Incorporating Traffic Patterns to Improve Delivery Performance

by

Melody J. Dickinson
B.A. International Relations, University of California, Davis, 2004

and

Jillian Leifer
M.S. Materials Science and Engineering, University of Florida, 2004
B.S. Materials Science and Engineering, University of Pennsylvania, 2002

Submitted to the Engineering Systems Division in Partial Fulfillment of the Requirements for the Degree of

Master of Engineering in Logistics

at the

Massachusetts Institute of Technology

June 2010

Signature of Authors...
Master of Engineering in Logistics Program, Engineering Systems Division
May 7, 2010

Certified by...
Dr. Jarrod Goentzel
Executive Director, Masters of Engineering in Logistics Program
Thesis Supervisor

Accepted by...
Prof. Yossi Sheffi
Professor, Engineering Systems Division
Professor, Civil and Environmental Engineering Department
Director, Center for Transportation and Logistics
Director, Engineering Systems Division

Incorporating Traffic Patterns to Improve Delivery Performance

by

Melody J. Dickinson and Jillian Leifer

Submitted to the Engineering Systems Division on May 7, 2010
in Partial Fulfillment of the Requirements for the Degree of
Master of Engineering in Logistics

ABSTRACT

Traffic, construction and other road hazards impact the on-time performance of companies that operate delivery fleets. This study examines how incorporating traffic patterns in vehicle route development compares with standard, deterministic methods. We seek to understand how using historical data improves both planning and overall delivery efficiency.

Our analysis contrasts manifests that were developed by an industry standard routing software tool with projections that use traffic data by benchmarking them against actual routes run by drivers. In addition to evaluating the differences between route planning tools, we explore why those differences exist, including how uncertainty is incorporated.

Evidence suggests that incorporating traffic patterns into vehicle routing does produce improved solutions. Needless to say, the delivery process needs to be evaluated holistically. Our recommendations involve the various steps for creating and executing a route. Operational considerations, the potential for improving customer service, and areas for further exploration are discussed.

This thesis is being conducted with sponsorship from a leading consumer products company and in coordination with the CarTel mobile sensing data project at Massachusetts Institute of Technology (MIT).

Thesis Supervisor: Dr. Jarrod Goentzel
Title: Executive Director, Master of Engineering in Logistics Program

ACKNOWLEDGEMENTS

We would like to thank our sponsor for their dedication, enthusiasm and constant support throughout the project. Their commitment was instrumental to our success.

Additionally, we would like to thank our thesis advisor, our friends and our families for their constant support in our academic endeavors.

Lastly, we want to thank each other for our teamwork, support and ability to make each other laugh during times of stress.

Melody and Jillian

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

1 INTRODUCTION

Traffic, construction and other road hazards impact the on-time performance of companies that operate delivery fleets. These factors are frequently treated as incidents beyond anyone's control. However, with some events, such as rush hour, being correlated to time of day, and others, such as construction, being seasonal, by collecting data on traffic patterns, potential exists for incorporating this information into delivery operations. One would anticipate that using real traffic data would improve route planning, as well as overall delivery efficiency.

Everyday the beverage distribution company, BevCo,[1] fulfills its customers' orders by delivering its products from its warehouses to retail outlets. This transaction occurs downstream in the company's supply chain and is part of a customer service component that begins with sales representatives taking orders the day prior.

Indeed, ordering sets the delivery process in motion. As orders are taken, they are forwarded and collected in a database. At a cutoff point in the late afternoon, the day's orders are transmitted to the company's vehicle routing software. By using this sales data, the software formulates route manifests that provide drivers with the information needed to make their deliveries.

BevCo uses an off-the-shelf optimization tool, OptiRoute[1], which is the industry standard in vehicle routing systems. OptiRoute employs crude approximations of average speeds on road segments using a combination of speed limits and type of road (interstate, state highway, etc.) to estimate the travel time from an origin to a destination. Instead of accepting the deterministic value for each road segment as embedded within the current vehicle routing system, we seek to understand how historical traffic data could be used to plan routes.

[1] Some of the names and data presented in this thesis have been disguised.

The CarTel group within the Computer Science and Artificial Intelligence Laboratory (CSAIL) at the Massachusetts Institute of Technology (MIT) has been compiling measurements of travel times in coordination with a private livery[2] fleet in the Greater Boston Metropolitan area for the past three years. The data has been carved into hourly time buckets from which the mean travel time for a given road segment can be calculated. This enables route planning that accounts for historical traffic patterns regardless of cause. At this level, solely planning in advance—for BevCo, this would be the day before routes are executed—is being examined.

Our research will evaluate the accuracy of OptiRoute against CarTel using actual routes as driven as a benchmark. After the initial analysis by segment and cumulative route is complete, follow-on work to examine how uncertainty is incorporated into the routing process and how this may be mitigated is outlined.

The introduction explains the motivation for seeking to improve the accuracy of routing solutions, summarizes current practices and finally considers alternative approaches to route planning and operations. A review of relevant literature pertaining to the Vehicle Routing Problem (VRP) and network effects follows. After the background has been presented, our research methodology and corresponding data analysis are presented. We conclude with our recommendations for how BevCo can improve delivery performance.

1.1 MOTIVATION

Many of BevCo's customers have time windows when they accept deliveries, and on-time delivery is a priority with numerous service implications. If a truck is late, it must wait at the end of a queue of other trucks, or if the window is missed completely, the truck may have to return on an emergency route another day. For some customers, drivers are also responsible for merchandising, which entails stocking the shelves and ensuring the customer is satisfied with service. If a driver is running late, merchandising suffers, as does the customer relationship.

[2] A for-hire vehicle, similar to a taxi but requiring a reservation.

This may have long-term consequences. Unsatisfied customers may switch to using a competitor, resulting in lost sales. Therefore, ensuring that delivery is on-time is a pressing operational aim for BevCo.

We answer three questions concerning the potential for improvements in both the delivery process, and overall employee efficiency:

- How significantly do traffic and other driving-related obstacles impact the delivery process?
- Can using travel time data create more accurate route plans, thus reducing the amount of stress that frontline staff (drivers, dispatchers, sales representatives and others) have in their daily jobs?
- By using more accurate data, can uncertainty be better mitigated throughout the delivery process?

1.2 COMPANY PROFILE

BevCo is the beverages unit of an international consumer products company. Its product assortment includes soft drinks, bottled water, iced teas, juices and sports and energy drinks. BevCo manufactures and distributes its own products. It also distributes other companies' branded products on a contract basis.

1.3 LOCATION

This thesis focuses on the Greater Boston Metropolitan area, primarily North Boston and Cambridge, Massachusetts. Not only does BevCo have a warehouse that serves this specific geographic region, but CarTel has collected considerable data on travel times for the locale.

1.4 CURRENT DISTRIBUTION PRACTICES

At present, BevCo relies on the combination of vehicle routing software and the expertise of its workforce to ensure efficiency throughout the delivery process. The company employs different operational configurations, which vary by region. The practices for Boston are detailed below; the Specialized Dispatch Operations (SDO) described in Section 1.8 represents an alternative outside of New England. To best serve its diverse customer base, BevCo deploys two types of trucks: bay and bulk. Our research focuses on bay trucks, as the routes served by bay trucks are characterized by more stops (and therefore customers) per route and fewer average cases per customer. This translates into more complex route plans, and BevCo anticipates a greater benefit from improved routing for this type of truck as opposed to a bulk truck.

1.4.1 Direct Store Delivery

BevCo employs Direct Store Delivery (DSD), a process in which a distributor bypasses its customers' warehouses. Instead of transferring product from its warehouse to a customer's warehouse, the product is delivered directly to a customer's retail store. Many times, the driver stocks the product on the shelves or a refrigerated display. In addition to merchandising, with DSD, drivers also assume responsibility for collecting payment, reverse logistics of out-of date product and customer service.

1.4.2 Bay Trucks

Bay trucks employ a unique loading process called "build-to-load." Orders for all customers in a route are aggregated and like products are placed into bays. At the time of delivery, the drivers disaggregate product from each bay to assemble customer orders on site. Multiple sizes of bay trucks are used in the field, and they range from 10 to 20 bays.

1.4.3 Vehicle Route Planning

The routing process begins with BevCo's sales force, which interfaces directly with customers. Orders input by representatives are transmitted to a backend system where they are collected and batched. Information collected includes customer name, location, order size (BevCo uses units of cases or a case approximation) and any delivery requirements. The cutoff time is 3:00 pm; after that new orders are collected for the next business day. The system produces an order summary, which is reviewed by an employee and pulled into OptiRoute.

Next, the software system then projects route plans. Around 4:00 pm, the North Boston dispatcher receives the output from OptiRoute and begins localized route planning. The dispatcher reviews the routes, including potential conflicts that the software has flagged (exceptions are explained in more detail in Section 1.5.3). The dispatcher has the authority to modify routes to eliminate exceptions and incorporate any last minute changes.

After routes are finalized, the batched order and route plans are sent to a truck mapping specialist, whose role is to determine how the product is allocated across the bays. High volume products have been assigned to a standard compartment for ease of driver recollection. Compartments that hold common sizes of product (e.g. 2 liters or 20 ounce glass bottles) have also been designated. This process has added complexity due to the different sizes and shapes of each product, as each bay can hold different quantities of each sized product. Once the product allocation for each bay is mapped, the product is loaded onto the truck. The trucks are then ready for the next day's deliveries.

1.4.4 Operational Priorities

BevCo strives to reduce travel time for three reasons. First, less time spent driving translates to increased time available for servicing customers. Second, by decreasing the amount of time spent in transit, more deliveries can be made in one day. If driving time is reduced below a certain threshold, BevCo can reduce the number of routes its runs. Finally, if driving time can be estimated with more certainty, calculating labor hours will be more accurate and overtime more predictable and manageable within legal constraints.

The delineation between time spent traveling (non-value added time) and time spent delivering or servicing a customer (valuable time) provides insight into BevCo's operational priorities. In effect, minimizing time has a higher priority than minimizing distance in the objective function. Ideally, reducing driving time will realize significant cost savings for BevCo.

Although the creation of routes by the software system may not account for instances that are localized or change daily such as traffic, holidays or special events, the system is undoubtedly more effective than an individual creating the routes from scratch each day. Accordingly, individuals making judgment calls about stop order or the road network traversed are more prone to make errors under stress, and providing the dispatcher and drivers with the flexibility to change paths makes troubleshooting of local phenomena more effective. While BevCo recognizes the margin for improvement of its current system, it works.

1.5 OPTIROUTE

OptiRoute is a commercial off-the-shelf vehicle routing software. Known as an industry-standard and widely used by companies managing delivery fleets, OptiRoute is typically configured to minimize time by minimizing distance. Routes are made up of two components: drive time[3] and stop time[4]. Delivery constraints, such as time windows, can also be entered into the system and accounted for in route plan development.

1.5.1 Underlying Road Network and Drive Time Model

Drive times are calculated by route segment[5] using an underlying road network provided by Navteq, one of two leading companies that specialize in mapping and navigation services. Embedded data includes road segment[6] distances, average speeds, turn restrictions, one-way

[3] The time that it takes a driver to travel from an origin to a destination.
[4] The time that a driver spends making a delivery or providing service to a customer.
[5] The distance between two customers, this consists of an origin and a destination.
[6] The distance traveled on a single road with a uniform travel speed.

streets and restricted access road heights[7]. Navteq issues quarterly updates and releases a new version of its map network every two years. While updates are available quarterly, BevCo updates their road database every two years due to the difficulty and time required. In addition, BevCo can modify the maps to add constraints as needed.

OptiRoute has fixed and variable components in its drive time model. A fixed drive time is assigned to each route segment with a distance greater than zero plus a variable component based on distance and road speeds. BevCo can adjust the fixed time.

1.5.2 Stop Time Model

OptiRoute uses criteria input by BevCo to model the stop time to calculate the amount of time that a driver will spend providing service to a customer. There are two components of the stop time. The first is a fixed amount of time that varies by customer type. Different types of customers (e.g. retail, restaurants and offices) have different types of merchandising, receiving and customer service requirements. The fixed time incorporates these requirements. The second component is a variable rate that corresponds to the number of case equivalents being delivered. The variable rate accounts for the amount of time to pick and deliver the order. Criteria such as customer channel may also be incorporated into the variable rate.

Table 1 provides an example of estimated stop times for four distinct customer types. For ABC Pharmacy, drivers may be contractually required to merchandise the product in the store, which requires a longer variable time per case; whereas, at Becker Accounting, drivers may be asked to simply leave the product in a cafeteria storeroom.

---

[7] A detailed description of the maps and underlying data supplied by Navteq is available on the company's website, http://www.navteq.com/.

TABLE 1: EXAMPLE OF STOP TIME PROJECTIONS BY CUSTOMER

Customer	Number of Cases	Fixed Time (min)	Variable Time (min)	Est. Stop Time (min)
ABC Pharmacy	92	14	.5	64.5
Ray's Steakhouse	45	10	.3	23.5
MU Convenience	70	12	.3	33
Becker Accounting	22	10	.2	14.4

Modifications to the stop time model can be input at the customer or order level. They are typically made at the customer level.

1.5.3 OptiRoute Constraints and BevCo Guidelines

To ensure routes are feasible and effective, OptiRoute incorporates constraints into the routing process[8]. In addition, BevCo has some guidelines the dispatchers use to finalize plans. These include the following:

- **Truck Capacity**: The dispatcher routes for both bulk and bay trucks; all have different case equivalent capacities.
- **Time Windows**: Some customers only accept deliveries during certain periods of the day.
- **Number of Routes and Case Load**: To optimize use of resources and ensure feasibility of routes, the company sets a target number of cases per truck. By modifying assigned stops, the number of routes and case allocation can be balanced.

[8] Dispatchers are also cognizant of these criteria constraints when modifying route plans.

- **Geographic Distribution of Stops**: We know that occasionally the routing system includes outlying stops and that this is an important consideration for the dispatcher in rebalancing stops.

- **Number of Driver Hours**: Drivers work an average of eight to ten hours, with ten hours being the maximum allowable by the US Department of Transportation regulation.

It should also be noted that, as a component of customer service, BevCo tries to provide drivers with similar routes and customers each day, although this is not one of the key routing criteria.

1.5.4 Manifests

The manifest is the primary means for communicating information about a day's route to drivers. It lists all of the stops in the order determined by OptiRoute and the dispatchers. In addition to the stops, the customer information field and salesperson notes are the two most important pieces of information. These fields include information related to collection of payment (a driver responsibility) and time windows.

1.6 The Driver Experience

As mentioned in Section 1.4, BevCo relies on a combination of OptiRoute and the expertise of its front line workforce to execute the delivery process. Understanding how OptiRoute works, we now seek to understand the driver experience—what a driver's typical day looks like and how route plans are used in practice.

Drivers have some flexibility as to when they begin their day, with some arriving as early as 4:30 am. Each morning, drivers receive their route plans as generated by OptiRoute. Start times are staggered to decrease the amount of time a driver spends at the depot. This prevents a bottleneck as trucks exit the warehouse and move along local roads. Drivers have developed their own processes, the most common being detailed below.

Manifest compliance is low, as few drivers follow the stop order as provided. Upon receiving the manifest, the first criterion that drivers have for determining stop order is time window. Information on time windows originates from three sources: salespeople's notes, an exception input to OptiRoute and drivers' field knowledge about customers.

After the drivers get a sense of the constraints placed on the route by time windows, they proceeded to sequence the stops. Wherever possible, stops located close to one another are grouped together, except where a time window takes precedence. For example, a driver may begin the day by stopping at a large chain hotel because this type of customer accepts deliveries at any time of day. Even if several stops are located on the same street, if they do not open until later, those deliveries must be postponed. Drivers can call the dispatch for support while out in the field.

1.7 CarTel Travel Times

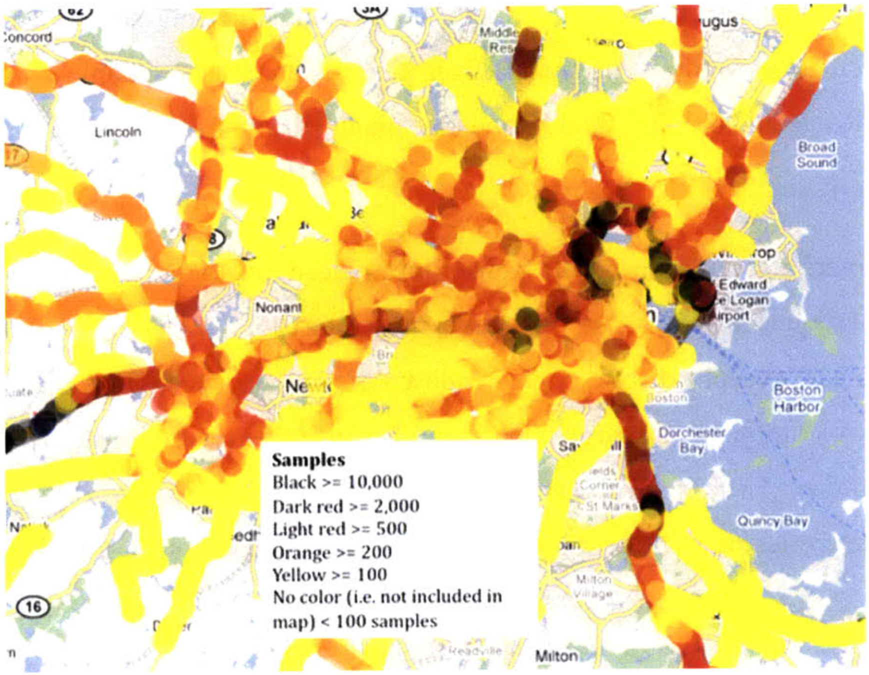
The alternative to traditional vehicle routing being explored is the use of historical data on road segment travel times to plan routes. Here we present background on how data is collected and how it can be parsed.

CarTel[9] is a computing system designed to collect, process, deliver and visualize data from mobile sensors. The CarTel group, which is part of MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL), has been collecting travel time data via units deployed on a private livery fleet and provided this data for our thesis.

The figure below shows a rough approximation of the samples taken around the Greater Boston Metropolitan area.

[9] For more information on CarTel, visit http://cartel.csail.mit.edu.

FIGURE 1: CARTEL DATA SAMPLING MAP

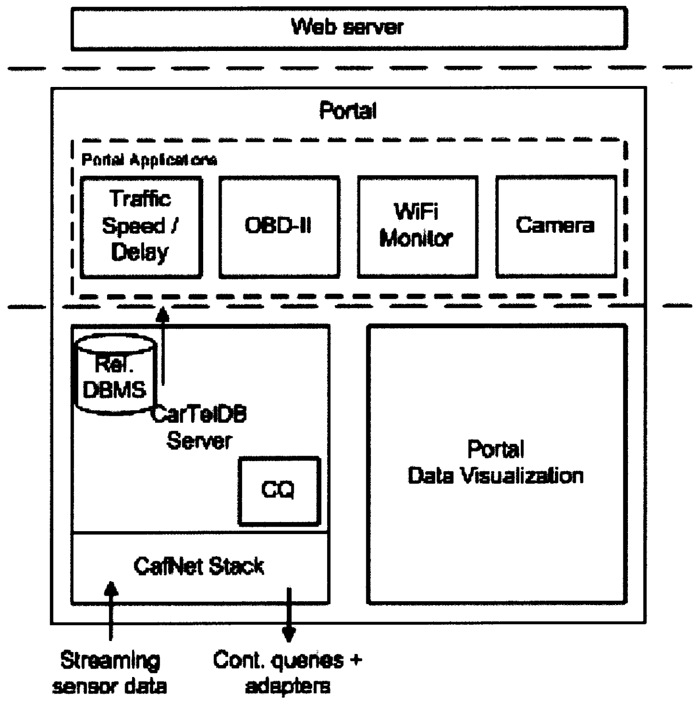


Source: CarTel

CarTel's data can be used to determine travel time profiles based on historical patterns along roads throughout the Boston Metropolitan area. Because the data set contains timestamps, each data point can be assigned to a time bucket, and CarTel has divided the day into hourly time buckets. By segmenting each day into 24 slices, descriptive statistics linked to time of day can be generated. This research uses the mean travel time for a given road segment at a given time of day to developing route plans.

There are three main components to the CarTel system as illustrated in Figure 2. The portal is the central location that hosts CarTel's applications and functions as the point of control and configuration for the distributed system. Data is sent from the mobile nodes to the portal. Second, an intermittently connected database (ICEDB) and third, CafNet (a carry-and-forward network), specify how the mobile units should collect, process and deliver sensor data. Data is organized into traces, which are sets of sensor readings collected by the mobile units during a drive.

FIGURE 2: CARTEL SYSTEM



Source: "CarTel: A Distributed Mobile Sensor Computing System"

The primary mode of the network access for the CarTel units is via opportunistic wireless, which is intermittent connectivity resulting from unsecured Wi-Fi networks. The system collects road traffic data, monitors the quality of Wi-Fi access points on routes, captures images along drives and gathers data from the On-Board Diagnostic (OBD-II) interface from vehicles.

1.8 AN ALTERNATIVE FOR BEVCO: SPECIALIZED DISPATCH OPERATIONS

BevCo employs different routing approaches regionally. As it may be determined that certain process improvements may increase the efficiency of vehicle routing and delivery, the SDO is being considered as an alternative to the warehouse operations in Boston. Located outside of New England, the Specialized Dispatch Operations (SDO) employs a holistic approach by breaking down delivery into three key factors: people, process and technology. The SDO facility houses dispatchers for multiple warehouses, and routing processes are more standardized along the three factors than compared to Boston.

1.8.1 People

A motivated and effective workforce is essential to efficient on-time delivery, and this is reflected in BevCo's organizational structure. People are divided into four roles: SDO Manager, Dispatchers, Service Managers and Drivers.

- **SDO Manager:** The role of the SDO Manager is to provide dedicated leadership through managing the SDO office, dispatchers and overall process. These managers are responsible for ensuring that performance is met and issues are resolved. The SDO manager even coaches dispatchers whose performance is suffering.
- **Dispatchers:** In this configuration, dispatchers work remotely at a centralized facility to review OptiRoute's output and finalize routing solutions. Dispatchers communicate with warehouse operations through a central point of contact, the Service Manager.
- **Service Managers:** Based at each warehouse, service managers oversee the drivers and manage communication between the dispatcher and the front line workforce. In this leadership role, Service Managers mentor drivers and supervise warehouse operations.
- **Drivers:** Drivers have the same responsibilities at the SDO as in Boston. However, instead of leveraging their own knowledge about routes and local areas, they are required to share that information with the Service Manager and use a 1-800 number to report time windows and other customer specific information.

Not only are roles clearly defined, but lines of communications are as well.

1.8.2 Process

There are four components to the delivery process: plan, prepare, execute, and closeout. Each role above has clearly defined tasks associated with each stage, and there is alignment between activities for each role.

Planning begins the day prior by reviewing the route plans and tasks to be completed. The necessary preparations are made—the prior day's results are viewed, conflicts are resolved, vehicles are loaded, and route plans are reviewed. The following day, employees are ready to hit the ground running and execute. Once all tasks are completed, then the closeout process begins, which includes debriefing and completing paperwork.

Two facets of the SDO's processes are distinct from those in Boston. First, manifest compliance is required of all drivers. This means that drivers are required to follow their route plans, and if they identify improvements that could be made, they must report them and obtain permission. Notably, the information reported by drivers is captured and incorporated into future route development. Second, for each role, specific tasks are assigned and simple checklists are utilized to ensure employees in each position complete each step. This promotes standardization.

1.8.3 Technology

The third and final element of the vehicle routing is technology. Technology tools can be classified as high and low tech.

High Tech

- **OptiRoute**: This enterprise-level software is also used in Boston.
- **Performance Scorecards**: These are used to communicate performance to employees.
- **1-800 Number**: Drivers can file incident reports and report information about customers that should be incorporated into future routing by calling a dedicated number.

Low Tech

- **Paper Manifests**: These are the route plans listing stops that drivers follow.
- **Daily Checklists**: These one-page checklists standardize processes for each role.

The technology used by BevCo supports effectiveness in its people and process goals.

2 LITERATURE REVIEW

The intent of this research is to understand the challenges facing delivery fleets, as well as the various heuristics that are applied to routing problems. By taking this approach, we hope to better understand the benefits and shortfalls of industry standards for routing, including the evolution of why deterministic data is used. We consider literature discussing the use of stochastic methods to understand the benefits and issues of incorporating travel time data. Finally, to become acquainted with traffic-related network effects, including the role that information plays, a brief review of relevant literature was also conducted.

2.1 EVOLUTION OF STUDY

The problem under consideration is a Vehicle Routing Problem (VRP). Research dates back to the late 1800s, when the Travelling Salesman Problem (TSP) and Messenger's Problem sought to employ algorithms to find the shortest possible tour between a series of geographically dispersed points (Menger, 1932). In the 1950s and 1960s, Dantzig (1955) and Fulkerson (1961) introduced the use of linear programming (LP) and developed an offshoot of the TSP called the Vehicle Routing Problem (VRP) (1959). The vast body of research on the VRP, in which a delivery truck is required to originate from and return to the same destination after servicing a predetermined number of stops, provides contextual background on what vehicle routing software seeks to accomplish.

The VRP seeks the optimal way to connect an origin and many destinations with the least amount of travel cost. While cost is typically measured in distance, it can be also measured in time, depending on the objective function under consideration. There are extensions to the VRP that add real world constraints, such as the Vehicle Routing Problem with Time Windows (VRPTW), Vehicle Routing Problem with Pickup and Delivery (VRPPD), Capacitated Vehicle Routing Problem (CVRP), and Stochastic Vehicle Routing Problem (SVRP). Of these, the VRPTW and SVRP (where time varying networks are stochastic) are most appropriate to our research.

Golden and Assad (1988) provide a comprehensive and classically regarded approach to modeling and implementation applications. Although their work provides the foundation for understanding vehicle routing systems, it must be noted that the bulk of literature solves variously constrained VRPs using a *deterministic* model, even though *stochastic* models more closely emulate traffic patterns.

In solving these problems, there has been an evolution from using algorithms to LPs to heuristics. The TSP and subsequent VRP are examples of NP-complete problems, where NP stands for nondetermininstic polynomial time. The number of steps required to reach a solution, regardless of method, increases rapidly with the addition of stops. LP solutions can approach computational intractability, where large problems require such immense processing power and time that for all practical purposes, they cannot be solved. This level of complexity favors the use of heuristics over LP methods. Heuristics arrive more quickly at solutions and have been widely adopted.

2.2 THE VEHICLE ROUTING PROBLEM

First, to create a fundamental understanding, we review deterministic heuristic models and their benefits, as well as crossover learning to the SVRP. This will be followed by a discussion of the SVRP and stochastically described decision variables. Stochastic variables take on many meanings in the VRP. The majority of literature considers demand as stochastic, that is, demand is unknown before arriving to the delivery point. We focus on literature that discusses travel times as the stochastic variable, that is, travel times are unknown before travelling.

Braysy & Gendreau (2005) survey and compare the numerous heuristic approaches proposed by researchers. The most common method of evaluating the solution quality of a heuristic algorithm is by empirical analysis. Most heuristics are not easily comparable, and there are two difficulties that affect the analysis: 1) the use of different computational machines and 2) only the best results from most studies are reported, because data on the efficiency of methods is lacking. Frequently, Solomon's (1987) 56 benchmark problems are used as the initial data set to which the various algorithms are applied. In their article on the VRPTW, Braysy & Gendreau discuss three sequential insertion heuristics in depth: 1) savings heuristic, 2) time-oriented nearest neighbor, and 3) "I1." The I1, which starts with a "seed" customer and adds the remaining un-routed customers to the route until it is full, is the most commonly used.

In the VRPTW, routes must be planned so that each geographic point is visited only once by a single vehicle within a selected time window. Also relevant is that the optimization is for the number of total routes. BevCo's vehicle routing software, monitored by a dispatcher, performs this function for the drivers.

Of the cost considerations for the objective function, the minimum number of tours and total distance traveled are more commonly used. Further, as heuristics, there are general methods of solving a common problem. Adding speed limits and similar deterministic constraints fits within the scope of many of the static solutions presented; however, incorporating travel time probabilities is an advanced approach.

Miller-Hooks and Mahmassani (2000) look at the least expected time (LET) paths in stochastic, time-varying transportation networks. They provide two algorithms for determining the least expected amount of travel time, as well as an expected range for those times. In time varying networks, future travel times are *a priori* with uncertainty, demanding that travel times be treated as random variables assigned a probability distribution function. The distribution allows the value of the variable to vary with time.

For this particular type of SVRP, solutions may produce multiple best paths associated with Pareto optimality/non-dominance. This contrasts with deterministic networks (and similarly, with deterministic software systems), where there is always one best path. We know that in the real world, where traffic is often linked to time of day, that the best path does indeed vary over time. This reinforces the argument that using real traffic data will provide improved routing solutions. In addition to the findings above, the important issue of computational efficiency due to the complexity of SVRPs is raised.

A survey by Bianchi (2006) of the performance of five common meta-heuristics used to solve the SVRP problem finds that hybridization between the methods and an estimation of local costs using a traveling salesman approximation both reduces some of the computational challenges and improves the performance of three of the five meta-heuristics. Highlighted in the survey were Simulation Annealing (SA), Tabu Search (TS), Iterated Local Search (ITS), Ant Colony Optimization (ACO), and Evolutionary Algorithm (EA); the performance of the TS, ILS, and EA improved.

Similar to real-life routing, the VRP and its various extensions seeks to take a set of geographically scattered points, group and connect them into the most efficient routes possible. Across the literature, classical local search methods, which are a class of heuristics that seek to iteratively improve a solution by exploring neighboring ones, are most frequently employed. Of these, the Tabu Search, a type of local search algorithm whereby once a potential solution has been identified, it (and the points included in it) are marked as taboo and removed from the greater set of data points being optimized, seems to be one of the dominant methods. This may be due in part to its ability to produce a faster computational speed with only a slight tradeoff on optimality.

2.3 NETWORK EFFECTS

Van Woensel et al (2007) explain the underlying traffic flow of an SVRP by using a queuing approach. This is a way to translate traffic flows into speeds that can be modeled. A key contribution is calculating road segment travel time depending on when the segment was traveled. The authors seek to use the queuing approach as a model for calculating expected travel times rather than use a model based on empirical data. While our research relies on empirical data to estimate travel segments, travel times may be calculated for the same road segment at different times of day. In fact, considerable research has been conducted on finding the optimal number of buckets. Van Woensel sliced the 24-hour day into 144 ten-minute time buckets; CarTel uses 24 one-hour time buckets. Fleishmann et al (2004) found the greatest improvement in travel time accuracy when increasing the number of buckets from three to five. Further improvement was found by increasing the number of buckets to ten, but limited improvement was seen beyond ten.

Gao and Chabini (2006) provide a framework for finding an optimal routing policy in a stochastic time dependent network. Their model seeks to minimize the estimated travel time to move from an origin to a single destination using both *a priori* and real time data. They discuss the Advanced Traveler Information System (ATIS), which provides real time data obtained via traffic cameras to assist travelers with making better travel decisions for themselves and collectively for the network.

In direct contrast, in his book *Urban Transportation Networks* (1985), Yosef Sheffi describes a system where drivers attempt to improve their own situation as a paradox. According to economic theory, attempts for each person to optimize their own situation without regard to the global optimal can worsen the situation for all. This investigation into network effects raises interesting issues as to how much individuals seeking to improve routing can do so, as well as the role that information plays in the network.

3 METHODOLOGY

To inform the hypothesis that a travel time-based routing approach will produce improved routing solutions, our approach was both quantitative and qualitative. On the quantitative end, a comparison of three sets of routing data was conducted:

- **Actual Data**: Used to benchmark the two route planning systems, two sources of actual data were used.
  - **Archived Manifests**: The manifests detail the routes in the actual order that stops were made based on sales transaction timestamps. This data set is our baseline due to the fact that the stops for virtually every route were made in a different order than the original route plan. We make two assumptions: 1) drivers deviate from their route plans intentionally, and 2) that they have some information that neither OptiRoute nor we have.
  - **Automatic Vehicle Location (AVL) Sequence**: These details include an engine on and engine off timestamp. These were used to calculate travel and stop times.
- **Resequenced OptiRoute Route Plans based on Archived Manifests**: OptiRoute created a time estimate based on the delivery sequence determined from the actual data.
- **CarTel Projected Travel Times**: The final part of the data analysis and benchmarking process will be to predict travel times for the manifest routes using CarTel's time estimates and the actual stop sequence.

The process for preparing and analyzing the data provided a basis for contrasting routes in terms of individual segments (origin to destination), as well as aggregated days.

In terms of qualitative methods, a series of interviews and site visits with BevCo was conducted to provide an understanding of the story behind the data. This background information informed the initial data scrubbing and subsequent analysis. The qualitative research included a site visit to the main distribution center for North Boston and Cambridge, two route rides with drivers as they made their daily deliveries, a site visit to the Specialized Dispatch Operations (SDO) facility and clarification from interviews with executives as needed.

3.1 PREPARING THE DATA FOR ANALYSIS

BevCo provided an Excel spreadsheet with ten full months of data on routes, segment times and stop sequences. Table 2 provides an account of the records and a description of each.

TABLE 2: DATA DESCRIPTIONS

Field	Description
Customer ID	A unique identifying number assigned to each customer.
Route number	A local, identifying number assigned to each route. This is an arbitrary assignment; the same route number can be used for a route allocated to a bay or bulk truck. Similarly, a driver is not assigned to a designated route number, nor are stops assigned to a specific route, although this may occur. Finally, a route can be run every day or on occasion as needed.
Case equivalents	Due to product differentiation, an order may include different sizes and quantities of products, including two liter bottles, eight ounce bottles and small, specialty beverages of varying shapes. To simplify the distribution process, a case equivalent is assigned to each type of product. The case equivalent is also used to project how long a driver will spend physically delivering the project to a customer[10].
Travel distance	Distance in miles of each route segment.
Estimated travel time	Estimated time in minutes to travel from stop to stop. Distance is

[10] This is accounted for in the stop time model.

Field	Description
	calculated by the software using underlying data on average speeds for a given road.
Estimated departure and arrival times	Based on the initial time that a driver left the warehouse, the estimated stop time and estimated travel time, these are projected for each stop.
Scheduled delivery sequence	The output sequence of the commercial routing software; also the sequence provided to drivers in their route plans.
Actual delivery sequence	The stop sequence in which drivers made their deliveries.
Transaction print time	When a delivery is made, a customer must sign for a product. This provides a time stamp. The assumption is made that this occurs at the end of a transaction.
Route break	Due to labor regulations, each driver is allotted a half-hour break every four hours. This break time is inserted into the drive time.

A second spreadsheet containing longitude and latitude coordinates (geocodes) for each customer ID was also provided, since OptiRoute and CarTel both use the customer geocodes in determining estimated travel time. For ease of analysis, the primary spreadsheet was updated to include this customer information.

3.1.1 Deriving the Baseline Data

Next, we had to determine what to do with erroneous entries such as duplicate entries and blank transaction timestamps. Consulting with our sponsor informed us that duplicate entries typically resulted from a skipped or missed stop that postponed service to another day. Blank transaction timestamps indicated the driver had another reason to make the stop other than to deliver product (e.g. to pick up a check). In both instances, the entry was not relevant and therefore removed.

Also of note, although descriptive and useful in helping us understand the more qualitative nature of driver behavior, some data fields such as scheduled delivery sequence were not used. The objective from this data set was to obtain a picture what did happen, not what should have happened. Another concern was that on routes for which a bay and bulk truck could be used, the data may include runs for both types of trucks. Further scrubbing the data was postponed until some analysis could be conducted and informed decisions made. It was agreed that as long as these routes were in the minority, they would be excluded. Upon completion of this process, we had the first of the three data sets—the actual routes as run by drivers.

The second part of the process was to prepare the second data set—OptiRoure's manifest projections using the actual stop sequence. The primary data set contained upwards of 65,000 entries; even though these were individual stops as opposed to entire routes, the decision was made to explore two weeks of data. The final two weeks of October 2009 were selected due to a lack of holidays and the assumption that this would be a relatively shock-free time period. The result was data set of 3000 entries, which was large enough to provide a picture but more manageable from our point of view.

After separating the two weeks of data from the larger set, we forwarded the spreadsheet to BevCo for resequencing in OptiRoute. This yielded route plans based on the same stop sequence that drivers actually took. These route plans were our second set of data.

The third step was to provide CarTel with the necessary data for its routing estimations. In order to extract this data, a simple spreadsheet that only included the data points of origin and destination, latitude and longitude, and the time of day the travel would take place. Similarly to the second data set, we only used the entries from the final two weeks of October. CarTel then used its own script to develop a travel time estimation using its own database. We collaborated with CarTel to extract a similar dataset for expected travel times between origin and destination pairs.

3.1.2 Using Automatic Vehicle Location Devices to Collect Additional Data on Actual Performance

While sponsoring this thesis, BevCo began piloting AVL devices on some of its bay delivery trucks. These devices are connected to the vehicles' on-board diagnostic (OBD-II) systems and also have a real-time monitoring capability through AVL. Data collected included vehicle start and stop times, as well as a trail of locations from the AVL satellite pings. Data could be accessed and downloaded through an online portal.

With support from BevCo, we decided to include an additional analysis of data collected from March 1 through March 26, 2010 using these devices. We were curious to understand the results from examining another "actual" set of data. Initially, the intention was to increase the rate at which the AVL sampled location and use that data. However, due to time constraints and the complexity in filtering delivery stop times versus travel-related stops, the vehicle start-stop data was used.

3.2 UNDERSTANDING THE STORY BEHIND THE DATA

In addition to the hard data analysis, it was important to gain an understanding of how the route planning and delivery occurs in practice. Not only did this provide context, but it increased our understanding of the company's operational policies and overall strategy. Therefore, the qualitative research included a site visit to the main distribution center for North Boston and Cambridge, two route rides with drivers making their daily deliveries, a site visit to a regional and out-of-state dispatch facility. Clarification from executives was also provided as needed.

3.2.1 Site Visit to the North Boston Warehouse

Vehicle routing for Northern Boston and Cambridge occurs onsite at the North Boston warehouse, so our first step in obtaining a clear picture of delivery operations was to tour the warehouse, observe the dispatcher preparing the routes for the following day and speak with employees.

3.2.2 Route Rides

Additionally, two route rides were observed with goals of understanding how route plans are used by drivers and what, if any, exceptions occur in the delivery process that may inform the analysis. One of the routes was located in the Greater Boston Metropolitan area; another was downtown.

The driving habits and experiences of an urban route (e.g. downtown Boston) are quite different from a suburban one. For a suburban driver with a concentrated route (one that traverses a single town) knowledge of the area and streets provides a significant advantage. For instance, knowing that traffic on Main Street in the direction towards Boston at 9 am is worse than traffic flowing away from Boston allows a driver to choose the less congested route, even if the manifest dictates otherwise. Another example is knowledge of back roads, which can be taken to bypass traffic or congestion. For a suburban route traversing many towns, traffic conditions play a much greater role in amount of time spent driving.

In contrast, for an urban route over a much smaller geographic area, traffic plays a less significant role. Geographically, the observed stops were located within a few blocks of one another, and traffic was not a significant source of uncertainty. In fact, little traffic was encountered. Anecdotally, the driver suggested that this route would be subject to seasonality, with more traffic encountered during the summer due to tourism.

The driver experience is also highly dependent on the type of customers serviced. Drivers build a rapport with their customers. Repeat customers with whom the driver had an established relationship were more likely to have payment ready at the time of delivery. Lowering transaction time is a priority, and the company has piloted a quicker order validation with one major pharmacy chain. The driver picks the order and brings a printout with a bar code. The contact in the store scans the barcode and quickly counts the number of case equivalents brought to the store. This saves time for both the driver and the retail outlet, helping to keep headcount to a minimum.

During the observation period, a number of exceptions to the planned routing process occurred. These instances were notable, as they could impact the analysis of the data:

- **Skipped Breaks:** Neither driver took a break. Manifests include a half hour break time.
- **Variability in Transaction Time Stamps:** In an attempt to create efficiencies at an office building with multiple deliveries, a driver created a transaction timestamp for all deliveries at one time, then loaded the product for all customers onto a cart, thus eliminating the need to return to the vehicle. In this case, the time stamp did not occur at the end of the stop. A similar process may be reflected in the data.
- **Reverse Logistics of Plastic Bottle Holders**: Cases of plastic bottles are carried by durable plastic trays. Drivers are responsible for collecting these from customers on return trips. This action is not accounted for in the stop time model.
- **Return for Payment:** For some customers, drivers must collect payment upon delivery. At one stop, the owner was offsite and had not left a check. The driver was told to return between the hours of 11 am and 3:30 pm. Therefore, the driver made a note to return to the store during that time window. Having to return for payment was reported as a common exception[11].
- **Mispick:** If errors occurred in the picking at the warehouse, the driver would not have the proper product assortment. In these instances, the driver communicated this with the customer and either made a note and had to return with the correct product at a later date or provided alternate product (what was picked and packed instead).
- **Queuing:** A series of stops were located down a small alley. Upon completing the deliveries, the driver had to wait until another truck in front of it finished making its deliveries to leave.

[11] When no transaction occurs, there is no timestamp. While we believe that this exception was scrubbed from the data, it is still worth noting.

As a result of the route rides, we confirmed that drivers do not follow their manifests and also received insight into delivery exceptions that may influence the data analysis.

3.2.3 Investigating an Alternative: Remote Routing Operations

BevCo's routing process differs by region; the final element of the qualitative analysis was a site visit to an out-of-state dispatch center. Called the Specialized Dispatch Operation (SDO), the facility houses dispatchers who route for multiple warehouses. While the process is similar to that at the Wilmington warehouse, there were marked differences.

In terms of contrasting the two operational configurations, two important observations were noted. First, processes were more standardized at the SDO than the warehouse. This translated into the second observation; knowledge about each routing function was more equally distributed with the SDO. For instance, information about routes that drivers had was collected and accessible centrally. Further, by implementing a checklist, another substitute dispatcher could more easily step in to perform the role in the case that the regular dispatcher was absent from work. This will be an important consideration in the recommendations.

4 ANALYSIS

Following the previously outlined methodology, we analyzed the three sets of data: 1) actual routes; 2) OptiRoute projections using the same sequence as actual; and 3) CarTel projections, also using the actual sequence.

In addition to evaluating the differences between route planning tools, it was important to understand why those differences exist. One hypothesis is that OptiRoute and CarTel deal with uncertainty differently. We also seek to explain the following:

- When CarTel projects longer routes, is it because the data indicates that a high likelihood that drivers will encounter traffic?
- When CarTel projects shorter routes, is it because CarTel has the capability of identifying shortcuts used as thoroughfares that OptiRoute's underlying map system identifies as neighborhood roads not commonly used?

Our findings are presented in this section.

4.1 MAPPING THE DELIVERY PROCESS

First to visually understand what the travel day looked like, we diagramed the delivery process. Figure 3 illustrates the process as captured by the actual data sets using timestamps, and Figure 4 depicts the process using AVL.

FIGURE 3: DEPICTION OF THE DELIVERY PROCESS USING TIMESTAMPS

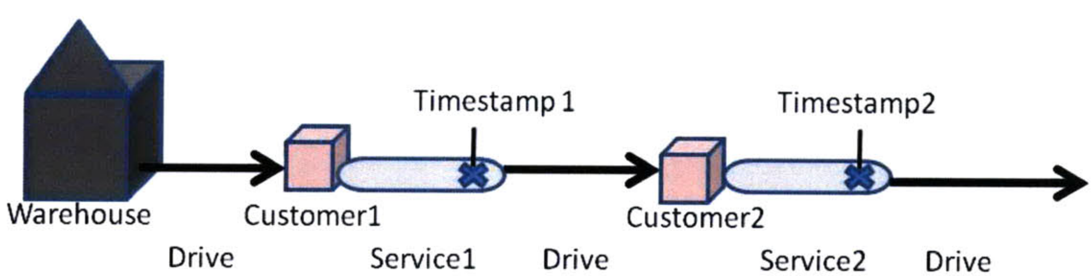
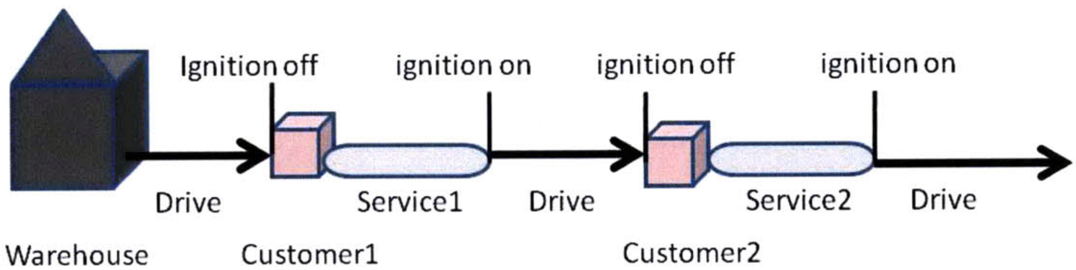


FIGURE 4: DEPICTION OF THE DELIVERY PROCESS USING AVL



In this way, we were able to understand how the actual data set translates into the aggregation of driving time segments and stop time estimates into a route by OptiRoute and CarTel. In the above figure, the driver leaves the warehouse and drives to his first customer. He spends time at that location, unloading product and merchandising shelves and then receives confirmation from the store that delivery was made. This confirmation is the transaction timestamp.

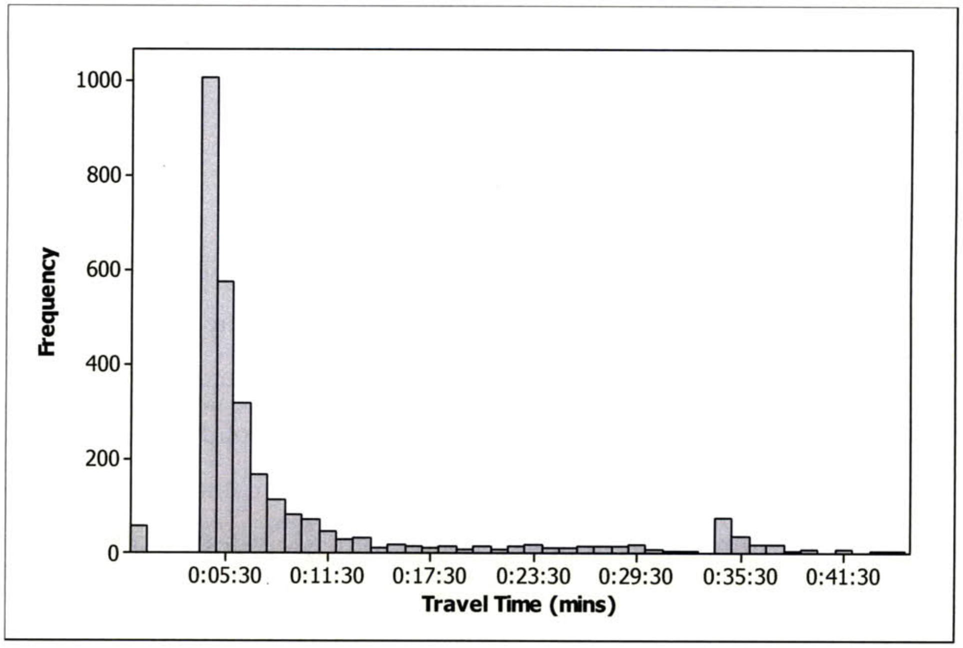
The segments in which the warehouse is the origin and the first customer is the destination (vice versa at the end of a route) are called the stem. Actual data on when the driver left and returned to the warehouse was not available. Since accurate stem times could not be calculated and benchmarked, they are not included[12]. Instead this analysis focuses on localized routing.

4.2 DATA PROFILE

Once the 10-month data set was distilled into a 2-week sample, we had approximately 3000 segments consisting of origin-destination pairs. To get a high level sense of how OptiRoute and CarTel project time, we compared the OptiRoute estimates to those provided by CarTel for identical segments. The distributions of these segments by length in minutes are illustrated in Figure 5 and Figure 6.

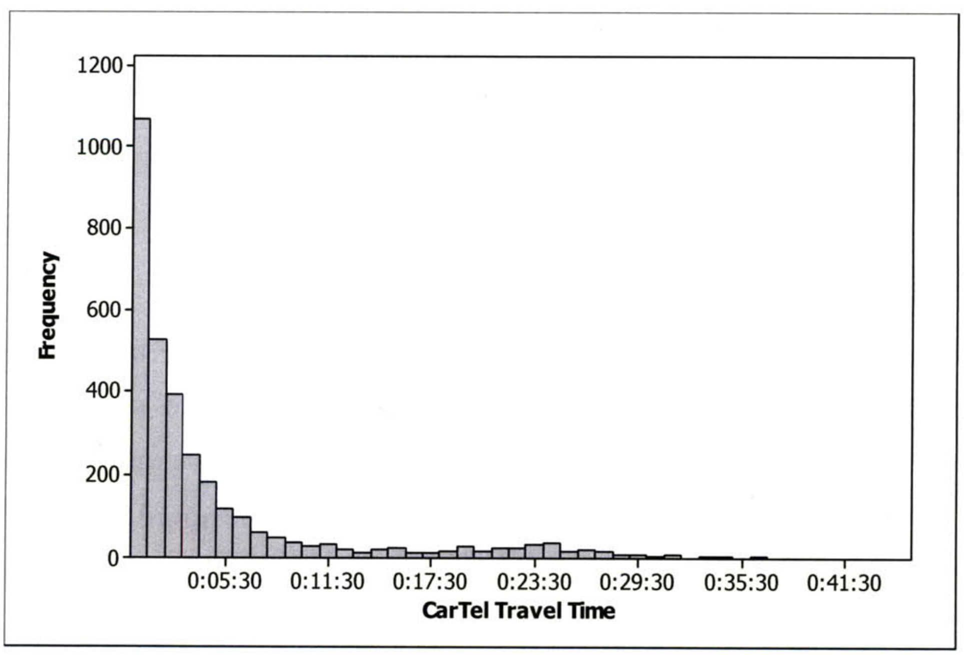
[12] Although stem times do represent longer driving segments, due to the location of the warehouse, the most efficient means for BevCo's drivers to reach the local delivery zone is via interstate.

FIGURE 5: HISTOGRAM OF OPTIROUTE SEGMENT TRAVEL TIMES



The highest frequency of segment times for OptiRoute occurs between 4 and 5 minutes, with another bump occurring between 34 and 35 minutes. The longest segment is close to 45 minutes.

FIGURE 6: HISTOGRAM OF CARTEL SEGMENT TRAVEL TIMES



In contrast, the highest frequency of driving segments for the CarTel data occurs at less than one minute. The longest segment is between 36 and 37 minutes and there is no spike at 34 minutes. Immediately, it was clear that OptiRoute and CarTel differ in their estimation of very short (less than five minutes) travel segments.

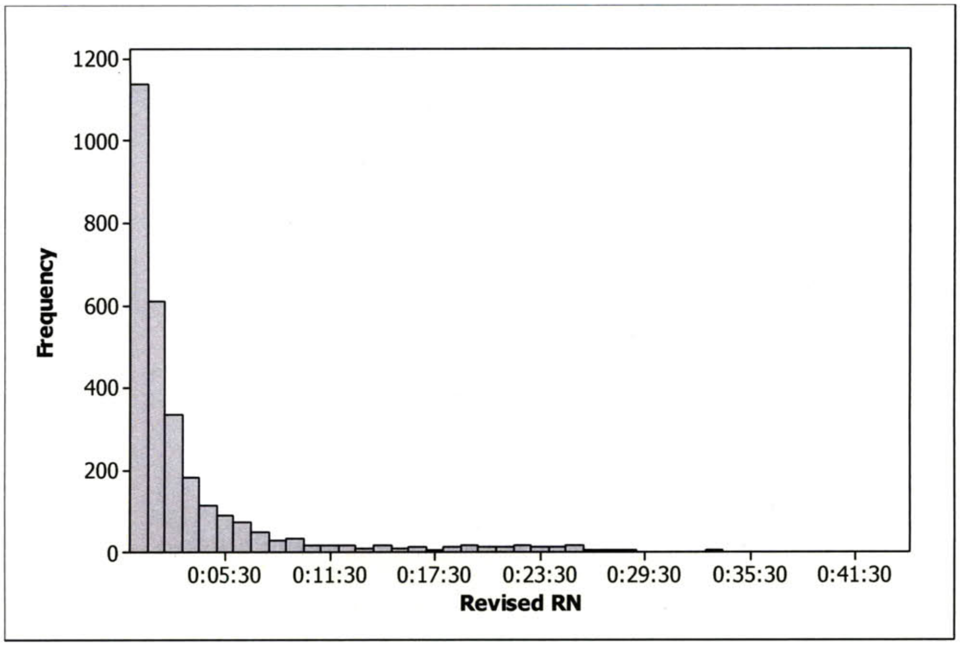
This seemed counterintuitive, so we proceeded to look into the data. OptiRoute estimates a travel time when consecutive deliveries are made at the same location if there is any differentiation in customer geocode (latitude and longitude coordinates). This means that even if the physical address is the same—in large buildings it is not uncommon to have more than one account—the geocode is the default criteria for sequencing. From our conversations with BevCo executives, we have confirmed that accuracy issues with geocodes exist. Given this, CarTel more accurately accounts for zero driving distance as zero driving time.

To further understand the four-minute OptiRoute segment peak, we proceeded to examine OptiRoute's drive time model. As discussed in Section 1.5.1, the system calculates drive time using a fixed and variable driving components. A four-minute fixed drive time is assigned to each segment with a geocode differential greater than zero plus a variable component based on distance and road speeds.

A second observation was a spike in the number of road segments at the 34-minute mark by OptiRoute; there was no spike with CarTel. This discrepancy can be attributed to OptiRoute's assignment of 30 minute breaks, which are required by the Department of Transportation.

After looking into these two discrepancies, it became clear that the 4-minute fixed time and 30-minute break were skewing the data. The 4-minute fixed time serves as a buffer for uncertainty in drive time, and the 30-minute break is entirely unrelated to drive time. Our objective was to isolate the expected drive time (only) for OptiRoute. We removed the break time from all subsequent analysis and note when four-minute fixed time is used. While still differentiated, Figure 7 shows how OptiRoute's "generalized" histogram resembles CarTel's more closely.

FIGURE 7: HISTOGRAM OF GENERALIZED OPTIROUTE SEGMENT TRAVEL TIMES



The most striking difference between the generalized OptiRoute and CarTel segments appears to be the tail, specifically segments over ten minutes. To explore the projections, we zoomed in on the histogram. Decreasing the frequency scale from 0 to 100 occurrences shows how CarTel and OptiRoute estimate longer travel times. This is depicted in Figure 8 and Figure 9.

FIGURE 8: HISTOGRAM OF GENERALIZED OPTIROUTE SEGMENT TRAVEL TIMES (TAIL)

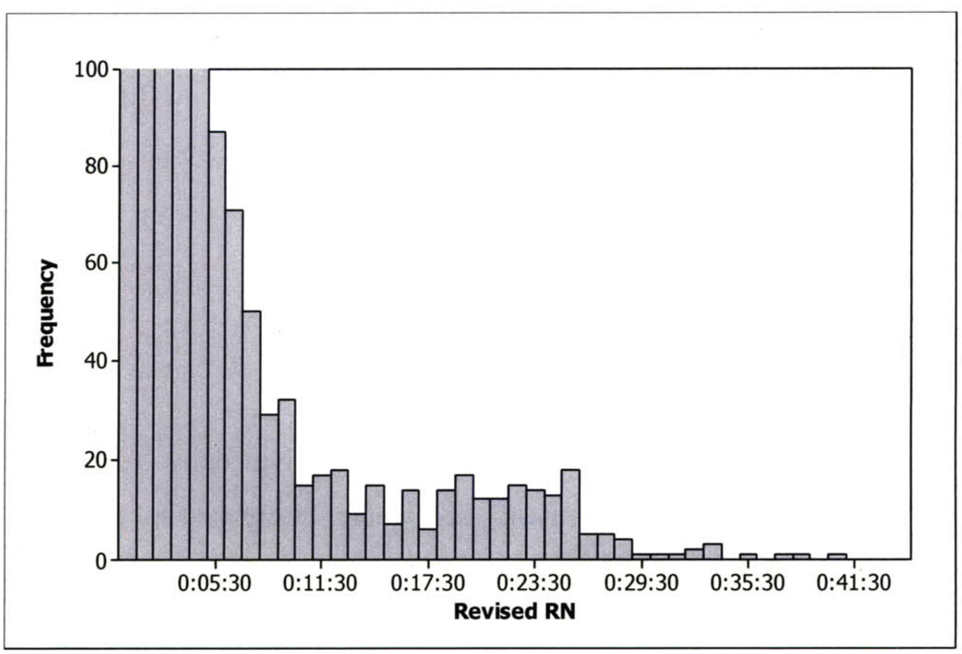
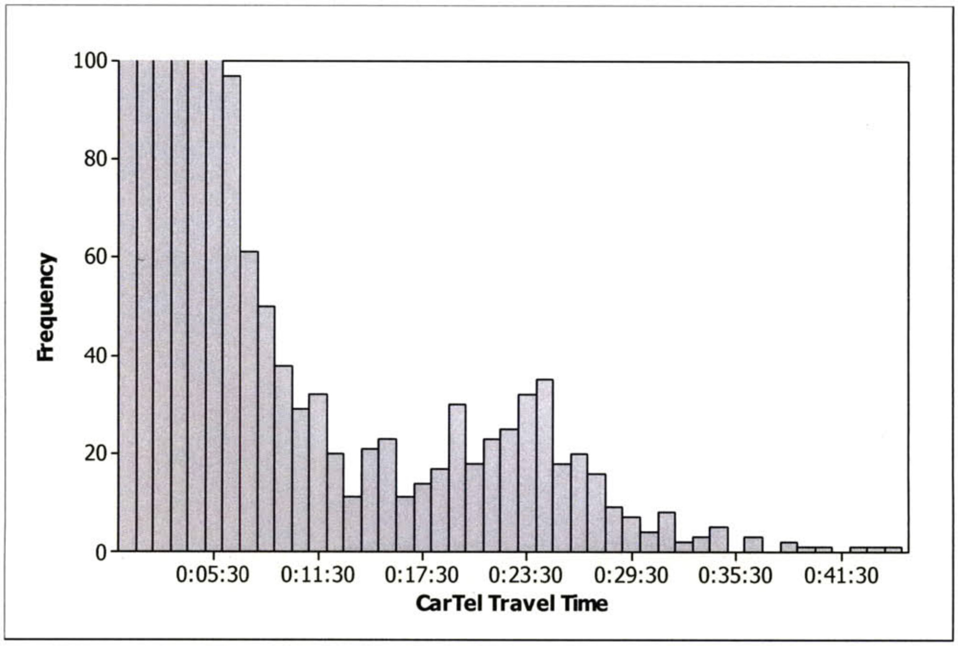


FIGURE 9: HISTOGRAM OF CARTEL SEGMENT TRAVEL TIMES (TAIL)



OptiRoute projects 8.2% of segments to be 10 minutes or longer; whereas, CarTel projects 13.6% of segments to be 10 minutes or longer. Immediate speculation for the difference revolved around how the two systems deal with uncertainty. OptiRoute projects segments based on average road speed, then adds the four minute buffer for uncertainty. By generalizing the data, not only are the projected segment times shorter, but all measures for dealing with uncertainty have been stricken from the data. On the other hand, CarTel's data incorporates uncertainty through its travel time probability distribution; there is no differentiation between expected driving time and uncertainty. The difference in tail segment frequencies calls for further investigation.

4.3 DAILY ROUTE SEGMENTATION

Next, segments were aggregated to investigate cumulative daily routes by individual methodology (OptiRoute, CarTel and actual). Each route, designated by a route number, was compared the in terms of time spent driving, time providing service customers, as well as other variables such as case equivalent load. Our research does not explore the stop time model[13], so the stop projections were taken as given.

4.3.1 OptiRoute Estimation of Total Route Travel Time

To obtain more detailed statistics and confidence intervals, we expanded the 2-week data sample to include 8-weeks of data with 10,000 data points. The additional data points were necessary to perform a one-way ANOVA[14] against OptiRoute's daily travel times by route, which was the next step. The intention was to see if the driving times were planned to be fairly consistent among routes. Indeed, they were. The ANOVA test shows with a p-value of 0.611 that we are unable to distinguish the routes based on estimated total driving time. This takes the dispatcher's human element of routing out of the equation as a cause for long or short routes.

4.3.2 CarTel Estimation of Travel Time

The traffic data estimates travel times by time of day that each segment was traversed, and naturally the aggregate routes reflect this information. In most cases, CarTel's cumulative route travel time estimation was greater than OptiRoute's, on average less than 15 minutes. Although notably consistent, this is a relatively short amount of time.

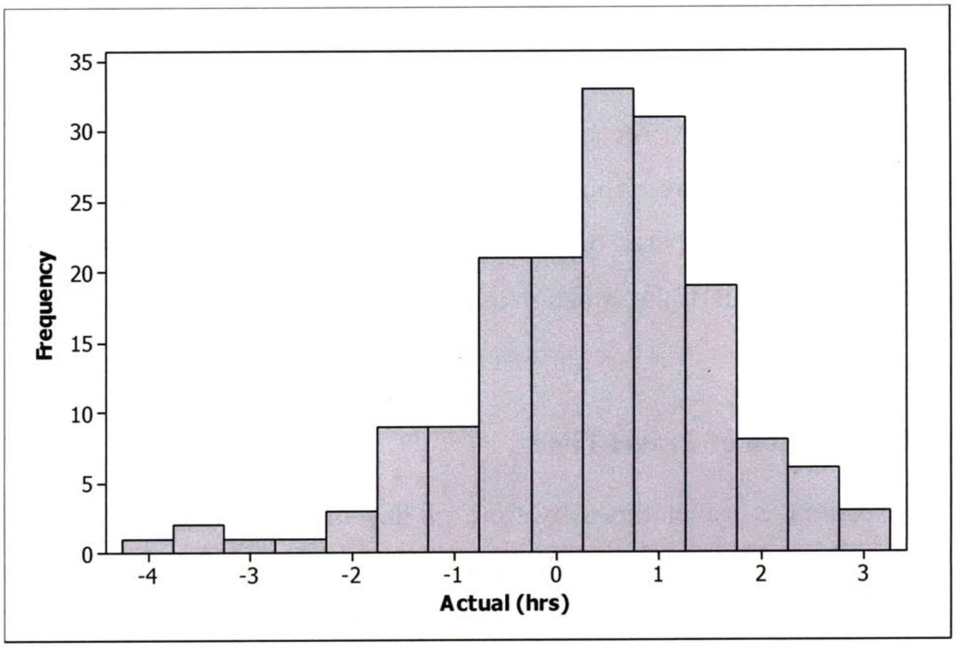
4.3.3 Actual Travel Time Using Transaction Timestamps

To estimate travel time, we isolated amount of time the driver was on the road. This was calculated by the taking difference of the original estimated departure from the depot and the calculated expected travel time needed to return to the depot added to the last transaction timestamp. Once the total amount of time the driver was away from the depot was calculated, the stop time as modeled by OptiRoute manifest was subtracted.

[13] Of note, the time estimate for providing service to customers, including unloading product, delivering to the stockroom, merchandising and receiving payment, is modeled in the OptiRoute software through the stop time.
[14] ANOVA stands for analysis of variance and is used to compare the means of two or more variables.

The results were surprising. A fraction of travel times were calculated to be negative. What that means is that the driver was estimated to spend more time servicing customers than they spent away from the depot. This finding is depicted below in Figure 10.

FIGURE 10: ACTUAL CALCULATED TRAVEL TIME IN HOURS



It is impossible for drivers to spend more time servicing customers than the total number of hours they spend at work, leaving us to question the accuracy of the stop time model.

4.3.4 Actual Travel Time Using Engine Data

As a second method of comparing routing data, we analyzed data collected using automatic vehicle location devices (AVL) linked into trucks' onboard diagnostic systems (OBD-II). The driving times and service times were calculated based on engine position. For each location, ignition off indicated the beginning of a service stop and ignition on indicated the end of the service stop. The time between ignition on and off was determined to be due to travel.

BevCo provided data for four weeks in March 2010; we compared the actual AVL routes with OptiRoute and CarTel estimations.

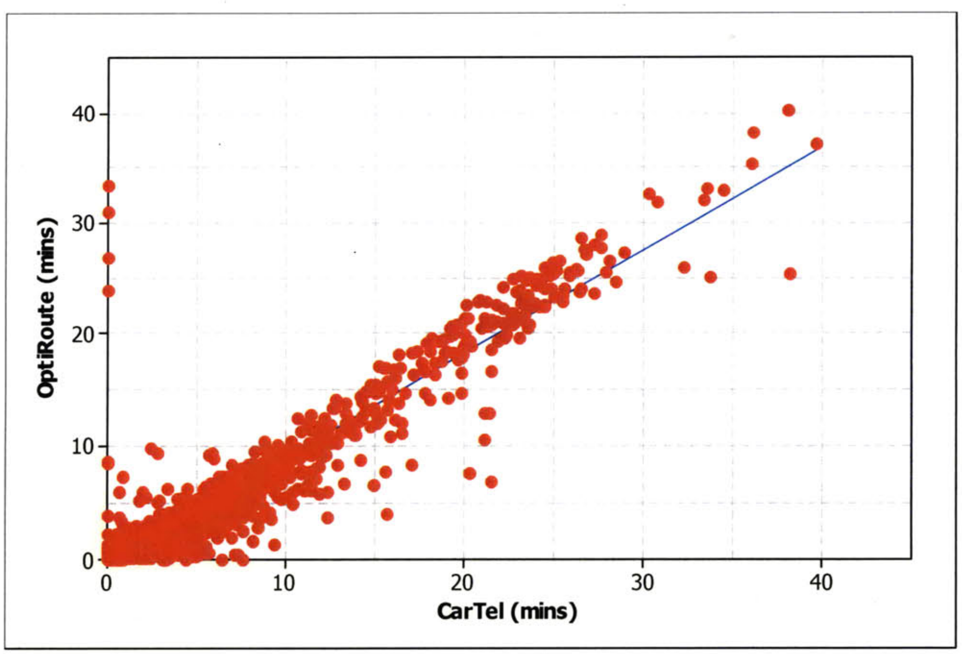
4.3.5 Outliers

Performing a one-way ANOVA test, led to the discovery that the average case load of route 570 was double that of the rest of the routes. Additionally, its standard deviation was four times that experienced by the other routes. Upon further investigation, it was determined that this route was served by both bulk and bay trucks, resulting in its data points producing outliers.

4.4 BENCHMARKING

In order to make recommendations on underlying road travel times, we compared OptiRoute and CarTel's time estimation for both road segments and entire routes with our October 2009 data set. By looking at each route segment, a very close correlation is seen between OptiRoute and CarTel's travel time estimations. Additionally, a regression was performed and the p-value was 0.000, confirming the statistical significance. The regression is illustrated in Figure 11.

FIGURE 11: CARTEL VS OPTIROUTE TRAVEL SEGMENTS



Breaking down the segments by route, the revised scatter plot shows many of the outliers pertain to route 570. Shown as black circles in the scatter plot below, the bulk of the outlying routes can be explained by the fact that both bay and bulk trucks are deployed on this route. Due to the high volumes of product and limited number of stops, bulk trucks skew the comparison with bay trucks. Figure 12 contains route 570 and can be compared to Figure 13 which does not.

FIGURE 12: OPTIROUTE VS. CARTEL SEGMENTS BY ROUTE NUMBER

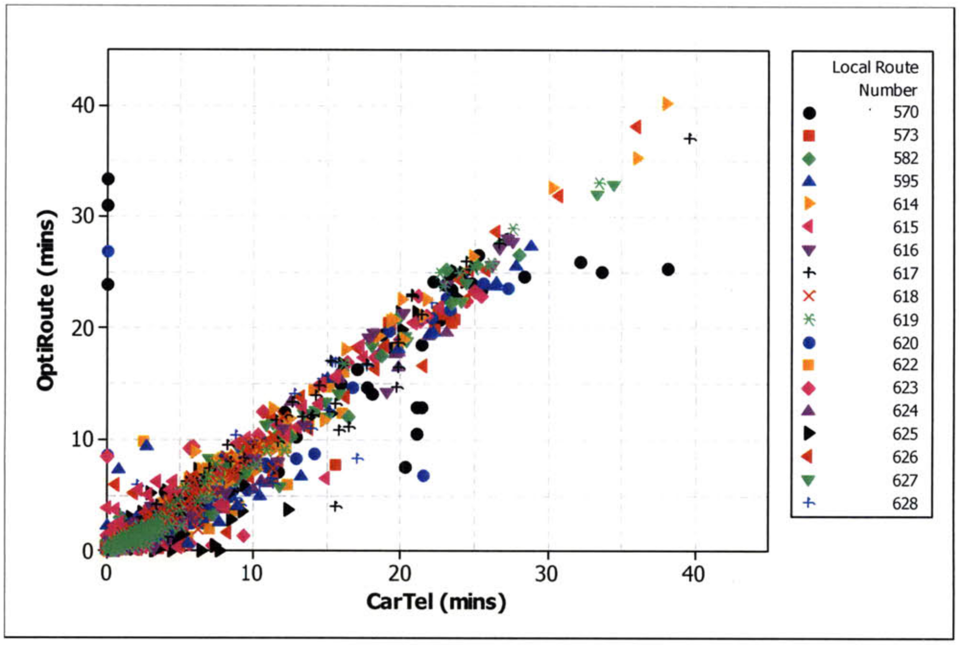
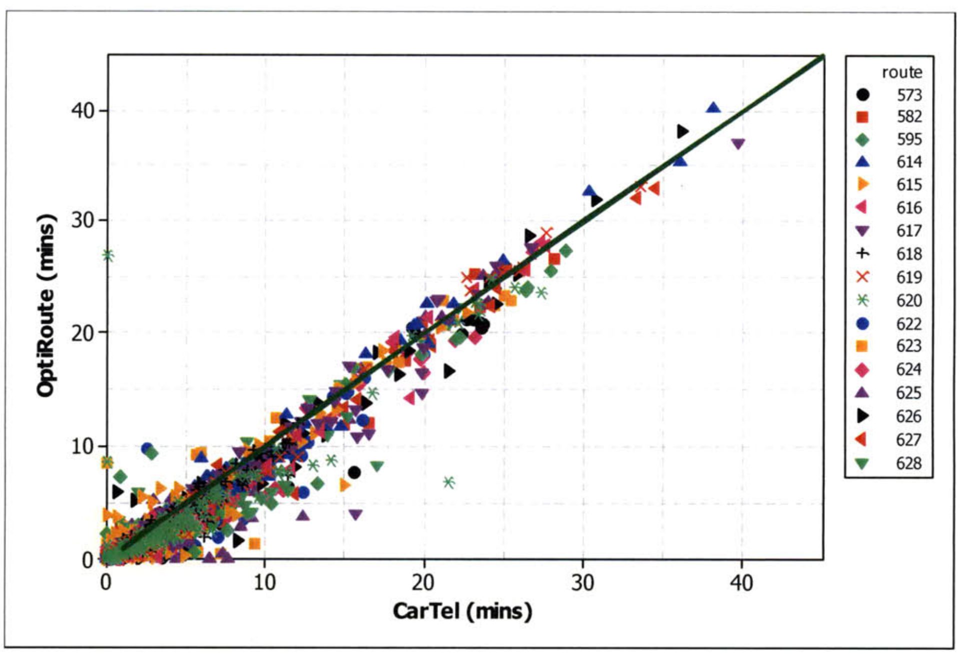


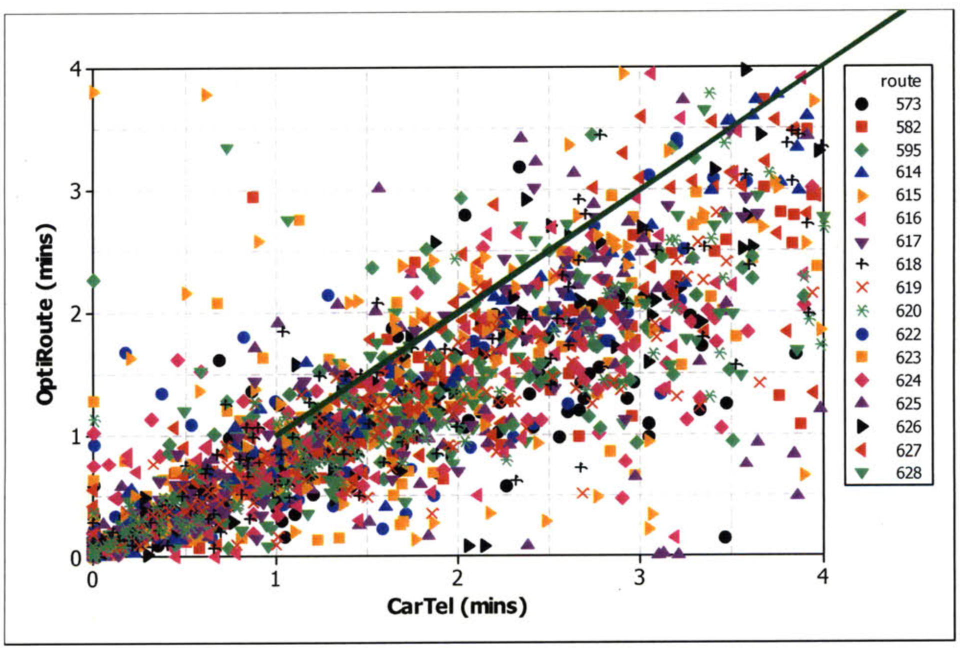
FIGURE 13: OPTIROUTE VS CARTEL SEGMENTS BY ROUTE NUMBER (NO 570)



Removing route 570 allows a regression to be achieved with a p-value of 0.000 and R-squared value of 93.8%.

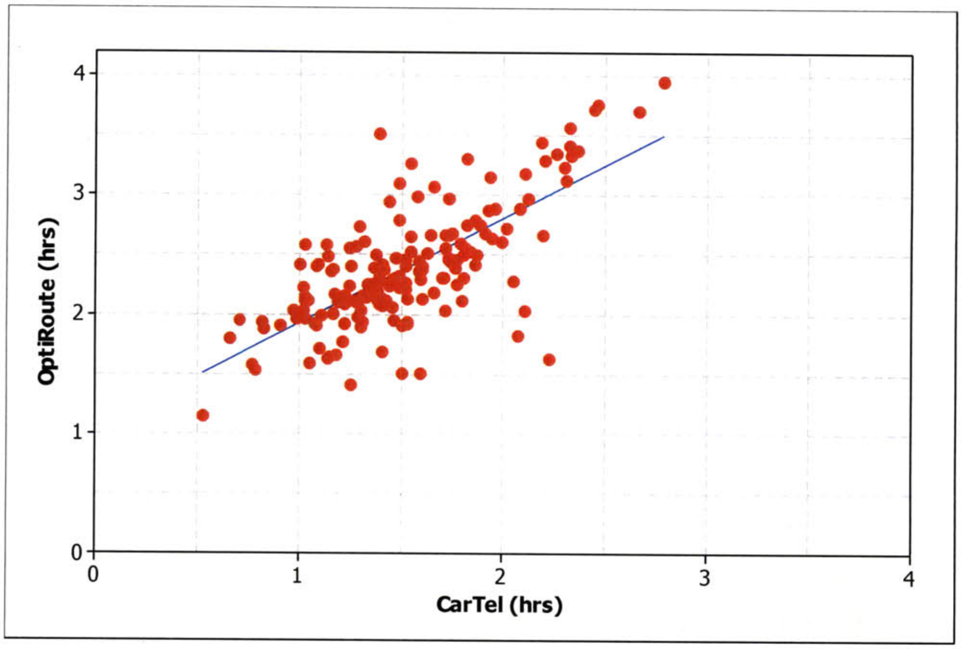
Values above the solid green 45-degree reference line are those segments for which OptiRoute estimated a longer travel time than CarTel. Points below the reference line indicate that OptiRoute estimated shorter travel times than CarTel. Our earlier investigation of segment time using histograms revealed that many of CarTel's segment projections above 10 minutes were longer than OptiRoute's. Figure 13 and Figure 14 illustrate that this also holds true for a majority of the shorter routes. By zooming in closer to the shorter time segments, it is clear that the bulk of segments appear below the reference line.

FIGURE 14: SUBSET OF OPTIROUTE VS CARTEL SEGMENTS



The spread between OptiRoute and CarTel's estimated segment times is dramatic. To obtain further insight, we aggregated the segments into daily routes. This shows a less obvious correlation between OptiRoute and CarTel, as depicted in Figure 15.

FIGURE 15: CARTEL AND OPTIROUTE TOTAL ROUTE TIME (WITH 4-MINUTE FIXED TIME)



Again, the scatterplot was broken down by routes, and route 570 contained many of the outliers. Figure 16 shows route 570 as black circles, while Figure 17 shows OptiRoute vs CarTel aggregated daily routes without route 570.

FIGURE 16: OPTIROUTE (WITH 4-MINUTES) VS CARTEL ROUTES BY NUMBER

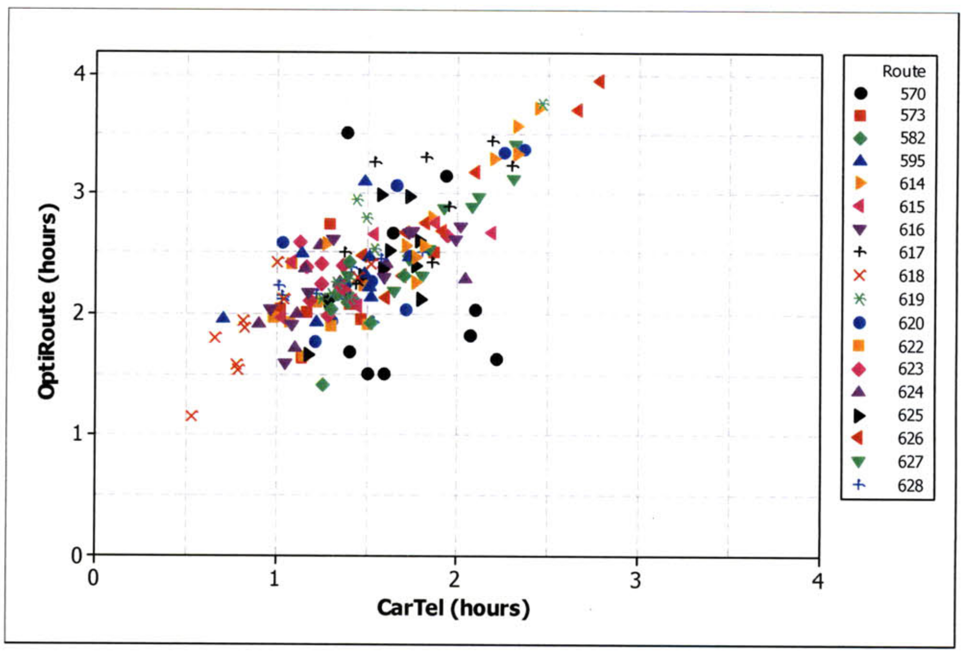
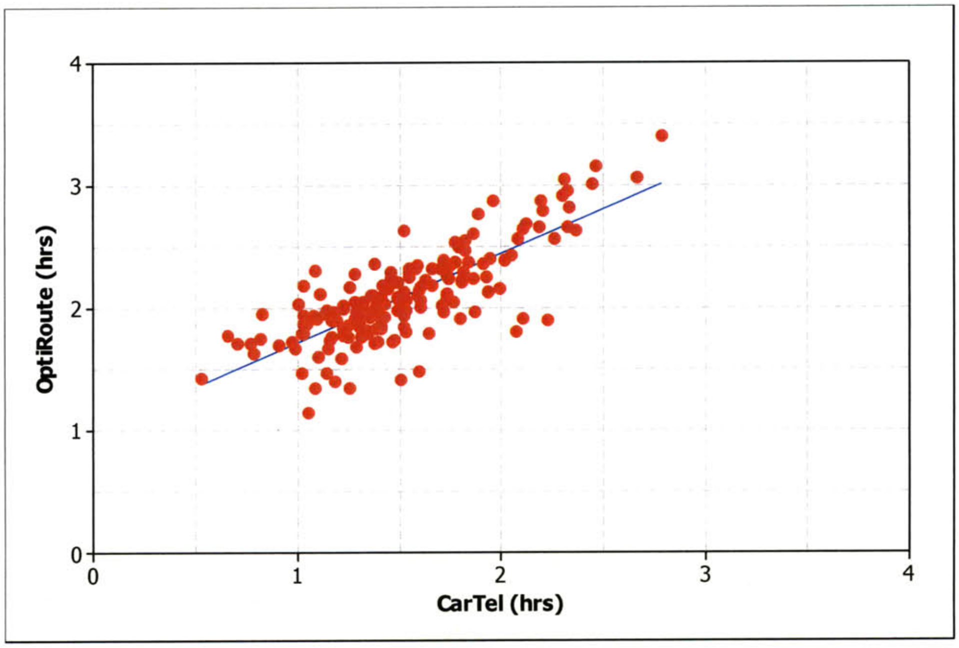


FIGURE 17: OPTIROUTE VS CARTEL ROUTE TIMES (ROUTE 570 REMOVED)

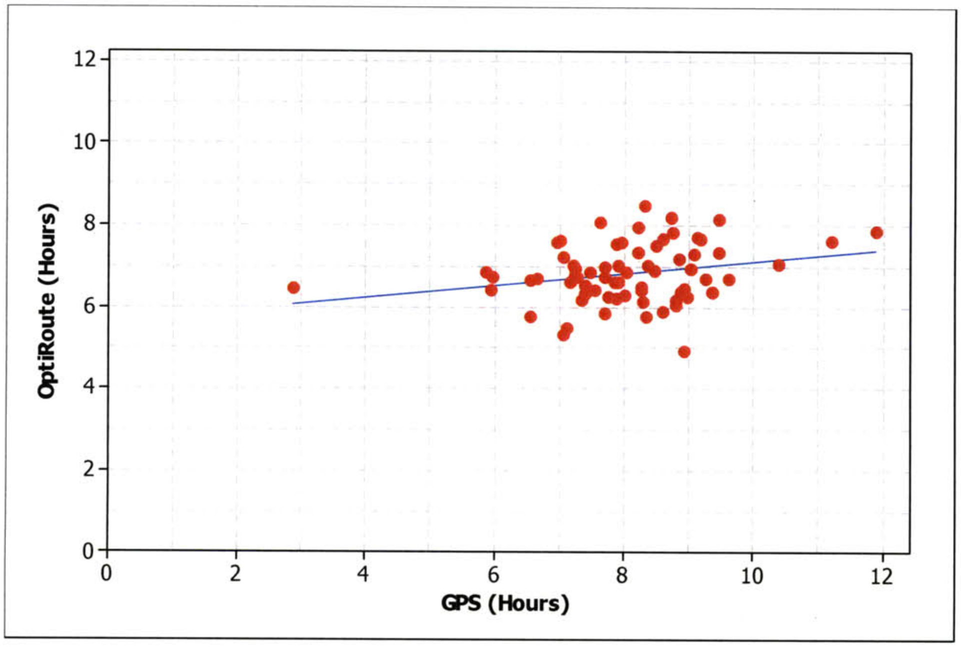


4.5 OPTIROUTE PERFORMANCE

Our final dataset contains AVL date from five of the trucks in BevCo's fleet. We compared OptiRoute's estimation of drive time and stop time models to actual driving behavior.

By first looking at the entire day, OptiRoute gives a decent approximation of how long drivers will spend out on the road. This is important as there are travel restrictions on the number of hours a driver can be at work. We looked at our generalized OptiRoute data, that is, no breaks were included and the fixed travel time of 4-minutes per segment was removed.

FIGURE 18: GENERALIZED OPTIROUTE VS AVL TOTAL TRAVEL TIME



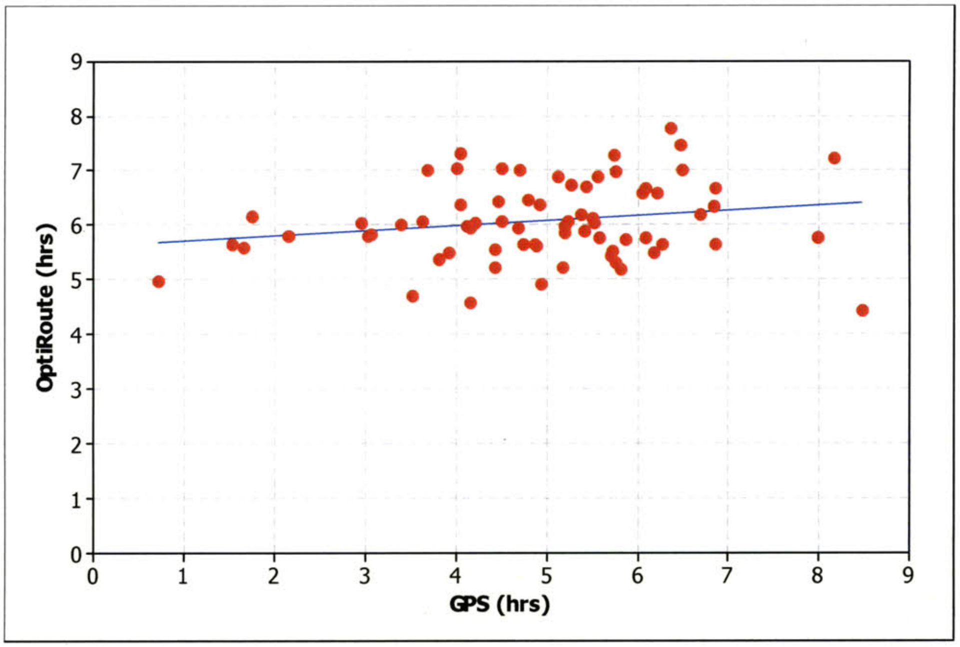
The regression between total travel day as estimated by OptiRoute and as measured by AVL devices gives a p-value of 0.027 and an R-squared value of 7.1%. The p-value indicates that the two data sets are correlated, but the R-squared value indicates that correlation is quite low. Then calculating the regression using our generalized OptiRoute data gives a p-value of 0.019 and an R-squared of 7.9%, both slightly better. These results are summarized in Table 3.

TABLE 3: TOTAL ROUTE STATISTICS

TOTAL ROUTE TIME			
	OptiRoute	Opti Gen	AVL
min	5.911	4.911	2.867
max	9.738	8.472	11.883
avg	8.042	6.823	8.013
std dev	0.798	0.716	1.288
N	70	70	70

By looking at the parts of the day – driving and servicing – OptiRoute doesn't perform these individual forecasting tasks well at all. This raises the concern that there is little understanding in the models of what actually is driving the time spent traveling from stop to stop and servicing customers.

FIGURE 19: AVL VS OPTIROUTE SERVICE TIMES

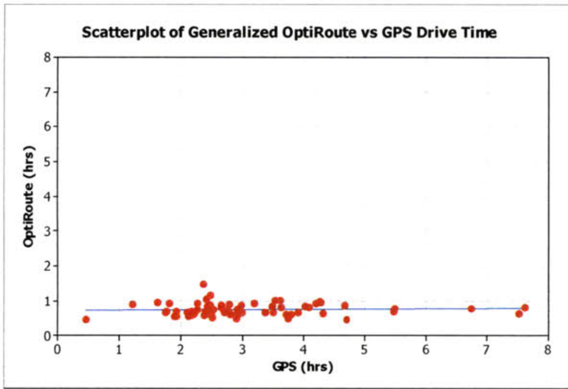


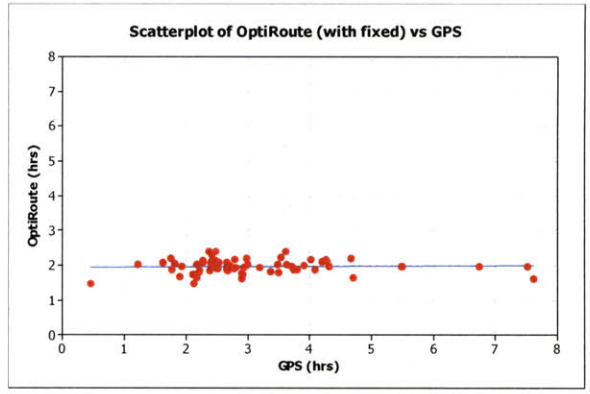
The regression for service time yields a p-value of 0.379 and an R-squared value of 1.1% indicating that OptiRoute's current service model is a poor predictor for actual service time. From the empirical AVL data, it appears that service time varies widely from route to route as seen in the scatter-plot in Figure 19. In our dataset of 70 routes, OptiRoute's estimated service time averaged close to six hours, while our AVL collected data showed an average service time closer to five hours. This data is summarized in Table 4.

TABLE 4: STOP TIME STATISTICS

SERVICE TIME		
	OptiRoute	AVL
min	4.440	0.717
max	7.785	8.467
avg	6.079	4.945
std dev	0.711	1.496
N	70	70

In contrast, the drive time shows the opposite connection.





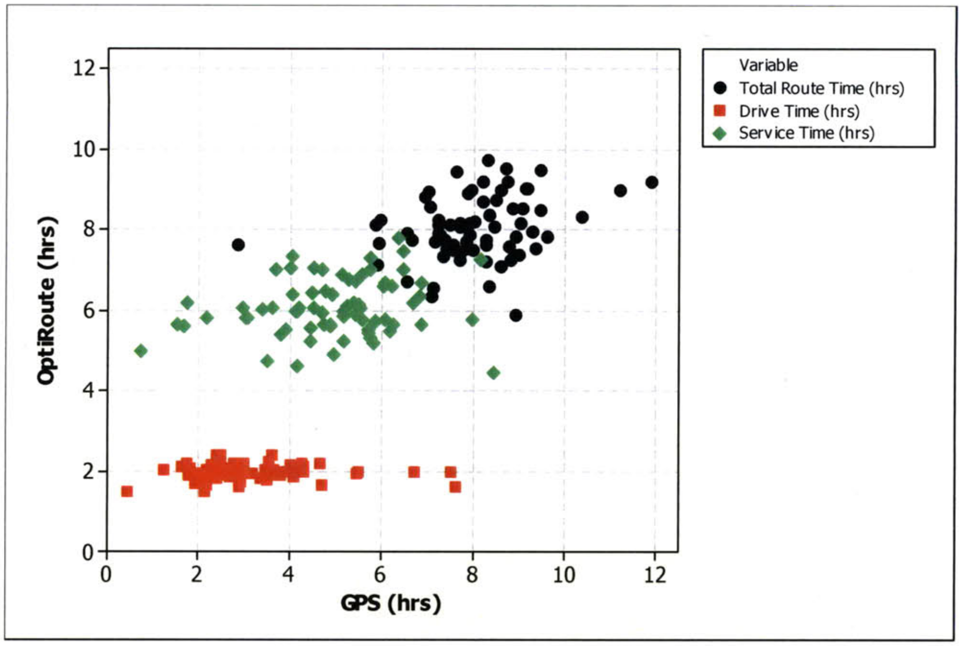
The regression for OptiRoute's buffered drive time (left) achieves a p-value of 0.76 and an R-squared value of 0.1%. Using OptiRoute's generalized drive time (right) achieves a p-value of 0.583 and an R-squared value of 0.4%. For this time estimation, OptiRoute's generalized data averages less than one hour, its as modeled estimate is close to two hours, while the AVL data average is close to three hours. These data are summarized in Table 5 below.

TABLE 5: ROUTE DRIVING TIME STATISTICS

DRIVE TIME			
	Opti	Opti Gen	AVL
min	1.471	0.455	0.450
max	2.411	1.478	7.617
avg	1.963	0.743	3.067
std dev	0.197	0.176	1.299
N	70	70	70

Understanding how the estimation of travel by OptiRoute compares with the actual driving experience is critical to improving the routing process. By looking at Figure 20, the amount of time a driver spends actually driving is a small percentage of his day.

FIGURE 20: DRIVER ACTIVITY BREAKDOWN FOR OPTIROUTE VS AVL



The next logical step to explore in driving time comparisons is to see if OptiRoute is always worse than CarTel, or if there are times of day when OptiRoute performs well. CarTel has time estimates for road segments for each hour of the day.

4.5.1 Time of Day Performance

Our initial hypothesis was that using CarTel's real traffic data would improve route planning accuracy. Therefore, we needed to understand to what extent was time of day is a factor. By using the October data set[15], we compared segment travel times estimated by OptiRoute to CarTel for each hour of the day between 5 am and 3 pm.

We found that OptiRoute's travel time estimations deviate from CarTel's most significantly from 8-10 am, during morning rush hour. Table 6 shows the R-squared values of the regressions between OptiRoute and CarTel sinking during rush hour.

---

[15] Route 570 was omitted.

TABLE 6: OPTIROUTE AND CARTEL CORRELATION BY TIME OF DAY

Hour	Count	R-squared
5	50	0.9806
6	354	0.9791
7	399	0.9593
8	413	0.8757
9	451	0.8994
10	460	0.8879
11	361	0.9478
12	213	0.9230
13	85	0.9658
14	22	0.9921
15	15	0.9816

Interestingly, during the 12 noon hour, when people dining out on their lunch break travel to restaurants, a second dip in the R-squared values occurs. We would expect traffic during this period to be different than rush hour in that people mainly travel in one direction during rush hour, whereas they are likely to cause general congestion during lunch.

4.6 DATA INTEGRITY

Having explored what differentiates OptiRoute from CarTel, our final goal was to make a determination as to which is more accurate and if applicable, in what circumstances. The final step was to perform a data integrity check to ensure the AVL data provided the same picture of a daily route as our OptiRoute sequence. We did this by confirming that the travel pattern seen by AVL was the same as given by the timestamps captured on drivers' handheld devices. The task sounds simple; however, the large number of AVL data points made finding a one-to-one data connection difficult. We looked at each route by day and determined the shortest distance between AVL data and our customer locations. The shortest distance identified a customer match and informed the order in which customers were served. For the most part, the transaction timestamps and matched the AVL engine signatures. However, in a few cases the order was incorrect. Most likely, drivers' traversing the same road in both directions throughout the day caused this discrepancy.

5 RECOMMENDATIONS

Our study evaluated the impact of using historical traffic data to plan delivery routes. Based on our approach and findings, evidence suggests that incorporating traffic patterns into vehicle routing does produce improved solutions. Needless to say, the delivery process needs to be evaluated holistically. As such, our recommendations involve the various steps for creating and executing a route. We discuss operational considerations, the potential for improving customer service, and areas for further exploration.

5.1 OPERATIONAL CONSIDERATIONS

While BevCo employs OptiRoute to generate efficient routes, there are two areas where the company could effectively improve on its current operational processes. One of BevCo's most dynamic challenges in the vehicle routing process is balancing empowering frontline staff to proactively improve day-to-day delivery with standardizing each job function and centralizing knowledge. Drivers are given the flexibility to "self-optimize" routes, and the core knowledge about their job is not accessible by others. The second area deals with improving customer location information through more accurate geocodes. As demonstrated in the analysis, inaccurate geocodes can negatively impact the quality of routing solutions.

5.1.1 Standardization and the Delivery Process

BevCo takes pride in empowering its frontline staff—drivers, dispatchers, other employees who work in distribution—to use knowledge acquired on the job to make decisions that will improve delivery. For this reason, in the Greater Boston warehouse, drivers are not required to run routes as dictated by OptiRoute. However, lack of enforced route compliance creates additional variability in the distribution process.

In the current system, core knowledge about each job function lies with individual employees, and when someone is sick or moves on to another role, the information that they posses is lost.

This means that the person who steps in to fill that position must begin relearning the intricacies of the position from square one.

Regardless of the routing system employed, more standardized processes are needed on two fronts. First, if drivers do not follow route plans, no level of improvement (incorporating traffic or otherwise) will improve actual performance. Therefore, some type of manifest compliance requirement needs to be instituted. Second, a mechanism needs to be developed whereby drivers and dispatchers share information: the more centralized and readily available, the better.

We suggest that the Greater Boston warehouse implement some of the procedures instituted by the Specialized Dispatch Operations (SDO). These incremental changes can lay the foundation for improving delivery performance through more advanced routing systems, including:

- **Route compliance policy:** Drivers would be required to follow route plans and would need get approval to deviate. In obtaining approval, they would be required to report the reason for their request. This would allow for the information to be captured and if applicable, applied to future routes.
- **Communication process**: In order to develop a method for amassing the customer information collected by drivers, a toll-free number has been implemented. Drivers call and leave a message, which is then logged. Information reported can include time window updates or other incidents; urgent messages can be flagged. The information is stored in a centralized database.
- **Utilizing a Checklist:** Dispatchers have a short checklist of items to ensure the routes they create match the assets of the warehouse (trucks) and the resources available (people). This helps ensure that in the instance that the regular dispatcher is not present in the office; there is a smooth transition for the replacement.

5.1.2 Metrics: How Should Manifest Compliance Be Measured?

In the SDO model, both driver and dispatcher performance is measured, creating incentives for working collaboratively. Compliance factors into the score for each, and other key performance indicators (KPIs) were developed with the managers. A dashboard tool was developed to communicate performance.

The two most important criteria for measuring driver performance are as follows:

- Whether all assigned stops were completed
- Whether the route was completed within the time allotted in a workday (per regulation)

However, the metrics above incentivize speed; therefore, additional metrics—those that focus on efficiency and/or customer service, for example—could be added. One such KPI could be performing all stops in the manifest sequence. This might ensure a driver hits all of the optimal time windows, as opposed to possible time windows.

There was also some discussion as to the best way to weight these metrics. Should they each be counted equally or should the two most important get the most weight, with a smaller percentage allocated for meeting the others.

5.1.3 Customer Location Accuracy

One of the early findings of the route compliance pilot was that the customer locations in OptiRoute vary from where the actual delivery point for two reasons. First, the actual delivery location may be different from a customer's address. For instance, the street address may be on a main highway, while the place where deliveries are accepted needs to be accessed via another street around the block. Second, the degree of precision of the geocodes varies. This may place a customer either across the street or in the middle of a highway, not in the usual parking spot he occupies when deliveries are made.

These two common incidents impact routing and make manifest development less accurate. As part of the commitment by drivers to follow the manifests, this information is being updated in the database so that the manifests become more accurate. While this is a labor-intensive process, the benefits are evident, especially with the utilization of the OptiRoute software.

5.2 IMPROVING CUSTOMER SERVICE

As a consumer products company, BevCo continuously works to improve service to its customers. First, by making accurate and timely deliveries, BevCo builds loyalty with its customers and therefore would like to increase the certainty of hitting times window dictated by stores. Second, with quicker transactions translating to greater overall efficiency, BevCo should consider piloting a delivery guarantee program with customers. The company has noted that as its customers' supply chains have become more sophisticated, they have embraced undertaking more sophisticated initiatives with BevCo.

The customer is one of the largest sources of variability in the downstream distribution process. Observed behavior of collecting payment during the route rides provides an excellent example. For established relationships, customers often had payment ready at the time of delivery. However, for new customers and also some who simply forgot that they were scheduled to receive a delivery, drivers had to wait for payment or even return at a later time.

If an accurate delivery window can be provided to customers, then the customer will be ready to accept BevCo's delivery. This translates into a more efficient transaction time and creates the ability to spend an increased amount of value-added time with customers and also to complete routes in a faster time. There are two ways to pilot this:

- **Internal Focus**: Initial targets for hitting time windows would be set internally. Once BevCo is confident that they can be met, the company could work with customers and expand the pilot externally.
- **External Focus**: By recognizing that customer readiness translates directly to a more efficient transaction, and BevCo could select a core group of customers with whom to working from the beginning. Based on the success of the pilot, the initiative could be broadened to other groups.

By standardizing the people and processes for which it has control, BevCo can effectively reduce the amount of variability throughout the supply chain and improve overall customer service.

5.3 TOPICS FOR FUTURE INVESTIGATION

Outside the scope of this study lie two impactful areas for further study, both with potential to further improve vehicle routing efficiency. The first is using the data collected by CarTel more stochastically by incorporating actual travel time probabilities into route development. The second is to look beyond the drive time model currently being used by BevCo and improve the accuracy of the stop time.

5.3.1 Stochastic Vehicle Routing

The nature of the data collected by CarTel is such that it can be used to incorporate more stochastic calculations of travel times. To clarify, while this study employs the mean travel time for a designated road segment at a designated time of day, CarTel's data set also has the capability of incorporating the variance. Instead of looking at the average time, routing solutions would essentially evaluate the likelihood of a travel time falling within a certain distance of the mean, as well as what that distance would be.

The implications of taking this type of approach correlate with the operational priorities set by the entity using them. For instance, if a company's main objective is to make deliveries within a series of tight time windows, it may chose to limit the variance of travel times on road segments to a predetermined threshold. This would enable drivers to arrive with degree of certainty; the tradeoff would be that this may involve traveling a longer distance. For BevCo, this would align with their objective of minimizing travel time over minimizing distance.

By using a distribution of empirical times travelled on a road segment to determine expected travel time and variance, the stochastic data can be used to create approximations of how long a delivery route should take. We feel that further investigation into this subject would add value by improving routing solutions while better hedging against uncertainty.

5.3.2 Road Network

The distributions that CarTel's traffic data uses are linked to specific road segments. In this study, CarTel provided travel time estimates but not information on the road network on which vehicles were routed. It would be interesting to study the role that distance plays in respect to time—are there certain times of day when CarTel selects roads that traverse a longer distance in order to maximize time?

5.3.3 Stop Time Model Revisited

Considering that our examination of methods to estimate non-value added driving time is a primary objective of this study, the stop times[16] projected by OptiRoute were taken as given—how they were derived was not taken into account. Nonetheless, we have reason to believe that evaluating the stop time model may be just as worthwhile.

[16] The time that a driver spends making a delivery, merchandising and providing customer service.

OptiRoute assigns fixed and variable components of each delivery. Fixed time includes but is not limited to parking the truck and receiving payment. Right now, the fixed time is set to a uniform ten minutes. Variable time is assigned by customer type based on activities such as picking an order from the bays, physically delivering it, and performing any merchandising within the store. The variable component is related only to product type.

Two shortcomings with this estimation reduce its accuracy. First, there is no factor for the expected number of trips the driver makes from a truck to a store. This is significant because a driver using a dolly to move product is limited by the number of cases that can be moved at one time. This should be incorporated in the stop time model. Additionally, customers require different levels of merchandising. Improving how merchandising is incorporated into the stop time model will make it considerably more accurate.

Returning to implications on the overarching delivery process, any improvements to the accuracy of predicting the drive time and/or stop time will increase efficiency. If the optimization decisions for vehicle routing are tied to the time of day, then holistically improving the accuracy of estimating each segment and each stop will have a compounding effect. Each part of a driver's route will be more likely to fall within its planned window.

5.4 CLOSING THOUGHTS

Through this study, we demonstrated that incorporating traffic data into vehicle route planning produces manifests that account for variations in travel times by hour of day. While we believe that the route plans using CarTel data are more accurate than OptiRoute further investigation and piloting is necessary before that hypothesis can be verified.

In addition, our investigation provided insight on the need to treat the delivery process holistically, examining the people, process and technology involved, just as the SDO does in its approach. Incremental improvements are possible; however, the biggest value add will come from ensuring there is alignment throughout the delivery process when considering implementing any new vehicle routing system or service program.

By working with BevCo, we were able to explore a problem with widespread implications for companies that operate delivery fleets. Instead of treating traffic, construction and other road hazards as factors beyond one's control, new methods exist for better managing for this uncertainty.

REFERENCES

Bianchi, L., Birattari, M., Chiarandini, M., Manfrin, M., Mastrolilli, M., Paquete, L., Rossi-Doria, O., & Schiavinotto, T. (2006). Hybrid Metaheurstics for the Vehicle Routing Problem with Stochastic Demands, *Journal of Mathematical Modeling and Algorithms*, 5: 91-110.

Braysy, O., & Gendreau, M. (2005). Vehicle Routing Problem with Time Windows, Part I: Route Construction and Local Search Algorithms. *Transportation Science*, 39, 104-118.

Braysy, O., & Gendreau, M. (2005). Vehicle Routing Problem with Time Windows, Part II: Metaheuristics. *Transportation Science*, 39, 119-139.

Dantzig, George B. (1955). Linear Programming under Uncertainty. *Management Science*, Vol. 1, No. 3/4 pp. 197-206.

Dantzig, G. B., Ramser, J.H. (1959). The Truck Dispatching Problem. *Management Science* 6(1) 80–91.

Fleischmann, Bernhard, Gietz, Martin, Gnutzman, Stefan (2004) Time-Varying Travel Time in Vehicle Routing. *Transportation Science*, Vol. 38, No. 3, 160-173.

Fulkerson, D. R. (1961). A Network Flow Computation for Project Cost Curves. *Management Science*, Vol. 7, No. 2, 167-178.

Fu, L. (2001). An Adaptive Routing Algorithm for In-Vehicle Route Guidance Systems with Real-Time Information. *Transportation Research Part B: Methodological, 35*(8), 749-765.

Gao, S. & Chabini, I. (2006). Optimal Routing Policy Problems in Stochastic Time-Dependent Networks. *Transportation Research Part B: Methodological, 40*(2), 93-122.

Golden, B. L. & Assad, A. A. (1988). *Vehicle Routing: Methods and Studies*. Elsevier Science Publishers, Amsterdam, Netherlands.

Golob, T. F. & Regan, A. C. (2003). Traffic congestion and trucking managers' use of automated routing and scheduling. *Transportation Research Part E: Logistics and Transportation Review, 39*(1), 61-78.

Hull, B., Bychkovsky, V., Zhang, Y. & Chen K. (2006). CarTel: A Distributed Mobile Sensor Computing System. Retrieved from http://nms.lcs.mit.edu/papers/cartel-sensys06.pdf.

Logi, F. & Ritchie, S. G. (2001). Development and Evaluation of a Knowledge-Based System for Traffic Congestion Management and Control. *Transportation Research Part C: Emerging Technologies, 9*(6), 433-459.

Menger, K. (1932) Das Botenproblem, in: Ergebmisse eines Mathematischen Kolloquiums 2, ed.

Miller-Hooks, E., & Mahmassani, H. S. (2000). Least Expected Time Paths in Stochastic, Time-Varying Transportation Networks. *Transportation Science, 34*(2), 198.

Peeta, S. & Zhou, C. (1999). Robustness of the Off-line a priori Stochastic Dynamic Traffic Assignment Solution for On-line Operations. *Transportation Research Part C: Emerging Technologies, 7*(5), 281-303.

Sheffi, Y. (1992). *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, NJ, Prentice-Hall Inc, Englewood Hills, NJ.

Sadek, A. W., Smith, B. L. & Demetsky, M. J. (2001). A prototype case-based reasoning system for real-time freeway traffic routing. *Transportation Research Part C: Emerging Technologies, 9*(5), 353-380.

T. Van Woensel, L. Kerbache, H. Peremans, N. Vandaele (2008). Vehicle Routing with Dynamic Travel Times: A Queueing Approach. *European Journal of Operational Research,* 186, 990–1007.

LIST OF ACRONYMS

Analysis of Variance (ANOVA)

Automatic Vehicle Location (AVL)

Advanced Traveler Information System (ATIS)

Ant Colony Optimization (ACO)

Capacitated Vehicle Routing Problem (CVRP)

Computer Science and Artificial Intelligence Lab (CSAIL)

Evolutionary Algorithm (EA)

Global Positioning System (AVL)

Key Performance Indicator (KPI)

Intermittently connected database (ICEDB)

Iterated Local Search (ILS)

Massachusetts Institute of Technology (MIT)

On-Board Diagnostic (OBD-II)

Simulation Annealing (SA)

Specialized Dispatch Operations (SDO)

Stochastic Vehicle Routing Problem (SVRP).

Tabu Search (TS)

Vehicle Routing Problem (VRP)

Vehicle Routing Problem with Stochastic Demand (VRPSD)

Vehicle Routing Problem with Time Windows (VRPTW)