleets for dairy and construction case studies with previous works (Tarantilis & Kiranoudis, 2007). In 2015, Dayarian et al. published a paper...
that is the first to explore multi-attribute vehicle routing with heterogeneous fleets and multiple depots. They developed a path-based set partitioning model for their case study (Dayarian et al., 2015).

The Table 2 is quite revealing in several ways. First, as seen from the Table 2, it is found that most of the MCPs are MILP-based single commodity investigations. CPLEX and Gurobi optimization tools are popular among researchers. The MCP base's objective function can be classified as maximizing the profit or minimizing the cost. Interestingly, no dominant algorithm is found in the MCP; it ranges from the exact algorithm like branch-and-cut to a metaheuristic algorithm like Tabu search. Regarding the constraint, the heterogeneous fleet with a time window and demand and capacity constraints are the most influential among the previous works of literature.
<table>
<thead>
<tr>
<th>References</th>
<th>Title</th>
<th>Objective function</th>
<th>Single/ Multicommodity</th>
<th>Algorithm</th>
<th>Mathematical model</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Paredes-Belmar, Montero, Villagra, et al., 2022)</td>
<td>Vehicle routing for milk collection with gradual blending: A case arising in Chile</td>
<td>Maximizing the profit</td>
<td>Multi-commodity</td>
<td>Branch and cut</td>
<td>MILP</td>
</tr>
<tr>
<td>(Nguyen et al., 2022)</td>
<td>Modeling and solving a multi-trip multi-distribution center vehicle routing problem with lower-bound capacity constraints</td>
<td>Minimization of the cost</td>
<td>Single</td>
<td>Adaptive large neighborhood search</td>
<td>MILP</td>
</tr>
<tr>
<td>(Polat et al., 2022)</td>
<td>Modeling and solving the milk collection problem with realistic constraints</td>
<td>Minimization of the cost</td>
<td>Multi-commodity</td>
<td>Variable neighborhood search</td>
<td>MILP</td>
</tr>
<tr>
<td>(Belenguer et al., 2016)</td>
<td>A branch-and-cut algorithm for the single truck and trailer routing</td>
<td>Minimization of cost</td>
<td>Single</td>
<td>Branch-and-cut</td>
<td>MILP</td>
</tr>
<tr>
<td>(Dayarian et al., 2015)</td>
<td>A column generation approach for a problem with satellite depots</td>
<td>Minimization of the cost</td>
<td>Single</td>
<td>Price</td>
<td>MILP</td>
</tr>
<tr>
<td>(Claassen &amp; Hendriks, 2007)</td>
<td>An application of Special Ordered Sets to a periodic milk collection problem</td>
<td>Minimization of the deviation of the duration of the trips of the trips</td>
<td>Single</td>
<td>Branch and bound</td>
<td>MILP</td>
</tr>
<tr>
<td>(Tarantilis &amp; Kiranoudis, 2007)</td>
<td>A flexible adaptive memory-based algorithm for real-life transportation operations: Two case studies from A flexible adaptive memory-based algorithm for real-life transportation operations: Two case studies from</td>
<td>Minimizing the cost</td>
<td>Single</td>
<td>Tabu search</td>
<td>MILP</td>
</tr>
<tr>
<td>(Paredes-Belmar, Montero, Leonardini, 2022)</td>
<td>A milk transportation problem with multi-commodity center and vehicle routing</td>
<td>Minimizing the cost</td>
<td>Single</td>
<td>Iterated local search</td>
<td>MILP</td>
</tr>
</tbody>
</table>
## 1.2 Research Gaps

The literature review found that different echelons of vehicle routing problems have been explored, like the study conducted by Jie et al. in 2019 (Jie et al., 2019). These kinds of models need a more holistic approach issue. Furthermore, the multi-echelon model, where the solution of the first model is utilized in the second model, results in both inefficiencies and hierarchical error. To summarize the results of the literature review, to the best of our knowledge, no previous research on MCP focused on the combined SCP and TSP algorithm with milk tanker cleaning cost every time the vehicle visits a depot to find the objective function of their mathematical model.

## 1.3 Objectives

The following points have been summarized in the contribution of this study:

1. The novelty of this work is introducing the MCP with the combined algorithm of TSP and SCP. A real-world dairy problem inspires the problem in West Virginia, USA. This

<table>
<thead>
<tr>
<th>Optimization software</th>
<th>CPLEX</th>
<th>CPLEX</th>
<th>CPLEX</th>
<th>CPLEX</th>
<th>N/A</th>
<th>N/A</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major Constraints</td>
<td>Multi commodity, somatic cell count, heterogeneous truck</td>
<td>Heterogeneous fleet, Demand and capacity constraint, Loading and service time constraint</td>
<td>Heterogenous Fleet; Fixed Fleet; Time Window; Time Limit; Divisible Demand; Location Requirement; Multi-Tank (Multi-compartment); Multi Product; Tank Loading Ratio; Fuel Consumption Rate</td>
<td>Tank capacity, multi depot</td>
<td>Heterogeneous fleet vehicle, multiple depots, time window, capacity</td>
<td>Time, supply and demand</td>
<td>Heterogeneous fleet, Time window, Demand and capacity</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
research determines the required number of Milk Processing Centers (MPCs) and their locations with the specific milk tanker routes at an operative level decision.

2. The study proposes a MILP based model that considers for determining the number of MPC/depot needs and assigning farms to the nearest MPC/depot with specific vehicle tanker routes to minimize the total cost of the problem.

3. Finally, a real-world scenario with 33 farms in West Virginia, USA has been investigated and solved using the developed MILP model. Moreover, the model has been further analyzed with the Multi-Vehicle Routing (MVR). This study also evaluated the proposed model by comparing its result with the steepest ascent Hill climbing algorithm, a mathematical optimization problem in Artificial Intelligence (AI) and Nearest Neighbor heuristics.

2 Traveling Salesman Problem (TSP)

TSP is one of the most studied research problems that researchers are investigating in machine learning and operation research (Yang & Rajgopal, 2020). Since 1930, the combinational optimization problem of TSP has been found, and it was Merril M. Flood who first developed a mathematical formulation for the school bus routing problem (Pop et al., 2023). TSP has been used across many domains, ranging from manufacturing, planning, neuroscience, genetics, telecommunication, supply chains, and logistics to healthcare (Marecek, 2008). In 1964, Clarke and Wright described a “Truck Dispatching Problem” in a linear optimization model, which was first introduced by Dantzig and Ramser (1959) (Clarke & Wright, 1964) (Dantzig & Ramser, 1959). In their model, they formulated how to serve a set of customers using a fleet of trucks with varying capacities from a central depot. Later, this problem became known as a Vehicle Routing Problem (VRP), a heavily studied topic in the operation research field.

One of the earliest studies in the 18th century focused on the TSP was conducted by the two mathematicians Sir William Rowan Hamilton and Thomas Pennington Kirkman from Ireland and England, respectively. TSP can be defined as the best route to visit all the locations from a starting point and return to the locations where it started with minimization of the travel distance when a set of cities and distance matrix among the pairs has been given (Davendra, 2010). This route is called the Hamilton cycle. A path given in a graph with an M number of the vertex is known as the Hamilton Path of M. A graph where the cycle contains every vertex of M is called the
Hamiltonian Cycle (Brucato, 2013). Figure 2 shows the Hamilton path, and the path sequence is {1-2-3-4}. {1-2-3-4-1} is a Hamilton Cycle that is shown in Figure 3.

Despite TSP being a simple problem to describe, it poses a significant challenge to solve, making it a classification of NP-hard problems. NP-hard problems include not only the inequalities but also recurrence relations. If there are n number of locations, covering all locations with a feasible set of locations or Hamiltonian Cycles TSP model can be represented as (n-1)!/2 with three or more vertices (Davendra, 2010). So, as the value of n increases the TSP problem is also getting more difficult to solve (Plaisted, 1984). It can be inferred that a computer can easily solve the TSP. However, a computer tries to find every possible Hamilton cycle for each Hamilton Path. As the n gets larger, the possible number of Hamilton cycles (n-1)!/2 also increases, as shown in Table 3.
Table 3: All possible solutions (Brute force)

<table>
<thead>
<tr>
<th>n</th>
<th>((n-1))/2</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>60</td>
</tr>
<tr>
<td>10</td>
<td>181,440</td>
</tr>
<tr>
<td>20</td>
<td>6.08E+16</td>
</tr>
<tr>
<td>30</td>
<td>4.42E+30</td>
</tr>
</tbody>
</table>

So, if the value of \(n\) increases, the computer will need much more time to solve the problem (Brucato, 2013). This type of problem can be solved by an exhaustive search. It becomes more challenging when the instances grow, and the time needed for this exhaustive search also increases. With the increased advancement of computer speed, a TSP problem can be solved effectively under some circumstances, but it is still not enough for a large number of locations (Woeginger, 2003). This problem can be referred as the heart of any logistic management scenario.

The TSP lies on the complexity class where the complexity class can be defined based on the problem, resource, and computational time (Brucato, 2013). Turing introduced the concept of the Turing Machine before the invention of the computer. The Turing machine is an automatic machine that can read the entry on the paper. If there is a previously determined action, it can execute that by moving the paper either right or left in each iteration. Different variations exist of Turing Machines, such as Binary machines, Non-writing machines, multiple tapes, deterministic, nondeterministic, weak, and semi-weak machines (De Mol, 2018). The complexity of class P can be defined as a problem that can be solved with the deterministic Turing Machine and represented in a polynomial function (Brucato, 2013). This can be explained by a set of \(n\) letters \(\{b,d,c,a\}\) to put it in an alphabetical sequence where the algorithm can explore two letters at a time and decide if they are right in order or need to switch. For example, in the first step, they need not switch their place among the letters \(b\) and \(d\) as they are in the right order. In the follow-up step, the algorithm will look at the alphabetical sequence between \(d\) and \(c\), where it will switch the order of the letters. The algorithm will iterate these steps as long as it does not alphabetically make the whole sequence
of the letters (De Mol, 2018). The whole iterations of this algorithm are shown below, and the comparison in each step is shown in bold.

$$\{b,d,c,a\} \rightarrow \{b,d,c,a\}$$
$$\{b,d,c,a\} \rightarrow \{b,c,d,a\}$$
$$\{b,c,d,a\} \rightarrow \{b,c,a,d\}$$
$$\{b,c,a,d\} \rightarrow \{b,c,a,d\}$$
$$\{b,c,a,d\} \rightarrow \{b,a,c,d\}$$
$$\{b,a,c,d\} \rightarrow \{a,b,c,d\}$$

On the other hand, a decision problem can be answered with binary variables such as either “yes” or “no”. Now, suppose a decision problem can be solved in a polynomial time with a nondeterministic Turing Machine, which can be verified by the deterministic Turing machine. In that case, it is known as the complexity class NP. A TSP is one kind of decision problem where, under a certain number of n locations, it has been asked if there exists a route that is less than a predefined distance like D (total distance) (Brucato, 2013). Hence, the complexity of TSP lies under the NP complexity class. NP complexity class problem can be aid NP-complete if it can be reduced to a problem in a polynomial time where P ≠ NP (Aaronson, 2005). One of the simplified versions of the NP-complete problem is TSP if there exists a Hamiltonian Cycle under a given graph G (Brucato, 2013). A problem can be said to be NP-hard if it is as difficult to solve as NP-complete (Knuth, 1974). In the earliest work on defining NP-hard problems, Garey and Johnson stated that a decision problem, whether a member of NP or not, if it can transform an Np-complete problem which cannot be solved in polynomial time except P=NP, then it is an NP-hard problem (Garey & Johnson, 1990). The TSP is classified as an NP-hard problem since it is at least as hard as an NP-complete problem, where it tries to find the shortest path among all the possible Hamilton Cycles for a given set of locations n on the graph G (Brucato, 2013). Therefore, if it is possible to have an efficient polynomial time-based algorithm for the TSP, it can be applied to all other NP-complete class problems (Hoffman et al., n.d.). Nevertheless, there is no such polynomial time-based TSP algorithm without proof that P ≠ NP is one of the most explored research problems among researchers (Davendra, 2010) (Zhou et al., 2015) (Ouaarab et al., 2014).

Nowadays, investigating the VRP model is different from the study of Dantzig & Ramser (1959) and Clarke & Wright (1964) as more investigation is going to present real-world scenarios
such as time window for pick-up and delivery, time-dependent travel time, and other variables like demand that changes over the time horizon. The single vehicle routing problem is commonly known as TSP (Braekers et al., 2016). In 2009, Eksioglu et al. showed that the research on VRP is increasing exponentially, with a 6% growth rate every year, making it difficult to track the advancement of this field. In their analysis, they described that there is no exact taxonomy that can describe the VRP problem (Eksioglu et al., 2009). There are different reviews have been found on the vehicle routing scenario where a different aspect of VRP is discussed such as dynamic and stochastic VRPs (Ritzinger et al., 2016) (Pillac et al., 2013) (Ojeda Rios et al., 2021), multiple depots (Montoya-Torres et al., 2015), capacitated VRPs (Jie et al., 2019), Green VRPs (Moghdani et al., 2021) (Lin et al., 2014), multi objective VRPs (Jozefowiez et al., 2008), time windows (Bräysy & Gendreau, 2005) (El-Sherbeny, 2010), heuristics aspects of VRPs (Laporte et al., 2000).

2.1 Classification of TSP

TSP can be categorized broadly into three sections. The first one is the Symmetric Traveling Salesman Problem (STSP), the second one is the Asymmetric Traveling Salesman Problem (ATSP), and the final one is the Multi Traveling Salesman Problem (MTSP) (Davendra, 2010).

2.1.1 STSP

If there are a set of locations \( L = \{L_1, L_2, \ldots, L_n\} \) and the edge set is \( V = \{(p, q): p, q \in L\} \), here is a cost measure related to the edge \( (p, q) \in V \) is the Euclidean distance between the location \( p \) and \( q \), and the location \( L_i \in L \), where their co-ordinates are denoted by \( (x_i, y_i) \). So, in the STSP, the distance between the two locations on graph \( G \) will be the same in either direction (Gutin & Punnen, 2002).

2.1.2 ATSP

If the distance between two locations among any set of locations \( L = \{L_1, L_2, \ldots, L_n\} \) is not Euclidean or the distance is not the same in either direction of the two vertex, then the problem is referred to as ATSP. Hence, in the ATSP, for at least one \( (p, q) \), \( d_{pq} \neq d_{qp} \) (Davendra, 2010).

2.1.3 MTSP

Let there be a single depot with \( m \) salesmen in a given set of locations under Graph \( G \). The remaining locations other than the depot will be considered intermediate locations. In MTSP, all the salesmen will start from the depot and return to the depot after visiting all the intermediate locations exactly once, minimizing the total cost of visiting all the locations. In MTSP, the cost
can be defined in different metrics based on distance, time, etc. Further, the MTSP can be further categorized based on different scenarios.

For example, in a single depot MTSP problem, all the salesmen complete the tour at a single point. On the other hand, in multiple depot MTSP, the salesman can return to any depot as long as all the depots have the same number of salesmen. It can also be classified based on the number of salesmen (can be fixed or bound), cost (cost associated with the number of salesmen), time (referred to as MTSP with specified timeframe where some locations need to be traveled in a particular time period), constraints (number of locations can visit, maximum or minimum distance a salesman can travel). Again, if there is a situation of only one salesman, the MTSP becomes the traveling salesman problem (Davendra, 2010). For a real-life scenario to optimize the path of vehicles, even for Unmanned Aerial Vehicles (UAVs: Operating aircraft without human pilot), MTSP has been widely used to optimize the vehicle routing path. Nowadays, it is not only limited to the application of transportation and delivery but also in the field of disaster management, monitoring and surveillance, precision agriculture, multi-robot task allocation, and scheduling, as shown in Figure 4.

![Application of MTSP](image)

Figure 4: Application of MTSP (Cheikhrouhou & Khoufi, 2021)
3 Set Covering Problem (SCP)

Facilities location plays a crucial role in the strategic planning of different organizations. Hence, evaluating the distance and time from the demand points is important, which is associated with costs (Owen & Daskin, 1998). Very few problems are as popular as covering problems among the facility location problem for its application in the real world (Farahani et al., 2012). For a quick response time, it is important to determine the service facilities, such as fire or ambulance stations, within a certain distance from the customer or the demand locations, known as the covering problem (García & Marín, 2015). In 1957, Claude Berge who first introduced the covering problem to find a minimum cover on a graph (Berge, 1957). For the protection of a highway network, later in 1965, Hakimi investigated the number of minimum police patrols necessary (Hakimi, 1965). Toregas et al. (1971) first developed the mathematical formulation of the location area problem.

The model can be described as follows:

- \( i \): index of customer points.
- \( j \): index of facilities.
- \( d_{ij} \): the distance from location \( i \) to facility point \( j \).
- \( x_j \): binary variables: 1 if the facility established at point \( j \), 0 otherwise.

\[
\text{Min} \quad Z = \sum_{j=1}^{n} x_j \quad (1)
\]

Subject to,

\[
\sum_{j \in N_j} x_j \geq 1 \quad \forall i = 1, \ldots, m \quad (2)
\]

\[
x_j \in \{0,1\} \quad \forall j = 1, \ldots, n \quad (3)
\]

Equation (1) is the objective function to minimize the total number of facility locations. Constraints (2) and (3) describe the service requirements at node \( i \) and the integrity constraint, respectively (Toregas et al., 1971).

In a literature review by Schilling et al. (1993) on covering problems in facility location, they categorized the concept of covering problems in two methods (Schilling et al., 1993). The first one is the set covering problem establishing the minimum number of facilities in order to cover every customer or demand point. Here, the coverage is required, and the total number of facilities are
minimized if all the facilities have a similar cost of establishment (García & Marín, 2015). The second one is the maximal covering problem introduced by Church and Velle (1974) (Church & Velle, 1974). In business, it might not be possible to establish facilities based on the requirements for the budget constraint. That is why minimizing the non-covered demand locations or maximizing the covered demand locations is needed, and these properties all belong to the maximal covering problem (Laporte et al., 2015). In their study, Balas and Padberg (1976) argued that the set covering problem is one of the most applied models, along with the traveling salesman problem and set partitioning in integer programming (Balas & Padberg, 1976). In set covering problem, which covers a matrix of m rows and n columns with zero one (a_{ij}) matrix of a subset of the columns at minimum cost.

Then the model can be represented associated with cost as follows:

\[ x_j = 1 \text{ if the facility established at point } j \text{ (cost } c_j) , 0 \text{ otherwise.} \]

\[
\text{Min} \quad Z = \sum_{j=1}^{n} c_j x_j
\]

Subject to,

\[
\sum_{j=1}^{n} a_{ij} x_j \geq 1, \quad \forall i = 1, \ldots, m \tag{2}
\]

\[
x_j \in \{0,1\} \quad \forall j = 1, \ldots, n \tag{3}
\]

In the above model, constraint (3) is an integrality constraint. Constraint (2) confirms that each row is covered by at least one column. If the inequalities notation is replaced with equalities, the problem will be considered a set partitioning problem (Beasley, 1987).

4 Problem formulation

For analyzing the routing problem, one of the most studied paradigms in the operations research fields, the distance matrix is the first step to start with. The primary data for this research is collected in collaboration with the West Virginia Department of Agriculture. Thirty-three dairy farms (represented as black dots on the map) are currently operational in West Virginia as depicted in Figure 5 among the fifty five counties (Bureau, n.d.).
For measurements of the travel distance between the farms are calculated by the Google Maps API. The acronym API stands for Application Programming Interface. According to Stylos & Myers (2007), API is the collection of codes and packages that helps other programmers use them to achieve their goals (Stylos & Myers, 2007). A google API key is first generated through the google map platform. The transportation mode is important for calculating the distances among the farms. For this analysis, the driving mode is considered, which indicates the distances using the road network. Although the driving mode is selected for measuring the road distances, the avoid parameter is also included. The routes of ferries are avoided while the start time is to calculate the distances set as the present time. The departure time cannot be set as past; it must be either the present or future. The result can be varied over time due to network changes, distribution of the nature of the service, and updated average traffic conditions.

Furthermore, the duration of the travel time between the farms has been calculated. When calculating the distances, the Direction API calculates the nearest transportation routes among the origin and destinations; for this analysis, it is the road as the transportation mode is selected as driving.

Figure 5: Dairy farms’ location at West Virginia
After that, several depots or the milk collection centers must be determined to cover certain farms. This gives the unique opportunity to model a combined TSP and SCP.

Two types of algorithms need to follow for this scenario. First, the farms need to be clustered based on the maximal coverage distance from the depot location, which is the SCP, and each node or collection point, here it is farm, needs to be visited exactly once in an optimized way so that the total cost will be minimized which represents the TSP. Figure 6 shows different farms' locations with circles (a…..g) where the triangular nodes (a,e,h) simultaneously are the depot locations and farm's location. A farm can be considered a depot based on the set covering algorithm with the maximal distance it covers. Here three clusters have been shown where the vehicle will be covered each cluster and then move forward to the next clusters to collect and deposit the milk in the depot location. It is imperative to note that the VRP solely pertains to the collection of milk from the dairy farms' premises and the primary milk depot located therein. The silos or storage points at the processing sites are not included in this study.

![Diagram](image)

**Figure 6: Visual representation of the vehicle routing problem**

A multi-stop truckload with a maximum volume of 9,000 gallons is used to haul the milk collection with a three-hour ability to maintain the quality of the stored milk. The tanker has an average speed of 55 miles/hr, resulting in a maximum coverage distance of 165 miles from the depot location. The milk tanker should be cleaned every time it enters back in the depot.

**4.1 Mathematical Model**

The assumptions that are considered here are as follows:

- The variable cost of assigning a farm location j to a depot location i is known and constant.
Throughout the planning horizon, the fixed cost for running a depot and the average transportation cost will be constant.

The milk production of each farm location will be constant, and the maximum vehicle capacity will be predefined.

The formulation is assumed for single-vehicle transportation, where the maximal coverage distance (D) is constant and determines whether a farm location j can be assigned to a depot location i based on the distance between them.

The distance matrix only represents the transportation time from one node to another node, excluding any loading and unloading times for milk.

The vehicle has followed an order to visit the depot locations during the trip to ensure that it returns to the location where the trip started after each trip is completed.

The cost of vehicle tanker is fixed, which is incurred every time the vehicle visits a depot before a new trip.

The distance between the location of farms is known and constant throughout the planning horizon, which is also the property of non-negativity values, and the distance from a farm location to itself is zero.

All the decision variables are binary, taking the value of either zero or one.

This study considers each farm as a single node with an \( n = 1 \ldots 33 \) farms locations. To avoid the symmetric problem, \( n+1 \) is defined as the depot location of the farms for returns. The variable cost for assigning a farm's location as a depot location and the fixed cost of running a depot are denoted by the terms \( C_{ij} \) and \( f_i \), respectively. \( p \) and \( g \) are the average cost of traveling per distance per gallon and the cleaning cost of the vehicle tanker. \( \sum_i Y_i \) is the total number of depots among the \( n \) number of farms. \( d_{ij} \) is the distance of farm location i to farm location j with a maximal coverage distance from the depot for the vehicle is D. \( M_i \) is the milk production at each farm location i where the vehicle tanker capacity is defined by the notation V.

The problem scenario can be represented by the following mathematical programming model:

4.2 Parameters

\( n \): Total number of targeted farm locations, where \( n = S \cup Q \)

\( \ell \): Index of locations; \( 1 \leq \ell \leq n \) for targeted farm locations; \( i = 0, n+1 \) for the origin
$f_i$: Fixed cost of running a depot at location $i$

$k$: Fixed cost of cleaning a vehicle at depot location $i$

$C_{ij}$: Variable cost of assigning a farm location $j$ to a depot at location $i$; $C_{ii} := 0$

$p$: Average transportation cost per unit of distance ($$/gallon/miles$$)

$d_{ij}$: Distance between location $i$ and location $j$ with $d_{ii} = 0$, and $d_{ij} = d_{ji}$

$D$: Maximal coverage distance (MCD) = 165 miles

$M_j$: Milk production of farm location $j$

$V$: Maximum vehicle capacity

$S$: set of depots

$Q$: Set of nodes served by each depot

$a_{ij} \in \{0,1\}$: 1 if location $j$ is within a distance $D$ from location $i$

### 4.3 Variables

$X_{ij} \in \{0,1\}$: 1 if location $j$ is assigned to depot at location $i$, 0 otherwise.

$Y_i \in \{0,1\}$: 1 if there is a depot at location $i$, 0 otherwise.

$Z_{ij} \in \{0,1\}$: 1 if location $i$ is followed by location $j$ on the trip route.

$U_i$: Continuous variable representing the order of visiting node $i$

### Objective function:

\[
\begin{align*}
\text{Min} & \quad \sum_{i} \sum_{j \leq n} \frac{M_j}{12} X_{ij} + \sum_{1 \leq i \leq n} f_i Y_i + \sum_{i \in S} \sum_{j \in N, j \neq i} p d_{ij} Z_{ij} + k \sum_{i} Y_i \\
\text{Subject to,} & \quad X_{ij} \leq a_{ij} \quad \forall i, j \\
& \quad a_{ij} \cdot Y_i \geq 1 \quad \forall 0 \leq i \leq n, 1 \leq j \leq n
\end{align*}
\]
\[ X_{ij} \leq Y_i \quad \forall 0 \leq i \leq n, \ 1 \leq j \leq n \quad (4) \]

\[ \sum_{i \in N} X_{ij} = 1 \quad \forall 1 \leq j \leq n \quad (5) \]

\[ \sum_{i \in N, i \neq j} Z_{ij} = 1 \quad \forall j \in Q \quad (6) \]

\[ \sum_{i \in N, i \neq j} Z_{ji} = 1 \quad \forall j \in Q \quad (7) \]

\[ U_i - U_j + |Q| Z_{ij} \leq |Q| - 1 \quad \forall i, j \quad (8) \]

\[ M_i \cdot Z_{ij} \leq V \quad 1 \leq i \text{ and } j \leq n \quad (9) \]

\[ 0 \leq Y_i \leq 1 \quad \forall i \in Q \quad (10) \]

\[ X_{ij} \in \{0, 1\} \quad \forall i, j \quad (11) \]

\[ Y_i \in \{0, 1\} \quad \forall i \quad (12) \]

\[ Z_{ij} \in \{0, 1\} \quad \forall i, j \quad (13) \]

\[ U_i \geq 0 \quad \forall i \quad (14) \]

The objective function (1) minimizes the sum of the variable cost of the facility, assignment cost of the depots, the transportation cost, and the cleaning cost of the vehicle tanker. Constraints (2)-(5) represents the set covering problem where, constraint (2) guarantees that a farm can only be assigned to a depot if the farm’s maximum location is within the maximum coverage distance \( D (=165 \text{ miles}) \). If so, \( a_{ij} = 1 \), otherwise 0. Constraint (3) ensures that each farm is assigned by at least one depot, and a depot can serve a farm only if there is a depot or the depot is open which is described by the constraint (4). Each farm \( (1 \leq j \leq n) \) is assigned to exactly one depot location \( i \) which is ensured by constraint (5).

Constraints (6)-(7) confirm the node visit constraint, where constraint (6) ensures there is only one incoming path from node \( i \) to node \( j \), and constraint (7) provides that for each farm \( j \), which is served by depot \( D \), there will be only one outgoing path from farm location \( j \) to location \( i \).
Finally, the Miller Tucker and Zemlin sub tour elimination constraint has been included (Miller et al., 1960), representing constraint (8). Note that if location i is followed by location j on the trip route, it is predefined that Q is set of nodes served by each depot.

The vehicle's capacity is ensured by constraint (9), where the amount of milk collected at location j after visiting location i must not exceed the tanker capacity of 9,000 gallons. Constraints set (10)-(14) are evident and self-explanatory.

5 Results and Discussion

In the mathematical model, the combined SCP and TSP constraints are defined. PuLP (version 2.7.0), a free, open-source library and Gurobi academic solver (version 10.0.3), has been used to solve this MILP model. An Intel(R) Xeon(R) CPU @ 2.20GHz with 12.68 GB RAM and T4 GPU is used to solve the model in the Python Colab environment. Four different scenarios have been investigated for this study to investigate the model’s applications. The first one is to find the optimal scenario with the exact algorithm. It has been investigated the number of required depots to be established and the routing path between them and farm locations. In the second scenario, the sensitivity analysis has been explored by changing the input parameters to evaluate the model response. It provides the focal point where the model is moderately sensitive to certain input parameters to identify and improve the model's reliability and robustness. In the third scenario, the heuristic algorithm of the nearest neighbor and steepest ascent hill climbing algorithm has been explored for the TSP part of the model. Lastly, the multi-vehicle analysis has been explored.

5.1 Scenario 1

To solve the exact algorithm, it is assumed that the assignment, or the establishment cost of each depot among the thirty-three firms, is constant and does not vary based on the location of the farms. The operating or the overhead cost of the depot, which encompasses the cost of labor, capital recovery of machinery and equipment, taxes and insurance, and general farm overhead, has been considered. The total allocated overhead cost for every 12 gallons of milk is considered $7.32 based on the data found by the U.S. Department of Agriculture (USDA ERS - Milk Cost of Production Estimates, n.d.). Each farm has several cows, which range from 10 to 100, so the daily milk production of the farms varies. For this study, 7.2 gallons of raw milk per day has been considered from each cow. The daily raw milk production of each farm has been described in the appendix, which is regarded as constant for this model. The SCP part of the model identified two
optimal locations of the depot, as shown in Figure 7. The blue dots indicate the depot's location, whereas the red dots on the map describe the farms' location. It is found that Hunter's Diary farm and Windy Acres Farm are the two geographically optimal locations to establish the depot among the thirty-three farms, which incur one million dollars of fixed cost as shown in Table 5.

It is apparent from Figure 7 that some farms are near to the one depot, although that depot does not cover it. For example, the River View Farms is closer to Hunter's Diary farms based on the Euclidean distance, although the Windy Acres Farm covers it. This is due to the time-based distance matrix that the Google Map API develops. While calculating the time required for traveling from one node to another node, factors like the interstate path proximity, traffic conditions, and how close the depot location is to the nearby street are also considered while calculating the travel time of the distance matrix from one place to another. Another problem arises when placing a depot in a densely populated node that can lead to redundancy. That is reflected here in the optimal number of depots and selecting their locations.

Figure 7: Optimal scenario of the depot location
The allocated overhead cost of depots, the variable cost of the SCP part, is $8,814. Among the 33 farms, as a depot location of the milk collection scenario, Hunter's Diary Farm and Windy Acres Farm covered 11 farms (nodes) and 22 farms (nodes), respectively. So, there will be two times the vehicle will be cleaned after completing the visit of each of the farms from the starting farms (depot).

The transportation cost per mile, fuel cost, tire cost, repair and maintenance cost, driver pay, and benefits have been considered while calculating the cost of TSP, as shown in Table 4. The maximum vehicle speed is assumed to be 55 miles/hr. The asymmetrical distance matrix is based on the hourly time needed for one node to another, which results in the per-hour transportation cost of $76.78.

<table>
<thead>
<tr>
<th>Cost type</th>
<th>Cost/mile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuel cost</td>
<td>$0.437</td>
</tr>
<tr>
<td>Tire costs</td>
<td>$0.047</td>
</tr>
<tr>
<td>Repair and maintenance cost</td>
<td>$0.103</td>
</tr>
<tr>
<td>Drivers pay and benefits</td>
<td>$0.809</td>
</tr>
<tr>
<td>Total variable cost</td>
<td>$1.396</td>
</tr>
<tr>
<td>Vehicle average speed per hour</td>
<td>55</td>
</tr>
<tr>
<td>per hour variable cost</td>
<td>76.78</td>
</tr>
</tbody>
</table>

The optimum routing path from the depot locations has been visualized in the Figure 8, where the gray node represents individual farms, and the green node represents the farms considered for the depot locations. For daily milk production, the optimal vehicle routing duration is found to be 28.28 hours.
After completing the cycle from the first depot location (node 9), it completed the second cycle from the second depot (node 14) and ended the route by returning to node 14 from node 10. The total cost of the optimal scenario has been described in Table 5.

Table 5: Total cost of the optimal scenario

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Depot Assignment cost</th>
<th>Overhead cost of depot</th>
<th>Transportation cost</th>
<th>Cleaning cost</th>
<th>Total cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal Scenario</td>
<td>$1,000,000</td>
<td>$8,814</td>
<td>$2,171</td>
<td>$300</td>
<td>$1,011,286</td>
</tr>
</tbody>
</table>

To find the optimal route, a total of 1,432,432 nodes have been explored with 11,659,085 simplex iterations in 1,079.3 seconds. The optimal path of the TSP route is 9-12-7-3-8-29-5-27-25-24-23-9-14-13-30-26-20-19-22-31-0-15-17-1-6-32-11-28-4-16-21-2-18-10-14.

5.2 Scenario 2

In the developed model, three major parameters can affect the solution's output in terms of cost and solution time. The first one is the depot assignment cost, and the second parameter is the overhead cost of the depot for the SCP part. The third parameter is the transportation cost, which is associated with the routing duration. In scenario 1, two of the farms' locations have been
identified as the optimal location where the depot assignment costs are constant regardless of the farm's geographical locations. The analysis is further conducted by randomly changing the assignment or the depot establishment cost within a range from half to one million to investigate whether the assignment cost affects the optimal solution. In the result, it is found that three different nodes are seen as a suggested optimal location to establish the depot, as shown in Figure 9 below, which is marked as blue color.

Figure 9: Sensitivity analysis of depot location

Node allocation is important for resource optimization. It is found that node 18 (Mountain View Dairy) covers nodes 1, 2, 4, 6, 10, 11, 16, 17, 18, 21, 25, 28, and 32. Node 9 (Windy Acres Farm) covers nodes 3, 5, 7, 8, 9, 12, 23, 24, and 29, and lastly, node 31 (Jamestown Dairy, LLC) covers nodes 0, 13, 14, 15, 19, 20, 22, 26, 27, and 30, 31. From this result, cost plays an important role in determining where to establish the depot, and the tradeoff between coverage and cost is significant. However, the relaxation of Maximal Coverage Distance (MCD) is also carried out to know how far the farms can be assigned from a depot location. It is apparent from Table 6 that if the value of
MCD increases the required number of depots also decreases. Interestingly, if the MCD is 3.5 or more, then there only needs one depot to cover all the farms’ locations.

Table 6: Variable hours in the MCD

<table>
<thead>
<tr>
<th>MCD (in hours)</th>
<th>Number of depots</th>
<th>Node numbers</th>
<th>Covered locations/nodes per depot (farms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>5</td>
<td>15, 19, 24, 27, 32</td>
<td>15 [10, 14, 15, 17, 25] 19 [13, 19, 20, 26, 30] 24 [3, 5, 7, 8, 9, 12, 23, 24, 29] 27 [0, 22, 27, 31] 32 [1, 2, 4, 6, 11, 16, 18, 21, 28, 32]</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>9, 27, 30, 32</td>
<td>9 [3, 5, 7-9, 12, 23, 29] 27 [0, 22, 25, 27, 31] 30 [10, 13-15, 19, 20, 24, 26, 30] 32 [1, 2, 4, 6, 11, 16, 17, 18, 21, 28, 32]</td>
</tr>
<tr>
<td>2.5</td>
<td>3</td>
<td>9, 25, 32</td>
<td>9 [3, 5, 7, 8, 9, 12, 23, 29] 25 [0, 6, 13-15, 17, 19, 20, 22, 24-27, 30, 31] 32 [1, 2, 4, 10, 11, 16, 18, 21, 28, 32]</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>9, 14</td>
<td>9 [3, 5, 7-9, 12, 23, 24, 25, 27, 29] 14 [0-2, 4, 6, 10, 11, 13-22, 26, 28, 30-32]</td>
</tr>
<tr>
<td>3.5</td>
<td>1</td>
<td>25</td>
<td>0-32 (all farms)</td>
</tr>
</tbody>
</table>
These results suggest that MCD affects the model's solution and cost. So, a vehicle with a higher refrigerating duration can reduce the cost as the depot establishment cost is prominent compared to other variable costs in the model. For the overhead cost of the depot, the different cost is considered based on various states along with the total scenario of the USA, as shown in Table 7.

Table 7: Different overhead cost of the depot based on the state.

<table>
<thead>
<tr>
<th>state</th>
<th>Total Allocated overhead cost per 12 gallons</th>
<th>Execution time (seconds)</th>
<th>Depot overhead Cost</th>
<th>Selected depot node number</th>
</tr>
</thead>
<tbody>
<tr>
<td>California</td>
<td>5.35</td>
<td>10</td>
<td>6,442</td>
<td>9, 14</td>
</tr>
<tr>
<td>Indiana</td>
<td>7.94</td>
<td>10</td>
<td>9,561</td>
<td>9, 14</td>
</tr>
<tr>
<td>Iowa</td>
<td>6.54</td>
<td>10</td>
<td>7,875</td>
<td>9, 14</td>
</tr>
<tr>
<td>Kentucky</td>
<td>9.74</td>
<td>10</td>
<td>11,728</td>
<td>9, 14</td>
</tr>
<tr>
<td>Michigan</td>
<td>10.44</td>
<td>10</td>
<td>12,571</td>
<td>9, 14</td>
</tr>
<tr>
<td>U.S. Total</td>
<td>7.32</td>
<td>10</td>
<td>8,814</td>
<td>9, 14</td>
</tr>
</tbody>
</table>

Although the depot overhead cost changed based on the different costs of the state, the optimal depot locations remained constant. It suggests that the depot overhead cost does not influence the optimal location of the depot. Furthermore, by changing the transportation cost per mile, the current study found that the optimal route (shortest path) remains the same in the TSP. Hence, the optimal route is solely dependent on the travel duration rather than the cost associated with traveling those distances.

5.3 Scenario 3

To evaluate the TSP part of the developed model, the result of the exact algorithm has been further compared with the steepest ascent Hill climbing algorithm, a mathematical optimization problem in Artificial Intelligence, and Nearest Neighbor heuristics. Due to the characteristics of Brute force, the number of Hamilton cycles is increased exponentially with the number of nodes. So, to find the shortest path within a shorter time, it is impossible for a large number of nodes with an exact algorithm rather than using the heuristics.
Nearest Neighbor algorithm is one of the first models to find a solution of the TSP. The nearest Neighbor algorithm does not need high computational time and hardware requirements. It can be applied to a large TSP problem with a shorter time frame. Although the Nearest Neighbor algorithm does not guarantee an optimal solution, it provides a good standing point to apply another optimization algorithm to reduce the search space. The flow chart of the Nearest Neighbor algorithm is shown in Figure 10.

![Flow chart of the Nearest Neighbor algorithm](image)

**Figure 10: Nearest neighbor algorithm**

In this method, the vehicle starts its journey from a random node and repeatedly visits the nearest node until all the nodes are visited. This heuristic method can easily find a short tour length, but that may only sometimes be the best path due to its greedy nature. The shortest path has been depicted in Figure 11, where the depots are denoted as green nodes. The duration of the route is 36.28 hours, and the total cost is $2,785 which is around 28.283% higher than the optimal cost. The shortest path is found 9-23-12-29-8-3-7-5-27-24-25-9-14-10-17-1-15-30-13-20- 26-19-31- 22-0-6-32-18-2-16-4-28-21-11-14.
The second heuristic approach used here is the steepest ascent Hill climbing algorithm, where its purpose is to find the best solution to a problem that has a large set of possible solutions. It is a local search optimizing algorithm that explores among the neighboring solutions and can converge faster to a local optimum with a moderate computational resource compared to other algorithms like the Genetic algorithm. Thus, for a small to medium-sized TSP problem, this method is well-suited to apply and is considered for this study as well. The steps of the steepest ascent Hill climbing algorithm have been described below.

1. Fixed the starting node as the depot node number

2. Generate a random initial solution

3. Determine the route duration of the initial solution

4. By swapping the nodes, generating a list of neighboring solutions

5. Find the best solution among the neighbors

6. Compare the route duration of the best neighboring solution with the initial solution

7. If the best neighboring solution is better, update the current solution

8. Steps 5-7 are repeated until no better neighbor is found

9. Within the time limit, run the above steps
a. Setting a time limit.

c. While the time limit is not exceeded.

   i. Run the steepest ascent hill climbing algorithm.

   ii. Determine route duration.

   iii. Upgrade the best route duration if a better solution is found.

10. Complete the algorithm

For the asymmetric distance matrix, the Hill climbing algorithm has been run for a predefined time frame repeatedly and choose the best outcome to overcome getting stuck in the local optima. The algorithm has been run for 7.5 seconds, and the route duration of each run of the algorithm concerning the time has been shown in Figure 12.

![Figure 12: Route durations of Steepest Ascent Hill climbing (SAHC) algorithm within a time frame.](image)
After a run time of 6.14s, the lowest route duration is found as 28.34 hours, resulting in $2,175.97 which is almost the optimal cost, only 0.215% higher than the optimal transportation cost. The route path is shown in Figure 13.

Figure 13: Routing path of steepest ascend Hill climbing.

Here, the minimal route sequence is found 9-7-3-29-5-27-25-24-23-12-9-14-13-30-26-20-19-22-31-0-15-17-1-6-32-11-28-4-16-21-2-18-10-14 by the Hill climbing algorithm. In comparison of speedup ratio, the steepest ascent Hill climbing is 175.78 times faster than the exact algorithm of the TSP.

6 Multi-vehicle Routing (Scenario 4)

In the capacity routing model for a single vehicle, the optimized route takes more than 28 hours to complete, which is not practically feasible. So, in this section, the MVR model is analyzed to overcome the shortcomings of previous scenarios. The available time for the reference vehicle (milk tanker), which can store the milk for a maximum of three hours without any system loss. To align more closely with the vehicle's system function, the MCD considers 2.5 hours, and three nodes (depots) are identified from the set covering part as described in the optimal location in Table 6. The assumptions that are considered for the transportation part of the model are considered here as below:

- All routes start and return to the single depot, where the depot index is identified by D.
- The distance between the two nodes, i and j, is constant.
• The average transportation cost will be constant on the planning horizon.

• All fleets are homogeneous and have the same capacity.

• Each farm is visited by a single vehicle once to collect the milk.

• The distance matrix only represents the transportation time, excluding any loading and unloading times for milk.

• The transportation cost is fixed and does not vary based on route or vehicle.

• Two vehicles are assigned from each depot location to complete the cycle.

Parameters

\( D \): Index of the depot

\( J \): Set of farms

\( K \): Set of vehicles

\( N \): Total number of nodes (depot and farms, \( D \subseteq J \))

\( k \): route index

\( d_{ij} \): Distance between location \( i \) and location \( j \) with \( d_{ii} = 0 \), and \( d_{ij} = d_{ji} \)

\( p \): Average transportation cost per unit of distance ($/gallon/miles)

\( M_i \): Milk production at node \( i \),

\( V \): Capacity of the vehicle \( k \)

Decision variables:

\( Z_{ijk} \in \{0,1\} \): 1 if location \( i \) is followed by location \( j \) on the trip route by vehicle \( k \); 0 otherwise.

\( U_{ik} \): Continuous variable representing the order of visiting node \( i \) by vehicle \( k \).

\[
\text{Min} \quad \sum_{k \in K} \sum_{i=0}^{N-1} \sum_{j=0, j \neq i}^{N-1} p d_{ij} Z_{ijk} \tag{1}
\]
The objective function minimizes the total transportation cost described in equation (1). Constraint (2) represents that each farm is visited by any vehicle once. Constraint (3) confirms that the vehicle capacity does not exceed the predefined value. Constraint (4)-(5) ensures that each vehicle enters and departs from the depot, and constraint (6) is the flow conservation at each farm/node location. Subtour elimination has been ensured by constraint (7) where vehicle index k, represents that the routing order is maintained separately for each vehicle. Constraints (8)-(9) are the non-negativity and initialization of sequence variables.

Two vehicles are assigned for each of the depots to complete the cycle. The total cycles of the vehicles are shown in the Table 8 and Figure 14 below:

<table>
<thead>
<tr>
<th>Depot/node</th>
<th>Vehicle 1 route cycle</th>
<th>Route duration (hrs)</th>
<th>Vehicle 2 route cycle</th>
<th>Route Duration (hrs)</th>
<th>Total route time (each depot) (hrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>9&gt;23&gt;9</td>
<td>0.08</td>
<td>9&gt;12&gt;7&gt;3&gt;8&gt;5&gt;9</td>
<td>3.07</td>
<td>3.15</td>
</tr>
</tbody>
</table>
In this scenario, the maximum route duration is found 11.38 hours, which is 59.76% more time efficient than the route’s duration of scenario 1 (TSP). This result shows that the TSP model in scenario 1 takes 2.48 times longer to complete than the multivehicle models in scenario 4. The transportation routes for each depot with two assigned vehicles have been shown in the Figure 14 below. The green nodes represent the depot location, and the terms v1 and v2 stand for vehicles 1 and 2.
Different scenarios of the transportation model with their computational time have been shown in the Table 9. The SAHC, TSP, and MVR with different combinations of vehicles, number of nodes (n) and depots are analyzed and observed. It is found that MVR model has much more advantage even based on the solution time as well.

Table 9: Computational time of different scenarios

<table>
<thead>
<tr>
<th>Instances</th>
<th>Number of vehicles</th>
<th>Number of depots</th>
<th>n</th>
<th>solution time (second)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSP-1</td>
<td>1</td>
<td>2</td>
<td>33</td>
<td>1,079.30</td>
</tr>
<tr>
<td>TSP-2</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>0.40</td>
</tr>
<tr>
<td>TSP-3</td>
<td>1</td>
<td>1</td>
<td>20</td>
<td>1,130.67</td>
</tr>
<tr>
<td>SAHC</td>
<td>1</td>
<td>1</td>
<td>33</td>
<td>6.14</td>
</tr>
<tr>
<td>MVR-1</td>
<td>2</td>
<td>1</td>
<td>8</td>
<td>0.14</td>
</tr>
<tr>
<td>MVR-2</td>
<td>2</td>
<td>1</td>
<td>10</td>
<td>0.37</td>
</tr>
<tr>
<td>MVR-3</td>
<td>2</td>
<td>1</td>
<td>15</td>
<td>3.23</td>
</tr>
<tr>
<td>MVR-4</td>
<td>2</td>
<td>1</td>
<td>20</td>
<td>181.54</td>
</tr>
<tr>
<td>MVR-5</td>
<td>3</td>
<td>1</td>
<td>8</td>
<td>0.04</td>
</tr>
<tr>
<td>MVR-6</td>
<td>3</td>
<td>1</td>
<td>10</td>
<td>0.44</td>
</tr>
<tr>
<td>MVR-7</td>
<td>3</td>
<td>1</td>
<td>15</td>
<td>4.79</td>
</tr>
<tr>
<td>MVR-8</td>
<td>3</td>
<td>1</td>
<td>20</td>
<td>307.96</td>
</tr>
</tbody>
</table>
7 Economic Feasibility Analysis

In the current scenario of milk distribution, the raw milk has been hauled from the firm's location to the nearby distribution center. United Dairy Processing Center is among them. The dairy processing center at United WV Charleston (38.36353624235745, -81.64604463815094), United OH Martins Ferry (40.098858351438, -80.7213569582406), and United PA Uniontown (39.9008173450056, -79.71656856918187) have been considered for this study. The daily milk production amount, considered constant, is 14,450.5 gallons. Equation 1 shows that the total transportation cost for the 33 firms' locations is $6,521.62, or $0.45 per gallon.

\[
\text{Per gallon transportation cost} = \frac{\text{Total transportation cost}}{\text{Milk production amount per day}} = \frac{6,521.62}{14,450.5} \approx 0.45 \quad (1)
\]

7.1 For scenario 1

After running the developed algorithm, Hunter's Dairy and Windy Acres have been identified as the best locations for the interim milk processing center/depot. With an optimized route, the transportation cost among the 33 locations is $2,171.3. Moreover, establishing the two depots will cost $1,000,000. With a depreciation rate of 3.636% (Publication 527 (2023), Residential Rental Property | Internal Revenue Service, n.d.), the annual depreciation is found to be $36,360 (depot establishment cost * annual depreciation rate), and the total depreciation cost for 10 years would be $363,600. So, the daily cost of the ownership of the depot is found as shown in Equation 2.

\[
\text{Daily cost of the depots} = \frac{\text{Total cost}}{\text{Days in 10 years}} = \frac{1,000,000 + 3 \times 6500}{3 \times 365} \approx 373.58 \quad (2)
\]

According to the Equation 1, it gives the per gallon transportation cost is $0.176, a 61.80% cost reduction in the milk transportation per gallon of raw milk from the firm’s location to the destination compared to the existing method of transportation. Furthermore, with a developed algorithm each day, it will save $3,976.82 for the constant amount of raw milk transportation. The break-even point of the depot cost is found in only 342.9 days or 0.94 years, which validates the feasibility of establishing the two depots in the assigned firm location.

7.2 For scenario 4

In the scenario 4, the developed multivehicle routing transportation model, the transportation cost results in $2,132.18 with three depots as the optimal locations. So, with three depots the daily cost of depots including the depreciation cost will be $560.38 as shown in the equation 3. According to the equation (1), the per gallon transportation cost is $0.186, and the savings of each day will
be $3,829.06 with the constant amount of milk production. Again, the break-even point for the depot cost is 1.35 years within a 10 years’ time frame which shows the feasibility to establish even the three depots with multivehicle routing model. The comparative economic scenario has been shown in the table 10 below.

\[
\text{Daily cost of the depots} = \frac{\text{Total cost}}{\text{Days in 10 years}} = \frac{1.5 \times 0.0000 + 5 \times 0.00}{3 \times 365.0} \approx $560.38 \quad (3)
\]

<table>
<thead>
<tr>
<th>Number of depots</th>
<th>Cost/gallon milk</th>
<th>Transportation cost/daily</th>
<th>Daily savings</th>
<th>Break-even point of the depots</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current transportation scenario</td>
<td>2</td>
<td>$0.45</td>
<td>$6,521.62</td>
<td>-</td>
</tr>
<tr>
<td>Scenario 1</td>
<td>2</td>
<td>$0.18</td>
<td>$2,544.88</td>
<td>$3,976.82</td>
</tr>
<tr>
<td>Scenario 4</td>
<td>3</td>
<td>$0.19</td>
<td>$2,692.56</td>
<td>$3,829.06</td>
</tr>
</tbody>
</table>

**8 Conclusion, limitation, and future works**

There are several limitations of this initial work, which pave the way for future research works. First, the developed distance matrix is only based on the driving mode and represents the time from one place to another when running the code. It considers the real-time scenario of the road, which might differ from time to time due to the traffic situation and different times of the day. Furthermore, the driving mode considered in the code does not reflect the raw milk transportation truck with a conventional tanker.

Second, the milk production capacity of each farm is considered constant on a daily basis. The milk production varies depending on the season, weather conditions, and feed availability. The milk quality with somatic cell count can be further investigated as a constraint in the model.

Third, an arguable weakness of this study is that the model is considered for the single depot multiple vehicle routing problem. But some of the farm’s location are on remote places where the larger types of fleets do not have the accessibility to collect the milk. So, it can be improved through multi-depot and heterogeneous fleet constraints to align more with the real-world problem.
Fourth, the Nearest neighbor and Hill climbing algorithm is limited to the larger data set, which may result in local optima, depending on the initial solution. Furthermore, the Hill climbing algorithm does not remember previously visited solutions. Advanced meta-heuristics algorithms like parallel metaheuristics, simulated annealing, partial Swarm optimization, and so on can be utilized for better efficiency with larger instances.

Lastly, additional real-life situations to obtain mechanisms consistent with larger and different nodes need to be explored where the historical optimization results can be applied in various machine learning algorithms.

This research intends to investigate the SCP and TSP, two classical problems in operation research. The combination of SCP and TSP constraints is studied in a Mixed Integer Linear problem. This study explored the exact algorithms that are dedicated to solving the SCP and TSP separately but are aimed at minimizing the overall cost. The development of a time-dependent distance matrix with Google Map API and a feasible solution for the model has been analyzed. We then present a detailed case study of the milk collection problem in West Virginia, USA. The sensitivity analysis is carried out to find out the parameters that significantly affect the models’ outcomes. It is found that the model is sensitive to the depot assignment cost and the MCD per hour parameters, which affects the objective function of the model. MVR scenario shows the potential to apply the model compared to the other scenario (TSP-based), still, there is a door to future research to incorporate the heterogeneous fleet and multi-depot constraints in the developed model. The Nearest Neighbor and Steepest ascend algorithms have been further studied for the TSP part of the model to have a better solution in terms of resource efficiency scalability and compared with the optimal solution. In terms of computational performance observed, the Hill climbing algorithm gives a promising result without sacrificing significant quality, although the model requires further development with more real-life constraints and machine learning approaches.

References


### Appendix A

Farms and their locations

<table>
<thead>
<tr>
<th>Node numbers</th>
<th>Farm names</th>
<th>Latitude and longitude</th>
<th>Daily milk production capacity (gallons)</th>
<th>Random assignment cost ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>J Lynn Farm</td>
<td>38.8707976522668,-81.3611466256723</td>
<td>662.4</td>
<td>3,943,729</td>
</tr>
<tr>
<td>1</td>
<td>Bachtel Dairy</td>
<td>39.2839467923414,-79.4904738826689</td>
<td>324</td>
<td>7,496,167</td>
</tr>
<tr>
<td>2</td>
<td>Shepherd's Whey Creamery</td>
<td>39.4617095306957,-79.295628654672</td>
<td>475.2</td>
<td>2,371,532</td>
</tr>
<tr>
<td>3</td>
<td>William Beiler Dairy</td>
<td>37.562346025191,-80.4081180283161</td>
<td>684</td>
<td>4,673,945</td>
</tr>
<tr>
<td>4</td>
<td>James T. Blue &amp; Sons Inc.</td>
<td>39.3836461513556,-77.8566479878076</td>
<td>568.8</td>
<td>1,663,303</td>
</tr>
<tr>
<td>5</td>
<td>Walnut Ridge Dairy Inc.</td>
<td>37.4444020254517,-80.7005196594359</td>
<td>331.2</td>
<td>7,576,652</td>
</tr>
<tr>
<td>6</td>
<td>Brookedale Holsteins</td>
<td>39.4842250310171,-78.8012613566908</td>
<td>640.8</td>
<td>2,354,400</td>
</tr>
<tr>
<td>7</td>
<td>Windspring Farms, Inc.</td>
<td>37.5884034680266,-80.3125456575795</td>
<td>468</td>
<td>9,774,217</td>
</tr>
<tr>
<td>8</td>
<td>Ecokee Holsteins</td>
<td>37.5872407742899,-80.4862153152518</td>
<td>144</td>
<td>5,959,038</td>
</tr>
<tr>
<td>9</td>
<td>Windy Acres Farm</td>
<td>37.896082614029,-80.359760730589</td>
<td>237.6</td>
<td>996,211</td>
</tr>
<tr>
<td>10</td>
<td>Mason Run Farm</td>
<td>39.7017883793537,-79.5749519152003</td>
<td>698.4</td>
<td>9,245,151</td>
</tr>
<tr>
<td>11</td>
<td>Gruber Farms</td>
<td>39.2654285729122,-77.9816654845221</td>
<td>504</td>
<td>9,213,011</td>
</tr>
<tr>
<td>12</td>
<td>Lotus Hill farm</td>
<td>37.9480615379475,-80.427567287349</td>
<td>316.8</td>
<td>2,234,075</td>
</tr>
<tr>
<td>13</td>
<td>Jon-Davy Farm</td>
<td>40.0650401230395,-80.5484139575189</td>
<td>424.8</td>
<td>4,962,441</td>
</tr>
<tr>
<td>14</td>
<td>Hunter's Diary Farm</td>
<td>39.7017493339065,-79.9142757998558</td>
<td>475.2</td>
<td>6,836,789</td>
</tr>
<tr>
<td>15</td>
<td>J D Farms LLC</td>
<td>39.299536689256,-79.947499057538</td>
<td>223.2</td>
<td>3,728,249</td>
</tr>
<tr>
<td>16</td>
<td>Maple Lawn Jersey Farm</td>
<td>39.4441794530105,-77.8376026710262</td>
<td>187.2</td>
<td>7,525,952</td>
</tr>
<tr>
<td>17</td>
<td>Lipscomb's Dairy</td>
<td>39.3013252258039,-79.5190911816045</td>
<td>180</td>
<td>9,816,325</td>
</tr>
<tr>
<td></td>
<td>Dairy Name</td>
<td>Latitude</td>
<td>Longitude</td>
<td>Acres</td>
</tr>
<tr>
<td>---</td>
<td>--------------------------------</td>
<td>------------------------------</td>
<td>-------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>18</td>
<td>Mountain View Dairy</td>
<td>39.4945433000993, -78.0053239728777</td>
<td>100.8</td>
<td>519,947</td>
</tr>
<tr>
<td>19</td>
<td>Miller View Farms, LLC</td>
<td>39.8469986849689, -80.7550929481203</td>
<td>655.2</td>
<td>6,926,792</td>
</tr>
<tr>
<td>20</td>
<td>Minch Dairy</td>
<td>40.0819002133266, -80.649783358822</td>
<td>158.4</td>
<td>8,042,581</td>
</tr>
<tr>
<td>21</td>
<td>Rolling Acres Farm</td>
<td>39.5206068270233, -77.8822591728771</td>
<td>712.8</td>
<td>5,746,953</td>
</tr>
<tr>
<td>22</td>
<td>Edwin S. Peachy Farm</td>
<td>38.9560268225127, -82.0198310017271</td>
<td>511.2</td>
<td>8,310,249</td>
</tr>
<tr>
<td>23</td>
<td>Perk Farm Organic Dairy</td>
<td>37.9051474449985, -80.3684070864081</td>
<td>144</td>
<td>9,302,598</td>
</tr>
<tr>
<td>24</td>
<td>River View Farms</td>
<td>39.2351665693312, -80.3894410305564</td>
<td>475.2</td>
<td>1,000,000</td>
</tr>
<tr>
<td>25</td>
<td>MR Dairy Farm</td>
<td>39.1822332794272, -80.2947800017216</td>
<td>489.6</td>
<td>3,661,625</td>
</tr>
<tr>
<td>26</td>
<td>Rip-Vale Farm</td>
<td>40.2201245948704, -80.5830480863511</td>
<td>460.8</td>
<td>7,876,115</td>
</tr>
<tr>
<td>27</td>
<td>Bridgewater Farm</td>
<td>38.8949732790045, -81.6930858017286</td>
<td>619.2</td>
<td>9,631,767</td>
</tr>
<tr>
<td>28</td>
<td>Locust Grove Farm</td>
<td>39.4054431198641, -77.8851287152076</td>
<td>640.8</td>
<td>6,069,364</td>
</tr>
<tr>
<td>29</td>
<td>Belle Vue Dairy, LLC</td>
<td>37.5894757796944, -80.5301389152518</td>
<td>201.6</td>
<td>2,673,477</td>
</tr>
<tr>
<td>30</td>
<td>Winter-Ayr Farm</td>
<td>40.1109591116835, -80.5466806286817</td>
<td>460.8</td>
<td>3,015,684</td>
</tr>
<tr>
<td>31</td>
<td>Jamestowne Dairy, LLC</td>
<td>38.8839086340813, -81.9959737440566</td>
<td>648</td>
<td>1,780,012</td>
</tr>
<tr>
<td>32</td>
<td>Triple E Farm</td>
<td>39.4111185901844, -78.629710935521</td>
<td>626.4</td>
<td>4,955,138</td>
</tr>
</tbody>
</table>
### Appendix B

#### Distance matrix (Based on hrs)

<table>
<thead>
<tr>
<th>Farm</th>
<th>Brother Dairy</th>
<th>Rip-Vale Farm</th>
<th>MR Dairy Farm</th>
<th>Edwin S. Peachy Farm</th>
<th>Rolling Acres Farm</th>
<th>Hunter's Diary Farm</th>
<th>Gruber Farms</th>
<th>Windy Acres Farm</th>
<th>William Beiler Dairy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brother Dairy</td>
<td>2.87</td>
<td>3.16</td>
<td>4.86</td>
<td>2.42</td>
<td>3.55</td>
<td>5.26</td>
<td>4.86</td>
<td>3.16</td>
<td>4.86</td>
</tr>
<tr>
<td>Rip-Vale Farm</td>
<td>3.16</td>
<td>2.87</td>
<td>3.16</td>
<td>3.55</td>
<td>5.26</td>
<td>2.87</td>
<td>3.16</td>
<td>4.86</td>
<td>2.87</td>
</tr>
<tr>
<td>MR Dairy Farm</td>
<td>4.86</td>
<td>2.87</td>
<td>3.16</td>
<td>5.26</td>
<td>2.87</td>
<td>3.16</td>
<td>3.55</td>
<td>5.26</td>
<td>3.16</td>
</tr>
<tr>
<td>Edwin S. Peachy Farm</td>
<td>2.42</td>
<td>3.55</td>
<td>5.26</td>
<td>2.87</td>
<td>3.16</td>
<td>3.55</td>
<td>5.26</td>
<td>2.87</td>
<td>3.16</td>
</tr>
<tr>
<td>Rolling Acres Farm</td>
<td>3.55</td>
<td>5.26</td>
<td>2.87</td>
<td>3.16</td>
<td>3.55</td>
<td>5.26</td>
<td>2.87</td>
<td>3.16</td>
<td>5.26</td>
</tr>
<tr>
<td>Hunter's Diary Farm</td>
<td>2.87</td>
<td>3.16</td>
<td>4.86</td>
<td>2.42</td>
<td>3.55</td>
<td>5.26</td>
<td>2.87</td>
<td>3.16</td>
<td>2.42</td>
</tr>
<tr>
<td>Gruber Farms</td>
<td>3.16</td>
<td>4.86</td>
<td>2.87</td>
<td>3.16</td>
<td>5.26</td>
<td>2.87</td>
<td>3.16</td>
<td>3.55</td>
<td>5.26</td>
</tr>
<tr>
<td>Windy Acres Farm</td>
<td>4.86</td>
<td>3.16</td>
<td>2.87</td>
<td>3.16</td>
<td>3.55</td>
<td>5.26</td>
<td>2.87</td>
<td>3.16</td>
<td>2.87</td>
</tr>
<tr>
<td>William Beiler Dairy</td>
<td>2.42</td>
<td>3.55</td>
<td>5.26</td>
<td>2.87</td>
<td>3.16</td>
<td>3.55</td>
<td>5.26</td>
<td>2.87</td>
<td>3.16</td>
</tr>
</tbody>
</table>

#### Additional Farms

<table>
<thead>
<tr>
<th>Farm</th>
<th>1.29</th>
<th>0.18</th>
<th>3.09</th>
<th>3.66</th>
<th>0.00</th>
<th>5.05</th>
<th>3.53</th>
<th>2.26</th>
<th>1.17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Holsteins</td>
<td>3.09</td>
<td>3.66</td>
<td>0.00</td>
<td>3.53</td>
<td>2.26</td>
<td>1.17</td>
<td>5.05</td>
<td>3.53</td>
<td>2.26</td>
</tr>
<tr>
<td>Lotus Hill</td>
<td>1.40</td>
<td>3.05</td>
<td>3.74</td>
<td>3.74</td>
<td>0.00</td>
<td>3.85</td>
<td>3.55</td>
<td>2.43</td>
<td>1.59</td>
</tr>
<tr>
<td>J D Farms</td>
<td>3.05</td>
<td>3.74</td>
<td>3.74</td>
<td>3.85</td>
<td>0.00</td>
<td>3.55</td>
<td>2.43</td>
<td>1.59</td>
<td>2.59</td>
</tr>
<tr>
<td>Maple Lawn</td>
<td>1.49</td>
<td>0.00</td>
<td>1.49</td>
<td>1.49</td>
<td>0.00</td>
<td>3.85</td>
<td>3.55</td>
<td>2.43</td>
<td>1.59</td>
</tr>
<tr>
<td>Nettles Dairy</td>
<td>0.00</td>
<td>3.85</td>
<td>3.85</td>
<td>0.00</td>
<td>3.55</td>
<td>2.43</td>
<td>1.59</td>
<td>2.59</td>
<td>1.68</td>
</tr>
<tr>
<td>Rock Rest</td>
<td>3.85</td>
<td>3.55</td>
<td>3.55</td>
<td>2.43</td>
<td>1.59</td>
<td>2.59</td>
<td>1.68</td>
<td>3.85</td>
<td>0.00</td>
</tr>
<tr>
<td>Tobacco Farms</td>
<td>3.55</td>
<td>2.43</td>
<td>2.43</td>
<td>1.59</td>
<td>2.59</td>
<td>1.68</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
</tr>
<tr>
<td>Brookside Farms</td>
<td>1.68</td>
<td>1.49</td>
<td>1.49</td>
<td>1.49</td>
<td>3.85</td>
<td>3.55</td>
<td>2.43</td>
<td>1.59</td>
<td>2.59</td>
</tr>
<tr>
<td>J H Farms</td>
<td>1.68</td>
<td>3.85</td>
<td>3.55</td>
<td>2.43</td>
<td>1.59</td>
<td>2.59</td>
<td>1.49</td>
<td>1.49</td>
<td>3.85</td>
</tr>
<tr>
<td>Locust Grove</td>
<td>1.49</td>
<td>1.49</td>
<td>1.49</td>
<td>1.49</td>
<td>3.85</td>
<td>3.55</td>
<td>2.43</td>
<td>1.59</td>
<td>2.59</td>
</tr>
<tr>
<td>J D Farms</td>
<td>3.85</td>
<td>3.55</td>
<td>3.55</td>
<td>2.43</td>
<td>1.59</td>
<td>2.59</td>
<td>1.49</td>
<td>1.49</td>
<td>3.85</td>
</tr>
<tr>
<td>Maple Lawn</td>
<td>3.55</td>
<td>2.43</td>
<td>2.43</td>
<td>1.59</td>
<td>2.59</td>
<td>1.68</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
</tr>
<tr>
<td>Rock Rest</td>
<td>2.43</td>
<td>1.59</td>
<td>1.59</td>
<td>2.59</td>
<td>1.68</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
</tr>
<tr>
<td>Tobacco Farms</td>
<td>1.59</td>
<td>2.59</td>
<td>2.59</td>
<td>1.68</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
</tr>
<tr>
<td>Brookside Farms</td>
<td>2.59</td>
<td>1.68</td>
<td>1.68</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
</tr>
<tr>
<td>J H Farms</td>
<td>3.85</td>
<td>0.00</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
</tr>
<tr>
<td>Locust Grove</td>
<td>0.00</td>
<td>3.85</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
</tr>
<tr>
<td>Tobacco Farms</td>
<td>3.85</td>
<td>0.00</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
</tr>
<tr>
<td>J H Farms</td>
<td>0.00</td>
<td>3.85</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
</tr>
<tr>
<td>Tobacco Farms</td>
<td>3.85</td>
<td>0.00</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
</tr>
<tr>
<td>R J Farms</td>
<td>0.00</td>
<td>3.85</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
<td>0.00</td>
<td>3.85</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Farm</th>
<th>3.97</th>
<th>4.86</th>
<th>4.98</th>
<th>4.98</th>
<th>2.55</th>
<th>5.57</th>
<th>5.57</th>
<th>2.55</th>
<th>5.57</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rip-Vale Farm</td>
<td>4.86</td>
<td>3.97</td>
<td>4.98</td>
<td>4.98</td>
<td>2.55</td>
<td>5.57</td>
<td>5.57</td>
<td>2.55</td>
<td>5.57</td>
</tr>
<tr>
<td>Hallow Farm</td>
<td>4.98</td>
<td>4.98</td>
<td>3.97</td>
<td>4.98</td>
<td>2.55</td>
<td>5.57</td>
<td>5.57</td>
<td>2.55</td>
<td>5.57</td>
</tr>
<tr>
<td>Lion's Dairy</td>
<td>2.55</td>
<td>2.55</td>
<td>2.55</td>
<td>2.55</td>
<td>3.97</td>
<td>4.98</td>
<td>4.98</td>
<td>3.97</td>
<td>4.98</td>
</tr>
<tr>
<td>Rip-Vale Farm</td>
<td>5.57</td>
<td>5.57</td>
<td>5.57</td>
<td>5.57</td>
<td>2.55</td>
<td>4.98</td>
<td>4.98</td>
<td>2.55</td>
<td>4.98</td>
</tr>
<tr>
<td>Hallow Farm</td>
<td>2.55</td>
<td>2.55</td>
<td>2.55</td>
<td>2.55</td>
<td>3.97</td>
<td>4.98</td>
<td>4.98</td>
<td>3.97</td>
<td>4.98</td>
</tr>
<tr>
<td>Lion's Dairy</td>
<td>4.98</td>
<td>4.98</td>
<td>3.97</td>
<td>3.97</td>
<td>2.55</td>
<td>5.57</td>
<td>5.57</td>
<td>2.55</td>
<td>5.57</td>
</tr>
</tbody>
</table>

The distances are calculated based on the linear distance between the farms. The values in the matrix represent the time it takes to travel from one farm to another, assuming no change in travel time. The distances are expressed in hours (hrs).